ISSN 2308-3840 (online)

# World Journal of *Meta-Analysis*

2014 Bound Volume 2 Issue 1-4: 1-227





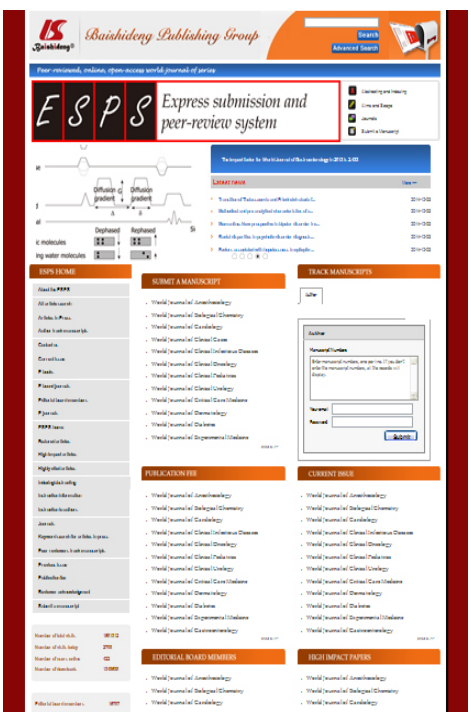











Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

A peer-reviewed, online, open-access journal of meta-analysis

# Editorial Board

2013-2018

The *World Journal of Meta-Analysis* Editorial Board consists of 400 members, representing a team of worldwide experts in clinical meta-analysis research. They are from 41 countries, including Argentina (2), Australia (3), Austria (1), Belgium (5), Brazil (11), Canada (16), Chile (3), China (121), Croatia (1), Egypt (1), Finland (4), France (2), Germany (9), Greece (11), Hungary (2), India (12), Iran (2), Ireland (1), Israel (2), Italy (41), Japan (5), Lithuania (1), Malaysia (1), Netherlands (8), New Zealand (1), Norway (1), Peru (1), Poland (4), Portugal (6), Romania (1), Saudi Arabia (4), Singapore (3), South Africa (1), South Korea (7), Spain (8), Sri Lanka (2), Sudan (1), Switzerland (2), Thailand (4), Turkey (3), United Kingdom (22), and United States (64).

**EDITOR-IN-CHIEF**

Giuseppe Biondi-Zoccai, *Latina*

**GUEST EDITORIAL BOARD MEMBERS**

Bo-Ying Bao, *Taichung*
Hsing-Yi Chang, *Maoli*
Ching-Chi Chi, *Chiayi*
Kuo-Liong Chien, *Taipei*
Chien-Chang Lee, *Doliou*
Hung-Chang Lee, *Hsinchu*
Henry WC Leung, *Taoyuan*
YC Su, *Chiayi*
Jauyih Tsauo, *Taipei*

**MEMBERS OF THE EDITORIAL BOARD**



**Argentina**

J Mariani, *Ciudad Autónoma de Buenos Aires*
Marcelo Signorini, *Provincia de Santa Fe*



**Australia**

Mark Boschen, *Southport*
Terry Boyle, *Perth*
Andy Kim Ho Lim, *Melbourne*



**Austria**

Patrick Sadoghi, *Graz*



**Belgium**

Marc Arbyn, *Brussels*
Christophe Demoulin, *Liege*
Sascha Colen, *Pellenberg*
P Lehert, *Mons*
Steve Majerus, *Liege*



**Brazil**

Euclides Castilho, *Sao Paulo*
Luciana Tricai Cavalini, *Rio de Janeiro*
Regina El Dib, *Botucatu*
Alexandre Fachini, *Araraquara*
Guilherme Francisco, *Sao Paulo*
Bruno Gualano, *Sao Paulo*
FC Paes-Barbosa, *Campo Grande*
Rachel Riera, *Sao Paulo*
Inajara Rotta, *Curitiba*
Marcos Sousa, *Belo Horizonte*
Felipe Francisco Tuon, *Curitiba*



**Canada**

Caroline Barakat-Haddad, *Toronto*
A Baranchuk, *Kingston*
Mohammad Bashashati, *Calgary*
Alonso Carrasco-Labra, *Hamilton*
Eugene Crystal, *Toronto*
Ediriweera Desapriya, *Vancouver*
A Fairchild, *Edmonton*
Alejandro Lazo-Langner, *London*
Michel Lucas, *Québec*
Alex Soroceanu, *Halifax*
Mohamed Ali Tagin, *Winnipeg*
Siamak Bashardoust Tajali, *London*
Sam Michael Wiseman, *Vancouver*
Rebecca KS Wong, *Toronto*
Clement Zai, *Toronto*
Konstantine K Zakzanis, *Toronto*



**Chile**

Alonso Carrasco-Labra, *Hamilton*
Romina Brignardello Petersen, *Santiago*
Luis A Quiñones, *Santiago*



**China**

Yi-Xi Bao, *Chongqing*
Janita Chau, *Hong Kong*
Jia-Xu Chen, *Beijing*
Shao-Jie Chen, *Chongqing*
Jin-Fei Chen, *Nanjing*
Hao-Yu Chen, *Shantou*
Ching-Lung Cheung, *Hong Kong*
Wen-Peng Cui, *Changchun*
Cong Dai, *Shenyang*
Bo Deng, *Chongqing*
Qiang Du, *Shenyang*
Jian Fei, *Shanghai*
Chun Gao, *Beijing*
Wei-Hong Ge, *Nanjing*
Ai-Hua Gu, *Nanjing*
Xiao-Xiang Guan, *Nanjing*
Zhi-Yong Guo, *Guangzhou*
Chuan-Yong Guo, *Shanghai*
Zhi-Wei He, *Dongguan*
Ben He, *Shanghai*

Guo-Wei He, *Tianjin*
G Huang, *Shanghai*
Bing-Yang Ji, *Beijing*
Jing Jiang, *Changchun*
Joey Sum Wing Kwong, *Hong Kong*
Wei-Dong Leng, *Shiyan*
Le-Qun Li, *Nanning*
Xiao-Ping Li, *Chengdu*
Jian-Sheng Li, *Zhengzhou*
Jing-Cheng Li, *Chongqing*
Jun-Sheng Li, *Nanjing*
Yan-Yan Li, *Nanjing*
Hua Liu, *Nanchong*
Tong Liu, *Tianjin*
Ai-Ping Lu, *Hong Kong*
Ying Luo, *Kunming*
Chao Ma, *Shanghai*
Jie Ma, *Xi'an*
Xiang-Yu Ma, *Chongqing*
Yan-Lei Ma, *Shanghai*
Wei Nie, *Shanghai*
Wen-Quan Niu, *Shanghai*
Wen-Sheng Pan, *Hangzhou*
Shi-Qiang Shen, *Wuhan*
Xiang-Chun Shen, *Guiyang*
Rui-Hua Shi, *Nanjing*
Yong-Bing Shi, *Suzhou*
Ke-Qing Shi, *Wenzhou*
Marcelo Signorini, *Hong Kong*
Zhi-Yuan Song, *Chongqing*
Qing-Min Sun, *Nanjing*
Yi-Hong Sun, *Beijing*
Shi-Qiao Tan, *Chengdu*
SQ Tan, *Chengdu*
Yong Tang, *Tianjin*
Jiu-Lai Tang, *Hefei*
Na-Ping Tang, *Shanghai*
Jian-Cheng Tu, *Wuhan*
Bin Wang, *Beijing*
Dao-Rong Wang, *Yangzhou*
Jing Wang, *Changshu*
Yu-Ting Wang, *Chengdu*
Xi-Shan Wang, *Harbin*
Fu-Zhou Wang, *Nanjing*
Bing Xia, *Wuhan*
Hong-Xia Wang, *Shanghai*
Zhen-Ning Wang, *Shenyang*
Na Wang, *Shijiazhuang*
Shu-Kui Wang, *Wuhan*
Xing-Huan Wang, *Wuhan*
Wei Wang, *Wuxi*
Cong-Xia Wang, *Xi'an*
Feng Xie, *Shanghai*
Zi-Qiang Xin, *Beijing*
Dan Xing, *Tianjin*
Jun Xiong, *Nanchang*
Xi-Ping Xu, *Guangzhou*
Lin Xu, *Nanjing*
Zhuo-Qun Xu, *Wuxi*
Hui-Ping Xue, *Shanghai*
Tian Yang, *Changchun*
Shuan-Ying Yang, *Xi'an*
Yi-Cong Ye, *Beijing*
Yan-Wei Yin, *Beijing*
Zi Yin, *Guangzhou*
Ym Yin, *Nanjing*
Bin Yu, *Guangzhou*
Yun-Xian Yu, *Hangzhou*
Ling Zhang, *Beijing*
Bei-Bei Zhang, *Chengdu*
Li-Li Zhang, *Chongqing*
Qiu Zhang, *Hefei*
Shuo Zhang, *Shenyang*
You-Cheng Zhang, *Lanzhou*
Jian Zhang, *Shanghai*
Jun-Hua Zhang, *Tianjin*
Yu-Rong Zhang, *Xi'an*
Zhong-Heng Zhang, *Jinhua*
Hai-Tao Zhao, *Beijing*
Pan Zhao, *Beijing*
Yu-Lan Zhao, *Shanghai*
Jie-Jiao Zheng, *Shanghai*
Xue-Sheng Zheng, *Shanghai*
Guo-Qing Zheng, *Wenzhou*
Ming-Hua Zheng, *Wenzhou*
Cui-Hong Zheng, *Wuhan*
Lai-Ping Zhong, *Shanghai*
Tian-Biao Zhou, *Nanning*
Peng Zhou, *Shanghai*
Ping Zhou, *Wuhan*
Kun-Ju Zhu, *Zhanjiang*
Xiao-Ping Zou, *Nanjing*



**Croatia**

Miljenko Franic, *Zagreb*



**Egypt**

Ashraf Fawzy Nabhan, *Cairo*



**Finland**

Jouni JK Jaakkola, *Oulu*
Ville Kyto, *Turku*
Jouko Miettunen, *Oulu*
R Quansah, *Oulu*



**France**

Alain Braillon, *Amiens*
Francesco Fiorica, *Ferrara*



**Germany**

Tonio Ball, *Freiburg*
Robert Bergholz, *Hamburg*
Jan Brunkwall, *Cologne*
Holger Cramer, *Essen*
Joseph Kambeitz, *Munich*
Sascha Meyer, *Homburg*
Thomas Nickl-Jockschat, *Aachen*
Martin Pinquart, *Marburg*
Robert Schier, *Cologne*



**Greece**

Vangelis G Alexiou, *Athens*
Stefanos Bonovas, *Athens*
Athanasios Papatsoris, *Athens*
Irini Chatziralli, *Athens*
Dimitrios Daoussis, *Patras*
Pagona Lagiou, *Athens*
John Goudakos, *Thessaloniki*
Savas Grigoriadis, *Thessaloniki*
Konstantinos A Toulis, *Thessaloniki*
Sotirios Tsiodras, *Athens*
Nikolaos Tsoukalas, *Athens*



**Hungary**

Balazs Gyorffy, *Budapest*
Istvan Wittmann, *Pecs*



**India**

Ritesh Agarwal, *Chandigarh*
Giridhara R Babu, *Bangalore*
Subho Chakrabarti, *Chandigarh*
Y Madhavi, *New Delhi*
Tanu Midha, *Kanpur*
Kameshwar Prasad, *New Delhi*
Kaushal Kishor Prasad, *Chandigarh*
Krishna Undela, *Mysore*
Singh Rajender, *Lucknow*
Vinod Ravindran, *Kozhikode*
V Shetty, *Mumbai*
R Umaya Suganthi, *Bangalore*



**Iran**

Nejat Mahdieh, *Ilam*
Sadeghi Ramin, *Mashhad*



**Ireland**

Ian Conrick-Martin, *Dublin*



**Israel**

Uri Kopylov, *Ramat Gan*
Meir Lotan, *Kfar-Saba*



**Italy**

Umberto Aguglia, *Catanzaro*
Alessandro Antonelli, *Pisa*
Annalisa Blasetti, *Chieti*
Francesco Brigo, *Merano*
Emanuele Cereda, *Pavia*
Emanuele Cigna, *Rome*
Roberto Cirocchi, *Terni*
Bernardo Cortese, *Milano*

Alessandro Cucchetti, *Bologna*
Gianfranco Damiani, *Rome*
Fabrizio D'Ascenzo, *Turin*
Massimo Del Fabbro, *Milano*
Valeria Fadda, *Florence*
Alessandro Fancellu, *Sassari*
Giuseppe Ferrante, *Milan*
Virginia Festa, *Rome*
Francesco Fiorica, *Ferrara*
Guglielmo Giraldi, *Rome*
Jenny Guidi, *Bologna*
Lorenzo Loffredo, *Rome*
Andrea Messori, *Firenze*
Eliano Pio Navarese, *Bydgoszcz*
Stefano Omboni, *Solbiate Arno*
Alvisa Palese, *Udine*
Stefano Palomba, *Reggio Emilia Stefano*
Carlo Perricone, *Rome*
Mario Petretta, *Naples*
Alessandro Pezzini, *Brescia*
Gianluca Pontone, *Milan*
Paolo Emilio Puddu, *Rome*
Andrea Rognoni, *Novara*
Giuseppe Scalabrino, *Milan*
Città della Salute e della Scienza, *Turin*
Fabrizio Sgolastra, *L'Aquila*
Maria Lucia Specchia, *Rome*
Stefano Trastulli, *Terni*
Fabio Tine, *Palermo*
Nereo Vettoretto, *Chiari*
Alberto Vianello, *Perugia*
Luigi Zorcolo, *Cagliari*



**Japan**

Nguyen Tien Huy, *Nagasaki*
Hiroharu Kamioka, *Tokyo*
Koji Kawakami, *Kyoto*
Keitaro Matsuo, *Nagoya*
Kazushi Okamoto, *Nagoya*



**Lithuania**

Edmundas Kadusevicius, *Kaunas*



**Malaysia**

SP Pani, *Ipoh*



**Netherlands**

Michel van den Bekerom, *Amsterdam*
Dimitra Dodou, *Delft*
D Haverkamp, *Amsterdam*
Vassilios Koussoulas, *Drachten*
BJ Polder, *Nijmegen*
Theo Stijnen, *Leiden*
RNM Weijers, *Amsterdam*
Joost de Winter, *Delft*



**New Zealand**

Shaofeng Li, *Auckland*



**Norway**

Eivind Berge, *Oslo*



**Peru**

Rafael Bolaños Díaz, *Lima*



**Poland**

Maciej Banach, *Lodz*
Krzysztof Jonderko, *Sosnowiec*
Jolanta Lissowska, *Warsaw*
Maciej Plaszewski, *Warsaw*



**Portugal**

Daniel Caldeira, *Lisboa*
J Costa, *Lisbon*
Ana Miguel, *Coimbra*
Manuel Morgado, *Covilh*
Bárbara Peleteiro, *Porto*
Rui Torres, *Paredes*



**Romania**

Fratila Ovidiu, *Oradea*



**Saudi Arabia**

Hazem M Al-Mandeel, *Riyadh*
Ezzeldin M Ibrahim, *Jeddah*
Mutahir A Tunio, *Riyadh*
Hayfaa A Wahabi, *Riyadh*



**Singapore**

Nikos LD Chatzisarantis, *Singapore*
Edwin Choon Wyn Lim, *Singapore*
Roger Ho, *Singapore*



**South Africa**

Alaine Umubyeyi Nyaruhirira, *Pretoria*



**South Korea**

Jung-Hee Kim, *Cheonan*
Hyangsook Lee, *Seoul*
Myeong Soo Lee, *Daejeon*
Chi-Un Pae, *Bucheon*
Yong Hyun Park, *Yangsan*
Jae Hong Seo, *Seoul*
Yong Sang Song, *Seoul*



**Spain**

Pablo Avanzas, *Oviedo*
Joan Cid, *Barcelona*
Joaquin de Haro, *Madrid*
Joan Guardia-Olmos, *Barcelona*
Nabil Halaihlel, *Zaragoza*
JA Monge-Argilés, *Alicante*
Raul Moreno, *Madrid*
Inés Velasco, *Aracena*



**Sri Lanka**

Ranil Jayawardena, *Colombo*
Priyanga Ranasinghe, *Colombo*



**Switzerland**

Jay P Singh, *Zurich*
Giorgio Treglia, *Bellinzona*



**Sudan**

Samir MH Shaheen, *Khartoum*



**Thailand**

Chuenjid Kongkaew, *Phitsanulok*
Manop Pithukpakorn, *Bangkok*
Piyamitr Sritara, *Bangkok*
Surasak Saokaew, *Phayao*



**Turkey**

Nese Demirturk, *Afyonkarahisar*
Nilüfer Ozabaci, *Eskisehir*
Ilke Sipahi, *Istanbul*



**United Kingdom**

Omar M Aboumarzouk, *Wales*
Abeer Al-Namankany, *London*
Ernest A Azzopardi, *Cardif*
Umberto Benedetto, *Cambridge*
Joanne Brooke, *London*
Noriko Cable, *London*
David Chan, *Cardiff*
YC Cheong, *Southampton*
Andrew Currie, *Harrow*
G Nabi, *Dundee*

Lesley A Anderson, *Belfast*
Valentina Gallo, *London*
Gianpiero Gravante, *Leicester*
Peter N Lee, *Surrey*
Igho Onakpoya, *Oxford*
Ashish Pradhan, *Huntingdon*
Evridiki Patelarou, *London*
Jian-Qing Shi, *Newcastle*
Surendra P Singh, *Wolverhampton*
Natalie Taylor, *Leeds*
Yousef Shahin, *Hull*
Zheng Ye, *Cambridge*



**United States**

L Joseph Su, *Rockville*
Olusola Adesope, *Pullman*
Mike Allen, *Milwaukee*
Bhupinder Anand, *Bellaire*
Stephen Aronoff, *Philadelphia*
KoKo Aung, *San Antonio*
William L Baker, *Farmington*
Moritz C Wyler von Ballmoos, *Milwaukee*
Matthew L Bechtold, *Columbia*
Atul Bhardwaj, *Hershey*
Somjot S Brar, *Los Angeles*
Hui Cai, *Nashville*
Subhash Chandra, *Towson*
Wen-Pin Chang, *Omaha*
Yong Chen, *North Wales*
Myunghan Choi, *Phoenix*
John Coverdale, *Houston*
Prakash C Deedwania, *Fresno*
Eugene Demidenko, *Hanover*
Hong-Wen Deng, *New Orleans*
Eric M Deshaies, *Syracuse*
Tao Fan, *Whitehouse Station*
Shinga Feresu, *Bloomington*
Janvier Gasana, *Miami*
Kaveh Hajifathalian, *Boston*
Mohammad Obaidul Hoque, *Baltimore*
Larissa R Brunner Huber, *Charlotte*
Imran H Iftikhar, *Columbia*
Vijayvel Jayaprakash, *Buffalo*
Xuezhi Jiang, *Weat Reading*
Shuo Jiao, *Seattle*
Evelyn Johnson, *Boise*
Le Kang, *Silver Spring*
SR Kapadia, *Cleveland*
Lior Katz, *Houston*
Daniel M Laskin, *Richmond*
Yu Liang, *Foster City*
Paul Ellis Marik, *Norfolk*
Lynne V McFarland, *Seattle*
Marcovalerio Melis, *New York*
Brian J Miller, *Augusta*
Pavlos Msaouel, *New York*
Joshua E Muscat, *Hershey*
Chee Yuan Ng, *Loma Linda*
Nghi C Nguyen, *Saint Louis*
Brandi S Niemeier, *Whitewater*
Thomas D Parsons, *Denton*
Nidal Abi Rafeh, *New Orleans*
Praveen Roy, *Marshfield*
Ali Salavati, *Philadelphia*
Ankur Sethi, *Chicago*
Tatyana A Shamliyan, *Minneapoli*
Qian Shi, *Rochester*
Zhongjie Shi, *Philadelphia*
Param Puneet Singh, *Chicago*
KV Slavin, *Chicago*
Ali El Solh, *Buffalo*
Jielin Sun, *Winston-Salem*
Richard G Trohman, *Chicago*
Laurah Turner, *Cincinnati*
Sheila Wilhelm, *Detroit*
Alex K Wong, *Los Angeles*
Xiaohui Xu, *Gainesville*
Lu Yin, *Nashville*

**ISSN** 2308-3840 (online)

# World Journal of *Meta-Analysis*

*World J Meta-Anal* 2014 February 26; 2(1): 1-16



Baishideng Publishing Group
www.wjgnet.com

WJMA

World Journal of Meta-Analysis

Contents

Quarterly Volume 2 Number 1 February 26, 2014

**REVIEW**

1 Role of nuclear cardiology for guiding device therapy in patients with heart failure

*Petretta M, Petretta A, Pellegrino T, Nappi C, Cantoni V, Cuocolo A*

*World Journal of Meta-Analysis*
**Volume 2 Number 1 February 26, 2014**

## Contents


| | | |
|---|---|---|
| **APPENDIX** | I-V | Instructions to authors |
| **ABOUT COVER** | | Editorial Board Member of *World Journal of Meta-Analysis*, Mario Petretta, Department of Translational Medical Sciences, University of Naples Federico Ⅱ, 80131 Naples, Italy |
| **AIM AND SCOPE** | | *World Journal of Meta-Analysis* (*World J Meta-Anal*, *WJMA*, online ISSN 2308-3840, DOI: 10.13105) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians, with a specific focus on meta-analysis, systematic review, mixed-treatment comparison, meta-regression, overview of reviews.<br>*WJMA* covers a variety of clinical medical fields including allergy, anesthesiology, cardiac medicine, clinical genetics, clinical neurology, critical care, dentistry, dermatology, emergency medicine, endocrinology, family medicine, gastroenterology and hepatology, geriatrics and gerontology, hematology, immunology, infectious diseases, internal medicine, obstetrics and gynecology, oncology, ophthalmology, orthopedics, otolaryngology, pathology, pediatrics, peripheral vascular disease, psychiatry, radiology, rehabilitation, respiratory medicine, rheumatology, surgery, toxicology, transplantation, and urology and nephrology, while maintaining its unique dedication to systematic reviews and meta-analyses. |
| **INDEXING/ABSTRACTING** | | *World Journal of Meta-Analysis* is now indexed in Digital Object Identifier. |
| **FLYLEAF** | I-IV | Editorial Board |


**EDITORS FOR THIS ISSUE**

**Responsible Assistant Editor:** *Xin-Xin Che*
**Responsible Science Editor:** *Huan-Huan Zhai*
**Responsible Electronic Editor:** *Cai-Hong Wang*
**Proofing Editor-in-Chief:** *Lian-Sheng Ma*

NAME OF JOURNAL
*World Journal of Meta-Analysis*

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Quarterly

EDITORS-IN-CHIEF
**Lesley A Anderson, PhD, Lecturer, PhD, MPHe, BSc (Hons), PGCHET, FHEA,** Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, Northern Ireland BT7 1NN, United Kingdom

**Giuseppe Biondi-Zoccai, MD, Assistant Professor,** Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
*World Journal of Meta-Analysis*
Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: bpgoffice@wjgnet.com
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Co., Limited
Flat C, 23/F., Lucky Plaza,
315-321 Lockhart Road, Wan Chai,
Hong Kong, China
Fax: +852-6555-7188
Telephone: +852-3177-9906
E-mail: bpgoffice@wjgnet.com
http://www.wjgnet.com

PUBLICATION DATE
February 26, 2014

COPYRIGHT
© 2014 Baishideng Publishing Group Co., Limited. Articles published by this Open Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

SPECIAL STATEMENT
All articles published in this journal represent the viewpoints of the authors except where indicated otherwise.

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Online Submissions: http://www.wjgnet.com/esps/
bpgoffice@wjgnet.com
doi:10.13105/wjma.v2.i1.1

*World J Meta-Anal* 2014 February 26; 2(1): 1-16
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Co., Limited. All rights reserved.

REVIEW

# Role of nuclear cardiology for guiding device therapy in patients with heart failure

Mario Petretta, Andrea Petretta, Teresa Pellegrino, Carmela Nappi, Valeria Cantoni, Alberto Cuocolo

Mario Petretta, Department of Translational Medical Sciences, University of Naples Federico Ⅱ, 80131 Naples, Italy
Andrea Petretta, Department of Arrhythmology, Maria Cecilia Hospital, 48010 Cotignola, Italy
Teresa Pellegrino, Institute of Biostructure and Bioimaging, National Council of Research, 80131 Napoli, Italy
Carmela Nappi, Alberto Cuocolo, Department of Advanced Biomedical Sciences, University of Naples Federico Ⅱ, 80131 Naples, Italy
Valeria Cantoni, SDN Foundation, Institute of Diagnostic and Nuclear Development, 80142 Naples, Italy
Author contributions: Petretta M and Cuocolo A designed the research; Petretta A and Pellegrino T performed the research; Nappi C and Cantoni V analyzed the data; Petretta M and Cuocolo A wrote the paper.
Correspondence to: Mario Petretta, MD, Department of Translational Medical Sciences, University of Naples Federico Ⅱ, Via Pansini 5, 80131 Naples, Italy. petretta@unina.it
Telephone: +39-81-7462233 Fax: +39-81-5457081
Received: October 25, 2013 Revised: November 20, 2013
Accepted: December 17, 2013
Published online: February 26, 2014

## Abstract

Heart failure is a dynamic condition with high morbidity and mortality and its prognosis should be reassessed frequently, particularly in patients for whom critical treatment decisions may depend on the results of prognostication. In patients with heart failure, nuclear cardiology techniques are useful to establish the etiology and the severity of the disease, while fewer studies have explored the potential capability of nuclear cardiology to guide cardiac resynchronization therapy (CRT) and to select patients for implantable cardioverter defibrillators (ICD). Left ventricular synchrony may be assessed by radionuclide angiography or gated single-photon emission computed tomography myocardial perfusion scintigraphy. These modalities have shown promise as predictors of CRT outcome using phase analysis. Combined assessment of myocardial viability and left ventricular dyssynchrony is feasible using positron emission tomography and could improve conventional response prediction criteria for CRT. Preliminary data also exists on integrated positron emission tomography/computed tomography approach for assessing myocardial viability, identifying the location of biventricular pacemaker leads, and obtaining left ventricular functional data, including contractile phase analysis. Finally, cardiac imaging with autonomic radiotracers may be useful in predicting CRT response and for identifying patients at risk for sudden cardiac death, therefore potentially offering a way to select patients for both CRT and ICD therapy. Prospective trials where imaging is combined with image-test driven therapy are needed to better define the role of nuclear cardiology for guiding device therapy in patients with heart failure.

© 2014 Baishideng Publishing Group Co., Limited. All rights reserved.

**Key words:** Heart failure; Cardiac resynchronization therapy; Implantable cardioverter defibrillators; Cardiovascular imaging; Single-photon emission-computed tomography; Positron emission tomography; Metaiodobenzylguanidine

**Core tip:** This article focuses on the potential capability of nuclear cardiology techniques to guide cardiac resynchronization therapy and to select patients for implantable cardioverter defibrillators. Radionuclide angiography and gated single-photon emission computed tomography myocardial perfusion imaging have shown promise as predictors of outcome after device therapies. Combined positron emission tomography/computed tomography may identify the location of biventricular pacemaker leads and obtain left ventricular functional data, including contractile phase analysis. Cardiac imaging with autonomic radiotracers may predict cardiac resynchronization therapy response and may also identify patients at risk for sudden cardiac death, therefore potentially offering a way to select patients for these device treatments.

Petretta M, Petretta A, Pellegrino T, Nappi C, Cantoni V, Cuocolo A. Role of nuclear cardiology for guiding device therapy in patients with heart failure. *World J Meta-Anal* 2014; 2(1): 1-16 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i1/1.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i1.1

## INTRODUCTION

The diagnostic and prognostic utility of nuclear cardiology techniques in patients with suspected or known coronary artery disease (CAD) is well established[1-5]. In patients with heart failure (HF), stress myocardial perfusion imaging is useful for establishing the etiology and the severity of the disease, while fewer studies have explored the potential capability of nuclear cardiology for guiding device therapy. Two aspects seem particularly relevant: (1) guiding cardiac resynchronization therapy (CRT) and (2) selecting patients for implantable cardioverter defibrillators (ICD).

In the United States, HF incidence has remained stable over the past several decades, with > 650000 new cases diagnosed annually[6-8]. It has been estimated that approximately 2% of the adult population in developed countries has HF; most patients will be aged > 70 years and about half will have a left ventricular (LV) ejection fraction (EF) < 50%[9]. Epidemiological and clinical trial data suggest that about 60% to 70% of HF patients have CAD[10]. An estimated 5.1 million Americans ≥ 20 years of age has HF[6]. Projections show that by 2030, the total cost of HF will increase almost 120% to $70 billion from the 2013 estimated total cost of $32 billion[6,11].

Patients with CAD-related LV systolic dysfunction have a worse prognosis than those with nonischemic cardiomyopathy[12,13]. In patients with LV systolic dysfunction, it may be useful to distinguish those with extensive CAD from those with more limited disease. The latter are considered to have coexisting but not causally related CAD and their prognosis seems to be similar to patients with nonischemic cardiomyopathy and better than those with extensive CAD[14]. However, in a cohort of 2331 patients enrolled in the HF-ACTION (Heart Failure: A Controlled Trial Investigating Outcomes of Exercise TraiNing) study, ischemic etiology was not a key determinant of a worse prognosis[15]. The analysis from the HF-ACTION clinical trial database represents the first risk prediction model for patients with HF due to systolic dysfunction that were treated with a high degree of evidence-based therapy (β-blockade, 95%; angiotensin-converting enzyme inhibitor, 74%; ICD, 40%; CRT, 18%). This analysis supports the consideration that because HF is a dynamic condition with high morbidity and mortality, its prognosis should be reassessed frequently, particularly in patients for whom critical treatment decisions may depend on the results of prognostication[16].

## CRT

In about one third of HF patients there is marked prolongation of the QRS interval[17-19]. This sign has been associated with ventricular electromechanical delay (dyssynchrony) and QRS duration and dyssynchrony are both predictors of worsening HF, sudden cardiac death and total death[20,21]. The worse prognosis cannot be only explained by a lower LVEF[22-25]. Several randomized clinical trials have shown that CRT is associated with decreases in recurrent HF hospitalization, improvement in quality of life, LV remodeling and survival[26-31]. In the REVERSE study (Resynchronization reverses Remodeling in Systolic left vEntricular dysfunction), which included New York Heart Association (NYHA) class I or Ⅱ HF patients and QRS prolongation, CRT-induced reverse remodeling was comparable in subjects with LVEF > 30% and in those with more severe LV systolic dysfunction[32]. However, about 30% of patients who receive CRT based on QRS duration do not derive symptom improvement or demonstrate reverse remodeling[33-36]. CRT seems to be useful only in a small proportion (5%-10%) of HF patients, nevertheless this is still a large number of individuals[37]. Based on data from two EuroHeart Failure surveys and extrapolating from hospital discharge statistics[38-40], it has been estimated that about 400 patients per million population per year might be suitable for CRT, or up to 400000 patients per year in European countries[41].

### *Guidelines for CRT*

The American College of Cardiology/American Heart Association/Heart Rhythm Society guidelines recommend CRT for patients who have LVEF ≤ 35%, sinus rhythm, left bundle-branch block (LBBB) with a QRS duration ≥ 150 ms, and NYHA class Ⅱ, Ⅲ, or ambulatory Ⅳ symptoms on guideline-directed medical therapy (GDMT) (class Ⅰ, level of evidence A)[42]. For these guidelines, class Ⅱa recommendations are: (1) CRT can be useful for patients who have LVEF ≤ 35%, sinus rhythm, LBBB with a QRS duration 120 to 149 ms, and NYHA class Ⅱ, Ⅲ or ambulatory Ⅳ symptoms on GDMT (level of evidence B); (2) CRT can be useful for patients who have LVEF ≤ 35%, sinus rhythm, a non-LBBB pattern with a QRS duration ≥ 150 ms, and NYHA class Ⅲ/ambulatory class Ⅳ symptoms on GDMT (level of evidence A); (3) CRT can be useful in patients with atrial fibrillation and LVEF ≤ 35% on GDMT if the patient requires ventricular pacing or otherwise meets CRT criteria and atrioventricular nodal ablation or pharmacological rate control will allow nearly 100% ventricular pacing with CRT (level of evidence B); and (4) CRT can be useful for patients on GDMT who have LVEF ≤ 35% and are undergoing new or replacement device placement with anticipated requirement for significant (> 40%) ventricular pacing (level of evidence C).

The European Society of Cardiology in collaboration with the European Heart Rhythm Association also developed guidelines for CRT[41]. According to these guidelines, CRT is recommended in chronic HF patients with LBBB and QRS duration > 150 ms and LVEF ≤ 35% who remain in NYHA functional class Ⅱ, Ⅲ and am-

**Table 1 Value of different imaging modalities for cardiac resynchronization therapy**

| Techniques | Advantages | Limitations |
|---|---|---|
| Echocardiography | Widely available; good spatial resolution; assessment of left and right ventricular volumes and function, scar burden and mitral regurgitation; no ionizing radiation; relatively low cost | Low reproducibility, limited by hemodynamic variations, operator experience, machine settings, available acoustic window, and angle of incidence; geometrical models may provide sources of error; complex interpretation due to too many indices; the time needed to perform extensive measurements may limit application in routine clinical practice |
| Cardiac computed tomography | Especially useful to guide endocardial left ventricular lead placement; pre-procedural use to characterize venous anatomy aids in lead placement; fusion imaging modalities available | Low temporal and spatial resolution, improved by the advent of dual-source multidetector; prolonged procedure times may increase risk for periprocedural complications and radiation exposure; low availability; limited clinical experience; intermediate cost |
| Cardiac magnetic resonance imaging | High spatial resolution and tissue characterization; accurate quantification of chamber size, ventricular function and 3-dimensional myocardial strain; high reproducibility with low operator dependency; no ionizing radiation; fusion imaging modalities available | Long acquisition times, potential magnetic resonance hazards of implanted cardiac devices; complex post-processing techniques; low availability; limited clinical experience; high cost |
| Radionuclide imaging | Widely available; simplicity of interpretation; provides data on scar burden and location, left ventricular function and site of latest contraction, and mechanical dyssynchrony from a single scan; fusion imaging modalities available | No role in identifying coronary venous anatomy; ionizing radiations; intermediate cost |

bulatory Ⅳ despite adequate medical treatment (class Ⅰ, level of evidence A). In chronic HF patients with LBBB with QRS duration 120-150 ms and LVEF ≤ 35% who remain in NYHA functional class Ⅱ, Ⅲ and ambulatory Ⅳ despite adequate medical treatment, CRT is also recommended (class Ⅰ, level of evidence B). For non-LBBB with QRS duration > 150 ms, CRT should be considered in chronic HF patients and LVEF ≤ 35% that remain in NYHA functional class Ⅱ, Ⅲ and ambulatory Ⅳ despite adequate medical treatment (class Ⅱa, level of evidence B). These recommendations represent a majority view of the task force but not all who contributed agreed. Several statements were based on subgroup analyses of randomized clinical trials that pose many problems with interpretation (interrelationship between QRS morphology and QRS duration, gender differences in response, prognostic benefit in ischemic *vs* non-ischemic patients) or with areas of uncertainty that are still the objects of investigation (potential role of echocardiographic dyssynchrony in narrow QRS). Thus, future studies might change knowledge and recommendations.

It must be outlined that, also using stringent criteria of selection, 20% to 40% of patients fail to respond to CRT[21,27,43-46]. It has been hypothesized that electrical dyssynchrony, as indicated by a prolonged QRS interval, does not necessarily reflect mechanical dyssynchrony, probably explaining why the patients in the above trials did not respond to CRT[47-49]. However, the recent findings do negate the use of echocardiographic measures of mechanical dyssynchrony in patients with QRS duration < 130 ms for identifying those who are most likely to have a good response to CRT[50]. In fact, in patients with systolic HF and a QRS duration < 130 ms, CRT does not reduce the rate of death or hospitalization for HF and may increase mortality[51].

### *Selection of patients for CRT*

It has been hypothesized that as a wide QRS duration (*i.e.*, electrical dyssynchrony) does not necessarily reflect mechanical dyssynchrony, characterization and quantification of resting mechanical dyssynchrony may help to improve the clinical response to CRT[52,53]. Accordingly, it has been reported that imaging techniques may be able to identify those patients who will respond favorably to CRT[54-56]; however, no modality can be considered the gold standard (Table 1). A multicenter study evaluated eight echocardiographic parameters of dyssynchrony, demonstrating a considerable variability among various techniques that assess prevalence of mechanical dyssynchrony and in identification of the latest mechanical LV contracting region[57]. Moreover, in the PROSPECT (Predictors Of Response to Cardiac Resynchronization Therapy) trial, echocardiographic parameters of cardiac mechanical dyssynchrony showed a modest accuracy to predict response to CRT, defined by improvement in the composite clinical score and ≥ 15% reduction in LV end-systolic volume[58] and only a sub-study of this trial showed that the extent of LV reverse remodeling was associated with the extent of baseline interventricular and intra-LV mechanical dyssynchrony[59]. However, a recent single center study reported that a multiparametric echocardiographic score, including LV end-diastolic dimension, LV global longitudinal strain, left atrium area, right ventricular end-diastolic area, right ventricular fractional area change and right atrium area, is helpful in selecting patients likely to undergo reverse remodeling post-CRT[60]. This score was also able to predict clinical outcomes, showing a direct inverse correlation with adverse cardiac events and all cause mortality.

Other imaging techniques, including cardiac magnetic resonance, speckle tracking echocardiography, tissue-Doppler strain imaging and nuclear cardiology, have been investigated, yielding several parameters of LV mechanical dyssynchrony that have been demonstrated to be independent predictors of CRT response and long-term outcome in observational studies[55,61-64]. Approaches

utilizing cardiac magnetic resonance and cardiac X-ray computed tomography (CT) have been used to assess LV dyssynchrony, also providing venous anatomy visualization and scar burden[65,66]. Therefore, these techniques have the potential for increased clinical use with further development and validation of adequate tools.

### Nuclear cardiology for CRT patients' selection

Radionuclide-based methods to measure LV synchrony include: (1) planar radionuclide angiography; (2) single-photon emission computed tomography (SPECT) radionuclide angiography; and (3) gated SPECT myocardial perfusion scintigraphy (MPS). These modalities have all shown promise as predictors of CRT outcome using phase analysis[62,67,68]. In particular, studies have shown that the degree of dyssynchrony on planar and SPECT radionuclide angiography has prognostic value and may be useful for predicting response to CRT[69,70]. However, these techniques are unable to provide simultaneous information on myocardial perfusion and scar location and extent that might compromise the response to CRT[55,71].

### Gated MPS

Phase analysis of gated MPS is a promising modality for several reasons: stress MPS is widely available and phase analysis can be performed on a study that is acquired using standard methods. Gated MPS also provides additional information on LV perfusion, function, scar location and extent[72]. Gated MPS produces a number of three-dimensional LV frames corresponding to different time points during the cardiac cycle. As these frames progress from the R wave, both location and intensity of each myocardial segment change periodically. The change in location of each myocardial segment allows assessment of regional wall motion and a change in the intensity indicates regional wall thickening as a result of the partial-volume effect[73,74]. Another technique of dyssynchrony calculation includes thickening and wall motion dyssynchrony calculations[75].

Phase analysis of gated MPS is based on Fourier transformation of the time-activity curve of each myocardial sample to derive the first harmonic function[76]. The first harmonic function is displayed as a continuous curve, representing thickening of the myocardial. The phase information is related to the time interval when a myocardial region starts to contract, providing information as to how uniform or inhomogeneous the distribution of these time intervals is for the entire LV. In particular, the temporal onset of mechanical contraction during the cardiac cycle of each myocardial sample is considered to be the phase of the inflection point of the thickening curve on a horizontal line representing the average myocardial count over a cardiac cycle. The phase and amplitude distribution from LV myocardial samples is displayed in histogram and polar map formats. The frequency distribution of Fourier phase angles is also tabulated in global and regional histograms, using amplitude information to eliminate phase measurements for which the corresponding amplitude was deemed too low to yield accurate measurements, as the Fourier phase angle is not well defined for a signal that exhibits low temporal variations[77]. Two validated quantitative indices of LV dyssynchrony (both expressed in degree) include the phase standard deviation (the standard deviation of the onset of mechanical contraction) and the phase histogram bandwidth (the width of the histogram band including 95% of the myocardium initiates contraction)[78]. Entropy is a further variability measure derived from information theory specifically addressed by Cedars-Sinai's QGS software[68]. For a normal subject, the entire LV starts contraction almost at the same time so that the phase polar map is nearly uniform and the phase histogram is narrow and highly peaked. Normal limits of phase standard deviation and phase histogram bandwidth have been reported[79]. Advantages are that the number of frames (usually 8 or 16) have no major effect on the calculations[76] and that reproducibility and repeatability are high due to the automated nature of the processing[80].

Scar defined by MPS, in terms of both overall scar burden and scar localized near the LV lead, has been shown to predict the lack of clinical response and failure to improve ventricular function after CRT in small studies with follow-up limited to 6 mo[81,82]. A subsequent study evaluated 620 NYHA classes Ⅲ-Ⅳ HF patients with LVEF $\leqslant$ 35% and QRS duration $\geqslant$ 120 ms referred for CRT[83]. During a mean follow-up of 2.1 years, ischemic cardiomyopathy patients had worse survival and less LVEF improvement than nonischemic patients. Ischemic patients with low scar burden had favorable survival and LVEF improvement, similar to nonischemic patients. Baseline echocardiographic dyssynchrony analysis, performed in a subgroup of patients, was not predictive of outcome. Thus, extensive scar burden unfavorably affects clinical and LV functional outcomes after CRT, regardless of baseline dyssynchrony measures. Ischemic patients and lower scar burden had significantly better outcomes, similar to nonischemic patients.

Although gated MPS may be used to quantify the scar burden, there are some limitations because it might overestimate the extent of necrosis due to low spatial resolution of the images, particularly when assessing viability in dilated ventricles with thin walls[72]. There is evidence that dyssynchrony during the relaxation phase of the LV, which causes LV filling abnormalities, may adversely affect CRT response and long-term clinical outcomes in HF patients[84,85]. Conversely, in patients with impaired relaxation, the improvement in echocardiographic assessed diastolic filling time after CRT implantation was associated with a significant reduction in all-cause mortality and HF admissions[86]. Diastolic dyssynchrony can also be assessed by nuclear cardiology techniques using a multi-harmonic fit of the thickening curve[87]. In patients with end-stage HF with LVEF of $\leqslant$ 35%, phase analysis on gated MPS showed a good intra- and interobserver reproducibility for the determination of diastolic phase

standard deviation and diastolic histogram bandwidth. Moreover, a good agreement was found between tissue Doppler imaging dyssynchrony and gated MPS diastolic phase histograms[88]. Diastolic dyssynchrony was also found to be more prevalent (65%) than systolic dyssynchrony (47%) in patients with end-stage renal disease[89].

### Positron emission tomography

Positron emission tomography (PET) imaging is performed with higher tracer counts and better spatial resolution[90,91]. However, there are limited data on using PET to quantify dyssynchrony and predict CRT response. A mismatch pattern (preserved glucose metabolism with reduced perfusion) is observed in ischemic, hibernating but viable myocardium[92], whereas a reverse mismatch pattern (reduced glucose metabolism compared to perfusion) is often observed in patients with recent myocardial infarction, chronic stable CAD or LBBB[93,94]. Also, if many patients with LBBB and LV dysfunction have a reverse mismatch pattern, the precise mechanism remains unclear and altered transmembrane glucose transport or phosphorylation kinetics has been proposed[95]. A reverse mismatch pattern is a sign of reversible myocardial dysfunction. In a pilot study, Inoue *et al*[96] found that a reverse mismatch pattern in the septum can predict a good prognosis. Birnie *et al*[97] evaluated the relationship between septal reverse mismatch and response to CRT in patients with pre-implant PET scanning and found that reduced septal glucose metabolism predicts the 3 mo response to CRT in the nonischemic cardiomyopathy subset. To evaluate if the probability of a CRT response increases with the presence of high amounts of viable and dyssynchronous myocardium, Lehner *et al*[98] studied patients who underwent gated PET before CRT device implantation, followed for 6 mo. A significantly higher amount of viable and dyssynchronous myocardium was found before CRT in responders than in nonresponders. These preliminary studies indicate that combined assessment of myocardial viability and LV dyssynchrony is feasible using PET and could improve conventional response prediction criteria for CRT.

Preliminary data also exists on integrated PET/CT approach for assessing LV viability, identifying the location of biventricular pacemaker leads, and obtaining LV functional data, including contractile phase analysis. Uebleis *et al*[99] compared seven consecutive CRT nonresponders with 7 age- and sex-matched CRT responders. Besides PET/CT, the authors measured brain natriuretic peptide levels and assessed dyssynchrony using transthoracic echocardiography. Compared with nonresponders, CRT responders showed significant differences in the declines of LV end-systolic volume and brain natriuretic peptide and in LV dyssynchrony, extent of the myocardial scar burden, and biventricular pacemaker leads positioned within viable myocardial regions. Among the nonresponders, further therapy management was guided by the PET/CT results in 4 of 7 patients. From this pilot study, it appears that cardiac hybrid imaging using gated PET/CT might identify potential reasons for nonresponse to CRT, which can guide subsequent therapy.

### Cardiac innervation imaging

Patients with HF show increased activation of the sympathetic nervous system, as reflected by an increase in plasma norepinephrine levels. In addition, neuronal uptake of norepinephrine is impaired in the failing myocardium. Both the enhanced release of norepinephrine and changes in its cardiac neuronal uptake may be responsible for the observed downregulation of adrenoreceptors in patients with HF[100]. Myocardial innervation imaging with $^{123}$I-metaiodobenzylguanidine (MIBG) scintigraphy provides a noninvasive tool for the investigation of cardiac sympathetic innervation. This technique can also demonstrate drug-induced changes in cardiac adrenergic activity[101]. MIBG is an analog of guanethidine and is taken up by the postganglionic pre-synaptic nerve endings of the adrenergic nervous system[102]. After depolarization, MIBG is released into the synaptic cleft like norepinephrine but is not metabolized. Labeling MIBG with $^{123}$I allows the visualization of adrenergic innervation *in vivo*; MIBG scintigraphy not only displays the presence of noradrenergic innervation but also its functional capability[103,104]. Thus, radiolabeled MIBG is a sympathetic neuron-imaging agent useful to study the organs richly innervated by the sympathetic nervous system, demonstrating storage, transport and reuptake characteristics similar to norepinephrine in sympathetic neurons. Uptake of MIBG into neurons from the synaptic cleft is achieved mainly through the human norepinephrine transporter 1 (hNET1). This homeostatic energy-dependent system responsible for the reuptake of norepinephrine is known as uptake-1[105]. Unlike norepinephrine, MIBG is not metabolized, allowing it to be imaged. The uptake-1 mechanism is one of the main norepinephrine disposal systems and its malfunction may lead to abnormal catecholamine concentration in the synaptic cleft. Cardiac hNET1 function appears to be reduced in common heart diseases, such as congestive HF, ischemic heart disease and stress-induced cardiomyopathy[106]. Increased norepinephrine turnover and pre-synaptic norepinephrine deficits can be identified as an increased MIBG washout rate from the heart and decreased MIBG activity quantified as the heart-to-mediastinum (H/M) ratio[107]. Whether in HF hNET1 dysfunction is only a consequence of progressive disease or if it may contribute to worsening of the disease is still unclear[102]. A standard MIBG imaging protocol is illustrated in Figure 1. Myocardial uptake and distribution of the tracer can be visually assessed and also semiquantified by calculating H/M ratio after drawing regions of interest over the heart and mediastinum (Figure 2). By comparing early and late activities, the washout rate from the myocardium can be derived, providing an index of retention of norepinephrine by sympathetic neurons[108]. In addition, polar maps of the myocardium can be constructed from the SPECT images, allowing assessment of the defect extent and severity (Figure 3). This approach pro-

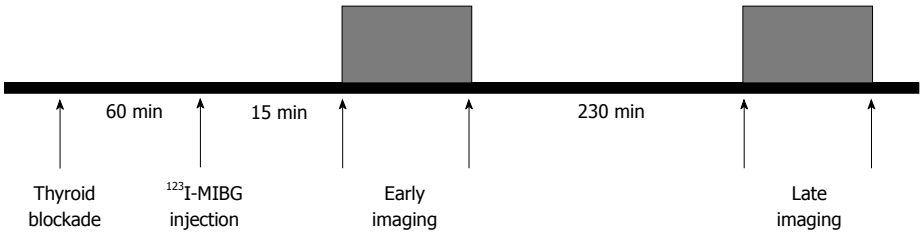



**Figure 1 Metaiodobenzylguanidine imaging protocol including planar and single-photon emission computed tomography images obtained 15 to 30 min (early) and 3 to 4 h (late) after intravenous injection of 111 to 370 MBq of $^{123}$I-metaiodobenzylguanidine.** MIBG: Metaiodobenzylguanidine.

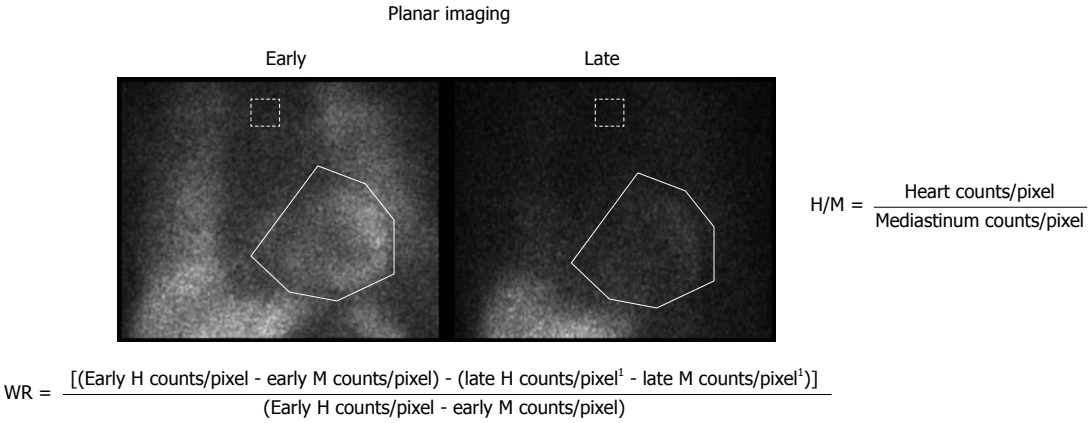



**Figure 2 Planar metaiodobenzylguanidine imaging with calculation of heart to mediastinum ratio and washout rate after drawing regions of interest over the heart (solid line) and mediastinum (dotted line).** H/M: Heart-to-mediastinum; WR: Washout rate. [1]Decay corrected.

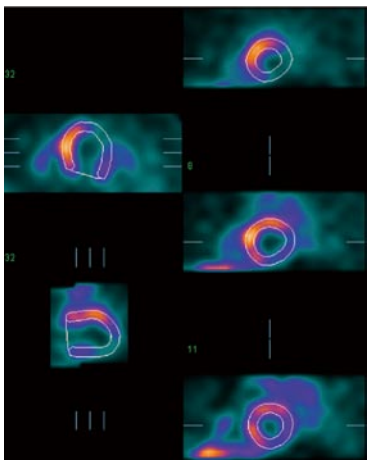



**Figure 3 Single-photon emission computed tomography metaiodobenzylguanidine imaging with construction of myocardial polar maps and assessment of innervation defect extent and severity.** SPECT: Single-photon emission computed tomography.

vides highly reproducible indices of cardiac sympathetic activity[109,110]. Figure 4 illustrates early and late MIBG imaging in a control subject and in patients with moderate or severe HF. As shown, H/M ratio progressively decreased and WR progressively accelerated from control subject to severe heart failure patient. MIBG SPECT image can also be compared with MPS to examine differences between regional innervation and perfusion, taking into account the differences between normal innervation and perfusion patterns, such as lower uptake of MIBG in the inferior wall, especially in elderly subjects[111,112].

Many studies have been designed to evaluate the role of MIBG cardiac scintigraphy in evaluating the effectiveness of pharmacological treatments in HF patients[113]. Moreover, in patients with severe cardiomyopathy, worsening of MIBG images 6 mo after optimal medical therapy was an independent predictor of cardiac death with brain natriuretic peptide [114]. Other studies also investigated the effects of CRT on myocardial innervation[103,115]. In patients with permanent dual-chamber pacemakers for complete heart block, the asynchronous ventricular activation from the apex of the right ventricle leads to regional disturbances of the adrenergic innervation of the LV myocardium, as assessed by MIBG activity[116]. Differently, Erol-Yilmaz *et al*[117] found in a very small number of HF patients that the significant improvement of NYHA class, QRS width and echocardiographic parameters after 6 mo of CRT was associated with an increase in late H/M ratio, as well as with a reduction in washout rate. Similarly, Gould *et al*[118] found that the activation of biventricular pacing was associated with an increased early and late H/M ratio in 10 patients with HF undergoing CRT. However, whether MIBG cardiac imaging has the potential to specifically direct CRT therapy in patients with HF remains uncertain. Tanaka *et al*[119] assessed H/M ratio in 50 HF patients who underwent CRT. Patients with dyssynchrony and high H/M ratio (*i.e.,* H/M ratio $\geqslant$ 1.6) had the highest frequency of response to CRT

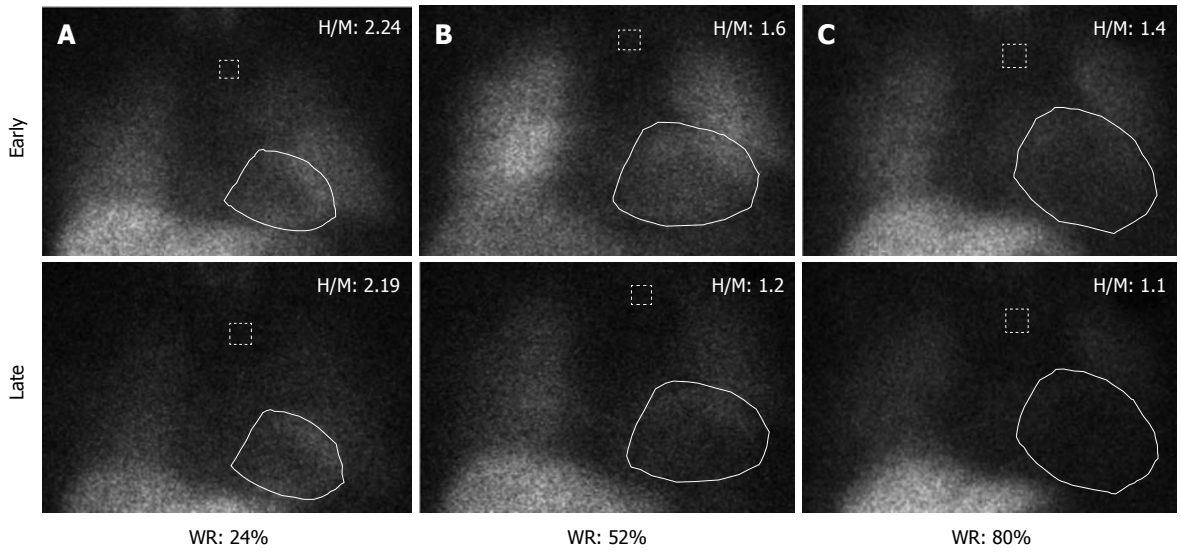



**Figure 4 Early and late planar metaiodobenzylguanidine imaging.** A: Control subject; B: Patient with moderate heart failure; C: Patient with severe heart failure. H/M: Heart-to-mediastinum; WR: Washout rate.

**Table 2 Risk of cardiac events according to different combinations of heart-to-mediastinum ratio, left ventricular ejection fraction, and β-type natriuretic peptide**

| Risk level | Predictors combination |
|---|---|
| Very high | H/M ratio < 1.6<br>LVEF < 30%<br>BNP > 140 ng/L |
| High | H/M ratio < 1.6<br>LVEF ⩾ 30%<br>BNP ⩽ 140 ng/L |
| Low | H/M ratio ⩾ 1.6<br>LVEF < 30%<br>BNP > 140 ng/L |
| Very low | H/M ratio ⩾ 1.6<br>LVEF ⩾ 30%<br>BNP ⩽ 140 ng/L |

H/M: Heart-to-mediastinum; LVEF: Left ventricular ejection fraction; BNP: β-type natriuretic peptide.

and a favorable outcome over 3 years. Based on the results of this study, it appears that dyssynchrony is associated with lower H/M ratio and higher washout rate and that MIBG scintigraphy may be valuable for predicting the response to CRT. However, this study covered a small number of patients in a single center. Thus, larger prospective studies are needed to address the role of MIBG in predicting CRT response and in appropriate selection of HF patients.

## ICD

Approximately 90% of HF patients die from cardiovascular causes. In particular, 50% die from progressive HF and the remainder die suddenly from ischemic events and/or arrhythmias, mainly due to ventricular tachycardia, ventricular fibrillation and bradycardia[120,121]. The decision to use an ICD is largely based on the patient's LVEF and NYHA functional class, the parameters most commonly used to determine the risk for sudden cardiac death[42,122,123].

### Guidelines for ICD

According to the European Society of Cardiology guidelines for the diagnosis and treatment of acute and chronic HF[9], to reduce the risk of sudden death, an ICD is recommended in secondary prevention in patients with ventricular arrhythmias causing hemodynamic instability who are expected to survive for at least 1 year with good functional status (class of recommendation I ; level of evidence A) and in primary prevention in patients with symptomatic HF (NYHA class Ⅱ-Ⅲ) and a LVEF ⩽ 35% despite ⩾ 3 mo of treatment with optimal pharmacological therapy who are expected to survive for at least 1 year with good functional status (class of recommendation I and level of evidence A for patients with ischemic etiology and > 40 d after acute myocardial infarction and B for patients with nonischemic etiology).

### Selection of patients for ICD

Widespread use of ICD therapy has increased the number of patients, medical costs and also unfavorable effects related to this device, especially in patients at a lower risk for sudden cardiac death[124-127] and in older patients, in whom underlying mortality may be increased because of comorbid conditions and competing mortality risks[128]. In many patients who receive an ICD on the basis of a reduced LVEF (*i.e.*, < 35%), the device never has to deliver therapy. An analysis from the MADIT-Ⅱ (Multicenter Automatic Defibrillator Implantation Trial) database[129] reported that of 720 patients with an ICD during an average follow-up of 21 mo, only 169 received 1 or more successful device therapies and among baseline characteristics, only NYHA class ⩾ Ⅱ and electrophysi-

ological inducibility distinguished patients who received or not appropriate ICD therapy for their first ventricular tachyarrhythmic episode. Conversely, most patients who die suddenly have a higher LVEF and thus by current guidelines do not qualify for ICD placement. Therefore, a parameter other than LVEF is needed to better select patients at risk for lethal ventricular arrhythmias who need an ICD[130].

### Nuclear cardiology for ICD patients' selection

A retrospective study investigated the relationship between LV dyssynchrony and cardiovascular events in 70 patients with ICD and LVEF < 40%, evaluated by gated MPS within 6 wk of the device implantation[131]. At 1 year, 8 patients died or had ICD shocks. The patients with events had significantly higher dyssynchrony than those without events. All patients with events had a phase standard deviation ⩾ 50°, while none of the patients with a phase standard deviation < 50° had an event. This study indicated that the severity of LV mechanical dyssynchrony by phase analysis in patients with LV dysfunction and ICD is associated with increased risk of death and appropriate ICD shock[78]. However, these preliminary findings need to be validated prospectively using a large patient population.

The autonomic nervous system plays a central role in the pathophysiology of HF and in the pathogenesis of potentially lethal ventricular arrhythmias[132] and the electrophysiological and potentially arrhythmogenic effects of catecholamines have been shown to be one of the main causes of ventricular tachycardia and sudden cardiac death in patients with autonomic dysfunction and sympathetic hyperactivity[133].

The importance of non-sudden cardiac death risk in predicting benefit from ICD therapy must be considered[134]. Considering the high economic cost of widespread ICD use, it is mandatory to identify a high-risk population who will benefit most from these devices. A number of non-LVEF risk stratification tests for predicting a variety of outcomes (overall mortality, arrhythmic events/mortality, ICD shocks and mortality benefit from ICD) have been evaluated[135,136]. These tests include signal-averaged electrocardiogram, microvolt T wave alternans, electrophysiological testing, serum markers (including brain natriuretic peptide), and autonomic function evaluation (including heart rate variability, baroreflex sensitivity, heart rate turbulence and deceleration capacity of heart rate). However, these tests have not demonstrated sufficiently high predictive value for arrhythmic death or arrhythmic events[137]. Thus, it has been hypothesized that cardiac radionuclide imaging may be useful for identifying patients at risk for sudden cardiac death from ventricular arrhythmias, therefore potentially offering a way to better select patients for ICD therapy.

The extent of myocardial scar and stress-induced ischemia, alone or in combination, are both predictive of sudden cardiac death[138,139], providing incremental prognostic information beyond the LVEF[140]. The extent of stress perfusion defects is also associated with an increased risk of sudden cardiac death in patients with CAD and LVEF > 35%[141]. Moreover, myocardial scar seems to provide a substrate for reentrant ventricular arrhythmias and sudden cardiac death in patients with nonischemic cardiomyopathy[142,143].

### Cardiac innervation imaging

As autonomic nervous system activation is often the trigger for life-threatening ventricular arrhythmias[144,145], an appealing hypothesis is that cardiac imaging with autonomic radiotracers may be useful for identifying patients at risk for sudden cardiac death, therefore potentially offering a way to better select patients for ICD therapy[146]. Several studies[147-150], a meta-analysis[151] and a pooled-data analysis of 1328 patients[152] have shown that the assessment of the cardiac autonomic state with MIBG scintigraphy can help to estimate the prognosis and to monitor the effects of therapeutic interventions in HF.

The ADMIRE-HF (AdreView Myocardial Imaging for Risk Evaluation in Heart Failure) study consisted of 2 identical open-label phase Ⅲ studies to provide prospective validation of the prognostic role of quantitation of sympathetic innervation of the myocardium using MIBG scintigraphy[150,153]. The combined primary efficacy findings from the two ADMIRE-HF studies provide the first large, prospective confirmation of the strong prognostic value of quantitation of cardiac adrenergic neuronal activity in HF patients. The results also suggest a potential application for the MIBG imaging for identifying HF patients at very low and very high risks for near-term morbidity and mortality. MIBG imaging also provided additional discrimination in analyses of interactions among β-type natriuretic peptide, LVEF and H/M ratio. Combining the values of these parameters, the highest rate events were among the subjects with all three abnormal parameters (*i.e.*, H/M ratio ⩽ 1.30, LVEF ⩽ 23%, BNP > 311 ng/L) and the lowest in the subjects with normal parameters (Table 2). During follow-up (median 17 mo), 9% of the subjects experienced nonfatal arrhythmic events or sudden cardiac death. These combined "arrhythmic" events were significantly more common in subjects with a late H/M ratio < 1.60 (reflecting both denervated myocardium and accelerated washout rate from increased sympathetic nerve activity). The highest prevalence of arrhythmic events was in the H/M range 1.30 to 1.39 and the highest H/M in a subject who experienced a fatal arrhythmic event was 1.60. Noteworthy, only 5 arrhythmic events occurred in the 191 subjects with H/M ratio ⩾ 1.6; of these, 2 of 137 subjects with no ICD experienced self-limited episodes of ventricular tachycardia, whereas 3 of 54 subjects with ICD had device activations (2 antitachycardia pacing, 1 direct current shock). These findings confirm a previous observation that abnormal MIBG uptake and washout is associated with increased incidence of sudden cardiac death and appropriate ICD discharges[154,155]. Post-hoc analyses of ADMIRE-HF study demonstrated that MIBG imaging has prognostic

value across a spectrum of LVEF[156] and that the combination of myocardial sympathetic innervation status by MIBG imaging and scar quantification by resting MPS provides risk stratification in patients with HF for the incidence of ventricular arrhythmia, in particular in patients with nonischemic cardiomyopathy[157].

It must be outlined that a cut-off point of 1.60 for H/M ratio, as found in the ADMIRE study for poor prognosis[150], may be too high for selecting patients with severe LV dysfunction who would benefit from ICD therapy. Preliminary data obtained in 47 consecutive patients with HF in NYHA functional class Ⅱ or Ⅲ, LVEF ≤ 35%, optimum pharmacological treatment, and class I indication for ICD who had undergone prior MIBG scintigraphy showed that only 5 patients had late H/M ratio > 1.6[158]. Nevertheless, the multivariable analysis, including QRS duration, creatinine levels, treatment with angiotensin converting enzyme inhibitors/angiotensin Ⅱ receptor antagonists, and early and late H/M ratios < 1.38 (median value), demonstrated an association between late H/M ratio ≤ 1.38 and creatinine levels and an increased risk of experiencing cardiac events at follow-up.

The presence of an innervation/perfusion mismatch, in which a peri-infarction zone of sympathetic denervation extends beyond the area of myocardial scar, as well as the presence of regional inhomogeneity of adrenergic innervation, also seems promising for identifying patients with HF at risk for sudden cardiac death. Hayashi *et al*[159] followed 40 patients with prior myocardial infarction, sustained ventricular arrhythmia and ICD placement for 2 years with serial noninvasive electrophysiological studies. In their study, the extent of innervation/perfusion mismatch was related to long-term variability in induced arrhythmias and was predictive of the emergence of spontaneous ventricular tachycardia or fibrillation. In patients with ventricular tachycardia and without underlying CAD, regional sympathetic denervation has been detected by MIBG imaging, whereas it was absent in patients without ventricular tachycardia[160].

### PET

The potential role of PET imaging has been evaluated in a small study[161] performed at the time of defibrillator implantation with an innervation tracer ($^{11}$C-hydroxyephedrine) imaging followed by electrophysiological mapping in patients with a history of sustained ventricular tachycardia or sudden cardiac death referred for placement of ICD. Regions of reduced hydroxyephedrine retention were detectable in each patient. The effective refractory period in areas of myocardium that demonstrated reduced hydroxyephedrine retention was longer than in areas of myocardium demonstrating normal retention. The PAREPET (Prediction of ARrhythmic Events with Positron Emission Tomography) study was designed to test the hypothesis that quantifying inhomogeneity in myocardial sympathetic innervation could identify patients at highest risk for sudden cardiac death[162]. This study prospectively enrolled 204 subjects with ischemic cardiomyopathy (LVEF ≤ 35%) eligible for primary prevention with ICD. PET was used to quantify myocardial sympathetic denervation, perfusion and viability. The primary end-point was sudden cardiac death defined as arrhythmic death or ICD discharge for ventricular fibrillation or ventricular tachycardia > 240 bpm. The results have been recently published[163]. After 4.1 years follow-up, cause-specific sudden cardiac death was 16.2%. Multivariate predictors of sudden cardiac death were PET sympathetic denervation, LV end-diastolic volume index, creatinine and no angiotensin inhibition. The study indicates that, in ischemic cardiomyopathy, sympathetic denervation assessed by PET predicts sudden cardiac death independently of LVEF and infarct volume. This may provide an improved approach for the identification of patients most likely to benefit from an ICD. Despite, some limitations[164], the PAREPET results are important because they show that cardiac PET innervation imaging is feasible in large clinical projects and that the presence of viable but denervated myocardium is harmful for the heart rhythm. Prospective trials where imaging is combined with image-test driven therapy are now the next steps.

## CONCLUSION

In patients with HF, nuclear cardiology techniques are valuable for establishing the etiology and the severity of the disease and may also be useful to guide CRT and select patients for ICD. Prospective trials where imaging is combined with image-test driven therapy are needed to better define the role of nuclear cardiology for guiding device therapy in patients with HF.

## REFERENCES


1 **de Jong MC**, Genders TS, van Geuns RJ, Moelker A, Hunink MG. Diagnostic performance of stress myocardial perfusion imaging for coronary artery disease: a systematic review and meta-analysis. *Eur Radiol* 2012; **22**: 1881-1895 [PMID: 22527375 DOI: 10.1007/s00330-012-2434-1]

2 **Beller GA**, Heede RC. SPECT imaging for detecting coronary artery disease and determining prognosis by noninvasive assessment of myocardial perfusion and myocardial viability. *J Cardiovasc Transl Res* 2011; **4**: 416-424 [PMID: 21732226 DOI: 10.1007/s12265-011-9290-2]

3 **Cuocolo A**, Petretta M, Acampa W, De Falco T. Gated SPECT myocardial perfusion imaging: the further improvements of an excellent tool. *Q J Nucl Med Mol Imaging* 2010; **54**: 129-144 [PMID: 20588210]

4 **Acampa W**, Petretta MP, Daniele S, Perrone-Filardi P, Petretta M, Cuocolo A. Myocardial perfusion imaging after coronary revascularization: a clinical appraisal. *Eur J Nucl Med Mol Imaging* 2013; **40**: 1275-1282 [PMID: 23604804 DOI: 10.1007/s00259-013-2417-8]

5 **Petretta M**, Cuocolo A. Prediction models for risk classification in cardiovascular disease. *Eur J Nucl Med Mol Imaging* 2012; **39**: 1959-1969 [PMID: 23053326 DOI: 10.1007/s00259-012-2254-1]

6 **Go AS**, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, Bravata DM, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Huffman MD, Kissela BM, Kittner SJ, Lackland DT, Li-

chtman JH, Lisabeth LD, Magid D, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER, Moy CS, Mussolino ME, Nichol G, Paynter NP, Schreiner PJ, Sorlie PD, Stein J, Turan TN, Virani SS, Wong ND, Woo D, Turner MB; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics--2013 update: a report from the American Heart Association. *Circulation* 2013; **127**: e6-e245 [PMID: 23239837 DOI: 10.1161/CIR.0b013e31828124ad]

7 **Curtis LH**, Whellan DJ, Hammill BG, Hernandez AF, Anstrom KJ, Shea AM, Schulman KA. Incidence and prevalence of heart failure in elderly persons, 1994-2003. *Arch Intern Med* 2008; **168**: 418-424 [PMID: 18299498 DOI: 10.1001/archinternmed.2007.80]

8 **Roger VL**, Weston SA, Redfield MM, Hellermann-Homan JP, Killian J, Yawn BP, Jacobsen SJ. Trends in heart failure incidence and survival in a community-based population. *JAMA* 2004; **292**: 344-350 [PMID: 15265849 DOI: 10.1001/jama.292.3.344]

9 **McMurray JJ**, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, Falk V, Filippatos G, Fonseca C, Gomez-Sanchez MA, Jaarsma T, Køber L, Lip GY, Maggioni AP, Parkhomenko A, Pieske BM, Popescu BA, Rønnevik PK, Rutten FH, Schwitter J, Seferovic P, Stepinska J, Trindade PT, Voors AA, Zannad F, Zeiher A; ESC Committee for Practice Guidelines. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2012; **33**: 1787-1847 [PMID: 22611136 DOI: 10.1093/eurheartj/ehs104]

10 **Gheorghiade M**, Sopko G, De Luca L, Velazquez EJ, Parker JD, Binkley PF, Sadowski Z, Golba KS, Prior DL, Rouleau JL, Bonow RO. Navigating the crossroads of coronary artery disease and heart failure. *Circulation* 2006; **114**: 1202-1213 [PMID: 16966596 DOI: 10.1161/CIRCULATIONAHA.106.623199]

11 **Heidenreich PA**, Trogdon JG, Khavjou OA, Butler J, Dracup K, Ezekowitz MD, Finkelstein EA, Hong Y, Johnston SC, Khera A, Lloyd-Jones DM, Nelson SA, Nichol G, Orenstein D, Wilson PW, Woo YJ; American Heart Association Advocacy Coordinating Committee; Stroke Council; Council on Cardiovascular Radiology and Intervention; Council on Clinical Cardiology; Council on Epidemiology and Prevention; Council on Arteriosclerosis; Thrombosis and Vascular. Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association. *Circulation* 2011; **123**: 933-944 [PMID: 21262990 DOI: 10.1161/CIR.0b013e31820a55f5]

12 **Levy WC**, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, Anand I, Maggioni A, Burton P, Sullivan MD, Pitt B, Poole-Wilson PA, Mann DL, Packer M. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation* 2006; **113**: 1424-1433 [PMID: 16534009 DOI: 10.1161/CIRCULATIONAHA.105.584102]

13 **Gorodeski EZ**, Chu EC, Chow CH, Levy WC, Hsich E, Starling RC. Application of the Seattle Heart Failure Model in ambulatory patients presented to an advanced heart failure therapeutics committee. *Circ Heart Fail* 2010; **3**: 706-714 [PMID: 20798278 DOI: 10.1161/CIRCHEARTFAILURE.110.944280]

14 **Felker GM**, Shaw LK, O'Connor CM. A standardized definition of ischemic cardiomyopathy for use in clinical research. *J Am Coll Cardiol* 2002; **39**: 210-218 [PMID: 11788209 DOI: 10.1016/S0735-1097(01)01738-7]

15 **O'Connor CM**, Whellan DJ, Wojdyla D, Leifer E, Clare RM, Ellis SJ, Fine LJ, Fleg JL, Zannad F, Keteyian SJ, Kitzman DW, Kraus WE, Rendall D, Piña IL, Cooper LS, Fiuzat M, Lee KL. Factors related to morbidity and mortality in patients with chronic heart failure with systolic dysfunction: the HF-ACTION predictive risk score model. *Circ Heart Fail* 2012; **5**: 63-71 [PMID: 22114101 DOI: 10.1161/CIRCHEARTFAILURE.111.963462]

16 **Aaronson KD**, Cowger J. Heart failure prognostic models: why bother? *Circ Heart Fail* 2012; **5**: 6-9 [PMID: 22253406 DOI: 10.1161/CIRCHEARTFAILURE.111.965848]

17 **Kashani A**, Barold SS. Significance of QRS complex duration in patients with heart failure. *J Am Coll Cardiol* 2005; **46**: 2183-2192 [PMID: 16360044 DOI: 10.1016/j.jacc.2005.01.071]

18 **Doval HC**, Nul DR, Grancelli HO, Perrone SV, Bortman GR, Curiel R. Randomised trial of low-dose amiodarone in severe congestive heart failure. Grupo de Estudio de la Sobrevida en la Insuficiencia Cardiaca en Argentina (GESICA) *Lancet* 1994; **344**: 493-498 [PMID: 7914611 DOI: 10.1016/S0140-6736(94)91895-3]

19 **Yancy CW**, Jessup M, Bozkurt B, Butler J, Casey DE, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation* 2013; **128**: e240-e327 [PMID: 23741058 DOI: 10.1161/CIR.0b013e31829e8776]

20 **Bleeker GB**, Schalij MJ, Molhoek SG, Verwey HF, Holman ER, Boersma E, Steendijk P, Van Der Wall EE, Bax JJ. Relationship between QRS duration and left ventricular dyssynchrony in patients with end-stage heart failure. *J Cardiovasc Electrophysiol* 2004; **15**: 544-549 [PMID: 15149423 DOI: 10.1046/j.1540-8167.2004.03604.x]

21 **Young JB**, Abraham WT, Smith AL, Leon AR, Lieberman R, Wilkoff B, Canby RC, Schroeder JS, Liem LB, Hall S, Wheelan K; Multicenter InSync ICD Randomized Clinical Evaluation (MIRACLE ICD) Trial Investigators. Combined cardiac resynchronization and implantable cardioversion defibrillation in advanced chronic heart failure: the MIRACLE ICD Trial. *JAMA* 2003; **289**: 2685-2694 [PMID: 12771115 DOI: 10.1001/jama.289.20.2685]

22 **Shamim W**, Francis DP, Yousufuddin M, Varney S, Pieopli MF, Anker SD, Coats AJ. Intraventricular conduction delay: a prognostic marker in chronic heart failure. *Int J Cardiol* 1999; **70**: 171-178 [PMID: 10454306]

23 **Xiao HB**, Roy C, Fujimoto S, Gibson DG. Natural history of abnormal conduction and its relation to prognosis in patients with dilated cardiomyopathy. *Int J Cardiol* 1996; **53**: 163-170 [PMID: 8682602]

24 **Gasparini M**, Bocchiardo M, Lunati M, Ravazzi PA, Santini M, Zardini M, Signorelli S, Passardi M, Klersy C; BELIEVE Investigators. Comparison of 1-year effects of left ventricular and biventricular pacing in patients with heart failure who have ventricular arrhythmias and left bundle-branch block: the Bi vs Left Ventricular Pacing: an International Pilot Evaluation on Heart Failure Patients with Ventricular Arrhythmias (BELIEVE) multicenter prospective randomized pilot study. *Am Heart J* 2006; **152**: 155.e1-155.e7 [PMID: 16824846 DOI: 10.1016/S0167-5273(99)00077-7]

25 **Ritter P**, Delnoy PP, Padeletti L, Lunati M, Naegele H, Borri-Brunetto A, Silvestre J. A randomized pilot study of optimization of cardiac resynchronization therapy in sinus rhythm patients using a peak endocardial acceleration sensor vs. standard methods. *Europace* 2012; **14**: 1324-1333 [PMID: 22549295 DOI: 10.1093/europace/eus059]

26 **Cazeau S**, Leclercq C, Lavergne T, Walker S, Varma C, Linde C, Garrigue S, Kappenberger L, Haywood GA, Santini M, Bailleul C, Daubert JC; Multisite Stimulation in Cardiomyopathies (MUSTIC) Study Investigators. Effects of multisite biventricular pacing in patients with heart failure and intraventricular conduction delay. *N Engl J Med* 2001; **344**: 873-880 [PMID: 11259720 DOI: 10.1056/NEJM200103223441202]

27 **Abraham WT**, Fisher WG, Smith AL, Delurgio DB, Leon AR, Loh E, Kocovic DZ, Packer M, Clavell AL, Hayes DL, Ellestad M, Trupp RJ, Underwood J, Pickering F, Truex C, McAtee P, Messenger J; MIRACLE Study Group. Multicenter InSync Randomized Clinical Evaluation. Cardiac resynchronization in chronic heart failure. *N Engl J Med* 2002; **346**: 1845-1853 [PMID: 12063368 DOI: 10.1056/NEJMoa013168]

28 **Auricchio A**, Stellbrink C, Sack S, Block M, Vogt J, Bakker P, Huth C, Schöndube F, Wolfhard U, Böcker D, Krahnefeld O, Kirkels H; Pacing Therapies in Congestive Heart Failure (PATH-CHF) Study Group. Long-term clinical effect of hemodynamically optimized cardiac resynchronization therapy in patients with heart failure and ventricular conduction delay. *J Am Coll Cardiol* 2002; **39**: 2026-2033 [PMID: 12084604 DOI: 10.1016/S0735-1097(02)01895-8]

29 **Auricchio A**, Stellbrink C, Butter C, Sack S, Vogt J, Misier AR, Böcker D, Block M, Kirkels JH, Kramer A, Huvelle E; Pacing Therapies in Congestive Heart Failure II Study Group; Guidant Heart Failure Research Group. Clinical efficacy of cardiac resynchronization therapy using left ventricular pacing in heart failure patients stratified by severity of ventricular conduction delay. *J Am Coll Cardiol* 2003; **42**: 2109-2116 [PMID: 14680736 DOI: 10.1016/j.jacc.2003.04.003]

30 **Bristow MR**, Saxon LA, Boehmer J, Krueger S, Kass DA, De Marco T, Carson P, DiCarlo L, DeMets D, White BG, DeVries DW, Feldman AM; Comparison of Medical Therapy, Pacing, and Defibrillation in Heart Failure (COMPANION) Investigators. Cardiac-resynchronization therapy with or without an implantable defibrillator in advanced chronic heart failure. *N Engl J Med* 2004; **350**: 2140-2150 [PMID: 15152059 DOI: 10.1056/NEJMoa032423]

31 **Cleland JG**, Daubert JC, Erdmann E, Freemantle N, Gras D, Kappenberger L, Tavazzi L; Cardiac Resynchronization-Heart Failure (CARE-HF) Study Investigators. The effect of cardiac resynchronization on morbidity and mortality in heart failure. *N Engl J Med* 2005; **352**: 1539-1549 [PMID: 15753115]

32 **Linde C**, Daubert C, Abraham WT, St John Sutton M, Ghio S, Hassager C, Herre JM, Bergemann TL, Gold MR. Impact of ejection fraction on the clinical response to cardiac resynchronization therapy in mild heart failure. *Circ Heart Fail* 2013; **6**: 1180-1189 [PMID: 24014828]

33 **Bleeker GB**, Schalij MJ, Holman ER, Steendijk P, van der Wall EE, Bax JJ. Cardiac resynchronization therapy in patients with systolic left ventricular dysfunction and symptoms of mild heart failure secondary to ischemic or nonischemic cardiomyopathy. *Am J Cardiol* 2006; **98**: 230-235 [PMID: 16828599 DOI: 10.1016/j.amjcard.2006.01.080]

34 **Leclercq C**, Kass DA. Retiming the failing heart: principles and current clinical status of cardiac resynchronization. *J Am Coll Cardiol* 2002; **39**: 194-201 [PMID: 11788207 DOI: 10.1016/S0735-1097(01)01747-8]

35 **Sutton MS**, Keane MG. Reverse remodelling in heart failure with cardiac resynchronisation therapy. *Heart* 2007; **93**: 167-171 [PMID: 16644861 DOI: 10.1136/hrt.2005.067967]

36 **Bax JJ**, Abraham T, Barold SS, Breithardt OA, Fung JW, Garrigue S, Gorcsan J, Hayes DL, Kass DA, Knuuti J, Leclercq C, Linde C, Mark DB, Monaghan MJ, Nihoyannopoulos P, Schalij MJ, Stellbrink C, Yu CM. Cardiac resynchronization therapy: Part 1--issues before device implantation. *J Am Coll Cardiol* 2005; **46**: 2153-2167 [PMID: 16360042 DOI: 10.1016/j.jacc.2005.09.019]

37 **Dickstein K**, Vardas PE, Auricchio A, Daubert JC, Linde C, McMurray J, Ponikowski P, Priori SG, Sutton R, van Veldhuisen DJ; Committee for Practice Guidelines of the European Society of Cardiology; ESC Committee for Practice Guidelines (CPG). 2010 focused update of ESC Guidelines on device therapy in heart failure: an update of the 2008 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure and the 2007 ESC Guidelines for cardiac and resynchronization therapy. Developed with the special contribution of the Heart Failure Association and the European Heart Rhythm Association. *Eur J Heart Fail* 2010; **12**: 1143-1153 [PMID: 20965877 DOI: 10.1093/eurjhf/hfq192]

38 **Khan NK**, Goode KM, Cleland JG, Rigby AS, Freemantle N, Eastaugh J, Clark AL, de Silva R, Calvert MJ, Swedberg K, Komajda M, Mareev V, Follath F; EuroHeart Failure Survey Investigators. Prevalence of ECG abnormalities in an international survey of patients with suspected or confirmed heart failure at death or discharge. *Eur J Heart Fail* 2007; **9**: 491-501 [PMID: 17218150 DOI: 10.1016/j.ejheart.2006.11.003]

39 **Baldasseroni S**, Opasich C, Gorini M, Lucci D, Marchionni N, Marini M, Campana C, Perini G, Deorsola A, Masotti G, Tavazzi L, Maggioni AP; Italian Network on Congestive Heart Failure Investigators. Left bundle-branch block is associated with increased 1-year sudden and total mortality rate in 5517 outpatients with congestive heart failure: a report from the Italian network on congestive heart failure. *Am Heart J* 2002; **143**: 398-405 [PMID: 11868043 DOI: 10.1067/mhj.2002.121264]

40 **Clark AL**, Goode K, Cleland JG. The prevalence and incidence of left bundle branch block in ambulant patients with chronic heart failure. *Eur J Heart Fail* 2008; **10**: 696-702 [PMID: 18501670 DOI: 10.1016/j.ejheart.2008.05.001]

41 **Brignole M**, Auricchio A, Baron-Esquivias G, Bordachar P, Boriani G, Breithardt OA, Cleland J, Deharo JC, Delgado V, Elliott PM, Gorenek B, Israel CW, Leclercq C, Linde C, Mont L, Padeletti L, Sutton R, Vardas PE; ESC Committee for Practice Guidelines (CPG), Zamorano JL, Achenbach S, Baumgartner H, Bax JJ, Bueno H, Dean V, Deaton C, Erol C, Fagard R, Ferrari R, Hasdai D, Hoes AW, Kirchhof P, Knuuti J, Kolh P, Lancellotti P, Linhart A, Nihoyannopoulos P, Piepoli MF, Ponikowski P, Sirnes PA, Tamargo JL, Tendera M, Torbicki A, Wijns W, Windecker S, Kirchhof P, Blomstrom-Lundqvist C, Badano LP, Aliyev F, Bänsch D, Baumgartner H, Bsata W, Buser P, Charron P, Daubert JC, Dobreanu D, Faerestrand S, Hasdai D, Hoes AW, Le Heuzey JY, Mavrakis H, McDonagh T, Merino JL, Nawar MM, Nielsen JC, Pieske B, Poposka L, Ruschitzka F, Tendera M, Van Gelder IC, Wilson CM. 2013 ESC Guidelines on cardiac pacing and cardiac resynchronization therapy: the Task Force on cardiac pacing and resynchronization therapy of the European Society of Cardiology (ESC). Developed in collaboration with the European Heart Rhythm Association (EHRA). *Eur Heart J* 2013; **34**: 2281-2329 [PMID: 23801822 DOI: 10.1093/eurheartj/eht150]

42 **Tracy CM**, Epstein AE, Darbar D, Dimarco JP, Dunbar SB, Estes NA, Ferguson TB, Hammill SC, Karasik PE, Link MS, Marine JE, Schoenfeld MH, Shanker AJ, Silka MJ, Stevenson LW, Stevenson WG, Varosy PD. 2012 ACCF/AHA/HRS focused update of the 2008 guidelines for device-based therapy of cardiac rhythm abnormalities: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2012; **60**: 1297-1313 [PMID: 22975230 DOI: 10.1016/j.jacc.2012.07.009]

43 **Bax JJ**, Van der Wall EE, Schalij MJ. Cardiac resynchronization therapy for heart failure. *N Engl J Med* 2002; **347**: 1803-1804; author reply 1803-1804 [PMID: 12456861 DOI: 10.1056/NEJM200211283472216]

44 **Bax JJ**, Bleeker GB, Marwick TH, Molhoek SG, Boersma E, Steendijk P, van der Wall EE, Schalij MJ. Left ventricular dyssynchrony predicts response and prognosis after cardiac resynchronization therapy. *J Am Coll Cardiol* 2004; **44**: 1834-1840 [PMID: 15519016 DOI: 10.1016/j.jacc.2004.08.016]

45 **Yu CM**, Fung JW, Zhang Q, Chan CK, Chan YS, Lin H, Kum LC, Kong SL, Zhang Y, Sanderson JE. Tissue Doppler imaging is superior to strain rate imaging and postsystolic shortening on the prediction of reverse remodeling in both ischemic and nonischemic heart failure after cardiac resynchronization therapy. *Circulation* 2004; **110**: 66-73 [PMID:

15197148 DOI: 10.1161/01.CIR.0000133276.45198.A5]
46 **Yu CM**, Zhang Q, Chan YS, Chan CK, Yip GW, Kum LC, Wu EB, Lee PW, Lam YY, Chan S, Fung JW. Tissue Doppler velocity is superior to displacement and strain mapping in predicting left ventricular reverse remodelling response after cardiac resynchronisation therapy. *Heart* 2006; **92**: 1452-1456 [PMID: 16621873 DOI: 10.1136/hrt.2005.083592]
47 **Leclercq C**, Faris O, Tunin R, Johnson J, Kato R, Evans F, Spinelli J, Halperin H, McVeigh E, Kass DA. Systolic improvement and mechanical resynchronization does not require electrical synchrony in the dilated failing heart with left bundle-branch block. *Circulation* 2002; **106**: 1760-1763 [PMID: 12356626 DOI: 10.1161/01.CIR.0000035037.11968.5C]
48 **Achilli A**, Sassara M, Ficili S, Pontillo D, Achilli P, Alessi C, De Spirito S, Guerra R, Patruno N, Serra F. Long-term effectiveness of cardiac resynchronization therapy in patients with refractory heart failure and "narrow" QRS. *J Am Coll Cardiol* 2003; **42**: 2117-2124 [PMID: 14680737 DOI: 10.1016/j.jacc.2003.08.024]
49 **Mollema SA**, Bleeker GB, van der Wall EE, Schalij MJ, Bax JJ. Usefulness of QRS duration to predict response to cardiac resynchronization therapy in patients with end-stage heart failure. *Am J Cardiol* 2007; **100**: 1665-1670 [PMID: 18036366 DOI: 10.1016/j.amjcard.2007.06.071]
50 **Yancy CW**, McMurray JJ. ECG--still the best for selecting patients for CRT. *N Engl J Med* 2013; **369**: 1463-1464 [PMID: 23998715 DOI: 10.1056/NEJMe1310406]
51 **Ruschitzka F**, Abraham WT, Singh JP, Bax JJ, Borer JS, Brugada J, Dickstein K, Ford I, Gorcsan J, Gras D, Krum H, Sogaard P, Holzmeister J; the EchoCRT Study Group. Cardiac-resynchronization therapy in heart failure with a narrow QRS complex. *N Engl J Med* 2013; **369**: 1395-1405 [PMID: 23998714 DOI: 10.1056/NEJMoa1306687]
52 **Strickberger SA**, Conti J, Daoud EG, Havranek E, Mehra MR, Piña IL, Young J; Council on Clinical Cardiology Subcommittee on Electrocardiography and Arrhythmias and the Quality of Care and Outcomes Research Interdisciplinary Working Group; Heart Rhythm Society. Patient selection for cardiac resynchronization therapy: from the Council on Clinical Cardiology Subcommittee on Electrocardiography and Arrhythmias and the Quality of Care and Outcomes Research Interdisciplinary Working Group, in collaboration with the Heart Rhythm Society. *Circulation* 2005; **111**: 2146-2150 [PMID: 15851622]
53 **Botvinick EH**, O'Connell JW, Badhwar N. Imaging synchrony. *J Nucl Cardiol* 2009; **16**: 846-848 [PMID: 19798545 DOI: 10.1161/01.CIR.0000161276.09685.4A]
54 **Bax JJ**, Gorcsan J. Echocardiography and noninvasive imaging in cardiac resynchronization therapy: results of the PROSPECT (Predictors of Response to Cardiac Resynchronization Therapy) study in perspective. *J Am Coll Cardiol* 2009; **53**: 1933-1943 [PMID: 19460606 DOI: 10.1016/j.jacc.2008.11.061]
55 **Delgado V**, Bax JJ. Assessment of systolic dyssynchrony for cardiac resynchronization therapy is clinically useful. *Circulation* 2011; **123**: 640-655 [PMID: 21321180 DOI: 10.1161/CIRCULATIONAHA.110.954404]
56 **Richardson M**, Freemantle N, Calvert MJ, Cleland JG, Tavazzi L; CARE-HF Study Steering Committee and Investigators. Predictors and treatment response with cardiac resynchronization therapy in patients with heart failure characterized by dyssynchrony: a pre-defined analysis from the CARE-HF trial. *Eur Heart J* 2007; **28**: 1827-1834 [PMID: 17540848 DOI: 10.1093/eurheartj/ehm192]
57 **Faletra FF**, Conca C, Klersy C, Klimusina J, Regoli F, Mantovani A, Pasotti E, Pedrazzini GB, De Castro S, Moccetti T, Auricchio A. Comparison of eight echocardiographic methods for determining the prevalence of mechanical dyssynchrony and site of latest mechanical contraction in patients scheduled for cardiac resynchronization therapy. *Am J Cardiol* 2009; **103**: 1746-1752 [PMID: 19539087 DOI: 10.1016/j.amjcard.2009.02.043]
58 **Chung ES**, Leon AR, Tavazzi L, Sun JP, Nihoyannopoulos P, Merlino J, Abraham WT, Ghio S, Leclercq C, Bax JJ, Yu CM, Gorcsan J, St John Sutton M, De Sutter J, Murillo J. Results of the Predictors of Response to CRT (PROSPECT) trial. *Circulation* 2008; **117**: 2608-2616 [PMID: 18458170 DOI: 10.1161/CIRCULATIONAHA.107.743120]
59 **van Bommel RJ**, Bax JJ, Abraham WT, Chung ES, Pires LA, Tavazzi L, Zimetbaum PJ, Gerritse B, Kristiansen N, Ghio S. Characteristics of heart failure patients associated with good and poor response to cardiac resynchronization therapy: a PROSPECT (Predictors of Response to CRT) sub-analysis. *Eur Heart J* 2009; **30**: 2470-2477 [PMID: 19717847 DOI: 10.1093/eurheartj/ehp368]
60 **Park JH**, Negishi K, Grimm RA, Popovic Z, Stanton T, Wilkoff BL, Marwick TH. Echocardiographic predictors of reverse remodeling after cardiac resynchronization therapy and subsequent events. *Circ Cardiovasc Imaging* 2013; **6**: 864-872 [PMID: 24084489]
61 **Bilchick KC**, Dimaano V, Wu KC, Helm RH, Weiss RG, Lima JA, Berger RD, Tomaselli GF, Bluemke DA, Halperin HR, Abraham T, Kass DA, Lardo AC. Cardiac magnetic resonance assessment of dyssynchrony and myocardial scar predicts function class improvement following cardiac resynchronization therapy. *JACC Cardiovasc Imaging* 2008; **1**: 561-568 [PMID: 19356481 DOI: 10.1016/j.jcmg.2008.04.013]
62 **Boogers MM**, Van Kriekinge SD, Henneman MM, Ypenburg C, Van Bommel RJ, Boersma E, Dibbets-Schneider P, Stokkel MP, Schalij MJ, Berman DS, Germano G, Bax JJ. Quantitative gated SPECT-derived phase analysis on gated myocardial perfusion SPECT detects left ventricular dyssynchrony and predicts response to cardiac resynchronization therapy. *J Nucl Med* 2009; **50**: 718-725 [PMID: 19403876 DOI: 10.2967/jnumed.108.060657]
63 **Gorcsan J**, Oyenuga O, Habib PJ, Tanaka H, Adelstein EC, Hara H, McNamara DM, Saba S. Relationship of echocardiographic dyssynchrony to long-term survival after cardiac resynchronization therapy. *Circulation* 2010; **122**: 1910-1918 [PMID: 20975000 DOI: 10.1161/CIRCULATIONAHA.110.954768]
64 **Hara H**, Oyenuga OA, Tanaka H, Adelstein EC, Onishi T, McNamara DM, Schwartzman D, Saba S, Gorcsan J. The relationship of QRS morphology and mechanical dyssynchrony to long-term outcome following cardiac resynchronization therapy. *Eur Heart J* 2012; **33**: 2680-2691 [PMID: 22351700 DOI: 10.1093/eurheartj/ehs013]
65 **Lardo AC**, Abraham TP, Kass DA. Magnetic resonance imaging assessment of ventricular dyssynchrony: current and emerging concepts. *J Am Coll Cardiol* 2005; **46**: 2223-2228 [PMID: 16360050 DOI: 10.1016/j.jacc.2005.09.015]
66 **Truong QA**, Singh JP, Cannon CP, Sarwar A, Nasir K, Auricchio A, Faletra FF, Sorgente A, Conca C, Moccetti T, Handschumacher M, Brady TJ, Hoffmann U. Quantitative analysis of intraventricular dyssynchrony using wall thickness by multidetector computed tomography. *JACC Cardiovasc Imaging* 2008; **1**: 772-781 [PMID: 19212461 DOI: 10.1016/j.jcmg.2008.07.014]
67 **Henneman MM**, Chen J, Dibbets-Schneider P, Stokkel MP, Bleeker GB, Ypenburg C, van der Wall EE, Schalij MJ, Garcia EV, Bax JJ. Can LV dyssynchrony as assessed with phase analysis on gated myocardial perfusion SPECT predict response to CRT? *J Nucl Med* 2007; **48**: 1104-1111 [PMID: 17574987 DOI: 10.2967/jnumed.107.039925]
68 **Van Kriekinge SD**, Nishina H, Ohba M, Berman DS, Germano G. Automatic global and regional phase analysis from gated myocardial perfusion SPECT imaging: application to the characterization of ventricular contraction in patients with left bundle branch block. *J Nucl Med* 2008; **49**: 1790-1797 [PMID: 18927331 DOI: 10.2967/jnumed.108.055160]

69 **Kerwin WF**, Botvinick EH, O'Connell JW, Merrick SH, DeMarco T, Chatterjee K, Scheibly K, Saxon LA. Ventricular contraction abnormalities in dilated cardiomyopathy: effect of biventricular pacing to correct interventricular dyssynchrony. *J Am Coll Cardiol* 2000; **35**: 1221-1227 [PMID: 10758964 DOI: 10.1016/S0735-1097(00)00555-6]

70 **Toussaint JF**, Lavergne T, Kerrou K, Froissart M, Ollitrault J, Darondel JM, Alonso C, Diebold B, Le Heuzey JY, Guize L, Paillard M. Basal asynchrony and resynchronization with biventricular pacing predict long-term improvement of LV function in heart failure patients. *Pacing Clin Electrophysiol* 2003; **26**: 1815-1823 [PMID: 12930495 DOI: 10.1046/j.1460-9592.2003.t01-1-00275.x]

71 **Riedlbauchová L**, Brunken R, Jaber WA, Popová L, Patel D, Lánská V, Civello K, Cummings J, Burkhardt JD, Saliba W, Martin D, Schweikert R, Wilkoff BL, Grimm R, Natale A. The impact of myocardial viability on the clinical outcome of cardiac resynchronization therapy. *J Cardiovasc Electrophysiol* 2009; **20**: 50-57 [PMID: 18803571 DOI: 10.1111/j.1540-8167.2008.01294.x]

72 **Friehling M**, Soman P. Newer applications of nuclear cardiology in systolic heart failure: detecting coronary artery disease and guiding device therapy. *Curr Heart Fail Rep* 2011; **8**: 106-112 [PMID: 21465127 DOI: 10.1007/s11897-011-0057-5]

73 **Cooke CD**, Garcia EV, Cullom SJ, Faber TL, Pettigrew RI. Determining the accuracy of calculating systolic wall thickening using a fast Fourier transform approximation: a simulation study based on canine and patient data. *J Nucl Med* 1994; **35**: 1185-1192 [PMID: 8014681]

74 **Ritt P**, Vija H, Hornegger J, Kuwert T. Absolute quantification in SPECT. *Eur J Nucl Med Mol Imaging* 2011; **38** Suppl 1: S69-S77 [PMID: 21484383 DOI: 10.1007/s00259-011-1770-8]

75 **van der Veen BJ**, Al Younis I, Ajmone-Marsan N, Westenberg JJ, Bax JJ, Stokkel MP, de Roos A. Ventricular dyssynchrony assessed by gated myocardial perfusion SPECT using a geometrical approach: a feasibility study. *Eur J Nucl Med Mol Imaging* 2012; **39**: 421-429 [PMID: 22113618 DOI: 10.1007/s00259-011-1991-x]

76 **Chen J**, Henneman MM, Trimble MA, Bax JJ, Borges-Neto S, Iskandrian AE, Nichols KJ, Garcia EV. Assessment of left ventricular mechanical dyssynchrony by phase analysis of ECG-gated SPECT myocardial perfusion imaging. *J Nucl Cardiol* 2008; **15**: 127-136 [PMID: 18242490 DOI: 10.1016/j.nuclcard.2007.11.004]

77 **O'Connell JW**, Schreck C, Moles M, Badwar N, DeMarco T, Olgin J, Lee B, Tseng Z, Kumar U, Botvinick EH. A unique method by which to quantitate synchrony with equilibrium radionuclide angiography. *J Nucl Cardiol* 2005; **12**: 441-450 [PMID: 16084433 DOI: 10.1016/j.nuclcard.2005.05.006]

78 **Chen J**, Garcia EV, Bax JJ, Iskandrian AE, Borges-Neto S, Soman P. SPECT myocardial perfusion imaging for the assessment of left ventricular mechanical dyssynchrony. *J Nucl Cardiol* 2011; **18**: 685-694 [PMID: 21567281 DOI: 10.1007/s12350-011-9392-x]

79 **Chen J**, Garcia EV, Folks RD, Cooke CD, Faber TL, Tauxe EL, Iskandrian AE. Onset of left ventricular mechanical contraction as determined by phase analysis of ECG-gated myocardial perfusion SPECT imaging: development of a diagnostic tool for assessment of cardiac mechanical dyssynchrony. *J Nucl Cardiol* 2005; **12**: 687-695 [PMID: 16344231 DOI: 10.1016/j.nuclcard.2005.06.088]

80 **Trimble MA**, Borges-Neto S, Honeycutt EF, Shaw LK, Pagnanelli R, Chen J, Iskandrian AE, Garcia EV, Velazquez EJ. Evaluation of mechanical dyssynchrony and myocardial perfusion using phase analysis of gated SPECT imaging in patients with left ventricular dysfunction. *J Nucl Cardiol* 2008; **15**: 663-670 [PMID: 18761269 DOI: 10.1016/j.nuclcard.2008.06.007]

81 **Adelstein EC**, Saba S. Scar burden by myocardial perfusion imaging predicts echocardiographic response to cardiac resynchronization therapy in ischemic cardiomyopathy. *Am Heart J* 2007; **153**: 105-112 [PMID: 17174647 DOI: 10.1016/j.ahj.2006.10.015]

82 **Ypenburg C**, Schalij MJ, Bleeker GB, Steendijk P, Boersma E, Dibbets-Schneider P, Stokkel MP, van der Wall EE, Bax JJ. Impact of viability and scar tissue on response to cardiac resynchronization therapy in ischaemic heart failure patients. *Eur Heart J* 2007; **28**: 33-41 [PMID: 17121757 DOI: 10.1093/eurheartj/ehl379]

83 **Adelstein EC**, Tanaka H, Soman P, Miske G, Haberman SC, Saba SF, Gorcsan J. Impact of scar burden by single-photon emission computed tomography myocardial perfusion imaging on patient outcomes following cardiac resynchronization therapy. *Eur Heart J* 2011; **32**: 93-103 [PMID: 20971745 DOI: 10.1093/eurheartj/ehq389]

84 **Cho GY**, Song JK, Park WJ, Han SW, Choi SH, Doo YC, Oh DJ, Lee Y. Mechanical dyssynchrony assessed by tissue Doppler imaging is a powerful predictor of mortality in congestive heart failure with normal QRS duration. *J Am Coll Cardiol* 2005; **46**: 2237-2243 [PMID: 16360052 DOI: 10.1016/j.jacc.2004.11.074]

85 **Schuster I**, Habib G, Jego C, Thuny F, Avierinos JF, Derumeaux G, Beck L, Medail C, Franceschi F, Renard S, Ferracci A, Lefevre J, Luccioni R, Deharo JC, Djiane P. Diastolic asynchrony is more frequent than systolic asynchrony in dilated cardiomyopathy and is less improved by cardiac resynchronization therapy. *J Am Coll Cardiol* 2005; **46**: 2250-2257 [PMID: 16360054 DOI: 10.1016/j.jacc.2005.02.096]

86 **Verbrugge FH**, Verhaert D, Grieten L, Dupont M, Rivero-Ayerza M, De Vusser P, Van Herendael H, Reyskens R, Vandervoort P, Tang WH, Mullens W. Revisiting diastolic filling time as mechanistic insight for response to cardiac resynchronization therapy. *Europace* 2013; **15**: 1747-1756 [PMID: 23821473]

87 **Shanks M**, Bertini M, Delgado V, Ng AC, Nucifora G, van Bommel RJ, Borleffs CJ, Holman ER, van de Veire NR, Schalij MJ, Bax JJ. Effect of biventricular pacing on diastolic dyssynchrony. *J Am Coll Cardiol* 2010; **56**: 1567-1575 [PMID: 21029873 DOI: 10.1016/j.jacc.2010.01.077]

88 **Boogers MJ**, Chen J, Veltman CE, van Bommel RJ, Mooyaart EA, Al Younis I, van der Hiel B, Dibbets-Schneider P, van der Wall EE, Schalij MJ, Garcia EV, Bax JJ, Delgado V. Left ventricular diastolic dyssynchrony assessed with phase analysis of gated myocardial perfusion SPECT: a comparison with tissue Doppler imaging. *Eur J Nucl Med Mol Imaging* 2011; **38**: 2031-2039 [PMID: 21850501 DOI: 10.1007/s00259-011-1870-5]

89 **Chen J**, Kalogeropoulos AP, Verdes L, Butler J, Garcia EV. Left-ventricular systolic and diastolic dyssynchrony as assessed by multi-harmonic phase analysis of gated SPECT myocardial perfusion imaging in patients with end-stage renal disease and normal LVEF. *J Nucl Cardiol* 2011; **18**: 299-308 [PMID: 21229401 DOI: 10.1007/s12350-010-9331-2]

90 **Schindler TH**, Schelbert HR, Quercioli A, Dilsizian V. Cardiac PET imaging for the detection and monitoring of coronary artery disease and microvascular health. *JACC Cardiovasc Imaging* 2010; **3**: 623-640 [PMID: 20541718 DOI: 10.1016/j.jcmg.2010.04.007]

91 **Parker MW**, Iskandar A, Limone B, Perugini A, Kim H, Jones C, Calamari B, Coleman CI, Heller GV. Diagnostic accuracy of cardiac positron emission tomography versus single photon emission computed tomography for coronary artery disease: a bivariate meta-analysis. *Circ Cardiovasc Imaging* 2012; **5**: 700-707 [PMID: 23051888 DOI: 10.1161/CIRCIMAGING.112.978270]

92 **Tillisch J**, Brunken R, Marshall R, Schwaiger M, Mandelkern M, Phelps M, Schelbert H. Reversibility of cardiac wall-motion abnormalities predicted by positron tomography. *N Engl J Med* 1986; **314**: 884-888 [PMID: 3485252]

93 **Yamagishi H**, Akioka K, Hirata K, Sakanoue Y, Takeuchi K, Yoshikawa J, Ochi H. A reverse flow-metabolism mismatch

pattern on PET is related to multivessel disease in patients with acute myocardial infarction. *J Nucl Med* 1999; **40**: 1492-1498 [PMID: 10492370]

94 **Zanco P**, Desideri A, Mobilia G, Cargnel S, Milan E, Celegon L, Buchberger R, Ferlin G. Effects of left bundle branch block on myocardial FDG PET in patients without significant coronary artery stenoses. *J Nucl Med* 2000; **41**: 973-977 [PMID: 10855620]

95 **Nowak B**, Sinha AM, Schaefer WM, Koch KC, Kaiser HJ, Hanrath P, Buell U, Stellbrink C. Cardiac resynchronization therapy homogenizes myocardial glucose metabolism and perfusion in dilated cardiomyopathy and left bundle branch block. *J Am Coll Cardiol* 2003; **41**: 1523-1528 [PMID: 12742293]

96 **Inoue N**, Takahashi N, Ishikawa T, Sumita S, Kobayashi T, Matsushita K, Matsumoto K, Taima M, Shimura M, Uchino K, Kimura K, Inoue T, Umemura S. Reverse perfusion-metabolism mismatch predicts good prognosis in patients undergoing cardiac resynchronization therapy: a pilot study. *Circ J* 2007; **71**: 126-131 [PMID: 17186990]

97 **Birnie D**, de Kemp RA, Tang AS, Ruddy TD, Gollob MH, Guo A, Williams K, Thomson K, DaSilva JN, Beanlands RS. Reduced septal glucose metabolism predicts response to cardiac resynchronization therapy. *J Nucl Cardiol* 2012; **19**: 73-83 [PMID: 22160630 DOI: 10.1007/s12350-011-9483-8]

98 **Lehner S**, Uebleis C, Schüßler F, Haug A, Kääb S, Bartenstein P, Van Kriekinge SD, Germano G, Estner H, Hacker M. The amount of viable and dyssynchronous myocardium is associated with response to cardiac resynchronization therapy: initial clinical results using multiparametric ECG-gated [18F]FDG PET. *Eur J Nucl Med Mol Imaging* 2013; **40**: 1876-1883 [PMID: 23903666]

99 **Uebleis C**, Ulbrich M, Tegtmeyer R, Schuessler F, Haserueck N, Siebermair J, Becker C, Nekolla S, Cumming P, Bartenstein P, Kääb S, Hacker M. Electrocardiogram-gated 18F-FDG PET/CT hybrid imaging in patients with unsatisfactory response to cardiac resynchronization therapy: initial clinical results. *J Nucl Med* 2011; **52**: 67-71 [PMID: 21149479 DOI: 10.2967/jnumed.110.078709]

100 **Böhm M**, La Rosée K, Schwinger RH, Erdmann E. Evidence for reduction of norepinephrine uptake sites in the failing human heart. *J Am Coll Cardiol* 1995; **25**: 146-153 [PMID: 7798493 DOI: 10.1016/0735-1097(94)00353-R]

101 **Flotats A**, Carrió I, Agostini D, Le Guludec D, Marcassa C, Schäfers M, Somsen GA, Unlu M, Verberne HJ; EANM Cardiovascular Committee; European Council of Nuclear Cardiology. Proposal for standardization of 123I-metaiodobenzylguanidine (MIBG) cardiac sympathetic imaging by the EANM Cardiovascular Committee and the European Council of Nuclear Cardiology. *Eur J Nucl Med Mol Imaging* 2010; **37**: 1802-1812 [PMID: 20577740 DOI: 10.1007/s00259-010-1491-4]

102 **Haider N**, Baliga RR, Chandrashekhar Y, Narula J. Adrenergic excess, hNET1 down-regulation, and compromised mIBG uptake in heart failure poverty in the presence of plenty. *JACC Cardiovasc Imaging* 2010; **3**: 71-75 [PMID: 20129534 DOI: 10.1016/j.jcmg.2009.11.002]

103 **Chirumamilla A**, Travin MI. Cardiac applications of 123I-mIBG imaging. *Semin Nucl Med* 2011; **41**: 374-387 [PMID: 21803188 DOI: 10.1053/j.semnuclmed.2011.04.001]

104 **Perrone-Filardi P**, Paolillo S, Dellegrottaglie S, Gargiulo P, Savarese G, Marciano C, Casaretti L, Cecere M, Musella F, Pirozzi E, Parente A, Cuocolo A. Assessment of cardiac sympathetic activity by MIBG imaging in patients with heart failure: a clinical appraisal. *Heart* 2011; **97**: 1828-1833 [PMID: 21917663 DOI: 10.1136/heartjnl-2011-300343]

105 **Schroeder C**, Jordan J. Norepinephrine transporter function and human cardiovascular disease. *Am J Physiol Heart Circ Physiol* 2012; **303**: H1273-H1282 [PMID: 23023867 DOI: 10.1152/ajpheart.00492.2012]

106 **Liang CS**. Cardiac sympathetic nerve terminal function in congestive heart failure. *Acta Pharmacol Sin* 2007; **28**: 921-927 [PMID: 17588326]

107 **Carrió I**. Cardiac neurotransmission imaging. *J Nucl Med* 2001; **42**: 1062-1076 [PMID: 11438630]

108 **Agostini D**, Carrio I, Verberne HJ. How to use myocardial 123I-MIBG scintigraphy in chronic heart failure. *Eur J Nucl Med Mol Imaging* 2009; **36**: 555-559 [PMID: 18989666 DOI: 10.1007/s00259-008-0976-x]

109 **Veltman CE**, Boogers MJ, Meinardi JE, Al Younis I, Dibbets-Schneider P, Van der Wall EE, Bax JJ, Scholte AJ. Reproducibility of planar (123)I-meta-iodobenzylguanidine (MIBG) myocardial scintigraphy in patients with heart failure. *Eur J Nucl Med Mol Imaging* 2012; **39**: 1599-1608 [PMID: 22790878 DOI: 10.1007/s00259-012-2180-2]

110 **Pellegrino T**, Petretta M, De Luca S, Paolillo S, Boemio A, Carotenuto R, Petretta MP, di Nuzzo C, Perrone-Filardi P, Cuocolo A. Observer reproducibility of results from a low-dose 123I-metaiodobenzylguanidine cardiac imaging protocol in patients with heart failure. *Eur J Nucl Med Mol Imaging* 2013; **40**: 1549-1557 [PMID: 23740373 DOI: 10.1007/s00259-013-2461-4]

111 **Estorch M**, Carrió I, Berná L, López-Pousa J, Torres G. Myocardial iodine-labeled metaiodobenzylguanidine 123 uptake relates to age. *J Nucl Cardiol* 1995; **2**: 126-132 [PMID: 9420777 DOI: 10.1016/S1071-3581(06)80022-X]

112 **Chen J**, Garcia EV, Galt JR, Folks RD, Carrio I. Optimized acquisition and processing protocols for I-123 cardiac SPECT imaging. *J Nucl Cardiol* 2006; **13**: 251-260 [PMID: 16580962 DOI: 10.1007/BF02971250]

113 **Treglia G**, Stefanelli A, Bruno I, Giordano A. Clinical usefulness of myocardial innervation imaging using Iodine-123-meta-iodobenzylguanidine scintigraphy in evaluating the effectiveness of pharmacological treatments in patients with heart failure: an overview. *Eur Rev Med Pharmacol Sci* 2013; **17**: 56-68 [PMID: 23329524]

114 **Matsui T**, Tsutamoto T, Maeda K, Kusukawa J, Kinoshita M. Prognostic value of repeated 123I-metaiodobenzylguanidine imaging in patients with dilated cardiomyopathy with congestive heart failure before and after optimized treatments--comparison with neurohumoral factors. *Circ J* 2002; **66**: 537-543 [PMID: 12074268 DOI: 10.1253/circj.66.537]

115 **Scholtens AM**, Braat AJ, Tuinenburg A, Meine M, Verberne HJ. Cardiac sympathetic innervation and cardiac resynchronization therapy. *Heart Fail Rev* 2013; Epub ahead of print [PMID: 23749044]

116 **Simantirakis EN**, Prassopoulos VK, Chrysostomakis SI, Kochiadakis GE, Koukouraki SI, Lekakis JP, Karkavitsas NS, Vardas PE. Effects of asynchronous ventricular activation on myocardial adrenergic innervation in patients with permanent dual-chamber pacemakers; an I(123)-metaiodobenzylguanidine cardiac scintigraphic study. *Eur Heart J* 2001; **22**: 323-332 [PMID: 11161951 DOI: 10.1053/euhj.2000.2482]

117 **Erol-Yilmaz A**, Verberne HJ, Schrama TA, Hrudova J, De Winter RJ, Van Eck-Smit BL, De Bruin R, Bax JJ, Schalij MJ, Wilde AA, Tukkie R. Cardiac resynchronization induces favorable neurohumoral changes. *Pacing Clin Electrophysiol* 2005; **28**: 304-310 [PMID: 15826264 DOI: 10.1111/j.1540-8159.2005.09508.x]

118 **Gould PA**, Kong G, Kalff V, Duffy SJ, Taylor AJ, Kelly MJ, Kaye DM. Improvement in cardiac adrenergic function post biventricular pacing for heart failure. *Europace* 2007; **9**: 751-756 [PMID: 17517804 DOI: 10.1093/europace/eum081]

119 **Tanaka H**, Tatsumi K, Fujiwara S, Tsuji T, Kaneko A, Ryo K, Fukuda Y, Matsumoto K, Shigeru M, Yoshida A, Kawai H, Hirata K. Effect of left ventricular dyssynchrony on cardiac sympathetic activity in heart failure patients with wide QRS duration. *Circ J* 2012; **76**: 382-389 [PMID: 22130315 DOI: 10.1253/circj.CJ-11-0752]

120 **Josephson M**, Wellens HJ. Implantable defibrillators and sudden cardiac death. *Circulation* 2004; **109**: 2685-2691 [PMID:

15184292 DOI: 10.1161/01.CIR.0000129322.97266.F3]
121 **Adabag AS**, Luepker RV, Roger VL, Gersh BJ. Sudden cardiac death: epidemiology and risk factors. *Nat Rev Cardiol* 2010; **7**: 216-225 [PMID: 20142817 DOI: 10.1038/nrcardio.2010.3]
122 **Epstein AE**, DiMarco JP, Ellenbogen KA, Estes NA, Freedman RA, Gettes LS, Gillinov AM, Gregoratos G, Hammill SC, Hayes DL, Hlatky MA, Newby LK, Page RL, Schoenfeld MH, Silka MJ, Stevenson LW, Sweeney MO, Smith SC, Jacobs AK, Adams CD, Anderson JL, Buller CE, Creager MA, Ettinger SM, Faxon DP, Halperin JL, Hiratzka LF, Hunt SA, Krumholz HM, Kushner FG, Lytle BW, Nishimura RA, Ornato JP, Page RL, Riegel B, Tarkington LG, Yancy CW; American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the ACC/AHA/NASPE 2002 Guideline Update for Implantation of Cardiac Pacemakers and Antiarrhythmia Devices); American Association for Thoracic Surgery; Society of Thoracic Surgeons. ACC/AHA/HRS 2008 Guidelines for Device-Based Therapy of Cardiac Rhythm Abnormalities: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the ACC/AHA/NASPE 2002 Guideline Update for Implantation of Cardiac Pacemakers and Antiarrhythmia Devices) developed in collaboration with the American Association for Thoracic Surgery and Society of Thoracic Surgeons. *J Am Coll Cardiol* 2008; **51**: e1-62 [PMID: 18498951 DOI: 10.1016/j.jacc.2008.02.032]
123 **Dagres N**, Hindricks G. Risk stratification after myocardial infarction: is left ventricular ejection fraction enough to prevent sudden cardiac death? *Eur Heart J* 2013; **34**: 1964-1971 [PMID: 23644180 DOI: 10.1093/eurheartj/eht109]
124 **Turakhia MP**. Sudden cardiac death and implantable cardioverter-defibrillators. *Am Fam Physician* 2010; **82**: 1357-1366 [PMID: 21121520]
125 **Bryant J**, Brodin H, Loveman E, Clegg A. Clinical effectiveness and cost-effectiveness of implantable cardioverter defibrillators for arrhythmias: a systematic review and economic evaluation. *Int J Technol Assess Health Care* 2007; **23**: 63-70 [PMID: 17234018 DOI: 10.1017/S0266462307051586]
126 **Das M**. Indications for ICD and cardiac resynchronization therapy for prevention of sudden cardiac death. *Expert Rev Cardiovasc Ther* 2009; **7**: 181-195 [PMID: 19210214 DOI: 10.1586/14779072.7.2.181]
127 **Atwater BD**, Daubert JP. Implantable cardioverter defibrillators: risks accompany the life-saving benefits. *Heart* 2012; **98**: 764-772 [PMID: 22422588 DOI: 10.1136/heartjnl-2012-301853]
128 **Sanders GD**, Kong MH, Al-Khatib SM, Peterson ED. Cost-effectiveness of implantable cardioverter defibrillators in patients & gt; or=65 years of age. *Am Heart J* 2010; **160**: 122-131 [PMID: 20598982 DOI: 10.1016/j.ahj.2010.04.021]
129 **Moss AJ**, Greenberg H, Case RB, Zareba W, Hall WJ, Brown MW, Daubert JP, McNitt S, Andrews ML, Elkin AD; Multicenter Automatic Defibrillator Implantation Trial-II (MADIT-II) Research Group. Long-term clinical course of patients after termination of ventricular tachyarrhythmia by an implanted defibrillator. *Circulation* 2004; **110**: 3760-3765 [PMID: 15583079 DOI: 10.1161/01.CIR.0000150390.04704.B7]
130 **Buxton AE**, Lee KL, Hafley GE, Pires LA, Fisher JD, Gold MR, Josephson ME, Lehmann MH, Prystowsky EN; MUSTT Investigators. Limitations of ejection fraction for prediction of sudden death risk in patients with coronary artery disease: lessons from the MUSTT study. *J Am Coll Cardiol* 2007; **50**: 1150-1157 [PMID: 17868806 DOI: 10.1016/j.jacc.2007.04.095]
131 **Aljaroudi WA**, Hage FG, Hermann D, Doppalapudi H, Venkataraman R, Heo J, Iskandrian AE. Relation of left-ventricular dyssynchrony by phase analysis of gated SPECT images and cardiovascular events in patients with implantable cardiac defibrillators. *J Nucl Cardiol* 2010; **17**: 398-404 [PMID: 20300907 DOI: 10.1007/s12350-009-9169-7]
132 **Braunwald E**. Research advances in heart failure: a compendium. *Circ Res* 2013; **113**: 633-645 [PMID: 23888056 DOI: 10.1161/CIRCRESAHA.113.302254]
133 **Akutsu Y**, Kaneko K, Kodama Y, Li HL, Kawamura M, Asano T, Tanno K, Shinozuka A, Gokan T, Kobayashi Y. The significance of cardiac sympathetic nervous system abnormality in the long-term prognosis of patients with a history of ventricular tachyarrhythmia. *J Nucl Med* 2009; **50**: 61-67 [PMID: 19091900 DOI: 10.2967/jnumed.108.055194]
134 **Scott PA**, Sterns LD, Tang AS. Do patients at high risk of nonsudden cardiac death benefit from prophylactic ICD therapy? *Curr Opin Cardiol* 2012; **27**: 1-7 [PMID: 22123603 DOI: 10.1097/HCO.0b013e32834d84bf]
135 **Stein KM**. Noninvasive risk stratification for sudden death: signal-averaged electrocardiography, nonsustained ventricular tachycardia, heart rate variability, baroreflex sensitivity, and QRS duration. *Prog Cardiovasc Dis* 2008; **51**: 106-117 [PMID: 18774010 DOI: 10.1016/j.pcad.2007.10.001]
136 **Stecker EC**, Chugh SS. Prediction of sudden cardiac death: next steps in pursuit of effective methodology. *J Interv Card Electrophysiol* 2011; **31**: 101-107 [PMID: 21384153 DOI: 10.1007/s10840-010-9535-z]
137 **Buxton AE**. Risk stratification for sudden death in patients with coronary artery disease. *Heart Rhythm* 2009; **6**: 836-847 [PMID: 19467514 DOI: 10.1016/j.hrthm.2009.02.016]
138 **Morishima I**, Sone T, Tsuboi H, Mukawa H, Uesugi M, Morikawa S, Takagi K, Niwa T, Morita Y, Murakami R, Numaguchi Y, Murohara T, Okumura K. Risk stratification of patients with prior myocardial infarction and advanced left ventricular dysfunction by gated myocardial perfusion SPECT imaging. *J Nucl Cardiol* 2008; **15**: 631-637 [PMID: 18761265 DOI: 10.1016/j.nuclcard.2008.03.009]
139 **Hoque A**, Maaieh M, Longaker RA, Stoddard MF. Exercise echocardiography and thallium-201 single-photon emission computed tomography stress test for 5- and 10-year prognosis of mortality and specific cardiac events. *J Am Soc Echocardiogr* 2002; **15**: 1326-1334 [PMID: 12415225 DOI: 10.1067/mje.2002.126109]
140 **Piccini JP**, Horton JR, Shaw LK, Al-Khatib SM, Lee KL, Iskandrian AE, Borges-Neto S. Single-photon emission computed tomography myocardial perfusion defects are associated with an increased risk of all-cause death, cardiovascular death, and sudden cardiac death. *Circ Cardiovasc Imaging* 2008; **1**: 180-188 [PMID: 19808541 DOI: 10.1161/CIRCIMAGING.108.776484]
141 **Piccini JP**, Starr AZ, Horton JR, Shaw LK, Lee KL, Al-Khatib SM, Iskandrian AE, O'Connor CM, Borges-Neto S. Single-photon emission computed tomography myocardial perfusion imaging and the risk of sudden cardiac death in patients with coronary disease and left ventricular ejection fraction& gt; 35%. *J Am Coll Cardiol* 2010; **56**: 206-214 [PMID: 20620740 DOI: 10.1016/j.jacc.2010.01.061]
142 **Assomull RG**, Prasad SK, Lyne J, Smith G, Burman ED, Khan M, Sheppard MN, Poole-Wilson PA, Pennell DJ. Cardiovascular magnetic resonance, fibrosis, and prognosis in dilated cardiomyopathy. *J Am Coll Cardiol* 2006; **48**: 1977-1985 [PMID: 17112987 DOI: 10.1016/j.jacc.2006.07.049]
143 **Gulati A**, Jabbour A, Ismail TF, Guha K, Khwaja J, Raza S, Morarji K, Brown TD, Ismail NA, Dweck MR, Di Pietro E, Roughton M, Wage R, Daryani Y, O'Hanlon R, Sheppard MN, Alpendurada F, Lyon AR, Cook SA, Cowie MR, Assomull RG, Pennell DJ, Prasad SK. Association of fibrosis with mortality and sudden cardiac death in patients with nonischemic dilated cardiomyopathy. *JAMA* 2013; **309**: 896-908 [PMID: 23462786 DOI: 10.1001/jama.2013.1363]
144 **Volders PG**. Novel insights into the role of the sympathetic nervous system in cardiac arrhythmogenesis. *Heart Rhythm* 2010; **7**: 1900-1906 [PMID: 20570754 DOI: 10.1016/j.hrthm.2010.06.003]
145 **Tan AY**, Verrier RL. The role of the autonomic nervous system in cardiac arrhythmias. *Handb Clin Neurol* 2013; **117**:

135-145 [PMID: 24095122 DOI: 10.1016/B978-0-444-53491-0.00012-2]
146 **Gerson MC**, Abdallah M, Muth JN, Costea AI. Will imaging assist in the selection of patients with heart failure for an ICD? *JACC Cardiovasc Imaging* 2010; **3**: 101-110 [PMID: 20129539 DOI: 10.1016/j.jcmg.2009.07.013]
147 **Merlet P**, Valette H, Dubois-Randé JL, Moyse D, Duboc D, Dove P, Bourguignon MH, Benvenuti C, Duval AM, Agostini D. Prognostic value of cardiac metaiodobenzylguanidine imaging in patients with heart failure. *J Nucl Med* 1992; **33**: 471-477 [PMID: 1552326]
148 **Agostini D**, Verberne HJ, Burchert W, Knuuti J, Povinec P, Sambuceti G, Unlu M, Estorch M, Banerjee G, Jacobson AF. I-123-mIBG myocardial imaging for assessment of risk for a major cardiac event in heart failure patients: insights from a retrospective European multicenter study. *Eur J Nucl Med Mol Imaging* 2008; **35**: 535-546 [PMID: 18043919 DOI: 10.1007/s00259-007-0639-3]
149 **Kasama S**, Toyama T, Sumino H, Nakazawa M, Matsumoto N, Sato Y, Kumakura H, Takayama Y, Ichikawa S, Suzuki T, Kurabayashi M. Prognostic value of serial cardiac 123I-MIBG imaging in patients with stabilized chronic heart failure and reduced left ventricular ejection fraction. *J Nucl Med* 2008; **49**: 907-914 [PMID: 18483106 DOI: 10.2967/jnumed.107.047548]
150 **Jacobson AF**, Senior R, Cerqueira MD, Wong ND, Thomas GS, Lopez VA, Agostini D, Weiland F, Chandna H, Narula J; ADMIRE-HF Investigators. Myocardial iodine-123 metaiodobenzylguanidine imaging and cardiac events in heart failure. Results of the prospective ADMIRE-HF (AdreView Myocardial Imaging for Risk Evaluation in Heart Failure) study. *J Am Coll Cardiol* 2010; **55**: 2212-2221 [PMID: 20188504 DOI: 10.1016/j.jacc.2010.01.014]
151 **Kuwabara Y**, Tamaki N, Nakata T, Yamashina S, Yamazaki J. Determination of the survival rate in patients with congestive heart failure stratified by $^{123}$I-MIBG imaging: a meta-analysis from the studies performed in Japan. *Ann Nucl Med* 2011; **25**: 101-107 [PMID: 21210256 DOI: 10.1007/s12149-010-0452-0]
152 **Nakata T**, Nakajima K, Yamashina S, Yamada T, Momose M, Kasama S, Matsui T, Matsuo S, Travin MI, Jacobson AF. A pooled analysis of multicenter cohort studies of (123)I-mIBG imaging of sympathetic innervation for assessment of long-term prognosis in heart failure. *JACC Cardiovasc Imaging* 2013; **6**: 772-784 [PMID: 23845574 DOI: 10.1016/j.jcmg.2013.02.007]
153 **Jacobson AF**, Lombard J, Banerjee G, Camici PG. 123I-mIBG scintigraphy to predict risk for adverse cardiac outcomes in heart failure patients: design of two prospective multicenter international trials. *J Nucl Cardiol* 2009; **16**: 113-121 [PMID: 19152136 DOI: 10.1007/s12350-008-9008-2]
154 **Nagahara D**, Nakata T, Hashimoto A, Wakabayashi T, Kyuma M, Noda R, Shimoshige S, Uno K, Tsuchihashi K, Shimamoto K. Predicting the need for an implantable cardioverter defibrillator using cardiac metaiodobenzylguanidine activity together with plasma natriuretic peptide concentration or left ventricular function. *J Nucl Med* 2008; **49**: 225-233 [PMID: 18199625 DOI: 10.2967/jnumed.107.042564]
155 **Tamaki S**, Yamada T, Okuyama Y, Morita T, Sanada S, Tsukamoto Y, Masuda M, Okuda K, Iwasaki Y, Yasui T, Hori M, Fukunami M. Cardiac iodine-123 metaiodobenzylguanidine imaging predicts sudden cardiac death independently of left ventricular ejection fraction in patients with chronic heart failure and left ventricular systolic dysfunction: results from a comparative study with signal-averaged electrocardiogram, heart rate variability, and QT dispersion. *J Am Coll Cardiol* 2009; **53**: 426-435 [PMID: 19179201 DOI: 10.1016/j.jacc.2008.10.025]
156 **Shah AM**, Bourgoun M, Narula J, Jacobson AF, Solomon SD. Influence of ejection fraction on the prognostic value of sympathetic innervation imaging with iodine-123 MIBG in heart failure. *JACC Cardiovasc Imaging* 2012; **5**: 1139-1146 [PMID: 23153914 DOI: 10.1016/j.jcmg.2012.02.019]
157 **Sood N**, Al Badarin F, Parker M, Pullatt R, Jacobson AF, Bateman TM, Heller GV. Resting perfusion MPI-SPECT combined with cardiac 123I-mIBG sympathetic innervation imaging improves prediction of arrhythmic events in non-ischemic cardiomyopathy patients: sub-study from the ADMIRE-HF trial. *J Nucl Cardiol* 2013; **20**: 813-820 [PMID: 23864400 DOI: 10.1007/s12350-013-9750-y]
158 **García-González P**, Cozar-Santiago P, Fabregat-Andrés O, Sánchez-Jurado R, Estornell-Erill J, Ridocci-Soriano F. Evaluation of Cardiac (123)I-MIBG Imaging in Patients With Severe Left Ventricular Dysfunction and Indication for Implantable Cardioverter Defibrillator. *Rev Esp Cardiol* 2013; **66**: 1000-1002 [PMID: 24041808 DOI: 10.1016/j.recesp.2013.06.009]
159 **Hayashi M**, Kobayashi Y, Morita N, Iwasaki YK, Ohmura K, Atarashi H, Katoh T, Takano T. Clinical significance and contributing factors of long-term variability in induced ventricular tachyarrhythmias. *J Cardiovasc Electrophysiol* 2003; **14**: 1049-1056 [PMID: 14521657 DOI: 10.1046/j.1540-8167.2003.03164.x]
160 **Mitrani RD**, Klein LS, Miles WM, Hackett FK, Burt RW, Wellman HN, Zipes DP. Regional cardiac sympathetic denervation in patients with ventricular tachycardia in the absence of coronary artery disease. *J Am Coll Cardiol* 1993; **22**: 1344-1353 [PMID: 8227790 DOI: 10.1016/0735-1097(93)90541-8]
161 **Calkins H**, Allman K, Bolling S, Kirsch M, Wieland D, Morady F, Schwaiger M. Correlation between scintigraphic evidence of regional sympathetic neuronal dysfunction and ventricular refractoriness in the human heart. *Circulation* 1993; **88**: 172-179 [PMID: 8319330 DOI: 10.1161/01.CIR.88.1.172]
162 **Fallavollita JA**, Luisi AJ, Michalek SM, Valverde AM, deKemp RA, Haka MS, Hutson AD, Canty JM. Prediction of arrhythmic events with positron emission tomography: PAREPET study design and methods. *Contemp Clin Trials* 2006; **27**: 374-388 [PMID: 16647885 DOI: 10.1016/j.cct.2006.03.005]
163 **Fallavollita JA**, Heavey BM, Luisi AJ, Michalek SM, Baldwa S, Mashtare TL, Hutson AD, Dekemp RA, Haka MS, Sajjad M, Cimato TR, Curtis AB, Cain ME, Canty JM. Regional myocardial sympathetic denervation predicts the risk of sudden cardiac arrest in ischemic cardiomyopathy. *J Am Coll Cardiol* 2014; **63**: 141-149 [PMID: 24076296 DOI: 10.1016/j.jacc.2013.07.096]
164 **Bengel FM**, Thackeray JT. Altered cardiac innervation predisposes to ventricular arrhythmia: targeted positron emission tomography identifies risk in ischemic cardiomyopathy. *J Am Coll Cardiol* 2014; **63**: 150-152 [PMID: 24076279 DOI: 10.1016/j.jacc.2013.08.1629]

**P- Reviewers**: Chen SJ, Liu T **S- Editor**: Ma YJ
**L- Editor**: Roemmele A **E- Editor**: Wang CH

World Journal of Meta-Analysis

Online Submissions: http://www.wjgnet.com/esps/
bpgoffice@wjgnet.com
www.wjgnet.com

*World J Meta-Anal* 2014 February 26; 2(1): I-V
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Co., Limited. All rights reserved.

*INSTRUCTIONS TO AUTHORS*

## GENERAL INFORMATION

*World Journal of Meta-Analysis* (*World J Meta-Anal*, *WJMA*, online ISSN 2308-3840, DOI: 10.13105) is a peer-reviewed open access (OA) academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians, with a specific focus on meta-analysis, systematic review, mixed-treatment comparison, meta-regression, overview of reviews.

### *Aim and scope*

*WJMA* covers a variety of clinical medical fields including allergy, anesthesiology, cardiac medicine, clinical genetics, clinical neurology, critical care, dentistry, dermatology, emergency medicine, endocrinology, family medicine, gastroenterology and hepatology, geriatrics and gerontology, hematology, immunology, infectious diseases, internal medicine, obstetrics and gynecology, oncology, ophthalmology, orthopedics, otolaryngology, pathology, pediatrics, peripheral vascular disease, psychiatry, radiology, rehabilitation, respiratory medicine, rheumatology, surgery, toxicology, transplantation, and urology and nephrology, while maintaining its unique dedication to systematic reviews and meta-analyses.

We encourage authors to submit their manuscripts to *WJMA*. We will give priority to manuscripts that are supported by major national and international foundations and those that are of great basic and clinical significance.

*WJMA* is edited and published by Baishideng Publishing Group (BPG). BPG has a strong professional editorial team composed of science editors, language editors and electronic editors. BPG currently publishes 43 OA clinical medical journals, including 42 in English, has a total of 15471 editorial board members or peer reviewers, and is a world first-class publisher.

### *Columns*

The columns in the issues of *WJMA* will include: (1) Editorial: The editorial board members are invited to make comments on an important topic in their field in terms of its current research status and future directions to lead the development of this discipline; (2) Frontier: The editorial board members are invited to select a highly cited cutting-edge original paper of his/her own to summarize major findings, the problems that have been resolved and remain to be resolved, and future research directions to help readers understand his/her important academic point of view and future research directions in the field; (3) Field of Vision: The editorial board members are invited to write commentaries on classic articles, hot topic articles, or latest articles to keep readers at the forefront of research and increase their levels of clinical research. Classic articles refer to papers that are included in Web of Knowledge and have received a large number of citations (ranking in the top 1%) after being published for more than 2 years, reflecting the quality and impact of papers. Hot topic articles refer to papers that are included in Web of Knowledge and have received a large number of citations after being published for no more than 2 years, reflecting cutting-edge trends in scientific research. Latest articles refer to the latest published high-quality papers that are included in PubMed, reflecting the latest research trends. These commentary articles should focus on the status quo of research, the most important research topics, the problems that have now been resolved and remain to be resolved, and future research directions. Basic information about the article to be commented will be provided as well (including authors, article title, journal name, year, volume, and inclusive page numbers; (4) Minireviews: The editorial board members are invited to write short reviews on recent advances and trends in research to provide readers; (5) Review: To make a systematic review to focus on the status quo of research, the most important research topics, the problems that have now been resolved and remain to be resolved, and future research directions; (6) Topic Highlight: The editorial board members are invited to write a series of articles (7-10 articles) to comment and discuss a hot topic; (7) Meta-Analysis: Covers the systematic review, mixed-treatment comparison, meta-regression, and overview of reviews. To summarize a given quantitative effect, *e.g.*, the clinical effectiveness and safety of clinical treatments by combining data from two or more randomized controlled trials, thereby providing more precise and externally valid estimates than those which would stem from each individual dataset if analyzed separately from the others; (8) Medical Ethics: The editorial board members are invited to write articles about medical ethics to increase readers' knowledge of medical ethics. The topic covers international ethics guidelines, animal studies, clinical trials, organ transplantation, *etc.*; (9) Letters to the Editor: To discuss and make reply to the contributions published in *WJMA*, or to introduce and comment on a controversial issue of general interest; (10) Book Reviews: To introduce and comment on quality monographs; and (11) Autobiography: The editorial board members are invited to write their autobiography to provide readers with stories of success or failure in their scientific research career. The topic covers their basic personal information and information about when they started doing research work, where and how they did research work, what they have achieved, and their lessons from success or failure.

### *Name of journal*

*World Journal of Meta-Analysis*

### *ISSN*

ISSN 2308-3840 (online)

### *Launch date*

May 26, 2013

### *Frequency*

Quarterly

### *Editors-in-Chief*

**Lesley A Anderson, PhD, Lecturer, PhD, MPHe, BSc (Hons), PGCHET, FHEA,** Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, Northern Ireland BT7 1NN, United Kingdom

**Giuseppe Biondi-Zoccai, MD, Assistant Professor,** Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

### *Editorial office*

Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
*World Journal of Meta-Analysis*

Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: bpgoffice@wjgnet.com
http://www.wjgnet.com

***Publisher***
Baishideng Publishing Group Co., Limited
Flat C, 23/F., Lucky Plaza, 315-321 Lockhart Road,
Wan Chai, Hong Kong, China
Telephone: +852-58042046
Fax: +852-31158812
E-mail: bpgoffice@wjgnet.com
http://www.wjgnet.com

***Production center***
Beijing Baishideng BioMed Scientific Co., Limited
Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381892
Fax: +86-10-85381893

***BPG US INC***
Baishideng Publishing Group Inc
8226 Regency Drive, Pleasanton,
CA 94588, United States

***Instructions to authors***
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm.

***Indexed and Abstracted in***
Digital Object Identifier.

## SPECIAL STATEMENT

All articles published in this journal represent the viewpoints of the authors except where indicated otherwise.

***Biostatistical editing***
Statistical review is performed after peer review. We invite an expert in Biomedical Statistics to evaluate the statistical method used in the paper, including *t*-test (group or paired comparisons), chi-squared test, Ridit, probit, logit, regression (linear, curvilinear, or stepwise), correlation, analysis of variance, analysis of covariance, *etc*. The reviewing points include: (1) Statistical methods should be described when they are used to verify the results; (2) Whether the statistical techniques are suitable or correct; (3) Only homogeneous data can be averaged. Standard deviations are preferred to standard errors. Give the number of observations and subjects (*n*). Losses in observations, such as drop-outs from the study should be reported; (4) Values such as ED50, LD50, IC50 should have their 95% confidence limits calculated and compared by weighted probit analysis (Bliss and Finney); and (5) The word 'significantly' should be replaced by its synonyms (if it indicates extent) or the *P* value (if it indicates statistical significance).

***Conflict-of-interest statement***
In the interests of transparency and to help reviewers assess any potential bias, *WJMA* requires authors of all papers to declare any competing commercial, personal, political, intellectual, or religious interests in relation to the submitted work. Referees are also asked to indicate any potential conflict they might have reviewing a particular paper. Before submitting, authors are suggested to read "Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Ethical Considerations in the Conduct and Reporting of Research: Conflicts of Interest" from International Committee of Medical Journal Editors (ICMJE), which is available at: http://www.icmje.org/ethical_4conflicts.html.

Sample wording: [Name of individual] has received fees for serving as a speaker, a consultant and an advisory board member for [names of organizations], and has received research funding from [names of organization]. [Name of individual] is an employee of [name of organization]. [Name of individual] owns stocks and shares in [name of organization]. [Name of individual] owns patent [patent identification and brief description].

***Statement of informed consent***
Manuscripts should contain a statement to the effect that all human studies have been reviewed by the appropriate ethics committee or it should be stated clearly in the text that all persons gave their informed consent prior to their inclusion in the study. Details that might disclose the identity of the subjects under study should be omitted. Authors should also draw attention to the Code of Ethics of the World Medical Association (Declaration of Helsinki, 1964, as revised in 2004).

***Statement of human and animal rights***
When reporting the results from experiments, authors should follow the highest standards and the trial should conform to Good Clinical Practice (for example, US Food and Drug Administration Good Clinical Practice in FDA-Regulated Clinical Trials; UK Medicines Research Council Guidelines for Good Clinical Practice in Clinical Trials) and/or the World Medical Association Declaration of Helsinki. Generally, we suggest authors follow the lead investigator's national standard. If doubt exists whether the research was conducted in accordance with the above standards, the authors must explain the rationale for their approach and demonstrate that the institutional review body explicitly approved the doubtful aspects of the study.

Before submitting, authors should make their study approved by the relevant research ethics committee or institutional review board. If human participants were involved, manuscripts must be accompanied by a statement that the experiments were undertaken with the understanding and appropriate informed consent of each. Any personal item or information will not be published without explicit consents from the involved patients. If experimental animals were used, the materials and methods (experimental procedures) section must clearly indicate that appropriate measures were taken to minimize pain or discomfort, and details of animal care should be provided.

## SUBMISSION OF MANUSCRIPTS

Manuscripts should be typed in 1.5 line spacing and 12 pt. Book Antiqua with ample margins. Number all pages consecutively, and start each of the following sections on a new page: Title Page, Abstract, Introduction, Materials and Methods, Results, Discussion, Acknowledgements, References, Tables, Figures, and Figure Legends. Neither the editors nor the publisher are responsible for the opinions expressed by contributors. Manuscripts formally accepted for publication become the permanent property of Baishideng Publishing Group Co., Limited, and may not be reproduced by any means, in whole or in part, without the written permission of both the authors and the publisher. We reserve the right to copy-edit and put onto our website accepted manuscripts. Authors should follow the relevant guidelines for the care and use of laboratory animals of their institution or national animal welfare committee. For the sake of transparency in regard to the performance and reporting of clinical trials, we endorse the policy of the ICMJE to refuse to publish papers on clinical trial results if the trial was not recorded in a publicly-accessible registry at its outset. The only register now available, to our knowledge, is http://www.clinicaltrials.gov sponsored by the United States National Library of Medicine and we encourage all potential contributors to register with it. However, in the case that other registers become available you will be duly notified. A letter of recommendation from each author's organization should be provided with the contributed article to ensure the privacy and secrecy of research is protected.

Authors should retain one copy of the text, tables, photographs and illustrations because rejected manuscripts will not be returned to the author(s) and the editors will not be responsible for loss or damage to photographs and illustrations sustained during mailing.

***Online submissions***
Manuscripts should be submitted through the Online Submission System at: http://www.wjgnet.com/esps/. Authors are highly recommended to consult the ONLINE INSTRUCTIONS TO AUTHORS (http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm) before attempting to submit online. For assistance, authors encountering problems with the Online Submission System may send an email describing the problem to bpgoffice@wjgnet.com, or by telephone: +86-10-85381892. If you submit your manuscript online, do not make a postal contribution. Repeated online submission for the same manuscript is strictly prohibited.

## MANUSCRIPT PREPARATION

All contributions should be written in English. All articles must be submitted using word-processing software. All submissions must be typed in 1.5 line spacing and 12 pt. Book Antiqua with ample margins. Style should conform to our house format. Required information for each of the manuscript sections is as follows:

***Title page***
**Title:** Title should be less than 12 words.

**Running title:** A short running title of less than 6 words should be provided.

**Authorship:** Authorship credit should be in accordance with the standard proposed by ICMJE, based on (1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; and (3) final approval of the version to be published. Authors should meet conditions 1, 2, and 3.

**Institution:** Author names should be given first, then the complete name of institution, city, province and postcode. For example, Xu-Chen Zhang, Li-Xin Mei, Department of Pathology, Chengde Medical College, Chengde 067000, Hebei Province, China. One author may be represented from two institutions, for example, George Sgourakis, Department of General, Visceral, and Transplantation Surgery, Essen 45122, Germany; George Sgourakis, 2nd Surgical Department, Korgialenio-Benakio Red Cross Hospital, Athens 15451, Greece

**Author contributions:** The format of this section should be: Author contributions: Wang CL and Liang L contributed equally to this work; Wang CL, Liang L, Fu JF, Zou CC, Hong F and Wu XM designed the research; Wang CL, Zou CC, Hong F and Wu XM performed the research; Xue JZ and Lu JR contributed new reagents/analytic tools; Wang CL, Liang L and Fu JF analyzed the data; and Wang CL, Liang L and Fu JF wrote the paper.

**Supportive foundations:** The complete name and number of supportive foundations should be provided, *e.g.*, Supported by National Natural Science Foundation of China, No. 30224801

**Correspondence to:** Only one corresponding address should be provided. Author names should be given first, then author title, affiliation, the complete name of institution, city, postcode, province, country, and email. All the letters in the email should be in lower case. A space interval should be inserted between country name and email address. For example, Montgomery Bissell, MD, Professor of Medicine, Chief, Liver Center, Gastroenterology Division, University of California, Box 0538, San Francisco, CA 94143, United States. montgomery.bissell@ucsf.edu

**Telephone and fax:** Telephone and fax should consist of +, country number, district number and telephone or fax number, *e.g.*, Telephone: +86-10-85381892 Fax: +86-10-85381893

**Peer reviewers:** All articles received are subject to peer review. Normally, three experts are invited for each article. Decision on acceptance is made only when at least two experts recommend publication of an article. All peer-reviewers are acknowledged on Express Submission and Peer-review System website.

***Abstract***
There are unstructured abstracts (no less than 200 words) and structured abstracts. The specific requirements for structured abstracts are as follows:

An informative, structured abstract should accompany each manuscript. Abstracts of original contributions should be structured into the following sections: AIM (no more than 20 words; Only the purpose of the study should be included. Please write the Aim in the form of "To investigate/study/…"), METHODS (no less than 140 words for Original Articles; and no less than 80 words for Brief Articles), RESULTS (no less than 150 words for Original Articles and no less than 120 words for Brief Articles; You should present *P* values where appropriate and must provide relevant data to illustrate how they were obtained, *e.g.*, $6.92 \pm 3.86$ *vs* $3.61 \pm 1.67$, $P < 0.001$), and CONCLUSION (no more than 26 words).

***Key words***
Please list 5-10 key words, selected mainly from *Index Medicus*, which reflect the content of the study.

***Core tip***
Please write a summary of less than 100 words to outline the most innovative and important arguments and core contents in your paper to attract readers.

***Text***
For articles of these sections, original articles and brief articles, the main text should be structured into the following sections: INTRODUCTION, MATERIALS AND METHODS, RESULTS and DISCUSSION, and should include appropriate Figures and Tables. Data should be presented in the main text or in Figures and Tables, but not in both.

***Illustrations***
Figures should be numbered as 1, 2, 3, *etc.*, and mentioned clearly in the main text. Provide a brief title for each figure on a separate page. Detailed legends should not be provided under the figures. This part should be added into the text where the figures are applicable. Keeping all elements compiled is necessary in line-art image. Scale bars should be used rather than magnification factors, with the length of the bar defined in the legend rather than on the bar itself. File names should identify the figure and panel. Avoid layering type directly over shaded or textured areas. Please use uniform legends for the same subjects. For example: Figure 1 Pathological changes in atrophic gastritis after treatment. A: ...; B: ...; C: ...; D: ...; E: ...; F: ...; G: …*etc.* It is our principle to publish high resolution-figures for the E-versions.

***Tables***
Three-line tables should be numbered 1, 2, 3, *etc.*, and mentioned clearly in the main text. Provide a brief title for each table. Detailed legends should not be included under tables, but rather added into the text where applicable. The information should complement, but not duplicate the text. Use one horizontal line under the title, a second under column heads, and a third below the Table, above any footnotes. Vertical and italic lines should be omitted.

***Notes in tables and illustrations***
Data that are not statistically significant should not be noted. $^{a}P < 0.05$, $^{b}P < 0.01$ should be noted ($P > 0.05$ should not be noted). If there are other series of *P* values, $^{c}P < 0.05$ and $^{d}P < 0.01$ are used. A third series of *P* values can be expressed as $^{e}P < 0.05$ and $^{f}P < 0.01$. Other notes in tables or under illustrations should be expressed as $^{1}$F, $^{2}$F, $^{3}$F; or sometimes as other symbols with a superscript (Arabic numerals) in the upper left corner. In a multi-curve illustration, each curve should be labeled with ●, ○, ■, □, ▲, △, *etc.*, in a certain sequence.

***Acknowledgments***
Brief acknowledgments of persons who have made genuine contributions to the manuscript and who endorse the data and conclusions should be included. Authors are responsible for obtaining written permission to use any copyrighted text and/or illustrations.

## REFERENCES

### *Coding system*

The author should number the references in Arabic numerals according to the citation order in the text. Put reference numbers in square brackets in superscript at the end of citation content or after the cited author's name. For citation content which is part of the narration, the coding number and square brackets should be typeset normally. For example, "Crohn's disease (CD) is associated with increased intestinal permeability[1,2]". If references are cited directly in the text, they should be put together within the text, for example, "From references[19,22-24], we know that..."

When the authors write the references, please ensure that the order in text is the same as in the references section, and also ensure the spelling accuracy of the first author's name. Do not list the same citation twice.

### *PMID and DOI*

Pleased provide PubMed citation numbers to the reference list, *e.g.*, PMID and DOI, which can be found at http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed and http://www.crossref.org/SimpleTextQuery/, respectively. The numbers will be used in E-version of this journal.

### *Style for journal references*

Authors: the name of the first author should be typed in bold-faced letters. The family name of all authors should be typed with the initial letter capitalized, followed by their abbreviated first and middle initials. (For example, Lian-Sheng Ma is abbreviated as Ma LS, Bo-Rong Pan as Pan BR). The title of the cited article and italicized journal title (journal title should be in its abbreviated form as shown in PubMed), publication date, volume number (in black), start page, and end page [PMID: 11819634 DOI: 10.3748/wjg.13.5396].

### *Style for book references*

Authors: the name of the first author should be typed in bold-faced letters. The surname of all authors should be typed with the initial letter capitalized, followed by their abbreviated middle and first initials. (For example, Lian-Sheng Ma is abbreviated as Ma LS, Bo-Rong Pan as Pan BR) Book title. Publication number. Publication place: Publication press, Year: start page and end page.

### *Format*

**Journals**

*English journal article (list all authors and include the PMID where applicable)*

1 **Jung EM**, Clevert DA, Schreyer AG, Schmitt S, Rennert J, Kubale R, Feuerbach S, Jung F. Evaluation of quantitative contrast harmonic imaging to assess malignancy of liver tumors: A prospective controlled two-center study. *World J Gastroenterol* 2007; **13**: 6356-6364 [PMID: 18081224 DOI: 10.3748/wjg.13.6356]

*Chinese journal article (list all authors and include the PMID where applicable)*

2 **Lin GZ**, Wang XZ, Wang P, Lin J, Yang FD. Immunologic effect of Jianpi Yishen decoction in treatment of Pixu-diarrhoea. *Shijie Huaren Xiaohua Zazhi* 1999; **7**: 285-287

*In press*

3 **Tian D**, Araki H, Stahl E, Bergelson J, Kreitman M. Signature of balancing selection in Arabidopsis. *Proc Natl Acad Sci USA* 2006; In press

*Organization as author*

4 **Diabetes Prevention Program Research Group**. Hypertension, insulin, and proinsulin in participants with impaired glucose tolerance. *Hypertension* 2002; **40**: 679-686 [PMID: 12411462 PMCID:2516377 DOI:10.1161/01.HYP.0000035706.28494.09]

*Both personal authors and an organization as author*

5 **Vallancien G**, Emberton M, Harving N, van Moorselaar RJ; Alf-One Study Group. Sexual dysfunction in 1, 274 European men suffering from lower urinary tract symptoms. *J Urol* 2003; **169**: 2257-2261 [PMID: 12771764 DOI:10.1097/01.ju.0000067940.76090.73]

*No author given*

6 21st century heart solution may have a sting in the tail. *BMJ* 2002; **325**: 184 [PMID: 12142303 DOI:10.1136/bmj.325.7357.184]

*Volume with supplement*

7 **Geraud G**, Spierings EL, Keywood C. Tolerability and safety of frovatriptan with short- and long-term use for treatment of migraine and in comparison with sumatriptan. *Headache* 2002; **42** Suppl 2: S93-99 [PMID: 12028325 DOI:10.1046/j.1526-4610.42.s2.7.x]

*Issue with no volume*

8 **Banit DM**, Kaufer H, Hartford JM. Intraoperative frozen section analysis in revision total joint arthroplasty. *Clin Orthop Relat Res* 2002; (**401**): 230-238 [PMID: 12151900 DOI:10.1097/00003086-200208000-00026]

*No volume or issue*

9 Outreach: Bringing HIV-positive individuals into care. *HRSA Careaction* 2002; 1-6 [PMID: 12154804]

**Books**

*Personal author(s)*

10 **Sherlock S**, Dooley J. Diseases of the liver and billiary system. 9th ed. Oxford: Blackwell Sci Pub, 1993: 258-296

*Chapter in a book (list all authors)*

11 **Lam SK**. Academic investigator's perspectives of medical treatment for peptic ulcer. In: Swabb EA, Azabo S. Ulcer disease: investigation and basis for therapy. New York: Marcel Dekker, 1991: 431-450

*Author(s) and editor(s)*

12 **Breedlove GK**, Schorfheide AM. Adolescent pregnancy. 2nd ed. Wieczorek RR, editor. White Plains (NY): March of Dimes Education Services, 2001: 20-34

*Conference proceedings*

13 **Harnden P**, Joffe JK, Jones WG, editors. Germ cell tumours V. Proceedings of the 5th Germ cell tumours Conference; 2001 Sep 13-15; Leeds, UK. New York: Springer, 2002: 30-56

*Conference paper*

14 **Christensen S**, Oppacher F. An analysis of Koza's computational effort statistic for genetic programming. In: Foster JA, Lutton E, Miller J, Ryan C, Tettamanzi AG, editors. Genetic programming. EuroGP 2002: Proceedings of the 5th European Conference on Genetic Programming; 2002 Apr 3-5; Kinsdale, Ireland. Berlin: Springer, 2002: 182-191

**Electronic journal** (list all authors)

15 Morse SS. Factors in the emergence of infectious diseases. Emerg Infect Dis serial online, 1995-01-03, cited 1996-06-05; 1(1): 24 screens. Available from: URL: http://www.cdc.gov/ncidod/eid/index.htm

**Patent** (list all authors)

16 **Pagedas AC**, inventor; Ancel Surgical R&D Inc., assignee. Flexible endoscopic grasping and cutting device and positioning tool assembly. United States patent US 20020103498. 2002 Aug 1

### *Statistical data*

Write as mean ± SD or mean ± SE.

### *Statistical expression*

Express *t* test as *t* (in italics), *F* test as *F* (in italics), chi square test as $\chi^2$ (in Greek), related coefficient as *r* (in italics), degree of freedom as υ (in Greek), sample number as *n* (in italics), and probability as *P* (in italics).

### *Units*

Use SI units. For example: body mass, m (B) = 78 kg; blood pressure, p (B) = 16.2/12.3 kPa; incubation time, t (incubation) = 96 h, blood glucose concentration, c (glucose) 6.4 ± 2.1 mmol/L; blood CEA mass concentration, p (CEA) = 8.6 24.5 μg/L; $CO_2$ volume fraction, 50 mL/L $CO_2$, not 5% $CO_2$; likewise for 40 g/L formaldehyde, not 10% formalin; and mass fraction, 8 ng/g, *etc.* Arabic numerals such as 23, 243, 641 should be read 23 243 641.

The format for how to accurately write common units and quantums can be found at: http://www.wjgnet.com/2308-3840/g_info_20100725073806.htm.

***Abbreviations***

Standard abbreviations should be defined in the abstract and on first mention in the text. In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Permissible abbreviations are listed in Units, Symbols and Abbreviations: A Guide for Biological and Medical Editors and Authors (Ed. Baron DN, 1988) published by The Royal Society of Medicine, London. Certain commonly used abbreviations, such as DNA, RNA, HIV, LD50, PCR, HBV, ECG, WBC, RBC, CT, ESR, CSF, IgG, ELISA, PBS, ATP, EDTA, mAb, can be used directly without further explanation.

***Italics***

Quantities: *t* time or temperature, *c* concentration, *A* area, *l* length, *m* mass, *V* volume.
Genotypes: *gyrA*, *arg* 1, *c myc*, *c fos*, *etc.*
Restriction enzymes: *Eco*RI, *Hin*dI, *Bam*HI, *Kbo* I, *Kpn* I, *etc.*
Biology: *H. pylori*, *E coli*, *etc.*

***Examples for paper writing***

All types of articles' writing style and requirement will be found in the link: http://www.wjgnet.com/esps/NavigationInfo.aspx?id=15

## RESUBMISSION OF THE REVISED MANUSCRIPTS

Authors must revise their manuscript carefully according to the revision policies of Baishideng Publishing Group Co., Limited. The revised version, along with the signed copyright transfer agreement, responses to the reviewers, and English language Grade A certificate (for non-native speakers of English), should be submitted to the online system *via* the link contained in the e-mail sent by the editor. If you have any questions about the revision, please send e-mail to esps@wjgnet.com.

***Language evaluation***

The language of a manuscript will be graded before it is sent for revision. (1) Grade A: priority publishing; (2) Grade B: minor language polishing; (3) Grade C: a great deal of language polishing needed; and (4) Grade D: rejected. Revised articles should reach Grade A.

***Copyright assignment form***

Please download a Copyright assignment form from http://www.wjgnet.com/2308-3840/g_info_20100725073726.htm.

***Responses to reviewers***

Please revise your article according to the comments/suggestions provided by the reviewers. The format for responses to the reviewers' comments can be found at: http://www.wjgnet.com/2308-3840/g_info_20100725073445.htm.

***Proof of financial support***

For papers supported by a foundation, authors should provide a copy of the approval document and serial number of the foundation.

***Statement about anonymous publication of the peer reviewers' comments***

In order to increase the quality of peer review, push authors to carefully revise their manuscripts based on the peer reviewers' comments, and promote academic interactions among peer reviewers, authors and readers, we decide to anonymously publish the reviewers' comments and author's responses at the same time the manuscript is published online.

## PUBLICATION FEE

*WJMA* is an international, peer-reviewed, OA online journal. Articles published by this journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium and format, provided the original work is properly cited. The use is non-commercial and is otherwise in compliance with the license. Authors of accepted articles must pay a publication fee. Publication fee: 600 USD per article. All invited articles are published free of charge.

**ISSN** 2308-3840 (online)

# World Journal of *Meta-Analysis*

*World J Meta-Anal* 2014 May 26; 2(2): 17-48



Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

## Contents

Quarterly Volume 2 Number 2 May 26, 2014

**REVIEW**

17 Varicocele: How this condition and its management affects men's health
*Baazeem A*

**MINIREVIEWS**

24 Prognostic significance of post percutaneous coronary intervention thrombocytopenia
*Schiariti M, Iannetta L, Torromeo C, De Gregorio M, Puddu PE*

**META-ANALYSIS**

29 Weekday *vs* weekend endoscopy: Is there a true difference in patient outcomes? A meta-analysis
*Hinds AM, Ahmad D, Lopez KT, Matteson-Kome ML, Bechtold ML*

36 Oerview and a meta-analysis of pancreatic melanoma based on case reports
*Yuan ZN, Liu FH, Tang Y*

42 Treatment strategy for gallstone pancreatitis and the timing of cholecystectomy
*Hu C, Shen SQ, Chen ZB*

**Contents**

*World Journal of Meta-Analysis*
**Volume 2 Number 2 May 26, 2014**

**APPENDIX** I-V Instructions to authors

**ABOUT COVER** Editorial Board Member of *World Journal of Meta-Analysis*, Dr. Wen-Peng Cui, PhD, Department of Nephrology, Second Hospital of Jilin University, Changchun 130041, Jilin Province, China

**AIM AND SCOPE**

*World Journal of Meta-Analysis* (*World J Meta-Anal*, *WJMA*, online ISSN 2308-3840, DOI: 10.13105) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians, with a specific focus on meta-analysis, systematic review, mixed-treatment comparison, meta-regression, overview of reviews.

*WJMA* covers a variety of clinical medical fields including allergy, anesthesiology, cardiac medicine, clinical genetics, clinical neurology, critical care, dentistry, dermatology, emergency medicine, endocrinology, family medicine, gastroenterology and hepatology, geriatrics and gerontology, hematology, immunology, infectious diseases, internal medicine, obstetrics and gynecology, oncology, ophthalmology, orthopedics, otolaryngology, pathology, pediatrics, peripheral vascular disease, psychiatry, radiology, rehabilitation, respiratory medicine, rheumatology, surgery, toxicology, transplantation, and urology and nephrology, while maintaining its unique dedication to systematic reviews and meta-analyses.

**INDEXING/ABSTRACTING** *World Journal of Meta-Analysis* is now indexed in Digital Object Identifier.

**FLYLEAF** I-IV Editorial Board

**EDITORS FOR THIS ISSUE**

**Responsible Assistant Editor:** *Xiang Li*
**Responsible Electronic Editor:** *Huan-Liang Wu*
**Proofing Editor-in-Chief:** *Lian-Sheng Ma*
**Responsible Science Editor:** *Fang-Fang Ji*
**Proofing Editorial Office Director:** *Xiu-Xia Song*

NAME OF JOURNAL
*World Journal of Meta-Analysis*

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Quarterly

EDITORS-IN-CHIEF
**Lesley A Anderson, PhD, Lecturer, PhD, MPHe, BSc (Hons), PGCHET, FHEA,** Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, Northern Ireland BT7 1NN, United Kingdom

**Giuseppe Biondi-Zoccai, MD, Assistant Professor,** Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
*World Journal of Meta-Analysis*
Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
May 26, 2014

COPYRIGHT
© 2014 Baishideng Publishing Group Inc. Articles published by this Open-Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

SPECIAL STATEMENT
All articles published in journals owned by the Baishideng Publishing Group (BPG) represent the views and opinions of their authors, and not the views, opinions or policies of the BPG, except where otherwise explicitly indicated.

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i2.17

*World J Meta-Anal* 2014 May 26; 2(2): 17-23
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

REVIEW

# Varicocele: How this condition and its management affects men's health

Abdulaziz Baazeem

Abdulaziz Baazeem, Department of Surgery, Umm Al-Qura University, Makkah 24382, Saudi Arabia
Author contributions: Baazeem A solely contributed to this paper.
Correspondence to: Dr. Abdulaziz Baazeem, MD, FRCS(C), FECSM, Assistant Professor of Urology, Department of Surgery, Umm Al-Qura University, PO Box 7607, Makkah 24382, Saudi Arabia. theurologist@gmail.com
Telephone: +966-1252-81189 Fax: +966-1252-700004125
Received: November 29, 2013 Revised: January 21, 2014
Accepted: February 18, 2014
Published online: May 26, 2014

## Abstract

Varicocele is a relatively common condition that can impact men's health in various ways. Unfortunately, its prevalence and the availability of various different methods of repairing it might lead to unnecessary treatment. An understanding of the various ways that this condition can impact men's health is necessary in order to manage it appropriately. At present, there is substantial evidence to support varicocele repair in men who present with infertility, abnormal semen parameters, clinical varicocele and a female partner with normal fertility (or one with a potentially correctable abnormality). Varicocele repair appears to improve seminal fluid quality and might improve pregnancy rates. It might also have a role in managing men with non-obstructive azoospermia. Varicocele can also be a cause of scrotal pain that is usually of a dull character. Varicocele repair is an effective method of managing this type of pain, especially once proper measures have been taken to exclude other possible causes of orchalgia. Conservative measures are generally not effective in managing varicocele-related scrotal pain. There is growing evidence to suggest that varicocele repair might have a role in improving the serum testosterone level in men with hypogonadism, especially in the subfertile population. Well-designed prospective studies are needed to support the utilization of varicocele repair in managing these patients, as well as in preventing testicular dysfunction on a prophylactic basis.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Varicocele; Varicocelectomy; Infertility; Azoospermia; Pain

**Core tip:** The prevalence of varicocele and the availability of various different methods of repairing it might lead to unnecessary treatment. An understanding of the various ways that this condition can impact men's health is necessary in order to manage it appropriately. An overview of this disorder is given, together with different indications of treating it in adult men.

Baazeem A. Varicocele: How this condition and its management affects men's health. *World J Meta-Anal* 2014; 2(2): 17-23 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i2/17.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i2.17

## INTRODUCTION

A varicocele is a tortuous spermatic vein that is abnormally dilated. It is a relatively common condition. A clinical varicocele has been described to be palpable in around 15% of all men[1]. However, in a more recent study, it was found to be present in 24.2% of healthy young men[2]. Furthermore, its prevalence has been described to be as high as about 35% of men who present for assessment of infertility[3]. Varicocele repair is currently considered to be the most commonly performed procedure to treat male infertility. This relatively high prevalence of varicocele, coupled with the availability of a multitude of relatively low-risk methods of repair, might lead to overtreating it[4]. The aim of this review is to examine how varicocele can affect men's health and the theories on how it exerts these effects. In addition, the potential out-

comes of repairing varicocele for various indications will be assessed.

## ASSOCIATIONS AND ETIOLOGY

Varicocele occurs predominantly on the left side. Several factors have been proposed to explain this disproportion in laterality, including the more frequent occurrence on the left side of absent or incompetent spermatic vein valves, a longer spermatic vein, retro-aortic left renal veins, the nut-cracker phenomenon (the compression of the left renal vein between the superior mesenteric artery and the aorta) as well as the insertion of the spermatic vein into the higher pressure left renal vein[5-8].

There appears to be an association between varicocele and conditions such as peripheral varicose veins and coronary artery ectasia, suggesting a diffuse vascular abnormality[9,10]. Its prevalence appears to rise with advancing age, increasing by about 10% with each decade[11]. The prevalence in the elderly was described to reach 42% in a study that involved men with a mean age of 60.7 years[12]. A higher prevalence of varicocele has been described in first-degree relatives, suggesting a hereditary component as well[13,14]. Varicocele also appears to be positively correlated with adolescent gynecomastia and inversely related to the body mass index[15,16].

## DIAGNOSIS AND CLASSIFICATION

Varicocele is typically diagnosed by physical examination. The classical description of it is that of a "bag of worms" when the patient is examined in the standing position, in a warm room. Varicoceles are usually classified according to the Dubin system[17]: Grade Ⅰ: palpable distension of intrascrotal veins only with the Valsalva maneuver; Grade Ⅱ: palpable, but not visible, distension of the intrascrotal veins without the Valsalva maneuver; Grade Ⅲ: distention of intrascrotal veins that is visible through the scrotal skin.

"Grade 0" is sometimes used to classify subclinical varicoceles that are not palpable, but rather are detected using ancillary diagnostic measures (*e.g.*, ultrasonography, venography, thermography and scintigraphy). Ultrasonography is one of the most popular of these ancillary measures and there are described systems of classifying varicocele using it[18]. Scintigraphy and thermography have been suggested as methods that are comparable or even superior to ultrasonography in detecting varicocele[19,20]. With the exception of venography, these techniques are considered simple and non-invasive.

Advocates of using ultrasonography contend that it helps avoid the risk of subjectivity that is associated with physical examination. However, no clear cut-off measurements have been agreed upon to establish the diagnosis of varicocele on ultrasonography[21]. Therefore, while it continues to be used widely to diagnose and classify varicocele, and is certainly indicated in cases where physical examination is inconclusive, some authorities caution that such procedures that detect subclinical varicoceles should be avoided[22].

## INDICATIONS FOR REPAIR

Varicocele repair in adults has been described to address several conditions. These include: (1) male infertility, including non-obstructive azoospermia (NOA); (2) pain; and (3) low serum testosterone.

## VARICOCELE AND MALE FACTOR INFERTILITY

In 1992, the World Health Organization assessed the influence of varicocele in more than nine thousand men who presented as partners in couples consulting for infertility across 24 countries[23]. A clinically palpable varicocele was detected in 11.7% of men with normal semen analyses, compared to 25.4% in those with abnormal analyses. Varicocele was associated with a reduced left testicular size and semen count. Interestingly, there was a wide range of varicocele detection among different centers (6%-47%), with the presence of a urologist or andrologist helping increase the detection rate. Also noteworthy is that the spontaneous pregnancy rate was not significantly different between men with and without varicocele.

The exact mechanism by which varicocele may negatively affect male fertility potential remains undetermined. However, several mechanisms that may contribute to this detrimental effect have been proposed. These include: (1) reflux of toxic adrenal and renal metabolites through the renal vein[24]; (2) elevated intrascrotal temperatures[25]; (3) elevated spermatozoal oxygen species[26,27]; (4) local hypoxia and ischemia[28,29]; (5) sperm chromosome aneuploidy[30]; and (6) sperm nuclear DNA damage[27,31].

Varicocele repair can be considered when a man has a clinically palpable varicocele, abnormal semen analysis results and his female partner has normal (or correctable) fertility potential[22]. The role of varicocele repair in managing men with male factor infertility has been a matter of heavy debate. A meta-analysis of randomized controlled trials (RCTs) to assess the influence of varicocele repair on pregnancy rates concluded that the treatment of varicocele in managing men from couples with unexplained infertility cannot be recommended[32]. This study was criticized for including men with subclinical varicocele as well as those with normal semen parameters. Several meta-analyses proposed a contrary perspective. A meta-analysis of the same RCTs in the above-mentioned study, after exclusion of those including men with subclinical varicocele and/or normal semen parameters showed that pregnancy rates in the treatment group were significantly higher in the treatment group on "as-treated" analysis[33]. A meta-analysis that included RCTs and observational studies that were scored for bias concluded that surgical repair of varicocele significantly improved the odds of spontaneous pregnancy[34]. An update of the aforementioned Cochrane Collaboration meta-analysis

found pregnancy outcomes that favored varicocele treatment over observation[35]. These results were even more pronounced when men with subclinical varicocele or normal semen parameters were excluded from analysis. However, the authors emphasized, as in most of these meta-analyses, the heterogeneity of the studies involved. Finally, a meta-analysis of RCTs involving only surgical repair of varicocele compared to observation found an even higher odds ratio (OR) of 4.15 favoring surgical repair after excluding men with normal semen parameters and subclinical varicocele[36].

Although pregnancy is the obvious desired principal outcome when managing men with infertility, there is a disadvantage to using it when assessing the effect of therapeutic interventions involving the male partner. This outcome inherently involves potential female partner factors that might impact the observed effect of treatment. Therefore, the role of semen analysis as a predictor of male fertility potential is probably just as important as that of pregnancy rates when assessing the benefit of varicocele repair on male fertility. There is evidence to suggest an improvement in sperm concentration, motility and morphology following varicocele repair[37-39]. According to a meta-analysis involving 22 prospective studies that addressed sperm concentration before and after varicocele repair, the combined post-repair improvement in sperm concentration postoperatively was 12.32 $\times 10^6$ sperm/mL[40]. Furthermore, in the same study, the combined improvements in total and progressive sperm motility were 10.86% and 10.67%, respectively. All of these findings were statistically significant. Significant improvements in seminal oxidative stress and sperm DNA damage following varicocele repair were also found in this meta-analysis. Sperm ultramorphology has also been shown to improve following surgical varicocele repair in a non-randomized prospective controlled study[41].

A review that compared the various methods of varicocele repair in more than 5000 patients concluded that the microsurgical approaches (subinguinal or inguinal) offer better outcomes in terms of pregnancy, recurrence and complication rates[42]. Retrograde radiographic-guided embolization or sclerotherapy carries with it several potential advantages. It minimizes the risk of post-procedure hydrocele formation and of testicular artery injury and subsequent testicular atrophy. There is also no risk of injury to the vas deferens and the recovery time is rapid. However, it also comes with its unique set of potential complications, including those related to intravenous contrast media administration, puncture site complications (*e.g.*, hematoma formation) and coil migration. A recent review of 158 patients who underwent varicocele embolization revealed an 18.9% failure rate when treating bilateral varicocele[43]. This relatively high failure rate was mainly due to technical difficulties when treating the right side. This brought the authors to suggest a microsurgical approach when treating bilateral varicocele. However, embolization remains a suitable option when treating left-sided varicocele.

## VARICOCELE AND NON-OBSTRUCTIVE AZOOSPERMIA

Patients with NOA represent possibly the most challenging patients with male factor infertility in terms of achieving pregnancy. They might require sperm retrieval procedures which, if successful, would be followed by assisted reproductive techniques (ARTs), such as *in vitro* fertilization (IVF)/intracytoplasmic sperm injection (ICSI). The aim when managing these patients would be to help achieve a return of sperm to the ejaculate with enough quality that would help achieve the goal of spontaneous pregnancy or improve semen quality in order to "upgrade" the couple to be considered for intra-uterine insemination, or help the male partner avoid sperm retrieval procedures by restoring sperm in the ejaculate for use in IVF/ICSI, or at least potentially improve spermatogenesis to a degree that might hopefully improve the success of future sperm retrieval attempts.

The return of sperm to the ejaculate following varicocele repair has been described in several series of men with a clinical varicocele and NOA, with success rates ranging between 20.8%-56.2%[44-48]. A meta-analysis that involved 233 patients from 11 retrospective studies found that motile sperm returned to the ejaculate in 39.1% following varicocele repair[49]. Fourteen patients (6%) achieved spontaneous pregnancy and 10 (4%) IVF/ICSI pregnancies. After an initial return of sperm to the ejaculate, 11 patients (4.6%) reverted back to azoospermia within 2-6 mo. When available, the histopathological picture can help predict the outcomes of varicocele repair, with hypospermatogenesis and maturation arrest associated with significantly higher success rates than Sertoli cell-only. Several studies have found an increase in sperm retrieval procedure success rates following varicocele repair when compared to proceeding with retrieval attempts without repairing the varicocele[50-52]. However, a retrospective study that included men with subclinical varicocele showed similar sperm retrieval rates between the two groups[48].

## PROPHYLACTIC VARICOCELECTOMY

Progressive, duration-dependant, deterioration in semen quality has been described in infertile men with a clinical varicocele who originally had normal semen parameters and were followed over time, albeit in a small-sized study[53]. Furthermore, the incidence of varicocele in men with secondary infertility is significantly higher than those with primary infertility, suggesting that prior fertility does not offer protection from possible varicocele-induced impairment of testicular function[54]. This potentially deleterious effect of varicocele on male fertility, coupled with the relative safety and low complication rates of modern microsurgical procedures, has led to consideration by some of varicocele repair in men with normal semen parameters and a clinically palpable varicocele, especially if there is ipsilateral testicular hypotrophy. However, there

is not enough evidence to strongly support the use of this theory in the clinical setting. No RCTs or prospective studies have been conducted to address the use of prophylactic varicocelectomy in this setting.

## VARICOCELE AND SCROTAL PAIN

Varicocele has been reported as a cause of scrotal pain[55]. The character of the pain has been described as dull and one that is aggravated by exercise and straining[56]. It has been estimated to occur in 2%-14% of men with varicocele[57]. Conservative measures that have been traditionally used to manage varicocele-associated pain include utilizing scrotal support, non-steroidal anti-inflammatory analgesics and limiting physical activity which are impractical and of limited benefit[56,58]. In a study of 140 patients who were treated conservatively for up to 8 wk, only 5 patients had an improvement in symptoms, and even those 5 patients eventually opted for surgical management for recurrent pain[58]. A low success rate (0.04%) for conservative management was also seen in another study when it was attempted for a mean of 4 wk[59]. There is some evidence, based on the experience with a small number of patients, to suggest a possible benefit from the venotonic drug micronized purified flavonoid fraction in reducing varicocele-associated pain in men with normal sperm concentration[60].

Studies, the vast majority of which are retrospective, that assess the effect of varicocele repair on pain are relatively few. Most of these series showed a relatively high rate of post-procedure complete or partial pain resolution, ranging between 76.7%-94.8%[56-59,61-67]. However, one of the earliest studies to address this issue described failure to resolve pain in 52% of the patients who underwent high ligation[55]. To date, there are no studies that assess the effect of subclinical varicocele repair on scrotal pain.

Other potential causes of scrotal pain should be ruled out prior to treating varicocele in order to help reduce treatment failures. The character of pain experienced by the patient can provide a clue to the cause of the patient's symptoms. Atypical pain properties such as sharp pain might be due to causes other than varicocele. Indeed, the presence of a dull aching or dragging pain has been reported as a predictor of treatment success, while sharp or throbbing pain was associated with a higher rate of recurrence, persistence or even worsening of symptoms[62]. In another study, post-operative success was significantly and independently higher in patients who were operated on for dull rather than "dragging" or "aching" pain and for mild or moderate rather than severe pain ($\geqslant$ 7 on the 11-point numeric rating scale)[68]. Yaman *et al*[59] suggested that the grade of a painful varicocele might also influence the outcome of repair, as did Kim *et al*[57]. However, this association was not observed in other studies[56,62,66,68]. Further support for the notion that a cause other than varicocele might be a reason behind treatment failure is noted in the finding that failure rates were significantly higher in patients with a duration of symptoms of less than 3 mo[66], but this association with symptom duration was not supported by findings in other studies[56,68]. In one of the few prospective studies that assessed varicocele repair for pain management, neither preoperative varicocele grade, pain character or duration of symptoms were predictive of post-operative outcomes[64]. In this study, among the patients who were available for their 3 mo post-operative assessment, 95% stated that they would recommend the procedure for their relatives.

Various techniques have been reported when treating varicocele for pain, including high ligation, subinguinal and inguinal microsurgery, conventional inguinal (Ivanissevich), scrotal and laparoscopic approaches, as well as percutaneous interventional radiological procedures such as sclerotherapy or embolization with reasonable results. Microsurgical approaches might have a higher post-operative pain resolution rate[69]. No difference in outcome was detected in a study that compared non-microsurgical approaches (inguinal, subinguinal and high ligation)[62].

Failure of varicocele resolution and its recurrence after treatment have been reported with varying rates. This might be a cause of persistent or recurrent scrotal pain following treatment[57]. Chawla *et al*[70] reported a series of 11 patients who presented with recurrent varicocele and pain and were treated with microsurgical subinguinal varicocelectomy. Ten of the patients had complete or partial resolution of pain and nine of them stated that they would undergo the same procedure for their presenting complaint.

## VARICOCELE AND HYPOGONADISM

Although the potential effect of varicocele and varicocelectomy on serum testosterone levels is one that has been studied in the past, this topic is garnering increasing interest more recently[71]. A possible dysfunction in testosterone production in infertile men has been described[72]. Tanrikut *et al*[73] described significantly lower serum testosterone levels in men with varicocele than a control group of men undergoing vasectomy reversal. The effect of varicocele repair on serum testosterone levels has been debated. Some studies showed no increase in serum testosterone levels following repair[74-76]. However, the baseline serum testosterone level in these studies was normal or somewhat higher, making it difficult to document an improvement following repair.

A meta-analysis involving 9 studies (one prospective and 8 retrospective observational studies) in which men with clinical varicocele underwent surgical repair showed a significant post-operative increase in serum testosterone from baseline by 97.48 ng/dL[77]. In a retrospective review of men who underwent microsurgical subinguinal varicocelectomy, men who had a baseline serum testosterone level of 400 ng/dL or less had a significant increase in their levels after surgery in all age groups[78]. In an interesting non-randomized prospective study, a hundred infertile men with clinical varicocele and low serum testosterone (< 280 ng/dL) who underwent microsurgi-

cal subinguinal repair were compared to a control group of hypogonadal varicocele patients who opted for ART utilization[79]. While there was a slight decrease in serum testosterone level in the control group at follow-up, the varicocelectomy group showed a significant increase in serum testosterone. Indeed, 78% of the men in the treatment arm went on to reach normal serum testosterone levels at follow-up, compared to only 16% in the control group. Similarly, serum testosterone and erectile function significantly improved from baseline after microsurgical subinguinal varicocelectomy in hypogonadal (serum testosterone < 300 ng/dL) infertile men with clinical varicocele when compared to a similar control group who refused surgery and opted for ARTs[80]. It is worth emphasizing that most of these studies involved men who also complained of infertility and may not represent the general population. Findings in this group of patients would be difficult to extrapolate to the general population.

## REFERENCES


1 **Clarke BG**. Incidence of varicocele in normal men and among men of different ages. *JAMA* 1966; **198**: 1121-1122 [PMID: 5953394 DOI: 10.1001/jama.1966.03110230137039]
2 **Soylemez H**, Atar M, Ali Sancaktutar A, Bozkurt Y, Penbegul N. Varicocele among healthy young men in Turkey; prevalence and relationship with body mass index. *Int Braz J Urol* 2012; **38**: 116-121 [PMID: 22397773 DOI: 10.1590/S1677-55382012000100016]
3 **Greenberg SH**, Lipshultz LI, Wein AJ. Experience with 425 subfertile male patients. *J Urol* 1978; **119**: 507-510 [PMID: 25971]
4 **Zargooshi J**. Sperm count and sperm motility in incidental high-grade varicocele. *Fertil Steril* 2007; **88**: 1470-1473 [PMID: 17451695 DOI: 10.1016/j.fertnstert.2007.01.016]
5 **Karazincir S**, Balci A, Görür S, Sumbas H, Kiper AN. Incidence of the retroaortic left renal vein in patients with varicocele. *J Ultrasound Med* 2007; **26**: 601-604 [PMID: 17460002]
6 **Wishahi MM**. Detailed anatomy of the internal spermatic vein and the ovarian vein. Human cadaver study and operative spermatic venography: clinical aspects. *J Urol* 1991; **145**: 780-784 [PMID: 2005700]
7 **Beinart C**, Sniderman KW, Tamura S, Vaughan ED, Sos TA. Left renal vein to inferior vena cava pressure relationship in humans. *J Urol* 1982; **127**: 1070-1071 [PMID: 7087011]
8 **Carl P**, Stark L, Ouzoun N, Reindl P. Venous pressure in idiopathic varicocele. *Eur Urol* 1993; **24**: 214-220 [PMID: 8375442]
9 **Kiliç S**, Aksoy Y, Sincer I, Oğuz F, Erdil N, Yetkin E. Cardiovascular evaluation of young patients with varicocele. *Fertil Steril* 2007; **88**: 369-373 [PMID: 17307174 DOI: 10.1016/j.fertnstert.2006.11.119]
10 **Yetkin E**, Kilic S, Acikgoz N, Ergin H, Aksoy Y, Sincer I, Aktürk E, Beytur A, Sivri N, Turhan H. Increased prevalence of varicocele in patients with coronary artery ectasia. *Coron Artery Dis* 2005; **16**: 261-264 [PMID: 16000882 DOI: 10.1097/00019501-200508000-00001]
11 **Levinger U**, Gornish M, Gat Y, Bachar GN. Is varicocele prevalence increasing with age? *Andrologia* 2007; **39**: 77-80 [PMID: 17683466 DOI: 10.1111/j.1439-0272.2007.00766.x]
12 **Canales BK**, Zapzalka DM, Ercole CJ, Carey P, Haus E, Aeppli D, Pryor JL. Prevalence and effect of varicoceles in an elderly population. *Urology* 2005; **66**: 627-631 [PMID: 16140091 DOI: 10.1016/j.urology.2005.03.062]
13 **Mokhtari G**, Pourreza F, Falahatkar S, Kamran AN, Jamali M. Comparison of prevalence of varicocele in first-degree relatives of patients with varicocele and male kidney donors. *Urology* 2008; **71**: 666-668 [PMID: 18279919 DOI: 10.1016/j.urology.2007.11.116]
14 **Raman JD**, Walmsley K, Goldstein M. Inheritance of varicoceles. *Urology* 2005; **65**: 1186-1189 [PMID: 15913726 DOI: 10.1016/j.urology.2004.12.057]
15 **Nielsen ME**, Zderic S, Freedland SJ, Jarow JP. Insight on pathogenesis of varicoceles: relationship of varicocele and body mass index. *Urology* 2006; **68**: 392-396 [PMID: 16904459 DOI: 10.1016/j.urology.2006.02.005]
16 **Kumanov P**, Deepinder F, Robeva R, Tomova A, Li J, Agarwal A. Relationship of adolescent gynecomastia with varicocele and somatometric parameters: a cross-sectional study in 6200 healthy boys. *J Adolesc Health* 2007; **41**: 126-131 [PMID: 17659215 DOI: 10.1016/j.jadohealth.2007.03.010]
17 **Dubin L**, Amelar RD. Varicocele size and results of varicocelectomy in selected subfertile men with varicocele. *Fertil Steril* 1970; **21**: 606-609 [PMID: 5433164]
18 **Hirsh AV**, Cameron KM, Tyler JP, Simpson J, Pryor JP. The Doppler assessment of varicoceles and internal spermatic vein reflux in infertile men. *Br J Urol* 1980; **52**: 50-56 [PMID: 7426951 DOI: 10.1111/j.1464-410X.1980.tb02919.x]
19 **Gat Y**, Bachar GN, Zukerman Z, Belenky A, Gorenish M. Physical examination may miss the diagnosis of bilateral varicocele: a comparative study of 4 diagnostic modalities. *J Urol* 2004; **172**: 1414-1417 [PMID: 15371858 DOI: 10.1097/01.ju.0000138540.57137.5f]
20 **Ortapamuk H**, Tekdogan UY, Naldoken S, Bulut S, Atan A. Hemodynamic evaluation of varicocele: the role of scrotal scintigraphy and Doppler ultrasonography in the prediction of postoperative seminal improvement. *Ann Nucl Med* 2005; **19**: 529-534 [PMID: 16363616 DOI: 10.1007/BF02985044]
21 **Lee J**, Binsaleh S, Lo K, Jarvi K. Varicoceles: the diagnostic dilemma. *J Androl* 2008; **29**: 143-146 [PMID: 18077824 DOI: 10.2164/jandrol.107.003467]
22 **Practice Committee of American Society for Reproductive Medicine**. Report on varicocele and infertility. *Fertil Steril* 2008; **90**: S247-S249 [PMID: 19007639 DOI: 10.1016/j.fertnstert.2008.08.050]
23 The influence of varicocele on parameters of fertility in a large group of men presenting to infertility clinics. World Health Organization. *Fertil Steril* 1992; **57**: 1289-1293 [PMID: 1601152]
24 **Comhaire F**, Vermeulen A. Varicocele sterility: cortisol and catecholamines. *Fertil Steril* 1974; **25**: 88-95 [PMID: 4810205]
25 **Zorgniotti AW**, Macleod J. Studies in temperature, human semen quality, and varicocele. *Fertil Steril* 1973; **24**: 854-863 [PMID: 4742006]
26 **Hendin BN**, Kolettis PN, Sharma RK, Thomas AJ, Agarwal A. Varicocele is associated with elevated spermatozoal reactive oxygen species production and diminished seminal plasma antioxidant capacity. *J Urol* 1999; **161**: 1831-1834 [PMID: 10332447 DOI: 10.1016/S0022-5347(05)68818-0]
27 **Smith R**, Kaune H, Parodi D, Madariaga M, Rios R, Morales I, Castro A. Increased sperm DNA damage in patients with varicocele: relationship with seminal oxidative stress. *Hum Reprod* 2006; **21**: 986-993 [PMID: 16361286 DOI: 10.1093/humrep/dei429]
28 **Donohue RE**, Brown JS. Blood gases and pH determinations in the internal spermatic veins of subfertile men with varicocele. *Fertil Steril* 1969; **20**: 365-369 [PMID: 5767279]
29 **Chakraborty J**, Hikim AP, Jhunjhunwala JS. Stagnation of blood in the microcirculatory vessels in the testes of men with varicocele. *J Androl* 1985; **6**: 117-126 [PMID: 3988623]
30 **Baccetti BM**, Bruni E, Capitani S, Collodel G, Mancini S, Piomboni P, Moretti E. Studies on varicocele III: ultrastructural sperm evaluation and 18, X and Y aneuploidies. *J Androl* 2006; **27**: 94-101 [PMID: 16400084 DOI: 10.2164/jandrol.05081]

31 **Saleh RA**, Agarwal A, Sharma RK, Said TM, Sikka SC, Thomas AJ. Evaluation of nuclear DNA damage in spermatozoa from infertile men with varicocele. *Fertil Steril* 2003; **80**: 1431-1436 [PMID: 14667879 DOI: 10.1016/S0015-0282(03)02211-8]

32 **Evers JH**, Collins J, Clarke J. Surgery or embolisation for varicoceles in subfertile men. *Cochrane Database Syst Rev* 2009; **(1)**: CD000479 [PMID: 19160180]

33 **Ficarra V**, Cerruto MA, Liguori G, Mazzoni G, Minucci S, Tracia A, Gentile V. Treatment of varicocele in subfertile men: The Cochrane Review--a contrary opinion. *Eur Urol* 2006; **49**: 258-263 [PMID: 16426727 DOI: 10.1016/j.eururo.2005.11.023]

34 **Marmar JL**, Agarwal A, Prabakaran S, Agarwal R, Short RA, Benoff S, Thomas AJ. Reassessing the value of varicocelectomy as a treatment for male subfertility with a new meta-analysis. *Fertil Steril* 2007; **88**: 639-648 [PMID: 17434508 DOI: 10.1016/j.fertnstert.2006.12.008]

35 **Kroese AC**, de Lange NM, Collins J, Evers JL. Surgery or embolization for varicoceles in subfertile men. *Cochrane Database Syst Rev* 2012; **10**: CD000479 [PMID: 23076888]

36 **Kim KH**, Lee JY, Kang DH, Lee H, Seo JT, Cho KS. Impact of surgical varicocele repair on pregnancy rate in subfertile men with clinical varicocele and impaired semen quality: a meta-analysis of randomized clinical trials. *Korean J Urol* 2013; **54**: 703-709 [PMID: 24175046 DOI: 10.4111/kju.2013.54.10.703]

37 **Kibar Y**, Seckin B, Erduran D. The effects of subinguinal varicocelectomy on Kruger morphology and semen parameters. *J Urol* 2002; **168**: 1071-1074 [PMID: 12187225 DOI: 10.1016/S0022-5347(05)64577-6]

38 **Madgar I**, Weissenberg R, Lunenfeld B, Karasik A, Goldwasser B. Controlled trial of high spermatic vein ligation for varicocele in infertile men. *Fertil Steril* 1995; **63**: 120-124 [PMID: 7805900]

39 **Boman JM**, Libman J, Zini A. Microsurgical varicocelectomy for isolated asthenospermia. *J Urol* 2008; **180**: 2129-2132 [PMID: 18804226 DOI: 10.1016/j.juro.2008.07.046]

40 **Baazeem A**, Belzile E, Ciampi A, Dohle G, Jarvi K, Salonia A, Weidner W, Zini A. Varicocele and male factor infertility treatment: a new meta-analysis and review of the role of varicocele repair. *Eur Urol* 2011; **60**: 796-808 [PMID: 21733620 DOI: 10.1016/j.eururo.2011.06.018]

41 **Reichart M**, Eltes F, Soffer Y, Zigenreich E, Yogev L, Bartoov B. Sperm ultramorphology as a pathophysiological indicator of spermatogenesis in males suffering from varicocele. *Andrologia* 2000; **32**: 139-145 [PMID: 10863968 DOI: 10.1046/j.1439-0272.2000.00355.x]

42 **Diegidio P**, Jhaveri JK, Ghannam S, Pinkhasov R, Shabsigh R, Fisch H. Review of current varicocelectomy techniques and their outcomes. *BJU Int* 2011; **108**: 1157-1172 [PMID: 21435155 DOI: 10.1111/j.1464-410X.2010.09959.x]

43 **Cassidy D**, Jarvi K, Grober E, Lo K. Varicocele surgery or embolization: Which is better? *Can Urol Assoc J* 2012; **6**: 266-268 [PMID: 23093537]

44 **Kadioglu A**, Tefekli A, Cayan S, Kandirali E, Erdemir F, Tellaloglu S. Microsurgical inguinal varicocele repair in azoospermic men. *Urology* 2001; **57**: 328-333 [PMID: 11182347 DOI: 10.1016/S0090-4295(00)00908-0]

45 **Gat Y**, Bachar GN, Everaert K, Levinger U, Gornish M. Induction of spermatogenesis in azoospermic men after internal spermatic vein embolization for the treatment of varicocele. *Hum Reprod* 2005; **20**: 1013-1017 [PMID: 15618245 DOI: 10.1093/humrep/deh706]

46 **Matthews GJ**, Matthews ED, Goldstein M. Induction of spermatogenesis and achievement of pregnancy after microsurgical varicocelectomy in men with azoospermia and severe oligoasthenospermia. *Fertil Steril* 1998; **70**: 71-75 [PMID: 9660424 DOI: 10.1016/S0015-0282(98)00108-3]

47 **Pasqualotto FF**, Sobreiro BP, Hallak J, Pasqualotto EB, Lucon AM. Induction of spermatogenesis in azoospermic men after varicocelectomy repair: an update. *Fertil Steril* 2006; **85**: 635-639 [PMID: 16500331 DOI: 10.1016/j.fertnstert.2005.08.043]

48 **Schlegel PN**, Kaufmann J. Role of varicocelectomy in men with nonobstructive azoospermia. *Fertil Steril* 2004; **81**: 1585-1588 [PMID: 15193481 DOI: 10.1016/j.fertnstert.2003.10.036]

49 **Weedin JW**, Khera M, Lipshultz LI. Varicocele repair in patients with nonobstructive azoospermia: a meta-analysis. *J Urol* 2010; **183**: 2309-2315 [PMID: 20400156 DOI: 10.1016/j.juro.2010.02.012]

50 **Zampieri N**, Bosaro L, Costantini C, Zaffagnini S, Zampieri G. Relationship between testicular sperm extraction and varicocelectomy in patients with varicocele and nonobstructive azoospermia. *Urology* 2013; **82**: 74-77 [PMID: 23680120 DOI: 10.1016/j.urology.2013.03.037]

51 **Inci K**, Hascicek M, Kara O, Dikmen AV, Gürgan T, Ergen A. Sperm retrieval and intracytoplasmic sperm injection in men with nonobstructive azoospermia, and treated and untreated varicocele. *J Urol* 2009; **182**: 1500-1505 [PMID: 19683732 DOI: 10.1016/j.juro.2009.06.028]

52 **Haydardedeoglu B**, Turunc T, Kilicdag EB, Gul U, Bagis T. The effect of prior varicocelectomy in patients with nonobstructive azoospermia on intracytoplasmic sperm injection outcomes: a retrospective pilot study. *Urology* 2010; **75**: 83-86 [PMID: 19913887 DOI: 10.1016/j.urology.2009.09.023]

53 **Chehval MJ**, Purcell MH. Deterioration of semen parameters over time in men with untreated varicocele: evidence of progressive testicular damage. *Fertil Steril* 1992; **57**: 174-177 [PMID: 1730313]

54 **Gorelick JI**, Goldstein M. Loss of fertility in men with varicocele. *Fertil Steril* 1993; **59**: 613-616 [PMID: 8458466]

55 **Biggers RD**, Soderdahl DW. The painful varicocele. *Mil Med* 1981; **146**: 440-441 [PMID: 6787479]

56 **Peterson AC**, Lance RS, Ruiz HE. Outcomes of varicocele ligation done for pain. *J Urol* 1998; **159**: 1565-1567 [PMID: 9554356 DOI: 10.1097/00005392-199805000-00043]

57 **Kim SO**, Jung H, Park K. Outcomes of microsurgical subinguinal varicocelectomy for painful varicoceles. *J Androl* 2012; **33**: 872-875 [PMID: 22174387 DOI: 10.2164/jandrol.111.014993]

58 **Yeniyol CO**, Tuna A, Yener H, Zeyrek N, Tilki A. High ligation to treat pain in varicocele. *Int Urol Nephrol* 2003; **35**: 65-68 [PMID: 14620287]

59 **Yaman O**, Ozdiler E, Anafarta K, Göğüş O. Effect of microsurgical subinguinal varicocele ligation to treat pain. *Urology* 2000; **55**: 107-108 [PMID: 10654904 DOI: 10.1016/S0090-4295(99)00374-X]

60 **Söylemez H**, Kiliç S, Atar M, Penbegül N, Sancaktutar AA, Bozkurt Y. Effects of micronised purified flavonoid fraction on pain, semen analysis and scrotal color Doppler parameters in patients with painful varicocele; results of a randomized placebo-controlled study. *Int Urol Nephrol* 2012; **44**: 401-408 [PMID: 21805085 DOI: 10.1007/s11255-011-0038-3]

61 **Lee HJ**, Cheon SH, Ji YH, Moon KH, Kim KS, Park S, Park S. Clinical characteristics and surgical outcomes in adolescents and adults with varicocele. *Korean J Urol* 2011; **52**: 489-493 [PMID: 21860771 DOI: 10.4111/kju.2011.52.7.489]

62 **Al-Buheissi SZ**, Patel HR, Wazait HD, Miller RA, Nathan S. Predictors of success in surgical ligation of painful varicocele. *Urol Int* 2007; **79**: 33-36 [PMID: 17627165 DOI: 10.1159/000102910]

63 **Park HJ**, Lee SS, Park NC. Predictors of pain resolution after varicocelectomy for painful varicocele. *Asian J Androl* 2011; **13**: 754-758 [PMID: 21102471 DOI: 10.1038/aja.2010.87]

64 **Karademir K**, Senkul T, Baykal K, Ateş F, Işeri C, Erden D. Evaluation of the role of varicocelectomy including external spermatic vein ligation in patients with scrotal pain. *Int J Urol* 2005; **12**: 484-488 [PMID: 15948749 DOI: 10.1111/j.1442-2042.2005.01063.x]

65 **Ribé N**, Manasia P, Sàrquella J, Grimaldi S, Pomerol JM. Clinical follow-up after subinguinal varicocele ligation to treat pain. *Arch Ital Urol Androl* 2002; **74**: 51-53 [PMID:

12161934]
66 **Altunoluk B**, Soylemez H, Efe E, Malkoc O. Duration of preoperative scrotal pain may predict the success of microsurgical varicocelectomy. *Int Braz J Urol* 2010; **36**: 55-59 [PMID: 20202236 DOI: 10.1590/S1677-55382010000100009]
67 **Maghraby HA**. Laparoscopic varicocelectomy for painful varicoceles: merits and outcomes. *J Endourol* 2002; **16**: 107-110 [PMID: 11962551 DOI: 10.1089/089277902753619627]
68 **Kim HT**, Song PH, Moon KH. Microsurgical ligation for painful varicocele: effectiveness and predictors of pain resolution. *Yonsei Med J* 2012; **53**: 145-150 [PMID: 22187245 DOI: 10.3349/ymj.2012.53.1.145]
69 **Shridharani A**, Lockwood G, Sandlow J. Varicocelectomy in the treatment of testicular pain: a review. *Curr Opin Urol* 2012; **22**: 499-506 [PMID: 22965318 DOI: 10.1097/MOU.0b013e328358f69f]
70 **Chawla A**, Kulkarni G, Kamal K, Zini A. Microsurgical varicocelectomy for recurrent or persistent varicoceles associated with orchalgia. *Urology* 2005; **66**: 1072-1074 [PMID: 16286126 DOI: 10.1016/j.urology.2005.05.052]
71 **Su LM**, Goldstein M, Schlegel PN. The effect of varicocelectomy on serum testosterone levels in infertile men with varicoceles. *J Urol* 1995; **154**: 1752-1755 [PMID: 7563339 DOI: 10.1016/S0022-5347(01)66776-4]
72 **Weiss DB**, Rodriguez-Rigau LJ, Smith KD, Steinberger E. Leydig cell function in oligospermic men with varicocele. *J Urol* 1978; **120**: 427-430 [PMID: 702664]
73 **Tanrikut C**, Goldstein M, Rosoff JS, Lee RK, Nelson CJ, Mulhall JP. Varicocele as a risk factor for androgen deficiency and effect of repair. *BJU Int* 2011; **108**: 1480-1484 [PMID: 21435152 DOI: 10.1111/j.1464-410X.2010.10030.x]
74 **Di Bisceglie C**, Bertagna A, Baldi M, Lanfranco F, Tagliabue M, Gazzera C, Gandini G, Manieri C. Varicocele sclerotherapy improves serum inhibin B levels and seminal parameters. *Int J Androl* 2007; **30**: 531-536 [PMID: 17376219 DOI: 10.1111/j.1365-2605.2007.00747.x]
75 **Rodriguez Peña M**, Alescio L, Russell A, Lourenco da Cunha J, Alzu G, Bardoneschi E. Predictors of improved seminal parameters and fertility after varicocele repair in young adults. *Andrologia* 2009; **41**: 277-281 [PMID: 19737275 DOI: 10.1111/j.1439-0272.2009.00919.x]
76 **Zheng YQ**, Gao X, Li ZJ, Yu YL, Zhang ZG, Li W. Efficacy of bilateral and left varicocelectomy in infertile men with left clinical and right subclinical varicoceles: a comparative study. *Urology* 2009; **73**: 1236-1240 [PMID: 19371942 DOI: 10.1016/j.urology.2008.11.050]
77 **Li F**, Yue H, Yamaguchi K, Okada K, Matsushita K, Ando M, Chiba K, Fujisawa M. Effect of surgical repair on testosterone production in infertile men with varicocele: a meta-analysis. *Int J Urol* 2012; **19**: 149-154 [PMID: 22059526 DOI: 10.1111/j.1442-2042.2011.02890.x]
78 **Hsiao W**, Rosoff JS, Pale JR, Greenwood EA, Goldstein M. Older age is associated with similar improvements in semen parameters and testosterone after subinguinal microsurgical varicocelectomy. *J Urol* 2011; **185**: 620-625 [PMID: 21168880 DOI: 10.1016/j.juro.2010.09.114]
79 **Sathya Srini V**, Belur Veerachari S. Does varicocelectomy improve gonadal function in men with hypogonadism and infertility? Analysis of a prospective study. *Int J Endocrinol* 2011; **2011**: 916380 [PMID: 22190930]
80 **Zohdy W**, Ghazi S, Arafa M. Impact of varicocelectomy on gonadal and erectile functions in men with hypogonadism and infertility. *J Sex Med* 2011; **8**: 885-893 [PMID: 20722780 DOI: 10.1111/j.1743-6109.2010.01974.x]

**P- Reviewers:** Faintuch S, Kulis T **S- Editor:** Wen LL
**L- Editor:** Roemmele A **E- Editor:** Wu HL

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i2.24

*World J Meta-Anal* 2014 May 26; 2(2): 24-28
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.



# Prognostic significance of post percutaneous coronary intervention thrombocytopenia

Michele Schiariti, Loredana Iannetta, Concetta Torromeo, Michele De Gregorio, Paolo Emilio Puddu

Michele Schiariti, Loredana Iannetta, Concetta Torromeo, Paolo Emilio Puddu, Laboratory of Biotechnologies Applied to Cardiovascular Medicine, Department of Cardiovascular, Respiratory, Nephrological, Anesthesiological and Geriatrical Sciences, Sapienza University of Rome, 00161 Rome, Italy
Michele De Gregorio, Department of Medicine, St. Joseph Mercy Oakland Hospital, Pontiac, MI 48341, United States
Author contributions: All the authors contributed equally to this work.
Correspondence to: Paolo Emilio Puddu, MD, PhD, FESC, FACC, Laboratory of Biotechnologies Applied to Cardiovascular Medicine, Department of Cardiovascular, Respiratory, Nephrological, Anesthesiological and Geriatrical Sciences, Sapienza University of Rome, Viale del Policlinico, 155, 00161 Rome, Italy. paoloemilio.puddu@uniroma1.it
Telephone: +39-6-49972654 Fax: +39-6-4453891
Received: November 5, 2013 Revised: January 25, 2014
Accepted: February 18, 2014
Published online: May 26, 2014

## Abstract

Several definitions of post percutaneous coronary intervention (PCI) thrombocytopenia (TC) were formulated. Recent studies demonstrated that a relative drop in platelet count ⩾ 25% is the most appropriate criterion. By this definition a population is detected that is exposed not only to increased risk of hemorrhagic complications but also to increased risk of ischemic events, which may appear a paradox. In patients with acute coronary syndromes undergoing PCI, several conditions might be associated with TC: cardiopulmonary bypass and the presence of extra corporeal membrane oxygenators, intra aortic balloon pump (IABP), cardiogenic shock, thrombolytic drugs and anticoagulant or antiplatelet drugs. Several studies demonstrated that TC and ischemic outcomes are related although it is unclear whether this is a direct relationship or TC is just a secondary effect of another cryptic protagonist. It is suggested that further investigations determine whether there is a real link between TC, a probably well defined covariate, and ischemic outcomes or whether IABP is the joining link between these two variables and whose presence needs in any case be considered in multivariable statistics. Post-PCI TC could be only a secondary effect of IABP use. On turn, the prolonged use of heparin necessarily accompanying the use of IABP, and producing a paradoxical pro-thrombotic TC, might also be implicated.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Thrombocytopenia; Outcome; Percutaneous coronary intervention; Intra aortic balloon pump

**Core tip:** This minireview suggested that further investigations are needed to determine whether there is a real link between thrombocytopenia (TC), a probably well defined covariate, and ischemic outcomes or whether intra-aortic balloon (IABP) is the joining link between these two variables and whose presence needs in any case be considered in multivariable statistics. Post-percutaneous coronary intervention TC could be only a secondary effect of IABP use. On turn, the prolonged use of heparin necessarily accompanying the use of IABP, and producing a paradoxical pro-thrombotic TC, might also be implicated.

Schiariti M, Iannetta L, Torromeo C, De Gregorio M, Puddu PE. Prognostic significance of post percutaneous coronary intervention thrombocytopenia. *World J Meta-Anal* 2014; 2(2): 24-28 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i2/24.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i2.24

## THROMBOCYTOPENIA: DETERMINATION AND DEFINITION

Thrombocytopenia (TC) is a clinical condition related to decreased production, sequestration or increased de-

**Table 1 An overview of the studies whereby post-percutaneous coronary intervention thrombocytopenia was differently defined and analysed as an outcome variable**

| Ref. | Number of thrombocytopenic patients *n* (%) | Definition of thrombocytopenia | Mortality rate[1] | Hemorrhagic outcomes[1] | Composite Ischemic outcomes[1] |
|---|---|---|---|---|---|
| Berkowitz *et al*[10] | 81/2099 (3.9) | Nadir PLT count < $100 \times 10^9$/L | 12.1% *vs* 1.1% at 1 mo | 48.7% *vs* 9.5% at 1 mo | NSC |
| McClure *et al*[8] | 633/10948 (7.0) | Nadir PLT count < $100 \times 10^9$/L or delta PLT > 50% | 7.4% *vs* 3.3% at 1 mo | 75.8% *vs* 27.8% at 1 mo | 30% *vs* 14.1% at 1 mo |
| Eikelboom *et al*[3] | 87/10141 (1) | Nadir PLT count < $100 \times 10^9$/L | 11.50% at 6 mo | 8% *vs* 1.4% at 6 mo | 33.3% *vs* 30.8%[2] at 6 mo[3] |
| Merlini *et al*[4] | 99/4809 (2.9) | Nadir PLT count < $100 \times 10^9$/L or delta PLT ⩾ 25% | 2.0% *vs* 0.4% at 1 mo | 5.1% *vs* 0.7% | 12.2% *vs* 6.6% at 1 mo |
| Nikolsky *et al*[9] | 50/2082 (2.5) | Nadir PLT count < $100 \times 10^9$/L | 10% *vs* 3.9% at 12 mo | 10% *vs* 2.7% | 26.1% *vs* 17.4% at 12 mo |
| Shenoy *et al*[32] | 41/1302 (3.1) | Nadir PLT count < $100 \times 10^9$/L or delta PLT > 50% | 10% *vs* 2% at 6 mo | 37% *vs* 3% | 39% *vs* 15% at 6 mo |
| Wang *et al*[6] | 4697/36182 (13) | Nadir PLT count < $150 \times 10^9$/L and/or delta PLT > 50% | 6.9% *vs* 2.6% in hospital stay | 26.7% *vs* 10% | NC |
| De Labriolle *et al*[11] | 1644/10146 (16.95) | Delta PLT ⩾ 25% | 11% *vs* 4.2% at 12 mo | 21% *vs* 2.6% | 22.9% *vs* 16.01% at 12 mo |
| Caixeta *et al*[7] | 740/10836 (6.8) | Nadir PLT count < $150 \times 10^9$/L | 6.5% *vs* 3.4% at 12 mo | 14.0% *vs* 4.3% | 22.8% *vs* 15.1% at 12 mo |
| Kiviniemi *et al*[35] | 46/929 (9.7) | Nadir PLT count < $100 \times 10^9$/L or delta PLT > 50% | 12% *vs* 11%[2] at 12 mo | 23% *vs* 22%[2] at 12 mo | 24% *vs* 19%[2] at 12 mo |

[1]Thrombocytopenic *vs* non thrombocytopenic patients; [2]Not statistically different; [3]There was a significant difference at 7 d only. Inclusion and exclusion criteria were different in all studies. NC: Not considered; NSC: Not statistically considered; PCI: Percutaneous coronary intervention; PLT: Platelets.

struction of platelets. Isolated TC is more likely due to peripheral destruction[1]. In the setting of acute coronary syndromes (ACS) the use of heparin[2,3], platelet glycoprotein Ⅱb/Ⅲa receptor inhibitors[4] and intra-aortic balloon pump (IABP)[5] is associated to improved outcomes but, on the other hand, it exposes patients to higher risk of bleeding and decline in platelet count. When the magnitude of the platelet drop is large enough, one of the main definitions of "acquired" TC may be met. A summary of the main definitions is available in Table 1. Classically, acquired TC was defined as a decrease of platelet levels from normal baseline count (> $150 \times 10^9$/L) to thrombocytopenic levels (< $150 \times 10^9$/L)[6,7] or as a nadir platelet count < $100 \times 10^9$/L[8-10] during the in-hospital phase or as a ⩾ 50% drop in platelet count from pre-intervention platelet values[6,8].

More recently, De Labriolle *et al*[11] proposed a novel threshold of ⩾ 25% decrease in platelet count to define acquired TC. The latter definition is important to detect a population that is exposed not only to increased risk of hemorrhagic complications but also, as might be easily foreseen, to higher risk of ischemic events[11]. However, although the first correlation is intuitive, the second one seems a real paradox. The purpose of this review is to examine the subset of thrombocytopenic patients in the setting of ACS and to define the clinical relevance of TC for prognosis.

## POST-PCI THROMBOCYTOPENIA: POSSIBLE CAUSES

Several causes of TC are known but, in many instances, they cannot be identified[12]. In patients with ACS undergoing percutaneous coronary intervention (PCI), several conditions might be associated with TC.

### *Cardiopulmonary by-pass and extracorporeal membrane oxygenation*

These conditions may cause TC and platelet dysfunction related to the artificial membrane oxygenators used in these settings[1,13].

### *Intra aortic balloon pump*

Platelet count decrease is correlated primarily to the device itself, independently from the use of intravenous heparin in one study[5]. It is a predictable platelet count decrease characterized by distinct features: it rapidly recovers if the device is removed or it stabilizes after 4 d if IABP remains in place[5,14].

### *Cardiogenic shock*

Several studies demonstrated that TC is common in shocked patients[5,15]. This could be related to decreased bone marrow productive activity in severely ill patients, but these patients will also more likely receive an IABP[5,15,16]. Therefore some investigators in post-PCI TC *de facto* excluded shocked patients from the study population[4,9,11] and probably made a major selection bias.

### *Thrombolytic drugs*

In the setting of ACS, thrombolytic therapy is widely used since it improves mortality rate, although it is associated with TC and therefore to increased risk of hemorrhage and death[17]. Since this is a well-known cause of platelet count reduction, some studies in post-PCI TC excluded all patients undergone thrombolysis[8,9,11].

### *Heparin*

Since Silver, Rhodes and Dixon first identified central features of Heparin Induced TC (HIT) where TC and thrombosis have an immune pathogenesis, HIT has been

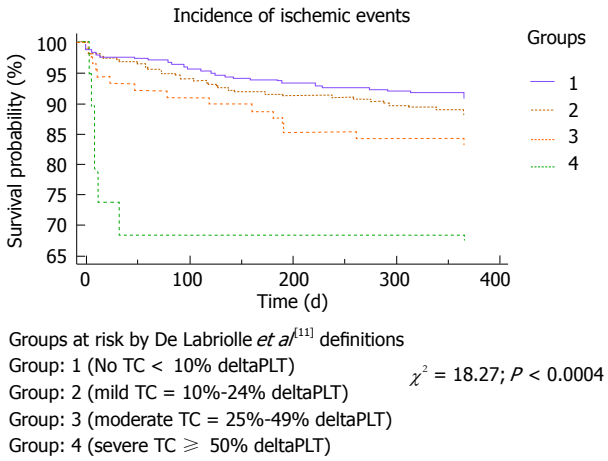



**Figure 1 Kaplan-Meier curves for the incidence of composite ischemic events during 1 yr among thrombocytopenic groups.** The ROC curves subdivided in accordance with De Labriolle *et al*[11], modified from Schiariti *et al*[34]. TC: Thrombocytopenia.

deeply investigated[18]. HIT is due to a peculiar heparin-dependent and platelet-activating antibody (IgG)[18,19] and seems related to heparin preparations[2] with typical features: it is usually moderate and more frequently begins later than 5-14 d (typical onset) after the drug administration, although it can occur after 7-40 d (delayed onset) or, in patients who received it in the previous 100 d, within 24 h (rapid onset)[18,20].

### *Glycoprotein Ⅱb/Ⅲa receptor inhibitors*

Patients undergone triple antiplatelet therapy based on glycoprotein Ⅱb/Ⅲa receptor inhibitors (GPI) have a better prognosis in the setting of ACS. However, based on a recent meta-analysis[21], GPI therapy seems to have no beneficial effect in elective PCI. The latter study does not clarify whether patients undergoing elective PCI were or not pretreated with a dual antiplatelet therapy before receiving a downstream GPI treatment[21]. GPI-related TC mostly occurs rapidly, within few hours from drug administration and, differing from HIT, it can be profound (platelet count < $20 \times 10^9$/L). It is related to the presence of drug-dependent antiplatelet antibodies[1,22,23]. This type of TC, independently from the specific GPI molecule administered, seems to follow a distinct clinical course when compared to other etiologies: it is more profound but has a quicker recovery[12,24,25]. Nevertheless, it seems also associated to increased rate of ischemic events[4].

### *P2Y-12 inhibitors*

Thienopyridines are a rare but possible cause of TC[26]. Ticlopidine-associated TC can occur 2-12 wk after its first administration, whereas clopidogrel-related TC is a rarer but earlier event[27,28].

### *Other causes*

In the setting of ACS other medications, particularly ACE-inhibitors and statins, can be a rare cause of TC[29,30]. On the other hand, TC has also to be differentiated from pseudo TC, a laboratory artifact resulting from platelet aggregation[31].

## FEATURES OF THROMBOCYTOPENIC PATIENTS

Thrombocytopenic patients have some peculiar features: they are more likely men[6-11], with a lower body mass index[6,8-11], diabetics[6,9] and with arterial hypertension[6]. Medical history of thrombocytopenic patients is typically characterized by previous myocardial infarction[6], previous PCI[6-8], previous stroke[6] and heart failure[6,11]. Looking at their home therapy they more likely are on statins and oral hypoglycemic therapy[9,11] and less likely treated with aspirin[9]. Looking at laboratory parameters, these patients have a lower baseline platelet count[6,7,9,11] and baseline renal insufficiency[7,11]. All these parameters should be therefore taken into account when defining the independent role of TC.

## POST-PCI THROMBOCYTOPENIA: CLINICAL OUTCOMES

Since its identification, post-PCI TC was classically related to hemorrhagic events[1,10,17,24]. Nevertheless, platelet count decline was also related to ischemic outcomes[3,4,7-9,11,32]. Several hypotheses were formulated to explain this relationship. Early discontinuation of antiplatelet agents for the onset of acute TC was considered causative of unfavorable clinical outcome[6,7], but other studies confuted this hypothesis[10]. It was also hypothesized that platelet consumption was associated with platelet activation, leading to downstream thrombotic events[6]. It was demonstrated that mean platelet volume was significantly increased despite concomitant decrease in platelet count in patients with ACS or stroke and animal models supported this, pointing to the fact that large platelets seem to have a higher thrombotic potential[33]. Moreover, TC was suspected to have a direct pathogenetic influence on adverse ischemic events[11] or it may be considered a simple marker of poor prognosis[6,32].

A clear-cut evolution in the definition of acquired TC and its relationship with ischemic outcomes may be seen between 1998 and 2013 (Table 1). Contrary to early studies defining TC only on the basis of absolute platelet count drop[7,9,10], more recent investigations clearly showed that post-PCI TC should be defined only based on relative drop in platelet count[11,34]. The different definitions of acquired TC are variously related to mortality and to ischemic or hemorrhagic outcomes (Table 1).

It is thus possible to identify individuals with substantial falls in platelet count, not reaching thrombocytopenic levels, that are at increased risk of ischemic events. We showed evidence[34] illustrated by Kaplan-Meyer curves (Figure 1) for a strong reverse relationship between platelet count[11] and the incidence of cumulative ischemic outcome. In 2013 Kiviniemi *et al*[35], assessing hemorrhagic and ischemic complications in patients on oral anticoagulation undergoing PCI, found a similar risk of adverse

outcome in patients with acquired TC compared with the control group. On the other hand, we indicated[34] that in the setting of TC there might be a cryptic protagonist: IABP. Indeed, in ACS, previous studies did not consider the important association with IABP itself, apart that from the concomitant presence of shock, among the determinants of TC[3,6,7] or interpreted it as a simple predictor of acquired TC[8,11].

Since there is a well known relationship between ischemic events and IABP use[36], the question is whether post-PCI TC is causative of ischemic events or whether the other protagonist (IABP) may have a con-causal (mechanic) role. Under these perspectives, cardiogenic shock should be considered a further etiologic factor for TC[5,14,16,34]. On the other hand, as IABP is necessarily associated with concomitant heparin utilization, it is for further study to assess and quantify the impact of prolonged heparin on TC and long-term ischemic outcome.

## IS POST-PCI THROMBOCYTOPENIA A COVARIATE?

An important question is whether TC, defined following the more recent investigations (Table 1) pointing to a $\geqslant$ 25% drop in platelet count from baseline to post-PCI, might be considered a covariate to be fed into multivariable statistics for prediction of ischemic or hemorrhagic outcomes. We tried to reply to this question by re-analyzing our data on 873 ACS patients treated by PCI[34] and of whom 867 had complete data (Figure 1): 107 (12%) of these were with TC according to the above-mentioned definition[11]. We then run a logistic regression in forced mode and the area under the ROC curve was 75% (log likelihood -267, model $r^2$ = 0.17) by fitting 13 covariates (age, gender, presence of shock, the occurrence of primary or rescue PCI, the number of vessels with significant coronary stenoses, the values of baseline platelet count, hemoglobin, red blood cells, creatine kinase, glomerular filtration rate by Modification of Diet in Renal Disease formula, absence of GPI inhibitors and presence of IABP). There were only 4 contributors to predict TC: age ($P$ = 0.06008), the occurrence of primary ($P$ = 0.00207) or rescue ($P$ =0.00098) PCI and the presence of IABP ($P$ = 0.00001) with a significant intercept ($P$ = 0.02168). From these data it might be concluded that TC is indeed a potential covariate, although in order to fully disclose its predictive role in post-PCI investigations aimed at assessing long-term impact on ischemic outcome, it is essential to fit other confounders and among these latter: age, IABP and the type of PCI performed.

## CONCLUSION

Post-PCI TC was classically related to increased mortality and incidence of bleeding and large evidence exists. On the other hand, TC was also related to higher risk of ischemic events. Although several hypotheses were formulated to define the latter relationship, no evidence proved that ischemic outcomes are related to TC *per se* or they might be just an ancillary phenomenon. It is suggested that further investigations determine whether there is a real link between TC, a probably well defined covariate, and ischemic outcomes or whether IABP is the joining link between these two variables. In any case presence of IABP should be considered in multivariable statistics. Post-PCI TC could be only a secondary effect of IABP use, possibly also related to the concomitant use of heparin, although precise quantification of this is lacking.

## REFERENCES


1 **Konkle BA**. Acquired disorders of platelet function. *Hematology Am Soc Hematol Educ Program* 2011; **2011**: 391-396 [PMID: 22160063 DOI: 10.1182/asheducation-2011.1.391]

2 **Bell WR**, Royall RM. Heparin-associated thrombocytopenia: a comparison of three heparin preparations. *N Engl J Med* 1980; **303**: 902-907 [PMID: 6997743 DOI: 10.1056/NEJM198010163031602]

3 **Eikelboom JW**, Anand SS, Mehta SR, Weitz JI, Yi C, Yusuf S. Prognostic significance of thrombocytopenia during hirudin and heparin therapy in acute coronary syndrome without ST elevation: Organization to Assess Strategies for Ischemic Syndromes (OASIS-2)study. *Circulation* 2001; **103**: 643-650 [PMID: 11156874 DOI: 10.1161/01.CIR.103.5.643]

4 **Merlini PA**, Rossi M, Menozzi A, Buratti S, Brennan DM, Moliterno DJ, Topol EJ, Ardissino D. Thrombocytopenia caused by abciximab or tirofiban and its association with clinical outcome in patients undergoing coronary stenting. *Circulation* 2004; **109**: 2203-2206 [PMID: 15117843 DOI: 10.1161/01.CIR.0000127867.41621.85]

5 **Vonderheide RH**, Thadhani R, Kuter DJ. Association of thrombocytopenia with the use of intra-aortic balloon pumps. *Am J Med* 1998; **105**: 27-32 [PMID: 9688018 DOI: 10.1016/S0002-9343(98)00128-4]

6 **Wang TY**, Ou FS, Roe MT, Harrington RA, Ohman EM, Gibler WB, Peterson ED. Incidence and prognostic significance of thrombocytopenia developed during acute coronary syndrome in contemporary clinical practice. *Circulation* 2009; **119**: 2454-2462 [PMID: 19398666 DOI: 10.1161/CIRCULATIONAHA.108.827162]

7 **Caixeta A**, Dangas GD, Mehran R, Feit F, Nikolsky E, Lansky AJ, Aoki J, Moses JW, Steinhubl SR, White HD, Ohman EM, Manoukian SV, Fahy M, Stone GW. Incidence and clinical consequences of acquired thrombocytopenia after antithrombotic therapies in patients with acute coronary syndromes: results from the Acute Catheterization and Urgent Intervention Triage Strategy (ACUITY) trial. *Am Heart J* 2011; **161**: 298-306.e1 [PMID: 21315212 DOI: 10.1016/j.ahj.2010.10.035]

8 **McClure MW**, Berkowitz SD, Sparapani R, Tuttle R, Kleiman NS, Berdan LG, Lincoff AM, Deckers J, Diaz R, Karsch KR, Gretler D, Kitt M, Simoons M, Topol EJ, Califf RM, Harrington RA. Clinical significance of thrombocytopenia during a non-ST-elevation acute coronary syndrome. The platelet glycoprotein IIb/IIIa in unstable angina: receptor suppression using integrilin therapy (PURSUIT) trial experience. *Circulation* 1999; **99**: 2892-2900 [PMID: 10359733 DOI: 10.1161/01.CIR.99.22.2892]

9 **Nikolsky E**, Sadeghi HM, Effron MB, Mehran R, Lansky AJ, Na Y, Cox DA, Garcia E, Tcheng JE, Griffin JJ, Stuckey TD, Turco M, Carroll JD, Grines CL, Stone GW. Impact of in-hospital acquired thrombocytopenia in patients undergoing primary angioplasty for acute myocardial infarction. *Am J Cardiol* 2005; **96**: 474-481 [PMID: 16098296 DOI: 10.1016/j.amjcard.2005.04.005]

10 **Berkowitz SD**, Sane DC, Sigmon KN, Shavender JH, Harrington RA, Tcheng JE, Topol EJ, Califf RM. Occurrence and clinical significance of thrombocytopenia in a population un-

dergoing high-risk percutaneous coronary revascularization. Evaluation of c7E3 for the Prevention of Ischemic Complications (EPIC) Study Group. *J Am Coll Cardiol* 1998; **32**: 311-319 [PMID: 9708455 DOI: 10.1016/S0735-1097(98)00252-6]

11 **De Labriolle A**, Bonello L, Lemesle G, Roy P, Steinberg DH, Xue Z, Suddath WO, Satler LF, Kent KM, Pichard AD, Lindsay J, Waksman R. Decline in platelet count in patients treated by percutaneous coronary intervention: definition, incidence, prognostic importance, and predictive factors. *Eur Heart J* 2010; **31**: 1079-1087 [PMID: 20089516 DOI: 10.1093/eurheartj/ehp594]

12 **Viswanathan G**, Kidambi A, Nelson A, Mayurathan G, Hardy J, Kesteven P, Zaman A. Glycoprotein IIb/IIIa inhibitor associated severe thrombocytopenia in patients with coronary artery disease: Clinical course and outcomes. *Platelets* 2012; **23**: 224-228 [PMID: 21913813 DOI: 10.3109/09537104.2011.604804]

13 **Kestin AS**, Valeri CR, Khuri SF, Loscalzo J, Ellis PA, MacGregor H, Birjiniuk V, Ouimet H, Pasche B, Nelson MJ. The platelet function defect of cardiopulmonary bypass. *Blood* 1993; **82**: 107-117 [PMID: 7686785]

14 **Bream-Rouwenhorst HR**, Hobbs RA, Horwitz PA. Thrombocytopenia in patients treated with heparin, combination antiplatelet therapy, and intra-aortic balloon pump counterpulsation. *J Interv Cardiol* 2008; **21**: 350-356 [PMID: 18537872 DOI: 10.1111/j.1540-8183.2008.00363]

15 **Naqvi TA**, Ikhlaque N, Baumann MA. Thrombocytopenia due to hypotension unrelated to infection: shock marrow. *Int J Clin Pract* 2005; **59**: 782-784 [PMID: 15963204 DOI: 10.1111/j.1742-1241.2005.00418.x]

16 **Schiariti M**, Puddu PE, Nesta C, Grillo P, Missiroli B, Cassese M. Prophylactic intra-aortic balloon pump is useful to facilitate percutaneous coronary interventions: images from rescue cases. *Heart Surg Forum* 2007; **10**: E411-E414 [PMID: 17855209 DOI: 10.1532/HSF98.20071085]

17 **Harrington RA**, Sane DC, Califf RM, Sigmon KN, Abbottsmith CW, Candela RJ, Lee KL, Topol EJ. Clinical importance of thrombocytopenia occurring in the hospital phase after administration of thrombolytic therapy for acute myocardial infarction. The Thrombolysis and Angioplasty in Myocardial Infarction Study Group. *J Am Coll Cardiol* 1994; **23**: 891-898 [PMID: 8106694 DOI: 10.1016/0735-1097(94)90634-3]

18 **Kelton JG**, Warkentin TE. Heparin-induced thrombocytopenia: a historical perspective. *Blood* 2008; **112**: 2607-2616 [PMID: 18809774 DOI: 10.1182/blood-2008-02-078014]

19 **Stribling WK**, Slaughter TF, Houle TT, Sane DC. Beyond the platelet count: heparin antibodies as independent risk predictors. *Am Heart J* 2007; **153**: 900-906 [PMID: 17540189 DOI: 10.1016/j.ahj.2007.03.013]

20 **Warkentin TE**, Kelton JG. Temporal aspects of heparin-induced thrombocytopenia. *N Engl J Med* 2001; **344**: 1286-1292 [PMID: 11320387 DOI: 10.1056/NEJM200104263441704]

21 **Geeganage C**, Wilcox R, Bath PM. Triple antiplatelet therapy for preventing vascular events: a systematic review and meta-analysis. *BMC Med* 2010; **8**: 36 [PMID: 20553581 DOI: 10.1186/1741-7015-8-36]

22 **Shantsila E**, Lip GY, Chong BH. Heparin-induced thrombocytopenia. A contemporary clinical approach to diagnosis and management. *Chest* 2009; **135**: 1651-1664 [PMID: 19497901 DOI: 10.1378/chest.08-2830]

23 **Aster RH**. Immune thrombocytopenia caused by glycoprotein IIb/IIIa inhibitors. *Chest* 2005; **127**: 53S-59S [PMID: 15706031 DOI: 10.1378/chest.127.2_suppl.53S]

24 **Kereiakes DJ**, Berkowitz SD, Lincoff AM, Tcheng JE, Wolski K, Achenbach R, Melsheimer R, Anderson K, Califf RM, Topol EJ. Clinical correlates and course of thrombocytopenia during percutaneous coronary intervention in the era of abciximab platelet glycoprotein IIb/IIIa blockade. *Am Heart J* 2000; **140**: 74-80 [PMID: 10874266 DOI: 10.1067/mhj.2000.106615]

25 **Gurm HS**, Tamhane U, Meier P, Grossman PM, Chetcuti S, Bates ER. A comparison of abciximab and small-molecule glycoprotein IIb/IIIa inhibitors in patients undergoing primary percutaneous coronary intervention: a meta-analysis of contemporary randomized controlled trials. *Circ Cardiovasc Interv* 2009; **2**: 230-236 [PMID: 20031720 DOI: 10.1161/CIRCINTERVENTIONS.108.847996]

26 **Steinhubl SR**, Tan WA, Foody JM, Topol EJ. Incidence and clinical course of thrombotic thrombocytopenic purpura due to ticlopidine following coronary stenting. EPISTENT Investigators. Evaluation of Platelet IIb/IIIa Inhibitor for Stenting. *JAMA* 1999; **281**: 806-810 [PMID: 10071001 DOI: 10.1001/jama.281.9.806]

27 **Zakarija A**, Bandarenko N, Pandey DK, Auerbach A, Raisch DW, Kim B, Kwaan HC, McKoy JM, Schmitt BP, Davidson CJ, Yarnold PR, Gorelick PB, Bennett CL. Clopidogrel-associated TTP: an update of pharmacovigilance efforts conducted by independent researchers, pharmaceutical suppliers, and the Food and Drug Administration. *Stroke* 2004; **35**: 533-537 [PMID: 14707231 DOI: 10.1161/01.STR.0000109253.66918.5E]

28 **Bennett CL**, Connors JM, Carwile JM, Moake JL, Bell WR, Tarantolo SR, McCarthy LJ, Sarode R, Hatfield AJ, Feldman MD, Davidson CJ, Tsai HM. Thrombotic thrombocytopenic purpura associated with clopidogrel. *N Engl J Med* 2000; **342**: 1773-1777 [PMID: 10852999 DOI: 10.1056/NEJM200006153422402]

29 **Grosbois B**, Milton D, Beneton C, Jacomy D. Thrombocytopenia induced by angiotensin converting enzyme inhibitors. *BMJ* 1989; **298**: 189-190 [PMID: 2538175 DOI: 10.1136/bmj.298.6667.189-b]

30 **Srinivasa Rao K**, Prasad T, Mohanta GP, Manna PK. An overview of statins as hypolipidemic drugs. *Int J Pharm Sci Rev Res* 2011; **3**: 178-183

31 **Payne BA**, Pierre RV. Pseudothrombocytopenia: a laboratory artifact with potentially serious consequences. *Mayo Clin Proc* 1984; **59**: 123-125 [PMID: 6422167 DOI: 10.1016/S0025-6196(12)60247-X]

32 **Shenoy C**, Orshaw P, Devarakonda S, Harjai KJ. Occurrence, predictors, and outcomes of post-percutaneous coronary intervention thrombocytopenia in an unselected population. *J Interv Cardiol* 2009; **22**: 156-162 [PMID: 19245383 DOI: 10.1111/j.1540-8183.2009.00424.x]

33 **Yusuf SW**, Iliescu C, Bathina JD, Daher IN, Durand JB. Antiplatelet therapy and percutaneous coronary intervention in patients with acute coronary syndrome and thrombocytopenia. *Tex Heart Inst J* 2010; **37**: 336-340 [PMID: 20548817]

34 **Schiariti M**, Saladini P, Cuturello D, Iannetta L, Torromeo C, Puddu PE. Decline in platelet count and long-term post-PCI ischemic events: implication of the intra-aortic balloon pump. *Vascul Pharmacol* 2014; **60**: 25-31 [PMID: 24239797 DOI: 10.1016/j.vph.2013.11.002]

35 **Kiviniemi T**, Karjalainen P, Rubboli A, Schlitt A, Tuomainen P, Niemelä M, Laine M, Biancari F, Lip GY, Airaksinen KE. Thrombocytopenia in patients with atrial fibrillation on oral anticoagulation undergoing percutaneous coronary intervention. *Am J Cardiol* 2013; **112**: 493-498 [PMID: 23672991 DOI: 10.1016/j.amjcard.2013.04.007]

36 **Schiariti M**, Saladini A, Missiroli B, Papalia F, Cuturello D, Puddu PE, Gaudio C. Safety of downstream high-dose tirofiban bolus among 1578 patients undergoing percutaneous coronary intervention: the Sant'ANna TIrofiban Safety study. *J Cardiovasc Med* (Hagerstown) 2010; **11**: 250-259 [PMID: 19952776 DOI: 10.2459/JCM.0b013e328334c7b9]

**P- Reviewer:** Lin GM **S- Editor:** Zhai HH **L- Editor:** A
**E- Editor:** Wu HL

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i2.29

*World J Meta-Anal* 2014 May 26; 2(2): 29-35
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Weekday *vs* weekend endoscopy: Is there a true difference in patient outcomes? A meta-analysis

Alisha M Hinds, Dina Ahmad, Kristi T Lopez, Michelle L Matteson-Kome, Matthew L Bechtold

Alisha M Hinds, Dina Ahmad, Kristi T Lopez, Michelle L Matteson-Kome, Matthew L Bechtold, Division of Gastroenterology and Hepatology, University of Missouri School of Medicine, Columbia, MO 65212, United States

Author contributions: Hinds AM, Ahmad D and Bechtold ML were responsible for the conception and design of the study; Hinds AM, Ahmad D and Lopez KT collected the data and organized data extraction sheets; Matteson-Kome ML and Bechtold ML statistically analyzed the data; Hinds AM, Ahmad D and Lopez KT drafted the manuscript with critical revision being performed by Matteson-Kome ML and Bechtold ML.

Correspondence to: Matthew L Bechtold, MD, FASGE, FACG, Division of Gastroenterology and Hepatology, University of Missouri Health Sciences Center, Five Hospital Drive, Columbia, MO 65212,
United States. bechtoldm@health.missouri.edu
Telephone: +1-573-8821013 Fax:+1-573-8844595
Received: November 6, 2013 Revised: March 5, 2014
Accepted: April 11, 2014
Published online: May 26, 2014

## Abstract

**AIM:** To evaluate the "weekend effect" on outcomes in patient admitted on the weekend for upper gastrointestinal bleeding (UGIB).

**METHODS:** A comprehensive search was performed (March 2014). Studies comparing weekend and weekday endoscopy in patients with UGIB were included. All studies had at least 2 of 3 primary outcomes which included: mortality, need for surgery, time to endoscopy, endoscopy on admission day, and length of hospital stay. Three authors individually extracted data. Meta-analysis was performed using pooled estimates with odds ratio or mean difference by fixed and random effects models.

**RESULTS:** Eleven studies met the inclusion criteria. Patients admitted with UGIB on the weekend exhibited a statistically significant increase in mortality (OR = 1.13; 95%CI: 1.06-1.20; $P < 0.01$), need for surgery (OR = 2.46; 95%CI: 1.51-3.99; $P < 0.01$), and time to endoscopy (MD 2.68; 95%CI: 0.17-5.20; $P = 0.04$) as compared to patients admitted with UGIB on a weekday. Furthermore, patients with UGIB admitted on weekend experienced statistically significant less endoscopy on day of admission (OR = 0.72; 95%CI: 0.62-0.85; $P < 0.01$). No difference was noted between the two groups for length of hospital stay (MD -1.29; 95%CI: -3.03-0.45; $P = 0.15$).

**CONCLUSION:** A weekend effect seems to be apparent in patients with UGIB with significantly poorer outcomes.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Weekend; Weekday; Upper gastrointestinal bleeding; Endoscopy; Meta-analysis

**Core tip:** Patients admitted on weekends have been suggested by multiple studies to have poorer outcomes, even in those with gastrointestinal bleeding. Therefore, we conducted a meta-analysis of observational studies to examine the effect of weekend *vs* weekday admissions for patients with upper gastrointestinal bleeding (UGIB). We discovered that patients with UGIB admitted on weekends had higher mortality, need for surgery, and time to endoscopy compared to those admitted on weekdays. This meta-analysis demonstrates the need for policies to decrease these poorer outcomes for our weekend patients with UGIB.

Hinds AM, Ahmad D, Lopez KT, Matteson-Kome ML, Bechtold ML. Weekday *vs* weekend endoscopy: Is there a true difference in patient outcomes? A meta-analysis. *World J Meta-Anal* 2014; 2(2): 29-35 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i2/29.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i2.29

## INTRODUCTION

Emergency room visits and hospital admissions for all causes of upper gastrointestinal bleeding (UGIB) have decreased over the last ten years but still account for around (50-170)/100000 patients per year[1-8]. Mortality from UGIB is approximated between 3%-7% in recent studies[3,7], higher with variceal hemorrhage (approximately 15%)[9,10]. Timely intervention in patients presenting with active bleeds has been shown to reduce mortality[3]. Over half of patients presenting with UGIB has been shown to be from peptic ulcer disease, an increasingly common diagnosis[1,7]. Approximately $750 million spent each year on hospitalizations of patients with UGIB, leading to the importance of timely intervention to not only reduce mortality but also hospital costs[1]. Multiple medical and surgical diagnoses have been shown to have increased mortality over the weekend, leading to a term known as the "weekend effect"[1-3,11-17].

Over the past decade, studies have suggested that patients admitted on the weekend for UGIB have a higher mortality rate, length of stay, time to endoscopy, and increase in hospital costs than patients originally admitted to the hospital on weekdays, thus displaying a "weekend effect". Multiple reasons for this weekend effect have been proposed, but no cause-effect relationship has been found to be significant[1,11,14]. The main proposed reasons for the weekend effect is the difference in staffing and access to critical procedures in an adequate amount of time; however, this issue seems to be the result of a combination of factors[1,2,18]. There have been multiple studies over the last ten years on this subject, but all have been observational studies with all but one being retrospective. No randomized controlled trial has been performed evaluating the weekend effect for patients with UGIB. Furthermore, of these observational studies, results have varied, adding the need for further examination into this issue. If a weekend effect is present, significant changes in policy must be considered to help reduce poorer outcomes in those patients admitted on the weekend with UGIB. Due to this important need for further examination and given the variation among the studies, a meta-analysis was completed comparing weekend verses weekday admission for UGIB.

## MATERIALS AND METHODS

### Literature search

A complete search of PubMed, Medline, Scopus, CINAHL and Cochrane databases was completed in March 2014. Search terms were used individually or in various combinations and included weekend endoscopy, weekday endoscopy, upper gastrointestinal bleed, upper gastrointestinal hemorrhage, variceal hemorrhage, non-variceal hemorrhage, esophagogastroduodenoscopy, peptic ulceration, duodenal ulceration, and gastric ulceration. Peer-reviewed studies that compared weekend *vs* weekday endoscopy with UGIB were selected and reviewed. References of relevant papers were searched as well for possible additional articles that were not identified in the original search. Three investigators reviewed all studies selected for inclusion criteria. Studies in children or in languages other than English were excluded from this meta-analysis.

### Data extraction

All studies which met the inclusion criteria were reviewed by three investigators. All prospective and retrospective studies on weekday verses weekend endoscopy for UGIB were selected. All studies had at least 2 of 3 primary outcomes which included: Mortality, need for surgery, time to endoscopy, endoscopy on admission day, and length of hospital stay. If a study had missing data on these subjects or clarification needed to be obtained, attempts were made to contact the authors to obtain the necessary information. Data from the studies chosen were extracted by three separate investigators individually with differences being settled by mutual agreement.

### Statistical analysis

The meta-analysis was performed comparing the outcomes of UGIB in weekend and weekday groups by calculating pooled estimates of mortality, need for surgery, time to endoscopy, endoscopy on admission day, and length of hospital stay by using OR or MD by fixed and random effects models. The meta-analysis was performed in accordance with the meta-analysis of observational studies in epidemiology (MOOSE) guidelines[19]. Publication bias was assessed by funnel plots. Heterogeneity among studies was assessed by calculating the $I^2$ measure of inconsistency, which was considered significant if $P$ value $< 0.05$ or $I^2 > 50\%$. If present, a sensitivity analysis was performed by eliminating one or more studies until heterogeneity was not present and results compared. RevMan 5.1 was used for statistical analysis.

### Study quality assessment

The quality of studies was assessed using the Effective Public Health Practice Project (EPHPP) model[20,21]. This scale assesses study quality as strong, moderate, or weak based upon criteria ratings for selection bias, study design, confounders, blinding, data collection methods, and withdrawal and dropout descriptions. The quality of the study is based upon how many weak ratings per category ($\geqslant$ 2 ratings is weak, one weak rating is moderate, and no weak ratings strong)[20,21].

## RESULTS

### Literature search

The initial search identified 90 articles through the electronic database search (Figure 1). Of these articles, 14 relevant peer-reviewed articles in English were selected for full-text review by three independent authors (AH, DA, MB). Of these studies, one article was found to be a letter to the editor and two articles had insufficient data for analysis. Therefore, 11 studies were identified

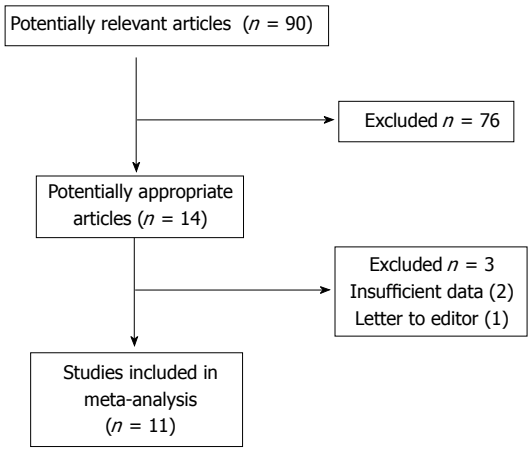



**Figure 1 Details of article search.**

that met the inclusion criteria and were included in the meta-analysis[1-4,8-11,14,18,22]. Of the 11 studies included (N = 870824), 10 were retrospective cohorts[1-3,8-11,14,18,22] and one was a prospective cohort[4]. The studies were performed at various locations in the United States and Europe and were published between 2004 and 2012. All studies examined the impact of weekend admissions on UGIB outcomes. Table 1 shows the details of the selected studies. Based upon the EPHPP model for quality assessment of studies, one study[4] was considered strong, eight studies[1-3,10,11,14,18,22] considered moderate, and two studies[8,9] considered weak. Table 2 given that all studies were retrospective except one, blinding and withdrawals/dropout description were not applicable. The study by de Groot *et al*[4] was the only prospective study.

### Mortality

Mortality was examined in 10 studies (N = 825207)[1-4,8-11,18,22]. Mortality was noted in 8817 of 194386 (4.5%) patients with UGIB bleeding admitted on weekends and 24762 of 630821 (3.9%) patients with UGIB admitted on weekdays. A statistically significant increase in mortality was noted for those patients with UGIB admitted on the weekend as compared to those admitted on the weekdays (OR = 1.13; 95%CI: 1.06-1.20; $P < 0.01$). Figure 2A statistically significant heterogeneity was observed ($I^2$ = 64%, $P < 0.01$). Upon sensitivity analysis, one study was eliminated[2] and revealed similar results with no significant heterogeneity (OR = 1.11; 95%CI: 1.04-1.17; $P < 0.01$ with $I^2 = 42\%$, $P = 0.09$).

### Need for surgery

The need for surgery was examined in six studies (N = 636413)[1-4,8,11]. On weekends, 2943 of 148810 (2%) patients with UGIB required surgery. On weekdays, 7562 of 487603 (1.6%) patients with UGIB required surgery. A statistically significant increase in the need for surgery was discovered in patients with UGIB admitted on weekends as compared to weekdays (OR = 2.46; 95%CI: 1.51-3.99; $P < 0.01$). Figure 2B statistically significant heterogeneity was observed ($I^2$ = 98%, $P < 0.01$). Upon sensitivity analysis, if two studies were eliminated[1,11], no heterogeneity was noted and similar results were observed (OR = 1.35; 95%CI: 1.24-1.46; $P < 0.01$ with $I^2$ = 0%, $P = 0.63$).

### Time to endoscopy

The time to endoscopy was analyzed in six studies (N = 382410)[1,3,8,9,14,18]. A statistically significant increase in the time to endoscopy was noted in patients with UGIB admitted on weekends as compared to weekdays (MD 2.68; 95% CI: 0.17-5.20; $P = 0.04$). Figure 2C statistically significant heterogeneity was observed ($I^2 = 100\%$, $P < 0.01$). Upon sensitivity analysis, if three studies were eliminated[3,8,18], no significant heterogeneity was noted and similar results were observed (OR = 3.96; 95% CI: 3.04-4.88; $P < 0.01$ with $I^2 = 41\%$, $P = 0.19$).

### Endoscopy on admission day

Endoscopy on admission day was evaluated in four studies (N = 700834)[1,2,10,11]. On weekends, 66876 of 165008 (40.5%) patients with UGIB underwent endoscopy on day of admission. On weekdays, 257067 of 535826 (48%) patients with UGIB underwent endoscopy on day of admission. A statistically significant decrease in endoscopy on day of admission was found in patients with UGIB admitted on weekends as compared to weekdays (OR = 0.72; 95%CI: 0.62-0.85; $P < 0.01$). Figure 2D statistically significant heterogeneity was observed ($I^2$ = 99%, $P < 0.01$). Given limited number of studies for this outcome, sensitivity analysis could not be performed.

### Length of hospital stay

The length of hospital stay was analyzed in four studies (N = 99831)[3,8,9,18]. No statistically significant difference in the length of hospital stay was observed in patients with UGIB admitted on weekends as compared to weekdays (MD -1.29; 95%CI: -3.03-0.45; $P = 0.15$). Figure 2E statistically significant heterogeneity was observed ($I^2$ = 100%, $P < 0.01$). Upon sensitivity analysis, if two studies were eliminated[3,18], no significant heterogeneity was noted and similar results were observed (MD -1.19; 95%CI: -2.78-0.41; $P < 0.01$ with $I^2 = 25\%$, $P = 0.25$).

### Publication bias

No publication bias was noted for any of the outcomes in this meta-analysis based upon funnel plots (Figure 3).

## DISCUSSION

Several studies have examined the effect of a weekend hospitalization on the outcomes of patients with various medical and surgical emergencies[12-14,23]. Many of the reports provided evidence of the so-called "weekend effect", a phenomenon in which worse outcomes have been shown in patients admitted during the weekend compared to weekdays[12-14,23]. This effect for patients admitted over the weekend has been thought to be due to lower hospital staffing, more provider workload, and

**Table 1 Details described on the studies included in meta-analysis**

| Ref. | Study type | Location | Time | # of Pts | Group | Pts/Group |
|---|---|---|---|---|---|---|
| Youn *et al*[3] | Retrospective 4 centers | South Korea | 1/07-12/09 | 388 | Weekend | 62 |
| | | | | | Weekday | 326 |
| de Groot *et al*[4] | Prospective 8 centers | The Netherlands | 10/09-9/11 | 571 | Weekend | 167 |
| | | | | | Weekday | 404 |
| Byun *et al*[9] | Retrospective 1 center | South Korea | 1/05-2/09 | 294 | Weekend | 74 |
| | | | | | Weekday | 220 |
| Jairath *et al*[11] | Retrospective 212 centers | United Kingdom | 5/07-6/07 | 6749 | Weekend | 1499 |
| | | | | | Weekday | 5250 |
| Button *et al*[22] | Retrospective 22 centers | Wales | 4/99-3/07 | 24421 | Weekend | 5686 |
| | | | | | Weekday | 18285 |
| Dorn *et al*[18] | Retrospective NIS database | United States | 1998-2003 | 98975 | Weekend | 23339 |
| | | | | | Weekday | 75636 |
| Shaheen *et al*[1] | Retrospective NIS database | United States | 1993-2005 | 237412 | Weekend | 57270 |
| | | | | | Weekday | 180142 |
| Myers *et al*[10] | Retrospective NIS database | United States | 1998-2005 | 36734 | Weekend | 9237 |
| | | | | | Weekday | 27497 |
| Ananthakrishnan *et al*[2] | Retrospective NIS database | United States | 2004 | 419939 | Weekend | 97002 |
| | | | | | Weekday | 322937 |
| Bell *et al*[14] | Retrospective 190 centers | Canada | 4/88-3/97 | 45167 | Weekend | 3602 |
| | | | | | Weekday | 41565 |
| Haas *et al*[8] | Retrospective 1 center | United States | 1/08-10/08 | 174 | Weekend | 50 |
| | | | | | Weekday | 124 |

**Table 2 Quality assessment of the studies based upon the effective public health practice project scale**

| Ref. | Selection bias | Study design | Confounders | Blinding | Data collection method | Withdrawals and dropouts | Overall assessment |
|---|---|---|---|---|---|---|---|
| Youn *et al*[3] | Moderate | Moderate | Weak | N/A | Strong | N/A | Moderate |
| de Groot *et al*[4] | Strong | STRONG | Moderate | N/A | Strong | Strong | Strong |
| Byun *et al*[9] | Weak | Moderate | Weak | N/A | Strong | N/A | Weak |
| Jairath *et al*[11] | Strong | Moderate | Weak | N/A | Strong | N/A | Moderate |
| Button *et al*[22] | Strong | Moderate | Weak | N/A | Strong | N/A | Moderate |
| Dorn *et al*[18] | Strong | Moderate | Weak | N/A | Strong | N/A | Moderate |
| Shaheen *et al*[1] | Strong | Moderate | Weak | N/A | Strong | N/A | Moderate |
| Myers *et al*[10] | Strong | Moderate | Weak | N/A | Strong | N/A | Moderate |
| Ananthakrishnan *et al*[2] | Strong | Moderate | Weak | N/A | Strong | N/A | Moderate |
| Bell *et al*[14] | Strong | Moderate | Weak | N/A | Strong | N/A | Moderate |
| Haas *et al*[8] | Weak | Moderate | Weak | N/A | Strong | N/A | Weak |

N/A: Not applicable.

sicker patients[13,15]. Patients admitted on weekends with UGIB may be impacted as well. Many studies have been performed in evaluating this "weekend effect" for patients with UGIB admitted on weekends. These studies comprise of 10 retrospective studies[1-3,8-11,14,18,22] and one prospective study[4]. Unfortunately, no randomized controlled trials have been performed on this subject, limiting the literature to a few prospective and many retrospective studies. Ananthakrishnan *et al*[2] conducted a respective study on patients with acute variceal hemorrhage and acute non-variceal hemorrhage (ANVH) in which the weekend effect was demonstrated in the ANVH group. In the study performed by de Groot *et al*[4], weekend admissions were at higher risk of adverse outcomes although the quality of care did not differ between weekday or weekend admissions. Shaheen *et al*[1] found that mortality and delay in endoscopy were higher in weekend admissions, but the delay in endoscopy did not appear to have affected the patient outcomes. The weekend effect was also demonstrated in studies by Dorn *et al*[18], Youn *et al*[3] and Button *et al*[22] studies. Four studies did not show that weekend admissions influenced in-hospital mortality in patients with UGIB[8-11]. Given the variability in results of the studies, this meta-analysis was performed.

Our meta-analysis demonstrates an increased mortality, need for surgery, and time to endoscopy in patients with UGIB admitted over the weekend as compared to being admitted on the weekday. Furthermore, patients with weekend admissions for UGIB experienced less endoscopy on admission day than those with weekday admissions. Despite these outcomes, the length of hospital stay did not differ between the two groups. Therefore, based upon these results, patients with UGIB who are admitted on the weekend appear to have poorer outcomes than those admitted on the weekday; however, evaluation of the strengths and limitations of this meta-analysis is necessary to clarify impact of results.

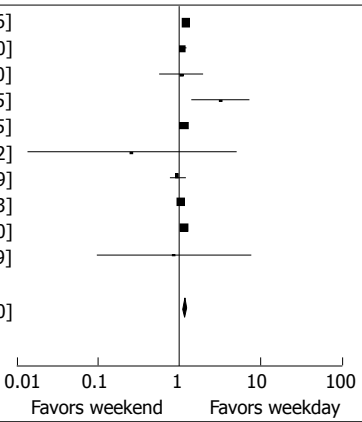
The strengths of our meta-analysis are numerous.

**A**

| Study or subgroup | Weekend Events | Weekend Total | Weekday Events | Weekday Total | Weight | Odds ratio M-H, random, 95%CI |
|---|---|---|---|---|---|---|
| Ananthakrishnan *et al* | 4156 | 97002 | 11503 | 322937 | 22.3% | 1.21 [1.17, 1.26] |
| Button *et al* | 602 | 5686 | 1792 | 18285 | 14.6% | 1.09 [0.99, 1.20] |
| Byun *et al* | 17 | 74 | 48 | 220 | 0.8% | 1.07 [0.57, 2.00] |
| de Groot *et al* | 15 | 167 | 12 | 404 | 0.6% | 3.22 [1.48, 7.05] |
| Dorn *et al* | 887 | 23339 | 2496 | 75636 | 17.0% | 1.16 [1.07, 1.25] |
| Haas *et al* | 0 | 50 | 4 | 124 | 0.0% | 0.27 [0.01, 5.02] |
| Jairath *et al* | 148 | 1499 | 527 | 5250 | 6.8% | 0.98 [0.81, 1.19] |
| Myers *et al* | 1044 | 9237 | 2970 | 27497 | 17.5% | 1.05 [0.98, 1.13] |
| Shaheen *et al* | 1947 | 57270 | 5404 | 180142 | 20.4% | 1.14 [1.08, 1.20] |
| Youn *et al* | 1 | 62 | 6 | 326 | 0.1% | 0.87 [0.10, 7.39] |
| Total (95%CI) | | 194386 | | 630821 | 100.0% | 1.13 [1.06, 1.20] |
| Total events | 8817 | | 24762 | | | |

Heterogeneity: Tau$^2$ = 0.00; $\chi^2$ = 24.77, df = 9 ($P$ = 0.003), $I^2$ = 64%

Test for overall effect: $Z$ = 4.03 ($P$ < 0.001)

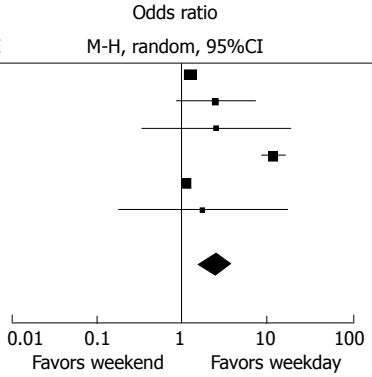



**B**

| Study or subgroup | Weekend Events | Weekend Total | Weekday Events | Weekday Total | Weight | Odds ratio M-H, random, 95%CI |
|---|---|---|---|---|---|---|
| Ananthakrishnan *et al* | 817 | 89762 | 2049 | 301357 | 27.4% | 1 .34 [1.24, 1.46] |
| de Groot *et al* | 7 | 167 | 7 | 404 | 11.8% | 2.48 [0.86, 7.19] |
| Haas *et al* | 2 | 50 | 2 | 124 | 4.9% | 2.54 [0.35, 18.56] |
| Jairath *et al* | 152 | 1499 | 49 | 5250 | 24.5% | 11.98 [8.63, 16.62] |
| Shaheen *et al* | 1964 | 57270 | 5452 | 180142 | 27.5% | 1.14 [1.08, 1.20] |
| Youn *et al* | 1 | 62 | 3 | 326 | 3.9% | 1.77 [0.18, 17.25] |
| Total (95%CI) | | 148810 | | 487603 | 100.0% | 2.46 [1.51, 3.99] |
| Total events | 2943 | | 7562 | | | |

Heterogeneity: Tau$^2$ = 0.22; $\chi^2$ = 200.40, df = 5 ($P$ < 0.00001), $I^2$ = 98%

Test for overall effect: $Z$ = 3.64 ($P$ = 0.0003)

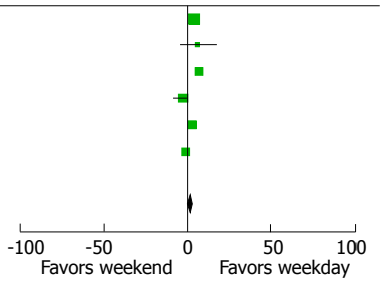



**C**

| Study or subgroup | Weekend Mean | Weekend SD | Weekend Total | Weekday Mean | Weekday SD | Weekday Total | Weight | Mean difference IV, random, 95%CI |
|---|---|---|---|---|---|---|---|---|
| Bell *et al* | 50.4 | 38.4 | 3602 | 45.6 | 45.6 | 41565 | 19.0% | 4.80 [3.47,6.13] |
| Byun *et al* | 20.9 | 43 | 74 | 14.7 | 29.1 | 220 | 4.4% | 6.20 [-4.32, 16.72] |
| Dorn *et al* | 74.4 | 33.6 | 23339 | 67.2 | 36 | 75636 | 19.9% | 7.20 [6.70, 7.70] |
| Haas *et al* | 7.52 | 7.02 | 50 | 10.82 | 9.26 | 124 | 16.6% | -3.30 [-5.84, -0.76] |
| Shaheen *et al* | 53.04 | 0.24 | 57270 | 49.44 | 0.24 | 180142 | 20.0% | 3.60 [3.60, 3.60] |
| Youn *et al* | 5.11 | 0.76 | 62 | 5.65 | 0.42 | 326 | 20.0% | -0.54 [-0.73, -0.35] |
| Total (95%CI) | | | 84397 | | | 298013 | 100.0% | 2.68 [0.17, 5.20] |

Heterogeneity: Tau$^2$ = 8.21; $\chi^2$ = 1968.20, df = 5 ($P$ < 0.00001), $I^2$ = 100%

Test for overall effect: $Z$ = 2.09 ($P$ = 0.04)

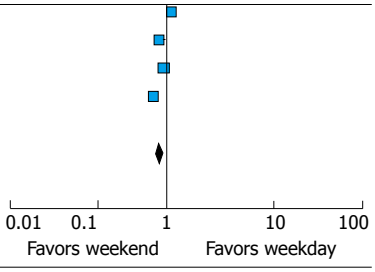



**D**

| Study or subgroup | Weekend Events | Weekend Total | Weekday Events | Weekday Total | Weight | Odds ratio M-H, random, 95%CI |
|---|---|---|---|---|---|---|
| Ananthakrishnan *et al* | 48622 | 97002 | 192547 | 322937 | 25.9% | 0.68 [0.67, 0.69] |
| Jairath *et al* | 570 | 1499 | 2888 | 5250 | 22.8% | 0.50 [0.45, 0.56] |
| Myers *et al* | 3972 | 9237 | 12374 | 27497 | 25.4% | 0.92 [0.88, 0.97] |
| Shaheen *et al* | 13712 | 57270 | 49258 | 180142 | 25.9% | 0.84 [0.82, 0.85] |
| Total (95%CI) | | 165008 | | 535826 | 100.0% | 0.72 [0.62, 0.85] |
| Total events | 66876 | | 257067 | | | |

Heterogeneity: Tau$^2$ = 0.03; $\chi^2$ = 371.90, df = 3 ($P$ < 0.00001), $I^2$ = 99%

Test for overall effect: $Z$ = 3.94 ($P$ < 0.0001)



**E**

| Study or subgroup | Weekend Mean | SD | Total | Weekday Mean | SD | Total | Weight | Mean difference IV, random, 95%CI |
|---|---|---|---|---|---|---|---|---|
| Byun *et al* | 12.6 | 12 | 74 | 15.5 | 15.6 | 220 | 13.8% | -2.90 [-6.32, 0.52] |
| Dorn *et al* | 4.24 | 3.2 | 23339 | 4.22 | 3.4 | 75636 | 29.6% | 0.02 [-0.03, 0.07] |
| Haas *et al* | 3.1 | 2.2 | 50 | 3.9 | 4.5 | 124 | 27.0% | -0.80 [-1.80, 0.20] |
| Youn *et al* | 6.2 | 0.5 | 62 | 8.5 | 0.3 | 326 | 29.6% | -2.30 [-2.43, -2.17] |
| Total (95%CI) | | | 23525 | | | 76306 | 100.0 | -1.29 [-3.03, 0.45] |

Heterogeneity: Tau$^2$ = 2.64; $\chi^2$ = 1101.87, df = 3 ($P$ < 0.00001), $I^2$ = 100%
Test for overall effect: $Z$ = 1.46 ($P$ = 0.15)

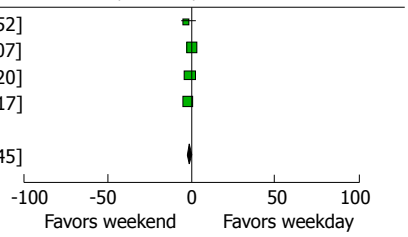



**Figure 2 Forest plot demonstrating comparison of weekend *vs* weekday admission for patients with upper gastrointestinal bleeding.** A: Mortality; B: Need for surgery; C: Time to endoscopy; D: Endoscopy on day of admission; E: Length of hospital stay.

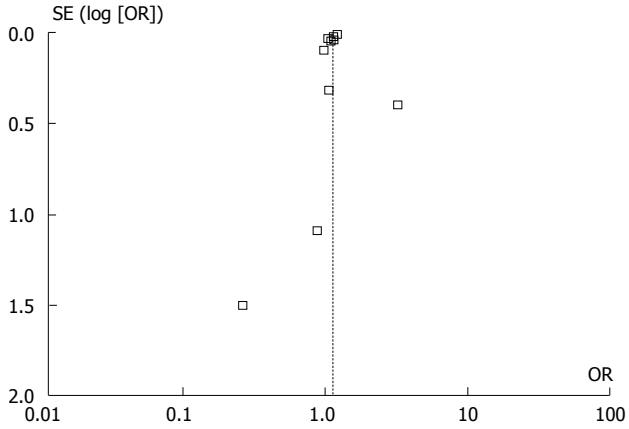



**Figure 3 Funnel plot showing no publication bias.**

First, a comprehensive article screening process with an extensive search technique was used to maximize article recognition. Second, a large number of patients in various populations were examined. Third, all of the included studies examined the impact of weekend admissions on patients with UGIB outcome. Fourth, no publication bias was noted. Finally, this represents the first meta-analysis to-date that assesses the weekend effect on patients admitted with UGIB. However, a few limitations were also observed. First, only cohort studies (retrospective and prospective) were included. However, no randomized controlled trials have been performed on this particular subject. Second, statistically significant heterogeneity was observed for all outcomes. To minimize the effect of heterogeneity, a random effects model was utilized. Also, a sensitivity analysis was performed by eliminating various studies resulting in no statistically significant heterogeneity showing similar results on all outcomes but one (endoscopy on admission day). Despite these adjustments, the results of this meta-analysis must be interpreted with some caution given the heterogeneity. Third, no abstracts were included in this meta-analysis. Due to abstracts not being peer reviewed and given the number of very large peer-reviewed studies on this subject, abstracts were deemed less quality and were not included. Fourth, selection bias may be apparent in this meta-analysis though not in traditional sense. Multiple studies utilized the same database for selection of the patients. However, search terms, such as variceal bleeding and non-variceal bleeding, and years were slightly different between the studies. Given this information and in an effort to avoid missing patients, all studies were included with the premise that a given patient may be counted twice in this meta-analysis. In conclusion, this meta-analysis demonstrates that patients with UGIB admitted on weekends have increased mortality, need for surgery, and time to endoscopy while decreased endoscopy on the admission day. Given the limitations in this meta-analysis, the results must be interpreted with caution but does strongly suggest a weekend effect is present in patients with UGIB.

## COMMENTS

### *Background*

The "weekend effect" showing poorer outcomes in patient admitted on the weekend rather than the weekday has been demonstrated in many different specialties. Upper gastrointestinal bleeding (UGIB) is a common reason for endoscopy and may occur any day of the week. Multiple studies have been performed evaluating the "weekend effect" in patients with UGIB requiring endoscopy with varied results.

### *Research frontiers*

Given the differences in the results of studies on weekend *vs* weekday outcomes in patients admitted with UGIB, the authors performed a meta-analysis comparing weekday to weekend admissions for UGIB to assess for mortality, need for surgery, time to endoscopy, endoscopy on admission day, and length of hospital stay.

### *Innovations and breakthroughs*

This is the first meta-analysis comparing weekend *vs* weekday admissions for UGIB. The authors found that a "weekend effect" does exist with patients admitted on the weekend having increased mortality, need for surgery, and time to endoscopy while decreased endoscopy on the admission day.

### *Applications*

Polices at local and regional levels are necessary to minimize the "weekend effect" in patients admitted with UGIB.

### *Terminology*

Odds ratio: Statistical term for the odds an event did or did not occur. Mean difference: Statistical term of difference between the means for a given variable. Heterogeneity: Test for uniformity in composition of studies included. Publication bias: Phenomenon where positive studies are more likely to be published than negative studies, leading to possible misrepresentation of data in meta-analysis.

### *Peer review*

The authors provide a study to evaluate the "weekend effect" on outcomes in patient admitted on the weekend for UGIB. This article is interesting.

## REFERENCES


1 **Shaheen AA**, Kaplan GG, Myers RP. Weekend versus weekday admission and mortality from gastrointestinal hemorrhage caused by peptic ulcer disease. *Clin Gastroenterol Hepatol* 2009; **7**: 303-310 [PMID: 18849015 DOI: 10.1016/j.cgh.2008.08.033]

2 **Ananthakrishnan AN**, McGinley EL, Saeian K. Outcomes of weekend admissions for upper gastrointestinal hemorrhage: a nationwide analysis. *Clin Gastroenterol Hepatol* 2009; **7**: 296-302e1 [PMID: 19084483]

3 **Youn YH**, Park YJ, Kim JH, Jeon TJ, Cho JH, Park H. Weekend and nighttime effect on the prognosis of peptic ulcer bleeding. *World J Gastroenterol* 2012; **18**: 3578-3584 [PMID: 22826623 DOI: 10.3748/wjg.v18.i27.3578]

4 **de Groot NL**, Bosman JH, Siersema PD, van Oijen MG, Bredenoord AJ. Admission time is associated with outcome of upper gastrointestinal bleeding: results of a multicentre prospective cohort study. *Aliment Pharmacol Ther* 2012; **36**: 477-484 [PMID: 22747509 DOI: 10.1111/j.1365-2036.2012.05205.x]

5 **van Leerdam ME**, Vreeburg EM, Rauws EA, Geraedts AA, Tijssen JG, Reitsma JB, Tytgat GN. Acute upper GI bleeding: did anything change? Time trend analysis of incidence and outcome of acute upper GI bleeding between 1993/1994 and 2000. *Am J Gastroenterol* 2003; **98**: 1494-1499 [PMID: 12873568 DOI: 10.1111/j.1572-0241.2003.07517.x]

6 **Lanas A**, García-Rodríguez LA, Polo-Tomás M, Ponce M, Quintero E, Perez-Aisa MA, Gisbert JP, Bujanda L, Castro M, Muñoz M, Del-Pino MD, Garcia S, Calvet X. The changing face of hospitalisation due to gastrointestinal bleeding and perforation. *Aliment Pharmacol Ther* 2011; **33**: 585-591 [PMID: 21205256 DOI: 10.1111/j.1365-2036.2010.04563.x]

7 **Laine L**, Yang H, Chang SC, Datto C. Trends for incidence of hospitalization and death due to GI complications in the United States from 2001 to 2009. *Am J Gastroenterol* 2012; **107**: 1190-1195; quiz 1196 [PMID: 22688850 DOI: 10.1038/ajg.2012.168]

8 **Haas JM**, Gundrum JD, Rathgaber SW. Comparison of time to endoscopy and outcome between weekend/weekday hospital admissions in patients with upper GI hemorrhage. *WMJ* 2012; **111**: 161-165 [PMID: 22970530]

9 **Byun SJ**, Kim SU, Park JY, Kim BK, Kim do Y, Han KH, Chon CY, Ahn SH. Acute variceal hemorrhage in patients with liver cirrhosis: weekend versus weekday admissions. *Yonsei Med J* 2012; **53**: 318-327 [PMID: 22318819 DOI: 10.3349/ymj.2012.53.2.318]

10 **Myers RP**, Kaplan GG, Shaheen AM. The effect of weekend versus weekday admission on outcomes of esophageal variceal hemorrhage. *Can J Gastroenterol* 2009; **23**: 495-501 [PMID: 19623333]

11 **Jairath V**, Kahan BC, Logan RF, Hearnshaw SA, Travis SP, Murphy MF, Palmer KR. Mortality from acute upper gastrointestinal bleeding in the United kingdom: does it display a "weekend effect"? *Am J Gastroenterol* 2011; **106**: 1621-1628 [PMID: 21606977 DOI: 10.1038/ajg.2011.172]

12 **Barba R**, Losa JE, Velasco M, Guijarro C, García de Casasola G, Zapatero A. Mortality among adult patients admitted to the hospital on weekends. *Eur J Intern Med* 2006; **17**: 322-324 [PMID: 16864005 DOI: 10.1016/j.ejim.2006.01.003]

13 **Becker DJ**. Do hospitals provide lower quality care on weekends? *Health Serv Res* 2007; **42**: 1589-1612 [PMID: 17610439 DOI: 10.1111/j.1475-6773.2006.00663.x]

14 **Bell CM**, Redelmeier DA. Mortality among patients admitted to hospitals on weekends as compared with weekdays. *N Engl J Med* 2001; **345**: 663-668 [PMID: 11547721 DOI: 10.1056/NEJMsa003376]

15 **Cram P**, Hillis SL, Barnett M, Rosenthal GE. Effects of weekend admission and hospital teaching status on in-hospital mortality. *Am J Med* 2004; **117**: 151-157 [PMID: 15276592 DOI: 10.1016/j.amjmed.2004.02.035]

16 **Kostis WJ**, Demissie K, Marcella SW, Shao YH, Wilson AC, Moreyra AE. Weekend versus weekday admission and mortality from myocardial infarction. *N Engl J Med* 2007; **356**: 1099-1109 [PMID: 17360988 DOI: 10.1056/NEJMoa063355]

17 **Saposnik G**, Baibergenova A, Bayer N, Hachinski V. Weekends: a dangerous time for having a stroke? *Stroke* 2007; **38**: 1211-1215 [PMID: 17347472 DOI: 10.1161/01.STR.0000259622.78616.ea]

18 **Dorn SD**, Shah ND, Berg BP, Naessens JM. Effect of weekend hospital admission on gastrointestinal hemorrhage outcomes. *Dig Dis Sci* 2010; **55**: 1658-1666 [PMID: 19672711 DOI: 10.1007/s10620-009-0914-1]

19 **Stroup DF**, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008-2012 [PMID: 10789670]

20 **Thomas BH**, Ciliska D, Dobbins M, Micucci S. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid Based Nurs* 2004; **1**: 176-184 [PMID: 17163895 DOI: 10.1111/j.1524-475X.2004.04006.x]

21 **Armijo-Olivo S**, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract* 2012; **18**: 12-18 [PMID: 20698919 DOI: 10.1111/j.1365-2753.2010.01516.x]

22 **Button LA**, Roberts SE, Evans PA, Goldacre MJ, Akbari A, Dsilva R, Macey S, Williams JG. Hospitalized incidence and case fatality for upper gastrointestinal bleeding from 1999 to 2007: a record linkage study. *Aliment Pharmacol Ther* 2011; **33**: 64-76 [PMID: 21128984 DOI: 10.1111/j.1365-2036.2010.04495.x]

23 **Redelmeier DA**, Bell CM. Weekend worriers. *N Engl J Med* 2007; **356**: 1164-1165 [PMID: 17360995 DOI: 10.1056/NEJMe068310]


**P- Reviewer:** Zhang LL **S- Editor:** Song XX
**L- Editor:** A **E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i2.36

*World J Meta-Anal* 2014 May 26; 2(2): 36-41
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.



# Overview and a meta-analysis of pancreatic melanoma based on case reports

Zhan-Na Yuan, Feng-Hua Liu, Yong Tang

Zhan-Na Yuan, Yong Tang, Department of Pancreatic Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300202, China
Feng-Hua Liu, Department of Gynecology, Key Laboratory of Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300202, China
Author contributions: Yuan ZN reviewed the clinical medical records; Liu FH and Tang Y revised this work; Tang Y revised the paper, participated in the design of the study and performed the statistical analysis.
Supported by National Natural Science Foundation of China, No. 81341068
Correspondence to: Tang Yong, MD, PhD, Department of Pancreatic Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute and Hospital, Hexi District, Huanhuxi Road, Tianjin 300202, China. tli001@163.com
Telephone: +86-22-23540123 Fax: +86-22-23540123
Received: November 24, 2013 Revised: January 3, 2014
Accepted: February 16, 2014
Published online: May 26, 2014

## Abstract

**AIM:** To investigate the clinical characteristics of pancreatic malignant melanoma.

**METHODS:** Here we report a case of pancreatic malignant melanoma and present a meta-analysis of pancreatic melanoma based on data from well-documented case reports.

**RESULTS:** A 32-year-old male presented with abdominal pain and jaundice. Computed tomography showed a non-discrete mass in the head of the pancreas. A pancreatoduodenectomy was performed. Histology revealed pancreatic malignant melanoma. Immunostaining showed that the neoplastic cells were positive for Melan A and S100, and that they expressed HMB-45. A review of 13 publications reporting 15 cases identified main clinical symptoms as abdominal pain (nine cases) and jaundice (three cases). Nine patients had a history of extra-abdominal malignant melanoma. The overall median survival time was 10 mo.

**CONCLUSION:** Pancreatic malignant melanoma is not typified by any special clinical presentation. The diagnosis depends on biopsy or specimen histology. Surgical resection remains the best treatment choice.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Pancreatic malignant melanoma; Survival; Diagnosis; Treatment; Review

**Core tip:** Pancreatic malignant melanoma is rare and only few well-documented cases have been reported. Consequently only sparse information regarding its clinical features and biologic behavior is available. Here we report a case of pancreatic malignant melanoma and present a meta-analysis of pancreatic melanoma based on data from well-documented case reports.

Yuan ZN, Liu FH, Tang Y. Overview and a meta-analysis of pancreatic melanoma based on case reports. *World J Meta-Anal* 2014; 2(2): 36-41 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i2/36.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i2.36

## INTRODUCTION

Malignant melanoma is a form of neoplasm which originates from melanocytes. Melanocytes are pigmented cells located in the dermis or epidermis of the skin, squamous mucous membranes, the retina and leptomeninges. Most melanomas occur at the junction of the skin and mucous membranes[1]. Pancreatic melanoma occurs in approximately 50% of cases in which there is a disseminated neoplastic disease[2]. The formation of melanoma from isolated pancreatic metastasis occurs in no more than 2% of patients with visceral metastases[1,3,4]. Consequently the

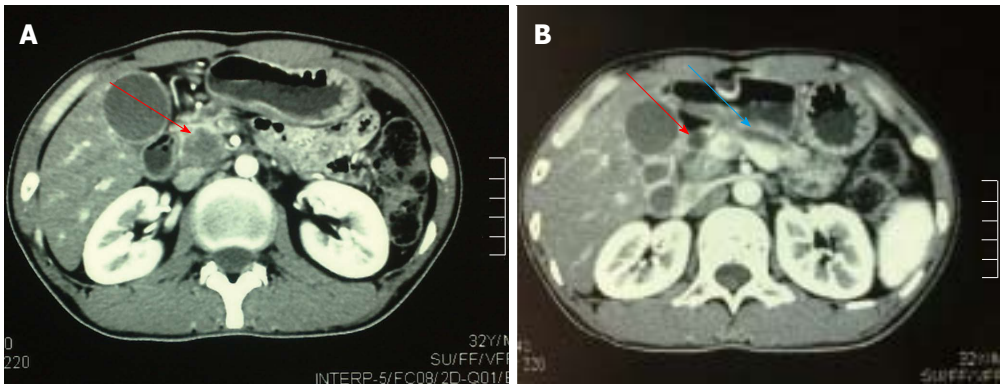



**Figure 1 Computed tomography characteristics of a pancreatic melanoma.** A computed tomography scan demonstrating a hypodense mass in the pancreatic head (A) (red arrow) and dilation of the common bile duct and pancreatic duct (B) (red and blue arrow).

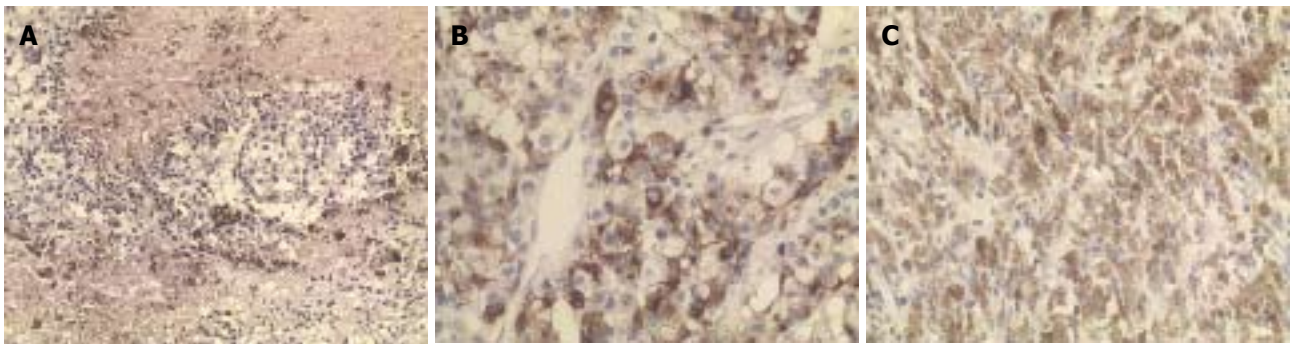



**Figure 2 HE and IHC staining of pancreatic melanoma tissue.** A: HE section showing a spindle cell appearance with melanin pigment production in a patient with a dilated pancreatic duct (arrow) (HE × 100); B: Immunostaining image showing expression of HMB-45 (× 400); C: Immunostaining image showing expression of S100 (× 200).

clinical features of pancreatic melanoma remain unclear.

Here we report a case of pancreatic melanoma treated at our own center. We also provide an overview and meta-analysis of data of published case reports.

## MATERIALS AND METHODS

A review of the literature was performed using MEDLINE (1966 to November 2012), CNKI (1994 to November 2012) and Springer Link (1989 to November 2012) databases, using the key words "pancreas" and "melanoma" to search for case reports and reviews. References cited in case series and case reports were searched to obtain all relevant case reports of pancreatic melanoma. From these, we compiled data on age, gender, symptoms, diagnostic methods, treatment protocols, and survival time. Statistical analysis was performed using Kaplan-Meier curves.

## RESULTS

### *Case report*

A 32-year-old man presented with a 2-wk history of abdominal pain and jaundice and was admitted to Tianjin Cancer Hospital. In 2011, the patient had been diagnosed with malignant skin melanoma which was successfully treated with two cycles of chemotherapy (BUCN and DTIC) at another center. One year ago, ultrasound and computed tomography (CT) revealed a mass in the pancreatic head. At that time the patient showed no sign of jaundice. The mass was successfully treated using gamma knife radiation.

Laboratory tests showed that total bilirubin and direct bilirubin were both elevated (372.3 μmol/L and 257.9 μmol/L, respectively), ALT was 284 μmol/L, AST was 93 μmol/L and alkaline phosphate was 745 μmol/L. Carcinoembryonic antigen and carbohydrate antigen 19.9 (CA 19.9) were normal. Abdominal CT showed a hypodense mass in the head of the pancreas together with dilation of both the common bile duct and pancreatic duct (Figure 1A and B). There was no evidence of lymphadenopathy or vascular encasement.

Based on these findings, pancreatoduodenectomy was performed. Intraoperatively, a lesion 3.5 cm in diameter was found in the pancreatic head, but there were no signs of other abdominal metastases. Histopathology of the resected specimen confirmed the presence of metastatic melanoma.

Immunostaining showed that the neoplastic cells were positive for Melan A, S100 and HMB-45 (Figure 2A, B and C). All resection margins were clear and there was no lymph node involvement. The patient achieved complete remission and was given one cycle of adjuvant chemotherapy (CCNU and DTIC). Liver metastases were found

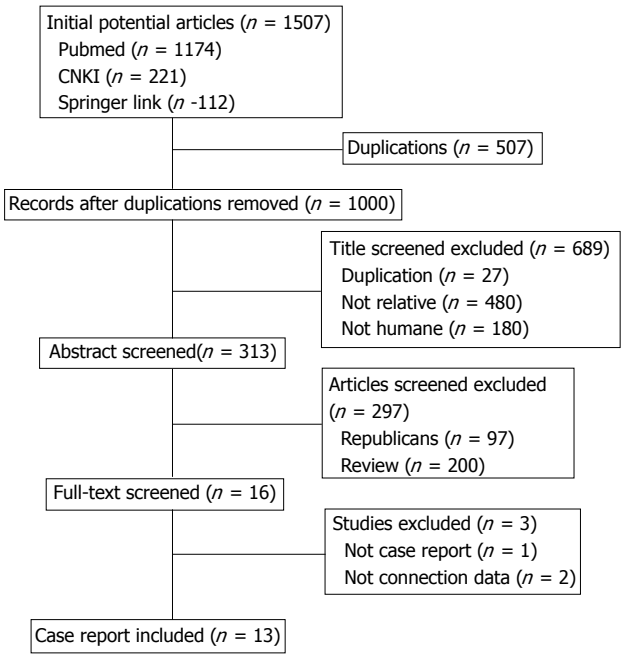



**Figure 3 Process of selection of articles.** The process of searching for articles for inclusion in this article: 1507 articles were found and 1000 articles remained after removal of duplicates. After review of the complete texts, 13 articles met the inclusion criteria and were included.

2 mo postoperatively. The patient was still alive with tumors 6 mo later.

### *Meta-analysis of case reports*

A total of 313 articles reporting pancreatic malignant melanoma were initially identified. Thirteen of them were case reports that included complete data (Figure 3). The 13 publications provided case data on 15 cases of malignant pancreatic melanoma (Table 1). The mean age of the patients was 53.2 years (range, 22 to 83 years) and the male-to-female ratio was 1.2:1. The most common clinical symptoms were abdominal pain (nine cases) and jaundice (three cases). Nine of the 15 had a history of extra-abdominal malignant melanoma.

The tumor imaging features were similar to those of other types of pancreatic carcinoma. In eight cases the lesions were located in the head, and three patients had lesions in the pancreatic body or tail. In two cases the lesions diffused throughout the entire pancreas and in two other patients pancreatic lesions were accompanied by other organ metastasis.

Ten of the 15 patients underwent surgery: four had local resections, three had Whipple operations, and the remaining three underwent total pancreatectomy. The overall median survival was 10 mo (Figure 4).

## DISCUSSION

As the pancreatic melanoma is rare, we have attempted to review its clinical presentations by analysis of published cases. The results of our meta-analysis show that the clinical presentations of pancreatic malignant melanoma

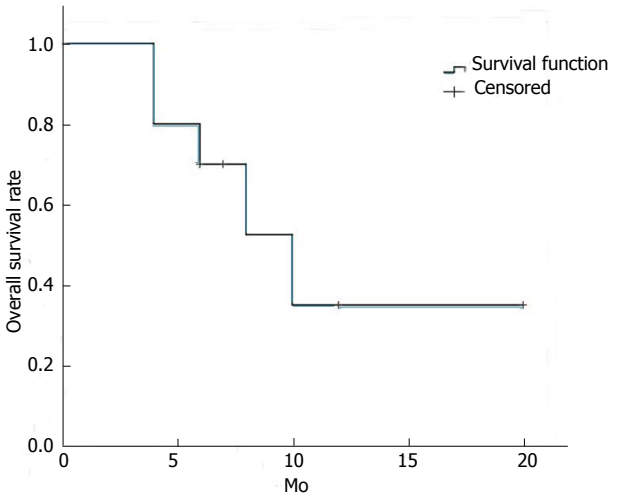



**Figure 4 Kaplan Meier survival curves.** The overall median survival was 10 mo.

are similar to those of other pancreatic malignant tumors and that diagnosis is reliant on biopsy or pathology after surgery. There is evidence to suggest that a history of malignant melanoma is associated with the diagnosis

Malignant melanoma rarely affects intra-abdominal organs, and most cases of visceral melanoma are the result of the spread of advanced malignant melanoma. A previously published autopsy review reported that the most common sites for metastatic gastrointestinal melanoma were the liver (68%), small bowel (58%), colon (22%) and stomach (20%). The duodenum, rectum, esophagus and anus accounted for 1% to 12% of cases[15]. Other researchers have shown that single organ involvement is present in approximately one third of all cases[2,4] of gastrointestinal melanoma, whereas extra-abdominal melanoma spreading to the pancreas only accounts for about 2% of all cases[14].

The lesions of pancreatic malignant melanoma are often located in the head of the pancreas, and the symptoms often mimic those of other pancreatic tumors (including abdominal pain, lack of strength and jaundice). The initial diagnosis of pancreatic malignant melanoma is, therefore, dependent on CT, magnetic resonance imaging or ultrasonography[8,9]. Recently, positron emission tomography has been used to identify sites of metastatic melanoma and for qualitative diagnosis[18,19].

A confirmatory diagnosis of pancreatic malignant melanoma depends on biopsy or pathology of the resected specimen. On gross inspection the pancreatic tumor is often present as a sharply demarcated mass. Histologically, the tumor displays white and brownish, grey or black microcystic areas[7], and is composed of a heterogeneous mix of spindle-shaped cells, plump, polygonal cells with an epithelioid appearance, and multinucleated giant cells. The cells are arranged in fascicles, whorls, nodular patterns, or are of a myxoid-cystic appearance. Most of the tumor cells contain melanin pigmentation. However, there is no recognizable growth pattern.

Immunohistochemical staining is useful in confirming the diagnosis of metastatic melanoma, and often reveals

**Table 1 Case reports of pancreatic malignant melanoma**

| Ref. | Age (yr) | Gender | Symptoms | History | Imaging | Treatment | Prognosis (mo) |
|---|---|---|---|---|---|---|---|
| Lei *et al*[5] | 64 | M | Abdominal pain, fatigue, weight loss | - | CT and MR showed a lesion located in the body and tail of the pancreas. | Surgical resection | - |
| Wu *et al*[6] | 57 | M | Abdominal pain, weight loss | - | CT showed a mass in the head of the pancreas | TP | 6 + |
| Bergmann *et al*[7] | 46 | M | - | - | CT and MRI showed a lesion in the head of the pancreas | Whipple operation | 4 |
| Yang *et al*[8] | 58 | M | Abdominal pain, fatigue, weight loss | - | Ultrasound and CT showed atumor in the pancreas | TP | - |
| Solomons *et al*[9] | 60 | F | Nausea, abdominal pain, obstructive jaundice | - | Ultrasound and CT and MRI revealed a mass in the region of the pancreatic head | Bile duct stenting | 8 |
| Nikfarjam *et al*[10] | 45 | F | Abdominal discomfort | Ocular melanoma 12 yr ago | CT, PET-CT and MRI showed a mass in the head of pancreas | Pylorus-preserving pancreato-duodenectomy and segmental liver resection | 6 + |
| | 55 | M | Abdominal pain | Ocular melanoma 13 yr ago | CT and PET-CT showed a mass in the head of the pancreas | TP | 7 + |
| Kitamura *et al*[11] | 74 | M | - | - | Ultrasound and ERCP disclosed a cyst in the pancreas, communicating with a crater in the stomach. | Local resection | - |
| Stauch *et al*[12] | 68 | M | Obstructive jaundice | Melanoma in right ear 13 yr ago | Ultrasound. ERCP and CT revealed a mass in the head of the pancreas | Whipple operation | - |
| DeWitt *et al*[13] | 33 | M | Abdominal discomfort, obstructive jaundice | | Ultrasound and CT demonstrated a mass in the head of the pancreas | Laparotomy | 6 |
| | 83 | F | Abdominal pain, hematochezia | | CT and EUS found multiple lesions in the body and the tail of the pancreas | Supportive treatment | 10 |
| Dumitrascu *et al*[14] | 43 | F | - | Melanoma in the right eye 4 yr ago | Ultrasound found three metastases: (one in the body of pancreas and two in the right inferior pulmonary lobe) | Central pancreatectomy, right inferior pulmonary lobectomy | 12 + |
| Minoguchi *et al*[15] | 35 | F | _ | Melanoma in the vagina 2 yr ago, and metastasis to the lung 1 yr ago | CT and MRI found a mass in the head of the pancreas | Chemotherapy | 4 |
| Belagyi *et al*[16] | 22 | F | - | Melanoma in the back 9 yr ago, and metastasis to the right ovary 4 yr ago | CT showed the metastases tothe right ovary, the body of the pancreas, bowel and distal stomach, respectively | Resection of all metastases | 4 |
| Camp *et al*[17] | 56 | F | Right upper abdominal pain | Uveal melanoma of the left eye 6 yr ago | CT and PET-CT revealed a hepatic metastasis and a metastasis in the midbody of the pancreas | DP, mesohepatectomy and biochemotherapy with dacarbazine and interferon alpha | 20 + |

TP: Total pancreatectomy; DP: Distal pancreatectomy. M: Month; ST: Survival time; TP: Total Pancreatectomy; CT: Computed tomography; MRI: Magnetic resonance imaging; ERCP: Endoscopic retrograde cholangiopancreatography; PET: Positron emission tomography; EUS: Endoscopic ultrasonography.

the presence of focal staining for S100 and HMB-45 within the pigmented areas[8-10].

Currently there is no satisfactory non-surgical treatment for metastatic melanoma. All of the metastatic tissue needs to be excised if surgery is to offer survival benefit. Complete pancreatic resection has been shown to improve survival even in patients with multiple metastases and in some cases is curative[14-16,20,21-24].

It is estimated that only 1% of all pancreatic resections are performed for metastases[20]. A study of 60 patients with metastatic diseases involving intra-abdominal solid organs showed that the 5-year overall survival (OS) was significantly improved after complete resection (24%) *vs* incomplete resection (0%), but median OS after complete resection was not significantly different for single-site *vs* synchronous multisite metastases[25]. Median disease-free survival (DFS) after complete resection was 15 mo. The 2-year DFS rate after complete resection was

53% for synchronous multi-site metastases *vs* 26% for single-site metastases[25].

It is unclear whether adjuvant therapy helps to improve the survival benefit achieved with surgery. Evidence from uncontrolled phase 2 studies suggests a modest survival benefit with such an approach, but the survival benefit of adjuvant therapy has yet to be demonstrated in randomized trials[26].

For patients unable to undergo resection, combination chemotherapy with dacarbazine, cisplatin, interferon-α-2α and interleukin-2 has been shown to be effective, resulting in a partial response rate of 24% with stable disease being maintained in 48% of cases[27].

The survival outcome of patients with visceral metastases from melanoma is dismal, regardless of histopathological features, with a median survival of 6-12 mo[25,28,29]. Five-year survival rates from several series is less than 10%[19,30-32].

In summary, the present case report and the review of literature show the difficulty in diagnosing a metastatic melanoma in the pancreas, due to its insidious clinical manifestations . In patients with a history of melanoma, a high index of suspicion for metastasis needs to be maintained if they present with seemingly unrelated symptoms. Diagnosis requires histology after surgery or biopsy with special immunohistochemical staining (HMB-45 and S100). This malignancy has a poor prognosis with a median survival of 6 to 12 mo. Surgery is the only effective treatment, and it is unclear whether an adjuvant therapy helps to improve the survival benefit.

## COMMENTS

### Background

Malignant melanoma is a form of neoplasm which originates from melanocytes. Most melanomas occur at the junction of the skin and mucous membranes. Pancreatic melanoma occurs in approximately 50% of cases in which there is a disseminated neoplastic disease. The formation of melanoma from isolated pancreatic metastasis occurs in no more than 2% of patients with visceral metastases. Consequently the clinical features of pancreatic melanoma remain unclear.

### Research frontiers

Malignant melanoma rarely affects intra-abdominal organs, and most cases of visceral melanoma are the result of the spread of advanced malignant melanoma. A previously published autopsy review reported that the most common sites for metastatic gastrointestinal melanoma were the liver (68%), small bowel (58%), colon (22%) and stomach (20%). The duodenum, rectum, esophagus and anus accounted for 1% to 12% of cases. Other researchers have shown that single organ involvement is present in approximately one third of all cases of gastrointestinal melanoma, whereas extra-abdominal melanoma spreading to the pancreas only accounts for about 2% of all cases.

### Innovations and breakthroughs

A confirmatory diagnosis of pancreatic malignant melanoma depends on biopsy or pathology of the resected specimen. On gross inspection the pancreatic tumor is often present as a sharply demarcated mass. Histologically, the tumor displays white and brownish, grey or black microcystic areas, and is composed of a heterogeneous mix of spindle-shaped cells, plump, polygonal cells with an epithelioid appearance, and multinucleated giant cells. The cells are arranged in fascicles, whorls, nodular patterns, or are of a myxoid–cystic appearance. Most of the tumor cells contain melanin pigmentation. However, there is no recognizable growth pattern. Immunohistochemical staining is useful in confirming the diagnosis of metastatic melanoma, and often reveals the presence of focal staining for S100 and HMB-45 within the pigmented areas.

### Applications

The study results suggest that a high index of suspicion for metastasis needs to be maintained if they present with seemingly unrelated symptoms. Diagnosis requires histology after surgery or biopsy with special immunohistochemical staining (HMB-45 and S100). This malignancy has a poor prognosis with a median survival of 6-12 mo. Surgery is the only effective treatment, and it is unclear whether an adjuvant therapy helps to improve the survival benefit.

### Terminology

Malignant melanoma is a form of neoplasm which originates from melanocytes. These are pigmented cells located in the dermis or epidermis of the skin, in squamous mucous membranes, the retina and leptomeninges.

### Peer review

The topic of the research is very relevant and interesting to a wide range of readers around the world, because presentations on pancreatic melanoma are very limited.

## REFERENCES


1 **Patel JK**, Didolkar MS, Pickren JW, Moore RH. Metastatic pattern of malignant melanoma. A study of 216 autopsy cases. *Am J Surg* 1978; **135**: 807-810 [PMID: 665907 DOI: 10.1016/0002-9610(78)90171-X]

2 **Dasgupta TK**, Brasfield rd. metastatic melanoma of the gastrointestinal tract. *Arch Surg* 1964; **88**: 969-973 [PMID: 14132002 DOI: 10.1001/archsurg.1964.01310240065013]

3 **Collaborative Ocular Melanoma Study Group**. Assessment of metastatic disease status at death in 435 patients with large choroidal melanoma in the Collaborative Ocular Melanoma Study (COMS): COMS report no. 15. *Arch Ophthalmol* 2001; **119**: 670-676 [PMID: 11346394 DOI: 10.1001/archopht.119.5.670]

4 **Brodish RJ**, McFadden DW. The pancreas as the solitary site of metastasis from melanoma. *Pancreas* 1993; **8**: 276-278 [PMID: 8460104 DOI: 10.1097/00006676-199303000-00023]

5 **Lei JQ**, Wang XH, Chen Y, Guo SL. Malignant melanoma of the pancreas: case report. Chin J Med Imaging Technol. *Chin J Med Imaging Technol* 2009; **25**: 1329. [DOI: 10.3321/j.issn:1003-3289.2009.07.074]

6 **Wu GJ,** Yang CQ, Huang MX, Wang ZT. Malignant melanoma of the pancreas: case report. *Chin J Hepatobiliary Surg* 2001; **7**: 265 [DOI: 10.3760/cma.j.issn.1007-8118.2001.05.028]

7 **Bergmann F**, Hackert T, Mechtersheimer G, Penzel R, Bläker H, Berger I, Esposito I, Büchler MW, Otto HF. Differential diagnosis of non-epithelial tumors of the pancreas: malignant non-epithelial pancreatic tumor with focal pigmentation. *Virchows Arch* 2004; **444**: 190-193 [PMID: 15046038 DOI: 10.1007/s00428-003-0940-x]

8 **Yang Y**, Li HL, Cao HG. Ultrasonic features of pancreatic malignant melanoma: case report. *Chin J Ultrasonogr* 2000; **9**: 45 [DOI: 10.1046/j.1440-1673.2000.00858.x]

9 **Solomons G**, Gibson RN, Tello RJ. Pancreatic head melanoma producing obstructive jaundice: magnetic resonance characteristics aiding differentiation from adenocarcinoma. *Australas Radiol* 2000; **44**: 471-473 [PMID: 11103552]

10 **Nikfarjam M**, Evans P, Christophi C. Pancreatic resection for metastatic melanoma. *HPB* (Oxford) 2003; **5**: 174-179 [PMID: 18332980 DOI: 10.1080/13651820310015284]

11 **Kitamura Y**, Sakue M, Nishiyama K, Matsumoto S, Tamagawa M, Okada K, Chinzei T, Yamashiro K, Bando K, Fujimoto T. A case of metastatic malignant melanoma mimicking pancreatic pseudocyst. *Gan No Rinsho* 1987; **33**: 748-752 [PMID: 3599484]

12 **Stauch S**, Rösch W. Obstructive jaundice caused by metastasis to the head of the pancreas by a malignant melanoma. *Bildgebung* 1987-1989; **56**: 115-117 [PMID: 3077069]

13 **DeWitt JM**, Chappo J, Sherman S. Endoscopic ultrasound-guided fine-needle aspiration of melanoma metastatic to the pancreas: report of two cases and review. *Endoscopy* 2003; **35**: 219-222 [PMID: 12584640 DOI: 10.1055/s-2003-37258]

14 **Dumitraşcu T**, Dima S, Popescu C, Gheonea DI, Ciurea T, Saftoiu A, Popescu I. An unusual indication for central pancreatectomy--late pancreatic metastasis of ocular malignant melanoma. *Chirurgia* (Bucur) 2008; **103**: 479-485 [PMID: 18780624]

15 **Minoguchi M**, Yanagawa N, Ishikawa C, Sasajima J, Goto M, Okamoto M, Saito Y, Murakami M, Orii Y, Yaosaka T. Pancreatic metastasis of malignant melanoma diagnosed by EUS-guided fine needle aspiration (EUS-FNA). *Nihon Shokakibyo Gakkai Zasshi* 2007; **104**: 1082-1087 [PMID: 17611386]

16 **Belágyi T**, Zsoldos P, Makay R, Issekutz A, Oláh A. Multiorgan resection (including the pancreas) for metastasis of cutaneous malignant melanoma. *JOP* 2006; **7**: 234-240 [PMID: 16525211]

17 **Camp R**, Lind DS, Hemming AW. Combined liver and pancreas resection with biochemotherapy for metastatic ocular melanoma. *J Hepatobiliary Pancreat Surg* 2002; **9**: 519-521 [PMID: 12483277 DOI: 10.1007/s005340200066]

18 **Harrison LE**, Merchant N, Cohen AM, Brennan MF. Pancreaticoduodenectomy for nonperiampullary primary tumors. *Am J Surg* 1997; **174**: 393-395 [PMID: 9337160]

19 **Damian DL**, Fulham MJ, Thompson E, Thompson JF. Positron emission tomography in the detection and management of metastatic melanoma. *Melanoma Res* 1996; **6**: 325-329 [PMID: 8873053 DOI: 10.1097/00008390-199608000-00008]

20 **Ollila DW**, Essner R, Wanek LA, Morton DL. Surgical resection for melanoma metastatic to the gastrointestinal tract. *Arch Surg* 1996; **131**: 975-999[PMID: 8790168]

21 **Prichard RS**, Hill AD, Skehan SJ, O'Higgins NJ. Positron emission tomography for staging and management of malignant melanoma. *Br J Surg* 2002; **89**: 389-396

22 **Hiotis SP**, Klimstra DS, Conlon KC, Brennan MF. Results after pancreatic resection for metastatic lesions. *Ann Surg Oncol* 2002; **9**: 675-679 [PMID: 12167582 DOI: 10.1007/BF02574484]

23 **Roland CF**, van Heerden JA. Nonpancreatic primary tumors with metastasis to the pancreas. *Surg Gynecol Obstet* 1989; **168**: 345-347 [PMID: 2928909]

24 **Hirota T**, Tomida T, Iwasa M, Takahashi K, Kaneda M, Tamaki H. Solitary pancreatic metastasis occurring eight years after nephrectomy for renal cell carcinoma. A case report and surgical review. *Int J Pancreatol* 1996; **19**: 145-153 [PMID: 8723558 DOI: 10.1007/BF02805229]

25 **Wood TF**, DiFronzo LA, Rose DM, Haigh PI, Stern SL, Wanek L, Essner R, Morton DL. Does complete resection of melanoma metastatic to solid intra-abdominal organs improve survival? *Ann Surg Oncol* 2001; **8**: 658-662 [PMID: 11569781 DOI: 10.1007/s10434-001-0658-4]

26 **Spitler LE**, Grossbard ML, Ernstoff MS, Silver G, Jacobs M, Hayes FA, Soong SJ. Adjuvant therapy of stage III and IV malignant melanoma using granulocyte-macrophage colony-stimulating factor. *J Clin Oncol* 2000; **18**: 1614-1621 [PMID: 10764421]

27 **Proebstle TM**, Scheibenbogen C, Sterry W, Keilholz U. A phase II study of dacarbazine, cisplatin, interferon-alpha and high-dose interleukin-2 in 'poor-risk' metastatic melanoma. *Eur J Cancer* 1996; **32A**: 1530-1533 [PMID: 8911113 DOI: 10.1016/0959-8049(96)00135-9]

28 **Sharpless SM**, Das Gupta TK. Surgery for metastatic melanoma. *Semin Surg Oncol* 1998; **14**: 311-318 [PMID: 9588724 DOI: 10.1002/(SICI)1098-2388(199806)14]

29 **Overett TK**, Shiu MH. Surgical treatment of distant metastatic melanoma. Indications and results. *Cancer* 1985; **56**: 1222-1230 [PMID: 4016710 DOI: 10.1002/1097-0142(19850901)56]

30 **Schuchter LM**, Green R, Fraker D. Primary and metastatic diseases in malignant melanoma of the gastrointestinal tract. *Curr Opin Oncol* 2000; **12**: 181-185 [PMID: 10750731 DOI: 10.1097/00001622-200003000-00014]

31 **Balch CM**, Soong SJ, Murad TM, Smith JW, Maddox WA, Durant JR. A multifactorial analysis of melanoma. IV. Prognostic factors in 200 melanoma patients with distant metastases (stage III). *J Clin Oncol* 1983; **1**: 126-134 [PMID: 6668496]

32 **Barth A**, Wanek LA, Morton DL. Prognostic factors in 1,521 melanoma patients with distant metastases. *J Am Coll Surg* 1995; **181**: 193-201 [PMID: 7670677]


**P- Reviewers:** Bashashati M, Kadusevicius E, Wang XH
**S- Editor:** Qi Y **L- Editor:** Wang TQ **E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i2.42

*World J Meta-Anal* 2014 May 26; 2(2): 42-48
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Treatment strategy for gallstone pancreatitis and the timing of cholecystectomy

Chao Hu, Shi-Qiang Shen, Zu-Bing Chen

Chao Hu, Shi-Qiang Shen, Zu-Bing Chen, Department of General Surgery, Renmin Hospital of Wuhan University, Wuhan 430060, Hubei Province, China

Author contributions: Hu C conceived and designed the review, conducted the statistical analyses and drafted the manuscript; Shen SQ critically revised the manuscript, and approved the final version of the manuscript submitted for publication; Chen ZB identified and acquired reports of trials and analyzed the data.

Correspondence to: Shi-Qiang Shen, MD, Department of General Surgery, Renmin Hospital of Wuhan University, Ziyang Road 99 #, Wuhan 430060, Hubei Province, China. swsw2218@hotmail.com
Telephone: +86-27-88041911 Fax: +86-27-88042292
Received: November 23, 2013 Revised: January 18, 2014
Accepted: March 3, 2014
Published online: May 26, 2014

## Abstract

**AIM:** To reviewed the literature and evaluated the scope and timing of the application of endoscopic retrograde cholangiopancreatography (ERCP)/endoscopic sphincterotomy (ES) and cholecystectomy.

**METHODS:** A pooled odds ratio (OR) and a pooled mean difference with the 95%CI were used to assess the enumeration data of included studies. A pooled weighted mean difference (WMD) and a pooled mean difference with the 95%CI were used to assess the measurement data of included studies. Statistical heterogeneity was tested with the $\chi^2$ test. According to forest plots, heterogeneity was not significant, so the fixed effect model was adopted. The significance of the pooled OR was determined by the $Z$ test and statistical significance was considered at $P < 0.05$.

**RESULTS:** Data were collected from two studies (353 patients, 142 in the early cholecystectomy group and 211 in the delayed cholecystectomy group) regarding the length of hospital stay [The WMD was -2.87 (95%CI: -3.36--2.39, $P < 0.01$). Data were collected from four studies (618 patients, 211 in the early cholecystectomy group and 408 in the delayed cholecystectomy group) regarding perioperative complications (OR = 0.94, 95%CI: 0.41-2.12, $P > 0.05$). Data were collected from four studies (618 patients, 211 in the early cholecystectomy group and 408 in the delayed cholecystectomy group) on the number of patients who underwent ERCP± ES postoperatively (OR = 0.80, 95%CI: 0.45-1.41, $P > 0.05$).

**CONCLUSION:** Cholecystectomy offers better protection than ES against further bouts of pancreatitis in patients with gallstone pancreatitis, although ES is an acceptable alternative.

**Key words:** Gallstone pancreatitis; Endoscopic retrograde cholangiopancreatography; Endoscopic sphincterotomy; Cholecystectomy

**Core tip:** In this study we reviewed the literature and evaluated the scope and timing of the application of endoscopic retrograde cholangiopancreatography (ERCP)/endoscopic sphincterotomy (ES) and cholecystectomy. We also carried out a meta-analysis regarding the timing of cholecystectomy and found that early cholecystectomy administered to patients with mild gallstone pancreatitis could reduce the length of hospital stay with no increase in perioperative complications or the incidence of postoperative ERCP ± ES.

Hu C, Shen SQ, Chen ZB. Treatment strategy for gallstone pancreatitis and the timing of cholecystectomy. *World J Meta-Anal* 2014; 2(2): 42-48 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i2/42.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i2.42

## INTRODUCTION

Acute pancreatitis is associated with considerable morbidity and mortality[1]. In fact, approximately 20% of patients with acute pancreatitis will present a severe disease

course, with a 10%-30% mortality rate[2]. The most common causes of acute pancreatitis include gallstones and alcohol abuse which together account for more than 80% of cases, and gallstones alone cause 30%-75% of cases[3,4]. Apart from these two main causes, hyperlipidemia, medication use, hypercalcemia, trauma, familial pancreatitis, and other rare origins also can lead to pancreatitis[5].

The theories of pathogenesis of gallstone pancreatitis (GSP) are still controversial now. The common channel theory proposes that a stone lodged at the ampulla of Vater creates a common channel between the bile and pancreatic ducts, which allows bile reflux into the pancreatic duct (PD) and subsequent activation of pancreatic enzymes[6]. In contrast, the duodenal reflux theory postulates that the passage of biliary stones creates incompetence of the sphincter of Oddi, which results in the reflux of duodenal contents into the PD and subsequently leads to the activation of pancreatic proenzymes and the development of pancreatitis[7]. The third prominent theory is the obstruction theory, which holds that a gallstone or inflammation with edema consequent to gallstone passage may impact or transiently obstruct the common bile duct (CBD) and PD at the level of the ampulla of Vater, which leads to extravasation of digestive enzymes into the gland parenchyma[8]. Recently, many studies have reported that inflammatory mediators, including cytokines, interleukins (IL-6 and IL-8), platelet activating factor, and tumor necrosis factor (TNF-α), also play an important role in the systemic complications of acute pancreatitis[9,10].

Most of the patients with gallstone pancreatitis demonstrate mild to moderate clinical disease, and some can recover with conservative management. The conservative management of gallstone pancreatitis includes bowel rest, fluid resuscitation, total parenteral nutrition or enteral nutrition, and parenteral analgesia, which is similar to the management approach for any form of pancreatitis. However, in contrast to other forms of pancreatitis, for instance, gallstone pancreatitis often requires surgery. According to the pathogenesis of gallstone pancreatitis, we know that removing the gallstones and resolving the obstruction of CBD and PD is the key point for management, although the optimal method for removing the gallstones and resolving the obstruction remains controversial. In particular, the application and timing of endoscopic retrograde cholangiopancreatography (ERCP)/endoscopic sphincterotomy (ES) or cholecystectomy remain the subject of current debate.

### Publication search

To assess the method and timing of surgical treatment for gallstone pancreatitis, we searched the relevant literature from the Cochrane Library and the MEDLINE and EMBASE databases up to september 20th, 2013. The search strategy included the following MESH terms: "gallstone pancreatitis" and "endoscopic retrograde cholangiopancreatography" and "cholecystectomy".

## MATERIALS AND METHODS

The inclusion criteria were as follows: (1) study population: patients with mild gallstone acute pancreatitis (less than 3 Ranson criteria); (2) intervention: ERCP with or without ES *vs* cholecystectomy; (3) outcome measures: the length of hospital stay, the incidence of perioperative complication and the mortality; and (4) study design: randomized controlled trial. The exclusion criteria were as follows: (1) patients with severe gallstone pancreatitis (3 or more Ranson criteria); (2) pancreatitis from other causes; (3) suspected concomitant acute cholangitis; and (4) articles reporting non-comparative studies.

**Table 1 Chief characteristics of the included trials**

| Ref. | Group | POT | No. of patients | LOS | ERCP ± ES postoperative | Complications |
|---|---|---|---|---|---|---|
| Aboulian *et al*[44], 2010 | Early | 35.1 | 25 | 3.5 ± 2.04 | 6 | 0 |
| | Delayed | 77.8 | 25 | 5.8 ± 5.35 | 4 | 0 |
| Rosing *et al*[42], 2007 | Early | 48 | 43 | 4 | 2 | 2 |
| | Delayed | 120 | 177 | 7 | 24 | 9 |
| Taylor *et al*[43], 2004 | Early | 43.2 | 26 | 3.5 | 6 (1) | 3 |
| | Delayed | 55.2 | 20 | 4.7 | 7 (2) | 2 |
| Falor *et al*[45], 2012 | Early | < 48 | 117 | 3.4 ± 2.20 | 19 (12) | 5 |
| | Delayed | > 48 | 186 | 6.3 ± 2.08 | 51 (19) | 9 |

POT: Preoperative time, hours; LOS: Length of hospital stay, days; ERCP: Endoscopic retrograde cholangiopancreatography; ES: Endoscopic sphincterotomy.

### Statistical analysis

For statistical analysis, Review Manager (RevMan) software Version 4.2 was used (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark). A pooled odds ratio (OR) and a pooled mean difference with the 95%CI were used to assess the enumeration data of included studies. A pooled weighted mean difference (WMD) and a pooled mean difference with the 95%CI were used to assess the measurement data of included studies. Statistical heterogeneity was tested with the $\chi^2$ test. According to forest plots, heterogeneity was not significant, so the fixed effect model was adopted. The significance of the pooled OR was determined by the $Z$ test and statistical significance was considered at $P < 0.05$.

## RESULTS

### Eligible studies

We found no randomized controlled trial or retrospective study concerning the application of ERCP with or without ES *vs* cholecystectomy. Therefore, we reviewed the previously published literature for studies evaluating ERCP/ES and cholecystectomy and then assessed the scope of the application of ERCP/ES and cholecystectomy.

Four studies showing that early cholecystectomy for mild gallstone pancreatitis may decrease hospital stay without increasing morbidity or mortality were identified; among these studies, 1 was a randomized controlled trial study and 3 were retrospective studies. All of these studies were single-center studies, and their main characteristics are listed in Table 1. Based on these 4 studies, we

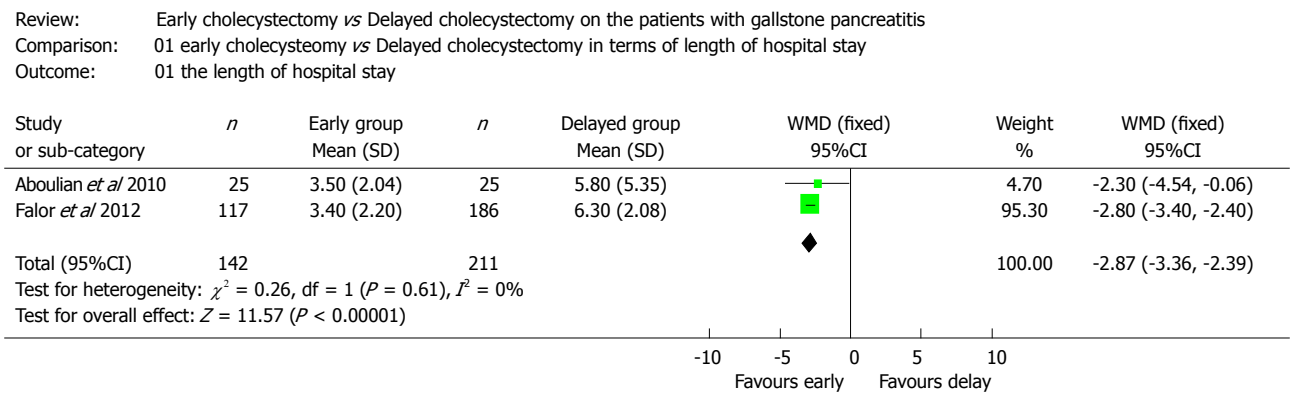



**Figure 1 Forest plot comparing early cholecystectomy *vs* delayed cholecystectomy in terms of length of hospital stay.** WMD: Weighted mean difference.

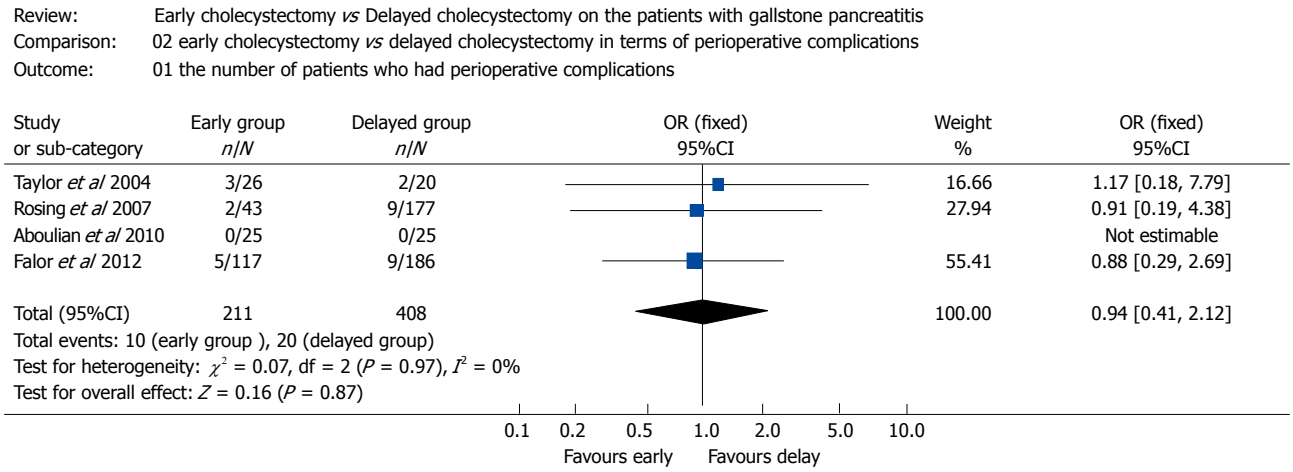



**Figure 2 Forest plot comparing early cholecystectomy *vs* delayed cholecystectomy in terms of perioperative complications.**

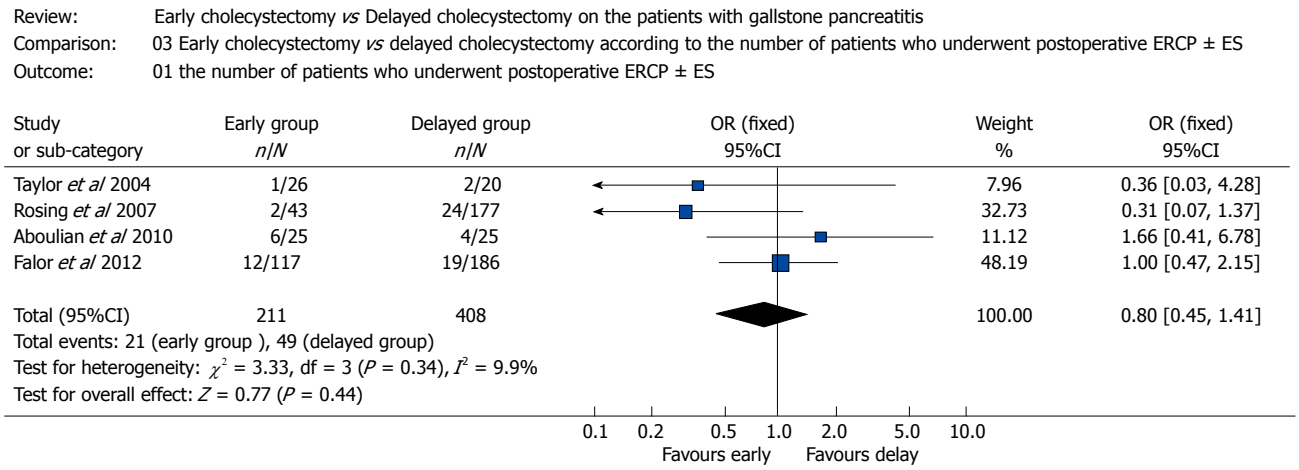



**Figure 3 Forest plot comparing early cholecystectomy vs delayed cholecystectomy according to the number of patients who underwent postoperative endoscopic retrograde cholangiopancreatography ± endoscopic sphincterotomy.** ERCP: Endoscopic retrograde cholangiopancreatography; ES: Endoscopic sphincterotomy.

carried out a meta-analysis regarding the timing of cholecystectomy in patients with mild gallstone pancreatitis.

### *Meta-analysis results*

**Length of hospital stay:** Data were collected from two studies (353 patients, 142 in the early cholecystectomy group and 211 in the delayed cholecystectomy group) regarding the length of hospital stay (LOS). The overall meta-analysis (Figure 1) indicated that early cholecystectomy for the patients with mild gallstone pancreatitis could significantly decrease the LOS compared with delayed cholecystectomy; the WMD was -2.87, (95%CI:

-3.36--2.39, $P < 0.01$).

**Perioperative complications:** Data were collected from four studies (618 patients, 211 in the early cholecystectomy group and 408 in the delayed cholecystectomy group) regarding perioperative complications. The overall meta-analysis (Figure 2) showed that the effect of early cholecystectomy on perioperative complications in patients with mild gallstone pancreatitis was inconclusive (OR = 0.94, 95%CI: 0.41-2.12, $P > 0.05$).

**Postoperative ERCP ± ES:** Data were collected from four studies (618 patients, 211 in the early cholecystectomy group and 408 in the delayed cholecystectomy group) on the number of patients who underwent ERCP ± ES postoperatively. The overall meta-analysis (Figure 3) indicated that the effect of early cholecystectomy on the number of patients who underwent ERCP ± ES postoperatively was also inconclusive (OR = 0.80, 95%CI: 0.45-1.41, $P > 0.05$).

## DISCUSSION

ERCP/ES is one of the most technically challenging endoscopic procedures performed by gastroenterologists and surgeons with special training. This approach is well suited for the evaluation of diseases of the bile ducts and pancreas. However, it has been reported that the rate of complications due to ERCP ranges from 5% to 7%, with pancreatitis and bleeding as the most common complications[11-14]. Moreover, death directly related to the procedure occurs in approximately 0.4% of cases in large prospective studies[13,14]. Therefore, many surgeons do not advocate for use ERCP/ES as the preferred treatment strategy for gallstone pancreatitis. A meta-analysis conducted by Uy *et al*[15] showed a trend that the mortality of patients receiving early ERCP with or without sphincterotomy in the setting of acute gallstone pancreatitis without cholangitis was higher than that of patients who received conservative management (RR = 1.92, 95%CI: 0.86-4.32 $P = 0.11$). A study conducted by Sees *et al*[16] also suggested that in patients with mild to moderate gallstone pancreatitis, preoperative ERCP/ES led to a higher rate of postoperative pancreatitis and an increased length of hospitalization, especially in patients with choledocholithiasis. In addition, numerous other studies have also shown that in patients with acute biliary pancreatitis who received early ERCP/ES presented more severe complications and did not benefit from the procedure[17-19].

Although ERCP/ES has limitations related due to potential morbidity and mortality, ERCP/ES also has its incomparable advantage. For example, a recent large-sample retrospective cohort study conducted by Hwang *et al*[20] analyzed the risk of developing recurrent gallstone pancreatitis in patients who never received a cholecystectomy. In this study, a total of 1119 patients were identified, including 802 who received no intervention and 317 who received ERCP. After a median follow-up of 2.3 years, the risk of recurrent pancreatitis was 8.2% and 17.1% in patients who received ERCP and those who received no intervention, respectively ($P < 0.001$). The median time to recurrence was also longer in the ERCP group(11.3 *vs* 10.1 mo), and the recurrence rates at 1, 2 and 5 years in the ERCP group were all lower than those in the non-intervention group (5.2%, 7.4%, 11.1% *vs* 11.3%, 16.1%, 22.7%). These results suggested that ERCP could mitigate the risk of recurrent pancreatitis and should be considered during initial hospitalization if cholecystectomy was not performed. Kaw *et al*[21] also found that the recurrence of pancreatitis after ERCP with ES alone for the treatment of gallstone pancreatitis was rare and the complication rate, as well as the rate of serious complications, was higher in comparison to cholecystectomy but not significant. Also recently, the most acceptable range for the use of ERCP was shown to be in high-risk surgical patients and the elderly[22-25], as cholecystectomy may lead to a higher rate of morbidity and mortality in these patients. Another condition for which ERCP/ES is recommended is severe gallstone pancreatitis, and it is reported that up to 25% of patients with gallstone pancreatitis experienced severe pancreatitis[6]. In addition to intensive care monitoring for the management of associated systemic complications, most of these patients are initially managed with ERCP and ES if common bile duct stones are found. Moreover, surgery is delayed for approximately 3 wk because numerous studies have shown that early cholecystectomy is associated with increased morbidity and mortality[10,26-28]. There are also other conditions that may benefit from ERCP/ES, such as cholangitis or obstructive jaundice associated with gallstone pancreatitis[29,30].

The United Kingdom guidelines for the management of gallstone pancreatitis published by the British Society of Gastroenterology in 1998[31] and 2005[32] suggest that all patients with mild GSP should be offered definitive treatment, *i.e.*, cholecystectomy, if they are fit for surgery. Several studies have also reported that cholecystectomy can reduce the risk of recurrence or further attack for patients with GSP[33-36]. Moreover, cholecystectomy is a lower-risk procedure, and the biliary complications attributable to gallstones in the gallbladder, such as cholecystitis, are avoided[37]. Therefore, for the treatment of GSP, it is generally agreed that all patients with acute GSP who are sufficiently fit should undergo cholecystectomy.

Recent reports have also presented controversial findings regarding the timing of surgery. According to several experts cholecystectomy for GSP can be delayed from the inciting pancreatitis. In their opinion, postponing cholecystectomy 6-8 wk may reduce the acute inflammation and decrease the technical difficulty of cholecystectomy and possibly lower the conversion rate[38]. However, in the last decade, the standard timing for cholecystectomy has supported earlier intervention. For example, a large-sample (8631 patients) observational study in 2009 divided study patients into four groups according to the treatment policy: cholecystectomy during index stay (group 1); cholecystectomy within 30 d of index admission (group 2); sphincterotomy, but not cholecystectomy within 30 d of index admission (group 3); and neither

cholecystectomy nor sphincterotomy within 30 d of index admission (group 4). The results showed that group 1 required a slightly longer length of stay in the hospital compared to group 2, although these patients had had much fewer readmissions for biliary complications. Similar to the study mentioned above, several observational studies have supported the use of doing cholecystectomy within 2 wk of index admission to prevent gallstone pancreatitis recurrence[39,40].

Some studies have even suggested that 2 wk after index admission is too long to wait, as patients presented with unacceptably high rates for recurrence of gallstone pancreatitis. For example, Ito *et al*[41], found that even following the guidelines that performing cholecystectomy within 2 wk of index admission (according to the guidelines), there remained a 33% rate of biliary related complications including a 13% recurrence rate for pancreatitis rate. As a result, so some experts have suggested that cholecystectomy should be performed within 48 h of admission for mild gallstone pancreatitis before serum pancreatic enzymes are normal and before abdominal pain has completely resolved. Indeed, some retrospective studies have confirmed that early cholecystectomy can shorten the length of hospital stay with no increase in complications or mortality or technical difficulty[42-45].

The timing of cholecystectomy in patients with severe pancreatitis is different from that in patients with mild disease, as the former patients often have contraindications to surgery due to various physiologic derangements or the inflammatory response associated with severe pancreatitis. Thus, the treatment strategy for these patients is early ERCP/ES and cholecystectomy is delayed until the signs of lung injury and systemic disturbance have resolved, generally between 3 and 4 wk.Various studies have shown less controversy over the timing of cholecystectomy for these patients.

Our present meta-analysis also obtained the conclusion that early cholecystectomy in patients with mild gallstone pancreatitis could reduce the length of hospital stay with no increase in perioperative complications or the incidence of postoperative ERCP ± ES. However, there were some limitations to our meta-analysis. First, only four studies were included in our study, of which only one study was a randomized controlled trial and the other three were retrospective studies. Second, individual patient data were not used. Third, the publication bias may have been present because the quantity of the studies included was so small. Therefore, to confirm that early cholecystectomy is indeed both superior to delayed cholecystectomy and safe, a multicenter randomized controlled trial is needed.

In conclusion, cholecystectomy offers better protection than ES against further bouts of pancreatitis in patients with GSP, although ES is an acceptable alternative. Patients with mild to moderate gallstone pancreatitis should receive cholecystectomy during index admission, although patients with more severe disease will require early ERCP/ES and cholecystectomy should therefore be delayed, depending on the clinical circumstances. For patients at a patients have high risk for surgery or the elderly, ERCP and ES should represent first-choice treatments.

## COMMENTS

### Background

Acute pancreatitis is associated with considerable morbidity and mortality. Approximately 30%-75% of cases are caused by gallstones and the management of gallstone pancreatitis remains challenging.

### Research frontiers

The role and timing of endoscopic retrograde cholangiopancreatography (ERCP)/endoscopic sphincterotomy (ES) and cholecystectomy in the management of gallstone pancreatitis have been a subject of much debate in recent years.

### Innovations and breakthroughs

In this study the authors reviewed the literature and evaluated the scope and timing of the application of ERCP/ES and cholecystectomy. The authors also carried out a meta-analysis regarding the timing of cholecystectomy and found that early cholecystectomy administered to patients with mild gallstone pancreatitis could reduce the length of hospital stay with no increase in perioperative complications or the incidence of postpoerative ERCP ± ES.

### Applications

For patients at a patients have high risk for surgery or the elderly, ERCP and ES should represent first-choice treatments.

### Peer review

Hu *et al* evaluated the management of gallstone pancreatitis. They demonstrated that cholecystectomy offers better protection than ES against further bouts of pancreatitis in patients with gallstone pancreatitis, but ES is an acceptable alternative. Patients with mild to moderate gallstone pancreatitis should have cholecystectomy during index admission, but patients with more severe disease will require early ERCP/ES and cholecystectomy should be delayed, depending on the clinical circumstances. If patients have high risk of surgery or the elderly, ERCP and ES is the first choice.

## REFERENCES


1 **Mitchell RM**, Byrne MF, Baillie J. Pancreatitis. *Lancet* 2003; **361**: 1447-1455 [PMID: 12727412 DOI: 10.1016/S0140-6736(03)13139-X]

2 **Whitcomb DC**. Clinical practice. Acute pancreatitis. *N Engl J Med* 2006; **354**: 2142-2150 [PMID: 16707751 DOI: 10.1056/NEJMcp054958]

3 **Steer ML**. Classification and pathogenesis of pancreatitis. *Surg Clin North Am* 1989; **69**: 467-480 [PMID: 2658159]

4 **Agrawal S**, Jonnalagadda S. Gallstones, from gallbladder to gut. Management options for diverse complications. *Postgrad Med* 2000; **108**: 143-146, 149-153 [PMID: 11004941 DOI: 10.3810/pgm.2000.09.1.1212]

5 **Sakorafas GH**, Tsiotou AG. Etiology and pathogenesis of acute pancreatitis: current concepts. *J Clin Gastroenterol* 2000; **30**: 343-356 [PMID: 10875461 DOI: 10.1097/00004836-200006000-00002]

6 **Frakes JT**. Biliary pancreatitis: a review. Emphasizing appropriate endoscopic intervention. *J Clin Gastroenterol* 1999; **28**: 97-109 [PMID: 10078816 DOI: 10.1097/00004836-199903000-00004]

7 **McCutcheon AD**. Gallstone pancreatitis. *Lancet* 1988; **2**: 166 [PMID: 2899218 DOI: 10.1016/S0140-6736(88)90719-2]

8 **Senninger N**, Moody FG, Coelho JC, Van Buren DH. The role of biliary obstruction in the pathogenesis of acute pancreatitis in the opossum. *Surgery* 1986; **99**: 688-693 [PMID: 2424109]

9 **Kusske AM**, Rongione AJ, Reber HA. Cytokines and acute pancreatitis. *Gastroenterology* 1996; **110**: 639-642 [PMID: 8566616 DOI: 10.1053/gast.1996.v110.agast960639]

10 **Mergener K**, Baillie J. Endoscopic treatment for acute biliary

pancreatitis. When and in whom? *Gastroenterol Clin North Am* 1999; **28**: 601-613, ix [PMID: 10503139 DOI: 10.1016/S0889-8553(05)70076-5]
11 **Mallery JS**, Baron TH, Dominitz JA, Goldstein JL, Hirota WK, Jacobson BC, Leighton JA, Raddawi HM, Varg JJ, Waring JP, Fanelli RD, Wheeler-Harbough J, Eisen GM, Faigel DO. Complications of ERCP. *Gastrointest Endosc* 2003; **57**: 633-638 [PMID: 12709688 DOI: 10.1053/ge.2003.v57.amge030576633]
12 **Cotton PB**, Lehman G, Vennes J, Geenen JE, Russell RC, Meyers WC, Liguory C, Nickl N. Endoscopic sphincterotomy complications and their management: an attempt at consensus. *Gastrointest Endosc* 1991; **37**: 383-393 [PMID: 2070995 DOI: 10.1016/S0016-5107(91)70740-2]
13 **Freeman ML**, Nelson DB, Sherman S, Haber GB, Herman ME, Dorsher PJ, Moore JP, Fennerty MB, Ryan ME, Shaw MJ, Lande JD, Pheley AM. Complications of endoscopic biliary sphincterotomy. *N Engl J Med* 1996; **335**: 909-918 [PMID: 8782497 DOI: 10.1056/NEJM199609263351301]
14 **Loperfido S**, Angelini G, Benedetti G, Chilovi F, Costan F, De Berardinis F, De Bernardin M, Ederle A, Fina P, Fratton A. Major early complications from diagnostic and therapeutic ERCP: a prospective multicenter study. *Gastrointest Endosc* 1998; **48**: 1-10 [PMID: 9684657 DOI: 10.1016/S0016-5107(98)70121-X]
15 **Uy MC**, Daez ML, Sy PP, Banez VP, Espinosa WZ, Talingdan-Te MC. Early ERCP in acute gallstone pancreatitis without cholangitis: a meta-analysis. *JOP* 2009; **10**: 299-305 [PMID: 19454823]
16 **Sees DW**, Martin RR. Comparison of preoperative endoscopic retrograde cholangiopancreatography and laparoscopic cholecystectomy with operative management of gallstone pancreatitis. *Am J Surg* 1997; **174**: 719-722 [PMID: 9409604 DOI: 10.1016/S0002-9610(97)00216-X]
17 **Fölsch UR**, Nitsche R, Lüdtke R, Hilgers RA, Creutzfeldt W. Early ERCP and papillotomy compared with conservative treatment for acute biliary pancreatitis. The German Study Group on Acute Biliary Pancreatitis. *N Engl J Med* 1997; **336**: 237-242 [PMID: 8995085 DOI: 10.1056/NEJM199701233360401]
18 **Chang L**, Lo S, Stabile BE, Lewis RJ, Toosie K, de Virgilio C. Preoperative versus postoperative endoscopic retrograde cholangiopancreatography in mild to moderate gallstone pancreatitis: a prospective randomized trial. *Ann Surg* 2000; **231**: 82-87 [PMID: 10636106 DOI: 10.1097/00000658-200001000-00012]
19 **Himal HS**. Preoperative endoscopic retrograde cholangiopancreatography (ERCP) is not necessary in mild gallstone pancreatitis. *Surg Endosc* 1999; **13**: 782-783 [PMID: 10430684 DOI: 10.1007/s004649901098]
20 **Hwang SS**, Li BH, Haigh PI. Gallstone pancreatitis without cholecystectomy. *JAMA Surg* 2013; **148**: 867-872 [PMID: 23884515 DOI: 10.1001/jamasurg.2013.3033]
21 **Kaw M**, Al-Antably Y, Kaw P. Management of gallstone pancreatitis: cholecystectomy or ERCP and endoscopic sphincterotomy. *Gastrointest Endosc* 2002; **56**: 61-65 [PMID: 12085036 DOI: 10.1067/mge.2002.125544]
22 **Pezzilli R**. Endoscopic sphincterotomy in acute biliary pancreatitis: A question of anesthesiological risk. *World J Gastrointest Endosc* 2009; **1**: 17-20 [PMID: 21160646 DOI: 10.4253/wjge.v1.i1.17]
23 **Trust MD**, Sheffield KM, Boyd CA, Benarroch-Gampel J, Zhang D, Townsend CM, Riall TS. Gallstone pancreatitis in older patients: Are we operating enough? *Surgery* 2011; **150**: 515-525 [PMID: 21878238 DOI: 10.1016/j.surg.2011.07.072]
24 **Welbourn CR**, Beckly DE, Eyre-Brook IA. Endoscopic sphincterotomy without cholecystectomy for gall stone pancreatitis. *Gut* 1995; **37**: 119-120 [PMID: 7672659 DOI: 10.1136/gut.37.1.119]
25 **Bignell M**, Dearing M, Hindmarsh A, Rhodes M. ERCP and endoscopic sphincterotomy (ES): a safe and definitive management of gallstone pancreatitis with the gallbladder left in situ. *J Gastrointest Surg* 2011; **15**: 2205-2210 [PMID: 22005898 DOI: 10.1007/s11605-011-1729-x.]
26 **Tang E**, Stain SC, Tang G, Froes E, Berne TV. Timing of laparoscopic surgery in gallstone pancreatitis. *Arch Surg* 1995; **130**: 496-499; discussion 496-500 [PMID: 7748087]
27 **Neoptolemos JP**, Carr-Locke DL, London NJ, Bailey IA, James D, Fossard DP. Controlled trial of urgent endoscopic retrograde cholangiopancreatography and endoscopic sphincterotomy versus conservative treatment for acute pancreatitis due to gallstones. *Lancet* 1988; **2**: 979-983 [PMID: 2902491 DOI: 10.1016/S0140-6736(88)90740-4]
28 **Fan ST**, Lai EC, Mok FP, Lo CM, Zheng SS, Wong J. Early treatment of acute biliary pancreatitis by endoscopic papillotomy. *N Engl J Med* 1993; **328**: 228-232 [PMID: 8418402 DOI: 10.1056/NEJM199301283280402]
29 **Byrne MF**. Gallstone pancreatitis--who really needs an ERCP? *Can J Gastroenterol* 2006; **20**: 15-17 [PMID: 16432554]
30 **West DM**, Adrales GL, Schwartz RW. Current diagnosis and management of gallstone pancreatitis. *Curr Surg* 2002; **59**: 296-298 [PMID: 16093150 DOI: 10.1016/S0149-7944(01)00615-8]
31 United Kingdom guidelines for the management of acute pancreatitis. British Society of Gastroenterology. *Gut* 1998; **42** Suppl 2: S1-13 [PMID: 9764029 DOI: 10.1136/gut.42.2008.S1]
32 UK guidelines for the management of acute pancreatitis. *Gut* 2005; **54** Suppl 3: iii1-iii9 [PMID: 15831893]
33 **Moreau JA**, Zinsmeister AR, Melton LJ, DiMagno EP. Gallstone pancreatitis and the effect of cholecystectomy: a population-based cohort study. *Mayo Clin Proc* 1988; **63**: 466-473 [PMID: 3361956 DOI: 10.1016/S0025-6196(12)65644-4]
34 **Judkins SE**, Moore EE, Witt JE, Barnett CC, Biffl WL, Burlew CC, Johnson JL. Surgeons provide definitive care to patients with gallstone pancreatitis. *Am J Surg* 2011; **202**: 673-677; discussion 677-678 [PMID: 22014646 DOI: 10.1016/j.amjsurg.2011.06.031.]
35 **Frei GJ**, Frei VT, Thirlby RC, McClelland RN. Biliary pancreatitis: clinical presentation and surgical management. *Am J Surg* 1986; **151**: 170-175 [PMID: 2418700 DOI: 10.1016/0002-9610(86)90028-0]
36 **El-Dhuwaib Y**, Deakin M, David GG, Durkin D, Corless DJ, Slavin JP. Definitive management of gallstone pancreatitis in England. *Ann R Coll Surg Engl* 2012; **94**: 402-406 [PMID: 22943329 DOI: 10.1308/003588412X13171221591934]
37 **Sandzén B**, Haapamäki MM, Nilsson E, Stenlund HC, Oman M. Cholecystectomy and sphincterotomy in patients with mild acute biliary pancreatitis in Sweden 1988 - 2003: a nationwide register study. *BMC Gastroenterol* 2009; **9**: 80 [PMID: 19852782 DOI: 10.1186/1471-230X-9-80]
38 **Tate JJ**, Lau WY, Li AK. Laparoscopic cholecystectomy for biliary pancreatitis. *Br J Surg* 1994; **81**: 720-722 [PMID: 8044561 DOI: 10.1002/bjs.1800810533]
39 **McCullough LK**, Sutherland FR, Preshaw R, Kim S. Gallstone pancreatitis: does discharge and readmission for cholecystectomy affect outcome? *HPB* (Oxford) 2003; **5**: 96-99 [PMID: 18332964 DOI: 10.1080/13651820310001135]
40 **Monkhouse SJ**, Court EL, Dash I, Coombs NJ. Two-week target for laparoscopic cholecystectomy following gallstone pancreatitis is achievable and cost neutral. *Br J Surg* 2009; **96**: 751-755 [PMID: 19526610 DOI: 10.1002/bjs.6644]
41 **Ito K**, Ito H, Whang EE. Timing of cholecystectomy for biliary pancreatitis: do the data support current guidelines? *J Gastrointest Surg* 2008; **12**: 2164-2170 [PMID: 18636298 DOI: 10.1007/s11605-008-0603-y]
42 **Rosing DK**, de Virgilio C, Yaghoubian A, Putnam BA, El Masry M, Kaji A, Stabile BE. Early cholecystectomy for mild to moderate gallstone pancreatitis shortens hospital stay. *J Am Coll Surg* 2007; **205**: 762-766 [PMID: 18035259 DOI: 10.1016/j.jamcollsurg.2007.06.291]
43 **Taylor E**, Wong C. The optimal timing of laparoscopic cholecystectomy in mild gallstone pancreatitis. *Am Surg* 2004; **70**:

971-975 [PMID: 15586508]

44 **Aboulian A**, Chan T, Yaghoubian A, Kaji AH, Putnam B, Neville A, Stabile BE, de Virgilio C. Early cholecystectomy safely decreases hospital stay in patients with mild gallstone pancreatitis: a randomized prospective study. *Ann Surg* 2010; **251**: 615-619 [PMID: 20101174 DOI: 10.1097/SLA.0b013e3181c38f1f]

45 **Falor AE**, de Virgilio C, Stabile BE, Kaji AH, Caton A, Kokubun BA, Schmit PJ, Thompson JE, Saltzman DJ. Early laparoscopic cholecystectomy for mild gallstone pancreatitis: time for a paradigm shift. *Arch Surg* 2012; **147**: 1031-1035 [DOI: 10.1001/archsurg.2012.1473]

**P- Reviewers:** Biondi-Zoccai G, Cui WP, Messori A
**S- Editor:** Qi Y **L- Editor:** A **E- Editor:** Wu HL



**ISSN** 2308-3840 (online)

# World Journal of *Meta-Analysis*

*World J Meta-Anal* 2014 August 26; 2(3): 49-134



Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

Contents

Quarterly Volume 2 Number 3 August 26, 2014

**REVIEW**

49 Stroke and depression: A bidirectional link
*Del Zotto E, Costa P, Morotti A, Poli L, de Giuli V, Giossi A, Volonghi I, Callea A, Padovani A, Pezzini A*

**OBSERVATIONAL STUDY**

64 Development of the first Arabic cognitive dental anxiety scale for children and young adults
*Al-Namankany A, Ashley P, Petrie A*

**META-ANALYSIS**

71 Interventions to improve treatment adherence among adolescents: A meta-analysis
*Miller TA, Bannon BL, DiMatteo MR*

78 Efficacy of therapeutic ultrasound *vs* sham ultrasound on pain and physical function in people with knee osteoarthritis: A meta-analysis of randomized controlled trials
*MacIntyre NJ, Negm A, Loyola-Sánchez A, Bhandari M*

91 Role of statin on mortality outcome in pneumonia patients: A meta-analysis
*Saha L, Kumar N, Khosla P, Kaur S*

98 Robot-assisted *vs* laparoscopy-assisted gastrectomy for gastric cancer: A meta-analysis based on 3518 subjects
*Lin ZD, Liu M, Tang D, Li H, Zhang BM*

107 Comparison of functional outcomes after retropubic, laparoscopic and robot-assisted radical prostatectomy: A meta-analysis
*Shi MJ, Yang J, Meng XY, Li S, Liu T, Fang ZH, Cao R, Wang XH*

127 Decision-tree analysis for cost-effective management of solitary pulmonary nodules in China
*Lu B, Sun LX, Yan X, Ai ZZ, Xu JZ*

**Contents**

*World Journal of Meta-Analysis*
**Volume 2 Number 3 August 26, 2014**

**APPENDIX** I-V Instructions to authors

**ABOUT COVER** Editorial Board Member of *World Journal of Meta-Analysis*, Alessandro Pezzini, MD, Dipartimento di Scienze Mediche e Chirurgiche, Clinica Neurologica, Università degli Studi di Brescia, P.le Spedali Civili, 1, 25123 Brescia, Italy

**AIM AND SCOPE**

*World Journal of Meta-Analysis* (*World J Meta-Anal*, *WJMA*, online ISSN 2308-3840, DOI: 10.13105) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians, with a specific focus on meta-analysis, systematic review, mixed-treatment comparison, meta-regression, overview of reviews.

*WJMA* covers a variety of clinical medical fields including allergy, anesthesiology, cardiac medicine, clinical genetics, clinical neurology, critical care, dentistry, dermatology, emergency medicine, endocrinology, family medicine, gastroenterology and hepatology, geriatrics and gerontology, hematology, immunology, infectious diseases, internal medicine, obstetrics and gynecology, oncology, ophthalmology, orthopedics, otolaryngology, pathology, pediatrics, peripheral vascular disease, psychiatry, radiology, rehabilitation, respiratory medicine, rheumatology, surgery, toxicology, transplantation, and urology and nephrology, while maintaining its unique dedication to systematic reviews and meta-analyses.

**INDEXING/ABSTRACTING** *World Journal of Meta-Analysis* is now indexed in Digital Object Identifier.

**FLYLEAF** I-IV Editorial Board

**EDITORS FOR THIS ISSUE**

**Responsible Assistant Editor:** *Xiang Li*
**Responsible Electronic Editor:** *Ya-Jing Lu*
**Proofing Editor-in-Chief:** *Lian-Sheng Ma*
**Responsible Science Editor:** *Yuan Qi*
**Proofing Editorial Office Director:** *Jin-Lei Wang*

NAME OF JOURNAL
*World Journal of Meta-Analysis*

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Quarterly

EDITORS-IN-CHIEF
**Giuseppe Biondi-Zoccai, MD, Assistant Professor,** Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
*World Journal of Meta-Analysis*
Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
August 26, 2014

COPYRIGHT
© 2014 Baishideng Publishing Group Inc. Articles published by this Open-Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

SPECIAL STATEMENT
All articles published in journals owned by the Baishideng Publishing Group (BPG) represent the views and opinions of their authors, and not the views, opinions or policies of the BPG, except where otherwise explicitly indicated.

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i3.49

*World J Meta-Anal* 2014 August 26; 2(3): 49-63
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

REVIEW

# Stroke and depression: A bidirectional link

Elisabetta Del Zotto, Paolo Costa, Andrea Morotti, Loris Poli, Valeria de Giuli, Alessia Giossi, Irene Volonghi, Antonio Callea, Alessandro Padovani, Alessandro Pezzini

Elisabetta Del Zotto, U.O. di Recupero e Rieducazione Funzionale, IRCCS Fondazione Don Gnocchi, 20148 Milano, Italy
Paolo Costa, Andrea Morotti, Loris Poli, Valeria de Giuli, Irene Volonghi, Alessandro Padovani, Alessandro Pezzini, Dipartimento di Scienze Cliniche e Sperimentali, Clinica Neurologica, Università degli Studi di Brescia, 25123 Brescia, Italy
Alessia Giossi, U.O Neurologia, Istituto Clinico "S. Anna", 25127 Brescia, Italy
Antonio Callea, Dipartimento Fisiopatologia Medico-Chirurgica e dei Trapianti, Sezione di Neuroscienze, Università degli Studi di Milano, 20122 Milano, Italy
Author contributions: Del Zotto E, Costa P, Morotti A, Poli L, de Giuli V, Giossi A, Volonghi I, Callea A, Padovani A and Pezzini A contributed to this paper; with substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; drafting the article or revising it critically for important intellectual content; and with final approval of the version to be published.
Correspondence to: Alessandro Pezzini, MD, Dipartimento di Scienze Cliniche e Sperimentali, Clinica Neurologica, Università degli Studi di Brescia, P.le Spedali Civili, 1, 25123 Brescia, Italy. ale_pezzini@hotmail.com
Telephone: +39-30-3384086 Fax: +39-30-3384086
Received: March 10, 2014 Revised: May 23, 2014
Accepted: June 10, 2014
Published online: August 26, 2014

## Abstract

A number of studies have assessed the influence of depression on the risk of cardiovascular disease. A growing literature indicates a link between depression and cerebrovascular events, although the direction of this association remains unclear. Numerous data have emerged suggesting an association between depressive symptoms and subsequent risk of stroke, thus leading to the hypothesis that a direct causality between depression and stroke exists. Notwithstanding, how depression may act as a risk factor for stroke is still unclear. Depression might be linked to stroke *via* neuroendocrine and inflammation effects, through correlation with major comorbidities such as hypertension and diabetes or by intervention of lifestyle behavioral mediators. Finally, antidepressant medications have recently drawn attention for a possible association with increased risk of stroke, although such findings remain uncertain. Depression has been also established as an important consequence after stroke, exerting a significant adverse impact on the course of motor recovery, social functioning and, overall, on quality of life. Post stroke depression occurs in nearly one third of stroke cases, but the exact mechanism leading to depression after stroke is still incompletely understood. In this article, we will review contemporary epidemiologic studies, discuss potential mechanisms and specific aspects of the complex relation between depression and stroke.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Depression; Mood disorders; Stroke; Post-stroke depression; Antidepressant medications; Cerebrovascular disease

**Core tip:** A number of studies have assessed the influence of depression on the risk of cardiovascular disease. A growing literature indicates a link between depression and cerebrovascular events, although the direction of this association remains unclear. Numerous data have emerged suggesting an association between depressive symptoms and subsequent risk of stroke, thus leading to the hypothesis that a direct causality between depression and stroke exists. Depression has been also established as an important consequence after stroke, affecting functional recovery and quality of life. Contemporary epidemiologic studies, potential mechanisms and specific aspects of the complex relation between depression and stroke will be discussed.

Del Zotto E, Costa P, Morotti A, Poli L, de Giuli V, Giossi A, Volonghi I, Callea A, Padovani A, Pezzini A. Stroke and depression: A bidirectional link. *World J Meta-Anal* 2014; 2(3): 49-63 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i3/49.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i3.49

## INTRODUCTION

According to definition of the Diagnostic and Statistical Manual of Mental Disorders-Ⅳ (DSM-Ⅳ), major depressive disorder consists of two or more episodes of depressed mood, loss of interest or diminished sense of pleasure in usual activities for more than 2 wk, in addition to other depressive features sufficient to cause clinically important psychological or physical distress, or functional impairment, that is atypical for usual behavior and not attributable to a medical condition or bereavement[1].

Depression is a highly prevalent condition worldwide, more common among women, with a lifetime prevalence of more than 16% in the general population, and a consequent impact on public health[2]. In the last decades, a number of studies have investigated the influence of depression on the risk of developing chronic condition such as diabetes[3] and hypertension[4], and also cardiovascular disease[5]. Adding to these findings, a growing literature indicates a link between depression and cerebrovascular events, although the exact mechanisms of this association remain unclear.

The presence of depression has been established as an important consequence after stroke[6], affecting functional recovery and quality of life[7]. In addition to this evidence, numerous data have emerged pointing toward a relation between depressive symptoms and subsequent risk of stroke as well as common predisposing conditions such as hypertension, diabetes or lifestyle behavioral factor[8], thus leading to the hypothesis that a direct causality between depression and stroke exists[9].

## IS DEPRESSION A RISK FACTOR FOR STROKE?

Several observational studies investigated the relation between depression and the risk of subsequent stroke, with conflicting results (Table 1).

Data from 10 studies published before 2005 were pooled in a first meta-analysis which detected an association between depression and risk of stroke, but with a significant heterogeneity among the studies included[38]. Subsequently, many other studies were published and recently summarized in two more detailed meta-analyses[39,40], strengthening the evidence of a possible role of depression as a modifiable risk factor for stroke. According to the meta-analysis of Pan and coworkers, including prospective cohort studies, the pooled HR of stroke among patients with depression is 1.45 (95%CI: 1.29-1.63). Stratifying analysis by pathological stroke subtype, the pooled HR for ischemic stroke is 1.25 (95%CI: 1.11-1.40), while there is no significant influence of depression on the pooled risk of hemorrhagic stroke[39]. However, since very few studies have analyzed the association by stroke subtype we cannot draw any conclusion in this regard.

Likewise, a predisposing effect of depression was also observed by Dong *et al*[40] in their meta-analysis (pooled RR = 1.34; 95%CI: 1.17-1.54 in depressed subjects compared to non-depressed). Even, the INTERSTROKE study, an international multicenter case-control study designed to establish the association of traditional and emerging risk factors with stroke in countries of high, middle, and low income, reported a similar magnitude of the association between depression and an increased risk of all stroke and ischemic stroke (OR = 1.35; 99%CI: 1.10-1.66, and 1.47; 99%CI: 1.19-1.83, respectively), but not intracerebral haemorrhagic stroke[41].

Both meta-analyzes found no difference in pooled risk stratifying data by gender[39,40]. Although few studies reported age-stratified results, the age difference in the depression-stroke association was also evaluated, resulting in a increased risk in younger subjects [mean age < 65 years, HR = 1.77 (95%CI: 1.30-2.41); mean age ⩾ 65 years, HR = 1.30 (95%CI: 1.18-1.44)][39]. In this regard, the Framingham study was the first to find evidence of effect modification by age, examining the elderly and non-elderly groups separately and documenting the association between depressive symptoms and stroke risk in those aged ⩽ 65 years[26]. Similarly, the Established Populations for Epidemiologic Studies of the Elderly (EPESE) which examined an older population with subgroups aged 65-74 years, and 75 years or older, observed this relation with age, with depression associated with increased stroke risk in younger but not in older participants[21]. Recently, in line with these data, the Intervention Project on Cerebrovascular Diseases and Dementia in the District of Ebersberg (INVADE trial), a population-based longitudinal study, corroborated the association between depression and the risk of ischemic stroke, particularly in women and patients younger than 65 years[35]. These findings suggest the hypothesis that differences may exist in the depression-associated stroke risk in sub-groups of subjects defined by age and sex, which needs to be investigated and confirmed in further studies.

Is this enough to establish a causal association between depression and subsequent risk of stroke? As pointed out by many authors, most of the studies reported so far present several methodological limitations, and there is evidence of a significant heterogeneity.

First, this depends on differences in study design, sample size and population characteristics. Exclusion of patients with history of stroke is important to avoid the possibility of reverse causality and bias in risk estimation. Both meta-analyzes observed a temporal relationship between depression and stroke by including only first-time stroke events that occurred after baseline assessments of depression[39,40]. However the possibility that undiagnosed stroke may have caused depression remains, despite efforts to exclude enrollment of participants with preexisting stroke at study entry. Furthermore, the reverse causality hypothesis might apply to the association with stroke as well, because depressive symptoms might be markers of preexisting cerebrovascular disease, as suggested by the vascular depression hypothesis[42]. In this regard, the results of two recent analyses from the Rotterdam

**Table 1 Risk of stroke according to depression status in prospective studies**

| Ref. | Population | Follow-up years, (period) | No. of stroke cases | Risk of total stroke (95%CI) | Risk of ischemic stroke (95%CI) | Risk of hemorrhagic stroke (95%CI) | Adjustment for covariates and confounders | Depression assessment | Stroke ascertainment |
|---|---|---|---|---|---|---|---|---|---|
| Vogt *et al*[10] | 2573 Men and women aged ⩾ 18 yr | 15 (1970/71-1985) | NA | HR 0.84 (0.57-1.22) | | | Age, sex, socioeconomic status, length of health plan membership, subjective health status, smoking | Depressive index | Death index or vital records |
| Wassertheil-Smoller *et al*[11] | 4367 Men and women aged ⩾ 60 yr | 4.5 (1985-1990) | 204 | RR 0.85 (0.45-1.64) | | | Age, baseline depression, race, education, smoking, diabetes, history of CVD, activities of daily living | CES-D | Medical records |
| Everson *et al*[12] | 6676 Men and women aged 16-94 yr | 29 (1965-1983) | 169 | RR 1.54 (1.06-2.22) | | | Age, sex, race, education, hypertension, diabetes, smoking status, alcohol, BMI | HPL Depression Scale | Death certificates |
| Whooley *et al*[13] | 7518 Women aged ⩾ 65 yr | 6 (1988-1994) | 94 | HR 1.7 (0.8-3.5) | | | Age, history of MI, stroke, hypertension, diabetes, smoking, perceived health, cognitive function | GDS | Medical records |
| Jonas *et al*[14] | 6095 Men and women aged 25-74 yr | 16 (1971/75-1992) | 483 | RR 1.73 (1.30-2.31) | | | Age, SBP, education, smoking, BMI, alcohol, physical activity, cholesterol, diabetes, history of heart disease | GWS | Hospital records and death certificates |
| Larson *et al*[15] | 1703 Men and women aged ⩾ 18 yr | 13 (1980/1983-1993/1996) | 95 | OR 2.67 (1.08-6.63) | | | Age, sex, race, educational attainment, smoking status, history of diabetes, heart problems, high blood pressure | DIS | Self-report and death certificates |
| Ohira *et al*[16] | 879 Men and women aged 40-78 yr | 10.3 (1985-1996) | 69 | RR 1.9 (1.1–3.5) | RR 2.7 (1.2–6.0) | RR 0.9 (0.3–3.1) | Age, sex, smoking status, alcohol use, BMI, SBP, serum total cholesterol level, current treatment with antihypertensive medication, and history of diabetes | Zung SDS | National insurance claims, medical records, clinical diagnosis and death certificates |
| Ostir *et al*[17] | 2478 Men and women aged ⩾ 65 yr | 6 (1986-1992) | 340 | RR 1.30 (0.85-1.99) | | | Age, sex, marital status, household income, education, smoking status, BMI, heart attack, diabetes, hypertension | CES-D | Self-report and death certificates |
| May *et al*[18] | 2124 Men aged 45-59 yr | 14 (1984/1988-1998) | 130 | | HR 1.26 (0.85–1.85) | | Age, social class, marital status, smoking status, alcohol use, BMI, SBP, comorbidity (ischemic heart disease, diabetes, respiratory disease, or retirement due to ill health) | GHQ | Medical records |
| Wassertheil-Smoller *et al*[19] | 93676 Women aged 50-79 yr | 4.1 (1993/1998-2000) | 751 | RR 1.01 (0.78-1.3)[1]; 1.45 (1.11-1.9)[2] | | | Age, race, education, income, smoking status, BMI, physical activity, hormone use, high cholesterol level requiring medications, diabetes, and hypertension | CES-D and DIS | Self-report and medical redords |
| Gump *et al*[20] | 11216 Men aged 35-57 yr | 18.4 (1979/1981-1999) | 167 | HR 1.48 (0.93-2.36) | | | Age, intervention group, race, educational attainment, smoking at baseline and visit 6, SBP, alcohol consumption, fasting cholesterol and the occurrence of nonfatal cardiovascular events during the trial | CES-D | Death certificates |
| Avendano *et al*[21] | 2812 Men and women aged ⩾ 65 yr | 12 (1982-1994) | 270 | HR age 65-74: 3.05 (1.63–5.7); age ⩾ 75: 0.95 (0.46–1.98) | | | Age, sex, race, education and income | CES-D | Self report and medical records |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Stürmer *et al*[22] | 3920 Men and women aged 40-65 yr | 8.5 (1992/1995-2002/2003) | 62 | RR 1.53 (0.83-2.8) | | Age, sex, education, smoking status, alcohol consumption, BMI, exercise, family histories of MI, stroke and cancer, comorbidity | Standardized personality questionnaires | Medical records and death certificates |
| Kamphuis *et al*[23] | 799 Men aged 70-90 yr | 7.4 (1989/1991-2000) | 66 | HR 3.41 (1.69-6.9) | | Age, country, education, BMI, smoking, alcohol intake, blood pressure, cholesterol, and physical activity | Zung SDS | Death certificates |
| Arbelaez *et al*[24] | 5525 Men and women aged ⩾ 65 yr | 11 (1989-2000) | 611 | HR 1.25 (1.02–1.53) | | Age, sex, race, occupation, income, education level, marital status, smoking status, BMI, blood cholesterol, highdensity lipoprotein cholesterol, lowdensity lipoprotein cholesterol, triglycerides levels, comorbidity (hypertension, diabetes, CHD), CRP | CES-D | Self report and medical records |
| Kawamura *et al*[25] | 535 Men and women aged ⩾ 65 yr | 6.3 (1985-2000) | 103 | RR 1.25 (0.82–1.90) | | NA | Zung SDS and physician diagnosis | Death certificates |
| Salaycik *et al*[26] | 4120 Men and women aged ⩾ 29 yr | 8 (1990/1998-1998/2006) | 228 | HR Age < 65: 3.59 (1.76–7.3); Age ⩾ 65: 0.93 (0.59–1.47) | | Age, sex, smoking status, blood pressure, diabetes, atrial fibrillation, CVD, left ventricular hypertrophy on the ECG | CES-D or ADM use | Medical records (Stroke+TIA) |
| Bos *et al*[27] | 4424 Men and women aged ⩾ 61 yr | 5.8 (1997/1999-2005) | 291 | HR 1.21 (0.80–1.83) | HR 1.43 (0.87-2.35) | Age, sex, smoking status, SBP, intima–media thickness, medication use (antithrombotic drug, antihypertensive drug, cholesterol lowering drug, psycholeptic drug and psychoanaleptics drug), comorbidity (diabetes, MI, PTCA or CABG, TIA) | CES-D | Medical records |
| Lee *et al*[28] | 4962 Men and women 18-44 yr | 5 (1998-2003) | 98 | HR 5.43 (3.47-8.5) | | Age, sex, geographic location, income, urbanization, substance abuse, comorbidity (diabetes, hypertension, hyperlipidemia, renal disease) | Physician diagnosis | Medical records |
| Liebetrau *et al*[29] | 401 Men and women aged 85 yr | 3 (1986/1987-1989/1990) | 56 | HR 2.6 (1.5-4.6) | | NA | DSMMD-III | Self report and medical records |
| Surtees *et al*[30] | 20627 Men and women aged 41-80 yr | 8.5 (1996/2000-2006) | 595 | HR 1.08 (0.67–1.75) | | Age, sex, social class, education, smoking status, obesity, SBP, total cholesterol, diabetes, hypertension treatment, MI, family history of stroke, and ADM use | HLEQ | Medical records |
| Whooley *et al*[8] | 20627 Men and women aged mean 63 yr and 68 yr[3] | 4.8 (2000/2002-2008) | 47 | HR 1.47 (0.70-3.11) | | Age, smoking, medication adherence, physical activity, CRP, left ventricular ejection fraction, history of MI, stroke, heart failure, diabetes | PHQ | Self report and medical records |
| Glymour *et al*[31] | 19087 Men and women aged ⩾ 50 yr | 8.1 (1996-2006) | 1864 | HR 1.25 (1.12- 1.39) | | Age, overweight, ethnicity, birth place, parental educations, self education, income, wealth and marital status, year of enrollment, alcohol and smoking status, and self-reported diagnoses of hypertension, diabetes, or heart disease | CES-D | Self report |
| Nabi *et al*[32] | 23282 Men and women aged 20-54 yr | 7 (1998-2005) | 129 | HR 0.87 (0.57–1.32) | | Age, sex, education, smoking status, alcohol use, sedentary lifestyle, obesity,comorbidity (hypertension, diabetes, incident CHD) | BDI | Medical records |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Peters *et al*[33] | 2656 Men and women aged ⩾ 80 yr | 2.1 (2001-2007) | 97 | HR 1.82 (1.19–2.78) | | | Age, sex, treatment group, treatment allocation, country area, educational level, living status, number of comorbidities, previous CVD, previous treatment for hypertension and hypertension | GDS | Self report and medical records |
| Pan *et al*[34] | 80574 Women aged 54-79 yr | 6 (2000-2006) | 1033 | RR 1.29 (1.13-1.48) | RR 1.11 (0.91-1.35) | RR 1.20 (0.80-1.79) | Age, marital status, ethnicity, smoking status, alcohol consumption, BMI, physical activity level, menopausal status, postmenopausal hormone therapy, current aspirin or multivitamin use, Dietary Approaches to Stop Hypertension dietary score, parental history of MI | MHI or self reported diagnosis or ADM use | Self report, medical records and death certificates |
| Seifert *et al*[35] | 3852 Men and women aged > 55 yr | 6.1 (2001/2003-2009) | 156 | HR 1.3 (0.9-1.89) | | | Age, sex, BMI, smoking, hypertension, diabetes, hyperlipidemia, previous MI, previous TIA, previous stroke, history of atrial fibrillation, physical activity | GDS or ADM use | Medical reports |
| Majed *et al*[36] | 9601 Men aged 48-64 yr | 10 (1991/1993-2003) | 136 | HR 1.41 (0.95-2.11) | HR 1.65 (1.07-2.55) | | Age; study centers; socioeconomic factors, including marital status, education level, employment status, physical activity, smoking status, daily alcohol intake, SBP, use of anti-hypertensive drugs, BMI, total and high-density lipoprotein cholesterol, treatment for diabetes, ADM use. | CES-D | Hospital records and general practitioner records |
| Jackson *et al*[37] | 10547 Women aged 47-52 yr | 12 (1998-2010) | 177 | OR 1.94 (1.37- 2.74) | | | Age, education, homeownership, hypertension, diabetes mellitus, heart disease, hysterectomy/oophorectom, smoking, alcohol use, physical activity, BMI | CES-D or ADM use | Self report and death certificates |

[1]No baseline CVD; [2]With baseline CVD; [3]Group with depression mean age 63 years, group without depression mean age 68 years. n Scale: Human Population Laboratory Depression Scale; GDS: Geriatric. CES-D: Center for Epidemiologic Studies Depression Scale; HPL Depression Depression Scale; GWS: General Well-Being Schedule; DIS: diagnostic interview schedule; Zung SDS: Zung Self-Rating Depression Scale; GHQ: General Health Questionnaire; ADM: Antidepressant medication; DSMMD: Diagnostic and Statistical Manual of Mental Disorders; HLEQ: Health and Life Experiences Questionnaire; PHQ: Patient Health Questionnaire; BDI: Beck Depression Inventory; MHI: Mental Health Index; CI: Confidence interval; NA: Not available; HR: Hazard ratio; RR: Relative risk; OR: Odds ratio; CVD: Cardiovascular disease; BMI: Body mass index; MI: Myocardial infarction; CHD: Coronary heart disease; CRP: C reactive protein; SBP: Systolic blood pressure; ECG: Electrocardiogram; PTCA: Percutaneous transluminal angioplasty; CABG: Coronary artery bypass graft.

Study[43] and the Health and Retirement Study[31] did not support this hypothesis.

Second, in most studies the influence of several confounders on the final results was not considered. These include, for example, behavioral and lifestyle factors such as alcohol consumption, dietary factors and physical activity, as well as socioeconomic status, which raise the possibility that residual confounders may be partially responsible for the relation between depression and stroke. In this regard, data from a prospective study including a cohort of 1017 subjects with stable coronary disease, aimed at evaluating the influence of baseline disease severity and potential biological or behavioral mediators on the association of depressive symptoms with subsequent cardiovascular events, point towards a role of poor health behaviors, particularly physical inactivity[8].

Third, in most of the studies reported so far there are differences in stroke and depression measures. Stroke case ascertainment is based on a variety of sources, including medical records, death certificates, clinical diagnoses and self-reports. The assessment of depression is more often based on self-rating questionnaires than on a psychiatric structured interview and the DSM criteria.

Although the link between depression and stroke is evident using diagnostic measures of depression[15,28], it is more consistent when using measures of depressive symptoms[14,16-17,26-27,30]. Most studies used the Center for Epidemiological Studies Depression scale (CES-D), a self-rating screening test for depression with an adequate internal consistency and reliability. The pooled RR for studies that used CES-D for depression assessment has been estimated to be 1.23 (95%CI: 1.13-1.34)[40]. However, we should also consider other types of mood disorder, such as bipolar disorder, which are not measured. In this regard, bipolar disorder has been suggested to be a potential risk factor for vascular disease[44] and stroke (HR = 1.24, 95%CI: 1.12-1.38) after adjusting for patients' demographic characteristics, comorbid medical disorders and socioeconomic status[45].

These drawbacks support the need for further studies aimed at investigating the link between affective disorders and stroke. Moreover, both stroke and depression have a high prevalence and incidence. It was estimated that approximately 3.9% ($n$ = 273000) of stroke cases in the United States could be related to depression[39]. This raises important clinical and public health implications because of the potential of reducing stroke risk by prevention and treatment of mood disorders.

## HOW CAN DEPRESSION LEAD TO STROKE?

So far, various and no fully convincing evidence has been produced to explain how depression may act as a risk factor for vascular disease, including stroke.

A first hypothesis is based on the physiological disturbances linked with depression. These changes have mostly been investigated in relation to cardiovascular disease, and not specifically to stroke. Depression has known neuroendocrine effects, that is, an enhanced activity of the hypothalamic pituitary adrenocortical (HPA) axis and sympathoadrenal hyperactivity[46,47]. HPA axis disturbances predict increased circulating catecholamines, endothelial dysfunction, platelet activation and reduced heart rate variability, which could influence stroke risk[46,48]. The evidence for an involvement of autonomic cardiovascular dysregulation in depressed patients is revealed also by an increased heart rate response to physical stressors, baroreceptor sensitivity and ventricular instability[46,48]. Autonomic dysfunction is also known to influence the risk of atrial fibrillation and depression has been shown to predict atrial fibrillation recurrence after cardioversion[49], suggesting a possible challenging link between depression and stroke that deserves further investigations.

Another mechanism whereby depression can affect the risk of stroke is an inflammation effect[50]. Inflammatory markers such as C-reactive protein, IL-1 and IL-6 have been suggested to be associated with depression[51,52]. Sparse results also pointed towards a role of some genetic[53] and biological markers of thrombotic risk, such as increased levels of fibrinogen[54] and increased serotonin and platelet activation[55]. The hypercoagulable, platelet-activating and inflammatory effects of depression may all be operant in increasing the risk of cardiovascular events in depressed patients. Notwithstanding, the degree to which these predisposing conditions might explain a significant alteration of the risk of stroke in depressed patients is unknown.

Since depression is correlated with major comorbidities, such as hypertension[4] and diabetes[3] probably trough increased adrenergic activity, a third hypothesis is that depression influences stroke risk through the development of hypertension or diabetes or both. In this regard, some authors have suggested that depression may be a sign of preexisting cerebrovascular disease[56]. According to the vascular depression hypothesis, a small-vessel disease secondary to hypertension or diabetes might predispose or at least exacerbate depressive symptoms through impairment of brain regions involved in the regulation of emotions as a direct result of disruption of frontal-subcortical circuits[42]. Data from Rotterdam Study indicate an association between the presence of depressive symptoms and the risk of stroke in the general elderly population, but only in men and not in women and the authors discussed the possibility that depressive symptoms could be the expression of a cerebral vascular damage[27]. In this view, the association between depressive symptoms and later onset ischemic stroke can be considered an epiphenomenon. The vascular depression hypothesis is also supported by the presence and severity of white matter hyperintensities in elderly groups with depression[57,58]. By contrast, data from the Health and Retirement Study, a large national study, provide evidence that depressive symptoms predicted an increased risk of stroke independently of memory impairment, considered a probable early manifestation of cerebral vascular injury[31]. This finding suggests that depression is independently associated with stroke rather than a marker of cerebrovascular damage, possibly through other mechanisms that increase stroke risk.

A fourth hypothesis is that the depression-stroke pathway is modulated by the intervention of behavioral mediators, such as smoking, physical inactivity, poor diet and lack of medication adherence, all of which are modifiable factors. In a prospective cohort study of more than 1000 outpatients with stable coronary heart disease, the association between depressive symptoms and cardiovascular events, including stroke, resulted non significant after adjustment for physical activity and other health behaviors (HR = 1.05; 95%CI: 0.79-1.40), suggesting that the increased risk of cardiovascular events associated with depression could potentially be prevented by behavior modifications, especially physical exercise[8]. Moreover, in depressed patients the control of vascular risk factors may be suboptimal because of non-adherence to medical treatment[59]. In this regard, a meta-analysis showed depression to be a risk factor for reduced medication compliance, with an OR = 3.03 (95%CI: 1.96-4.89)[59].

Finally, depression and stroke might be linked *via* effects of antidepressant medications (ADM).

## ANTIDEPRESSANT MEDICATION USE AND STROKE RISK

The trend of antidepressant use has increased in many countries, including the United States and Europe[60,61]. ADM use has recently drawn attention for a possible association with increased risk of cardiovascular events, although such findings remain uncertain[62-65]. There is evidence that ADM exposure is correlated with bleeding complication[66], increased inflammation[67], weight gain[68], cardiac toxic effects[69] and hypertension[70] thus resulting in a possible effects on vascular outcome. The link of antidepressants with stroke development has also been investigated in several studies, with different results. A 20% to 50% increased risk of stroke associated with ADM was shown in a large case-control study[71] and in a case-crossover study[72]. Recently, data from 6 prospective studies published before 2011 were pooled in a secondary analysis of the meta-analysis of Pan and coworkers, which showed a positive association between ADM use and stroke risk, with an estimated HR = 1.41 (95%CI: 1.25-1.59)[39]. Negative findings in randomized trials[73] and case-control studies[74,75] had also been reported. However, misclassification of depression and the absence of information on dose and duration of treatment in

many studies limit the possibility to understand the link between antidepressant use and stroke. A further confounder is the prescription of ADM for conditions other than depression such as insomnia, headaches and neuropathic pain.

Overall, these findings point out a "paradox" related to the fact that depression is a potential risk factor for stroke, but so it appears to be the use of ADM prescribed to treat depression[76]. A key point is whether antidepressants exposure may be considered a surrogate for depression severity, rather than a causal mechanism. The drug-disease association may be expression of underlying differences in vascular risk factor, including depression, among the exposed patients. A recent population-based cohort study found a significant association between depression and risk of stroke regardless of exposure to antidepressant while in patients using only antidepressants an increased risk of stroke was not observed[77]. The authors suggested as explanation the possibility of "confounding by indication", that is the situation in which "the indication for drug use could confound the drug-disease association so that it appears as if the drugs causes the disease"[77].

Among the different classes of antidepressants, selective serotonin reuptake inhibitors (SSRIs) are the first line agents for management of depression today, owing to their relative safety in overdose, tolerability and well-established efficacy. Nevertheless, data from a well-characterized cohort of > 80000 United States middle-aged and elderly women of the Nurses' Health Study during 6 years of follow-up showed that ADM use was associated with an increased stroke risk (HR = 1.30; 95%CI: 1.08-1.55), with a significant association for selective serotonin reuptake inhibitors (HR = 1.39; 95%CI: 1.13-1.72), but not for other ADM[34], suggesting different effects for SSRIs on the risk of stroke in line with what observed in other studies[63,72].

Despite the antiplatelet effect, SSRIs exposure was associated with an increased risk of ischemic stroke[71,72,77]. A possible biological explanation for this association is that serotonin receptors may act on smooth muscle cells leading to a vasoconstriction and, thus, favoring the thrombotic process in atherosclerotic cerebral arteries[78-80].

Consistent with the evidence that SSRIs may increase the risk of bleeding complications owing to the blockade of serotonin reuptake and secondary depletion of platelet serotonin, which may inhibit platelet aggregation[81], a possible association between SSRIs exposure and risk of hemorrhagic stroke has been also reported[63,72], though with mixed results[71]. Recently, a systematic review and meta-analysis of 16 observational studies was performed to determine the association between SSRIs use and risk of brain hemorrhage, showing a significant association with both intracranial (RR = 1.51; 95%CI: 1.26-1.81) and intracerebral bleeding (RR = 1.42; 95%CI: 1.23-1.65)[82]. Moreover, SSRIs use increased significantly the risk of hemorrhagic stroke in patients concomitantly using oral anticoagulants compared with patients receiving only oral anticoagulants (RR = 1.56, 95%CI: 1.33-1.83)[82]. This raises the issue, when considering SSRIs prescription, of an appropriate patient selection. In particular, caution should be used in those subjects with intrinsic risk factors for intracerebral hemorrhage, such as oral anticoagulant exposure, cerebral amyloid angiopathy or severe alcohol abuse[83]. However, more data are needed in this setting.

## POST-STROKE DEPRESSION

Mood disorders are common and important sequelae of stroke. Depression, anxiety disorder, apathy, catastrophic reactions and psychosis are frequently observed after stroke[84]. In stroke patients, neuropsychiatric complications may exert a significant adverse impact on the course of motor recovery, social functioning and, overall, on quality of life[84].

## DIAGNOSIS OF PSD AND METHODOLOGICAL PROBLEMS

Post stroke depression (PSD) has been associated with increased disability[85,86], impaired rehabilitation outcomes[87,88], and mortality[88,89]. Although many studies have investigated the occurrence of depression among patients with stroke, a real estimate of its prevalence is difficult as a consequence of the wide variability across studies. This is due in part to methodological aspects, such as study population and timing of assessment, and in part to complexity in recognition, assessment, and diagnosis of depression. Furthermore, a contribution to reported differences in the prevalence of PSD may also arise from diagnostic tools used for detection of this disorder. The diagnosis of PSD was assessed on the basis of structured interviews using the diagnostic standards defined by the DSM, while in other studies the assessment is based on the use of cutoff scores in different rating scales. A recent review suggested that, among depression scales, the CES-D, the Hamilton Depression Rating Scale and the Patient Health Questionnaire-9 are adequate options, but they should not be used without a more detailed clinical assessment for an accurate identification of depression in stroke patients[90]. Moreover, there is an obvious risk of under or overestimation in the diagnosis of PSD[91,92]. In fact, stroke may produce somatic symptoms such as fatigue, sleep disturbance, appetite disturbances that might lead to an overdiagnosis of PSD, while post-stroke neurological disabilities, including aphasia or cognitive impairment, may cause under-recognition of PSD. An under-diagnosis of PSD may be also observed when its assessment is made by non-psychiatrists[93].

As a consequence of this, the issue of how to diagnose depression in stroke patients has been the focus of a large number of studies. The presence of physiological symptoms such as psychomotor retardation, and disturbances in appetite, sleep, and sexual interest that can be

related either to stroke or PSD may affect the diagnosis. Stroke is one of the few conditions listed in the DSM-IV as "directly" causing depression. In this case, PSD is classified within the group of "Mood disorders due to stroke, with depressive features" or "with major depressive-like episode"[1].

The validity of DSM diagnostic criteria for depression among patients with stroke has been assessed in different studies. Depressive syndrome in patients with post-stroke major depression is similar to that observed in patients with major depression without a known medical cause[94]. Furthermore, all the symptoms used for the diagnosis of major depression following stroke are significantly more common among depressed patients compared with non-depressed[91,95,96]. However, differences between symptoms in major PSD and primary major depression were also described, with more likely catastrophic reactions, hyper-emotionalism, and diurnal mood variations in patients with PSD[97,98]. In a recent study, Cumming and coworkers confirmed that PSD has a phenomenological profile similar to that of depression unrelated to brain injury regarding psychological and somatic symptoms. This suggests that the diagnosis of depression based on DSM criteria is valid in patients with stroke[99].

## EPIDEMIOLOGICAL ASPECTS AND COURSE OF PSD

According to the meta-analysis of Hackett *et al*[6] including studies published up to 2004, the pooled estimated frequency of depression was of 33% (95%CI: 29%-36%) at any time after acute stroke. The assessment of prevalence rates of PSD is complicated by a considerable variation across studies because of the variability in mood assessment, difference in the selection of cases (*i.e.*, variation in stroke features, clinical characteristics, source of patient recruitment), and timing of assessment. A lower prevalence rates has been generally observed in population studies than in studies conducted in acute hospital setting, rehabilitation hospitals or outpatients clinics, suggesting a potential selection bias. Moreover, the risk of depression is expected to be higher in the first few months after stroke. Quite surprisingly, the meta-analysis showed consistency in the overall frequency of depression across population-based, hospital-based and rehabilitation-based studies and different time intervals from stroke onset[6] (Table 2).

A more recent meta-analysis, including data from studies conducted between 1983 and 2011, confirmed that PSD occurs in nearly one third of cases and that this prevalence is independent of time-interval after stroke and study setting[100]. The incidence of PSD has been poorly investigated. Recent data from the South London Stroke Register showed an incidence of depression of 16% in the first year after stroke, of 7%-21% in the 15 years after stroke, and a cumulative incidence of 55%[101]. Moreover, the few studies comparing the incidence of depression in cohorts of stroke patients with that in appropriately matched non-stroke cohorts reported a doubled risk in the former group[102,103].

The natural course of PSD seems to be dynamic and dependent on the timing of onset. Longitudinal studies observed that most patients who have depression after stroke became depressed shortly after the acute event, with a greatest increase in the prevalence of PSD during the first months post-stroke despite the overall disability decreases over time[101,104-106]. Moreover, although a significant proportion of these patients recovered from depression, the occurrence of new cases made the overall prevalence of depression stable over time. About 15%-50% of patients with early onset PSD has been reported to recover in subsequent assessments within 1 year[100], and to have a higher probability of remission in comparison to patients with later onset depressive episodes[106]. Data from a rehabilitation-based study indicates that, at 1 year, 60% of the patients with early depression (0 to 3 mo) had recovered and that those who had not recovered at this follow-up time had a high risk of developing chronic depression[107]. In line with these observations, data from a longitudinal study with a 15-year follow-up showed that half of the patients who were depressed at 3 mo had recovered from depression at 1 year, while the other half recovered gradually between years 2 and 9 and that the proportion of recurrent cases rose from 38% at 2 years to 100% at 15 years[101]. Therefore, depression is often persistent after stroke, with high risk of relapse even after remission over a long period of time.

Stroke survivors have more than six-fold higher risk of developing clinically overt depression even two or more years after index stroke compared to age-matched controls[108]. This suggests that stroke survivors remain at elevated risk for clinically significant depressive symptoms for years after the incident stroke.

According to the results of the Depression in Stroke patients multicenter observational study group, early onset depression appears to be distinct from later onset depression (after 6$^{th}$ month) regarding not only time course but also clinical features. Actually, patients with early occurrence of PSD presented more severe symptoms of depression, assessed using the Montgomery-Asberg Depression Rating Scale, than those developing PSD later[106].

## WHAT IS THE PATHOGENESIS OF POSTSTROKE DEPRESSION?

The exact mechanism leading to depression after stroke is still incompletely understood (Figure 1).

In the mid-seventies, the hypothesis that PSD might depend on the anatomic location of brain lesions led to the view of this disorder as a clinical condition related to the interruption of specific pathways involved in mood regulation[109]. Subsequently, many reports suggested that left hemispheric lesions involving frontal region, basal ganglia and those the frontal pole are correlated with an increased risk of PSD[110-114]. The association between le-

**Table 2 Pooled prevalence of post stroke depression stratified by study setting and timing assessment in the meta-analysis of Hackett *et al*[6,147] and Ayerbe *et al*[100,101]**

| | Hackett *et al*[6,147] | Ayerbe *et al*[100,101] |
|---|---|---|
| Publication period of included studies | 1977-2002 | 1983-2011 |
| Number of included studies (population: *n*, hospital: *n*, rehabilitation: n) | 51 (population: 6, hospital: 16, rehabilitation: 29) | 43 (population: 6, hospital: 15, rehabilitation: 22) |
| Pooled prevalence | | |
| Overall | 33% (95%CI: 29%-36%) | 29% (95%CI: 25%-32%) |
| Study setting | | |
| Population-based | AP: 33% (95%CI: 29%-37%) MTP: 33% (95% CI: 0%-72%) LTP: 33% (95%CI: 29%-36%) | 22% (95%CI: 17%-28%) |
| Hospital-based | AP: 36% (95%CI: 0%-73%) MTP: 32% (95% CI: 23%-41%) LTP: 34% (95%CI: 24%-45%) | 30% (95%CI: 24%-36%) |
| Rehabilitation-based | AP: 30% (95%CI: 16%-44%) MTP: 36% (95% CI: 20%-39%) LTP: 34% (95%CI: 26%-42%) | 30% (95%CI: 25%-36%) |
| Timing of assessment | | |
| Acute phase (< 1 mo) | 32% (95%CI: 19%-44%) | 28% (95%CI: 23%-34%) |
| Medium-term phase (1-6 mo) | 34% (95%CI: 20%-39%) | 31% (95%CI: 24%-39%) |
| Long-term phase (6 mo or more)[1] (6 mo to 1 yr)[2] | 34% (95%CI: 29%-39%) | 33% (95%CI: 23%-43%) |
| Very long-term phase (> 1 yr) | n.d. | 25% (95%CI: 19%-32%) |

AP: Acute phase; MTP: Medium-term phase; LTP: Long-term phase; n.d.: Not determined.

sions involving these anatomic regions and depression was even stronger during the first months after an acute stroke[110,115]. Though interesting, these data have not been consistently replicated and some studies were also reported showing that depression might be associated with right-hemisphere lesions[106,115-118]. Moreover, the hypothesis that depression is influenced by the site of the cerebral lesion was not confirmed in a systematic review by Carson and coworkers. Thus, there is no definitive conclusion on the hypothesis of lateralization and risk for depression[119].

Methodological limitations have been considered to explain these inconsistencies. Boghal and coworkers suggested that the heterogeneity of the results regarding lesions located in the left hemisphere and PSD might depend on whether patients were sampled as inpatients or from the community[120]. Moreover, a significant association of PSD with lesions located in the left hemisphere was found in the first month after stroke, in the right hemisphere after 6 mo[120]. Thus, it is possible that acute PSD and late PSD might be due to different mechanisms. In the acute phase variations in biogenic amines and modulations of serotonin (5HT) receptor may be involved, while in chronic phase PSD may reflect a failure to adapt to changes secondary to stroke, such as impairment in daily activities[121,122].

Several studies have also explored the potential impact of small vascular lesions and chronic ischemic damage in triggering PSD. Based on the concept of vascular depression, chronic ischemic damage could predispose, precipitate or perpetuate depression in the elderly as a consequence of affecting frontal-subcortical circuits responsible for mood control[42]. White matter lesions are conceptualized as a marker of underlying cerebral vascular pathology and they have been described associated with late-onset depression, possibly affecting severity and outcome[123]. In line with this, Brodaty and coworkers found that PSD may be related to accumulation of vascular lesions rather than site and severity of single stroke, supporting the view that biological factors might be an important determinant of PSD[124]. Moreover, in a neuropathological studies of 41 consecutive autopsy cases of patients with stroke it was observed that the vascular burden depending on progressive accumulation of lacunar infarcts within the thalamus and basal ganglia, and of microvascular lesions in deep white matter might have a role in the prediction of PSD[125].

Further support to these findings comes from the growing evidence supporting an association between cerebral microbleeds (CMBs) and occurrence of PSD[126-129]. CMBs are common in ischemic stroke and considered as an indicator of underlying small vessel vasculopathy[130]. There is some evidence indicating that CMBs could affect not only the risk of PSD but also its severity[128]. Furthermore in a recent study, the possibility of non-remission of depression at 1 year follow-up was associated to the presence of lobar CMBs in patients with well-established cerebrovascular disease, suggesting also an influence on outcome[129]. Research on the biology of CBMs may provide useful information on the mechanisms of PSD.

Vascular risk factors might also influence the risk of PSD. In particular, hypertension was found to impact significantly the development of post-stroke depressive symptoms[131,132]. In line with previous findings, hypertension may be linked to depression following stroke through a development of small vessel vasculopathy.

Additionally, several lines of evidence have shown that stroke determines a perturbation of proinflammatory cytokines, which might influence the inflammatory responses implicated in the pathophysiology of depression[133,134], through a physiological dysfunction of brain structures involved in mood control, such as the limbic system[133,135].

However, feeble and often contrasting results have

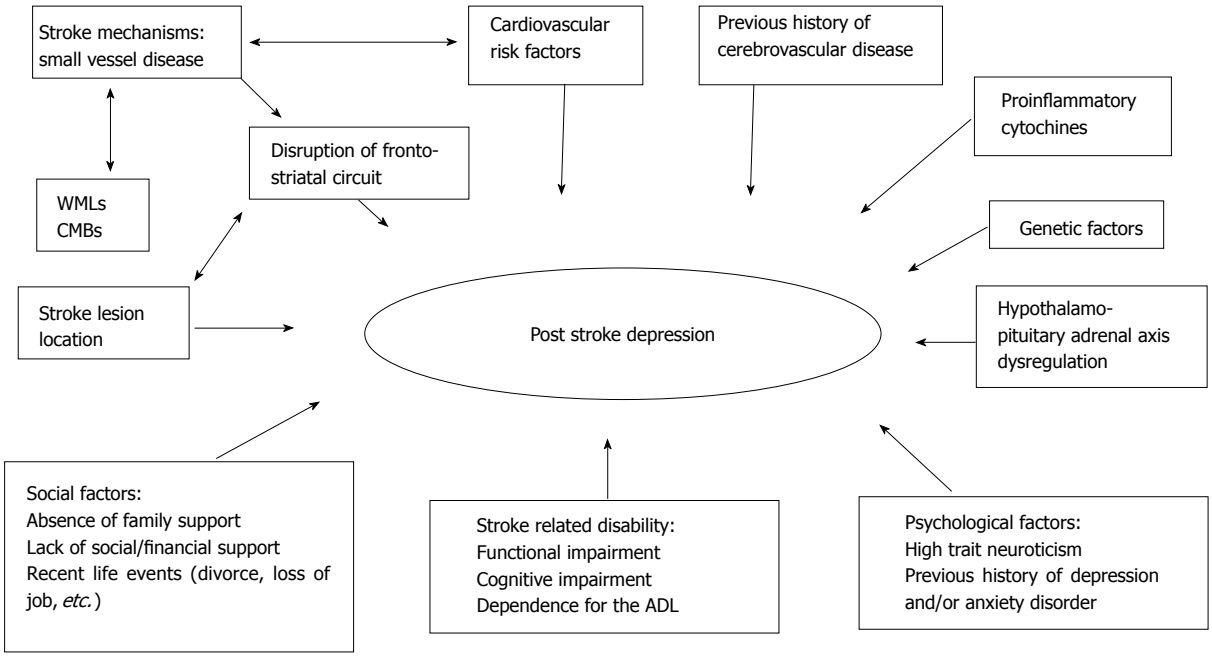



**Figure 1 Potential pathogenic pathways of post-stroke depression.** WMLs: White matter lesions; CMBs: Cerebral microbleeds; ADL: Activities of daily living.

been reported for the association of PSD with a variety of inflammatory mediators, in particular interleukin (IL)-1beta, IL-6, IL 18, tumor necrosis factor alpha or C-reactive protein[136-139]. A significant association between high serum leptin levels and PSD have also been found at 1-month after stroke[138].

Inconsistent results have also been reported for association between neurotrophic factors such as brain-derived neurotrophic factor (BDNF) and the development of PSD at the acute stage of ischemic stroke[138,140].

The contribution of genetic factors has been also investigated in PSD. A common genetic variant in the promoter region of the serotonin transporter (5-HTT) gene, the short variation length in the 5-HTT-linked polymorphic region (5-HTTLPR, s-allele) has been found to be significantly associated to PSD[141]. Patients carrying the s/s genotype have been also reported to have a better response to psychological intervention for PSD[142]. The val66met polymorphism of BDNF is another variant which may be implicated in PSD[143].

Other predictive factors for PSD, including female gender[144] and previous stroke[118], have been considered with inconsistent results.

Psychosocial factors have been reported to play a role in development of PSD. These include specific personality traits such as premorbid neuroticism[145], previous history of depression[118], living alone, and social isolation with lack of support[107,126]. Major recent life events seem to be a strong risk factor for PSD[118]. In these cases, the overwhelming psychological nature of stroke can trigger a depressive episode in predisposed individuals or in subjects with inadequate social relations. One year after stroke the persistence of few social contacts outside the immediate family contributes to depression[107]. Dependence in the activities of daily living is another important predictor of depression after the first three months[107,146]. In a systematic review of observational studies, Hackett and coworkers found that, despite a wide range of predictive factors, only physical disability, stroke severity, and cognitive impairment resulted consistently associated to PSD[147]. As observed by the authors, however, methodological heterogeneity and the limited number of studies on this topic do not allow firm conclusions on how to identify those patients at the greatest risk of developing depression following a stroke. Depression and physical and cognitive impairment in stroke patients may be associated by a bidirectional causal link. Stroke-related disability may trigger depression which, in turn, may reduce patients' compliance to rehabilitation treatments leading to unfavorable functional outcome[126,147-149].

## OUTCOME OF PSD

Depression can exert significant negative impacts on stroke recovery and impair outcome leading to a worsening of cognitive functions, motor abilities and quality of life. It also increases mortality. In a recent prospective population-based study, Ayerbe and coworkers recently reported that the occurrence of depression 3 mo after acute stroke was significantly associated with higher disability, anxiety and a lower quality of life up to 5 years after stroke[150]. Moreover, mortality rate during the 5 years following stroke was higher for patients depressed at 3 mo, and recovering from depression at 1 year did not improve prognosis. These patients, in fact, showed a higher mortality risk during the 5 years after stroke (HR = 1.69; 95%CI: 1.09-2.62) compared with those non-depressed[150]. Additionally, patients with acute PSD were 3.4 times more likely to die during a 10-year follow-up, compared to patients who were non-depressed after

acute stroke[151].

## CONCLUSION

Strong evidence supports the view that depression and vascular diseases are deeply related, especially in the elderly. The link between depression and stroke observed in epidemiological analyses appears to be bidirectional, being depression both a precursor and an important consequence of stroke. In spite of a wide literature, the mechanisms underlying this association have not been completely clarified. In this regard, it is necessary that future studies use common methodological approaches, based on accurate description and validated scales for depression. Understanding how depression leads to stroke would allow the development of targeted prevention strategies and interventions aimed at reducing depression-related morbidity and mortality. Identifying the subgroup of stroke patients at highest risk of depression should be the first step. Early identification and treatment of PSD may improve stroke rehabilitation outcomes and decrease mortality.

## REFERENCES


1 **American Psychiatric Association.** Diagnostic and statistical manual of mental disorders. Washington (DC): APA, 1994

2 **Kessler RC**, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS; National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 2003; **289**: 3095-3105 [PMID: 12813115]

3 **Mezuk B**, Eaton WW, Albrecht S, Golden SH. Depression and type 2 diabetes over the lifespan: a meta-analysis. *Diabetes Care* 2008; **31**: 2383-2390 [PMID: 19033418 DOI: 10.2337/dc08-0985]

4 **Patten SB**, Williams JV, Lavorato DH, Campbell NR, Eliasziw M, Campbell TS. Major depression as a risk factor for high blood pressure: epidemiologic evidence from a national longitudinal study. *Psychosom Med* 2009; **71**: 273-279 [PMID: 19196807 DOI: 10.1097/PSY.0b013e3181988e5f]

5 **Nicholson A**, Kuper H, Hemingway H. Depression as an aetiologic and prognostic factor in coronary heart disease: a meta-analysis of 6362 events among 146 538 participants in 54 observational studies. *Eur Heart J* 2006; **27**: 2763-2774 [PMID: 17082208]

6 **Hackett ML**, Yapa C, Parag V, Anderson CS. Frequency of depression after stroke: a systematic review of observational studies. *Stroke* 2005; **36**: 1330-1340 [PMID: 15879342]

7 **Carod-Artal J**, Egido JA, González JL, Varela de Seijas E. Quality of life among stroke survivors evaluated 1 year after stroke: experience of a stroke unit. *Stroke* 2000; **31**: 2995-3000 [PMID: 11108762]

8 **Whooley MA**, de Jonge P, Vittinghoff E, Otte C, Moos R, Carney RM, Ali S, Dowray S, Na B, Feldman MD, Schiller NB, Browner WS. Depressive symptoms, health behaviors, and risk of cardiovascular events in patients with coronary heart disease. *JAMA* 2008; **300**: 2379-2388 [PMID: 19033588 DOI: 10.1001/jama.2008.711]

9 **Evans DL**, Charney DS, Lewis L, Golden RN, Gorman JM, Krishnan KR, Nemeroff CB, Bremner JD, Carney RM, Coyne JC, Delong MR, Frasure-Smith N, Glassman AH, Gold PW, Grant I, Gwyther L, Ironson G, Johnson RL, Kanner AM, Katon WJ, Kaufmann PG, Keefe FJ, Ketter T, Laughren TP, Leserman J, Lyketsos CG, McDonald WM, McEwen BS, Miller AH, Musselman D, O'Connor C, Petitto JM, Pollock BG, Robinson RG, Roose SP, Rowland J, Sheline Y, Sheps DS, Simon G, Spiegel D, Stunkard A, Sunderland T, Tibbits P, Valvo WJ. Mood disorders in the medically ill: scientific review and recommendations. *Biol Psychiatry* 2005; **58**: 175-189 [PMID: 16084838]

10 **Vogt T**, Pope C, Mullooly J, Hollis J. Mental health status as a predictor of morbidity and mortality: a 15-year follow-up of members of a health maintenance organization. *Am J Public Health* 1994; **84**: 227-231 [PMID: 8296945]

11 **Wassertheil-Smoller S**, Applegate WB, Berge K, Chang CJ, Davis BR, Grimm R, Kostis J, Pressel S, Schron E; SHEP Cooperative Research Group (Systolic Hypertension in the elderly). Change in depression as a precursor of cardiovascular events. SHEP Cooperative Research Group (Systoloc Hypertension in the elderly). *Arch Intern Med* 1996; **156**: 553-561 [PMID: 8604962]

12 **Everson SA**, Roberts RE, Goldberg DE, Kaplan GA. Depressive symptoms and increased risk of stroke mortality over a 29-year period. *Arch Intern Med* 1998; **158**: 1133-1138 [PMID: 9605786]

13 **Whooley MA**, Browner WS; Study of Osteoporotic Fractures Research Group. Association between depressive symptoms and mortality in older women. Study of Osteoporotic Fractures Research Group. *Arch Intern Med* 1998; **158**: 2129-2135 [PMID: 9801180]

14 **Jonas BS**, Mussolino ME. Symptoms of depression as a prospective risk factor for stroke. *Psychosom Med* 2000; **62**: 463-471 [PMID: 10949089]

15 **Larson SL**, Owens PL, Ford D, Eaton W. Depressive disorder, dysthymia, and risk of stroke: thirteen-year follow-up from the Baltimore epidemiologic catchment area study. *Stroke* 2001; **32**: 1979-1983 [PMID: 11546884]

16 **Ohira T**, Iso H, Satoh S, Sankai T, Tanigawa T, Ogawa Y, Imano H, Sato S, Kitamura A, Shimamoto T. Prospective study of depressive symptoms and risk of stroke among japanese. *Stroke* 2001; **32**: 903-908 [PMID: 11283390]

17 **Ostir GV**, Markides KS, Peek MK, Goodwin JS. The association between emotional well-being and the incidence of stroke in older adults. *Psychosom Med* 2001; **63**: 210-215 [PMID: 11292267]

18 **May M**, McCarron P, Stansfeld S, Ben-Shlomo Y, Gallacher J, Yarnell J, Davey Smith G, Elwood P, Ebrahim S. Does psychological distress predict the risk of ischemic stroke and transient ischemic attack? The Caerphilly Study. *Stroke* 2002; **33**: 7-12 [PMID: 11779881]

19 **Wassertheil-Smoller S**, Shumaker S, Ockene J, Talavera GA, Greenland P, Cochrane B, Robbins J, Aragaki A, Dunbar-Jacob J. Depression and cardiovascular sequelae in postmenopausal women. The Women's Health Initiative (WHI). *Arch Intern Med* 2004; **164**: 289-298 [PMID: 14769624]

20 **Gump BB**, Matthews KA, Eberly LE, Chang YF; MRFIT Research Group. Depressive symptoms and mortality in men: results from the Multiple Risk Factor Intervention Trial. *Stroke* 2005; **36**: 98-102 [PMID: 15569872]

21 **Avendano M**, Kawachi I, Van Lenthe F, Boshuizen HC, Mackenbach JP, Van den Bos GA, Fay ME, Berkman LF. Socioeconomic status and stroke incidence in the US elderly: the role of risk factors in the EPESE study. *Stroke* 2006; **37**: 1368-1373 [PMID: 16690902]

22 **Stürmer T**, Hasselbach P, Amelang M. Personality, lifestyle, and risk of cardiovascular disease and cancer: follow-up of population based cohort. *BMJ* 2006; **332**: 1359 [PMID: 16687457]

23 **Kamphuis MH**, Kalmijn S, Tijhuis MA, Geerlings MI, Giampaoli S, Nissinen A, Grobbee DE, Kromhout D. Depressive symptoms as risk factor of cardiovascular mortality in older European men: the Finland, Italy and Netherlands Elderly

(FINE) study. *Eur J Cardiovasc Prev Rehabil* 2006; **13**: 199-206 [PMID: 16575273]
24 **Arbelaez JJ**, Ariyo AA, Crum RM, Fried LP, Ford DE. Depressive symptoms, inflammation, and ischemic stroke in older adults: a prospective analysis in the cardiovascular health study. *J Am Geriatr Soc* 2007; **55**: 1825-1830 [PMID: 17916124]
25 **Kawamura T**, Shioiri T, Takahashi K, Ozdemir V, Someya T. Survival rate and causes of mortality in the elderly with depression: a 15-year prospective study of a Japanese community sample, the Matsunoyama-Niigata suicide prevention project. *J Investig Med* 2007; **55**: 106-114 [PMID: 17481379]
26 **Salaycik KJ**, Kelly-Hayes M, Beiser A, Nguyen AH, Brady SM, Kase CS, Wolf PA. Depressive symptoms and risk of stroke: the Framingham Study. *Stroke* 2007; **38**: 16-21 [PMID: 17138952]
27 **Bos MJ**, Lindén T, Koudstaal PJ, Hofman A, Skoog I, Breteler MM, Tiemeier H. Depressive symptoms and risk of stroke: the Rotterdam Study. *J Neurol Neurosurg Psychiatry* 2008; **79**: 997-1001 [PMID: 18208858 DOI: 10.1136/jnnp.2007.134965]
28 **Lee HC**, Lin HC, Tsai SY. Severely depressed young patients have over five times increased risk for stroke: a 5-year follow-up study. *Biol Psychiatry* 2008; **64**: 912-915 [PMID: 18718571 DOI: 10.1016/j.biopsych.2008.07.006]
29 **Liebetrau M**, Steen B, Skoog I. Depression as a risk factor for the incidence of first-ever stroke in 85-year-olds. *Stroke* 2008; **39**: 1960-1965 [PMID: 18451342 DOI: 10.1161/STROKEAHA.107.490797]
30 **Surtees PG**, Wainwright NW, Luben RN, Wareham NJ, Bingham SA, Khaw KT. Psychological distress, major depressive disorder, and risk of stroke. *Neurology* 2008; **70**: 788-794 [PMID: 18316690 DOI: 10.1212/01.wnl.0000304109.18563.81]
31 **Glymour MM**, Maselko J, Gilman SE, Patton KK, Avendaño M. Depressive symptoms predict incident stroke independently of memory impairments. *Neurology* 2010; **75**: 2063-2070 [PMID: 21135381 DOI: 10.1212/WNL.0b013e318200d70e]
32 **Nabi H**, Kivimäki M, Suominen S, Koskenvuo M, Singh-Manoux A, Vahtera J. Does depression predict coronary heart disease and cerebrovascular disease equally well? The Health and Social Support Prospective Cohort Study. *Int J Epidemiol* 2010; **39**: 1016-1024 [PMID: 20360321 DOI: 10.1093/ije/dyq050]
33 **Peters R**, Pinto E, Beckett N, Swift C, Potter J, McCormack T, Nunes M, Grimley-Evans J, Fletcher A, Bulpitt C. Association of depression with subsequent mortality, cardiovascular morbidity and incident dementia in people aged 80 and over and suffering from hypertension. Data from the Hypertension in the Very Elderly Trial (HYVET). *Age Ageing* 2010; **39**: 439-445 [PMID: 20497949 DOI: 10.1093/ageing/afq042]
34 **Pan A**, Okereke OI, Sun Q, Logroscino G, Manson JE, Willett WC, Ascherio A, Hu FB, Rexrode KM. Depression and incident stroke in women. *Stroke* 2011; **42**: 2770-2775 [PMID: 21836097 DOI: 10.1161/STROKEAHA.111.617043]
35 **Seifert CL**, Poppert H, Sander D, Feurer R, Etgen T, Ander KH, Pürner K, Brönner M, Sepp D, Kehl V, Förstl H, Bickel H. Depressive symptoms and the risk of ischemic stroke in the elderly--influence of age and sex. *PLoS One* 2012; **7**: e50803 [PMID: 23226388 DOI: 10.1371/journal.pone.0050803]
36 **Majed B**, Arveiler D, Bingham A, Ferrieres J, Ruidavets JB, Montaye M, Appleton K, Haas B, Kee F, Amouyel P, Ducimetiere P, Empana JP; PRIME Study Group. Depressive symptoms, a time-dependent risk factor for coronary heart disease and stroke in middle-aged men: the PRIME Study. *Stroke* 2012; **43**: 1761-1767 [PMID: 22517599 DOI: 10.1161/STROKEAHA.111.645366]
37 **Jackson CA**, Mishra GD. Depression and risk of stroke in midaged women: a prospective longitudinal study. *Stroke* 2013; **44**: 1555-1560 [PMID: 23686976 DOI: 10.1161/STROKEAHA.113.001147]
38 **Van der Kooy K**, van Hout H, Marwijk H, Marten H, Stehouwer C, Beekman A. Depression and the risk for cardiovascular diseases: systematic review and meta analysis. *Int J Geriatr Psychiatry* 2007; **22**: 613-626 [PMID: 17236251]
39 **Pan A**, Sun Q, Okereke OI, Rexrode KM, Hu FB. Depression and risk of stroke morbidity and mortality: a meta-analysis and systematic review. *JAMA* 2011; **306**: 1241-1249 [PMID: 21934057 DOI: 10.1001/jama.2011.1282]
40 **Dong JY**, Zhang YH, Tong J, Qin LQ. Depression and risk of stroke: a meta-analysis of prospective studies. *Stroke* 2012; **43**: 32-37 [PMID: 22020036 DOI: 10.1161/STROKEAHA.111.630871]
41 **O'Donnell MJ**, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, Rangarajan S, Islam S, Pais P, McQueen MJ, Mondo C, Damasceno A, Lopez-Jaramillo P, Hankey GJ, Dans AL, Yusoff K, Truelsen T, Diener HC, Sacco RL, Ryglewicz D, Czlonkowska A, Weimar C, Wang X, Yusuf S. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* 2010; **376**: 112-123 [PMID: 20561675 DOI: 10.1016/S0140-6736(10)60834-3]
42 **Alexopoulos GS**, Meyers BS, Young RC, Campbell S, Silbersweig D, Charlson M. 'Vascular depression' hypothesis. *Arch Gen Psychiatry* 1997; **54**: 915-922 [PMID: 9337771]
43 **Newson RS**, Hek K, Luijendijk HJ, Hofman A, Witteman JC, Tiemeier H. Atherosclerosis and incident depression in late life. *Arch Gen Psychiatry* 2010; **67**: 1144-1151 [PMID: 21041615 DOI: 10.1001/archgenpsychiatry.2010.142]
44 **Fiedorowicz JG**, He J, Merikangas KR. The association between mood and anxiety disorders with vascular diseases and risk factors in a nationally representative sample. *J Psychosom Res* 2011; **70**: 145-154 [PMID: 21262417 DOI: 10.1016/j.jpsychores.2010.07.010]
45 **Wu HC**, Chou FH, Tsai KY, Su CY, Shen SP, Chung TC. The incidence and relative risk of stroke among patients with bipolar disorder: a seven-year follow-up study. *PLoS One* 2013; **8**: e73037 [PMID: 24023667 DOI: 10.1371/journal.pone.0073037]
46 **Musselman DL**, Evans DL, Nemeroff CB. The relationship of depression to cardiovascular disease: epidemiology, biology, and treatment. *Arch Gen Psychiatry* 1998; **55**: 580-592 [PMID: 9672048]
47 **Pariante CM**, Lightman SL. The HPA axis in major depression: classical theories and new developments. *Trends Neurosci* 2008; **31**: 464-468 [PMID: 18675469 DOI: 10.1016/j.tins.2008.06.006]
48 **Carney RM**, Freedland KE, Veith RC. Depression, the autonomic nervous system, and coronary heart disease. *Psychosom Med* 2005; **67** Suppl 1: S29-S33 [PMID: 15953797]
49 **Lange HW**, Herrmann-Lingen C. Depressive symptoms predict recurrence of atrial fibrillation after cardioversion. *J Psychosom Res* 2007; **63**: 509-513 [PMID: 17980224]
50 **Shimbo D**, Chaplin W, Crossman D, Haas D, Davidson KW. Role of depression and inflammation in incident coronary heart disease events. *Am J Cardiol* 2005; **96**: 1016-1021 [PMID: 16188535]
51 **Vaccarino V**, Johnson BD, Sheps DS, Reis SE, Kelsey SF, Bittner V, Rutledge T, Shaw LJ, Sopko G, Bairey Merz CN; National Heart, Lung, and Blood Institute. Depression, inflammation, and incident cardiovascular disease in women with suspected coronary ischemia: the National Heart, Lung, and Blood Institute-sponsored WISE study. *J Am Coll Cardiol* 2007; **50**: 2044-2050 [PMID: 18021871]
52 **Howren MB**, Lamkin DM, Suls J. Associations of depression with C-reactive protein, IL-1, and IL-6: a meta-analysis. *Psychosom Med* 2009; **71**: 171-186 [PMID: 19188531 DOI: 10.1097/PSY.0b013e3181907c1b]
53 **Kendler KS**, Gatz M, Gardner CO, Pedersen NL. A Swedish national twin study of lifetime major depression. *Am J Psychiatry* 2006; **163**: 109-114 [PMID: 16390897]

54 **Kop WJ**, Gottdiener JS, Tangen CM, Fried LP, McBurnie MA, Walston J, Newman A, Hirsch C, Tracy RP. Inflammation and coagulation factors in persons & gt; 65 years of age with symptoms of depression but without evidence of myocardial ischemia. *Am J Cardiol* 2002; **89**: 419-424 [PMID: 11835923]

55 **Schins A**, Hamulyák K, Scharpé S, Lousberg R, Van Melle J, Crijns H, Honig A. Whole blood serotonin and platelet activation in depressed post-myocardial infarction patients. *Life Sci* 2004; **76**: 637-650 [PMID: 15567189]

56 **Verdelho A**, Ferro JM. Late onset depressive symptoms can be a marker of cerebral vascular pathology. *J Neurol Neurosurg Psychiatry* 2008; **79**: 977 [PMID: 18708565 DOI: 10.1136/jnnp.2007.142364]

57 **Taylor WD**, MacFall JR, Payne ME, McQuoid DR, Steffens DC, Provenzale JM, Krishnan RR. Greater MRI lesion volumes in elderly depressed subjects than in control subjects. *Psychiatry Res* 2005; **139**: 1-7 [PMID: 15927454]

58 **Disabato BM**, Sheline YI. Biological basis of late life depression. *Curr Psychiatry Rep* 2012; **14**: 273-279 [PMID: 22562412 DOI: 10.1007/s11920-012-0279-6]

59 **DiMatteo MR**, Lepper HS, Croghan TW. Depression is a risk factor for noncompliance with medical treatment: meta-analysis of the effects of anxiety and depression on patient adherence. *Arch Intern Med* 2000; **160**: 2101-2107 [PMID: 10904452]

60 **Olfson M**, Marcus SC. National patterns in antidepressant medication treatment. *Arch Gen Psychiatry* 2009; **66**: 848-856 [PMID: 19652124 DOI: 10.1001/archgenpsychiatry.2009.81]

61 **Reid S**, Barbui C. Long term treatment of depression with selective serotonin reuptake inhibitors and newer antidepressants. *BMJ* 2010; **340**: c1468 [PMID: 20348175 DOI: 10.1136/bmj.c1468]

62 **Whang W**, Kubzansky LD, Kawachi I, Rexrode KM, Kroenke CH, Glynn RJ, Garan H, Albert CM. Depression and risk of sudden cardiac death and coronary heart disease in women: results from the Nurses' Health Study. *J Am Coll Cardiol* 2009; **53**: 950-958 [PMID: 19281925 DOI: 10.1016/j.jacc.2008.10.060.]

63 **Smoller JW**, Allison M, Cochrane BB, Curb JD, Perlis RH, Robinson JG, Rosal MC, Wenger NK, Wassertheil-Smoller S. Antidepressant use and risk of incident cardiovascular morbidity and mortality among postmenopausal women in the Women's Health Initiative study. *Arch Intern Med* 2009; **169**: 2128-2139 [PMID: 20008698 DOI: 10.1001/archinternmed.2009.436]

64 **Tata LJ**, West J, Smith C, Farrington P, Card T, Smeeth L, Hubbard R. General population based study of the impact of tricyclic and selective serotonin reuptake inhibitor antidepressants on the risk of acute myocardial infarction. *Heart* 2005; **91**: 465-471 [PMID: 15772201]

65 **Sauer WH**, Berlin JA, Kimmel SE. Effect of antidepressants and their relative affinity for the serotonin transporter on the risk of myocardial infarction. *Circulation* 2003; **108**: 32–36 [PMID: 12821544]

66 **Meijer WE**, Heerdink ER, Nolen WA, Herings RM, Leufkens HG, Egberts AC. Association of risk of abnormal bleeding with degree of serotonin reuptake inhibition by antidepressants. *Arch Intern Med* 2004; **164**: 2367-2370 [PMID: 15557417]

67 **Dawood T**, Lambert EA, Barton DA, Laude D, Elghozi JL, Esler MD, Haikerwal D, Kaye DM, Hotchkin EJ, Lambert GW. Specific serotonin reuptake inhibition in major depressive disorder adversely affects novel markers of cardiac risk. *Hypertens Res* 2007; **30**: 285-293 [PMID: 17541206]

68 **Kivimäki M**, Hamer M, Batty GD, Geddes JR, Tabak AG, Pentti J, Virtanen M, Vahtera J. Antidepressant medication use, weight gain, and risk of type 2 diabetes: a population-based study. *Diabetes Care* 2010; **33**: 2611-2616 [PMID: 20823343 DOI: 10.2337/dc10-1187]

69 **Roose SP**. Treatment of depression in patients with heart disease. *Biol Psychiatry* 2003; **54**: 262-268 [PMID: 12893102]

70 **Licht CM**, de Geus EJ, Seldenrijk A, van Hout HP, Zitman FG, van Dyck R, Penninx BW. Depression is associated with decreased blood pressure, but antidepressant use increases the risk for hypertension. *Hypertension* 2009; **53**: 631-638 [PMID: 19237679 DOI: 10.1161/HYPERTENSIONAHA.108.126698]

71 **Chen Y**, Guo JJ, Li H, Wulsin L, Patel NC. Risk of cerebrovascular events associated with antidepressant use in patients with depression: a population-based, nested case-control study. *Ann Pharmacother* 2008; **42**: 177-184 [PMID: 18212255 DOI: 10.1345/aph.1K369]

72 **Wu CS**, Wang SC, Cheng YC, Gau SS. Association of cerebrovascular events with antidepressant use: a case-crossover study. *Am J Psychiatry* 2011; **168**: 511-521 [PMID: 21406464 DOI: 10.1176/appi.ajp.2010.10071064]

73 **Swenson JR**, Doucette S, Fergusson D. Adverse cardiovascular events in antidepressant trials involving high-risk patients: a systematic review of randomized trials. *Can J Psychiatry* 2006; **51**: 923-929 [PMID: 17249635]

74 **Bak S**, Tsiropoulos I, Kjaersgaard JO, Andersen M, Mellerup E, Hallas J, García Rodríguez LA, Christensen K, Gaist D. Selective serotonin reuptake inhibitors and the risk of stroke: a population-based case-control study. *Stroke* 2002; **33**: 1465-1473 [PMID: 12052976]

75 **Kharofa J**, Sekar P, Haverbusch M, Moomaw C, Flaherty M, Kissela B, Broderick J, Woo D. Selective serotonin reuptake inhibitors and risk of hemorrhagic stroke. *Stroke* 2007; **38**: 3049-3051 [PMID: 17901378]

76 **Bushnell C**. Depression and the risk of stroke in women: an identification and treatment paradox. *Stroke* 2011; **42**: 2718-2719 [PMID: 21921282 DOI: 10.1161/STROKEAHA.111.626895]

77 **Rahman I**, Humphreys K, Bennet AM, Ingelsson E, Pedersen NL, Magnusson PK. Clinical depression, antidepressant use and risk of future cardiovascular disease. *Eur J Epidemiol* 2013; **28**: 589-595 [PMID: 23836399 DOI: 10.1007/s10654-013-9821-z]

78 **Singhal AB**, Caviness VS, Begleiter AF, Mark EJ, Rordorf G, Koroshetz WJ. Cerebral vasoconstriction and stroke after use of serotonergic drugs. *Neurology* 2002; **58**: 130-133 [PMID: 11781419]

79 **Ramasubbu R**. Cerebrovascular effects of selective serotonin reuptake inhibitors: a systematic review. *J Clin Psychiatry* 2004; **65**: 1642-1653 [PMID: 15641869]

80 **Noskin O**, Jafarimojarrad E, Libman RB, Nelson JL. Diffuse cerebral vasoconstriction (Call-Fleming syndrome) and stroke associated with antidepressants. *Neurology* 2006; **67**: 159-160 [PMID: 16832100]

81 **Li N**, Wallén NH, Ladjevardi M, Hjemdahl P. Effects of serotonin on platelet activation in whole blood. *Blood Coagul Fibrinolysis* 1997; **8**: 517-523 [PMID: 9491270]

82 **Hackam DG**, Mrkobrada M. Selective serotonin reuptake inhibitors and brain hemorrhage: a meta-analysis. *Neurology* 2012; **79**: 1862-1865 [PMID: 23077009 DOI: 10.1212/WNL.0b013e318271f848]

83 **McGrath ER**, O'Donnell MJ. Estimating treatment effects in observational studies. *Neurology* 2012; **79**: 1844-1845 [PMID: 23077011 DOI: 10.1212/WNL.0b013e318271f8b6]

84 **Chemerinski E**, Robinson RG. The neuropsychiatry of stroke. *Psychosomatics* 2000; **41**: 5-14 [PMID: 10665263]

85 **Ramasubbu R**, Robinson RG, Flint AJ, Kosier T, Price TR. Functional impairment associated with acute poststroke depression: the Stroke Data Bank Study. *J Neuropsychiatry Clin Neurosci* 1998; **10**: 26-33 [PMID: 9547463]

86 **Singh A**, Black SE, Herrmann N, Leibovitch FS, Ebert PL, Lawrence J, Szalai JP. Functional and neuroanatomic correlations in poststroke depression: the Sunnybrook Stroke Study. *Stroke* 2000; **31**: 637-644 [PMID: 10700497]

87 **Paolucci S**, Antonucci G, Grasso MG, Morelli D, Troisi

E, Coiro P, De Angelis D, Rizzi F, Bragoni M. Post-stroke depression, antidepressant treatment and rehabilitation results. A case-control study. *Cerebrovasc Dis* 2001; **12**: 264-271 [PMID: 11641594]
88 **Naess H**, Lunde L, Brogger J, Waje-Andreassen U. Depression predicts unfavourable functional outcome and higher mortality in stroke patients: the Bergen Stroke Study. *Acta Neurol Scand Suppl* 2010; **190:** 34-38 [PMID: 20586733 DOI: 10.1111/j.1600-0404.2010.01373.x]
89 **House A**, Knapp P, Bamford J, Vail A. Mortality at 12 and 24 months after stroke may be associated with depressive symptoms at 1 month. *Stroke* 2001; **32**: 696-701 [PMID: 11239189]
90 **Meader N**, Moe-Byrne T, Llewellyn A, Mitchell AJ. Screening for poststroke major depression: a meta-analysis of diagnostic validity studies. *J Neurol Neurosurg Psychiatry* 2014; **85**: 198-206 [PMID: 23385849 DOI: 10.1136/jnnp-2012-304194]
91 **Fedoroff JP**, Starkstein SE, Parikh RM, Price TR, Robinson RG. Are depressive symptoms nonspecific in patients with acute stroke? *Am J Psychiatry* 1991; **148**: 1172-1176 [PMID: 1882994]
92 **Williams CL**, Rittman MR, Boylstein C, Faircloth C, Haijing Q. Qualitative and quantitative measurement of depression in veterans recovering from stroke. *J Rehabil Res Dev* 2005; **42**: 277-290 [PMID: 16187241]
93 **Schubert DS**, Burns R, Paras W, Sioson E. Increase of medical hospital length of stay by depression in stroke and amputation patients: a pilot study. *Psychother Psychosom* 1992; **57**: 61-66 [PMID: 1584900]
94 **Lipsey JR**, Spencer WC, Rabins PV, Robinson RG. Phenomenological comparison of poststroke depression and functional depression. *Am J Psychiatry* 1986; **143**: 527-529 [PMID: 3953895]
95 **Paradiso S**, Ohkubo T, Robinson RG. Vegetative and psychological symptoms associated with depressed mood over the first two years after stroke. *Int J Psychiatry Med* 1997; **27**: 137-157 [PMID: 9565720]
96 **Spalletta G**, Ripa A, Caltagirone C. Symptom profile of DSM-IV major and minor depressive disorders in first-ever stroke patients. *Am J Geriatr Psychiatry* 2005; **13**: 108-115 [PMID: 15703319]
97 **Gainotti G**, Azzoni A, Razzano C, Lanzillotta M, Marra C, Gasparini F. The Post-Stroke Depression Rating Scale: a test specifically devised to investigate affective disorders of stroke patients. *J Clin Exp Neuropsychol* 1997; **19**: 340-356 [PMID: 9268809]
98 **Gainotti G**, Azzoni A, Marra C. Frequency, phenomenology and anatomical-clinical correlates of major post-stroke depression. *Br J Psychiatry* 1999; **175**: 163-167 [PMID: 10627800]
99 **Cumming TB**, Churilov L, Skoog I, Blomstrand C, Linden T. Little evidence for different phenomenology in poststroke depression. *Acta Psychiatr Scand* 2010; **121**: 424-430 [PMID: 20384602 DOI: 10.1111/j.1600-0447.2010.01558.x]
100 **Ayerbe L**, Ayis S, Wolfe CD, Rudd AG. Natural history, predictors and outcomes of depression after stroke: systematic review and meta-analysis. *Br J Psychiatry* 2013; **202**: 14-21 [PMID: 23284148 DOI: 10.1192/bjp.bp.111.107664ù]
101 **Ayerbe L**, Ayis S, Crichton S, Wolfe CD, Rudd AG. The natural history of depression up to 15 years after stroke: the South London Stroke Register. *Stroke* 2013; **44**: 1105-1110 [PMID: 23404719 DOI: 10.1161/STROKEAHA.111.679340]
102 House A, Dennis M, Mogridge L, Warlow C, Hawton K, Jones L. Mood disorders in the year after first stroke. *Br J Psychiatry* 1991; **158:** 83–92 [PMID: 2015456]
103 **Kase CS**, Wolf PA, Kelly-Hayes M, Kannel WB, Beiser A, D' Agostino RB. Intellectual decline after stroke: the Framingham Study. *Stroke* 1998; **29**: 805-812 [PMID: 9550515]
104 **Townend BS**, Whyte S, Desborough T, Crimmins D, Markus R, Levi C, Sturm JW. Longitudinal prevalence and determinants of early mood disorder post-stroke. *J Clin Neurosci* 2007; **14**: 429-434 [PMID: 17336529]
105 **Aben I**, Verhey F, Strik J, Lousberg R, Lodder J, Honig A. A comparative study into the one year cumulative incidence of depression after stroke and myocardial infarction. *J Neurol Neurosurg Psychiatry* 2003; **74**: 581-585 [PMID: 12700297]
106 **Paolucci S**, Gandolfo C, Provinciali L, Torta R, Toso V; DESTRO Study Group. The Italian multicenter observational study on post-stroke depression (DESTRO). *J Neurol* 2006; **253**: 556-562 [PMID: 16767539]
107 **Aström M**, Adolfsson R, Asplund K. Major depression in stroke patients. A 3-year longitudinal study. *Stroke* 1993; **24**: 976-982 [PMID: 8322398]
108 **Whyte EM**, Mulsant BH, Vanderbilt J, Dodge HH, Ganguli M. Depression after stroke: a prospective epidemiological study. *J Am Geriatr Soc* 2004; **52**: 774-778 [PMID: 15086660]
109 **Carson AJ**. Impact commentaries. Mood disorder as a specific complication of stroke. *J Neurol Neurosurg Psychiatry* 2012; **83**: 859 [PMID: 22291218 DOI: 10.1136/jnnp-2011-301854]
110 **Robinson RG**, Kubos KL, Starr LB, Rao K, Price TR. Mood disorders in stroke patients. Importance of location of lesion. *Brain* 1984; **107** (Pt 1): 81-93 [PMID: 6697163]
111 **Starkstein SE**, Robinson RG, Berthier ML, Parikh RM, Price TR. Differential mood changes following basal ganglia vs thalamic lesions. *Arch Neurol* 1988; **45**: 725-730 [PMID: 3390026]
112 **Starkstein SE**, Robinson RG, Price TR. Comparison of cortical and subcortical lesions in the production of poststroke mood disorders. *Brain* 1987; **110** (Pt 4): 1045-1059 [PMID: 3651794]
113 **Robinson RG**, Kubos KL, Starr LB, Rao K, Price TR. Mood changes in stroke patients: relationship to lesion location. *Compr Psychiatry* 1983; **24**: 555-566 [PMID: 6653097]
114 **Aström M**, Olsson T, Asplund K. Different linkage of depression to hypercortisolism early versus late after stroke. A 3-year longitudinal study. *Stroke* 1993; **24**: 52-57 [PMID: 8418550]
115 **Iacoboni M**, Padovani A, Di Piero V, Lenzi GL. Post-stroke depression: relationships with morphological damage and cognition over time. *Ital J Neurol Sci* 1995; **16**: 209-216 [PMID: 7591672]
116 **MacHale SM**, O'Rourke SJ, Wardlaw JM, Dennis MS. Depression and its relation to lesion location after stroke. *J Neurol Neurosurg Psychiatry* 1998; **64**: 371-374 [PMID: 9527152]
117 **House A**, Dennis M, Warlow C, Hawton K, Molyneux A. Mood disorders after stroke and their relation to lesion location. A CT scan study. *Brain* 1990; **113** (Pt 4): 1113-1129 [PMID: 2397385]
118 **Andersen G**, Vestergaard K, Ingemann-Nielsen M, Lauritzen L. Risk factors for post-stroke depression. *Acta Psychiatr Scand* 1995; **92**: 193-198 [PMID: 7484197]
119 **Carson AJ**, MacHale S, Allen K, Lawrie SM, Dennis M, House A, Sharpe M. Depression after stroke and lesion location: a systematic review. *Lancet* 2000; **356**: 122-126 [PMID: 10963248]
120 **Bhogal SK**, Teasell R, Foley N, Speechley M. Lesion location and poststroke depression: systematic review of the methodological limitations in the literature. *Stroke* 2004; **35**: 794-802 [PMID: 14963278]
121 **Robinson RG**, Bolduc PL, Price TR. Two-year longitudinal study of poststroke mood disorders: diagnosis and outcome at one and two years. *Stroke* 1987; **18**: 837-843 [PMID: 3629640]
122 **Shimoda K**, Robinson RG. The relationship between poststroke depression and lesion location in long-term follow-up. *Biol Psychiatry* 1999; **45**: 187-192 [PMID: 9951566]
123 **Herrmann LL**, Le Masurier M, Ebmeier KP. White matter hyperintensities in late life depression: a systematic review. *J Neurol Neurosurg Psychiatry* 2008; **79**: 619-624 [PMID: 17717021]

124 **Brodaty H**, Withall A, Altendorf A, Sachdev PS. Rates of depression at 3 and 15 months poststroke and their relationship with cognitive decline: the Sydney Stroke Study. *Am J Geriatr Psychiatry* 2007; **15**: 477-486 [PMID: 17545448]

125 **Santos M**, Gold G, Kövari E, Herrmann FR, Bozikas VP, Bouras C, Giannakopoulos P. Differential impact of lacunes and microvascular lesions on poststroke depression. *Stroke* 2009; **40**: 3557-3562 [PMID: 19696424 DOI: 10.1161/STROKEAHA.109.548545]

126 **Choi-Kwon S**, Han K, Choi S, Suh M, Kim YJ, Song H, Cho KH, Nah HW, Kwon SU, Kang DW, Kim JS. Poststroke depression and emotional incontinence: factors related to acute and subacute stages. *Neurology* 2012; **78**: 1130-1137 [PMID: 22459674 DOI: 10.1212/WNL.0b013e31824f8090]

127 **Tang WK**, Chen YK, Lu JY, Chu WC, Mok VC, Ungvari GS, Wong KS. Cerebral microbleeds and depression in lacunar stroke. *Stroke* 2011; **42**: 2443-2446 [PMID: 21757672 DOI: 10.1161/STROKEAHA.111.614586]

128 **Tang WK**, Chen YK, Lu JY, Chu WC, Mok VC, Ungvari GS, Wong KS. Cerebral microbleeds and symptom severity of post-stroke depression: a magnetic resonance imaging study. *J Affect Disord* 2011; **129**: 354-358 [PMID: 20817306 DOI: 10.1016/j.jad.2010.08.007]

129 **Tang WK**, Chen Y, Liang H, Chu WC, Mok VC, Ungvari GS, Wong KS. Cerebral microbleeds as a predictor of 1-year outcome of poststroke depression. *Stroke* 2014; **45**: 77-81 [PMID: 24178917 DOI: 10.1161/STROKEAHA.113.002686]

130 **Werring DJ**, Coward LJ, Losseff NA, Jäger HR, Brown MM. Cerebral microbleeds are common in ischemic stroke but rare in TIA. *Neurology* 2005; **65**: 1914-1918 [PMID: 16380612]

131 **Tennen G**, Herrmann N, Black SE, Levy KS, Cappell J, Li A, Lanctôt KL. Are vascular risk factors associated with post-stroke depressive symptoms? *J Geriatr Psychiatry Neurol* 2011; **24**: 215-221 [PMID: 22228828 DOI: 10.1177/0891988711422526]

132 **Chatterjee K**, Fall S, Barer D. Mood after stroke: a case control study of biochemical, neuro-imaging and socio-economic risk factors for major depression in stroke survivors. *BMC Neurol* 2010; **10**: 125 [PMID: 21192808 DOI: 10.1186/1471-2377-10-125]

133 **Pascoe MC**, Crewther SG, Carey LM, Crewther DP. Inflammation and depression: why poststroke depression may be the norm and not the exception. *Int J Stroke* 2011; **6**: 128-135 [PMID: 21371275 DOI: 10.1111/j.1747-4949.2010.00565.x]

134 **Raison CL**, Miller AH. Is depression an inflammatory disorder? *Curr Psychiatry Rep* 2011; **13**: 467-475 [PMID: 21927805 DOI: 10.1007/s11920-011-0232-0]

135 **Spalletta G**, Bossù P, Ciaramella A, Bria P, Caltagirone C, Robinson RG. The etiology of poststroke depression: a review of the literature and a new hypothesis involving inflammatory cytokines. *Mol Psychiatry* 2006; **11**: 984-991 [PMID: 16894392]

136 **Ormstad H,** Aass HC, Amthor KF, Lund-Sørensen N, Sandvik L. Serum levels of cytokines, glucose, and hemoglobin as possible predictors of poststroke depression, and association with poststroke fatigue. *Int J Neurosci* 2012; **122:** 682-690 [DOI: 10.3109/00207454.2012.709892]

137 **Yang L,** Zhang Z, Sun D, Xu Z, Zhang X, Li L. The serum interleukin-18 is a potential marker for development of post-stroke depression. *Neurol Res* 2010; **32:** 340-346 [DOI: 10.1179/016164110X12656393665080]

138 **Jiménez I,** Sobrino T, Rodríguez-Yáñez M, Pouso M, Cristobo I, Sabucedo M, Blanco M, Castellanos M, Leira R, Castillo J. High serum levels of leptin are associated with post-stroke depression. *Psychol Med* 2009; **39**: 1201-1209 [DOI: 10.1017/S0033291709005637]

139 **Noonan K,** Crewther SG, Carey LM, Pascoe MC, Linden T. Sustained inflammation 1.5 years post-stroke is not associated with depression in elderly stroke survivors. *Clin Interv Aging* 2013; **8:** 69-74 [DOI: 10.2147/CIA.S38547]

140 **Yang L,** Zhang Z, Sun D, Xu Z, Yuan Y, Zhang X, Li L. Low serum BDNF may indicate the development of PSD in patients with acute ischemic stroke. *Int J Geriatr Psychiatry* 2011; **26:** 495-502 [DOI: 10.1002/gps.2552]

141 **Mak KK**, Kong WY, Mak A, Sharma VK, Ho RC. Polymorphisms of the serotonin transporter gene and post-stroke depression: a meta-analysis. *J Neurol Neurosurg Psychiatry* 2013; **84**: 322-328 [PMID: 23236014 DOI: 10.1136/jnnp-2012-303791]

142 **Kohen R**, Cain KC, Buzaitis A, Johnson V, Becker KJ, Teri L, Tirschwell DL, Veith RC, Mitchell PH. Response to psychosocial treatment in poststroke depression is associated with serotonin transporter polymorphisms. *Stroke* 2011; **42**: 2068-2070 [PMID: 21847802]

143 **Kim JM**, Stewart R, Kim SW, Yang SJ, Shin IS, Kim YH, Yoon JS. BDNF genotype potentially modifying the association between incident stroke and depression. *Neurobiol Aging* 2008; **29**: 789-792 [PMID: 17222482]

144 **Poynter B**, Shuman M, Diaz-Granados N, Kapral M, Grace SL, Stewart DE. Sex differences in the prevalence of post-stroke depression: a systematic review. *Psychosomatics* 2009; **50**: 563-569 [PMID: 19996226 DOI: 10.1176/appi.psy.50.6.563]

145 **Aben I**, Denollet J, Lousberg R, Verhey F, Wojciechowski F, Honig A. Personality and vulnerability to depression in stroke patients: a 1-year prospective follow-up study. *Stroke* 2002; **33**: 2391-2395 [PMID: 12364726]

146 **De Ryck A**, Brouns R, Fransen E, Geurden M, Van Gestel G, Wilssens I, De Ceulaer L, Mariën P, De Deyn PP, Engelborghs S. A prospective study on the prevalence and risk factors of poststroke depression. *Cerebrovasc Dis Extra* 2013; **3**: 1-13 [PMID: 23626594 DOI: 10.1159/000345557]

147 **Hackett ML**, Anderson CS. Predictors of depression after stroke: a systematic review of observational studies. *Stroke* 2005; **36**: 2296-2301 [PMID: 16179565]

148 **Parikh RM**, Robinson RG, Lipsey JR, Starkstein SE, Fedoroff JP, Price TR. The impact of poststroke depression on recovery in activities of daily living over a 2-year follow-up. *Arch Neurol* 1990; **47**: 785-789 [PMID: 2357159]

149 **Paolucci S**, Antonucci G, Pratesi L, Traballesi M, Grasso MG, Lubich S. Poststroke depression and its role in rehabilitation of inpatients. *Arch Phys Med Rehabil* 1999; **80**: 985-990 [PMID: 10488996]

150 **Ayerbe L**, Ayis S, Crichton S, Wolfe CD, Rudd AG. The long-term outcomes of depression up to 10 years after stroke; the South London Stroke Register. *J Neurol Neurosurg Psychiatry* 2014; **85**: 514-521 [PMID: 24163430 DOI: 10.1136/jnnp-2013-306448]

151 **Morris PL**, Robinson RG, Andrzejewski P, Samuels J, Price TR. Association of depression with 10-year poststroke mortality. *Am J Psychiatry* 1993; **150**: 124-129 [PMID: 8417554]


**P- Reviewer:** Sobrino T **S- Editor:** Song XX **L- Editor:**A
**E- Editor:** Lu YJ



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i3.64

*World J Meta-Anal* 2014 August 26; 2(3): 64-70
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

OBSERVATIONAL STUDY

# Development of the first Arabic cognitive dental anxiety scale for children and young adults

Abeer Al-Namankany, Paul Ashley, Aviva Petrie

Abeer Al-Namankany, Dental Department, King Fahd Armed Forces Hospital, Jeddah 21159, Saudi Arabia
Paul Ashley, Aviva Petrie, Eastman Dental Institute, University College London, WC1X8LD London, United Kingdom
Author contributions: Al-Namankany A contributed to conducted the research; Ashley P was first supervisor and management support; Petrie A was second supervisor and statistical support.
Correspondence to: Dr. Abeer Al-Namankany, BDS, PGCert Sedation, Dip Hypnosis, MClinDent Paed, MPhil Paed, GCAP, PhD, Paediatric Dentist, Hypnotherapist and Dental Phobia treatment expert, Dental Department, King Fahd Armed Forces Hospital, Prince Mohammed bin Abdul Aziz street, POB 9862, Jeddah 21159, Saudi Arabia. a.alnamankany@yahoo.com
Telephone: +966-26-653000 Fax: +966-26-653000
Received: April 6, 2014 Revised: May 27, 2014
Accepted: June 27, 2014
Published online: August 26, 2014

## Abstract

**AIM:** To validate the Arabic version of abeer children dental anxiety scale.

**METHODS:** Two ethical approvals for this study were obtained from United Arab Emirates, Ministry of Health and Dubai Health Authority; reference number: 2011/57. The Abeer children dental anxiety scale (ACDAS) was translated from English to Arabic by the native speaker chief investigator, and then back translated by another native speaker in Dubai (AS) to ensure comparability with the original one. Part C of ACDAS was excluded for the schoolchildren because those questions were only applicable for children at the dentist with their parents or legal guardian. A total of 355 children (6 years and over) were involved in this study; 184 in Dubai, 96 from the Religious International Institute for boys and 88 from Al Khansaa Middle School for girls. A sample of 171 children was assessed for external validity (generalizability) from two schools in different areas of London in the United Kingdom.

**RESULTS:** Receiver operating characteristic curve showed that the cut-off $\geqslant$ 26 for ACDAS gave the optimal results for sensitivity = 90% (95%CI: 81.2%-95.6%), and specificity = 86.6% (95%CI: 78.2%-92.7%), with AUROC = 0.93 (95%CI: 0.90-0.97). Cronbach's Alpha ($\alpha$) was 0.90 which indicated good internal consistency. Results of the external validity assessing the agreement between ACDAS and dental subscale of the children's fear survey schedule was substantial for the East London school ($\kappa$ = 0.68, 95%CI: 0.53-0.843); sensitivity = 92.9% (95%CI: 82.7%-98.0%); specificity = 73.5% (95%CI: 55.6%-87.1%) and almost perfect for the Central London school ($\kappa$ = 0.79; 95%CI: 0.70-0.88); sensitivity = 96.4% (95%CI: 81.7%-99.9%); specificity = 65.9%, (95%CI: 57.4%-73.8%).

**CONCLUSION:** The Arabic ACDAS is a valid cognitive scale to measure dental anxiety for children age 6 years or over.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Anxiety; Cognition; Children

**Core tip:** The Abeer children dental anxiety scale (ACDAS) scale is different from existing scales as it is the first dental anxiety scale for children which correlate dental anxiety with cognitive status. It can recognise the stimuli for dental anxiety in a logical order, and has questions concerning the expectation of the child' s legal guardian about the behaviour of the child before the treatment, whether the child has any previous dental treatment experience and the dentist's rating for the child's behaviour at the end of the treatment at the same visit. Finally, when assessing the external validity of the binary ACDAS, it was shown that its results compared favourably with those of the main study ($\kappa$ = 0.79, sensitivity = 96.4%, specificity = 65.9%) when applied to children in a different London school ($\kappa$ =

0.68, sensitivity = 92.9%, specificity = 73.5%). Therefore, ACDAS was shown to work well in two different locations with different children, which suggests that it is a generalisable scale. Based on the findings of this study, it is proposed that the ACDAS encompasses the required criteria for the gold standard dental anxiety scale for children.

Al-Namankany A, Ashley P, Petrie A. Development of the first Arabic cognitive dental anxiety scale for children and young adults. *World J Meta-Anal* 2014; 2(3): 64-70 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i3/64.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i3.64

## INTRODUCTION

Children dental anxiety has been a matter of concern for many years but despite this the etiology is still not entirely understood[1]. Anxiety may occur without cause, or it may be based on a real situation that leads to a reaction that is out of proportion to what would normally be expected. Severe anxiety can have a serious impact on daily life and effect quality of life and its different dimensions, such as speaking, eating, and appearance, and through these also social intercourse[2]. Dental anxiety is cumulative over time, and its development is influenced by multiple variables. It is most likely to start in childhood[3]. There are three general sources of information that have been evaluated as measures of anxiety for children and adults: (1) the behavioral measures, which is what the patient does, such as overt distress, general behavior, or specific motor acts like gripping the chair arms tightly. The results of these measures tend to be more subjective than the objective ones; if two dentists are observing the behavior of the same patient on the same time there is no guarantee that both will score the patient in a similar way. The differences in scoring could depend on the time of the appointment, the experience, the temperament, the age, and the gender of the dentist. Hence, the reliability of these measures will not be strong enough if they are used for research purposes. However, it might be the only method that could be used with preschool children; (2) the physiological measures, which is the measurement of the patient' s responses to the dental anxiety, such as rapid breathing, profuse sweating, muscle tension, pulse rate, or heart rate. These measures were neither reliable nor practical in use with children because the scene, the sound, and the application of the equipment might increase the child's anxiety level[4]. The use of the physiological measures were found to be less appropriate for assessing dental fear in children[5] for several reasons: the standard normal reading for children will vary and depend on age of the child; the results of these measures could be overlapped with current medical problems; the requirement of knowledge and training on how to use these equipments; wrong results by faulty machines; it is not available in all dental clinics; the practicality of using it in terms of cost, time, maintenance and a space in the clinic; last but not least. It is not appropriate to be used for children of all age groups; and (3) the self reported measures, which is what the patient says about his/her fear *via* direct report or scaling, interview, or inventory[6,7]. These measures are the most reliable measures for children who are able to read and have the cognitive ability to understand how to report their anxiety on the scale. Previous studies found that, in adult patients, the self reported anxiety scale can distinguish between high or low dental anxiety in terms of avoidance or distress behaviors. This may not be the case in preschool children; the ability of the young children may not be fully developed and they tend to report more fears regardless of the situation and more likely to show anxiety at separation from the parent. For those young children, the use of the behavioral measures is the best option[8].

Studies of dental anxiety in children rather than in adults may allow us to more reliably investigate the causes and management of dental anxiety. This is due to the limited reliability and validity of adult dental anxiety studies and to the extensive time span between the onset of the anxiety during the childhood and these studies[9]. Although measurement of dental anxiety is important for research and delivery of high quality clinical care, it is the corner stone of dental anxiety management.

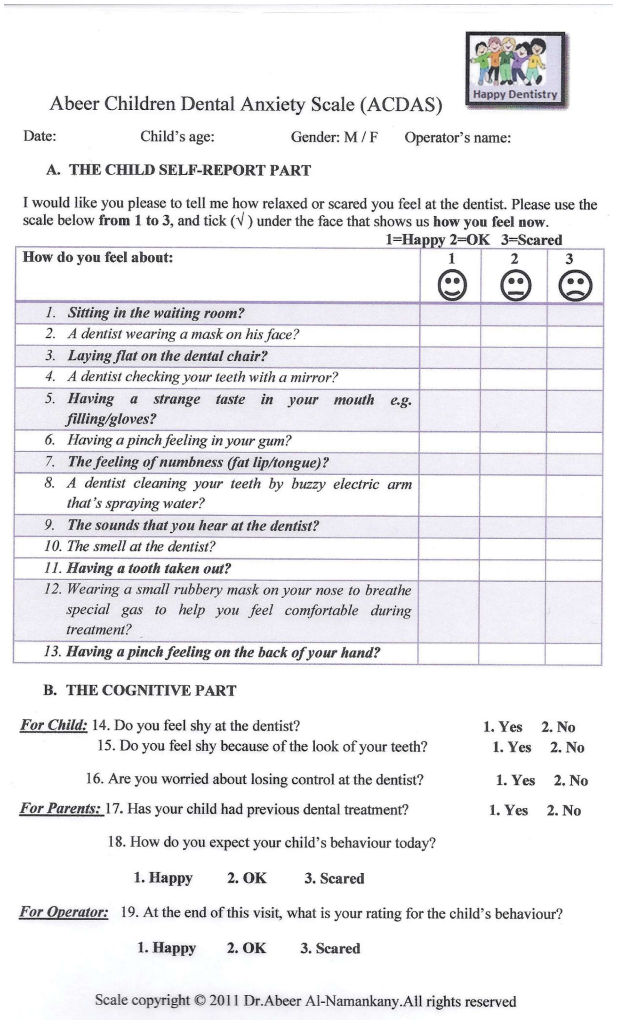
The development of self-reported measures was started in early 1960s and has continued up until the present. Dental anxiety measures have been developed in order to help the dentist detect anxious patients in order to provide better management and treatment. The degree of belief in negative cognition is associated with the severity of DA[10], the negative thinking patterns of the anxious individual is centered on danger and harm. The cognitive measures are widely used as self-report scales that request the patient to respond to list of statements or questions, these measures could be incorporated into pediatric DAM[11]. Abeer children dental anxiety scale (ACDAS) is the first children dental anxiety scale that incorporated the cognitive questions and is a valid cognitive scale to measure DA for children aged $\geqslant$ 6 years[12]. Although there are 14 different dental anxiety scales for children, some of them have been validated in many languages[13]; to date there is no DA scale that validated in Arabic language. Hence, the objective of this study was to validate the Arabic version of ACDAS in order to extend its benefits to more people and to be the first Arabic dental anxiety scale.

## MATERIALS AND METHODS

This study was made up of two parts, development of a new scale and then validation of this scale. According to the regulations of the United Arab of Emirates, two ethical approvals for this study were obtained from the Ministry of Health and Dubai Health Authority; Reference number: 2011/57.

### *Development of the scale*

The previously validated (Abeer Children Dental Anxiety



Abeer Children Dental Anxiety Scale (ACDAS)

Happy Dentistry

Date: Child's age: Gender: M / F Operator's name:

**A. THE CHILD SELF-REPORT PART**

I would like you please to tell me how relaxed or scared you feel at the dentist. Please use the scale below **from 1 to 3**, and tick (√) under the face that shows us **how you feel now**.

1=Happy 2=OK 3=Scared

| How do you feel about: | 1 | 2 | 3 |
|---|---|---|---|
| 1. *Sitting in the waiting room?* | | | |
| 2. *A dentist wearing a mask on his face?* | | | |
| 3. *Laying flat on the dental chair?* | | | |
| 4. *A dentist checking your teeth with a mirror?* | | | |
| 5. *Having a strange taste in your mouth e.g. filling/gloves?* | | | |
| 6. *Having a pinch feeling in your gum?* | | | |
| 7. *The feeling of numbness (fat lip/tongue)?* | | | |
| 8. *A dentist cleaning your teeth by buzzy electric arm that's spraying water?* | | | |
| 9. *The sounds that you hear at the dentist?* | | | |
| 10. *The smell at the dentist?* | | | |
| 11. *Having a tooth taken out?* | | | |
| 12. *Wearing a small rubbery mask on your nose to breathe special gas to help you feel comfortable during treatment?* | | | |
| 13. *Having a pinch feeling on the back of your hand?* | | | |

**B. THE COGNITIVE PART**

*For Child:* 14. Do you feel shy at the dentist? 1. Yes 2. No

15. Do you feel shy because of the look of your teeth? 1. Yes 2. No

16. Are you worried about losing control at the dentist? 1. Yes 2. No

*For Parents:* 17. Has your child had previous dental treatment? 1. Yes 2. No

18. How do you expect your child's behaviour today?

1. Happy 2. OK 3. Scared

*For Operator:* 19. At the end of this visit, what is your rating for the child's behaviour?

1. Happy 2. OK 3. Scared

Scale copyright © 2011 Dr.Abeer Al-Namankany.All rights reserved

**Figure 1 Abeer children dental anxiety scale-Arabic.** Scale copyright © 2011 Al-Namankany. All rights reserved.

Scale "ACDAS")[12] was translated from English to Arabic by the native speaker chief investigator, and then back translated by another native speaker in Dubai to ensure comparability with the original one (Figure 1). ACDAS is a 19 item, cognitive scale which can be used for children from age 6 years and over, we proposed the following name - Abeer Children Dental Anxiety Scale-Arabic (ACDAS-Arabic). It is made up of three parts (Figure 1): (1) this comprises 13 self-reported questions arranged in logical order. Each question uses three faces as a response set. Face "1" represents the feeling of a relaxed not scared "Happy"; face "2" represents a neutral/fair feeling "OK"; and face "3" represents the anxious feeling "Scared". The child is asked to tick under the face that best represents the child's response to the question and a mark (1, 2 or 3) is assigned accordingly. The range of values is therefore from 13 to 39; (2) this comprises three self-reported questions which afford a cognitive assessment, each question uses "Yes" or "No" as a response; and (3) this comprises three questions for further assessment of the child as reported by the legal guardian and the dentist, each question uses"Yes" or "No" as a response.

### *Validation of the scale*

The inclusion criteria for this study were children aged of 6 years or over, with no learning disability, and the ability to read Arabic. The children had to be at least 6 years of age, because younger children do not have the cognitive complexity required to report and react to dental situations accurately and they may not have the experience of dental situations[14]. A convenience sample of 184 students participated in this study; 96 males (The Religious International Institute), and 88 females (Al Khansaa Middle School). The study composed of two parts: assessment of reliability and validity, and assessment of generalizability or external validity.

### *Assessment of reliability and validity*

On the first visit, the local department of school health in Dubai gave permission for the study to be conducted on children in specific schools in their jurisdiction. In addition, permission from each school principal and a verbal consent by the students were also obtained prior to the start of the study. During the class time and in the presence of the teacher for each class, ACDAS-Arabic was completed by each child after being administered twice, once by each of two observers in order to measure the inter-observer reliability, and, in addition, the chief investigator administered ACDAS-Arabic twice, one week apart, to each child in order to measure the intra-observer reliability. Each child on the first visit also completed dental subscale of the children's fear survey schedule (CFSS-DS) after it was administered by the chief investigator in order to assess the validity of ACDAS-Arabic. On the second visit, seven students (4 males/3 females) who participated in the first visit were absent. Therefore there were seven missing from the total sample ($n$ = 184) which resulted in the 177 participants for the analysis.

**Statistical methods for numerical anxiety scores:** Initially the scores from Part-A of ACDAS-Arabic (the first 13 questions) were summed to provide a numerical anxiety score for each child at each visit. Intra-observer and inter-observer agreement were each assessed by performing a paired t-test to determine if there was a systematic effect, creating a Bland Altman diagram to assess whether the agreement was independent of the magnitude of the score, calculating the British Standards repeatability/reliability coefficient to provide the maximum likely difference between a pair of measurements, and determining Lin's concordance correlation coefficient as a measure of agreement. The Pearson correlation coefficient was determined between the scores of ACDAS-Arabic and CFSS-DS to investigate concurrent validity, and Cronbach's alpha evaluated to assess internal consistency.

**Statistical methods for the two anxiety categories (anxious/ not anxious):** Creating a categorical outcome facilitates the use of the ACDAS-Arabic scale for clinical and research purposes in terms of translating its numerical score into clinically relevant outcomes (anxious/not anxious). The

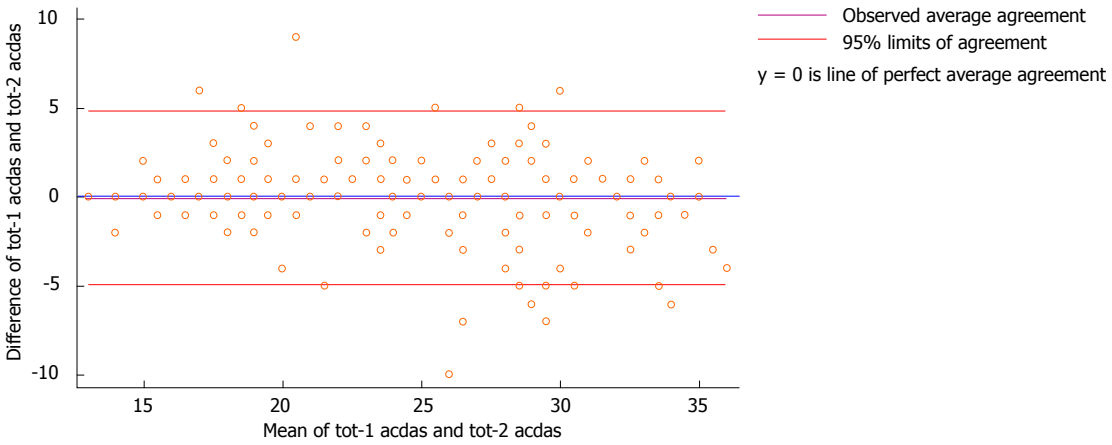



**Figure 2 Bland and Altman diagram for the 1st and 2nd scores from one observer.**

ACDAS-Arabic questionnaire results from the second visit of the 177 children were used to determine the sensitivity and the specificity for different cut-off values of the total score for Part A to distinguish anxious from not anxious children. The receiver operating characteristic (ROC) curve, plotting the sensitivity against 100, specificity for different cut-offs, was used to select an optimal cut-off value for the new scale. The classification of anxious and not anxious for these 177 children was also determined using the previously published optimal cut-off of ⩾ 36 for the CFSS-DS scale[9].

Cohen's kappa (k) with its confidence interval (CI) was evaluated to assess intra-observer and inter-observer reliability, and the discriminative validity when comparing the binary outcomes of ACDAS-Arabic and CFSS-DS. Convergent validity was assessed by using the Chi-squared test or Fisher's exact test when expected frequencies were small to compare dental anxiety with each of the other variables defined by the questions in Part B of ACDAS-Arabic.

### *Assessment of eexternal validity (generalizability)*

In order to know whether the scale and its dichotomized score will work well in populations that are different from the one from which it was derived, and to assess whether the cut-off "26" of ACDAS produces the similar results in terms of anxiety for different samples of children, a sample of 171 children was assessed for external validity (generalizability) from two schools in different areas of London in the United Kingdom; 81 from St. Alban' s Primary school in Central London and 90 from Cayley primary school in East London. In addition, bootstrapping[15] was used because data had not been collected at other schools on the visit to Dubai and it was not possible to travel to Dubai to collect additional data. Bootstrapping is a simulation process which involves estimating the parameter of interest from each of many random samples of size 177 (in this instance) by sampling with replacement from the original sample of size 177.

## RESULTS

### *Numerical anxiety scores*

The analysis of the numerical anxiety scores from ACDAS-Arabic indicated good reliability for both intra- and inter-observer agreement (Lin's concordance correlation coefficient 0.91 (95%CI: 0.89-0.94) and 0.92 (95%CI: 0.90-0.94), respectively; a value of 1 indicates perfect agreement. There was no evidence of a funnel effect in either of the Bland Altman diagrams assessing intra- and inter-observer reliability, and the limits of agreement for them were -4.93 to 4.84 and -3.87 to 5.47, respectively (Figures 2 and 3). The British Standards repeatability/reliability coefficient indicates the maximum likely differences between a pair of measurements were 4.9 and 4.5 for intra-and inter observer reliability, respectively. Using the first set of results for ACDAS-Arabic from the principal observer, the Pearson correlation coefficient between the ACDAS-Arabic and CFSS-DS indicated moderate concurrent validity ($r$ = 0.46, $P$ = 0.007). Cronbach's Alpha ($\alpha$) for ACDAS-Arabic was 0.90 which indicated a good internal consistency.

### *Two anxiety categories (anxious/ not anxious)*

The sensitivity and the specificity of the ACDAS-Arabic were determined for different cut-off points of the numerical anxiety scores (*i.e.*, the sum of the scores from Questions 1 to 13) as a means of distinguishing anxious from not anxious children. The cut-off point closest to the top left hand corner of the receiver operating characteristic (ROC curve) is circled in red (Figure 4). It gives the optimal results for sensitivity (90.0%, 95%CI: 81.2%-95.6%) and specificity (86.6%, 95%CI: 78.2%-92.7%). The area under the curve was 0.93 (95%CI: 0.90-0.97) as indicated in Figure 3. (A test which is perfect at discriminating between the two outcomes has an area under the curve of one).

There was almost perfect intra-observer agreement for the binary anxiety outcomes, using a cut-off of ⩾ 26 for ACDAS-Arabic to indicate anxiety, when the questionnaire was administered one week apart by the chief investigator ($\kappa$ = 0.91; 95%CI: 0.85-0.97), and almost perfect inter-observer agreement ($\kappa$ = 0.89; 95%CI: 0.82-0.96). There was substantial agreement between the two binary anxiety scales (ACDAS-Arabic with a cut-off of ⩾ 26 and CFSS-DS with a cut-off of ⩾ 36) ($\kappa$ = 0.76; 95%CI: 0.67-0.86),

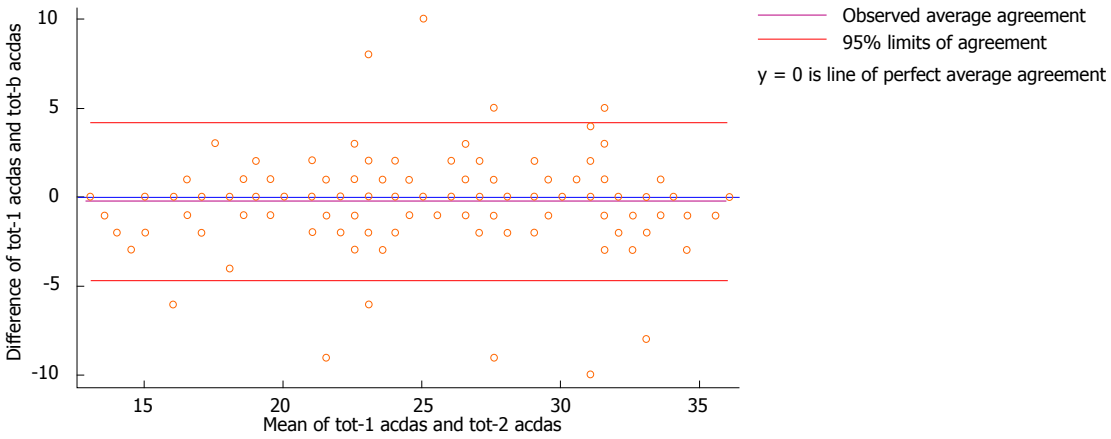



**Figure 3 Bland and Altman diagram for the scores by from the two observers (tot-1 and tot-b).**

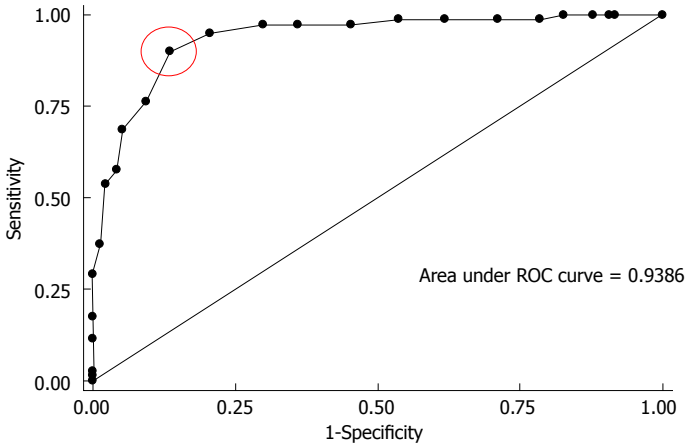



**Figure 4 Receiver operating characteristic curve for different cut-offs of Abeer children dental anxiety scale-Arabic.**

providing evidence of good discriminative validity.

Convergent validity indicated that there was a strong relationship between DA and cognition. The correlation coefficient is statistically significant ($P = 0.004$) for question 14: "Do you feel shy at the dentist?". There was no evidence of a linear relationship between the score of dental anxiety that was reported by the child and his/her answer to the question 15: "Do you feel shy because of the way your teeth look?" ($P = 0.25$). However, there was a highly significant relationship between child's DA and the cognitive question16: "Are you worried about losing control at the dentist?" ($P < 0.001$).

### *External validity*

One thousand bootstrap replications for 177 observations showed substantial agreement between the two binary scales (ACDAS-Arabic/CFSS-DS). This result compared favorably with the previous result that was obtained from the two schools in London, as shown in Table 1, which suggested that ACDAS-Arabic is working well in another location for another sample and it is a generalizable scale.

## DISCUSSION

Given the fact that there is currently no Arabic dental anxiety measure, the idea of the initiator of ACDAS, who is a native Arabic speaker, was to translate ACDAS to Arabic (ACDAS-Arabic) and validate it as the first cognitive and dental anxiety scale in the Arab world. ACDAS was validated as the first cognitive dental anxiety scale for children and adolescents; it included questions about the dental experience in a logical order and not only the most common feared items as the previous scales. Moreover, it included the perception of losing control; embarrassment; self-confidence and the cognitive nature of the child as important factors in anxiety provoking.

This study has shown almost perfect results for both numerical and categorical outcomes; the children had to be at least 6 years of age, because younger children do not have the cognitive complexity required to report and react to dental situations accurately and they may not have the experience of dental situations.

Given the significance of the crucial role of negative cognitive patterns in anxiety evocation that could make the person apprehensive and difficult to treat dentally and who also might not easily comply with anxiety treatment techniques, the present results were in line with the previous similar studies on adults. These demonstrated a strong relation between the negative thoughts and the level of dental anxiety[11,16,17]. The perception of losing control, embarrassment and self-confidence are impor-

**Table 1 Comparing Abeer children dental anxiety scale-Arabic (≥ 26) to dental subscale of the children's fear survey schedule (≥ 36)**

| | Central London | Dubai | East London |
|---|---|---|---|
| Kappa (95%CI) | K = 0.79 (0.70-0.88) | K = 0.76 (0.67 -0.86) | K = 0.68 (0.53-0.84) |
| Sensitivity (95%CI) | 96.40% (81.7%-99.9%) | 90% (81.2%-95.6%) | 92.90% (82.7%-98.0%) |
| Specificity (95%CI) | 65.90% (57.4%-73.8%) | 86.60% (78.2%-92.7%) | 73.50% (55.6%-87.1%) |

tant factors in anxiety provoking; these results suggest that 91% to 95% of the children who reported negative cognitions on questions 14, 15, and 16 were anxious; 98% were reported in other studies for adults[17]. The cut-off point for anxiety for ACDAS-Arabic (≥ 26) gave the optimal results for sensitivity (86.8%) and specificity (86.2%). These values suggested that ACDAS-Arabic has the ability to identify the anxious and non-anxious individual correctly. In addition, the area under the curve was 0.93: if the score discriminates perfectly, the AUROC equals 1[15].

The strong correlation between the ACDAS-Arabic and the CFSS-DS scores supports the validity of the ACDAS-Arabic in the dental setting, *i.e.*, the ACDAS-Arabic measures what it intends to measure, it includes items that are relevant to the most of children's dental experience and it asks about the child's five sensations (*i.e.*, sight, hearing, taste, smell, and touch). Moreover, it includes items that are relevant to treatment under inhalation and intravenous sedation. Treatment under general anesthesia was not included because the child will be asleep and will not really face the actual dental experience. ACDAS-Arabic is easy to administer and it took a very short time (3 min) to do so.

One of the limitations of this study was the use of CFSS-DS in its English version for validating the ACDAS-Arabic. To date, there is no Arabic DA scale that could be used instead. Therefore the English version of the CFSS-DS had to be translated into Arabic, and read and explained verbally by the chief investigator.

Another limitation was that the order of administration of the ACDAS-Arabic and the CFSS-DS for the school children was not randomized; it was impossible to do this because of the time restriction, as the administration was during the class time. Because of this time constraint, each child could not have a one-to-one interview with the observer in order to complete the ACDAS-Arabic and the CFSS-DS questionnaires. Instead the questionnaires were read to the class as a whole and the all the children in a class completed them at the same time. A third limitation was that the validation of this scale was planned for both a clinical and school setting but, because of the restrictions of cost and the time that the principal investigator would have had to spend in Dubai to obtain the information from a clinical setting, only school children were used.

It is crucial to understand the importance of measuring children dental anxiety and its correlation with the cognitive status of the child. ACDAS helps to highlight the unmet needs of many children who do not go to dentists because of fear of general anesthesia (GA). While some cases may still require GA, with appropriate anxiety management there is a significant number in whom it could be avoided.

Finally, although prevention is better than the treatment, to date there is no study that includes dental anxiety measurement in the list of preventative strategies which usually includes oral hygiene instruction, diet advice, fissure sealant, chlorhexidine and fluoride application[18]. Therefore, the first author suggests the inclusion of dental anxiety measure as a prevention item from the first visit and throughout the dental treatment. Assessing patients' thoughts could be a first step on the development of cognitive treatment strategies for dental anxiety.

The Arabic version of ACDAS is a valid and generalizable cognitive dental anxiety scale for children and adolescents.

## ACKNOWLEDGMENTS

The authors would like to thank Dr. Abdul Rahim Almarri, Dr Aisha Sultan, Ms. Fayza Albadri, and the head teachers at all schools in Dubai-UAE. Special thanks to Dr. Paul Ashley who encouraged me to conduct this study in Dubai, and to the children who participated in this study.

## COMMENTS

### *Background*

Dental anxiety is still remain as one of the main problems that caused avoiding visiting the dentists. Assessment of dental anxiety is the first step toward a better management.

### *Research frontiers*

The Abeer children dental anxiety scale (ACDAS) scale is different from existing scales as it is the first dental anxiety scale for children which correlate dental anxiety with cognitive status.

### *Innovations and breakthroughs*

It can recognise the stimuli for dental anxiety in a logical order, and has questions concerning the expectation of the child's legal guardian about the behaviour of the child before the treatment, whether the child has any previous dental treatment experience and the dentist's rating for the child's behaviour at the end of the treatment at the same visit.

### *Applications*

Finally, when assessing the external validity of the binary ACDAS, it was shown that its results compared favourably with those of the main study ($\kappa$ = 0.79, sensitivity = 96.4%, specificity = 65.9%) when applied to children in a different London school ($\kappa$ = 0.68, sensitivity = 92.9%, specificity = 73.5%).

### *Terminology*

Therefore, ACDAS was shown to work well in two different locations with dif-

ferent children, which suggests that it is a generalisable scale. Based on the findings of this study, it is proposed that the ACDAS encompasses the required criteria for the gold standard dental anxiety scale for children.

### *Peer review*

The work is well done and structured.

## REFERENCES


1 **Townend E**, Dimigen G, Fung D. A clinical study of child dental anxiety. *Behav Res Ther* 2000; **38**: 31-46 [PMID: 10645022 DOI: 10.1016/S0005-7967(98)00205-8]

2 **Luoto A**, Lahti S, Nevanperä T, Tolvanen M, Locker D. Oral-health-related quality of life among children with and without dental fear. *Int J Paediatr Dent* 2009; **19**: 115-120 [PMID: 19250394 DOI: 10.1111/j.1365-263X.2008.00943.x]

3 **Tickle M**, Jones C, Buchannan K, Milsom KM, Blinkhorn AS, Humphris GM. A prospective study of dental anxiety in a cohort of children followed from 5 to 9 years of age. *Int J Paediatr Dent* 2009; **19**: 225-232 [PMID: 19486376 DOI: 10.1111/j.1365-263X.2009.00976.x]

4 **Klingberg G**, Berggren U. Dental problem behaviors in children of parents with severe dental fear. *Swed Dent J* 1992; **16**: 27-32 [PMID: 1579885]

5 **Aartman I,** Everdingen T, Hoogstraten J, Schuurs A. Appraisal of Behavioral Measurement Techniques for Assessing Dental Anxiety and Fear in Children: A Review. *J Psychopathol Behav Assess* 1996; **18**: 153-171 [DOI: 10.1007/BF02229114]

6 **Atkins CO**, Farrington FH. Informed consent and behavior management. *Va Dent J* 1994; **71**: 16-20 [PMID: 9540755]

7 **McGrath PA**. Measurement issues in research on dental fears and anxiety. *Anesth Prog* 1986; **33**: 43-46 [PMID: 3458389]

8 **Melamed BG**. Assessment and management strategies for the difficult pediatric dental patient. *Anesth Prog* 1986; **33**: 197-200 [PMID: 3465253]

9 **Boman UW**, Lundgren J, Elfström ML, Berggren U. Common use of a Fear Survey Schedule for assessment of dental fear among children and adults. *Int J Paediatr Dent* 2008; **18**: 70-76 [PMID: 18086029 DOI: 10.1111/j.1365-263X.2007.00863.x]

10 **de Jongh A**, Muris P, ter Horst G, Duyx MP. Acquisition and maintenance of dental anxiety: the role of conditioning experiences and cognitive factors. *Behav Res Ther* 1995; **33**: 205-210 [PMID: 7887880 DOI: 10.1016/0005-7967(94)P4442-W]

11 **Ayer W**. Psychology and dentistry: mental health aspects of patient care. 1st ed. New York: Haworth Press Inc., 2005

12 **Al-Namankany A**, Ashley P, Petrie A. The development of a dental anxiety scale with a cognitive component for children and adolescents. *Pediatr Dent* 2012; **34**: e219-e224 [PMID: 23265158]

13 **Al-Namankany A**, de Souza M, Ashley P. Evidence-based dentistry: analysis of dental anxiety scales for children. *Br Dent J* 2012; **212**: 219-222 [PMID: 22402535 DOI: 10.1038/sj.bdj.2012.174]

14 **Alwin NP**, Murray JJ, Britton PG. An assessment of dental anxiety in children. *Br Dent J* 1991; **171**: 201-207 [PMID: 1910981 DOI: 10.1038/sj.bdj.4807661]

15 **Petrie A**, Sabin C. Medical Statistics at a Glance. 3rd ed. USA: Blackwell Publishing, 2009

16 **Briers S.** Brilliant Cognitive Behavioural Therapy. 1st ed. USA: Prentice Hall, 2009

17 **De Jongh A**. Dental Anxiety: A Cognitive Perspective. Thesis, 1995

18 **Sarmadi R**, Gahnberg L, Gabre P. Clinicians' preventive strategies for children and adolescents identified as at high risk of developing caries. *Int J Paediatr Dent* 2011; **21**: 167-174 [PMID: 20961342]


**P- Reviewer:** Lopez-Jornet P, Tomofuji T **S- Editor:** Ji FF
**L- Editor:** A **E- Editor:** Lu YJ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i3.71

*World J Meta-Anal* 2014 August 26; 2(3): 71-77
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.



# Interventions to improve treatment adherence among adolescents: A meta-analysis

Tricia A Miller, Brittany L Bannon, M Robin DiMatteo

Tricia A Miller, Brittany L Bannon, M Robin DiMatteo, Department of Psychology, University of California, Riverside, CA 92507, United States

Author contributions: Miller TA and DiMatteo MR were responsible for the conception and design of the study; Miller TA and Bannon BL collected the data, organized, and coded relevant empirical articles; Miller TA and Bannon BL statistically analyzed the data and performed moderator analysis; Miller TA and Bannon BL drafted the manuscript with critical revisions performed by DiMatteo MR.

Correspondence to: Tricia A Miller, MA, Department of Psychology, University of California, 900 University Avenue, Riverside, CA 92507, United States. trishmiller122@gmail.com
Telephone: +1-951-8275243 Fax: +1-951-8273985
Received: November 29, 2013 Revised: April 17, 2014
Accepted: July 17, 2014
Published online: August 26, 2014

## Abstract

**AIM:** To examine the overall effectiveness of interventions designed to improve medical treatment adherence among adolescent patients.

**METHODS:** PubMed and PsycINFO databases were searched to retrieve and analyze empirical journal articles (from 1948-2013). Only peer-reviewed, English language journals that defined a measure of adherence (or compliance), assessed an intervention aimed at improving adherence among adolescents, and provided information to calculate an *r* effect size were included. Studies were excluded if they lacked assessment of the effectiveness of interventions on improving adherence in adolescents as compared to no interventions or standard care. Case studies or journal articles that examined substance abuse or psychological disorders were also excluded. Analyses were conducted with fixed and random-effects methods, and moderators of intervention efficacy were also examined.

**RESULTS:** For each study that met the inclusion criteria ($n$ = 45), an effect size $r$, reflecting the strength and direction of the interventions' relationship to adherence was recorded; a positive $r$ indicated that the intervention increased adolescent adherence, whereas a negative $r$ indicated that the intervention decreased adolescent adherence. The overall effectiveness of adolescent adherence interventions was positive and significant (unweighted mean $r$ = 0.27, 95%CI: 0.21-0.33, $P$ = 0.001). Moderator analyses at the fixed effects level revealed that interventions were less effective when adolescents reported their adherence behaviors, when the type of adherence regimen was a medication regimen, and when the type of intervention was cognitive-modification based.

**CONCLUSION:** These findings contribute to understanding interventions for enhancing adolescent adherence. Future research should continue to examine the specific challenges faced by adolescents and create targeted interventions.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Adolescent; Adherence (compliance); Intervention; Meta-analysis

**Core tip:** Estimates of nonadherence among the adolescent population range from 25%-70%, depending on the disease or condition. Intervention components in patient samples vary widely across studies; thus, it is important to systematically identify elements of interventions that are most effective. Meta-analytic techniques were used in this study to provide a comprehensive, quantitative summary of empirical studies evaluating the effectiveness of interventions aimed at improving treatment adherence among adolescents. This meta-analysis showed that interventions were effective, specifically when the type of regimen was behavioral, whereas cognitive-based interventions were less effective.

Miller TA, Bannon BL, DiMatteo MR. Interventions to improve treatment adherence among adolescents: A meta-analysis. *World J Meta-Anal* 2014; 2(3): 71-77 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i3/71.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i3.71

## INTRODUCTION

Patient nonadherence (or noncompliance) to medical treatment involves the degree to which an individual fails to follow specific disease management activities as directed by his or her healthcare provider[1,2]. Nonadherence can occur in the context of a variety of treatment behaviors, such as medication use, electronic pill counts, diet and exercise, and disease management skills[1]. The prevalence of nonadherence can be close to 25% on average, across a variety of diseases, and specific patient populations[2]. For some chronic diseases such as diabetes, that require complex treatment regimens, rates of nonadherence can exceed 50%[2]. Moreover, it is estimated that nearly 240 million medical visits are wasted annually due to nonadherence[2]. Despite these alarming statistics, patients and healthcare providers remain largely unaware that one major cause of poor health outcomes may be patient nonadherence[3,4].

### Adolescent adherence relationship

Adolescence is a unique period of the lifespan in which individuals begin to explore their social identities and seek independence; however, among adolescents coping with illnesses, this autonomy-seeking can bring about confusion, frustration, and conflict with caregivers over prescribed treatment regimens, and consequently, nonadherence. In a study conducted by DiGirolamo *et al*[5] 30% of adolescents with cystic fibrosis reported completing less than half of their daily-prescribed airway clearance regimens that are essential for the prevention of further morbidity and even mortality. In another study of adolescent patients with Type 1 diabetes, 25% reported mismanagement behaviors, such as missing prescribed insulin injections. Similarly, between 50 and 60% of children and adolescents were found to underuse prescribed medications, and less than 10% overused prescribed medications[6,7]. A study by Chappuy *et al*[8] looked at prescription medication adherence in adolescents and found that only 36.2% actually completed their prescribed regimens. Lastly, Guilfoyle *et al*[9] noted that nonadherence to an oral immunosuppressant medication regimen, commonly used to prevent a patient's body from rejecting an organ transplant, is prevalent (approximately 70%) and has been found to significantly compromise the long-term graft survival and life span of adolescents with kidney transplants. Nonadherence is a prevalent and consequential issue for adolescent patients and their families; however, the development of effective interventions to improve adherence behaviors in this age group remains an ongoing challenge. The various types of interventions established to date and their benefits are reviewed below.

### Interventions to improve adherence in adolescents

Studies of current interventions to promote adherence among adolescents have consistently shown that educational interventions alone are not sufficient to change adherence behaviors[10]. In fact, a meta-analysis conducted by Dean *et al*[11] revealed that a multifaceted approach to interventions showed the greatest potential efficacy in improving adherence behaviors. The optimal combination of intervention elements remains unclear, however. Dean *et al*[11] suggest that the combination of education and behavioral methods (*e.g.*, reinforcement) for increasing motivation and providing problem-solving strategies can produce the greatest results. Additionally, in an intervention study conducted by Wysocki *et al*[10], behavioral family systems therapy improved both adherence to treatment and family relationship quality among adolescents, providing evidence for the importance of social support and family cohesion in disease management. Moreover, in a separate intervention conducted by Wysocki *et al*[12], researchers looked at adherence to self-monitoring of blood glucose, utilizing a behavioral intervention that compared two groups: a pill count meter-alone group and a pill count meter-plus-behavioral contract group. Results from this study indicated that both groups showed moderate improvement in measures of diabetic control, demonstrating the value of behavioral reminders and patients' commitment to their own care.

In addition to interventions combining educational and behavioral components, those incorporating cognitive-behavioral principles have also been successful. For example, research by van Es *et al*[13] found that adolescents with asthma demonstrated better treatment adherence if they received both education and group therapy for disease-focused issues, including attitudes toward disease and coping skills. In another study by Magyary *et al*[14], a cognitive-behavioral intervention significantly increased therapeutic adherence and self-responsibility for the management of health conditions in children and adolescents. Although the components of adolescent adherence interventions vary widely from study to study, and aspects have yielded results in particular patient samples, it is important that the overall efficacy of this wide range of interventions is subjected to systematic, quantitative review.

### Present study

The purpose of the present study is to utilize meta-analytic techniques to review and summarize research findings on the effectiveness of interventions designed to improve medical adherence in adolescents. Additionally, potential moderators of the effectiveness of these interventions will be examined. More specifically, this meta-analysis will test the primary hypothesis that there is a positive relationship between adherence interventions and adolescents' adherence to their medical treatment regimens, such that interventions would, on average,

improve adherence among adolescents. Exploratory moderator analyses will also be used to evaluate whether the following factors explain variability in the intervention studies' effect sizes: type of sample (adolescent only versus children and adolescents combined), type of intervention (*e.g.*, educational, behavioral), type of adherence measure (*e.g.*, self-report, electronic), type of treatment regimen (*e.g.*, medication-based, behavioral), type of illness, parental involvement in the intervention, patient gender, and patient ethnicity.

## MATERIALS AND METHODS

### Literature search

A "Top-down" literature search was conducted for the retrieval and analysis of empirical journal articles published from 1948 through 2013. PubMed and PsycINFO databases were searched using combinations of the following keywords: adherence (compliance), persistence AND adolescent, youth, teens, children, interventions, disease management, self-management, randomized control trial. In addition, the reference sections of obtained journal articles were examined for relevant studies.

### Inclusion and exclusion criteria

Studies were included if published in a peer-reviewed, English language journal, if they defined and explained a method of measuring adherence (or compliance), and if they assessed an intervention that aimed to improve adherence to medical treatment among adolescents. Studies were included only if they provided an effect size *r* or statistical information to calculate an effect size representing the magnitude and direction of the interventions' effect on adherence. Meta-analytic techniques were used to extract average *r* effect sizes and assess their significance. Furthermore, relevant journal articles were coded for moderators of the interventions' effectiveness, including: age or age range, total *N*, location of the study (United States or non-United States), type of patient illness, type of intervention, how adherence was measured, type of regimen, whether or not there was parental involvement in the management of care, patient gender, and patient ethnicity. Studies were excluded if they: measured the effectiveness of interventions on adherence in adult patients; did not ASSESS patient adherence to treatment; lacked an intervention to improve adherence; lacked assessment of the effectiveness of an intervention on improving adherence in adolescents; and/or provided no information to calculate an *r* effect size. Finally, case studies or journal articles that examined substance abuse or psychological disorders (*i.e.*, depression, bipolar disorder) were excluded, because treatment adherence in mental health and substance use is beyond the scope of the present research, although it is an important issue for future examination.

### Effect size calculation

An *r* effect size was calculated from Cohen's *d*, *P* or descriptive statistics. If a study reported a significant result but did not report a *P*-value, then the one-tailed *P*-value was assumed to be 0.025. If a study reported results that were non-significant and no exact *P*-value was provided, then the study was conservatively assigned $P = 0.5$ one-tailed, and $r = 0$[15]. An effect size of $r = 0$ indicated that the intervention did not have an effect on adolescent adherence. For studies in which there were multiple measures of adherence, the *r* for each measure was converted to a Fisher's $Z_r$ and they were averaged.

### Statistical analysis

The effect size *r* was used because *r* most clearly illustrates both the *strength* (from 0.00 to 1.00) and *direction* (positive or negative) of the relationship between variables[15,16]. In this meta-analysis, a positive *r* indicates an improvement in adherence as a result of the intervention, whereas a negative *r* indicates a decrease in adolescent adherence to medical treatment as a result of the intervention (relative to control or standard care groups). The effect size *r* was obtained for each of the studies. If statistics were provided that could be transformed into an *r* (*e.g.*, *t*, Z and *P*-value, chi-square, or 1 degree of freedom in the numerator *F*; or means and standard deviations), the appropriate statistical analyses were conducted to yield a Phi, Pearson Product-Moment, or point-biserial correlation coefficient[17]. All calculations involving *r* were performed by transforming *r* to the Fisher's *Z* transformation of *r* and then returning the results back to the scale of *r*.

The random effects model was used to combine effect size statistics using the unweighted mean *r* based on *k* (the total number of studies included). This method allows for the generalization of findings to other studies beyond those that were included in this meta-analysis[15,18]. The fixed effects model was also used to carry out weighted mean analyses and tests of heterogeneity based on *N* (the total number of participants across all studies). All analyses of moderators were first performed using the random effects model; if results were not significant at the random effects level, results from the fixed effects model were provided. Random effects tests of methodological and substantive moderators were conducted to examine the heterogeneity of the study effects. These include: sample type (adolescent and children versus adolescents only), total *N*, location of the study (United States or non-United States), patient illness, type of intervention, how adherence was measured, type of regimen, whether or not there was a parent involved in the management of care, patient gender and patient ethnicity. In addition, for the effects that were significant, the fail safe *N* was calculated (to address the file drawer problem) that indicated the number of studies, new, unpublished or unretrieved with no effect that would be needed in order for significant results to be declared non-significant at $P < 0.05$[17]. The standardized odds ratio and standardized relative risk (including 95%CI) were calculated from the unweighted mean *r* using the binomial effect size display

(BESD). The BESD is a useful tool for effect size estimation that can be used to display changes in success rates (*i.e.*, survival or improvement rates) that are attributable to specific treatment procedures[17]. Preliminary statistical analyses were conducted using SPSS 12.0 (*i.e.*, calculation of means, medians, standard deviations, correlations and *t*-tests). A T1-84 Plus graphing calculator and Excel 2008 v.12.2.3 were used for essential calculation verification.

## Results

Meta-analytic calculations were performed on 45 independent studies to examine the overall effectiveness of interventions aimed at improving adherence to medical treatment among adolescents. Twenty-four of the 45 studies combined the results of the interventions' effectiveness on adherence among both adolescents and pediatric patients. The remaining 21 studies looked at the effectiveness of interventions aimed at improving adherence among adolescents only. Therefore, meta-analytic computations were done for the total sample ($k$ = 45) and also separately for these two groups, constituting a moderator analysis for "type of sample." In addition, for each sample ($k$ = 45, and the subgroups of $k$ = 24 adolescents plus children, $k$ = 21 adolescents only) the following statistics were computed: the total number of subjects ($N$), the median $r$ and range, the fixed effects weighted mean $r$ with a 95%CI, the random effects model unweighted mean $r$ with a 95% confidence interval, the fail safe $N$, the standardized odds ratio with a 95%CI, and the standardized relative risk with a 95% confidence interval (Table 1).

### Interventions to improve adherence

Across 45 independent studies, with a total of 3890 participants, the average relationship between an adherence intervention and improvement in adolescent adherence (as compared to a control group or to a group receiving standard care) was positive and significant under the random effects model (unweighted mean $r$ = 0.27, $P$ < 0.001). This demonstrated that interventions aimed at improving adolescent adherence were effective. The median $r$ of 0.23 was close in magnitude to the unweighted mean and to the weighted (by sample size) mean $r$ of 0.18. Effect sizes in the positive direction indicated that interventions aimed at increasing adherence were effective; conversely, effect sizes in the negative direction indicated that these interventions reduced adherence. Within this sample of studies, there were only two negative $r$ effect sizes: -0.24 and -0.05.

Both the fixed (weighted) and random (unweighted) effects models indicated a positive and significant effect of interventions on improving adherence to medical treatment in adolescents (Table 1). The random effects model indicated that adherence interventions were effective [$r$ = 0.27, $t_{(44)}$ = 7.55, $P$ < 0.001]. Therefore, the effectiveness of interventions to improve adherence among adolescents can be generalized to studies outside the present sample. The fixed effects model also indicated that interventions aimed at improving adherence were effective ($r$ = 0.18, 95%CI: 0.15-0.21; $P$ < 0.001). The fail safe $N$ demonstrated that more than 2189 studies with non-significant results would have to be included in order for these results to be rendered non-significant; this number exceeds the tolerance level of 235 unpublished (or otherwise non-retrievable) studies with null results that might possibly exist. The standardized odds ratio, using the BESD, indicated that the odds of being adherent to medical treatment were 3.03 times higher if adolescents participated in an adherence intervention compared with the odds if he/she did not participate in an intervention (95%CI: 2.35-3.94; $P$ < 0.001). The standardized relative risk (also calculated using the BESD) indicated that the risk of poor adherence to medical treatment was 1.74 times higher if adolescents did not participate in an adherence intervention (95%CI: 1.53-1.99; $P$ < 0.001). Furthermore, the effect sizes were significantly heterogeneous ($\chi^2$ = 122.82, $P$ = 2.24 × $10^{-9}$), which indicated that moderators might account for this variation in effect sizes.

### Moderator analyses

Moderator analyses using the random effects model revealed no significant results in studies that combined children and adolescents and in studies of adolescents only. Meta-analytic calculations were applied to each sample type separately and are detailed below.

In the 24 studies that included a combined sample of adolescents and children, there was a moderate, significant and positive effect of interventions using both the fixed (weighted) and random (unweighted) effects models (Table 1). The random effects model indicated that interventions to improve adherence to medical treatment in adolescents were effective [$r$ = 0.32, $t_{(23)}$ = 5.58, $P$ < 0.001]. Additionally, the weighted mean (fixed effects model) yielded a similarly positive result ($r$ = 0.27, 95%CI: 0.22-0.32; $P$ < 0.001). The fail safe $N$ demonstrated that more than 800 studies with non-significant results would have to be included in order for these results to be rendered non-significant. However, the tolerance level suggested that 130 unpublished null studies might possibly exist. The BESD-based, standardized odds ratio indicated that the odds of adhering to medical treatment were 3.77 times higher if patients participated in an adherence intervention as compared to the odds if they had not (95%CI: 2.44-5.99; $P$ < 0.001). The standardized relative risk indicated that the risk of nonadherence to medical treatment was 1.94 times higher if the patient did not participate in an adherence intervention (95%CI: 1.56-2.45; $P$ < 0 001). Furthermore, the 24 effect sizes were heterogeneous, ($\chi^2$ = 53.36, $P$ = 3.28 × $10^{-4}$). Inspection of the distribution of $r$'s revealed a range from $r$ = -0.24 to $r$ = 0.71.

In the comparison subgroup of 21 studies that assessed adolescent patients only, analyses revealed a moderate, yet positive and significant, effect of interventions on improvements in adherence to medical treatment,

**Table 1 Summary of overall meta-analysis results**

| Effect of adolescent adherence interventions | K[4] | Total *n*[5] | Unweighted median *r* (range) | Weighted mean *r*[6] (95%CI) | Unweighted mean *r*[7] (95%CI) | Fail safe *n* | Standardized odds ratio[8] (95%CI) | Standardized relative risk[9] (95%CI) |
|---|---|---|---|---|---|---|---|---|
| Interventions[1] | 45 | 3890 | 0.23 (-0.24-0.71) | 0.18 (0.15-0.21)[b] | 0.27 (0.21-0.33)[b] | 2189 (tolerance level 235) | 3.03 (2.35, 3.94)[b] | 1.74 (1.53-1.99)[b] |
| Adolescent and Children Interventions[2] | 24 | 1476 | 0.325 (-0.24-0.71) | 0.27 (0.22-0.32)[b] | 0.32 (0.22-0.42)[b] | 800 (tolerance level 130) | 3.77 (2.44, 5.99)[b] | 1.94 (1.56, 2.45)[b] |
| Adolescent Only Interventions[3] | 21 | 2414 | 0.18 (-0.05-0.51) | 0.12 (0.08-0.16)[b] | 0.20 (0.13-0.27)[b] | 324 (tolerance level 115) | 2.25 (1.69-3.03)[b] | 1.5 (1.30-1.74)[b] |

[b]$P < 0.01$. [1]Analyses for the entire sample of intervention studies; [2]Analyses for the subgroup of studies that combined adolescent with pediatric samples; [3]Analyses for the subgroup of studies that included adolescent-only samples; [4]Number of samples; [5]Total *n* across all samples; [6]Effect size obtained from the fixed effects model, or weighted by the total number of participants across studies. [7]Effect size obtained from the random effects model, or based on the total number of studies included. [8]The standardized odds ratio depicts the odds of being adherent in the intervention group relative to the control group. Across all analyses, the intervention group had a higher likelihood of improved adherence than the control group (receiving no interventions or standard care). [9]The standardized relative risk can be interpreted as the control group's risk for nonadherence as compared to the intervention group's risk. In each set of analyses, the control group was at significantly greater risk for nonadherence. The Fail Safe *n* exceeds the level of tolerance for future null results making it unlikely that the "file drawer problem" is a source of bias. The binomial effect size display from the unweighted mean effects (random effects model) was used to obtain the standardized odds ratio and relative risk. The heterogeneity test ($k = 45$) for the overall adolescent adherence interventions was significant ($\chi^2 = 122.82$, $P < 0.001$). The heterogeneity test ($k = 24$; adolescent and children samples) for adolescent adherence interventions was significant ($\chi^2 = 53.36$, $P < 0.001$). The heterogeneity test ($k = 21$; adolescent only samples) for the second group of adolescent adherence interventions was also significant ($\chi^2 = 44.53$, $P < 0.001$).

using both the fixed (weighted) and random (unweighted) effects models (Table 1). First, the random effects model indicated that adolescent-only adherence interventions were effective [$r = 0.20$, $t_{(20)} = 5.18$, $P < 0.001$]. The fixed effects model revealed the same significant intervention efficacy, albeit with a slightly smaller effect ($r = 0.12$, 95%CI: 0.08-0.16; $P < 0.001$). In addition, the fail safe *N* demonstrated that more than 324 studies with non-significant results would have to be included in order for these results to be rendered non-significant. The tolerance level suggested that 115 unpublished null studies possibly exist. The standardized odds ratio (BESD-based) indicated that the odds of being adherent to medical treatment were 2.25 times higher if the adolescent participated in an adherence intervention compared with the odds if he/she had not participated in an intervention (95%CI: 1.69-3.03; $P < 0.001$). The standardized relative risk indicated that the risk of nonadherence to medical treatment was 1.50 times higher if the adolescent did not participate in an adherence intervention (95%CI: 1.30-1.74; $P < 0.001$). Furthermore, this set of 21 effect sizes was heterogeneous ($\chi^2 = 44.53$, $P = 1.28 \times 10^{-3}$).

Analysis of other potential moderators of the relationship between interventions and improvements in adherence was conducted. Moderator analysis at the fixed effects level revealed three significant moderators: type of regimen, intervention type, and self-report by adolescents of their own adherence behaviors. For type of regimen, the effectiveness of interventions aimed at improving adherence was moderated by whether or not the intervention was a medication regimen. Specifically, adherence interventions had a greater positive effect on adherence to health behaviors such as diet, exercise, appointment keeping or screening regimens than to medication regimens ($z = -1.77$, $P = 0.039$). In addition, the fixed effects analyses revealed that interventions were less successful at improving adherence in studies where adolescents reported their own adherence behaviors, as compared to having a parent or guardian report adolescent adherence behaviors ($z = -1.91$, $P = 0.038$). Finally, interventions that involved only cognitive modification were less effective in improving adherence than were approaches based on educational intervention, behavioral intervention or a combination of both ($z = -2.14$, $P = 0.16$).

## RESULTS

The present meta-analysis provided a comprehensive, quantitative summary of empirical studies evaluating the effectiveness of interventions aimed at improving medical treatment adherence among adolescents. The main hypothesis, that interventions focused on improving treatment adherence are indeed effective, was supported. Although moderator analyses using the random effects model revealed no significant overall moderators of studies' effect sizes, there were several significant results from the fixed effects approach. In studies where adolescent patients reported their own adherence behaviors, interventions demonstrated reduced efficacy, suggesting the possibility of measurement challenges in this work. Interventions improved non-medication regimen adherence (*e.g.*, diet, exercise, appointment keeping) more than medication adherence, perhaps due to the greater challenges of health behavior change, or because medication regimens may be more difficult to target. Finally, educational and behavioral interventions (both combined and in isolation) were more effective than cognitive approaches. These cognitive approaches often attempt to change adolescents' attitudes and beliefs about risk, and they may be less effective because adolescence is a time of heightened vulnerability to risk taking behaviors[19]. Research suggests that because of the temporal gap between puberty and the slower maturation of the cognitive-control system,

changing the context in which the risky behavior occurs maybe more successful than changing the way adolescents think about risk[19].

## DISCUSSION

These results provided compelling evidence for the success of efforts to address adolescent nonadherence through interventions designed to assist with the complexities of treating chronic illnesses in this age group. It should be noted that although there were no significant differences between studies that combined adolescents and children samples and those that included adolescents alone, future research should continue to look at these age groups separately, as barriers to adherence can differ between children and adolescents[20]. In doing so, interventions aimed at improving adherence can be targeted to better address the specific needs of each group.

### *Strengths, limitations and future directions*

This systematic quantitative review of adolescent adherence interventions sought to explain what aspects of these interventions were the most successful. With regards to comprehensiveness, several search strategies were utilized and all references were carefully cross-checked. Furthermore, although the mean effect sizes were moderate in size, they may be important in clinical application, and therefore should not be underestimated. Research in the medical field commonly reports small, but highly significant findings with major implications for health. For example, the relationship between consumption of aspirin and the occurrence of heart attacks is in the range of $r$ = 0.03-0.04[21,22]. In other words, there is a 3% to 4% risk difference in prevention of a serious health outcome due to consumption of a simple medication such as aspirin, making the application of this finding very important clinically. In the present research, an unweighted mean $r$ of 0.27 reflects a 27 percent difference in the risk of nonadherence between patients who receive an adherence enhancing intervention and those who do not. Adherence interventions, thus, can have a profound impact on improving adherence among adolescents.

Limitations of this research include the possibility that some empirical studies were missed unintentionally. For example, it is possible that statistically significant findings had greater likelihood of publication, but the large fail safe *N*'s in this review made it unlikely that the current results exhibited the "file drawer bias." Additionally, several studies in the meta-analysis combined the results of both adolescents and children in their reports of the interventions' effectiveness on improving adherence. Therefore, results from the present meta-analysis should be interpreted with caution. Future studies should assess adherence behaviors in children and adolescents separately to allow for exploration of potential age-specific factors that may influence adolescents' adherence and interventions' efficacy. Findings from this study (*i.e.*, the positive effects of multi-faceted, educational/behavioral interventions as compared to cognitive approaches, and of parental assessments of adherence) also underscore the importance of shared decision-making and the role that adolescents, healthcare providers, and parents or caregivers play in the management of disease.

In sum, future research should identify the psychological and behavioral aspects and determinants of adolescent adherence. The present review could assist in the development of specific interventions to enhance adolescent adherence to various medical treatments and types of treatment regimens. Future adherence interventions should also measure and seek to determine both mediators and moderators of adherence interventions' effectiveness in order to fine-tune their development and eliminate the less successful elements. Lastly, future studies should recruit adolescent-only samples, thereby providing meta-analytic opportunities to better understand the challenges (or facilitators) of treatment adherence specific to adolescents.

## ACKNOWLEDGEMENTS

The authors would like to thank Mia Montoya for helping to analyze effect sizes and moderators, in addition to, managing the data table for this meta-analysis. Thanks are also due to research assistants: Sang Nguyen, Shiromi Ratnayake, Amberly Navalta, Kim Pham, Vanessa Gutierrez, Shriraj Shah, and Punam Patel for their assistance.

## COMMENTS

### *Background*

For adolescents coping with illness, autonomy seeking and other aspects of this developmental transition can lead to confusion, frustration, and conflict with caregivers over prescribed treatment regimens, and consequently, nonadherence. Components of adolescent adherence interventions have been found to vary widely from study to study; thus it is important that the overall efficacy of this wide range of interventions is subjected to systematic, quantitative review.

### *Research frontiers*

Interventions are effective at improving adherence among adolescents. However, previous adolescent studies have treated adolescent and child patients as a homogenous group and reported findings based on the combination of adolescents and children. This limits the interpretability of findings; more quantitative reviews are needed that focus on the unique challenges of adolescent nonadherence.

### *Innovations and breakthroughs*

Previous studies have reported inconsistent results on the efficacy of interventions aimed at improving adherence among adolescents. Some research suggests that multifaceted interventions are more effective than single-approach interventions. The present meta-analysis found that cognitive-based interventions were less effective at improving adherence in adolescents than educational and behavioral approaches.

### *Applications*

The study results suggested that interventions aimed at improving treatment adherence in adolescents are effective, specifically when the type of regimen was behavioral. In addition, cognition-based interventions were less effective. Given the unique challenges adolescents face in coping with illness, future research should consider developmentally appropriate intervention aspects.

### *Terminology*

Patient adherence (or compliance) is the degree to which patients follow treatment directives given by clinicians or other health care providers.

### *Peer review*

Patient nonadherence is a worldwide public health concern, linked to poor clinical outcomes and increased medical costs. The paper was well written. It is interesting and novel topic for present social medical questions.

## REFERENCES


1 **DiMatteo MR**, Giordani PJ, Lepper HS, Croghan TW. Patient adherence and medical treatment outcomes: a meta-analysis. *Med Care* 2002; **40**: 794-811 [PMID: 12218770 DOI: 10.1097/01.MLR.0000024612.61915.2D]

2 **DiMatteo MR**. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. *Med Care* 2004; **42**: 200-209 [PMID: 15076819 DOI: 10.1097/01.mlr.0000114908.90348]

3 **Zolnierek KB**, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. *Med Care* 2009; **47**: 826-834 [PMID: 19584762 DOI: 10.1097/MLR.0b013e31819a5acc]

4 **DiMatteo MR**, Haskard-Zolnierek KB, Martin LR. Improving patient adherence: A three-factor model to guide practice. *Health Psychol Rev* 2011; 1-18 [DOI: 10.1080/17437199.2010.537592]

5 **DiGirolamo AM**, Quittner AL, Ackerman V, Stevens J. Identification and assessment of ongoing stressors in adolescents with a chronic illness: an application of the behavior-analytic model. *J Clin Child Psychol* 1997; **26**: 53-66 [PMID: 9118176]

6 **Chmelik F**, Doughty A. Objective measurements of compliance in asthma treatment. *Ann Allergy* 1994; **73**: 527-532 [PMID: 7998669]

7 **Coutts JA**, Gibson NA, Paton JY. Measuring compliance with inhaled medication in asthma. *Arch Dis Child* 1992; **67**: 332-333 [PMID: 1575560 DOI: 10.1136/adc.67.3.332]

8 **Chappuy H**, Tréluyer JM, Faesch S, Giraud C, Chéron G. Length of the treatment and number of doses per day as major determinants of child adherence to acute treatment. *Acta Paediatr* 2010; **99**: 433-437 [PMID: 19912146 DOI: 10.1111/j.1651]

9 **Guilfoyle SM**, Goebel JW, Pai AL. Efficacy and flexibility impact perceived adherence barriers in pediatric kidney post-transplantation. *Fam Syst Health* 2011; **29**: 44-54 [PMID: 21417523 DOI: 10.1037/a0023024]

10 **Wysocki T**, Harris MA, Buckloh LM, Mertlich D, Lochrie AS, Taylor A, Sadler M, Mauras N, White NH. Effects of behavioral family systems therapy for diabetes on adolescents' family relationships, treatment adherence, and metabolic control. *J Pediatr Psychol* 2006; **31**: 928-938 [PMID: 16401678 DOI: 10.1093/jpepsy/jsj098]

11 **Dean AJ**, Walters J, Hall A. A systematic review of interventions to enhance medication adherence in children and adolescents with chronic illness. *Arch Dis Child* 2010; **95**: 717-723 [PMID: 20522463 DOI: 10.1136/adc.2009.175125]

12 **Wysocki T**, Green L, Huxtable K. Blood glucose monitoring by diabetic adolescents: compliance and metabolic control. *Health Psychol* 1989; **8**: 267-284 [PMID: 2767019 DOI: 10.1037/0278-6133.8.3.267]

13 **van Es SM**, Nagelkerke AF, Colland VT, Scholten RJ, Bouter LM. An intervention programme using the ASE-model aimed at enhancing adherence in adolescents with asthma. *Patient Educ Couns* 2001; **44**: 193-203 [PMID: 11553420 DOI: 10.1016/S0738-3991(00)00195-6]

14 **Magyary D**, Brandt P. A school-based self-management program for youth with chronic health conditions and their parents. *Can J Nurs Res* 1996; **28**: 57-77 [PMID: 9128476]

15 **Rosenthal R**, DiMatteo MR. Meta-analysis: recent developments in quantitative methods for literature reviews. *Annu Rev Psychol* 2001; **52**: 59-82 [PMID: 11148299 DOI: 10.1146/annurev.psych.52.1.59]

16 **DiMatteo MR**, Haskard KB. Further challenges in adherence research: measurements, methodologies, and mental health care. *Med Care* 2006; **44**: 297-299 [PMID: 16565628]

17 **Rosenthal R,** Rosnow RL. Essentials of behavioral research: Methods and data analysis. 3rd ed. McGraw-Hill, 2008: 663-690

18 **Frattaroli J**. Experimental disclosure and its moderators: a meta-analysis. *Psychol Bull* 2006; **132**: 823-865 [PMID: 17073523 DOI: 10.1037/0033-2909.132.6.823]

19 **Steinberg L**. Adolescence. 8th ed. New York: McGraw-Hill, 2007

20 **DiMatteo MR**, Miller TA. Treatment adherence in adolescence. In: O'Donohue L, Benuto L, Tolle, editors. Handbook of adolescent health psychology. New York: Springer, 2013: 373-386

21 **Rosenthal R**, Rosnow RL. Essentials of behavioral research: Methods and data analysis. 2nd ed. New York: McGraw-Hill, 1991

22 **Rosenthal R,** DiMatteo MR. Meta-Analysis: Stevens' handbook of experimental psychology. 3rd ed. In: Pashler H, Wixted J, editors. Hoboken (NJ): John Wiley & Sons Inc, 2002


**P- Reviewer:** Deng Bo, Shen XC **S- Editor:** Wen LL
**L- Editor**: A **E- Editor:** Lu YJ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i3.78

*World J Meta-Anal* 2014 August 26; 2(3): 78-90
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Efficacy of therapeutic ultrasound *vs* sham ultrasound on pain and physical function in people with knee osteoarthritis: A meta-analysis of randomized controlled trials

Norma J MacIntyre, Ahmed Negm, Adalberto Loyola-Sánchez, Mohit Bhandari

Norma J MacIntyre, Ahmed Negm, Adalberto Loyola-Sánchez, School of Rehabilitation Science, McMaster University, Hamilton, Ontario L8S 1C7, Canada
Mohit Bhandari, Department of Surgery, McMaster University, Hamilton, Ontario L8L 8E7, Canada
Author contributions: All authors contributed to the design of the research study and interpretation of the data; MacIntyre NJ, Negm A and Loyola-Sánchez A contributed to the acquisition of data as described in the methods; MacIntyre NJ performed the search and data analyses and wrote the paper; Negm A, Loyola-Sánchez A and Bhandari M reviewed the paper critically for important intellectual content and approved the version to be published.
Supported by The Canadian Institutes of Health Research Randomized Controlled Trials Mentorship Program (NJM, MB), No. MTP 108229
Correspondence to: Norma J MacIntyre, PT, PhD, Associate Professor, School of Rehabilitation Science, McMaster University, IAHS 403-1400 Main St W., Hamilton, Ontario L8S 1C7, Canada. macint@mcmaster.ca
Telephone: +1-905-5259140-21166 Fax: +1-905-5240069
Received: November 27, 2013 Revised: May 5, 2014
Accepted: May 28, 2014
Published online: August 26, 2014

## Abstract

**AIM:** To determine the efficacy of therapeutic ultrasound vs sham for improving pain and physical function immediately post-intervention in people with knee osteoarthritis (OA).

**METHODS:** We hand searched meta-analyses on the topic published in 2010 and updated the search in three electronic databases (MEDLINE, EMBASE, CINAHL) January 1, 2009 to September 5, 2013 to identify relevant studies. The inclusion criteria were human randomized controlled trials published in the English language in which active therapeutic ultrasound was compared to sham ultrasound, data for people with knee OA were reported separately, participants were blinded to treatment allocation and outcomes assessed before and after treatment included pain, self-reported physical function and performance-based physical function. Two reviewers independently screened titles and abstracts retrieved in the search to identify trials suitable for full text review. Data extraction and risk of bias assessment of the identified trials were completed independently by two reviewers. Pooled analyses were conducted using inverse-variance random effects models.

**RESULTS:** We screened 1013 titles and abstracts. Meta-analysis of pain outcomes from 5 small trials (281 participants/OA knees) showed that, compared to sham ultrasound, therapeutic ultrasound improves pain [standardized mean difference (SMD) (95%CI) = -0.39 (-0.70--0.08); $P$ = 0.01] but not physical function [self-reported in 3 trials (130 participants/OA knees): SMD (95%CI) = -0.21 (-0.55-0.14), $P$ = 0.24; walking performance in 4 trials (130 participants/OA knees): SMD (95%CI) = -0.11 (-0.59-0.37), $P$ = 0.65). For the walking performance outcome, the dispersion of the estimated effects exceeded that expected due to sampling error ($\chi^2$ = 8.37, $P$ = 0.04, $I^2$ = 64%). Subgroup analyses of three trials that administered high dose ultrasound improved the consistency ($I^2$ = 28%) but the treatment effect remained insignificant.

**CONCLUSION:** Meta-analyzed double-blind placebo-controlled randomized trials provide low-strength evidence that therapeutic ultrasound decreases knee OA pain and very low-strength evidence that it does not improve physical function.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Ultrasonic therapy; Physical therapy modalities; Osteoarthritis; Double-blind method; Evidence-

based medicine

**Core tip:** Controversy exists regarding the efficacy of therapeutic ultrasound in the management of knee osteoarthritis (OA). Lack of participant blinding in effectiveness trials introduces bias known to exaggerate treatment effect estimates particularly for outcomes such as pain and self-reported physical function. We meta-analyzed data from double- and triple-blind trials only and high level evidence shows that therapeutic ultrasound decreases knee OA pain but does not increase physical function immediately following treatment. Due to the methodological quality of the included trials, we conclude that a large well-designed trial is required before this clinical question can be answered definitively.

MacIntyre NJ, Negm A, Loyola-Sánchez A, Bhandari M. Efficacy of therapeutic ultrasound *vs* sham ultrasound on pain and physical function in people with knee osteoarthritis: A meta-analysis of randomized controlled trials. *World J Meta-Anal* 2014; 2(3): 78-90 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i3/78.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i3.78

## INTRODUCTION

Knee osteoarthritis (OA) is a highly prevalent chronic condition and a leading cause of lower extremity disability in community-dwelling older adults in North America[1]. To date, no treatment exists which modifies the disease and, despite symptom management, pain and functional limitations may progress to the point where total joint replacement is required[2]. As the population ages and the prevalence of obesity increases, the associated economic and personal burden associated with knee OA is expected to rise[2]. Current clinical practice guidelines for managing knee OA recommend a combination of pharmacologic and nonpharmacologic treatment in order to decrease the need for surgical replacement of the damaged joint[3,4].

Therapeutic ultrasound is widely used in clinical settings for various musculoskeletal conditions[5,6]. In a provincial survey of 123 Canadian physical therapists treating clients with knee OA, 81% reported at least some use of ultrasound therapy in their multicomponent management of knee OA[7]. Recent meta-analyses demonstrate that 10 to 24 sessions of continuous or pulsed ultrasound reduces knee OA pain[8-10]. Using two different methods for investigating the effect of dose, two meta-analyses reported that low dose ultrasound achieved using the pulsed mode to administer low intensity (0.375-0.625 W/cm$^2$) sound waves produced greater pain relief than high dose ultrasound[8,9]. The authors of the most recent meta-analysis back transformed the effect estimate for pain to a visual analogue scale (VAS) score and found a difference of -16.3 [95%CI: -20.9-(-11.7)] mm which was judged to reflect a clinically important change[10]. The effect of therapeutic ultrasound on physical function (self-reported and walking performance) was not significant in the two meta-analyses published in 2010[8,9]. However, the meta-analysis published in 2012 (6 trials, 387 participants) found clinical and statistically significant effects on composite physical function and gait function outcomes[10]. Although these systematic reviews provide evidence for the efficacy of therapeutic ultrasound in the management of knee OA pain and physical disability, all reported that the few, small trials with low methodological quality eligible for inclusion limit the confidence in the effect estimates[8-10].

Meta-analyses of randomized trials provide the highest level of evidence for evaluating the effectiveness of clinical interventions such as therapeutic ultrasound. However the quality of this evidence depends on study design characteristics that yield comparable intervention and control groups. A meta-epidemiologic study of 1973 trials found that intervention effect estimates in trials using subjective outcomes such as self-reported measures were inflated and heterogeneity between trials was increased when double blinding was absent or unclear[11]. It appears that bias due to lack of double blinding exaggerates the effect estimates and heterogeneity for subjectively assessed outcomes more than other study design flaws such as inadequate/unclear random sequence generation and inadequate/unclear allocation concealment[11]. In at least half of the trials included in previous systematic reviews investigating the efficacy of ultrasound therapy on pain and physical function in people with knee OA, no attempt was made to administer sham ultrasound[8-10]. As a result, between-trial heterogeneity and the estimates of the effect of ultrasound on self-reported pain and physical function outcomes published in the highest level evidence available to date are likely to be inflated due to a lack of or unclear blinding of participants with knee OA.

Therefore the objective of this systematic review and meta-analysis was to determine if therapeutic ultrasound *vs* sham ultrasound is effective in improving pain and physical function immediately following the intervention period in people with knee OA blinded to treatment allocation. As a secondary objective, we determined treatment safety based on reported side effects and adverse events.

## MATERIALS AND METHODS

Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) recommendations were followed to ensure full and transparent reporting of this review[12]. This systematic review was conducted following a pre-determined protocol which is available from the authors upon request. This protocol was not registered and is not publicly available.

### *Search strategy*

We updated the search for relevant papers completed for previous systematic reviews on this topic (to February 2009)[8,9], by searching OVID MEDLINE, OVID EMBASE, and CINAHL (through EBSCOhost) electronic

**Table 1 Example search strategy: Strategy used to search the OVID Medline electronic database (January 1, 2009 to September 5 2013)**

| | |
|---|---|
| 1 | Exp Osteoarthritis/ |
| 2 | osteoarthr$.ti,ab,sh. |
| 3 | gonarthr$.ti,ab,sh. |
| 4 | coxarthr$.ti,ab,sh. |
| 5 | arthr$.ti,ab. |
| 6 | [(knee$ or hip$ or joint$) adj3 (pain$ or ach$ or discomfort$)].ti,ab. |
| 7 | [(knee$ or hip$ or joint$) adj3 stiff$].ti,ab. |
| 8 | Exp Ultrasonic Therapy/ |
| 9 | Exp Ultrasonography/ |
| 10 | us.fs. |
| 11 | (ultrasound$ or ultrasonic$).tw. |
| 12 | short wave therapy.tw. |
| 13 | ultrasonograph$.tw. |
| 14 | randomized controlled trial.pt. |
| 15 | controlled clinical trial.pt. |
| 16 | randomized controlled trial.sh. |
| 17 | random allocation.sh. |
| 18 | double blind method.sh. |
| 19 | single blind method.sh. |
| 20 | clinical trial.pt. |
| 21 | Exp Clinical Trial/ |
| 22 | (clin$ adj25 trial$).ti,ab. |
| 23 | [(singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)].ti,ab. |
| 24 | placebos.sh. |
| 25 | placebo$.ti,ab. |
| 26 | random$.ti,ab. |
| 27 | research design.sh. |
| 28 | comparative study.sh. |
| 29 | exp evaluation studies/ |
| 30 | follow up studies.sh. |
| 31 | prospective studies.sh. |
| 32 | (control$ or prospectiv$ or volunteer$).ti,ab. |
| 33 | 1 or 2 or 3 or 4 or 5 or 6 or 7 |
| 34 | 8 or 9 or 10 or 11 or 12 or 13 |
| 35 | 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 |
| 36 | 33 and 34 and 35 |
| 37 | animal/ |
| 38 | animal/ and human/ |
| 39 | 37 not 38 |
| 40 | 36 not 39 |
| 41 | Limit 40 to yr = "2009-Current" |

databases from January 1, 2009 to September 5, 2013. The search strategy combined Medical Subject Headings (MeSH) terms and text words related to design (randomized controlled trial), condition (knee osteoarthritis), and intervention (therapeutic ultrasound). We reproduced the specific search strategy used for each database published in the previous systematic review focusing on clinical outcomes[9]. (Although previously published, the replicated strategy used to search the Medline electronic database is shown in Table 1 to comply with the PRISMA checklist.) No language limit was placed on the search in order to increase sensitivity. Duplicates were removed after all databases were searched. We searched for published trials only and attempted to contact primary authors to request additional information if necessary.

### Study eligibility criteria and selection

Two reviewers (NM, AL) independently screened the titles and abstracts for all citations identified by the search and retrieved all parallel-group randomized sham-controlled trials assessing the effect of ultrasound on pain or physical function in people with knee OA published in English. Studies examining other joints were only included if the data for people with knee OA were reported separately. Trials that included other interventions in addition to active ultrasound were included as long as the sham ultrasound group received the same combination of interventions. Lack of blinding of study participants was an exclusion criterion because we considered this to be a major source of biased results regarding ultrasound efficacy on self-reported pain and physical function outcomes[11]. Abstracts published as conference proceedings and not as full trials were excluded due to insufficient reporting of data for extraction. Cohen's unweighted kappa ($\kappa$) was used to measure agreement between reviewers. Disagreement was solved by consensus including a third reviewer (AN). A final list of eligible studies was prepared for full text review after title and abstract screening.

### Data extraction and management

The data collection form used for our previous systematic review[8] was modified for this updated review to reflect the narrowed research question focusing on patient-centred outcomes of pain and physical function. The form was independently pilot-tested on two randomly-selected studies by two reviewers (NM, AN) to ensure consistency in coding instructions. A double extraction method was followed using the refined data collection form. A physical therapist with expertise in research methodology and OA rehabilitation (NM) and an orthopaedic surgeon with expertise in OA (AN) reviewed the papers and extracted the data in a standardized manner. As recommended in the Cochrane Handbook[13], the reviewers were not blinded to any aspect of the trials during data extraction. Any disagreement was resolved through consensus.

The information extracted from each study included: (1) characteristics of the study participants (age, gender, diagnostic criteria, joint involvement, and knee OA severity); (2) characteristics of the therapeutic ultrasound intervention (device, frequency, mode, intensity, effective radiating area, surface area treated, application protocol, number and length of treatment sessions); (3) description of co-interventions; (4) description of pain and physical function outcomes and corresponding data; (5) reported adverse events and reasons for loss to follow up; and (6) general characteristics of the studies (location, clinical setting and funding source). The ultrasound dose was calculated as for our previous review[8] using the following formula:

Energy ($J/cm^2$) = [(average temporal intensity) (time) (effective radiating area)]/treated surface area

Group means for outcomes at baseline, post treatment, change from baseline and standard deviations (SDs), or the information from which SDs could be derived, were extracted. For trials that included two groups receiving active ultrasound (continuous and pulsed

mode), the formulae for combining means and SDs for two groups published in the Cochrane Handbook were used for estimating the effect of active ultrasound on the clinical outcomes of interest. (See Table 7.7a in the Cochrane Handbook 2005[13]). When a trial presented outcomes at time points other than pre- and post-intervention, we extracted the data for all time points, however, the mean values at the end of the treatment period were used in the meta-analysis. Heterogeneity in trial outcomes was minimized by pooling the data for the same outcome measure where possible, or pooling those that were most similar in terms of constructs assessed (*e.g.*, pain with movement measured by VAS or numeric rating scale, NRS).

### Risk of bias and quality assessment

Two reviewers (NM, AN) independently assessed risk of bias for each study and the level of agreement was determined using linear weighted kappa (κ). Any disagreement regarding risk of bias was resolved by consensus. The methodological domains recommended by The Cochrane Collaboration were assessed: randomization, treatment allocation concealment, blinding of participants, care givers and outcome assessors, completeness of outcome data, selective outcome reporting and other potential threats to validity[14]. Table 2 illustrates the criteria for assessing risk of bias in these domains. An overall risk of bias was considered for each study and across all studies based on the criteria outlined in the Cochrane Handbook[14]. We assigned studies as having high risk of bias if at least one criterion was not met. The quality of the body of evidence for each outcome was determined considering within-trial risk of bias, directness of evidence, heterogeneity, risk of publication bias and precision of effect estimates as recommended by the Grading of Recommendations, Assessment Development and Evaluation (GRADE) working group[15]. We defined treatment effects as precise when pooled estimates had reasonably narrow 95%CIs and the pooled sample size was greater than 400[15].

### Statistical analysis

We focused on patient-centred outcomes of pain and physical function. Follow up duration was to the end of the ultrasound therapy treatment. Review Manager Version 5 was used for data analyses. Effect size for each outcome (by study and overall) was estimated using the standardized mean difference (SMD) calculated as the raw mean difference divided by the pooled variance for each outcome to allow comparison of estimates of effect. Where the same outcome measure was used, the mean difference was also calculated. We used inverse-variance random-effects models to pool results to account for the inevitable variation in patient populations, concomitant treatments, and specific components of the physical therapy intervention as recommended when the number of studies is small and the reasons for heterogeneity across the studies cannot be reliably evaluated[16]. We evaluated the consistency of findings by comparing the direction and strength of effect along with the degree of statistical heterogeneity (based on the $\chi^2$ and $I^2$ statistical tests) in effects across studies. $\chi^2$ values with $P \geqslant 0.1$ and $I^2 < 60\%$ were considered to be acceptable homogeneity for pooling the data[16]. We used subgroup analyses and planned to conduct meta-regression to evaluate the effects of *a priori*-defined clinical characteristics (OA severity) and study characteristics (ultrasound dose) on pain and physical function outcomes. These subgroups were of interest because our previous meta-analysis demonstrated that the effect estimates for pain reduction were increased for those receiving low dose therapeutic ultrasound and heterogeneity observed for physical function and walking performance was reduced in the subgroup of trials including participants with mild knee OA severity[8]. We hypothesize that therapeutic ultrasound will be more effective in people with less severe knee OA in whom tissue damage is less advanced[17]. Confidence intervals (CI) at the 95% level were calculated for pooled estimates for each outcome and the *Z* test was used for determining the treatment effect. Statistical significance was considered at $P \leqslant 0.05$.

For each study reporting each of the three clinical outcomes of interest, we evaluated funnel plots for asymmetry in the standard error of the SMD as a function of the SMD. We did not use statistical tests for publication bias given that we did not include unpublished data sources in our search strategy.

### Role of the funding source

The study was funded, in part, by CIHR (NM, MB). CIHR had no role in the development of the question or the review protocol, literature search, data extraction and analysis, interpretation of the results, or preparation of the manuscript.

## RESULTS

Figure 1 is a flowchart of the results of search strategy. The initial database and hand searches retrieved 1226 citations ($n$ = 213 duplicates). After title and abstract screening, 10 randomized controlled trials were identified as eligible and retrieved for full text review. Of these, six were eligible for inclusion in the systematic review and five provided published data that permitted pooling. Reasons for exclusion are shown in Figure 1. The agreement between the reviewers in identifying the studies eligible for full text review and for inclusion in this review was κ = 1.0.

### Study characteristics

Characteristics of the six parallel-group, randomized, placebo-controlled trials eligible for inclusion in the systematic review are summarized in Table 1[18-23]. The RCTs were published between 1992 and 2012. Five trials included two arms (active ultrasound and sham ultrasound)[18,19,21-23], whereas the trial by Tascioglu *et al*[20]

**Table 2 Criteria used to classify risk of bias in trials according to components of methodological quality**

| | |
|---|---|
| Random sequence generation | |
| Low risk | Referring to a random number table; Using a computer random number generator; Coin tossing; Shuffling cards or envelopes; Throwing dice; Drawing of lots |
| High risk | Sequence generated by odd or even date of birth; Sequence generated by some rule based on date (or day) of admission; Sequence generated by some rule based on hospital or clinic record number; Allocation by judgment of the clinician; Allocation by preference of the participant; Allocation based on the results of a laboratory test or a series of tests; Allocation by availability of the intervention |
| Unclear | Insufficient information |
| Allocation concealment | |
| Low risk | Central allocation (including telephone, web-based); Sequentially numbered, opaque, sealed envelopes; An equivalent method was used to conceal allocation |
| High risk | Using an open random allocation schedule (*e.g.*, a list of random numbers); Assignment envelopes were used without appropriate safeguards (*e.g.*, if envelopes were unsealed or nonopaque or not sequentially numbered); Alternation or rotation; Date of birth; Case record number; Any other explicitly unconcealed procedure |
| Unclear | Insufficient information |
| Blinding of each of: participant, care provider, outcome assessor | |
| Low risk | Unlikely that the blinding could have been broken |
| High risk | Likely that the blinding could have been broken |
| Unclear | Insufficient information |
| Completeness of outcome data collection | |
| Low risk | No missing outcome data; Reasons for missing outcome data unlikely to be related to true outcome; Missing outcome data balanced in numbers across all groups, with similar reasons for missing data across groups; Missing data have been imputed using appropriate methods; For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size |
| High risk | Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; "As-treated" analysis done with substantial departure of the intervention received from that assigned at randomization. Potentially inappropriate application of simple imputation. For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size |
| Unclear | Insufficient information |
| Completeness of outcome reporting | |
| Low risk | The study protocol is available and all of the study's pre-specified outcomes that are of interest in the review have been reported in the pre-specified way; The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified |
| High risk | Not all of the study's pre-specified primary outcomes have been reported; One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (*e.g.*, subscales) that were not pre-specified; One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; The study report fails to include results for a key outcome that would be expected to have been reported for such a study |
| Unclear | Insufficient information |
| Other potential sources | |
| Low risk | The study appears to be free of other sources of bias |
| High risk | There is at least one important risk of bias. For example, the study had a potential source of bias related to the specific study design used or has been claimed to be fraudulent, or has some other problem |
| Unclear | Insufficient information |

included three arms in which one arm received active pulsed ultrasound (low dose), one arm received active continuous ultrasound (high dose), and one arm received sham ultrasound. Three trials[19,20,23] were conducted in Turkey and one trial was conducted in each of the United States of America[18], Canada[22], and China[21]. The duration of the intervention varied from 5 d to 8 wk and the dosage varied from 26 J/cm$^2$ to 196.3 J/cm$^2$ (Table 3). The control groups in all the studies received sham ultrasound and the participants were blinded to the allocated treatment. Only one trial confirmed that treatment providers were blinded to the intervention allocation[22]. In the trial conducted by Yang *et al*[21] it was unclear how the sham ultrasound was delivered. Furthermore, the trial by Yang *et al*[21] reported a "curative effect score" for the 87 participants (100 knees) randomized to active or sham ultrasound groups rather than reporting group means (SD) for pain and physical function outcomes and attempts to contact the authors to secure the raw data were unsuccessful. This calculated efficacy index was significantly improved for pain ($P < 0.001$) and physical function ($P < 0.001$)[21]. Nevertheless, the data from this trial[21] could not be included in the meta-analysis. The five trials eligible for meta-analysis reported data for a total of 281 participants and OA knees[18-20,22,23]. Table 3 shows that the number of participants providing data for analyses was lower than the number of participants randomized (varying from 2 to 8 participants per trial) in all trials meta-analyzed. Reasons given for loss to follow up included protocol violation (used analgesics)[19,20], illness[18], dissatisfaction with treatment[18], lack of time to attend/lack of regular attendance at sessions[20,23], transportation problems[18], or incomplete baseline assessment before withdrawing from the study[22]. The average age across the studies was approximately 61 years with the majority of participants being female (Table 3). Four

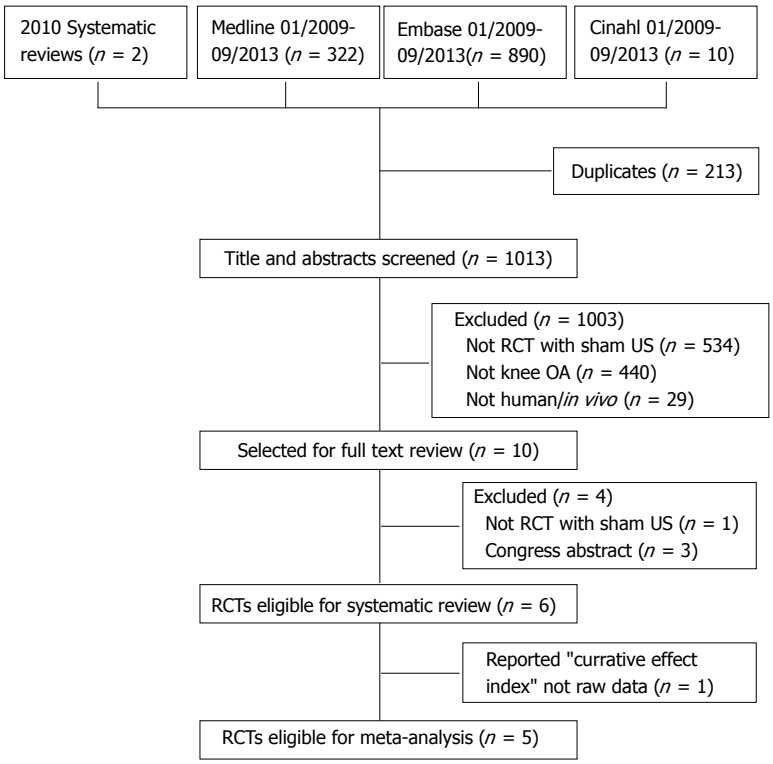



**Figure 1 Flow diagram for identification of eligible trials evaluating the effect of ultrasound vs sham on pain and physical function.**

trials[19,20,22,23] confirmed that participants had mild to moderate radiographic knee OA whereas one included trial[18] reported results for 69 participants/knees with unknown radiographic severity. (The latter trial included eight participants with a history of total knee replacement and all 69 had restricted knee range of motion for more than 6 mo[18].) Three trials reported that the majority of participants had bilateral knee OA[18,22,23] with one of these trials including only participants with bilateral knee OA[23]. Two of the five meta-analyzed trials applied active and sham ultrasound in combination with concurrent treatments (including isometric quadriceps exercises) that were identical for both study groups[18,23]. In contrast, two trials[19,20] did not allow any concurrent treatments including the use of analgesics during the intervention period. Three trials[18,20,22] administered low dose ultrasound ($< 150$ J/cm$^2$) and three trials[19,20,23] administered high dose ultrasound ($\geqslant 150$ J/cm$^2$). Three trials[19,22,23] reported scores on more than 1 pain outcome measure and two of these trials[22,23] reported scores on more than 1 self-reported physical function outcome measure. Table 3 specifies the pain and self-reported physical function outcomes selected for inclusion in the meta-analysis in order to reduce heterogeneity arising from differences in outcomes used across studies. Three trials[19,22,23] administered the WOMAC LK 3.1 questionnaire and reported scores on the physical function subscale. One trial[20], administering the WOMAC LK3.1 to assess the effect of ultrasound therapy on disability, reported only the total score which was not included in the meta-analysis for the self-reported physical function outcome. Another trial[21] administered the Lequesne Severity Index to assess the effect of ultrasound therapy on self-reported physical function but did not report the group means in a format that permitted pooling as mentioned previously. Walking performance was measured as walking speed (time to walk a specified distance) in four trials[18-20,23], and as distance walked (metres walked in six minutes) in one trial[22]. One trial[18] reported means for walking velocity (metres/s) pre and post treatment for the total sample only. We were unsuccessful in obtaining group means from the primary author and therefore these walking performance data could not be included in the meta-analysis. The units of measurement for the meta-analysed walking performance values were time (in seconds), where lower values indicate better scores, and distance (in metres), where higher values indicate better scores. Therefore, the group means for distance walked in six minutes were entered into the meta-analysis as negative values to adjust for the directional difference in interpreting better scores for this outcome.

### *Trial risk of bias and quality of the evidence*

The agreement between reviewers in determining risk of bias was $\kappa = 0.81$. Table 4 summarizes the methodological quality assessment of the six trials retrieved for full text review. Two trials[22,23] reported adequate random sequence generation and allocation concealment. Only one trial[21] did not report the blinding of participants adequately. Only one trial[22] confirmed that care providers were blinded (triple blind) and all five of the meta-analyzed trials confirmed that assessors were blinded to treatment allocation (double blind)[18-20,22,23]. In two included trials[19,20], it was unclear if complete data were collected. In one trial[18] included in the meta-analysis, data reporting was incomplete. Only one trial[21] (not included

**Table 3 Characteristics of the studies selected for full text review**

| | Falconer *et al*[18] | Ozgönenel *et al*[19] | Tascioglu *et al*[20] | Yang *et al*[21] | Loyola-Sánchez *et al*[22] | Ulus *et al*[23] |
|---|---|---|---|---|---|---|
| Trial registry number | Not available | Not available | Not available | Not available | NTC00931749 | Not available |
| Trial duration | 4-6 wk (12 sessions 2-3 x/wk) | 2 wk (10 sessions 5 x/wk) | 2 wk (10 sessions 5x/wk) | 5 d (5 sessions); + 1 mo follow up | 8 wk (24 sessions 3 x/wk) | 3 wk (15 sessions 5 x/wk) |
| Sample size | Randomized: 74; analyzed: 69 (35 CG) | Randomized: 67; analyzed: 65 (31 CG) | Randomized: 90; analyzed: 82 (27 CG) | Randomized and analyzed: 87 (100 knees; 50 CG) | Randomized: 27; analyzed: 25 (13 CG) | Randomized: 42; analyzed: 40 (20 CG) |
| Sample characteristics (Mean SD/*n* reported) | Age approximately 67.5 (11) yr; 50 F; All restricted knee ROM ⩾ 6 mo; 8 knee joint replacement; 51 bilateral OA | Age approximately 55 (7.5) yr; 54 F; Newly diagnosed; 31 mild OA, 36 moderate OA | Age approximately 60 (3) yr; 56 F; Disease duration 6.5 yr; 48 mild OA, 34 moderate OA | Age 58.3 yr; 72 F; Disease duration 2.8 yr | Age approximately 61.8 (10) yr; 21 F; 8 had mild OA, 19 had moderate OA; 24 bilateral OA | Age approximately 60.5 (9.5) yr; 34 F; disease duration 8.9(8.7) yr; 17 mild OA, 23 moderate OA; all bilateral OA |
| Ultrasound device | Chattanooga Intellect 200 | Peterson .250 | Sonopuls 434 | NERCUM | Chattanooga Intellect Mobile | Sonopuls 434 |
| Application protocol | 12 min; 1 MHz; intensity: 1.7 W/cm$^2$, continuous mode (*n* = 34) | 5 min; 1 MHz; intensity 1 W/cm$^2$; continuous mode (*n* = 34) | 5 min; 1 MHz; intensity 1 W/cm$^2$, continuous mode (*n* = 27); pulsed (duty cycle 20%, *n* = 28) | No details: 15min treatment model then 20min rehabilitation model (*n* = 50 knees) | 9.5 min; 1MHz; intensity 1 W/cm2; pulsed mode (duty cycle 20%, *n* = 12) | 10 min; 1 MHz; intensity 1 W/cm$^2$; continuous mode (*n* = 20) |
| Sham Application | Start button not pushed | Applicator disconnected from back of device | No output delivered | No output delivered | Ceramic crystal removed from soundhead | No output delivered and applicator disconnected from back of device |
| Application site | Knee flexed or extended per most restricted motion; treated surface area 100 cm$^2$ | Patellofemoral and tibiofemoral borders; treated surface area 25 cm$^2$ | Antero-medial and lateral parts of extended knee; treated surface area 60 cm$^2$ | Knee extended; 4 soundheads fixed on joint line; treated surface area not reported | Knee flexed to 90°; Soundhead fixed at antero-medial joint line. treated surface area 5 cm$^2$ | Antero-medial and lateral parts of extended knee; treated surface area 60 cm$^2$ |
| Dosage | 26 J/cm$^2$ | 150.7 J/cm$^2$ | 196.3 J/cm$^2$<br>39.3 J/cm$^2$ | Unable to calculate | 114 J/cm$^2$ | 196.3 J/cm$^2$ |
| Concurrent treatment | Stretching, joint mobilizations, exercises (ROM, bridging, isometric quads, home program) | none | none | none | None reported; use of analgesics not reported | Hotpacks (20 min); IFC (10 min); exercises (isometric quads); analgesics except during physio |
| Outcomes included in meta-analysis | Pain – 10 cm VAS | Pain on movement in past wk – 10 cm VAS<br>WOMAC LK 3.1 Physical Function subscale<br>Walking speed [time (s) to walk 50 m] | Pain on movement in past wk – 10 cm VAS<br>Walking speed [time (s) to walk 20 m] | none | Pain following walking test – 11 point NRS<br>WOMAC LK 3.1 Physical Function subscale<br>Distance (m) walked in 6 min | Pain on activity – 10 cm VAS<br>WOMAC LK 3.1 Physical Function subscale<br>Walking speed (time (s) to walk 50 m) |
| Funding source | Non-profit | Not reported | Not reported | NERCUM, institutional | Government | Unfunded |
| Comments | Trial author confirmed mode was continuous; pain data extracted from graphs; request for pain and walking data was unsuccessful | | Treated surface area not reported; estimated to be 3x the sound head size (based on parts of knee treated) | Attempts to contact authors for pain and physical function data that could be pooled were unsuccessful | 1 of the 2 reviewers co-authored the trial; outcomes pooled in this review were secondary outcomes in trial | Treated surface area not reported; estimated to be 3x the sound head size (based on parts of knee treated) |

OA: Knee osteoarthritis; CG: Control group; F: Female; NERCUM: National Engineering Research Centre of Ultrasound Medicine; ROM: Range of motion; IFC: Interferential current electrotherapy; VAS: Horizontal visual analogue scale (10 = worst pain); NRS: Numeric rating scale (10 = worst pain); WOMAC LK 3.1: Western Ontario and McMaster Universities Osteoarthritis Index (Likert scale version) self-report questionnaire physical function subscale (68 = worst limitation).

**Table 4 Methodological quality assessment of the randomized sham-controlled trials**

| Ref. | Random Sequence Generation | Allocation Concealment | Blinding of Participant | Blinding of Care Provider | Blinding of Assessor | Data Collection Complete | Complete Outcome Reporting | Free of Other Potential Bias | Risk of Bias[1] |
|---|---|---|---|---|---|---|---|---|---|
| Falconer *et al*[18] | Unclear | Unclear | Yes | No | Yes | Yes | No | Yes | High |
| Ozgönenel *et al*[19] | Unclear | Unclear | Yes | No | Yes | Unclear | Yes | Yes | High |
| Tascioglu *et al*[20] | Unclear | Unclear | Yes | No | Yes | Unclear | Yes | Yes | High |
| Yang *et al*[21] | Unclear | Unclear | Unclear | Unclear | Unclear | Yes | No | No | High |
| Loyola-Sánchez *et al*[22] | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Low |
| Ulus *et al*[23] | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | High |

[1]Overall risk of bias for each study assessed as Low if has a Yes answer for ALL the items or as High if has NO or Unclear answer for one or more items.

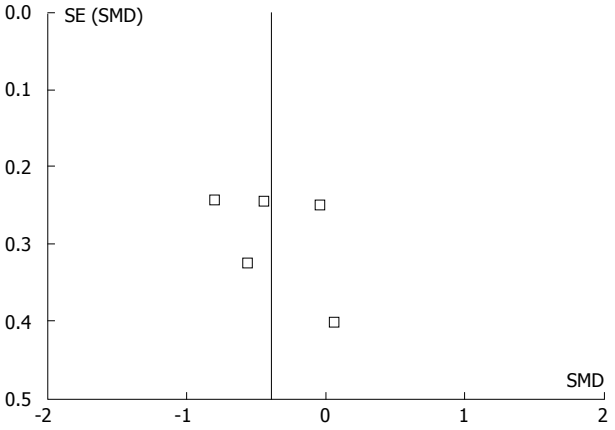



**Figure 2 Funnel plot illustrating the statistical precision plotted as a function of the standardized mean difference for the effect of therapeutic ultrasound on patient-reported knee pain.** Statistical inferences are not possible when fewer than 10 trials are available; therefore the pseudo 95% confidence limits around the summary treatment effect are not shown. SMD: Standardized mean difference.

in the meta-analysis) was deemed to have other potential sources of bias due to randomization of more than one knee from the same participant. Based on our criteria for judging the overall methodological quality assessment for each trial, risk of bias was low in one trial and high in five trials (Table 4). Therefore, risk of bias across the studies is high. Slight asymmetry is noted in the funnel plots for each of the outcomes as illustrated in Figure 2 for the pain outcome.

### *Effects of the intervention*

In the five trials included in the analysis assessing efficacy of therapeutic ultrasound on knee OA pain, a total of 281 participants were randomised: 155 to an active ultrasound group and 126 to a sham ultrasound group[18-20,22,23]. Figure 3A illustrates that therapeutic ultrasound (pulsed or continuous) was effective in reducing pain compared to sham ultrasound [SMD = -0.39 (-0.70, -0.08); $P = 0.01$]. The proportion of the dispersion in effect estimates across studies fell within that expected due to sampling error ($I^2 = 36\%$) and this low heterogeneity is unlikely to represent a true difference in effects in the studies ($\chi^2 = 6.26$, $P = 0.18$). Subgroup analyses based on dosage planned *a priori* demonstrated that low dose therapeutic ultrasound resulted in a significant decrease in pain [3 trials; SMD = -0.45 (-0.85, -0.05); $P = 0.03$; Figure 3B] and reduced the heterogeneity further ($\chi^2 = 2.82$, $P = 0.24$; $I^2 = 29\%$). In contrast, pooling the three studies administering high dose ultrasound *vs* sham ultrasound yielded a statistically insignificant decrease in pain and, although still acceptable, the heterogeneity increased [3 trials, SMD = -0.46 (-0.94, 0.03); $P = 0.06$; $\chi^2 = 4.49$, $P = 0.11$; $I^2 = 55\%$]. We were unable to perform subgroup analyses based on OA severity because all five either included participants with mild to moderate radiographic knee OA or did not report radiographic OA severity. Meta-regression analyses were not performed due to the small number of trials available for subgroup comparisons.

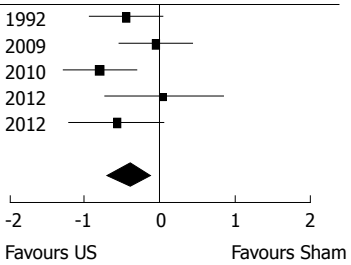
Three trials (130 participants and OA knees) assessed the effect of therapeutic ultrasound on self-reported physical function. Figure 3C illustrates that the point estimate for self-reported physical function in people with knee OA favours therapeutic ultrasound, however, the 95%CI crosses the point of no difference (SMD= -0.21, 95%CI: -0.55-0.14, $P = 0.24$). Homogeneity for pooling the data across studies was acceptable ($\chi^2 = 1.92$, $P = 0.24$, $I^2 = 0\%$). Due to the small numbers of studies that reported self-reported physical function outcomes, we could not perform meta-regression analyses.

Four trials assessed the effect of therapeutic ultrasound on physical function with respect to walking performance in people with knee OA. Figure 3D illustrates that walking performance in the group who received active ultrasound was not significantly different from that in the group who received sham ultrasound (SMD= -0.11, 95%CI: -0.59-0.37, $P = 0.65$). In this analyses, the dispersion of the estimated effects exceeded that expected due to sampling error ($\chi^2 = 8.37$, $P = 0.04$, $I^2 = 64\%$). Figure 3E illustrates that only the subgroup analysis of three trials administering high dose therapeutic ultrasound *vs* sham reduced the heterogeneity ($I^2 = 28\%$). The sample size was very small and the small effect estimate was not statistically significant (SMD = 0.04, 95%CI: -0.33-0.41, $P = 0.82$). Of the four trials available for meta-analysis, all included participants with mild to moderate knee OA. Meta-regression analyses planned *a priori* were not performed due to the small number of trials available for subgroup comparisons.

**A**

| Study or Subgroup | US Therapy Mean | SD | Total | Sham Mean | SD | Total | Weight | Std. Mean difference IV, Random, 95%CI | Year | Std. Mean difference IV, Random, 95%CI |
|---|---|---|---|---|---|---|---|---|---|---|
| Falconer | 2.6 | 2.91 | 34 | 3.9 | 2.95 | 35 | 23.9% | -0.44 [-0.92, 0.44] | 1992 | |
| Ozgonenel | 3.9 | 2 | 34 | 4 | 2.6 | 31 | 23.4% | -0.04 [-0.53, 0.44] | 2009 | |
| Tascioglu | 5.24 | 1.79 | 55 | 6.67 | 1.78 | 27 | 23.9% | -0.79 [-1.27, -0.32] | 2010 | |
| Loyola Sanchez | 1.58 | 2.02 | 12 | 1.46 | 2.02 | 13 | 12.1% | 0.06 [-0.73, 0.84] | 2012 | |
| Ulus | 3.3 | 1.83 | 20 | 4.45 | 2.16 | 20 | 16.7% | -0.56 [-1.20, 0.07] | 2012 | |
| Total (95%CI) | | | 155 | | | 126 | 100.0% | -0.39[-0.70, -0.08] | | |

Hetergeneity: Tau$^2$ = 0.04; $\chi^2$ = 6.26, df = 4 ($P$ = 0.18); $I^2$ = 36%
Test for overall effect: $Z$ = 2.49 ($P$ = 0.01)



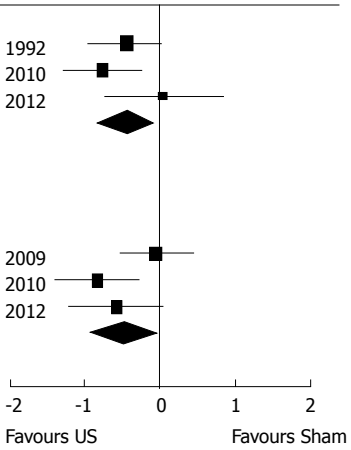

**B**

| Study or Subgroup | US Therapy Mean | SD | Total | Sham Mean | SD | Total | Weight | Std. Mean difference IV, Random, 95%CI | Year | Std. Mean difference IV, Random, 95%CI |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.2.1 US dose < 150 J/cm$^2$ | | | | | | | | | | |
| Falconer | 2.6 | 2.91 | 34 | 3.9 | 2.95 | 35 | 43.0% | -0.44 [-0.92, 0.04] | 1992 | |
| Tascioglu | 5.25 | 1.9 | 28 | 6.67 | 1.78 | 27 | 36.0% | -0.76 [-1.31, -0.21] | 2010 | |
| Loyola Sanchez | 1.58 | 2.02 | 12 | 1.46 | 2.02 | 13 | 21.0% | 0.06 [-0.73, 0.84] | 2012 | |
| Subtotal (95%CI) | | | 74 | | | 75 | 100.0% | -0.45[-0.85, -0.05] | | |

Hetergeneity: Tau$^2$ = 0.04; $\chi^2$ = 2.82, df = 2 ($P$ = 0.24); $I^2$ = 29%
Test for overall effect: Z = 2.21 ($P$ = 0.03)

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.2.2 US dose ⩾ 150 J/cm$^2$ | | | | | | | | | | |
| Ozgonenel | 3.9 | 2 | 34 | 4 | 2.6 | 31 | 37.2% | -0.04 [-0.53, 0.44] | 2009 | |
| Tascioglu | 5.22 | 1.7 | 27 | 6.67 | 1.78 | 27 | 33.3% | -0.82 [-1.38, -0.26] | 2010 | |
| Ulus | 3.3 | 1.83 | 20 | 4.45 | 2.16 | 20 | 29.5% | 0.56 [-1.20, 0.07] | 2012 | |
| Subtotal (95%CI) | | | 81 | | | 78 | 100.0% | -0.46 [-0.94, 0.03] | | |

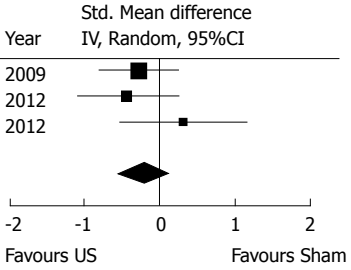
Hetergeneity: Tau$^2$ = 0.10; $\chi^2$ = 4.49, df = 2 ($P$ = 0.11); $I^2$ = 55%
Test for overall effect: $Z$ = 1.85 ($P$ = 0.06)




**C**

| Study or Subgroup | US Therapy Mean | SD | Total | Sham Mean | SD | Total | Weight | Std. Mean difference IV, Random, 95%CI | Year | Std. Mean difference IV, Random, 95%CI |
|---|---|---|---|---|---|---|---|---|---|---|
| Ozgonenel | 23.6 | 11.6 | 34 | 27.1 | 14 | 31 | 50.2% | -0.27 [-0.76, 0.22] | 2009 | |
| Tascioglu | 28.65 | 10.62 | 20 | 32.45 | 7.33 | 20 | 30.5% | -0.41 [-1.04, 0.22] | 2012 | |
| Loyola Sanchez | 23.92 | 11.3 | 12 | 20.38 | 13 | 13 | 19.3% | 0.28 [-0.51, 1.07] | 2012 | |
| Total (95%CI) | | | 66 | | | 64 | 100.0% | -0.21[-0.55, 0.14] | | |

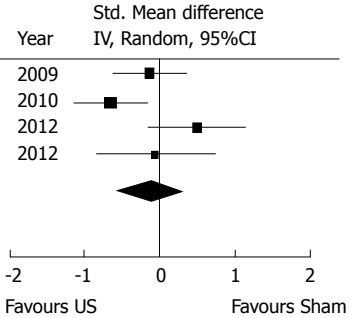
Hetergeneity: Tau$^2$ = 0.00; $\chi^2$ = 1.92, df = 2 ($P$ = 0.38); $I^2$ = 0%
Test for overall effect: $Z$ = 1.17 ($P$ = 0.24)




**D**

| Study or Subgroup | US Therapy Mean | SD | Total | Sham Mean | SD | Total | Weight | Std. Mean difference IV, Random, 95%CI | Year | Std. Mean difference IV, Random, 95%CI |
|---|---|---|---|---|---|---|---|---|---|---|
| Ozgonenel | 35.5 | 6.7 | 34 | 36.4 | 7.6 | 31 | 28.3% | -0.12 [-0.61, 0.36] | 2009 | |
| Tascioglu | 21.4 | 2.87 | 55 | 23.19 | 2.54 | 27 | 28.8% | -0.64 [-1.11, -0.17] | 2010 | |
| Ulus | 68.05 | 44.53 | 20 | 51.25 | 15.74 | 20 | 23.6% | 0.49 [-0.14, 1.12] | 2012 | |
| Loyola Sanchez | -414.55 | 68.53 | 12 | -411.85 | 89.8 | 12 | 19.3% | -0.03 [-0.82, 0.75] | 2012 | |
| Total (95%CI) | | | 121 | | | 91 | 100.0% | -0.11 [-0.59, 0.37] | | |

Hetergeneity: Tau$^2$ = 0.15; $\chi^2$ = 8.24, df = 3 ($P$ = 0.04); $I^2$ = 64%
Test for overall effect: $Z$ = 0.45 ($P$ = 0.65)

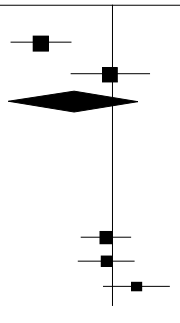



**E**

| Study or Subgroup | US Therapy Mean | SD | Total | Sham Mean | SD | Total | Weight | Std. Mean difference IV, Random, 95%CI | Year | Std. Mean difference IV, Random, 95%CI |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.5.1 US dose < 150 J/cm$^2$ | | | | | | | | | | |
| Tascioglu | 20 | 1.94 | 28 | 23.19 | 2.54 | 27 | 51.8% | -1.39 [-1.99, -0.80] | 2010 | |
| Loyola Sanchez | -414.55 | 68.53 | 12 | -411.85 | 89.8 | 13 | 48.2% | -0.03 [-0.82, 0.75] | 2012 | |
| Subtotal (95%CI) | | | 40 | | | 40 | 100.0% | -0.74 [-2.07, 0.60] | | |

Hetergeneity: Tau$^2$ = 0.80; $\chi^2$ = 7.36, df = 1 ($P$ = 0.007); $I^2$ = 86%
Test for overall effect: Z = 1.09 ($P$ = 0.28)

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.5.2 US dose ⩾ 150 J/cm$^2$ | | | | | | | | | | |
| Ozgonenel | 35.5 | 6.7 | 34 | 36.4 | 7.6 | 31 | 38.9% | -0.12 [-0.61, 0.36] | 2009 | |
| Tascioglu | 22.85 | 2.99 | 27 | 23.19 | 2.54 | 27 | 34.3% | -0.12 [-0.65, 0.41] | 2010 | |
| Ulus | 68.05 | 44.53 | 20 | 51.25 | 15.74 | 20 | 26.8% | 0.49 [-0.14, 1.12] | 2012 | |



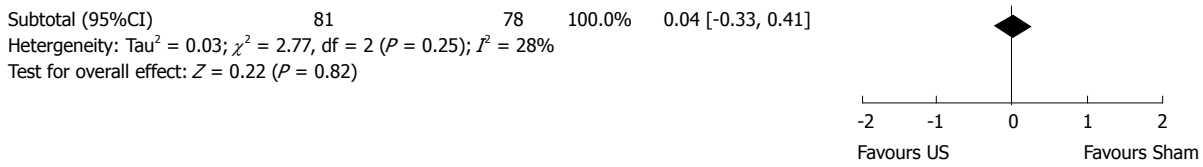



**Figure 3 Forest plot.** A: For meta-analysis of the effect of ultrasound (US) compared to sham on pain; B: For trials reporting the pain outcome subgrouped according to US dose; C: For meta-analysis of the effect of US compared to sham treatment on self-reported physical function; D: For meta-analysis of the effect of US compared to sham treatment on walking performance; E: For trials reporting a walking performance outcome subgrouped according to US dose.

### *Adverse events*

Two trials reported that no major complaints were reported[18,23]. Two trials reported no side effects although in one trial[19] two participants in the sham ultrasound group had to use analgesics and were dropped from the study and in another trial[20] six out of 90 participants (one in the active pulsed ultrasound group, two in the active continuous ultrasound group and three in the sham ultrasound group) had to use nonsteroidal anti-inflammatory drugs or analgeiscs because of intolerable pain and were withdrawn from the study. In one trial[21], adverse effects included dizziness ($n$ = 3), mental stress, palpitation or fatigue however it was not stated if these participants received active or sham ultrasound. In the triple blind trial[22], two participants reported a "stinging" sensation during active ultrasound treatment and two participants receiving sham ultrasound reported a similar 'stinging' sensation during treatment.

### *Strength of the body of evidence*

The strength of the body of evidence for pain reduction and improvement in self-reported physical function and walking performance is summarized in Table 5. The finding that therapeutic ultrasound reduces pain is based on low-strength evidence (Table 5). For the self-reported physical function outcome, the mean difference in score is shown in Table 5 since all three of the trials provided data for the WOMAC LK3.1 physical function subscale that permitted pooling. The finding that therapeutic ultrasound does not improve self-reported physical function is based on very low strength evidence (Table 5). We report the findings for walking performance obtained when we pooled all four trials that assessed this outcome although our criteria for pooling data were not met ($P$ = 0.04 for $\chi^2$ and $I^2$ = 64%). Acceptable homogeneity was achieved when we meta-analyzed the 3 trials (159 participants and knees) applying high dose ultrasound ($\chi^2$ = 2.77, $P$ = 0.25, $I^2$ = 28%). We did not report the results of this subgroup analysis in the Summary of Findings Table because we felt the very small sample size contributed to the imprecision of the effect estimate for walking performance to a similar degree as the inconsistency thus the strength of the body of evidence could not be upgraded.

## DISCUSSION

This systematic review provides a meta-analysis of the efficacy of therapeutic ultrasound *vs* sham ultrasound for decreasing pain and improving physical function in people with knee OA who were blinded to treatment allocation. The main finding is that therapeutic ultrasound treatment decreases pain but does not improve physical function significantly in this patient population. Our results confirm those of recent meta-analyses[8-10]. Confidence in the treatment effect estimates is increased by including only those studies in which participants were blinded to the treatment they received[11]. Whereas we did remove an important source of subjective assessment bias, the strength of the evidence for pain and physical function outcomes is low and very low, respectively. Few trial participants from either the active or sham ultrasound group reported adverse events or side effects. These findings demonstrate that therapeutic ultrasound is a safe and effective treatment for pain in people with knee OA and further research is needed to determine if physical function improves.

In our systematic review, the treatment effect estimate is lower compared to those reported previously [SMD = -0.39 (-0.70, -0.08); $P$ = 0.01] demonstrating the importance of study design characteristics of the trials included in the meta-analyses. Of the three new trails added to our meta-analysis, two had low risk of bias due to randomized sequence generation, allocation concealment and blinding of participants to this meta-analysis (Table 4). In one trial[22], the person administering the intervention was also blinded and the overall risk of bias was low. In contrast, trials published prior to 2011 had unclear risks of bias due to random sequence generation and allocation concealment and the care providers were not blinded. Given that the two outcomes of interest in our meta-analyses are self-reported, the estimates of treatment effect and between-trial heterogeneity may still be inflated[11] since these studies contribute 216 participants/OA knees to the total sample of 281 participants and OA knees. Our sample was too small to determine whether ultrasound dose influences the treatment effect estimates and all participants had mild to moderate OA so the influence of radiographic OA severity could not be investigated. Nevertheless, our results confirm the findings of previous meta-analyses[8-10] reporting that therapeutic ultrasound reduces knee OA pain and further research will clarify if there is a beneficial treatment effect with respect to physical function outcomes.

While recently published trials have ensured adequate

**Table 5 Summary of findings**

| Outcomes | Difference[1] in ultrasound group mean relative to the control group mean (95%CI) | No of Participants and knees (studies) | Strength of the body of evidence[2] | Inconsistency ($I^2$) | Outcome specific risk of bias |
|---|---|---|---|---|---|
| Pain<br>VAS; NRS<br>Follow-up: 2-8 wk | 0.39 standard deviations lower [-0.70-(-0.08)] | 281 (5 studies) | Low | 36% | High risk of bias of the included studies, imprecision due to small sample size and wide CI |
| Self-reported physical function<br>WOMAC® LK 3.1 Physical function; Follow-up: 2-8 wk | 2.49 points lower (-0.55-0.14) | 130 (3 studies) | Very low | 0% | High risk of bias of the included studies, imprecision due to very small sample size and wide CI |
| Walking performance<br>50 m walk speed (s); 20 m walk speed (s); 6MWT (m)<br>Follow-up: 2-8 wks | 0.11 standard deviations lower (-0.59-0.37) | 212 (4 studies) | Very low | 64% | High risk of bias in the included studies, imprecision due to small sample size and wide CI, inconsistent |

[1]Standardized mean difference reported for pain and walking performance; Mean difference reported for self-reported physical function outcomes; [2]GRADE Working Group grades of evidence, High quality: Further research is very unlikely to change our confidence in the estimate of effect. Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Very low quality: We are very uncertain about the estimate. VAS: Horizontal 10 cm visual analogue scale for pain (maximum score = 10, lower score is better); NRS: Pain with movement numeric rating scale (maximum score = 10, lower score is better); WOMAC LK 3.1 Physical function: Likert Scale version of the Western Ontario and McMaster Osteoarthritis Index physical function subscale (maximum score = 80, lower score is better); 6MWT: 6 min walk test (distance measured in metres, higher score is better).

blinding of the participants, the trials included in the current meta-analysis lack methodological rigour in terms of registering the protocol, assuring that risk of bias is low, and recruiting a sample size large enough to ensure precise estimates of efficacy. One trial[22] addressed the feasibility of recruitment and describes the burden of attending a centre for a passive treatment unproven in terms of efficacy as a barrier to participation and retention. Given the growing prevalence of knee OA, strategies to increase sample size and minimize loss to follow up is required in order to conduct a definitive high quality trial addressing this clinical question. If high quality evidence confirms that ultrasound is an effective conservative treatment for knee OA, the wide availability of the treatment ensures rapid translation of this evidence into clinical practice[6,7].

Apart from the walking performance outcome, between-trial homogeneity suggests that the interventions provided in the five trials were comparable. However, subgroup analyses pooling only the trials administering low dose ultrasound reduced the heterogeneity in the pain outcome from 36% to 29%. Table 3 highlights other differences in the interventions administered. Data from two trials that included co-interventions were pooled with data from three trials that did not describe co-interventions. Co-interventions included exercise in both trials[18,23] and the use of analgesics (except during physiotherapy) in one[23] of these trials. These treatments have proven efficacy in the treatment of knee OA and, clinical practice guidelines recommend multicomponent physical therapy for this population[3,4]. However, the interactions between these co-interventions and ultrasound therapy are not known. Another source of variability in the intervention was the application site. Of the five trials included in the meta-analyses, three trials[19,20,23] administered therapeutic ultrasound using a moving sound head applied to the anteromedial and lateral parts of the knee; one trial[18] used a moving sound head applied to the soft tissue around the knee which limited joint range of motion the most; one trial[22] used a stationary sound head applied to the anteromedial tibiofemoral joint line. Despite these differences, the data for pain and self-reported physical function outcomes had acceptable homogeneity for pooling. Taken together, these interventions were effective in reducing knee OA pain. However, further research is required to identify the critical components before a specific protocol for administering therapeutic ultrasound can be recommended.

We chose to focus this review on published trials that compared active and sham ultrasound in order to identify and synthesize the highest quality evidence available in the English-language literature. The funnel plots may be interpreted as evidence of publication bias however asymmetry is attenuated due to the small number of trials and the similar small sample size in each trial. We expected publication bias and therefore we did not plan to perform statistical tests to confirm this. Given that the Cochrane Handbook recommends a minimum of 10 studies in order to run the inferential analysis[24], we could not have conducted statistical testing anyway.

Whereas our systematic review provides high level evidence that therapeutic ultrasound reduces knee OA pain and may improve physical function based on trials in which participants did not know if they received active or sham ultrasound, this review was not without limitations. Random sequence generation and allocation concealment were inadequately described in three of the five meta-analyzed trials. However, it has been suggested that the selection bias which may be present as a result of these study design characteristics are most problematic when the participants' prognosis is easy to assess at the time of

randomization[11]. Since the progression of clinical symptoms in chronic knee OA is difficult to predict, we believe that the influence of selection bias on the effect estimates in the current review is minimal. Inadequate design and/or reporting of trial methods in the trials included in our systematic review limit the strength of the body of evidence which increases the uncertainty in the results. Finally, we chose to focus this review on double blind trials published in English; we did not request data for the three trials published only as abstracts in conference proceedings nor did we search trial registries to identify unpublished data. Furthermore, we were unsuccessful in our attempts to contact two of the primary authors to clarify trial methodology and secure study data suitable for pooling. The inclusion of these unpublished data may have yielded different results.

In conclusion, our meta-analysis suggests that therapeutic ultrasound decreases pain in people with knee OA; however, it is very likely that this conclusion will change when more research is conducted. Therapeutic ultrasound appears to be no better than sham ultrasound for improving self-reported physical function or walking performance. However, the very low strength of the body of evidence for the physical function outcomes leaves us very uncertain about these estimates and further research is necessary to answer this question.

## ACKNOWLEDGEMENTS

We thank Heather N deBoer for her assistance with updating the literature search. This review protocol was designed and initiated while NJ MacIntyre was completing the Canadian Institutes of Health Research Randomized Controlled Trials Mentorship Program under the mentorship of M. Bhandari.

## COMMENTS

### Background

Knee osteoarthritis (OA) is a common problem resulting in pain and mobility limitations in people around the globe. Often symptoms progress to the point where joint replacement is required. Interventions are required that will delay or alleviate the need for surgical interventions. Multiple non-surgical interventions are used to treat patient-reported symptoms of pain and activity limitations and the efficacy of a single treatment component is difficult to evaluate.

### Research frontiers

Past systematic reviews evaluating the efficacy of therapeutic ultrasound which pooled data from studies in which participants did not know if there received active ultrasound therapy (blinded) or received the same interventions with the exception of ultrasound therapy (not blinded) suggest ultrasound is effective in reducing knee OA pain whereas reductions in physical function are not statistically significant. Efficacy of ultrasound with respect to patient-reported pain and physical function outcomes should be evaluated in randomized controlled trials in which participants are blinded.

### Innovations and breakthroughs

Based on this meta-analysis, including only participants who were blinded to treatment allocation, knee OA pain decreases but physical function is not significantly improved following a course of therapeutic ultrasound treatment. These findings strengthen our confidence in similar findings reported in previous systematic reviews.

### Applications

Therapeutic ultrasound appears to be an effective treatment to include in the multi-component non-surgical management of the important clinical problem of knee OA pain. Physical function was not statistically improved immediately following a course of therapeutic ultrasound treatment. Large, methodologically rigorous randomized trials in which the participants, health care providers, and assessors are blinded to treatment allocation and followed up over a longer period are needed to provide a clear answer to this clinical question.

### Terminology

Knee osteoarthritis is characterized by degenerative changes to cartilage and other tissues in and around the joint. Therapeutic ultrasound delivers high frequency sound waves and is a physical agent commonly applied by physical therapists to reduce inflammation and/or enhance tissue healing. These effects are postulated to result in improvements in patient-reported pain and physical function (self-reported or performance-based).

### Peer review

This meta-analysis was used to determine the efficacy of therapeutic ultrasound versus sham for improving pain and physical function immediately post-intervention in people with knee OA. This is an interesting meta-analysis.

## REFERENCES

1 **Guccione AA**, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, Kelly-Hayes M, Wolf PA, Kreger BE, Kannel WB. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. *Am J Public Health* 1994; **84**: 351-358 [PMID: 8129049 DOI: 10.2105/AJPH.84.3.351]

2 **Felson DT**. An update on the pathogenesis and epidemiology of osteoarthritis. *Radiol Clin North Am* 2004; **42**: 1-9, v [PMID: 15049520 DOI: 10.1016/S0033-8389(03)00161-1]

3 **Hochberg MC**, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, Towheed T, Welch V, Wells G, Tugwell P. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis Care Res* (Hoboken) 2012; **64**: 465-474 [PMID: 22563589 DOI: 10.1186/1471-2474-14-320]

4 **Richmond J**, Hunter D, Irrgang J, Jones MH, Snyder-Mackler L, Van Durme D, Rubin C, Matzkin EG, Marx RG, Levy BA, Watters WC, Goldberg MJ, Keith M, Haralson RH, Turkelson CM, Wies JL, Anderson S, Boyer K, Sluka P, St Andre J, McGowan R. American Academy of Orthopaedic Surgeons clinical practice guideline on the treatment of osteoarthritis (OA) of the knee. *J Bone Joint Surg Am* 2010; **92**: 990-993 [PMID: 20360527 DOI: 10.2106/JBJS.I.00982]

5 **Bélanger AY.** Ultrasound. Evidence-based guide to therapeutic physical agents. In: Philadelphia PA, editor. USA: Lippincott Williams & Wilkins, 2003: 223-261

6 **Nussbaum EL**, Burke S, Johnstone L, Lahiffe G, Robitaille E, Yoshida K. Use of electrophysical agents: Findings and implications of a survey of practice in metro Toronto. *Physiother Can* 2007; **59:** 118-131 [DOI: 10.3138/ptc.59.2.118]

7 **MacIntyre NJ**, Busse JW, Bhandari M. Physical therapists in primary care are interested in high quality evidence regarding efficacy of therapeutic ultrasound for knee osteoarthritis: a provincial survey. *ScientificWorldJournal* 2013; **2013**: 348014 [PMID: 23844391 DOI: 10.1155.2013/348014]

8 **Loyola-Sánchez A**, Richardson J, MacIntyre NJ. Efficacy of ultrasound therapy for the management of knee osteoarthritis: a systematic review with meta-analysis. *Osteoarthritis Cartilage* 2010; **18**: 1117-1126 [PMID: 20637297 DOI: 10.1016/j.joca.2010.06.010]

9 **Rutjes AW**, Nüesch E, Sterchi R, Jüni P. Therapeutic ultrasound for osteoarthritis of the knee or hip. *Cochrane Database Syst Rev* 2010; **(1)**: CD003132 [PMID: 20091539 DOI: 10.1002/14651858.CD003132.pub2]

10 **Wang SY**, Olson-Kellogg B, Shamliyan TA, Choi JY, Ramakrishnan R, Kane RL. Physical therapy interventions for knee pain secondary to osteoarthritis: a systematic review. *Ann Intern Med* 2012; **157**: 632-644 [PMID: 23128863 DOI: 10.7326/0003-4819-157-9-201211060-00007]

11 **Savović J**, Jones H, Altman D, Harris R, Jűni P, Pildal J, Als-Nielsen B, Balk E, Gluud C, Gluud L, Ioannidis J, Schulz K, Beynon R, Welton N, Wood L, Moher D, Deeks J, Sterne J. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technol Assess* 2012; **16**: 1-82 [PMID: 22989478 DOI: 10.3310/hta16350]

12 **Liberati A**, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009; **62**: e1-34 [PMID: 19631507 DOI: 10.1016/j.jclinepi.2009.06.006]

13 **Higgins JPT**, Deeks JJ. Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: URL: http://www.cochrane-handbook.org

14 **Higgins JPT**, Altman DG. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: URL: http://www.cochrane-handbook.org

15 **Guyatt GH**, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; **336**: 924-926 [PMID: 18436948 DOI: 10.1136/bmj.39489.470347.AD]

16 **Deeks JJ,** Higgins JPT, Altman DG. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: URL: http://www.cochrane-handbook.org

17 **Gurkan I**, Ranganathan A, Yang X, Horton WE, Todman M, Huckle J, Pleshko N, Spencer RG. Modification of osteoarthritis in the guinea pig with pulsed low-intensity ultrasound treatment. *Osteoarthritis Cartilage* 2010; **18**: 724-733 [PMID: 20175971 DOI: 10.1016/j.joca.2010.01.006]

18 **Falconer J**, Hayes KW, Chang RW. Effect of ultrasound on mobility in osteoarthritis of the knee. A randomized clinical trial. *Arthritis Care Res* 1992; **5**: 29-35 [PMID: 1581369]

19 **Ozgönenel L**, Aytekin E, Durmuşoglu G. A double-blind trial of clinical effects of therapeutic ultrasound in knee osteoarthritis. *Ultrasound Med Biol* 2009; **35**: 44-49 [PMID: 18829151 DOI: 10.1016/j.ultrasmedbio.2008.07.009]

20 **Tascioglu F**, Kuzgun S, Armagan O, Ogutler G. Short-term effectiveness of ultrasound therapy in knee osteoarthritis. *J Int Med Res* 2010; **38**: 1233-1242 [PMID: 20925995 DOI: 10.1177/147323001003800404]

21 **Yang PF**, Li D, Zhang SM, Wu Q, Tang J, Huang LK, Liu W, Xu XD, Chen SR. Efficacy of ultrasound in the treatment of osteoarthritis of the knee. *Orthop Surg* 2011; **3**: 181-187 [PMID: 22009649]

22 **Loyola-Sánchez A**, Richardson J, Beattie KA, Otero-Fuentes C, Adachi JD, MacIntyre NJ. Effect of low-intensity pulsed ultrasound on the cartilage repair in people with mild to moderate knee osteoarthritis: a double-blinded, randomized, placebo-controlled pilot study. *Arch Phys Med Rehabil* 2012; **93**: 35-42 [PMID: 22200383 DOI: 10.1155/2013/348014]

23 **Ulus Y**, Tander B, Akyol Y, Durmus D, Buyukakıncak O, Gul U, Canturk F, Bilgici A, Kuru O. Therapeutic ultrasound versus sham ultrasound for the management of patients with knee osteoarthritis: a randomized double-blind controlled clinical study. *Int J Rheum Dis* 2012; **15**: 197-206 [PMID: 22462424 DOI: 10.1111/j.1756-185X.2012.01709.x]

24 **Sterne JAC**, Egger M, Moher, D. Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: URL: http://www.cochrane-handbook.org


**P- Reviewer:** Dai SM, Guo X **S- Editor:** Wen LL **L- Editor**: A **E- Editor:** Lu YJ



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i3.91

*World J Meta-Anal* 2014 August 26; 2(3): 91-97
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Role of statin on mortality outcome in pneumonia patients: A meta-analysis

Lekha Saha, Narinder Kumar, Pratibha Khosla, Sharonjeet Kaur

Lekha Saha, Pratibha Khosla, Sharonjeet Kaur, Department of Pharmacology, Post Graduate Institute of Medical Education and Research, Chandigarh 160012, India
Narinder Kumar, Department of Statistics, Panjab University, Chandigarh 160014, India
Author contributions: Saha L, Khosla P and Kaur S were equally involved in planning, data extraction, calculating and writing the manuscript; Kumar N is involved in data extraction, statistical analysis and writing of the manuscript.
Correspondence to: Dr. Lekha Saha, Assistant Professor, Department of Pharmacology, Post Graduate Institute of Medical Education and Research, Sector 12, Chandigarh 160012, India. lekhasaha@rediffmail.com
Telephone: +91-172-2645262 Fax: +91-172-2744401
Received: March 11, 2014 Revised: May 23, 2014
Accepted: June 18, 2014
Published online: August 26, 2014

## Abstract

**AIM:** To determine the role of statin on mortality outcome in patient with pneumonia.

**METHODS:** For the present meta-analysis, we search the published literatures online through PubMed, Embase, Scopus and the Cochrane Library databases and the search words used were "statins'", "bacteraemia", "pneumonia", and "ICU infections". During the online search our focus was on full text articles, peer-reviewed, observational cohort or case control studies and randomized controlled trials. Those studies were selected whose outcome was hospital mortality among patients with pneumonia whether or not on statins. In this meta-analysis, 30 d mortality was used as the primary outcome as it has been demonstrated in the previous research that 30 d mortality is primarily because of community acquired pneumonia. As all studies were observational, where statin users were compared with historical rather than randomized controls, odds ratio for in-hospital or all-cause 30 d mortality was used as the primary effect measure used in the meta-analysis.

**RESULTS:** We came across the total 25 studies comprising 35355 patients (2734 statin users and 32621 statin non-users) during the electronic search. Four studies out of 25 were included in the final analysis. In this meta-analysis, when data regarding the use of statin in pneumonia patients on mortality was pooled, its results showed the non-significant effect of the statin on mortality outcome.

**CONCLUSION:** Although statins seems to be useful in the treatment of pneumonia patients but for statistical conclusion, further randomized controlled trials needs to be done or their results still waited to be published of ongoing trials, with the conclusion that presently statins showing no clinical benefit in the pneumonia patients.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Mortality; Pneumonia; Statin; Meta-analysis

**Core tip:** The relation between statin treatment and improvement in outcomes in patients with pneumonia have been reported in published literatures. This is believed to be due to anti-inflammatory and immunoregulatory effects rather than an effect on cholesterol metabolism. Recently, however, the potential benefit of statins has been called into question and the study has suggested that there may be confounding factor responsible for the observed benefit. All the studies were observational. In this present meta-analysis of randomized controlled trials, results showed the non-significant effect of the statin on mortality outcome.

Saha L, Kumar N, Khosla P, Kaur S. Role of statin on mortality outcome in pneumonia patients: A meta-analysis. *World J Meta-Anal* 2014; 2(3): 91-97 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i3/91.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i3.91

## INTRODUCTION

Statins, an inhibitor of 3-hydroxy-3 methylglutaryl coenzyme A reductase (HMG CoA reductase), was used to reduce total cholesterol, low-density lipoprotein cholesterol, apolipoprotein B and triglyceride levels[1-5]. Statins are effective in reducing the cardiovascular mortality and morbidity in patients with or without coronary artery disease. Apart from its lipid lowering effect, statins also possess multiple pleiotropic effects like anti-inflammatory/immunomodulatory effects. These immunomodulatory effects appear to be mediated through the blockade of mevalonate synthesis[6]. The positive impact observed in the cardiovascular disorders is due to their antioxidant effect, plaque stability, favorable coagulation profile, normalization of sympathetic outflow and immunomodulatory effect[7]. By virtue of their pleiotropic effect they have got a role to play in various other disorders like sepsis, dementia, osteoporosis, bacteremia, venous thromboembolism[6-9].

There is an increased rate of morbidity and mortality in hospitalized Patients with pneumonia[10]. Antimicrobial therapy is the mainstay of treatment for pneumonia. Irrespective of the efforts towards more timely treatment of pneumonia and advances in antimicrobial therapy, mortality rates for patients hospitalized with bacterial pneumonia remains relatively unchanged[11]. Studies published in the literatures have demonstrated a link between statin treatment and improvements in patients with bacterial pneumonia. Several large retrospective observational studies have shown a reduced incidence of pneumonia and improvement in outcomes in pneumonia patients taking statins[12-14]. This is believed to be due to anti-inflammatory and immunoregulatory effects rather than an effect on cholesterol metabolism. Statin drugs are known to reduce cytokine levels in stable patients with coronary artery disease and hypercholesterolemia[15-17]. Excess pro-inflammatory cytokines are found to be associated with severity of pneumonia and the development of complications such as acute respiratory distress syndrome[18-22]. Agents with anti-inflammatory properties such as statins might therefore have the potential to improve pneumonia outcomes. Recently, however, the potential benefit of statins has been called into question and the study has suggested that the confounding factors might be the reason for the observed benefit[23]. A series of studies with contradictory results have been published in the literature that has made clinicians hopeful but confused. With this background, we planned to do a meta-analysis of studies using statins in patients with pneumonia to study the role of statin use in the management of pneumonia.

## MATERIALS AND METHODS

### Literature search

To get the studies published in the literature, we search the various databases online like PubMed, Embase, Scopus and the Cochrane Library. The search words used were "statins", "bacteraemia", "pneumonia"', and "ICU infections". The type of studies looked for were peer-reviewed, full text, observational cohort or case controls studies and randomized controlled trials. Those studies whose outcome were hospital mortality in patients with pneumonia, whether or not received statins. Those studies which included patients with pneumonia due to bacterial infection were included in the present meta-analysis. Those studies were excluded from the meta-analysis where the effects of statins on viral (*i.e.*, influenza and human immunodeficiency virus infection), fungal and protozoan infections were studied. Experimental and laboratory studies were also excluded from the meta-analysis. In our search, there were no time or language limits. For additional relevant information, all references from the identified articles were scanned and search. The following data were extracted and tabulated from the selected articles: study design, patient settings, type of infection, the number of patients included, primary and secondary outcomes. For randomized controlled trials (RCTs), we search online registry of RCTs through http://www.controlled-trial.com for relevant studies. The last date of our review process was July 31, 2013.

### Study selection

Those studies which were related to use of statin in pneumonia were first identified. Observational cohort studies whether prospective or retrospective or case control studies which compared 30 d mortality in pneumonia patients among statin users and statin non-users were included in this meta-analysis. We could not identify any randomized control trials.

### Statistical analysis

Patients who are taking statin at time of hospital admission were defined as current statin users. The primary outcome in the meta-analysis was 30 d mortality in patients with bacterial pneumonia because studies in the literatures have shown that 30 d mortality is primarily due to community acquired pneumonia rather than other coexisting co morbid conditions. Odds ratio (OR) for in-hospital or all-cause 30 d mortality was measured effect in this meta-analysis as all studies used for meta-analysis were observational, where statin users were compared with historical rather than randomized controls. The heterogeneity between reports was assessed by using $\chi^2$ test of heterogeneity. As significant heterogeneity was found between the reports ($P < 0.05$), we used a random effects model to calculate weighted summary (total) odds ratio and their 95% confidence intervals (CIs) to take the decision. The Mantel-Haenszel method was used to compute the total ORs and CIs. ORs with 95% confidence intervals of all the four studies and the overall odds ratio with 95%CI are depicted in the forest plot (Figure 1) and Peto plot (Figure 2).

## RESULTS

The total number of studies searched for literature was

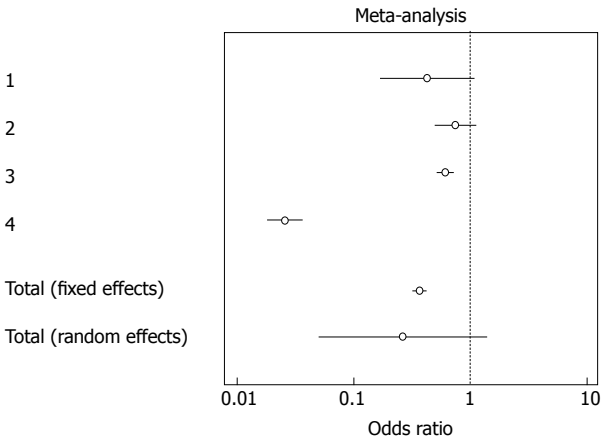



**Figure 1 Forest plot with fixed and random effects model comparing mortality among statin users and nonusers.**

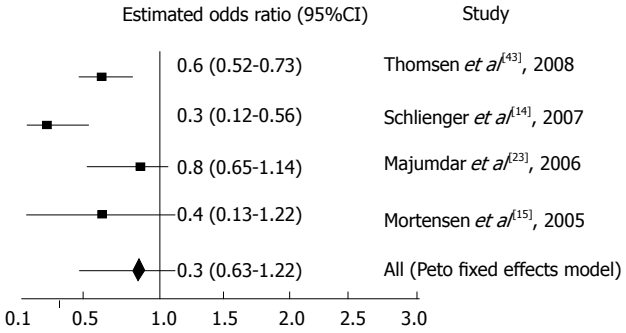



**Figure 2 Peto plot with fixed effects model comparing mortality among statin users and nonusers.**

twenty five. Among all these, two were review articles[7,24], another fifteen commented upon the infections like sepsis, bacteremia and other post operative settings[10,25-38], one trial is complete, but the results are still unpublished[39] and in an another trial, recruitment is still going on[40]. There were two studies which have evaluated the impact of statin use along with other drugs[41,42]. Four studies were actually evaluated to document the effect of statin in pneumonia[14,15,23,43] (Figure 3).

Four studies out of the total 25 studies comprising 35355 patients (2734 statin users and 32621 statin nonusers) were included in this meta-analysis and in the final analysis. The details of these 4 studies have been mentioned in Table 1. It included both the prospective and retrospective observational studies. All studies were published from 2005 to 2008 (Table 2). In this meta-analysis, when data regarding the use of statin in pneumonia patients on mortality was pooled, its results showed the non-significant effect of the statin on mortality outcome (Figure 1).

## DISCUSSION

Studies regarding the statins use have favourably focused on their cardiovascular outcome. Regarding their pleiotropic effects, studies have been conducted which

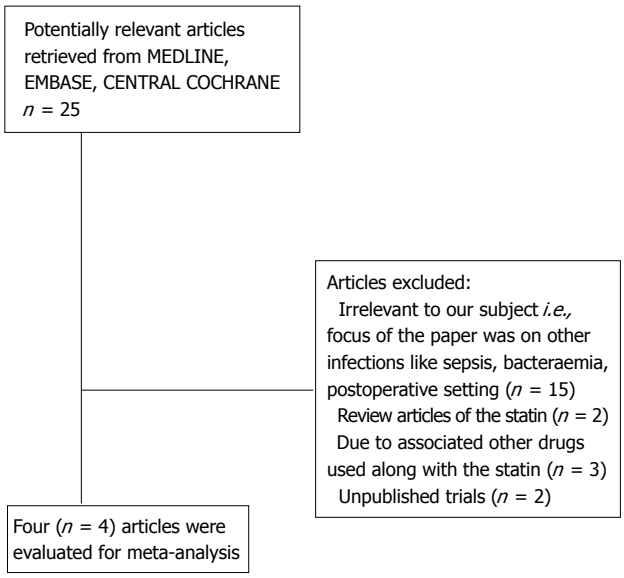



**Figure 3 Flow chart of literature search.**

have shown varying/mixed results. Different observational studies regarding statin use in pneumonia have too shown invariably mixed results. The calculated weighted summary (total) OR obtained in this meta-analysis suggests that there was no significant effect of the statin use on mortality outcome in patients with pneumonia.

The important known facts regarding the value of any meta-analysis is that it depends upon the quality of reporting of information in individual trials. If we individually comment on the results of these studies regarding their effect on mortality in the pneumonia patients, it was found that three studies reported lower mortality on pneumonia outcome[13,42,43] and one study reported the nonsignificant effect with its use[23]. The studies which we omitted are the ones in which other cardiovascular drugs have been used like aspirin, beta blockers and angiotensin converting enzyme (ACE) inhibitors, which can bias the results of the meta-analysis[42], in another study by Mortensen 2008[41], again along with stations, ACE inhibitors were added and results of starting a group and station along with ACE inhibitors were compared. The third conducted by Van de Garde[13] was excluded due to the reasons that it evaluated the risk outcome associated with the statin use on pneumonia patients rather than the mortality in them.

Chopra and Flanders[44] in their evidence based study demonstrated results favoring and against the use of statins in mortality outcome, with the ultimate conclusions that further randomized, controlled studies are needed to reach a valid conclusion[44]. Kopterides *et al*[45] also looked at the outcomes in a different infections, including the sepsis, bacteraemia and pneumonia, but without reaching the firm conclusions of clinical benefit of the statin therapy. A study conducted by Falagas *et al*[46] has also reported similar results, but no definitive conclusions could be drawn from the pooled data.

The major limitation of this meta-analysis is the in-

**Table 1 Summary of the different studies evaluating the effect of statins on mortality in pneumonia**

| Ref. | Study design | Patient's setting and study groups | Outcomes and odds ration (95%CI) | Conclusion |
|---|---|---|---|---|
| Thomsen *et al*[43], 2008 Preadmission Use of Statins and outcomes After Hospitalization With Pneumonia: Population-Based Cohort Study of 29900 Patients | Retrospective population based cohort study | 29900 adults hospitalized with pneumonia for the first time between January 1, 1997, and December 31, 2004 in northern Denmark. Data on statin and other medication use, comorbidities, socioeconomic markers, laboratory findings,bacteremia, pulmonary complications and death were obtained from medical databases. SU: 1371 NSU: 28529 | Of patients with pneumonia, 1371 (4.6%) were current statin users. Mortality among statin users was lower than among nonusers: 10.3% *vs* 15.7% after 30 d and 16.8% *vs* 22.4% after 90 d, corresponding to adjusted 30- and 90-d mortality rate ratios of 0.69 (95%CI: 0.58-0.82) and 0.75 (0.65-0.86). | The use of statins is associated with de creased mortality after hospitalization with pneumonia |
| Schlienger *et al*[14], 2007 Statins and the Risk of Pneumonia: A population-Based, Nested Case-Control Study | Population-based, retrospective, nested case-control analysis | The study population (134262 patients aged > 30 yr) consisted of 55118 patients who took statins and/or fibrates, 29144 patients with hyperlipidemia not taking lipid-lowering agents, and 50000 randomly selected patients without hyperlipidemia and without lipid-lowering treatment. Authors identified 1253 patients with pneumonia and matched them with 4838 control subjects based on age, sex, general practice, and index date.After adjusting for comorbidity and frequency of visits to general practitioners, we calculated the risks (OR with 95% confidence intervals) of uncomplicated pneumonia, hospitalization for pneumonia with survival, and fatal pneumonia in participants who used statins compared with those who did not. SU: 927 NSU: 326 | 30 d Mortality ratio 0.262 (0.182-0.377) | Significant lower mortality among statin users as compared to non users |
| Majumdar *et al*[23], 2006 Statins and outcomes in patients admitted to hospital with community acquired pneumonia: population based prospective cohort study | Prospective population based cohort study | 3415 adult patients admitted with CAP and categorised according to use of statins for at least one week before admission and during hospital stay. SU: 325 NSU: 3090 | Of 3415 patients with pneumonia admitted to hospital, 624 (18%) died or were admitted to an intensive care unit. Statin users were less likely to die or be admitted to an intensive care unit than non-users [50/325 (15%) *vs* 574/3090 (19%), OR = 0.80, *P* = 0.15]. After more complete adjustment for confounding, however, the OR changed from potential benefit (0.78, adjusted for age and sex) to potential harm (1.10, fully adjusted including propensity scores, 95%CI: 0.76-1.60) | Statins are not associated with reduced mortality or need for admission to an intensive care unit in patients with pneumonia; reports of benefit in the setting of sepsis may be a result of confounding. |
| Mortensen *et al*[15], 2005 The effect of prior statin use on 30-d mortality for patients hospitalized with community-acquired pneumonia | Retrospective cohort study | A retrospective cohort study conducted at two tertiary teaching hospitals. Eligible subjects were admitted with a diagnosis of, had a chest x-ray consistent with, and had a discharge ICD-9 diagnosis of pneumonia. Subjects were excluded if they were "comfort measures only" or transferred from another acute care hospital. Subjects were considered to be on a medication if they were taking it at the time of presentation. Data was abstracted on 787 subjects at the two hospitals. SU: 110 NSU: 677 | In the multivariable regression analysis, after adjusting for the propensity score and processes of care, the use of statins at presentation (OR = 0.36, 95%CI: 0.14-0.92) was significantly associated with decreased 30-d mortality | Prior outpatient statin use was associated with decreased mortality in patients hospitalized with community-acquired pneumonia despite their use being associated with comorbid illnesses likely to contribute to increased mortality |

SU: Statin users; NSU: Nonstatin users.

**Table 2 Individual and total evaluation number of statin users and nonusers on the mortality in pneumonia in various studies**

| Ref. | Mortality in SU | Mortality in NSU | OR | 95%CI |
|---|---|---|---|---|
| Thomsen *et al*[43] | 141/1372 | 4489/28528 | 0.61 | 0.514-0.732 |
| Schlienger *et al*[14] | 54/927 | 229/326 | 0.262 | 0.182-0.377 |
| Majumdar *et al*[23] | 25/325 | 309/3090 | 0.75 | 0.491-1.147 |
| Mortensen *et al*[15] | 5/110 | 67/677 | 0.434 | 0.171-1.101 |
| Total (fixed effects) | 225/2734 | 5094/32621 | 0.373 | 0.322 - 0.431 |
| Total (Random effects) | 225/2734 | 5094/32621 | 0.267 | 0.0499 - 1.428 |
| Test for heterogeneity | 246.1742 | | | |
| Computed value | | | | |
| Degrees of freedom | 3 | | | |
| Significance level | $P < 0.05$ | | | |

SU: Statin users; NSU: Nonstatin users.

clusion of only observational studies and marked heterogeneity seen in these clinical studies. The heterogeneity was related to the design of the studies along with the sample size, current statin use, associated Co morbidities, and identification of the clinical infection seen in these patients. Another important issue regarding statin use is a healthy user effect, which refers to those patients receiving statins also belongs to higher socioeconomic classes than patients who do not. These patients have the advantage of higher education, better awareness regarding their health which increases the complaint regarding their visits to doctors and treatment schedule, with the results for a better outcome in case of infection.

Based on currently available clinical evidence it can be concluded that statins use shows no clinical benefit in terms of mortality outcome in the pneumonia patients. This conclusion is despite the fact the three studies reported the positive outcome regarding the statin use in the pneumonia patients as compared to the only single one reporting the negative results. As the studies included were heterogeneous, therefore the results of ongoing trials and further randomized controlled trials could only provide a definitive evidence regarding the effect of statins on mortality in pneumonia.

## COMMENTS

### Background

Statins, an 3-hydroxy-3 methylglutaryl coenzyme A reductase (HMG CoA) reductase inhibitor, used to reduce total cholesterol, low-density lipoprotein cholesterol, apolipoprotein B and triglyceride levels. Statins are also effective in reducing the cardiovascular mortality and morbidity in patients with or without coronary artery disease. Apart from its lipid lowering effect, statins also possess multiple pleiotropic effects like anti-inflammatory/immunomodulatory effects. The association between statin use and improvement in outcomes in patients with pneumonia has been reported in published literatures. This is believed to be due to anti-inflammatory and immunoregulatory effects rather than an effect on cholesterol metabolism.

### Research frontiers

Several large retrospective observational studies have shown a reduced incidence of pneumonia and improvement in outcomes in pneumonia patients taking statins. Recently, however, the potential benefit of statins has been called into question and study has suggested that the confounding factors might be the reason for the observed benefit of those studies. A series of studies with contradictory results have been published in the literature that has made clinicians hopeful but confused. With this background, the authors planned to do a meta-analysis of studies using statins in patients with pneumonia to study the role of statin use in the management of pneumonia.

### Innovations and breakthroughs

Four studies out of the total 25 studies comprising 35355 patients (2734 statin users and 32621 statin non-users) were included in this meta-analysis and in the final analysis. It included both the prospective and retrospective observational studies. All studies were published from 2005 to 2008. Odds ratio (OR) for in-hospital or all-cause 30 d mortality was measured effect in this meta-analysis as all studies used for meta-analysis were observational, where statin users were compared with historical rather than randomized controls. The heterogeneity between reports was assessed by using $\chi^2$ test of heterogeneity. As significant heterogeneity was found between the reports ($P < 0.05$), the authors used a random effects model to calculate weighted summary (total) odds ratio and their 95%CIs to take the decision. The Mantel-Haenszel method was used to compute the total OR and CIs. OR with 95%CIs of all the four studies and the overall OR with 95%CI are depicted in the forest plot and Peto plot.

### Applications

In this meta-analysis, when data regarding the use of statin in pneumonia patients on mortality was pooled, its results showed the non-significant effect of the statin on mortality outcome. As the studies included were heterogeneous, therefore the results of ongoing trials and further randomized controlled trials could only provide a definitive evidence regarding the effect of statins on mortality in pneumonia.

### Terminology

HMG CoA reductase is 3-hydroxy-3 methylglutaryl coenzyme A reductase is the rate-controlling enzyme of the mevalonate pathway, the metabolic pathway that produces cholesterol and other isoprenoids. Pleiotropic effects of statin are those properties of statins other than cholesterol lowering effects like the effect on osteoporosis, dementia and so on.

### Peer review

Authors done a meta-analysis to identify the effectiveness of statin use on mortality due to pneumonia. Authors conducted study well.

## REFERENCES


1 **Simvastatin Survival Study Group.** Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994; **344**: 1383-1389 [PMID: 7968073]

2 **Sacks FM**, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG, Brown L, Warnica JW, Arnold JM, Wun CC, Davis BR, Braunwald E. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. *N Engl J Med* 1996; **335**: 1001-1009 [PMID: 8801446]

3 **The Long-term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group.** Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cho-

lesterol levels. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. *N Engl J Med* 1998; **339**: 1349-1357 [PMID: 9841303 DOI: 10.1056/NEJM199811053391902]
4 **West of Scotland Coronary Prevention Group.** West of Scotland Coronary Prevention Study: identification of High-risk Groups and Comparison with Other Cardiovascular Intervention Trials. *Lancet* 1996; **348**: 1339-1342 [DOI: 10.1016/S0140-6736(96)04292-4]
5 **Downs JR**, Clearfield M, Weis S, Whitney E, Shapiro DR, Beere PA, Langendorfer A, Stein EA, Kruyer W, Gotto AM. Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels: results of AFCAPS/TexCAPS. Air Force/Texas Coronary Atherosclerosis Prevention Study. *JAMA* 1998; **279**: 1615-1622 [PMID: 9613910]
6 **Terblanche M**, Almog Y, Rosenson RS, Smith TS, Hackam DG. Statins and sepsis: multiple modifications at multiple levels. *Lancet Infect Dis* 2007; **7**: 358-368 [PMID: 17448939]
7 **Tandon V,** Bano G, Khajuria V, Parihar A, Gupta S. Pleiotropic effects of statins. *Indian J Pharmacol* 2005; **37**: 77-85 [DOI: 10.4103/0253-7613.15106]
8 **Waldman A**, Kritharides L. The pleiotropic effects of HMG-CoA reductase inhibitors: their role in osteoporosis and dementia. *Drugs* 2003; **63**: 139-152 [PMID: 12515562 DOI: 10.2165/00003495-200363020-00002]
9 **Chopra V**, Choksi PU, Cavusoglu E. Beyond lipid lowering: the anti-hypertensive role of statins. *Cardiovasc Drugs Ther* 2007; **21**: 161-169 [PMID: 17468937]
10 **Kruger P**, Fitzsimmons K, Cook D, Jones M, Nimmo G. Statin therapy is associated with fewer deaths in patients with bacteraemia. *Intensive Care Med* 2006; **32**: 75-79 [PMID: 16283159]
11 **Thomsen RW**, Riis A, Nørgaard M, Jacobsen J, Christensen S, McDonald CJ, Sørensen HT. Rising incidence and persistently high mortality of hospitalized pneumonia: a 10-year population-based study in Denmark. *J Intern Med* 2006; **259**: 410-417 [PMID: 16594909 DOI: 10.1111/j.1365-2796.2006.01629.x]
12 **Fry AM**, Shay DK, Holman RC, Curns AT, Anderson LJ. Trends in hospitalizations for pneumonia among persons aged 65 years or older in the United States, 1988-2002. *JAMA* 2005; **294**: 2712-2719 [PMID: 16333006 DOI: 10.1001/jama.294.21.2712.]
13 **van de Garde EM**, Hak E, Souverein PC, Hoes AW, van den Bosch JM, Leufkens HG. Statin treatment and reduced risk of pneumonia in patients with diabetes. *Thorax* 2006; **61**: 957-961 [PMID: 16809409 DOI: 10.1136/thx.2006.062885]
14 **Schlienger RG**, Fedson DS, Jick SS, Jick H, Meier CR. Statins and the risk of pneumonia: a population-based, nested case-control study. *Pharmacotherapy* 2007; **27**: 325-332 [PMID: 17316144 DOI: 10.1592/phco.27.3.325]
15 **Mortensen EM**, Restrepo MI, Anzueto A, Pugh J. The effect of prior statin use on 30-day mortality for patients hospitalized with community-acquired pneumonia. *Respir Res* 2005; **6**: 82 [PMID: 16042797 DOI: 10.1186/1465-9921-6-82]
16 **Jialal I**, Stein D, Balis D, Grundy SM, Adams-Huet B, Devaraj S. Effect of hydroxymethyl glutaryl coenzyme a reductase inhibitor therapy on high sensitive C-reactive protein levels. *Circulation* 2001; **103**: 1933-1935 [PMID: 11306519 DOI: 10.1161/01.CIR.103.15.1933]
17 **Musial J**, Undas A, Gajewski P, Jankowski M, Sydor W, Szczeklik A. Anti-inflammatory effects of simvastatin in subjects with hypercholesterolemia. *Int J Cardiol* 2001; **77**: 247-253 [PMID: 11182189]
18 **Strandberg TE**, Vanhanen H, Tikkanen MJ. Effect of statins on C-reactive protein in patients with coronary artery disease. *Lancet* 1999; **353**: 118-119 [PMID: 10023901]
19 **Almirall J**, Bolíbar I, Toran P, Pera G, Boquet X, Balanzó X, Sauca G. Contribution of C-reactive protein to the diagnosis and assessment of severity of community-acquired pneumonia. *Chest* 2004; **125**: 1335-1342 [PMID: 15078743]
20 **Fernandez-Serrano S**, Dorca J, Coromines M, Carratala J, Gudiol F, Manresa F. Molecular inflammatory responses measured in blood of patients with severe community acquired pneumonia. *Clin Diagn Lab Immunol* 2003; **10**: 813-820 [PMID: 12965910 DOI: 10.1128/CDLI.10.5.813–820.2003]
21 **Glynn P**, Coakley R, Kilgallen I, Murphy N, O'Neill S. Circulating interleukin 6 and interleukin 10 in community acquired pneumonia. *Thorax* 1999; **54**: 51-55 [PMID: 10343632]
22 **Puren AJ**, Feldman C, Savage N, Becker PJ, Smith C. Patterns of cytokine expression in community-acquired pneumonia. *Chest* 1995; **107**: 1342-1349 [PMID: 7750329 DOI: 10.1378/chest.107.5.1342]
23 **Majumdar SR**, McAlister FA, Eurich DT, Padwal RS, Marrie TJ. Statins and outcomes in patients admitted to hospital with community acquired pneumonia: population based prospective cohort study. *BMJ* 2006; **333**: 999 [PMID: 17060337 DOI: 10.1136/bmj.38992.565972.7C]
24 **Daniel Pella**, Rafael Rybar and Viola Mechirova. Pleoitropic effect of statins. *Acta Cardiol Sin* 2005; **21**: 190-198
25 **Mortensen EM**, Restrepo MI, Copeland LA, Pugh JA, Anzueto A, Cornell JE, Pugh MJ. Impact of previous statin and angiotensin II receptor blocker use on mortality in patients hospitalized with sepsis. *Pharmacotherapy* 2007; **27**: 1619-1626 [PMID: 18041882 DOI: 10.1592/phco.27.12.1619]
26 **Yang KC**, Chien JY, Tseng WK, Hsueh PR, Yu CJ, Wu CC. Statins do not improve short-term survival in an oriental population with sepsis. *Am J Emerg Med* 2007; **25**: 494-501 [PMID: 17543651]
27 **Tseng MY**, Hutchinson PJ, Czosnyka M, Richards Hugh, Pickard JD, Kirkpatrick PJ. Effects of acute pravastatin treatment on intensity of rescue therapy, length of inpatient stay, and 6-month outcome in patients after aneurysmal subarachnoid hemorrhage. *Stroke* 2007; **38**: 1545–1550 [DOI: 10.1161/STROKEAHA.109.556332]
28 **Martin CP**, Talbert RL, Burgess DS, Peters JI. Effectiveness of statins in reducing the rate of severe sepsis: a retrospective evaluation. *Pharmacotherapy* 2007; **27**: 20-26 [PMID: 17192158 DOI: 10.1592/phco.27.1.20]
29 **Gupta R**, Plantinga LC, Fink NE, Melamed ML, Coresh J, Fox CS, Levin NW, Powe NR. Statin use and sepsis events [corrected] in patients with chronic kidney disease. *JAMA* 2007; **297**: 1455-1464 [PMID: 17405971 DOI: 10.1001/jama.297.13.1455]
30 **Almog Y**, Novack V, Eisinger M, Porath A, Novack L, Gilutz H. The effect of statin therapy on infection-related mortality in patients with atherosclerotic diseases. *Crit Care Med* 2007; **35**: 372-378 [PMID: 17205009]
31 **Fernandez R**, De Pedro VJ, Artigas A. Statin therapy prior to ICU admission: protection against infection or a severity marker? *Intensive Care Med* 2006; **32**: 160-164 [PMID: 16086178]
32 **Hackam DG**, Mamdani M, Li P, Redelmeier DA. Statins and sepsis in patients with cardiovascular disease: a population-based cohort analysis. *Lancet* 2006; **367**: 413-418 [PMID: 16458766]
33 **Almog Y**, Shefer A, Novack V, Maimon N, Barski L, Eizinger M, Friger M, Zeller L, Danon A. Prior statin therapy is associated with a decreased rate of severe sepsis. *Circulation* 2004; **110**: 880-885 [PMID: 15289367 DOI: 10.1161/01.CIR.0000138932.17956.F1]
34 **Thomsen RW**, Hundborg HH, Johnsen SP, Pedersen L, Sørensen HT, Schønheyder HC, Lervang HH. Statin use and mortality within 180 days after bacteremia: a population-based cohort study. *Crit Care Med* 2006; **34**: 1080-1086 [PMID: 16484926]
35 **Liappis AP**, Kan VL, Rochester CG, Simon GL. The effect of statins on mortality in patients with bacteremia. *Clin Infect Dis* 2001; **33**: 1352-1357 [PMID: 11565076 DOI:

10.1086/323334]
36 **Subramaniam K**, Koch CG, Bashour A, O'Connor M, Xu M, Gillinov AM, Starr NJ. Preoperative statin intake and morbid events after isolated coronary artery bypass grafting. *J Clin Anesth* 2008; **20**: 4-11 [PMID: 18346602 DOI: 10.1016/j.jclinane.2007.09.003]
37 **Coleman CI**, Lucek DM, Hammond J, White CM. Preoperative statins and infectious complications following cardiac surgery. *Curr Med Res Opin* 2007; **23**: 1783-1790 [PMID: 17597556]
38 **Hauer-Jensen M**, Fort C, Mehta JL, Fink LM. Influence of statins on postoperative wound complications after inguinal or ventral herniorrhaphy. *Hernia* 2006; **10**: 48-52 [PMID: 16151608]
39 Pravastatin and ventilatory associated pneumonia. Available from: URL: http://clinicaltrials.gov
40 Statin-Vap statin-Vap - Statins and Ventilator-Associated Pneumonia. Available from: URL: http: //clinicaltrials.gov
41 **Mortensen EM**, Pugh MJ, Copeland LA, Restrepo MI, Cornell JE, Anzueto A, Pugh JA. Impact of statins and angiotensin-converting enzyme inhibitors on mortality of subjects hospitalised with pneumonia. *Eur Respir J* 2008; **31**: 611-617 [PMID: 17959631 DOI: 10.1183/09031936.00162006]
42 **Chalmers JD**, Singanayagam A, Murray MP, Hill AT. Prior statin use is associated with improved outcomes in community-acquired pneumonia. *Am J Med* 2008; **121**: 1002-1007.e1 [PMID: 18954848]
43 **Thomsen RW**, Riis A, Kornum JB, Christensen S, Johnsen SP, Sørensen HT. Preadmission use of statins and outcomes after hospitalization with pneumonia: population-based cohort study of 29,900 patients. *Arch Intern Med* 2008; **168**: 2081-2087 [PMID: 18955636 DOI: 10.1001/archinte.168.19.2081]
44 **Chopra V**, Flanders SA. Does statin use improve pneumonia outcomes? *Chest* 2009; **136**: 1381-1388 [PMID: 19892677 DOI: 10.1378/chest.09-0941]
45 **Kopterides P**, Falagas ME. Statins for sepsis: a critical and updated review. *Clin Microbiol Infect* 2009; **15**: 325-334 [PMID: 19416304 DOI: 10.1111/j.1469-0691.2009.02750.x]
46 **Falagas ME**, Makris GC, Matthaiou DK, Rafailidis PI. Statins for infection and sepsis: a systematic review of the clinical evidence. *J Antimicrob Chemother* 2008; **61**: 774-785 [PMID: 18263570 DOI: 10.1093/jac/dkn019]

**P- Reviewer:** Surani S, Undela K **S- Editor:** Ji FF **L- Editor:** A **E- Editor:** Lu YJ



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i3.98

*World J Meta-Anal* 2014 August 26; 2(3): 98-106
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Robot-assisted *vs* laparoscopy-assisted gastrectomy for gastric cancer: A meta-analysis based on 3518 subjects

Zhi-Dong Lin, Mao Liu, Dan Tang, Huan Li, Bai-Meng Zhang

Zhi-Dong Lin, Dan Tang, Bai-Meng Zhang, Department of General Surgery, the Fifth Affiliated Hospital of Sun Yat-Sen University, Zhuhai 519000, Guangdong Province, China
Mao Liu, Department of Cardiology, the Fifth Affiliated Hospital of Sun Yat-Sen University, Zhuhai 519000, Guangdong Province, China
Huan Li, Sun Yat-Sen University Cancer Center, Guangzhou 510060, Guangdong Province, China
Author contributions: Lin ZD and Liu M contributed equally to this work; Lin ZD, Tang D and Zhang BM designed the research; Tang D, Li H and Liu M performed the research; Lin ZD and Liu M contributed analytic tools and analyzed the data; Lin ZD, Tang D, Liu M and Zhang BM wrote the paper; Lin ZD and Liu M contributed equally to this work.
Correspondence to: Dr. Zhi-Dong Lin, Department of General Surgery, the Fifth Affiliated Hospital of Sun Yat-Sen University, No.52 Meihua East Rd, Zhuhai 519000, Guangdong Province, China. linzhidongtd@21cn.com
Telephone: +86-756-2528811 Fax: +86-756-2528333
Received: January 8, 2014 Revised: May 16, 2014
Accepted: June 10, 2014
Published online: August 26, 2014

## Abstract

**AIM:** To compare the short-term clinical outcomes of robot-assisted gastrectomy (RAG) with laparoscopy-assisted gastrectomy (LAG) in gastric cancer patients.

**METHODS:** Articles were identified through a literature search of Pubmed, EMBASE, Scopus, Web of Science, Chinese National Knowledge Infrastructure and the Cochrane Library. Weighted mean differences (WMDs) and odds ratios (ORs) were selected as effect sizes for quantitative variables and qualitative variables, respectively. And 95%CIs were also calculated.

**RESULTS:** A total of 13 studies with 3518 patients were included. RAG was associated with longer operative time (WMD = 46.26 min, 95%CI: 31.89-60.63, $P$ < 0.00001), less blood loss [WMD = -37.19 mL, 95%CI: -60.16-(-14.23), $P$ = 0.002] and shorter postoperative hospital stay [WMD = -0.65 d, 95%CI: -1.24-(-0.05), $P$ = 0.03] than LAG. No significant difference in the numbers of retrieved lymph nodes was found between the two groups (WMD = 1.46, 95%CI: -0.19-3.10, $P$ = 0.08). There was no significant difference in mortality (OR = 1.55, 95%CI: 0.49-4.94, $P$ = 0.45), overall complications (OR = 1.00, 95%CI: 0.80-1.26, $P$ = 0.98), anastomosis leakage (OR = 1.02, 95%CI: 0.62-1.65, $P$ = 0.95) and anastomosis stenosis rates (OR = 0.54, 95%CI: 0.18-1.57, $P$ = 0.25).

**CONCLUSION:** RAG is effective and safe in the treatment of gastric cancer. RAG is a promising alternative to laparoscopic surgery. Long-term randomized controlled studies with large scale and improved designs are needed to further evaluate the long-term outcomes.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Gastric cancer; Robot; Laparoscopy; Gastrectomy; Meta-analysis

**Core tip:** A total of 13 studies with 3518 patients were included in this meta-analysis. The results indicated that robot-assisted gastrectomy was associated with longer operative time (WMD = 46.26 min, 95%CI: 31.89, 60.63, $P$ < 0.00001), less blood loss [WMD = -37.19 mL, 95%CI: -60.16-(-14.23), $P$ = 0.002] and shorter postoperative hospital stay [WMD = -0.65 d, 95%CI: -1.24-(-0.05), $P$ = 0.03] than laparoscopy-assisted gastrectomy. Robot-assisted gastrectomy is effective and safe in the treatment of gastric cancer and will be a promising alternative to laparoscopic surgery. Long-term randomized controlled studies with large scale and improved designs are needed to further evaluate the long-term outcomes.

Lin ZD, Liu M, Tang D, Li H, Zhang BM. Robot-assisted *vs* laparoscopy-assisted gastrectomy for gastric cancer: A meta-analysis

based on 3518 subjects. *World J Meta-Anal* 2014; 2(3): 98-106 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i3/98.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i3.98

## INTRODUCTION

Gastric cancer is the fourth leading cancer and second leading cause of cancer death in the world[1,2]. At present, radical gastrectomy with lymph node (LN) dissection is still the mainstay of treatment for gastric cancer[3]. Since 1994, the laparoscopy-assisted gastrectomy (LAG) has become widely accepted in Asian countries because it offers less invasiveness and pain, speedier recovery, milder morbidity and shorter hospital stay[4-7]. According to the report of Japanese Society of Endoscopic Surgery, the total number of patients who had undergone LAG for gastric cancer was 34645 until 2013[8]. However, the instruments of LAG have a limited range of motion and are usually associated with a long learning curve, especially in LN dissection[9].

Another minimally invasive approach for gastric cancer seems to be more promising. Hashizume *et al*[10] had performed distal gastrectomy successfully with the assistance of the da Vinci computer-enhanced surgical system in 2002. They found that the robotic system enhanced visualization of both the operative field and precision of the necessary techniques. It may therefore help surgeons overcome many of the difficulties associated with the endoscopic approach. Since then, several studies[11-15] have been conducted to evaluate the safety and efficacy of robot-assisted gastrectomy (RAG) for gastric cancer. However, most of them were case control studies and their sample sizes were rather small. Therefore, in this study, we conducted a meta-analysis to compare the short-term clinical outcomes of RAG with LAG in gastric cancer patients.

## MATERIALS AND METHODS

### Search strategy

We performed an electronic search of Pubmed, EMBASE, Scopus, Web of Science, Cochrane Library and Chinese National Knowledge Infrastructure from the inception to December 13th, 2013. The following search terms were used: gastric cancer, gastric carcinoma, gastrectomy, robotic, robot, laparoscopy and laparoscopic. Only English and Chinese articles were considered. We also searched additional articles through the reference lists of related papers. Two investigators screened the articles independently.

### Study selection

Two investigators identified appropriate articles and conducted data extraction independently. Eligible studies should match all of the following: (1) study design: prospective or retrospective cohort studies, randomized or nonrandomized controlled studies, case-control studies; (2) study population: gastric cancer patients who received RAG or LAG; (3) grouping: RAG group *vs* LAG group; and (4) outcomes: intraoperative outcome (operative time, blood loss, number of retrieved LNs, conversion to open gastrectomy) and postoperative outcome (overall complications, anastomosis leakage, anastomosis stenosis, bleeding, intestinal obstruction, mortality and postoperative hospital stay). Meeting abstracts, case reports, editorials and reviews were excluded.

### Data extraction and quality assessment

We extracted the study type, country, patient characteristics, age, clinical outcomes, operating cost and the number of cases for each article. The quality of the included studies was evaluated by Newcastle-Ottawa quality assessment (NOS) scale[16]. A study can be awarded a highest score of nine. Data extraction was completed independently by two investigators.

### Statistical analysis

All statistical tests were performed with Review Manager 5.1 software (Cochrane Collaboration, Copenhagen, Denmark). In this study, $I^2$ was used to investigate the heterogeneity. In the analysis process, if $I^2 \geqslant 50\%$, we ran a random-effect model. On the other hand, a fixed-effect model was chosen if $I^2 < 50\%$. In the analysis of quantitative variables (operative time, blood loss, number of retrieved LNs and postoperative hospital stay), we chose weighted mean difference (WMD) with 95%CI as summary statistics. As for qualitative variables (overall complications, anastomosis leakage, anastomosis stenosis, bleedings, intestinal obstruction, conversion to open gastrectomy and mortality), odds ratios (ORs) with 95% CIs were used accordingly. A value of $P < 0.05$ (two-tailed test) was considered statistically significant.

## RESULTS

### Study selection

At the beginning of search process, 535 publications were reviewed. After a screening process, 11 retrospective studies[1,3,6,9,11-15,17,18], one nonrandomized prospective study[19] and one randomized control trial[20] were included (Figure 1). Twelve studies[1,3,6,9,11-13,15,17-20] were from Asia and one[14] from Europe. Totally, 3518 patients with gastric cancer were included in this meta-analysis. Among them, 1143 cases were in RAG group, the other 2375 patients received LAG (Table 1). Ten studies[1,3,9,12-15,17-19] were published in English, and three[6,11,20] published in Chinese. No significant publication bias was found (Figure 2).

### Comparison of intraoperative outcomes between RAG and LAG group

In this pooled analysis, operative time, blood loss and number of retrieved LNs were included. In total, there were 13 studies[1,3,6,9,11-15,17-20] which reported of the operative time, 12 studies[1,3,6,9,11-14,17-20] reported of blood loss

**Table 1 Main characteristic of the included studies**

| Ref. | Country | Group | No. of patients | Age (yr) | Males (%) | BMI (kg/m$^2$) | TNM stage (Ⅰ/Ⅱ/Ⅲ/Ⅳ) | NOS score |
|---|---|---|---|---|---|---|---|---|
| Eom *et al*[17] | South Korea | RAG | 30 | 52.8 | 70 | 24.2 | 25/3/2/0 | 6 |
| | | LAG | 62 | 57.9 | 66.1 | 24.1 | 56/6/0/0 | |
| Huang *et al*[9] | Taiwan China | RAG | 39 | 65.1 | 48.7 | 24.2 | 29/7/3/0 | 6 |
| | | LAG | 64 | 65.6 | 67.2 | 24.7 | 55/9/0/0 | |
| Hyun *et al*[18] | South Korea | RAG | 38 | 54.2 | 65.8 | 23.8 | 30/5/3/0 | 7 |
| | | LAG | 83 | 60.3 | 66.3 | 23.8 | 67/9/7/0 | |
| Kang *et al*[12] | South Korea | RAG | 100 | 53.2 | 63 | 23.7 | 82/11/7/0 | 6 |
| | | LAG | 282 | 58.8 | 67.7 | 23.6 | NR | |
| Kim *et al*[3] | South Korea | RAG | 436 | 54.2 | 60.8 | 23.6 | 350/51/32/0 | 7 |
| | | LAG | 861 | 58.8 | 63.9 | 23.5 | 714/96/43/0 | |
| Kim *et al*[13] | South Korea | RAG | 16 | 53.8 | 62.5 | 21.3 | NR | 6 |
| | | LAG | 11 | 57.9 | 90.9 | 25.3 | NR | |
| Liu *et al*[6] | China | RAG | 48 | 51.8 | 85.4 | 21.2 | 14/5/27/2 | 7 |
| | | LAG | 48 | 52.1 | 83.3 | 21 | 16/6/23/3 | |
| Noshiro *et al*[19] | Japan | RAG | 21 | 66 | 66.7 | 22.8 | 18/-/-/- | 6 |
| | | LAG | 160 | 69 | 63.8 | 21.8 | 113/-/-/- | |
| Pugliese *et al*[14] | Italy | RAG | 16 | 71 | NR | 28.8 | NR | 7 |
| | | LAG | 48 | 71 | NR | 28.8 | NR | |
| Woo *et al*[1] | South Korea | RAG | 236 | 54 | 57.6 | 23.5 | 236/0/0/0 | 6 |
| | | LAG | 591 | 58.3 | 61.6 | 23.5 | 591/0/0/0 | |
| Yoon *et al*[15] | South Korea | RAG | 36 | 53.9 | 50 | 23.2 | 29/7/0/0 | 7 |
| | | LAG | 65 | 56.9 | 47.7 | 23.6 | 55/7/3/0 | |
| Zhang *et al*[11] | China | RAG | 97 | 56.1 | 68 | 22.5 | 23/22/52/0 | 7 |
| | | LAG | 70 | 54.8 | 70 | 21.7 | 8/17/45/0 | |
| Zhao *et al*[20] | China | RAG | 30 | 71.8 | 73.3 | 23.6 | 2/18/9/1 | 8 |
| | | LAG | 30 | 72.4 | 76.7 | 23.9 | 1/25/3/1 | |

BMI: Body mass index; TNM: Tumor node metastases; NOS: Newcastle-ottawa quality scale; RAG: Robot-assisted gastrectomy; LAG: Laparoscopy-assisted gastrectomy; NR: Not Reported.

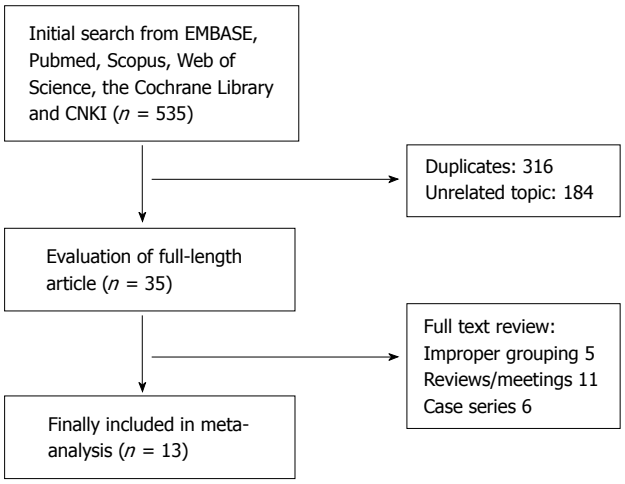



**Figure 1 Flow chart of literature search for meta-analysis.**

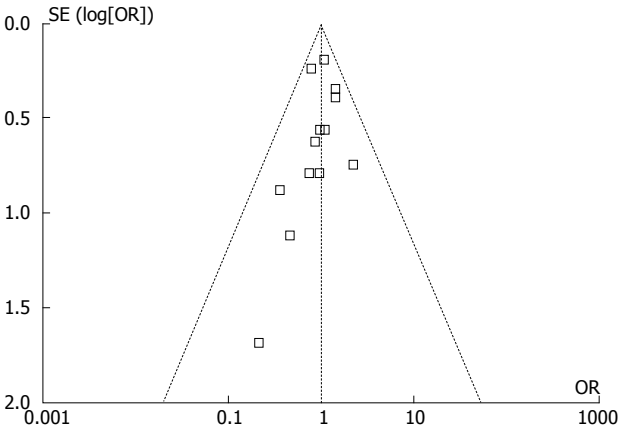



**Figure 2 Funnel plot of included studies in this meta-analysis.**

and 12 studies[1,3,6,9,11,13-15,17-20] reported of the number of retrieved LNs. In the heterogeneity tests of operative time, blood loss and number of retrieved LNs, $I^2$ were 91%, 94% and 70%, respectively. Accordingly, we chose the random-effect model.

As shown in Figure 3A, the mean operative time for the RAG group was on average 46 min longer than the LAG group (WMD = 46.26 min, 95%CI: 31.89-60.63, $P < 0.00001$), while mean blood loss was significantly less in the RAG group [WMD = -37.19 mL, 95%CI: -60.16-(-14.23), $P = 0.002$]. The pooled results also indicated that there was no significant difference in the number of retrieved LNs between the two groups (WMD = 1.46, 95%CI: -0.19-3.10, $P = 0.08$).

### *Comparison of postoperative outcomes between RAG and LAG group*

Overall complications, anastomosis leakage and anastomosis stenosis were included for analysis. Information in detail is shown in Figure 3B. Thirteen studies[1,3,6,9,11-15,17-20] which reported of the overall complications were included. No statistical heterogeneity was found in this analysis

**A**

| Study or Subgroup | RAG Mean | RAG SD | RAG Total | LAG Mean | LAG SD | LAG Total | Weight | Mean difference IV, Random, 95%CI | Year |
|---|---|---|---|---|---|---|---|---|---|
| 1.1.1 Operative time | | | | | | | | | |
| Kim | 259.2 | 38.9 | 16 | 203.9 | 36.4 | 11 | 6.9% | 55.30 [26.56, 84.04] | 2010 |
| Pugliese | 344 | 62 | 16 | 235 | 23 | 48 | 6.6% | 109.00 [77.93, 140.07] | 2010 |
| Woo | 219.5 | 46.8 | 236 | 170.7 | 55.8 | 591 | 9.1% | 48.80 [41.32, 56.28] | 2011 |
| Eom | 229.1 | 52.2 | 30 | 189.4 | 52.2 | 62 | 7.7% | 39.70 [16.95, 62.45] | 2012 |
| Huang | 415.9 | 101.2 | 39 | 362.2 | 72.8 | 64 | 5.9% | 53.70 [17.27, 90.13] | 2012 |
| Zhang | 272.3 | 46.1 | 97 | 240.3 | 89.1 | 70 | 7.7% | 32.00 [9.20, 54.80] | 2012 |
| Kang | 202.1 | 52.3 | 100 | 173.5 | 51.2 | 282 | 9.0% | 28.60 [16.73, 40.47] | 2012 |
| Kim | 226 | 54 | 436 | 176 | 63 | 861 | 9.4% | 50.00 [43.41, 56.59] | 2012 |
| Yoon | 305.8 | 115.8 | 36 | 210.2 | 57.7 | 65 | 5.5% | 95.60 [55.26, 135.94] | 2013 |
| Hyun | 234.4 | 48 | 38 | 220 | 60.6 | 83 | 8.1% | 14.40 [-5.67, 34.47] | 2013 |
| Zhao | 156.7 | 14.9 | 30 | 159 | 27.1 | 30 | 9.1% | -2.30 [-13.37, 8.77] | 2013 |
| Liu | 238 | 29.6 | 48 | 221.5 | 33.1 | 48 | 9.0% | 16.50 [3.94, 29.06] | 2013 |
| Noshiro | 439 | 86 | 21 | 315 | 90 | 160 | 5.6% | 124.00 [84.66, 163.34] | 2013 |
| Subtotal (95%CI) | | | 1143 | | | 2375 | 100.0% | 46.26 [31.89, 60.63] | |
| Heterogeneity: Tau$^2$ = 559.65; $\chi^2$ = 139.08, df = 12 ($P$ < 0.00001); $I^2$ = 91% | | | | | | | | | |
| Test for overall effect: $Z$ = 6.31 ($P$ < 0.00001) | | | | | | | | | |
| 1.1.2 Blood loss | | | | | | | | | |
| Pugliese | 90 | 48 | 16 | 148 | 53 | 48 | 8.6% | -58.00 [-85.89, -30.11] | 2010 |
| Kim | 30.3 | 15.1 | 16 | 44.7 | 37.1 | 11 | 9.0% | -14.40 [-37.54, 8.74] | 2010 |
| Woo | 91.6 | 152.6 | 236 | 147.9 | 269 | 591 | 8.5% | -56.30 [-85.44, -27.16] | 2011 |
| Eom | 152.8 | 173.5 | 30 | 88.3 | 173.5 | 62 | 4.8% | 64.50 [-11.13, 140.13] | 2012 |
| Huang | 93.9 | 89 | 39 | 163.7 | 154.7 | 64 | 7.0% | -69.80 [-116.88, -22.72] | 2012 |
| Kang | 93.3 | 84.6 | 100 | 173.5 | 145.2 | 282 | 8.9% | -80.20 [-103.91, -56.49] | 2012 |
| Kim | 85 | 160 | 436 | 112 | 229 | 861 | 9.0% | -27.00 [-48.44, -5.56] | 2012 |
| Zhang | 80.8 | 53.1 | 97 | 153.7 | 26.4 | 70 | 9.6% | -72.90 [-85.14, -60.66] | 2012 |
| Hyun | 131.3 | 10.1 | 38 | 130.5 | 17.8 | 83 | 9.8% | 0.80 [-4.20, 5.80] | 2013 |
| Noshiro | 91.4 | 108.6 | 21 | 109.5 | 165.7 | 160 | 6.5% | -18.10 [-71.17, 34.97] | 2013 |
| Liu | 107.8 | 55.9 | 48 | 132.7 | 51.5 | 48 | 8.9% | -24.90 [-48.41, -1.39] | 2013 |
| Zhao | 38 | 17.9 | 30 | 85 | 40.8 | 30 | 9.4% | -47.00 [-62.94, -31.06] | 2013 |
| Subtotal (95%CI) | | | 1107 | | | 2310 | 100.0% | -37.19 [-60.16, -14.23] | |
| Heterogeneity: Tau$^2$ = 1393.22; $\chi^2$ = 195.54, df = 11 ($P$ < 0.00001); $I^2$ = 94% | | | | | | | | | |
| Test for overall effect: $Z$ = 3.17 ($P$ = 0.002) | | | | | | | | | |
| 1.1.3 Number of retrieved LNs | | | | | | | | | |
| Pugliese | 25 | 4.5 | 16 | 31 | 8 | 48 | 9.5% | -6.00 [-9.16, -2.84] | 2010 |
| Kim | 41.1 | 10.9 | 16 | 37.4 | 10 | 11 | 3.3% | 3.70 [-4.27, 11.67] | 2010 |
| Woo | 39 | 15.2 | 236 | 37.4 | 14.2 | 591 | 11.5% | 1.60 [-0.65, 3.85] | 2011 |
| Kim | 40.2 | 15.5 | 436 | 37.6 | 13.9 | 861 | 12.6% | 2.60 [0.87, 4.33] | 2012 |
| Huang | 32 | 13.7 | 39 | 26 | 12.4 | 64 | 5.9% | 6.00 [0.74, 11.26] | 2012 |
| Eom | 30.2 | 9.9 | 30 | 33.4 | 9.9 | 62 | 7.3% | -3.20 [-7.52, 1.12] | 2012 |
| Zhang | 23.1 | 5.4 | 97 | 20 | 4.3 | 70 | 13.1% | 3.10 [1.63, 4.57] | 2012 |
| Yoon | 42.8 | 12.7 | 36 | 39.4 | 13.4 | 65 | 5.8% | 3.40 [-1.87, 8.67] | 2012 |
| Hyun | 32.8 | 123.8 | 38 | 32.6 | 13.3 | 83 | 5.9% | 0.20 [-5.04, 5.44] | 2013 |
| Noshiro | 44 | 19 | 21 | 40 | 15 | 160 | 3.0% | 4.00 [-4.45, 12.45] | 2013 |
| Liu | 34.7 | 5.2 | 48 | 32.5 | 4.7 | 48 | 12.1% | 2.20 [0.22, 4.18] | 2013 |
| Zhao | 30.8 | 6.3 | 30 | 28.1 | 5.1 | 30 | 10.1% | 2.70 [-0.20, 5.60] | 2013 |
| Subtotal (95%CI) | | | 1043 | | | 2093 | 100.0% | 1.46 [-0.19, 3.10] | |
| Heterogeneity: Tau$^2$ = 4.80; $\chi^2$ = 36.48, df = 11 ($P$ = 0.0001); $I^2$ = 70% | | | | | | | | | |
| Test for overall effect: $Z$ = 1.74 ($P$ = 0.08) | | | | | | | | | |

Mean difference IV, Random, 95%CI

-100 -50 0 50 100

Favours RAG Favours LAG

Test for subgroup differences: $\chi^2$ = 48.00, df = 2 ($P$ < 0.00001); $I^2$ = 95.8%

**B**

| Study or Subgroup | RAG Mean | RAG Total | LAG Mean | LAG Total | Weight | Odds ratio M-H, Fixed, 95%CI | Year |
|---|---|---|---|---|---|---|---|
| 2.1.1 Overall complications | | | | | | | |
| Pugliese | 1 | 16 | 6 | 48 | 1.9% | 0.47 [0.05, 4.20] | 2010 |
| Kim | 0 | 16 | 1 | 11 | 1.1% | 0.21 [0.01, 5.71] | 2010 |
| Woo | 26 | 236 | 81 | 591 | 27.4% | 0.78 [0.49, 1.25] | 2011 |
| Yoon | 6 | 36 | 10 | 65 | 4.0% | 1.10 [0.365, 3.32] | 2012 |
| Zhang | 6 | 97 | 5 | 70 | 3.6% | 0.86 [0.25, 2.93] | 2012 |
| Kang | 14 | 100 | 29 | 282 | 8.7% | 1.42 [0.72, 2.81] | 2012 |
| Kim | 44 | 436 | 81 | 861 | 32.6% | 1.08 [0.73, 1.59] | 2012 |
| Huang | 6 | 39 | 10 | 64 | 4.3% | 0.98 [0.33, 2.95] | 2012 |
| Eom | 4 | 30 | 4 | 62 | 1.5% | 2.23 [0.52, 9.62] | 2012 |
| Zhao | 2 | 30 | 5 | 30 | 3.1% | 0.36 [0.06, 2.01] | 2013 |
| Hyun | 18 | 38 | 32 | 83 | 7.0% | 1.43 [0.66, 3.11] | 2013 |

Odds ratio M-H, Fixed, 95%CI

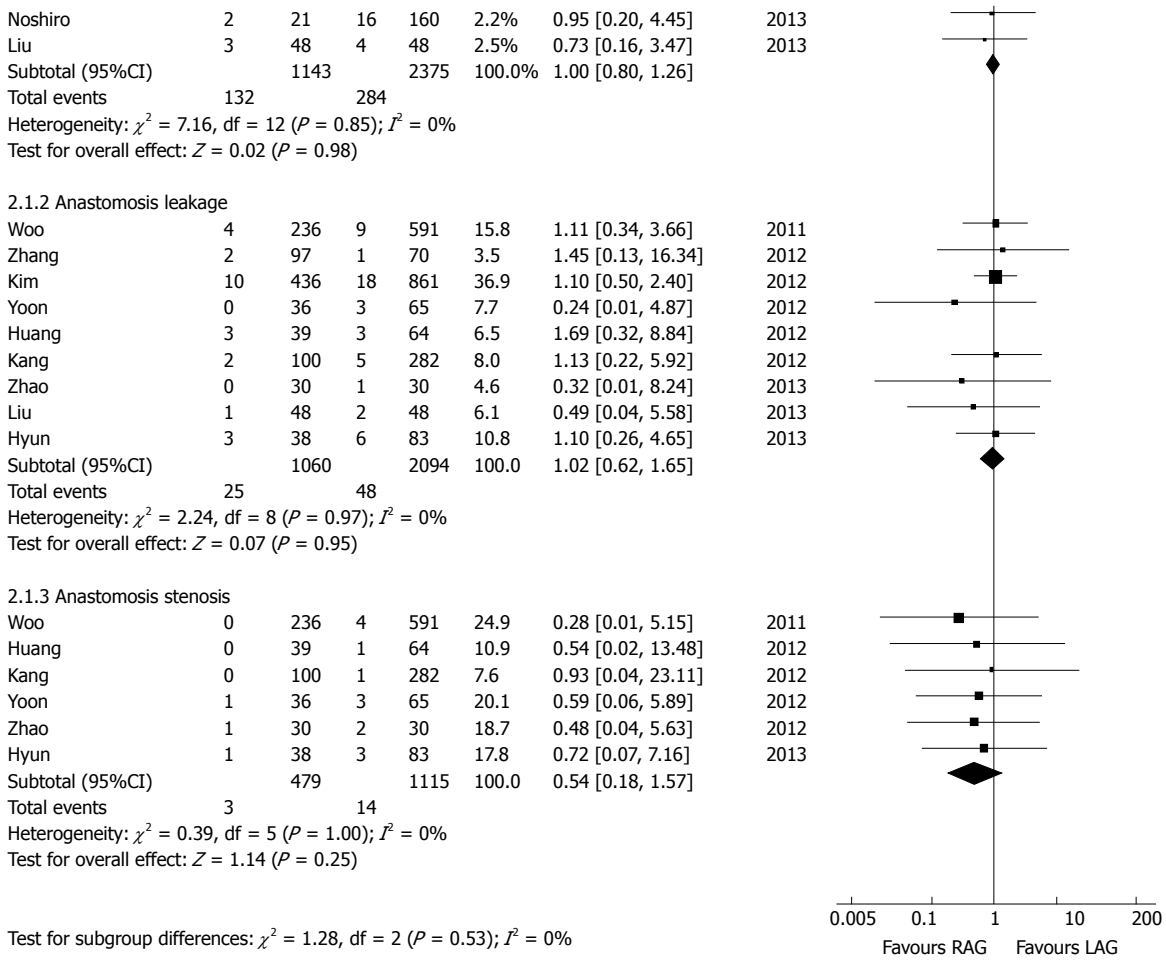



**Figure 3 Forest plot comparing intraoperative (A) and postoperative (B) outcomes between robot-assisted gastrectomy and laparoscopy-assisted gastrectomy.** RAG: Robot-assisted gastrectomy; LAG: Laparoscopy-assisted gastrectomy.

($I^2$ = 0%). A fixed-effect model was selected. No significant difference between RAG and LAG group was found in the comparison of the incidences of overall complications (11.5% *vs* 12.0%, OR = 1.00, 95%CI: 0.80-1.26, $P$ = 0.98).

The incidences of anastomosis leakages were reported in 9 studies[1,3,6,9,11,12,15,18,20]. In total, 1060 patients were treated with RAG and 2094 patients received LAG. No statistical heterogeneity was found ($I^2$ = 0%). There was no significant difference between RAG and LAG group in the comparison of the incidences of anastomosis leakages (2.4% *vs* 2.3%, OR = 1.02, 95%CI: 0.62-1.65, $P$ = 0.95).

Six studies[1,9,12,15,18,20] involving 1594 subjects were included in the analysis of anastomosis stenosis rates. No statistical heterogeneity was found in this analysis ($I^2$ = 0%) and a fixed-effect model was selected. The results didn' t indicate statistical difference between the two groups in the comparison of the anastomosis stenosis rates (0.6% *vs* 1.3%, OR = 0.54, 95%CI: 0.18-1.57, $P$ = 0.25).

### *Comparison of other clinical outcomes between RAG and LAG group*

There were 7 studies[1,6,13,14,17-19] which reported of numbers of patients in LAG or RAG group who converted to open gastrectomy (Figure 4A). No statistical heterogeneity was found ($I^2$ = 0%). There was no statistical difference in the comparison of conversion to open gastrectomy between the two groups (OR = 2.10, 95%CI: 0.47-9.40, $P$ = 0.33).

The incidences of bleeding events after operation were reported in 6 studies[1,3,9,11,12,18], involving 2897 subjects. No statistical heterogeneity was found in this analysis ($I^2$ = 29%) and a fixed-effect model was selected (Figure 4B). The results indicated that there was no significant difference in the comparison of bleeding rates (1.6% *vs* 1.4%, OR = 1.19, 95%CI: 0.63-2.24, $P$ = 0.59).

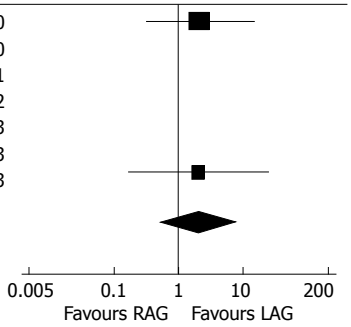
Seven studies[1,3,9,11,12,15,20] involving 2937 patients were included in the analysis of intestinal obstruction (Figure 4C). A fixed-effect model was selected ($I^2$ = 0%). No significant difference was found in the comparison of intestinal obstruction rates (1.1% *vs* 0.9%, OR = 1.17, 95%CI: 0.56-2.48, $P$ = 0.67).

Eleven studies[1,3,6,9,12-15,17,19,20] had reported the mortalities (Figure 4D). There were 3230 subjects included (1008 in RAG group and 2222 in LAG group). No statistical heterogeneity was found ($I^2$ = 0%). The results indicated no significant difference of mortality between the two

**A**

| Study or Subgroup | RAG Mean | RAG Total | LAG Mean | LAG Total | Weight | Odds ratio M-H, Fixed, 95%CI | Year |
|---|---|---|---|---|---|---|---|
| Pugliese | 2 | 16 | 3 | 48 | 57.8% | 2.14 [0.32, 14.14] | 2010 |
| Kim | 0 | 16 | 0 | 11 | | Not estimable | 2010 |
| Woo | 0 | 236 | 0 | 591 | | Not estimable | 2011 |
| Eom | 0 | 30 | 0 | 62 | | Not estimable | 2012 |
| Noshiro | 0 | 21 | 0 | 160 | | Not estimable | 2013 |
| Hyun | 0 | 38 | 0 | 83 | | Not estimable | 2013 |
| Liu | 2 | 48 | 1 | 48 | 42.2% | 2.04 [0.18, 23.32] | 2013 |
| Subtotal (95%CI) | | 405 | | 1003 | 100.0% | 2.10 [0.47, 9.40] | |
| Total events | 4 | | 4 | | | | |

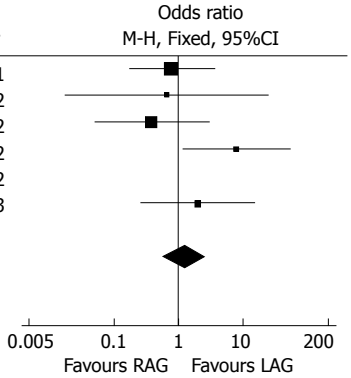
Heterogeneity: $\chi^2$ = 0.00, df = 1 ($P$ = 0.98); $I^2$ = 0%
Test for overall effect: $Z$ = 0.97 ($P$ = 0.33)




**B**

| Study or Subgroup | RAG Mean | RAG Total | LAG Mean | LAG Total | Weight | Odds ratio M-H, Fixed, 95%CI | Year |
|---|---|---|---|---|---|---|---|
| Woo | 4 | 236 | 12 | 591 | 39.5% | 0.83 [0.27, 2.61] | 2011 |
| Zhang | 1 | 97 | 1 | 70 | 6.7% | 0.72 [0.04, 11.69] | 2012 |
| Kim | 2 | 436 | 8 | 861 | 31.4% | 0.49 [0.10, 2.32] | 2012 |
| Kang | 5 | 100 | 3 | 282 | 8.8% | 4.89 [1.15, 20.87] | 2012 |
| Huang | 0 | 39 | 0 | 64 | | Not estimable | 2012 |
| Hyun | 3 | 38 | 4 | 83 | 13.6% | 1.69 [0.36, 7.97] | 2013 |
| Subtotal (95%CI) | | 946 | | 1951 | 100.0% | 1.19 [0.63, 2.24] | |
| Total events | 15 | | 28 | | | | |

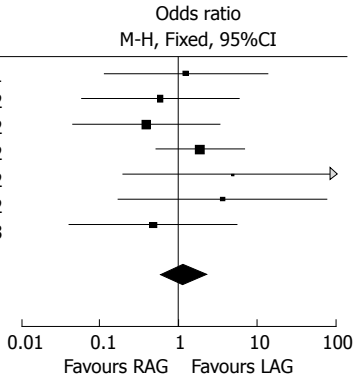
Heterogeneity: $\chi^2$ = 5.60, df = 4 ($P$ = 0.23); $I^2$ = 29%
Test for overall effect: $Z$ = 0.54 ($P$ = 0.59)




**C**

| Study or Subgroup | RAG Mean | RAG Total | LAG Mean | LAG Total | Weight | Odds ratio M-H, Fixed, 95%CI | Year |
|---|---|---|---|---|---|---|---|
| Woo | 1 | 236 | 2 | 591 | 9.1% | 1.25 [0.11, 13.89] | 2011 |
| Yoon | 1 | 36 | 3 | 65 | 16.7 | 0.59 [0.06, 5.89] | 2012 |
| Kim | 1 | 436 | 5 | 861 | 26.9 | 0.39 [0.05, 3.38] | 2012 |
| Kang | 4 | 100 | 6 | 282 | 24.2 | 1.92 [0.53, 6.94] | 2012 |
| Huang | 1 | 39 | 0 | 64 | 2.9 | 5.03 [0.20, 126.47] | 2012 |
| Zhang | 2 | 97 | 0 | 70 | 4.5 | 3.69 [0.17, 78.09] | 2012 |
| Zhao | 1 | 30 | 2 | 30 | 15.5% | 0.48 [0.04, 5.63] | 2013 |
| Subtotal (95%CI) | | 974 | | 1963 | 100.0% | 1.17 [0.56, 2.48] | |
| Total events | 11 | | 18 | | | | |

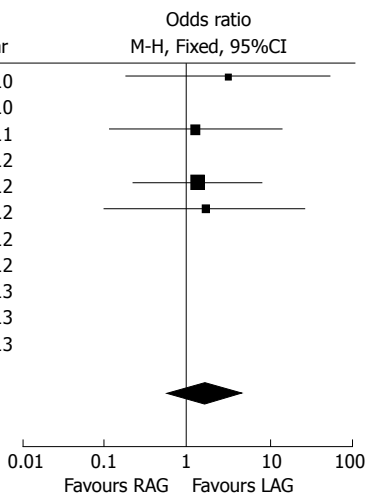
Heterogeneity: $\chi^2$ = 3.72, df = 6 ($P$ = 0.71); $I^2$ = 0%
Test for overall effect: $Z$ = 0.42 ($P$ = 0.67)




**D**

| Study or Subgroup | RAG Mean | RAG Total | LAG Mean | LAG Total | Weight | Odds ratio M-H, Fixed, 95%CI | Year |
|---|---|---|---|---|---|---|---|
| Pugliese | 1 | 16 | 1 | 48 | 10.8% | 3.13 [0.18, 53.21] | 2010 |
| Kim | 0 | 16 | 0 | 11 | | Not estimable | 2010 |
| Woo | 1 | 236 | 2 | 591 | 26.1% | 1.25 [0.11, 13.89] | 2011 |
| Yoon | 0 | 36 | 0 | 65 | | Not estimable | 2012 |
| Kim | 2 | 436 | 3 | 861 | 46.1% | 1.32 [0.22, 7.92] | 2012 |
| Huang | 1 | 39 | 1 | 64 | 17.0% | 1.66 [0.10, 27.29] | 2012 |
| Kang | 0 | 100 | 0 | 282 | | Not estimable | 2012 |
| Eom | 0 | 30 | 0 | 62 | | Not estimable | 2012 |
| Zhao | 0 | 30 | 0 | 30 | | Not estimable | 2013 |
| Liu | 0 | 48 | 0 | 48 | | Not estimable | 2013 |
| Noshiro | 0 | 21 | 0 | 160 | | Not estimable | 2013 |
| Subtotal (95%CI) | | 405 | | 1003 | 100.0% | 2.10 [0.47, 9.40] | |
| Total events | 4 | | 4 | | | | |

Heterogeneity: $\chi^2$ = 0.30, df = 1 ($P$ = 0.96); $I^2$ = 0%
Test for overall effect: $Z$ = 0.75 ($P$ = 0.45)




**E**

| Study or Subgroup | RAG Mean | RAG SD | RAG Total | LAG Mean | LAG SD | LAG Total | Weight | Mean difference IV, Random, 95%CI | Year |
|---|---|---|---|---|---|---|---|---|---|
| Pugliese | 10 | 3 | 16 | 10 | 2.6 | 48 | 7.0% | 0.00 [-1.64, 1.64] | 2010 |
| Kim | 5.1 | 0.3 | 16 | 6.5 | 0.8 | 11 | 13.5% | -1.40 [-1.90, -0.90] | 2010 |
| Woo | 7.7 | 17.2 | 236 | 7 | 5.7 | 591 | 4.8% | 0.70 [-1.54, 2.94] | 2011 |



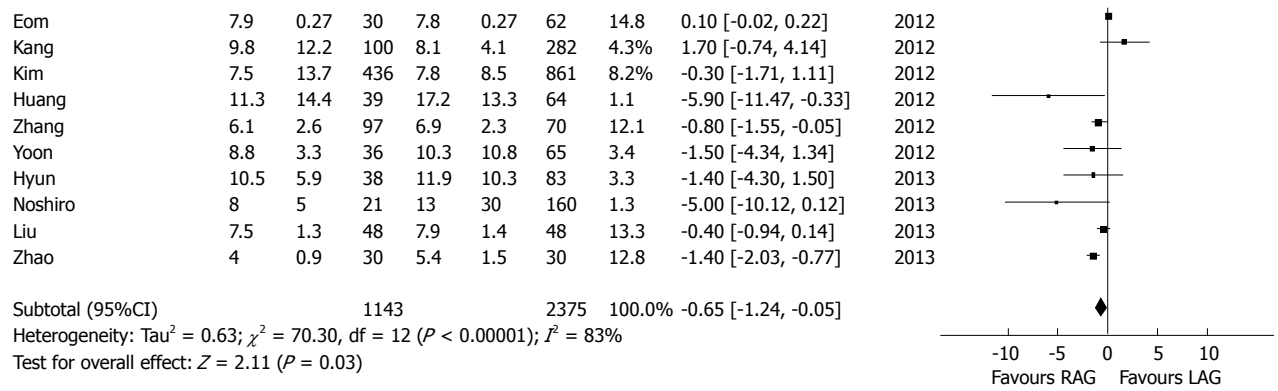



**Figure 4 Forest plot.** A: Comparing conversion to open gastrectomy between RAG and LAG; B: Comparing bleeding events between RAG and LAG; C: Comparing intestinal obstruction between RAG and LAG; D: Comparing mortality between RAG and LAG; E: Comparing postoperative hospital stay between RAG and LAG. RAG: Robot-assisted gastrectomy; LAG: Laparoscopy-assisted gastrectomy.

groups (0.5% *vs* 0.3%, OR = 1.55, 95%CI: 0.49-4.94, $P$ = 0.45).

Thirteen studies[1,3,6,9,11-15,17-20] involving 3518 patients were included in the analysis of postoperative hospital stay (Figure 4E). A random-effect model was selected ($I^2$ = 83%). The RAG group had a shorter mean postoperative hospital stay than the LAG group [WMD = -0.65 d, 95%CI: -1.24-(-0.05), $P$ = 0.03].

## DISCUSSION

The findings from our meta-analysis suggest that RAG is effective and safe for gastric cancer compared to LAG. Overall, combining the available data RAG was associated with longer operative time, less blood loss and shorter postoperative hospital stay than LAG. Moreover, there was no significant difference in mortality, conversion, overall complications, postoperative bleeding events, intestinal obstruction, anastomosis leakage and anastomosis stenosis rates. There was also no significant difference in the numbers of retrieved LNs during the operation between RAG and LAG.

Previous studies have reported the application of RAG for the treatment of gastric cancer. Yoon *et al*[15] included 36 patients who underwent RAG and 65 patients who underwent LAG at the National Cancer Center in South Korea. The operative data, postoperative morbidity, and pathologic data were analyzed. They found that the mean postoperative hospital stay was 8.8 ± 3.3 d in the RAG group and 10.3 ± 10.8 d in the LATG group ($P$ = 0.416). The mean operative time was 305.8 ± 115.8 min in the RAG group and 210.2 ± 57.7 min in the LAG group ($P$ < 0.001). No significant differences were found in the comparison of mean number of dissected LNs and incidence of postoperative complications. Some other studies[6,18,20] and meta-analysis[21,22] have reported similar results. However, these studies have limited samples and most of them were retrospective. Therefore, we pooled relevant studies and conducted a meta-analysis to compare the short-term clinical outcomes of RAG with LAG systematically. Finally, 13 studies involving 3815 subjects were included. The quality of these studies was relatively high because their NOS scores ranged from 6 to 8. There was a significant heterogeneity among the included studies in the analysis of intraoperative outcomes. This may be explained by the differences in the stage of gastric cancer, resection scope, operation skill, gastric resection approach, extension of LN dissection and the standards for discharge among the studies. Further, according to the funnel plot, the publication bias was acceptable.

According to the results of our analysis, the operative time is much longer in RAG group. It may be related to the increased set-up time to position and the inexpert skill of surgeons. RAG was also associated with less estimated blood loss compared with LAG. It's more convenient for hemostatic treatment because RAG provides an excellent and stable visualization of the operative field[9,11]. Even though the mean postoperative hospital stay is 0.65 d shorter in RAG, we think that it is of little practical significance because it's too short. Moreover, there are no differences between RAG and LAG in the comparison of retrieved LNs and postoperative outcomes. Briefly, the results in the current study indicate that RAG is as safe and effective as LAG in the treatment of gastric cancer.

However, the costs of RAG are much higher than those of LAG. The mean cost of RAG is about $6000 to $11400 for gastric cancer, while only $2000 to $6000 in LAG group[9,11,17]. Consequently, before surgeons and patients make the decision, patients' economic condition should also be taken into consideration.

However, this study had some potential limitations. Firstly, there might be a certain degree of language bias because only publications in Chinese or English were searched in the databases. And then, the number of included subjects was relatively few in this study, which may lead to low statistical power. Moreover, most of them were retrospective designed and long-term outcomes were not reported. More high-quality randomized clinical studies are deserved to better evaluate both short and long-term outcomes of RAG. Further, the end points predetermined in the included studies were different. We can only partly extract the information from these

studies. As for study population, most participants were Asian. Studies in Western countries were relatively rare. Lastly, the differences in population characteristics (stage of gastric cancer, age, gender ratio, diabetes mellitus, hypertension, *etc*.), device and the duration of follow up among the included studies may also lead to a bias in a certain degree.

In conclusion, the synthesis of available evidence indicates that RAG is effective and safe in the treatment of gastric cancer. RAG is a promising alternative to laparoscopic surgery. Long-term randomized controlled studies with large scales and improved designs are needed to further evaluate the long-term outcomes.

## ACKNOWLEDGEMENTS

The authors wish to thank Dan Huang for assistance with references.

## COMMENTS

### Background

Gastric cancer is the fourth leading cancer and second leading cause of cancer death in the world. Robot-assisted gastrectomy (RAG) is a new approach for gastric cancer and is reported to be safe and efficient. However, most of the studies were case control studies and their sample sizes were rather small.

### Research frontiers

The purpose of this study was to perform a meta-analysis to compare the short-term clinical outcomes of RAG with laparoscopy-assisted gastrectomy (LAG) in gastric cancer patients systematically.

### Innovations and breakthroughs

In the present study, the results indicate that RAG is associated with longer operative time [weighted mean difference (WMD) = 46.26 min, $P < 0.00001$], less blood loss (WMD = -37.19 mL, $P = 0.002$) and shorter postoperative hospital stay (WMD = -0.65 d, $P = 0.03$) than LAG. So far, this is a meta-analysis with most included studies and largest number of included subjects.

### Applications

RAG is effective and safe in the treatment of gastric cancer. RAG is a promising alternative to laparoscopic surgery.

### Terminology

RAG: Robot-assisted gastrectomy. Robot-assisted surgery is a kind of minimally invasive approaches. It can be used in gastrectomy and may enhance visualization of both the operative field and precision of the necessary techniques. The most popular one is da Vinci computer-enhanced surgical system.

### Peer review

Lin *et al* in their manuscript present an interesting meta-analysis. This is a well done study.

## REFERENCES

1 **Woo Y**, Hyung WJ, Pak KH, Inaba K, Obama K, Choi SH, Noh SH. Robotic gastrectomy as an oncologically sound alternative to laparoscopic resections for the treatment of early-stage gastric cancers. *Arch Surg* 2011; **146**: 1086-1092 [PMID: 21576595 DOI: 10.1001/archsurg.2011.114]

2 **Hohenberger P**, Gretschel S. Gastric cancer. *Lancet* 2003; **362**: 305-315 [PMID: 12892963 DOI: 10.1016/S0140-6736(03)13975-X]

3 **Kim KM**, An JY, Kim HI, Cheong JH, Hyung WJ, Noh SH. Major early complications following open, laparoscopic and robotic gastrectomy. *Br J Surg* 2012; **99**: 1681-1687 [PMID: 23034831 DOI: 10.1002/bjs.8924]

4 **Shehzad K**, Mohiuddin K, Nizami S, Sharma H, Khan IM, Memon B, Memon MA. Current status of minimal access surgery for gastric cancer. *Surg Oncol* 2007; **16**: 85-98 [PMID: 17560103 DOI: 10.1016/j.suronc.2007.04.012]

5 **Qiu J**, Pankaj P, Jiang H, Zeng Y, Wu H. Laparoscopy versus open distal gastrectomy for advanced gastric cancer: a systematic review and meta-analysis. *Surg Laparosc Endosc Percutan Tech* 2013; **23**: 1-7 [PMID: 23386142 DOI: 10.1097/SLE.0b013e3182747af7]

6 **Liu C**, Tang B, He Y, Shi Y, Zeng D, Luo H, Zhao Y, Qian F, Yu P. Surgical short-term outcomes of robotic gastrectomy vs laparoscopic gastrectomy: a case-control study. *Disan Junyi Daxue Xuebao* 2013, **35**: 1164-1166

7 **Kitano S**, Iso Y, Moriyama M, Sugimachi K. Laparoscopy-assisted Billroth I gastrectomy. *Surg Laparosc Endosc* 1994; **4**: 146-148 [PMID: 8180768]

8 **Etoh T**, Inomata M, Shiraishi N, Kitano S. Minimally invasive approaches for gastric cancer-Japanese experiences. *J Surg Oncol* 2013; **107**: 282-288 [PMID: 22504947 DOI: 10.1002/jso.23128]

9 **Huang KH**, Lan YT, Fang WL, Chen JH, Lo SS, Hsieh MC, Li AF, Chiou SH, Wu CW. Initial experience of robotic gastrectomy and comparison with open and laparoscopic gastrectomy for gastric cancer. *J Gastrointest Surg* 2012; **16**: 1303-1310 [PMID: 22450954 DOI: 10.1007/s11605-012-1874-x]

10 **Hashizume M**, Shimada M, Tomikawa M, Ikeda Y, Takahashi I, Abe R, Koga F, Gotoh N, Konishi K, Maehara S, Sugimachi K. Early experiences of endoscopic procedures in general surgery assisted by a computer-enhanced surgical system. *Surg Endosc* 2002; **16**: 1187-1191 [PMID: 11984681 DOI: 10.1007/s004640080154]

11 **Zhang X**, Jiang Z, Kun Z. Comparative study on clinical efficacy of robot-assisted and laparoscopic gastrectomy for gastric cancer. *Zhonhua Weichang Waike Zazhi* 2012, **15**: 804-806 [DOI: 10.3760/cma.j.issn.1671-0274.2012. 08.016]

12 **Kang BH**, Xuan Y, Hur H, Ahn CW, Cho YK, Han SU. Comparison of Surgical Outcomes between Robotic and Laparoscopic Gastrectomy for Gastric Cancer: The Learning Curve of Robotic Surgery. *J Gastric Cancer* 2012; **12**: 156-163 [PMID: 23094227 DOI: 10.5230/jgc.2012.12.3.156]

13 **Kim MC**, Heo GU, Jung GJ. Robotic gastrectomy for gastric cancer: surgical techniques and clinical merits. *Surg Endosc* 2010; **24**: 610-615 [PMID: 19688399 DOI: 10.1007/s00464-009-0618-9]

14 **Pugliese R**, Maggioni D, Sansonna F, Costanzi A, Ferrari GC, Di Lernia S, Magistro C, De Martini P, Pugliese F. Subtotal gastrectomy with D2 dissection by minimally invasive surgery for distal adenocarcinoma of the stomach: results and 5-year survival. *Surg Endosc* 2010; **24**: 2594-2602 [PMID: 20414682 DOI: 10.1007/s00464-010-1014-1]

15 **Yoon HM**, Kim YW, Lee JH, Ryu KW, Eom BW, Park JY, Choi IJ, Kim CG, Lee JY, Cho SJ, Rho JY. Robot-assisted total gastrectomy is comparable with laparoscopically assisted total gastrectomy for early gastric cancer. *Surg Endosc* 2012; **26**: 1377-1381 [PMID: 22083338 DOI: 10.1007/s00464-011-2043-0]

16 **Mao L**, Jian C, Changzhi L, Dan H, Suihua H, Wenyi T, Wei W. Cytochrome CYP2C19 polymorphism and risk of adverse clinical events in clopidogrel-treated patients: a meta-analysis based on 23,035 subjects. *Arch Cardiovasc Dis* 2013; **106**: 517-527 [PMID: 24080325 DOI: 10.1016/j.acvd.2013.06.055]

17 **Eom BW**, Yoon HM, Ryu KW, Lee JH, Cho SJ, Lee JY, Kim CG, Choi IJ, Lee JS, Kook MC, Rhee JY, Park SR, Kim YW. Comparison of surgical performance and short-term clinical outcomes between laparoscopic and robotic surgery in distal gastric cancer. *Eur J Surg Oncol* 2012; **38**: 57-63 [PMID: 21945625 DOI: 10.1016/j.ejso.2011.09.006]

18 **Hyun MH**, Lee CH, Kwon YJ, Cho SI, Jang YJ, Kim DH, Kim JH, Park SH, Mok YJ, Park SS. Robot versus laparoscopic gastrectomy for cancer by an experienced surgeon:

comparisons of surgery, complications, and surgical stress. *Ann Surg Oncol* 2013; **20**: 1258-1265 [PMID: 23080320 DOI: 10.1245/s10434-012-2679-6]

19 **Noshiro H**, Ikeda O, Urata M. Robotically-enhanced surgical anatomy enables surgeons to perform distal gastrectomy for gastric cancer using electric cautery devices alone. *Surg Endosc* 2014; **28**: 1180-1187 [PMID: 24202713]

20 **Zhao K**, Pan H, Wang G, Li M, Ruan H, Jiang Z, Li N, Li J. Contrast study of short-term effect between the Da Vinci surgical robot and laparoscopic technology in patients after distal gastric cancer surgery. *Zhonguo Shiyong Waike Zazhi* 2013, **33**: 325-327

21 **Marano A**, Hyung WJ. Robotic gastrectomy: the current state of the art. *J Gastric Cancer* 2012; **12**: 63-72 [PMID: 22792518 DOI: 10.5230/jgc.2013.13.3.136]

22 **Xiong J**, Nunes QM, Tan C, Ke N, Chen Y, Hu W, Liu X, Mai G. Comparison of short-term clinical outcomes between robotic and laparoscopic gastrectomy for gastric cancer: a meta-analysis of 2495 patients. *J Laparoendosc Adv Surg Tech A* 2013; **23**: 965-976 [PMID: 24093968 DOI: 10.1089/lap.2013.0279]

**P- Reviewer:** Bener A **S- Editor:** Wen LL **L- Editor**: A
**E- Editor:** Lu YJ



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i3.107

*World J Meta-Anal* 2014 August 26; 2(3): 107-126
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Comparison of functional outcomes after retropubic, laparoscopic and robot-assisted radical prostatectomy: A meta-analysis

Ming-Jun Shi, Jie Yang, Xiang-Yu Meng, Sheng Li, Tao Liu, Zhi-Hai Fang, Rui Cao, Xing-Huan Wang

Ming-Jun Shi, Sheng Li, Tao Liu, Zhi-Hai Fang, Rui Cao, Xing-Huan Wang, Department of Urology, Zhongnan Hospital, Wuhan University, Wuhan 430071, Hubei Province, China
Jie Yang, Department of Radiation and Medical Oncology, Zhongnan Hospital of Wuhan University, Wuhan 430071, Hubei Province, China
Xiang-Yu Meng, Department of Hematology, Zhongnan Hospital of Wuhan University, Wuhan 430071, Hubei Province, China
Author contributions: Shi MJ and Yang J contributed equally to this work; Shi MJ, Yang J, Li S and Meng XY performed the literature search, study selection and data extraction; Shi MJ, Wang XH, Liu T and Fang ZH analyzed the results; Shi MJ and Cao R did the statistical analysis; Shi MJ, Yang J and Meng XY drafted the manuscript; all authors provided input into the development of the manuscript and approved the final manuscript.
Supported by National Natural Science Foundation of China, No. 81172734
Correspondence to: Xing-Huan Wang, Professor, Department of Urology, Zhongnan Hospital, Wuhan University, 169 Donghu Road, Wuchang, Wuhan 430071, Hubei Province, China. wangxinghuan1965@gmail.com
Telephone: +86-27-67813104 Fax: +86-27-67813104
Received: March 26, 2014 Revised: May 10, 2014
Accepted: June 18, 2014
Published online: August 26, 2014

## Abstract

**AIM:** To assess the 6-mo and 12-mo functional outcomes after retropubic, laparoscopic and robot-assisted laparoscopic radical prostatectomy retropubic radical prostatectomy (RRP) laparoscopic radical prostatectomy (LRP); robot-assisted laparoscopic prostatectomy (RARP).

**METHODS:** A literature search was conducted using the PubMed, EMBASE, The Cochrane Library and the Web of Knowledge databases updated to March, 2014 for relevant published studies. After data extraction and quality assessment *via* the Newcastle-Ottawa Scale or the Cochrane collaboration's tool for assessing risk of bias, meta-analysis was performed using RevMan 5.1. Either a random-effects model or a fixed-effects model was used. Potential publication bias was assessed using visual inspection of the funnel plots, and verified by the Egger linear regression test.

**RESULTS:** Thirty-seven studies were identified in total: 14 articles comparing LRP with RRP, 12 articles comparing RARP with RRP, and 11 articles comparing RARP with LRP. For urinary continence, a statistically significant advantage was observed in RARP compared with LRP or RRP both at 6 mo [odds ratio (OR) = 1.93; $P < 0.01$, OR = 2.23; $P < 0.05$, respectively] and 12 mo (OR = 1.47; $P < 0.01$, OR = 2.93; $P < 0.01$, respectively) postoperatively. The continence recovery rates after LRP and RRP, with obvious heterogeneity (6-mo: $I^2$ = 74%; 12-mo: $I^2$ = 75%), were equivalent (6-mo: $P = 0.52$; 12-mo: $P = 0.75$). In terms of potency recovery, for the first time, we ranked the three surgical approaches into a superiority level: RARP > LRP > RRP, with a statistically significant difference at 12 mo [RARP *vs* LRP (OR = 1.99; $P < 0.01$); RARP *vs* RRP (OR = 2.66; $P < 0.01$); LRP *vs* RRP (OR = 1.34; $P < 0.05$)], respectively. Meta-regression and subgroup analyses according to adjustment of the age, body mass index, prostate volume, Gleason score or prostate-specific antigen did not vary significantly.

**CONCLUSION:** Current evidence suggests that minimally invasive approaches (RARP or LRP) are effective procedures for functional recovery. However, more high-quality randomized control trials investigating the long-term functional outcomes are needed.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Prostate cancer; Radical prostatectomy; Urinary continence; Potency; Meta-analysis

**Core tip:** This review directly compared the functional

outcomes after retropubic, laparoscopic and robot-assisted radical prostatectomy, both at 6-mo and 12-mo follow-up. Compared with the previous meta-analysis which reported a comparable potency recovery of robot-assisted laparoscopic prostatectomy (RARP) *vs* laparoscopic radical prostatectomy (LRP), our review obviously included more studies and ranked the three techniques into a superiority level: RARP > LRP > RRP (retropubic radical prostatectomy). In addition, we performed a quality assessment of the studies, separated evaluation of randomized control trials (RCTs) and non-RCTs, and subgroup analyses or meta-regression as a supplement, thus the risk of methodological bias was reduced considerably.

Shi MJ, Yang J, Meng XY, Li S, Liu T, Fang ZH, Cao R, Wang XH. Comparison of functional outcomes after retropubic, laparoscopic and robot-assisted radical prostatectomy: A meta-analysis. *World J Meta-Anal* 2014; 2(3): 107-126 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i3/107.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i3.107

## INTRODUCTION

Prostate cancer (PCa) is now recognized as one of the most important medical problems in the male population. PCa accounted for almost 28% (238590) of all newly diagnosed cancer cases and it is the second cause of male cancer death (after lung cancer) in the United States, while in Europe, data show an incidence rate of 22.8% and a mortality of 9.5%[1,2]. With combined application of prostate-specific antigen (PSA) test and prostate biopsy, the percentage of early diagnosed PCa cases has increased .

Radical prostatectomy (RP) is one of the recommended standard treatments for clinically localized prostate cancer (cT1-cT2) patients with a life expectancy of more than 10 years[3]. The retropubic radical prostatectomy (RRP), since its first introduction by Walsh *et al*[4] in 1982, soon became the gold standard and the most widely used treatment for patients with localized PCa[5]. Recently, we have witnessed the emergence of laparoscopic radical prostatectomy (LRP) and robot-assisted laparoscopic prostatectomy (RARP). Facing all these surgical options, both patients and surgeons hesitate when a best treatment choice should be made. Although several experts have demonstrated that when compared with RRP, LRP and RARP have obvious advantages such as lower blood loss, less need for transfusion and shorter hospital-stay[6,7], but the lack of high-quality evidence and randomized control trials (RCTs) available precluded us from proving the superiority of any surgical approaches in terms of postoperative functional outcomes.

The increase in life expectancy in patients with localized PCa has made the post-treatment quality of life a key issue for PCa survivors, but some negative functional outcomes such as urinary incontinence and erectile dysfunction make the health-related quality of life worse. Relevant comparative studies showed 12-mo urinary continence recovery rates ranging from 47% to 96%, 48% to 97% and 88% to 97% after RRP, LRP and RARP, respectively. The previously published surgical series showed 12-mo potency recovery rates ranging from 39% to 72%, 41% to 81% and 61% to 87% after RRP, LRP and RARP, respectively. This apparent difference can be attributed to multiple definitions of urinary continence and potency, variations in population baseline, differences among surgical techniques and diverged data collection as well. In comparison with the only two meta-analyses evaluating functional outcomes after different surgical approaches, reported by the same author Ficarra *et al*[8,9] in August 2011, obviously our review included more studies and excluded two studies[10,11] which appeared to be ineligible since the presence of preoperative adjuvant hormonal therapy. Moreover, powerful quality assessment tools were utilized in this initial comparison of three key techniques (RRP, LRP and RARP) both at 6-mo and 12-mo follow-up.

## MATERIALS AND METHODS

### Literature search

A literature search of the following databases was performed: the PubMed, EMBASE, The Cochrane Library and the Web of Knowledge databases up to March, 2014. We used the following limits: humans, gender (male), and no restriction for languages. For each database, the same search terms "radical prostatectomy", "urinary continence", "incontinence", "potency" and "erectile function" were used. Although we also paid attention to two unpublished studies (gray literature) with relevant outcomes reported on the website "Clinical Trials.gov" and tried to contact the experts by e-mail, there has been no response so far, and therefore in this review only published papers were included.

### Study selection

Our study followed the preferred reporting items for meta-analyses of observational studies in epidemiology (MOOSE) statement[12]. The inclusion criteria were as follows: (1) patient characteristics: localized PCa (cT1-cT2); comparable baseline demographics; preoperatively potent and continent; no obvious comorbidities; (2) surgical techniques: only pure RRP/RARP/LRP with or without modification; (3) methodologically: all studies comparing the postoperative outcomes as RRP/LRP, RRP/RARP or LRP/RARP and including at least one of the functional results; clear definition of urinary continence and potency; and (4) population-based studies, duplicated publications and meeting abstracts were excluded.

### Data extraction

All eligible records were extracted independently by two reviewers and selected according to the inclusion criteria. We extracted the details of authors and publishing date;

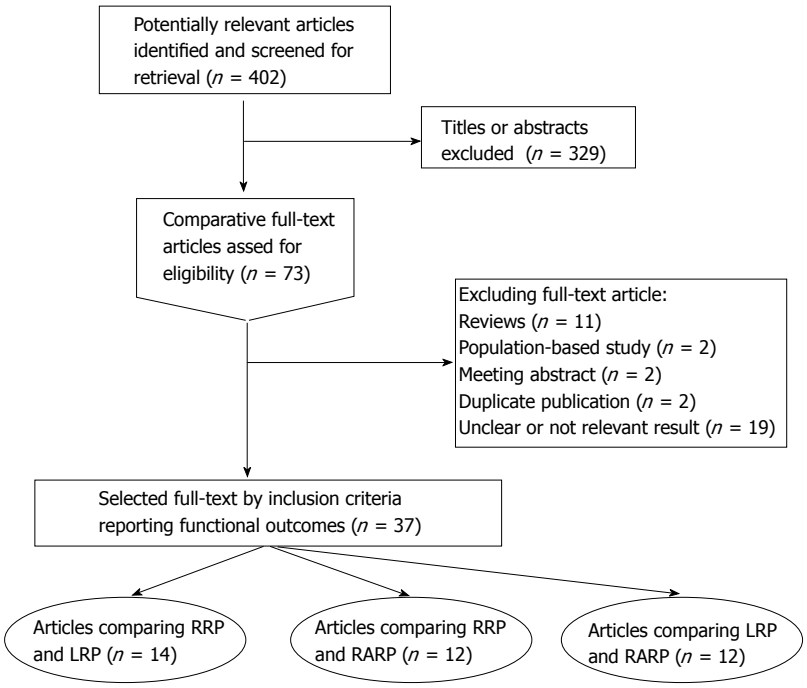



**Figure 1 Flow diagram of the systematic review.** RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy; RARP: Robot-assisted radical prostatectomy.

surgical techniques and number of patients; study design; baseline mean age; BMI value; prostate volume; PSA level; urinary continence and potency definition; and 6- and 12-mo recovery rates of urinary continence and potency. Any uncertainties or discrepancies between the two reviewers were resolved by open discussion or consultation with the third reviewer.

### Methodological quality assessment

The quality of cohort and case-control studies was assessed using the Newcastle-Ottawa quality scale (NOS) proposed by Wells *et al*[13]. This tool can be used either as a checklist or as a scale. The NOS scales were separately developed for cohort and case-control studies. Briefly, a star system was used for quality assessment of studies, and the NOS ranges from zero up to nine stars; studies were evaluated using items from three broad perspectives: selection of study groups (0-4 stars), comparability between groups (0-2 stars), and ascertainment of either the exposure or the outcome of interest (0-3 stars) for case-control or cohort studies, respectively.

The quality of each RCT was assessed using the Cochrane collaboration's tool for assessing risk of bias[14], which utilizes seven aspects: (1) details of randomization method; (2) allocation concealment; (3) blinding of participants and personnel; (4) blinding of outcome assessment; (5) incomplete outcome data; (6) selective outcome reporting; and (7) other sources of bias, to provide a qualification of risk of bias.

### Statistical analysis

Statistical analyses were performed using the Cochrane Review Manager (RevMan) Version 5.1 software. Odds ratios (ORs) with 95% confidence intervals (CIs) for dichotomous variables were computed as summary statistics. According to the Higgins' $I^2$ statistic, a statistical heterogeneity of < 25, 25-50, and > 50% was defined as low, moderate, and high, respectively[15]. If no heterogeneity was found, a fixed-effects model using the Mantel-Haenszel method would be used[16,17]. If statistically significant heterogeneity was revealed, a random-effects model would be used[18]. The sensitivity analysis was also performed by two methods: (1) subgroup analysis, and (2) exclusion of the study accounting for the largest proportion; if no difference was detected then it could be confirmed that the outcomes were stable and reliable. The meta-regression analyses were performed by modeling on binary continence and potency outcomes, adjusting the age, BMI, prostate volume, mean Gleason score, and PSA level by using the STATA SE 12.0. For all statistical analyses, a *P*-value < 0.05 was set as the level of significance. The publication bias was examined using the funnel plot, the results of which were further verified by Egger's test[19].

## RESULTS

### Study identification

Figure 1 shows the flowchart of this review and summarizes the number of potential citations (Figure 1). The authors selected 73 full-text articles after a comprehensive review of 402 potential relevant citations. Among these, 14 articles compared RRP with LRP, consisting of seven prospective and seven retrospective studies[20-33]; 12 articles compared RRP with RARP, which consisted of six prospective and six retrospective studies[10,11,34-43]; 12 articles compared LRP with RARP, including two RCTs, one prospective and nine retrospective studies[39,44-54].

### Quality of studies

Totally, there were 14 prospective studies and 21 retrospective studies included in this review. According to the

**Table 1 Comparative studies evaluating urinary continence recovery after retropubic radical prostatectomy or laparoscopic radical prostatectomy**

| Quality | Case, *n* | Ref. | Country | Age (yr) | BMI (kg/m$^2$) | Prostate volume (mL, g) | Gleason score (biopsy) | PSA (ng/mL) | Study design | Continence definition | Data collection | Loss of follow-up (N/Y, %) | Urinary continence recovery, %(*n*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | 6 mo | 12 mo |
| 3/2/2(H) | RRP, 70 | Anastasiadis *et al*[20], 2003 | France | 64.8 ± 6.4 | - | - | 6.1 ± 1.1 | 11.2 ± 9.7 | Prospective | 0 pad | Nonvalidated questionnaire | Y, > 20% | 43.3 (16/37) | 77.7 (26/33) |
| | LRP, 230 | | | 64.1 ± 6.4 | | | 5.8 ± 1.2 | 10.7 ± 8.8 | | | | | 59.2[a] (67/113) | 89.0 (94/106) |
| 2/2/3(H) | RRP, 77 | Roumeguere *et al*[21], 2003 | Belgium | 63.9 ± 5.5 | - | 42.0 ± 20.4 | 5.4 ± 1.5 | 10.5 ± 11.5 | Prospective | 0 pad | Interview | Y, > 20% | 62.5 (40/64) | 83.9 (47/56) |
| | LRP, 85 | | | 62.5 ± 6.0 | | 37.3 ± 15.6 | 5.4 ± 1.5 | 8.6 ± 5.2 | | | | | 50.6 (37/73) | 80.7 (42/52) |
| 3/1/3(H) | RRP, 41 | Remzi *et al*[22], 2005 | Austria | 60 ± 14 | - | 44 ± 18 | 4.7 ± 1.5 | 6.9 ± 4.4 | Prospective | 0 pad | Physician | N | - | 80.3 (33/41) |
| | (a)tLRP, 39 | | | 61 ± 11 | | 37 ± 16 | 5.1 ± 1.2 | 5.5 ± 3.7 | | | | | | 84.6 (33/39) |
| | (b)eLRP, 41 | | | 59 ± 12 | | 32 ± 14 | 5.5 ± 1.3 | 8.1 ± 6.1 | | | | | | 87.8 (36/41) |
| 3/2/3(H) | RRP, 75 | Wagner *et al*[23], 2007 | United States | 59 ± 6.9 | 29 ± 4.5 | - | - | 8.1 ± 6.27 | Prospective | 0 pad | EPIC | Y, < 20% | - | 47.0 (31/66) |
| | LRP, 75 | | | 58 ± 6.9 | 27 ± 3.0 | | | 6.2 ± 4.22 | | | | | | 64.0[a] (43/67) |
| 3/2/2(H) | RRP, 222 | Touijer *et al*[24], 2008 | United States | 59 (54, 64) | - | - | - | 5.3 (4.1, 7.3) | Prospective | 0-1 safety pad | Institutional questionnaire | N | - | 75.0[a] (167/222) |
| | LRP, 193 | | | 60 (55, 65) | | | | 5.3 (4.0, 7.5) | | | | | | 48.0 (93/193) |
| 3/2/3(H) | RRP, 150 | Greco *et al*[25], 2009 | Italy | 61.5 (49-74) | 29 (25-33) | - | 5 (3-7) | 6.95 (3.4-10) | Prospective | 0 pad | Validated questionnaire | N | 76.0 (114/150) | 91.0 (137/150) |
| | LRP, 150 | | | 60.5 (45-76) | 32 (26-38) | | 5 (3-7) | 6.3 (2.4-10) | | | | | 89.3 (134/150) | 97.0 (146/150) |
| 3/2/2(H) | RRP, 102 | Dahl *et al*[26], 2009 | United States | 59.9 | - | - | - | - | Prospective | 0 pad | Validated questionnaire | Y, > 20% | 49.0 (38/78) | 49.0 (35/72) |
| | LRP, 104 | | | 59.5 | | | | | | | | | 42.0 (31/74) | 53.0 (41/78) |
| 2/2/2(M) | RRP, 49 | Egawa *et al*[27], 2003 | Japan | 67.0 ± 0.7 | - | - | 6.0 ± 0.2 | 8.3 ± 1.4 | Retrospective | 0 pad | Interview | Y, > 20% | 84.1[a] (37/44) | 92.9[a] (39/42) |
| | LRP, 34 | | | 68.0 ± 0.9 | | | 5.0 ± 0.2 | 6.6 ± 0.6 | | | | | 46.9 (15/32) | 60.0 (12/20) |
| 3/1/2(M) | RRP, 50 | Artibani *et al*[28], 2003 | Italy | 64.28 ± 6.6 | - | - | 5.7 ± 1.2 | 11 ± 9 | Retrospective | 0 pad | Nonvalidated questionnaire | Y, > 20% | - | 64.0 (9/14) |
| | LRP, 71 | | | 63.14 ± 5.8 | | | 5.8 ± 1.3 | 15.7 ± 17 | | | | | | 40.0 (8/20) |
| 4/2/2(H) | RRP, 70 | Ghavamian *et al*[29], 2006 | United States | 57.8 ± 7.3 | 28.1 | 53.2 (19-135) | 6.7 ± 1.3 | 9.9 ± 7.1 | Retrospective | 0 pad | Physician | Y, < 20% | 71.4 (50/70) | 87.6 (57/65) |
| | LRP, 70 | | | 60.8 ± 6.1 | 27.5 | 40.8 (20-114) | 6.4 ± 0.8 | 7.6 ± 8.0 | | | | | 70.0 (49/70) | 90.0 (63/70) |
| 4/2/2(H) | RRP, 37 | Takenaka *et al*[30], 2008 | Japan | 67.1 ± 6.0 | 23.5 ± 3.0 | 30.1 ± 26.9 | 6.9 ± 1.0 | 14.7 ± 11.9 | Retrospective | 0 pad | Nonvalidated questionnaire | N | 77.0 (28/37) | 91.0 (34/37) |
| | LRP, 109 | | | 66.1 ± 6.3 | 23.8 ± 2.5 | 32.2 ± 16.5 | 6.6 ± 0.7 | 11.0 ± 8.4 | | | | | 65.0 (71/109) | 77.0 (84/109) |
| 2/2/3(H) | RRP, 188 | Simforoosh *et al*[31], 2009 | Iran | 62.1 (45-74) | - | - | - | 13.6 | Retrospective | 0 pad | Physician | N | 91.5 (172/188) | 95.2 (179/188) |
| | LRP, 136 | | | 62.5 (45-76) | | | | 12.7 | | | | | 89.0 (121/136) | 96.3 (131/136) |
| 2/1/1(M) | RRP, 128 | Springer *et al*[32], 2013 | Germany | 57.2 ± 7.4 | 28.3 ± 2.6 | - | - | 3.1 ± 1.7 | Retrospective | 0 pad | Validated questionnaire | N | 73.4 (94/128) | 86.4 (111/128) |
| | LRP, 125 | | | 56.8 ± 6.7 | 27.7 ± 3.8 | | | 3.2 ± 1.4 | | | | | 86.4 (108/125)[a] | 96.8[a] (121/125) |
| 3/2/2(H) | RRP, 168 | Magheli *et al*[33], 2014 | Germany | 62.6 ± 5.4 | - | 58 ± 22 | - | 10.1 ± 11.9 | Retrospective | 0-1 safety pad | Validated questionnaire | Y, > 20% | - | 83.2 (99/119) |
| | LRP, 171 | | | 62.3 ± 5.7 | | 53 ± 20 | | 9.2 ± 6.9 | | | | | | 82.8 (96/116) |

[a]$P < 0.05$. RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy; (a)tLRP: Transperitoneal laparoscopic radical; EPIC: Expanded prostate cancer index composite.

NOS scale (case-control studies) used for quality evaluation of the retrospective studies, 12 studies were in the high level (7-9 stars)[29-31,33,41-43,48,50-52,54], one study was in the low level (0-3 stars)[11], and the remaining eight studies were in the middle level (4-6 stars). As for the quality of the prospective studies, the NOS scale (cohort studies) was used, and 12 studies were in the high level (7-9 stars)[20-26,34-36,38,46], one study was in the middle level (4-6 stars)[37], and one study was in the low level (0-3 stars)[10].

The only available two RCTs were considered high quality by using the Cochrane collaboration's tool for assessing risk of bias.

### *Characteristics of included studies and meta-analyses on urinary continence recovery*

Table 1 summarizes the results of urinary continence recovery rate between LRP and RRP. Among the 12 studies[20-33], 1427 patients treated with RRP and 1633 patients treated with LRP were included. Most of the selected studies had a very strict urinary continence definition as no pad. Only seven studies[25-27,29-32] provided the 6-mo urinary continence

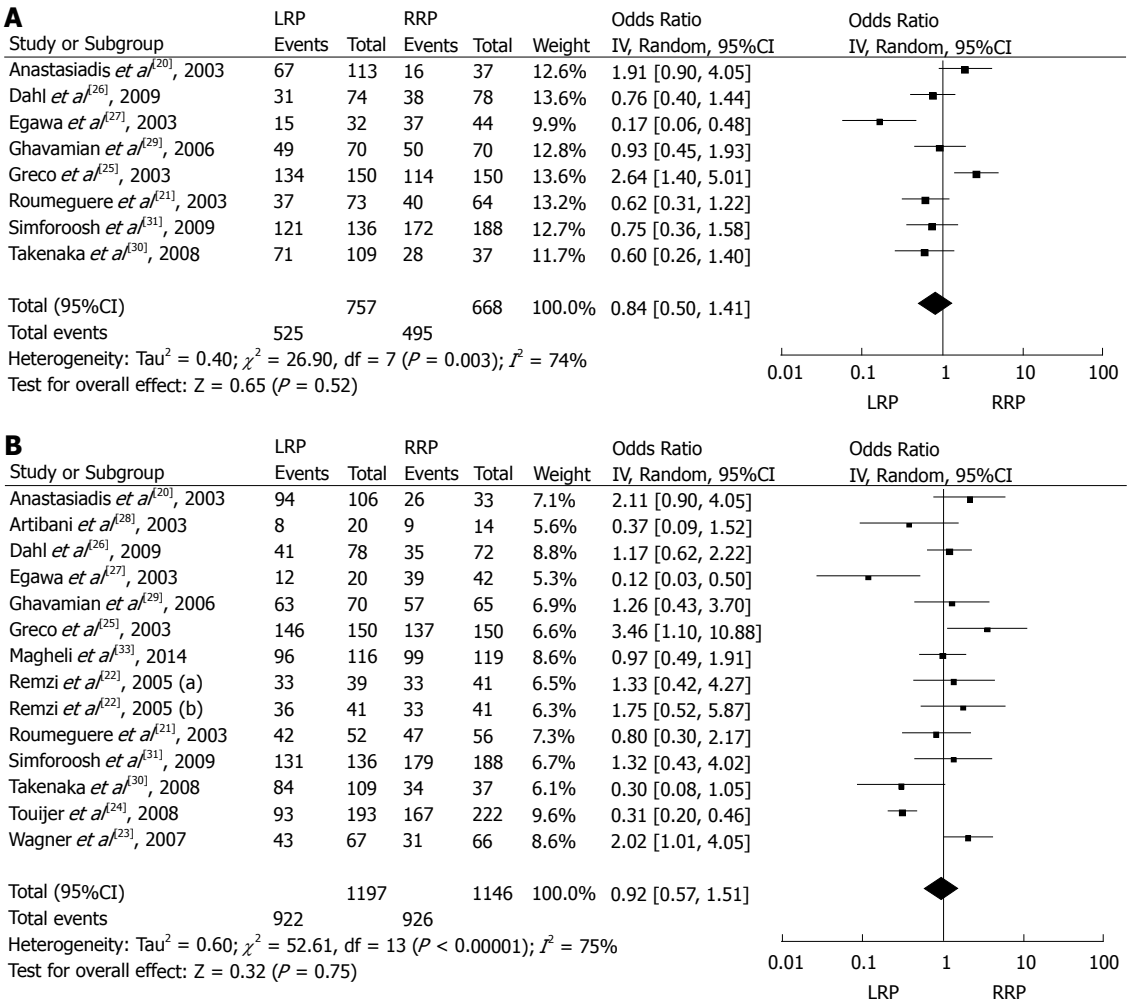



**Figure 2 Forest plots and meta-analyses of laparoscopic radical prostatectomy and retropubic radical prostatectomy.** A: 6-mo continence recovery; B: 12-mo continence recovery. RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy.

rate. The 12-mo loss to follow-up rate was > 20% in six studies[20-21,26-28,33]. Although Springer *et al*[32] reported demonstrated a significant better outcome of LRP than ORP (96.8% *vs* 86.4%, $P < 0.05$), we did not include it in because of the preoperatively performed transurethral resection of the prostate in that report, which could potentially be an inconsistent factor among the groups. The mean urinary continence recovery rates at 6 and 12 mo were 56.6% (42%-70%) and 84.3% (48.0%-96.3%) after LRP; and 64.9% (43.3%-84.1%) and 77.8% (47.0%-95.2%) after RRP, respectively.

**Six-month continence recovery after LRP and RRP**: Statistically high heterogeneity ($I^2 = 74\%$, $P < 0.05$) was observed among the eight studies[20-21,25-27,29-31] included. The meta-analysis with a random-effects model showed no significant difference between LRP and RRP (OR = 0.84; 95%CI: 0.50-1.41; $P = 0.52$) (Figure 2).

**Twelve-month continence recovery after LRP and RRP:** Fourteen studies were included in the meta-analysis[20-31,33], and there was a statistical heterogeneity ($I^2 = 75\%$, $P < 0.05$). No significant difference was found between LRP and RRP by using a random-effects model (OR = 0.92; 95%CI: 0.57-1.51; $P = 0.75$) (Figure 2).

Table 2 summarizes the results of urinary continence recovery rate between RARP and RRP. A total of 1942 patients who received RRP and 1882 patients who received RARP were included. Half of the included studies had a very strict urinary continence definition as no pad. Only two studies[37,40] had a high loss to follow-up rate (> 20%) at 12 mo. Tewari *et al*[34] reported that the median urinary continence recovery was significantly better after RARP compared with after RRP (44 d *vs* 160 d, $P < 0.05$), and Kim *et al*[10] drew the same conclusion, while Krambeck *et al*[11] presented an opposite result in the comparison of RARP and RRP (91.8% *vs* 93.7%, respectively). However, compared with the previous meta-analysis[8], Kim *et al*[10] and Krambeck *et al*[11] results were excluded in our review because of their preoperative adjuvant hormonal therapy, which would undoubtedly cause difference.

**Six-month continence recovery after RARP and RRP**: Statistically significant heterogeneity was observed among the eight included studies ($I^2 = 73\%$, $P < 0.05$)[36-43],

**Table 2 Comparative studies evaluating urinary continence recovery after retropubic radical prostatectomy or robot-assisted radical prostatectomy**

| Quality | Case, *n* | Ref. | Country | Age (yr) | BMI (kg/m$^2$) | Prostate volume (mL, g) | Gleason score (biopsy) | PSA (ng/mL) | Study design | Continence definition | Data collection | Loss of follow-up (N/Y,%) | Urinary continence recovery, %(*n*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | 6 mo | 12 mo |
| 3/2/3(H) | RRP, 100<br>RARP, 200 | Tewari *et al*[34], 2003 | United States | 63.1 (42.8-72)<br>59.9 (40-72) | 27.6 (17-41)<br>27.7 (19-38) | 48.4 (24.2-70)<br>58.8 (18-140) | - | 7.3 (1.9-35)<br>6.4 (0.6-41) | Prospective | 0-1 safety pad | Interview | - | Median:160 d<br>Median:44 d[a] | |
| 3/2/2(H) | RRP, 105<br>RARP, 103 | Ficarra *et al*[35], 2008 | Italy | 65 (61-69)<br>61 (57-67) | 26 (24-28)<br>26 (24-28) | 40 (30-47)<br>37.5 (30-48) | - | 6 (5-10)<br>6.4 (4.6-9) | Prospective | 0 pad | ICIQ-UI | N | - | 88.0 (92/105)<br>97.0[a] (100/103) |
| 3/2/3(H) | RRP, 110<br>RARP, 188 | Ham *et al*[36], 2008 | South Korea | 66.9 ± 6.0<br>67.3 ± 6.2 | 23.6 ± 1.8<br>23.6 ± 2.3 | - | - | 55.2 ± 23.7<br>22.3 ± 34.3 | Prospective | 0 pad | Validated questionnaire | N | 75.5 (83/110)<br>87.2 (164/188) | 81.8 (90/110)<br>92.0[a] (173/188) |
| 3/1/2(M) | RRP, 75<br>RARP, 75 | Di Pierro *et al*[37], 2010 | Switzerland | 64.3 (59.1-68.0)<br>62.8 (58.4-67.0) | - | - | - | 7.57 (5.1-10.4)<br>7.72 (5.6-12.1) | Prospective | 0 pad | Institutional questionnaire | Y, > 20% | 83.0 (62/75)<br>95.0[a] (71/75) | 80.0 (60/75)<br>89[a] (40/45) |
| 1/1/1(L) | RRP, 235<br>RARP, 528 | Kim *et al*[10], 2011 | South Korea | 66.5 ± 5.7<br>64.2 ± 7.3 | - | 18.2 ± 23.4<br>15.2 ± 20.2 | - | 14.6 ± 22.1<br>10.4 ± 16.0 | Prospective | 0 pad | Validated questionnaire | - | Median: 4.3 mo<br>Median: 3.7 mo | |
| 4/2/3(H) | RRP, 109<br>RARP, 61 | Geraerts *et al*[38], 2013 | Belgium | 62.22 ± 6.12<br>61.48 ± 6.08 | - | - | - | - | Prospective | 24h pad test | Validated questionnaire | N | 94.0 (102/109)<br>95.0 (58/61) | 96.0 (105/109)<br>97.0 (59/61) |
| 2/1/2(M) | RRP, 62<br>RARP, 60 | Caballero *et al*[39], 2008 | Spain | 66.5 (62-69)<br>56 (56-65.25) | - | 41 (30.15-52)<br>29.5 (23-40) | - | 9.66 (7-16.6)<br>7 (5.7-10) | Retrospective | 0 pad | Unspecified | Y, < 20% | 45.9 (28/61)<br>60.0 (30/50)[a] | - |
| 2/0/1(L) | RRP, 588<br>RARP, 294 | Krambeck *et al*[11], 2008 | United States | 61.0 (41.0-77.0)<br>61.0 (38.0-76.0) | - | - | - | 5.0 (0.6-39.7)<br>4.9 (0.5-33.5) | Retrospective | 0 pad | Institutional questionnaire | Y, < 20% | | 93.7 (446/476)<br>91.8 (224/244) |
| 3/1/2(M) | RRP, 240<br>RARP, 120 | Rocco *et al*[40], 2009 | Italy | 63 (46-77)<br>63 (47-76) | - | - | 6 (4-10)<br>6 (4-9) | 6.7 (0.7-22.0)<br>6.9 (0.4-23.0) | Retrospective | 0-1 safety pad | Interview | Y, > 20% | 83.0 (189/229)<br>93.0[a] (102/110) | 88.0 (191/217)<br>97.0[a] (77/79) |
| 3/1/3(H) | RRP, 30<br>RARP, 30 | Ou *et al*[41], 2009 | United States | 70.03 ± 6.10<br>67.27 ± 6.21 | 24.09 ± 3.28<br>24.22 ± 3.16 | 15.89 ± 14.15<br>16.45 ± 18.80 | 6.22 ± 1.62<br>6.13 ± 0.9 | - | Retrospective | 0-1 safety pad | Unspecified | N | 83.3 (25/30)<br>96.7 (29/30) | 96.6 (29/30)<br>100.0 (30/30) |
| 3/2/3(H) | RRP, 176<br>RARP, 77 | Choo *et al*[42], 2013 | South Korea | 67 ± 6.25<br>66 ± 7.75 | 24 ± 2.73<br>24 ± 2.55 | 42 ± 18.82<br>41 ± 15.77 | - | 7.6 ± 19.33<br>7.2 ± 13.19 | Retrospective | 0-1 safety pad | Validated questionnaire | N | 92.0 (162/176)<br>84.0 (65/77) | 96.0 (169/176)<br>94.0 (72/77) |
| 3/1/3(H) | RRP, 112<br>RARP, 146 | Son *et al*[43], 2013 | South Korea | 65.0 ± 6.1<br>65.5 ± 6.7 | 24.3 ± 2.4<br>24.5 ± 2.5 | 41.3 ± 30.0<br>45.9 ± 16.3 | - | - | Retrospective | 0 pad | Validated questionnaire | Y, < 20% | 51.7 (49/94)<br>87.5[a] (107/122) | 70.7 (66/94)<br>94.5[a] (115/122) |

[a] $P < 0.05$. RRP: Retropubic radical prostatectomy; RARP: Robot-assisted radical prostatectomy; ICIQ-UI: International consultation of incontinence questionnaire-urinary incontinence.

and the meta-analysis with a random-effects model showed a significant advantage after RARP than after RRP (OR = 2.23; 95%CI: 1.20-4.14; $P < 0.05$) (Figure 3).

**Twelve-month continence recovery after RARP and RRP:** Totally, eight studies were included to compare RARP and RRP[35-38,40-43]. No evidence of statistical heterogeneity was observed ($I^2 = 49\%$, $P = 0.06$) and pooled analysis with a fixed-effects model demonstrated a statistically significant advantage in favor of RARP (OR = 2.93; 95%CI: 1.99-4.32; $P < 0.01$) (Figure 3).

Table 3 summarizes the results of urinary continence recovery rate between LRP and RARP. A total of 2195 patients treated with LRP and 1940 patients treated with RARP were included. Both of the only two RCTs (high quality) revealed that the urinary continence recovery rates were significantly higher at 6 and 12 mo after RARP, in comparison with those after LRP ($P < 0.05$)[44,45]. The evidence with the largest sample size, reported by Ploussard *et al*[46], was the only prospective study of high quality (7 stars) and showed similar results with the two RCTs. Almost all of the remaining retrospective studies also indicated better outcomes after RARP. Only one study[52] had a high loss to follow-up rate (> 20%) at 12 mo. The most crucial difference between our pooled-analysis and the previous meta-analysis[8] was that the RCTs were evaluated separately with non-RCTs (NRCTs) in this review, since they had totally different data types.

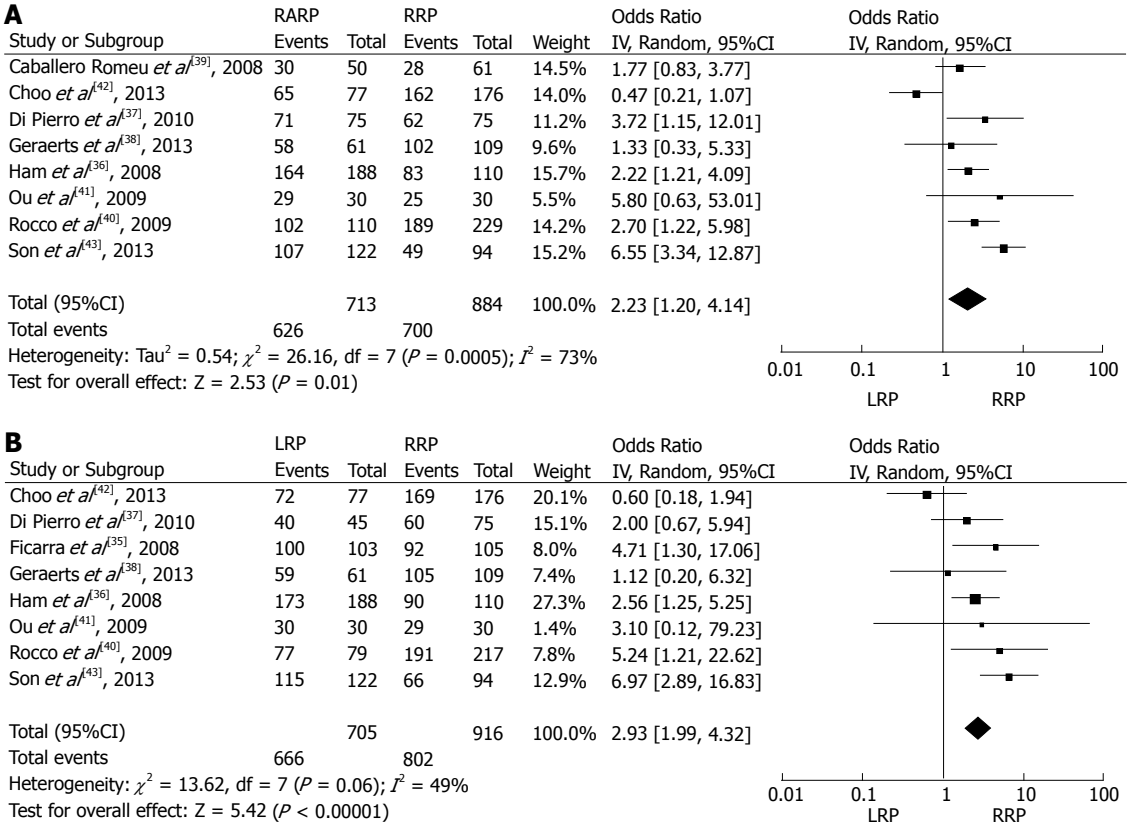



**Figure 3 Forest plots and meta-analyses of robot-assisted radical prostatectomy and retropubic radical prostatectomy.** A: 6-mo continence recovery; B: 12-mo continence recovery. RRP: Retropubic radical prostatectomy; RARP: Robot-assisted radical prostatectomy.

**Six-month continence recovery after RARP and LRP**: The two RCTs[44,45] showed no heterogeneity ($I^2$ = 0%, $P$ = 0.92), and supported the advantage after RARP with a fixed-effects model (OR = 2.66; 95%CI: 1.31-5.40; $P$ < 0.01) (Figure 4). In the cumulative analysis of 10 NRCTs[39,46-54], no heterogeneity was found ($I^2$ = 38%, $P$ = 0.11), so a fixed-effects model was performed. The result also demonstrated a statistically significant advantage in favor of RARP (OR = 1.93; 95%CI: 1.67-2.23; $P$ < 0.01) (Figure 4).

**Twelve-month continence recovery after RARP and LRP:** No evidence of statistical heterogeneity was observed in both of the two RCTs ($I^2$ = 0%, $P$ = 0.88) or the seven NRCTs ($I^2$ = 0%, $P$ = 0.44), and the pooled analyses with a fixed-effects model either for the RCTs or the NRCTs showed a statistically significant advantage in favor of RARP [(OR = 3.52; 95%CI: 1.36-9.13; $P$ < 0.05); (OR = 1.47; 95%CI: 1.25-1.74; $P$ < 0.01), respectively] (Figure 4).

### *Characteristics of included studies and meta-analyses on potency recovery*

Table 4 summarizes the results of potency recovery rate between LRP and RRP. Among the 10 studies, 907 patients treated with RRP and 1004 patients treated with LRP were included. Eight of them had a very strict potency definition as erection sufficient for intercourse (ESI). The 12-mo loss to follow-up rate was > 20% in three studies[20,26,33]. Springer *et al*[32] report was not included in the meta-analysis because of its preoperative surgery. The nerve sparing (NS) procedures were not clearly mentioned in two studies[24,26], and the remaining studies either used the bilateral or unilateral nerve sparing measures. The mean potency recovery rates at 6 and 12 mo were 30.6% (23.0%-38.1%) and 45.8% (32.0%-54.5%) after RRP; and 42.5% (37%-48%) and 55% (41%-66%) after LRP, respectively.

**Six-month potency recovery after LRP and LRP**: No statistical heterogeneity was observed ($I^2$ = 0%, $P$ = 0.67) in the included four studies[21,26,28,29]. The meta-analysis evaluating potency with a fixed-effects model suggested no statistically significant difference between LRP and RRP (OR = 1.48; 95%CI: 0.94-2.34; $P$ = 0.09) (Figure 5).

**Twelve-month potency recovery after LRP and LRP:** Eight studies were included[20-21,23-26,29,33] and no statistical heterogeneity was observed ($I^2$ = 0%, $P$ = 0.50). The pooled analysis with a fixed-effects model showed a statistically significant advantage in favor of LRP (OR = 1.34; 95%CI: 1.05-1.70; $P$ < 0.05) (Figure 5).

Table 5 summarizes the results of potency recovery rate after RARP and RRP. A total of 1278 patients treated with RRP and 1309 patients treated with RARP were included. In half of them, the NS procedures were not

**Table 3 Comparative studies evaluating urinary continence recovery after laparoscopic radical prostatectomy or robot-assisted radical prostatectomy**

| Quality | Case, *n* | Author, yr | Country | Age (yr) | BMI (kg/m$^2$) | Prostate volume (mL,g) | Gleason score (biopsy) | PSA (ng/mL) | Study design | Continence definition | Data collection | Loss of follow-up (N/Y, %) | Urinary continence recovery, %(*n*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | 6 mo | 12 mo |
| High | LRP, 60 | Asimakopoulos *et al*[44], 2011 | Italy | 61.1 ± 5.1 | 26.3 ± 2.2 | - | - | 7.37 (1.5-9.15) | RCT | 0 pad | ICS-MSF | N | 75.0 (45/60) | 83.0 (50/60) |
| | RARP, 52 | | | 59.6 ± 5.4 | 25.8 ± 2.6 | | | 8.9 (5.8-9.2) | | | | | 88.0 (46/52) | 94.0 (49/52) |
| High | LRP, 60 | Porpiglia *et al*[45], 2012 | Italy | 64.7 ± 5.9 | 26.8 ± 2.9 | 37.7 ± 14.1 | - | 8.3 ± 6.5 | RCT | 0-1 pad | EPIC | N | 73.3 (44/60) | 83.3 (50/60) |
| | RARP, 60 | | | 63.9 ± 6.7 | 26.2 ± 2.5 | 36.2 ± 12.6 | | 6.9 ± 4.2 | | | | | 88.3a (53/60) | 95.0[a] (57/60) |
| 3/1/3(H) | LRP, 1377 | Ploussard *et al*[46], 2012 | France | 62.7 | 26.6 | - | - | 9.8 | Prospective | 0 pad | Validated questionnaire | N | 58.9 (811/1377) | 68.5 (943/1377) |
| | RARP, 1009 | | | 62.7 | 26.5 | | | 9.2 | | | | | 72.0[a] (726/1009) | 75.4 (761/1009) |
| 2/1/2(M) | LRP, 50 | Joseph *et al*[47], 2005 | United Kingdom | 61.8 ± 1.6 | - | - | 6 ± 0.14 | 6.0 ± 0.83 | Retrospective | 0 pad | Interview | N | 92.0 (46/50) | - |
| | RARP, 50 | | | 59.6 ± 1.6 | | | 6 ± 0.15 | 7.3 ± 1.2 | | | | | 90.0 (45/50) | |
| 2/1/2(M) | LRP, 70 | Caballero *et al*[39], 2008 | Spain | 66.5 (62-69) | - | 41 (30.15-52) | - | 9.66 (7-16.6) | Retrospective | 0 pad | Unspecified | Y, < 20% | 36.4 (24/66) | - |
| | RARP, 60 | | | 56 (56-65.25) | | 29.5 (23-40) | | 7 (5.7-10) | | | | | 60.0 (30/50) | |
| 3/1/3(H) | LRP, 31 | Lee *et al*[48], 2009 | South Korea | 63.0 ± 8.52 | 25.2 ± 2.59 | 37.4 ± 13.05 | 6.5 ± 1.23 | 11.7 ± 13.72 | Retrospective | 0-1 safety pad | Institutional questionnaire | N | 80.6 (25/31) | - |
| | RARP, 21 | | | 64.6 ± 6.79 | 25.5 ± 2.64 | 39.9 ± 15.54 | 6.6 ± 0.97 | 8.1 ± 7.01 | | | | | 81.0 (17/21) | |
| 3/1/2(M) | LRP, 60 | Cho *et al*[49], 2009 | South Korea | 66.5 (57-75) | 23.65 (18.1-28.4) | 39.7 (19-72) | 6.81 (5-9) | 11.04 (2.72-36.6) | Retrospective | 0-1 safety pad | Interview | N | 71.7 (43/60) | 100.0 (60/60) |
| | RARP, 60 | | | 66.3 (50-77) | 24.61 (19.9-26.3) | 36.6 (22-92.8) | 6.83 (5-8) | 9.98 (2.91-26.3) | | | | | 93.3 (56/60) | 100.0 (60/60) |
| 4/2/3(H) | LRP, 75 | Hakimi *et al*[50], 2009 | United States | 59.6 (43-72) | - | - | - | 7.5 | Retrospective | 0 pad | IPSS | N | 65.3 (49/75) | 89.3 (67/75) |
| | RARP, 75 | | | 59.8 (42-71) | | | | 8.4 | | | | | 74.7 (56/75) | 93.3 (70/75) |
| 4/2/2(H) | LRP, 45 | Trabulsi *et al*[51], 2010 | United States | 58.1 (43-74) | - | - | - | 6.2 | Retrospective | 0-1 safety pad | Validated questionnaire | N | 71.0 (32/45) | 82.0 (37/45) |
| | RARP, 205 | | | 59.9 (42-76) | | | | 6.4 | | | | | 91.0[a] (187/205) | 94.0[a] (193/205) |
| 3/2/2(H) | LRP, 161 | Willis *et al*[52], 2011 | United States | 58.0 ± 6.7 | 27.0 ± 3.4 | 35.2 ± 10.1 | - | 5.7 ± 2.9 | Retrospective | 0 pad | Validated questionnaire | Y, > 20% | 55.0 (64/117) | 72.0 (84/116) |
| | RARP, 121 | | | 58.1 ± 6.3 | 26.7 ± 3.3 | 41.5 ± 15.2 | | 5.0 ± 2.2 | | | | | 66.0 (50/76) | 75.0 (33/44) |
| 3/1/2(M) | LRP, 62 | Park *et al*[53], 2011 | South Korea | 65.7 (38-77) | 24.6 (19.4-31.4) | 30.1 (12.0-56.0) | - | 9.14 (2.65-30.77) | Retrospective | 0-1 safety pad | Interview | N | 76.3 (47/62) | 95.0 (59/62) |
| | RARP, 44 | | | 62.7 (46-71) | 26.0 (19.7-39.4) | 32.9 (15.5-66.8) | | 6.32 (1.86-29.5) | | | | | 93.5 (41/44) | 94.4 (42/44) |
| 3/2/3(H) | LRP, 144 | Park *et al*[54], 2013 | South Korea | 67 (38-77) | 24.2 (17.2-31.4) | 28.8 (12.0-74.0) | - | 5.84 (0.08-41.26) | Retrospective | 0 pad | Interview | N | 65.5 (94/144) | 78.1 (112/144) |
| | RARP, 183 | | | 63 (44-75) | 24.7 (16.4-39.4) | 30.3 (15.5-82.8) | | 4.98 (0.05-51.46) | | | | | 83.5[a] (153/183) | 87.4 (160/183) |

[a]$P < 0.05$. LRP: Laparoscopic radical prostatectomy; RARP: Robot-assisted radical prostatectomy; RCT: Randomized controlled trial; IPSS: International prostate symptom score; EPIC: Expanded prostate cancer index composite; ICS-MSF: International continence society-male short form questionnaire.

clearly mentioned. Three studies[11,37,40] had a high loss to follow-up rate (> 20%) at 12 mo. Tewari *et al*[34] reported that the median potency recovery was significantly better after RARP than after RRP (180 d *vs* 440 d, $P < 0.05$). The mean 12-mo potency recovery rate ranged from 40% to 50% after RRP and from 54% to 87.5% after RARP.

**Six-month potency recovery after RARP and RRP**: A statistically significant heterogeneity was observed among the three included studies ($I^2 = 68\%$, $P = 0.05$)[37,40,42], and the pooled analysis with a random-effects model suggested a statistically significance in favor of RARP (OR = 2.77; 95%CI: 1.23-6.21; $P < 0.05$) (Figure 6).

**Twelve-month potency recovery after RARP and RRP:** Six studies were included[35-37,40-42] and no statistical heterogeneity was observed ($I^2 = 0\%$, $P = 0.61$). The cumulative analysis with a fixed- effects model showed a statistically significant advantage in favor of RARP (OR = 2.66; 95%CI: 1.96-3.60; $P < 0.01$) (Figure 6).

Table 6 summarizes the results of potency recovery rate after RARP with LRP. Among these eight studies[44-46,49,50,52-54], 1322 patients who received LRP and 1203 patients who received RARP were included, and all these studies performed the NS techniques (bilateral or unilateral) except the one by Asimakopoulos *et al*[44]. Most of the studies

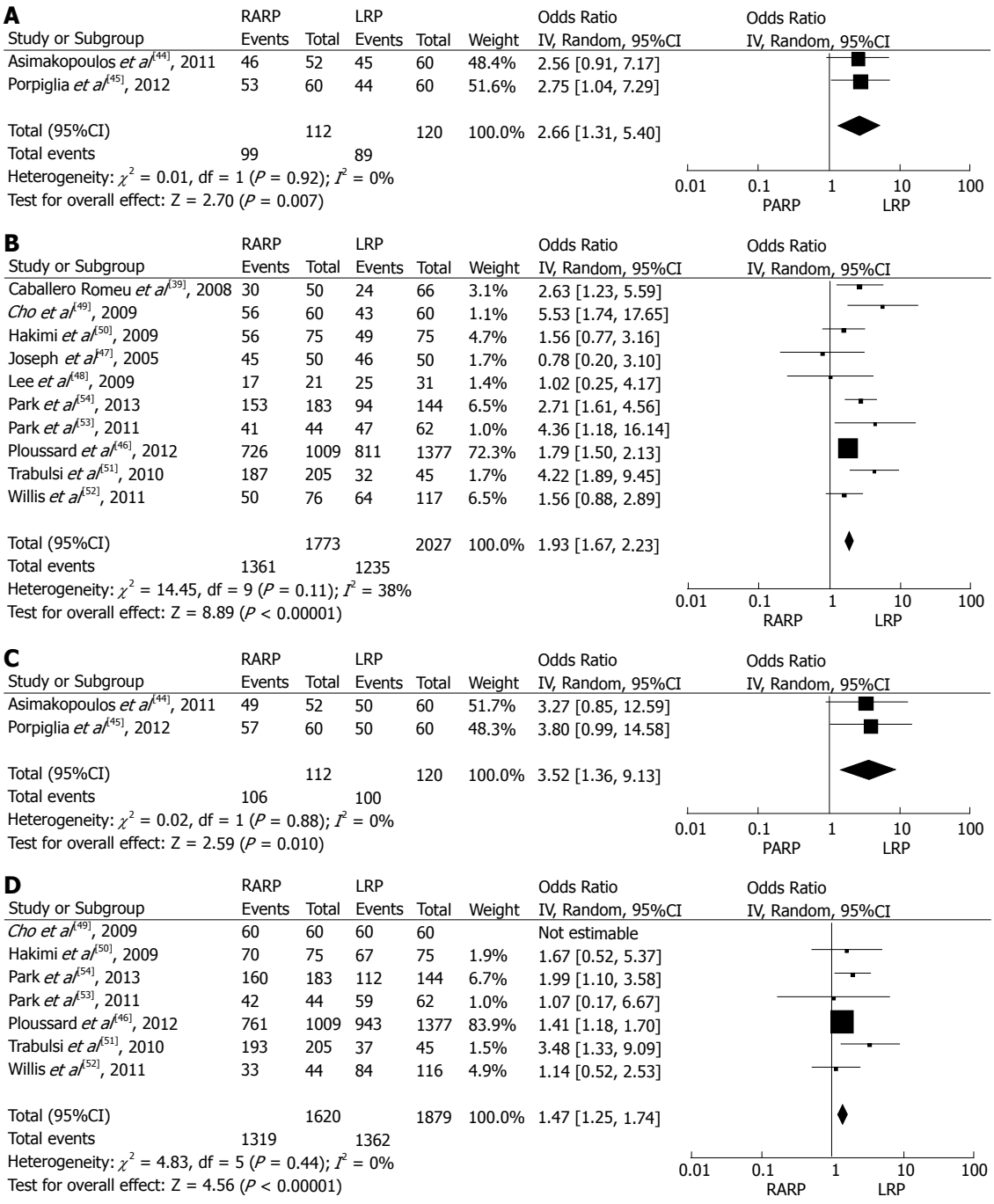



**Figure 4 Forest plots and meta-analyses of robot-assisted radical prostatectomy and laparoscopic radical prostatectomy.** A: 6-mo continence recovery based on randomized control trials (RCTs); B: 6-mo continence recovery based on non-randomized control trials (NRCTs); C: 12-mo continence recovery based on RCTs; D: 12-mo continence recovery based on NRCTs. RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy.

used a strict potency definition as ESI. In addition, two retrospective studies[52,53] had a high loss to follow-up rate (> 20%) at 12 mo. The RCTs were evaluated separately with NRCTs. For NRCTs, the mean potency recovery rates at 6 and 12 mo were 33.8% (20.4%-48.5%) and 43.2% (31.6%-65.5%) after LRP; and 55.5% (31.1%-75%) and 65.1% (36.5%-80.0%) after RARP.

**Six-month potency recovery after RARP and LRP**: The two RCTs[44,45] showed a statistical heterogeneity ($I^2$ = 84%, $P$ < 0.05), and demonstrated comparable result between RARP and LRP with a random-effects model (OR = 4.75; 95%CI: 0.92-24.54; $P$ = 0.06) (Figure 7). In the cumulative analysis of five NRCTs[46,49-50,52,54], no heterogeneity was found ($I^2$ = 0%, $P$ = 0.50), so a fixed-effects model was utilized. The result demonstrated a statistically significant advantage in favor of RARP (OR = 2.56; 95%CI: 2.11-3.10; $P$ < 0.01) (Figure 7).

**Twelve-month potency recovery after RARP and LRP:** No evidence of statistical heterogeneity was observed in the two RCTs ($I^2$ = 17%, $P$ = 0.27) and

**Table 4 Comparative studies evaluating potency recovery after retropubic radical prostatectomy or laparoscopic radical prostatectomy**

| Quality | Case, *n* | Author, yr | Country | Age (yr) | BMI (kg/m$^2$) | Prostate volume (mL, g) | Gleason score (biopsy) | PSA (ng/mL) | Study design | Potency definition | Data collection | Loss of follow-up (N/Y, %) | Potency recovery(UNS/BNS),%(*n*) | | Potency recovery(unclear NS), %(*n*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | 6 mo | 12 mo | 6 mo | 12 mo |
| 3/2/2(H) | RRP, 70 | Anastasiadis *et al*[20], 2003 | France | 64.8 ± 6.4 | - | - | 6.1 ± 1.1 | 11.2 ± 9.7 | Prospective | ESI | Nonvalidated questionnaire | Y, > 20% | - | 71.0 (23/33) | - | 30.0 (10/33) |
| | LRP, 230 | | | 64.1 ± 6.4 | | | 5.8 ± 1.2 | 10.7 ± 8.8 | | | | | | 98.0 (104/106) | | 41.0 (43/106) |
| 2/2/3(H) | RRP, 33 | Roumeguere *et al*[21], 2003 | Belgium | 63.9 ± 5.5 | - | 42.0 ± 20.4 | 5.4 ± 1.5 | 10.5 ± 11.5 | Prospective | ESI | IIEF-5 | N | 33.3 (11/33) | 54.5 (18/33) | - | - |
| | LRP, 26 | | | 62.5 ± 6.0 | | 37.3 ± 15.6 | 5.4 ± 1.5 | 8.6 ± 5.2 | | | | | 34.6 (9/26) | 65.3 (17/26) | | |
| 3/2/3(H) | RRP, 25 | Wagner *et al*[23], 2007 | United States | 59 ± 6.9 | 29 ± 4.5 | - | - | 8.1 ± 6.27 | Prospective | ESI | EPIC | N | - | 44.0 (11/25) | - | - |
| | LRP, 37 | | | 58 ± 6.9 | 27 ± 3.0 | | | 6.2 ± 4.22 | | | | | | 41.0 (15/37) | | |
| 3/2/2(H) | RRP, 164 | Touijer *et al*[24], 2008 | United States | 59 (54, 64) | - | - | - | 5.3 (4.1, 7.3) | Prospective | ESI | Institutional questionnaire | N | - | - | - | 58.5 (96/164) |
| | LRP, 132 | | | 60 (55, 65) | | | | 5.3 (4.0, 7.5) | | | | | | | | 56.2 (73/130) |
| 3/2/3(H) | RRP, 150 | Greco *et al*[25], 2009 | Italy | 61.5 (49-74) | 29 (25-33) | - | 5 (3-7) | 6.95 (3.4-10) | Prospective | ESI | IIEF-5 | N | - | 51.0 (77/150) | - | |
| | LRP, 150 | | | 60.5 (45-76) | 32 (26-38) | | 5 (3-7) | 6.3 (2.4-10) | | | | | | 66.0[a] (99/150) | | |
| 3/2/2(H) | RRP, 102 | Dahl *et al*[26], 2009 | United States | 59.9 | - | - | - | - | Prospective | ESI | Validated questionnaire | Y, > 20% | - | - | 23.0 (18/77) | 32.0 (23/73) |
| | LRP, 104 | | | 59.5 | | | | | | | | | | | 37.0 (28/75) | 43.0 (33/77) |
| 3/1/2(M) | RRP, 50 | Artibani *et al*[28], 2003 | Italy | 64.28 ± 6.6 | - | - | 5.7 ± 1.2 | 11 ± 9 | Retrospective | ESI | Nonvalidated questionnaire | Y, < 20% | - | - | 10.0 (4/40) | - |
| | LRP, 71 | | | 63.14 ± 5.8 | | | 5.8 ± 1.3 | 15.7 ± 17 | | | | | | | 8.8 (5/57) | |
| 4/2/2(H) | RRP, 42 | Ghavamian *et al*[29], 2006 | United States | 57.8 ± 7.3 | 28.1 | 53.2 (19-135) | 6.7 ± 1.3 | 9.9 ± 7.1 | Retrospective | ESI | IIEF-5 | N | 38.1 (16/42) | 52.5 (21/40) | - | - |
| | LRP, 50 | | | 60.8 ± 6.1 | 27.5 | 40.8 (20-114) | 6.4 ± 0.8 | 7.6 ± 8.0 | | | | | 48.0 (24/50) | 64.0 (32/50) | | |
| 2/1/1(M) | RRP, 128 | Springer *et al*[32], 2013 | Germany | 57.2 ± 7.4 | 28.3 ± 2.6 | - | - | 3.1 ± 1.7 | Retrospective | IIEF-5 > 22 | IIEF-5 | N | - | 53.1 (68/128) | - | - |
| | LRP, 125 | | | 56.8 ± 6.7 | 27.7 ± 3.8 | | | 3.2 ± 1.4 | | | | | | 74.4[a] (93/125) | | |
| 3/2/2(H) | RRP, 143 | Magheli *et al*[33], 2014 | Germany | 62.6 ± 5.4 | - | 58 ± 22 | - | 10.1 ± 11.9 | Retrospective | IIEF-5 > 17 | Validated questionnaire | Y, > 20% | - | 29.0 (18/62) | - | - |
| | LRP, 79 | | | 62.3 ± 5.7 | | 53 ± 20 | | 9.2 ± 6.9 | | | | | | 28.0 (7/25) | | |

[a]$P < 0.05$. RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy; ESI: Erection sufficient for intercourse; IIEF: International index of erectile function; EPIC: Expanded prostate cancer index composite; UNS: Unilateral nerve sparing; PSA: Prostate-specific antigen; BNS: Bilateral nerve sparing.

the pooled analyses with a fixed-effects model showed a statistically significant advantage in favor of RARP (OR = 5.35; 95%CI: 2.77-10.31; $P < 0.01$) (Figure 7). In the six NRCTs[46,49-50,52-54], a statistical heterogeneity ($I^2 = 52\%$, $P = 0.27$) was found, and the cumulative analysis also demonstrated a statistically significant advantage in favor of RARP by using a random-effects model (OR = 1.99; 95%CI: 1.35-2.93; $P < 0.01$) (Figure 7).

### *Sensitivity analysis and meta-regression analysis*

Sensitivity analysis was preformed to verify the reliability and stability of the evidence when a statistical heterogeneity existed. The subgroup analyses of the 6- or 12-mo urinary continence recovery following LRP and RRP did not vary significantly by source of country ($P > 0.05$), continence definition ($P > 0.05$), study design ($P > 0.05$) and loss of follow-up rate ($P > 0.05$) (Tables 7 and 8). While in the subgroup analyses of 6-mo urinary continence recovery following RARP and RRP, the results were unstable, with Western country and strict definition indicating better outcomes in favor of RARP (OR = 2.32; 95%CI: 1.47-3.67; $P < 0.01$ and OR = 3.09; 95%CI: 1.65-5.80; $P < 0.01$, respectively) (Table 9). Table 10 independently evaluates the most important factors (NS procedures) for 12-mo potency recovery among different techniques. Since all the included studies comparing RARP and LRP had taken unilateral or bilateral NS procedures, only the subgroups comparing LRP/RRP and RARP/RRP were analyzed. Our results again

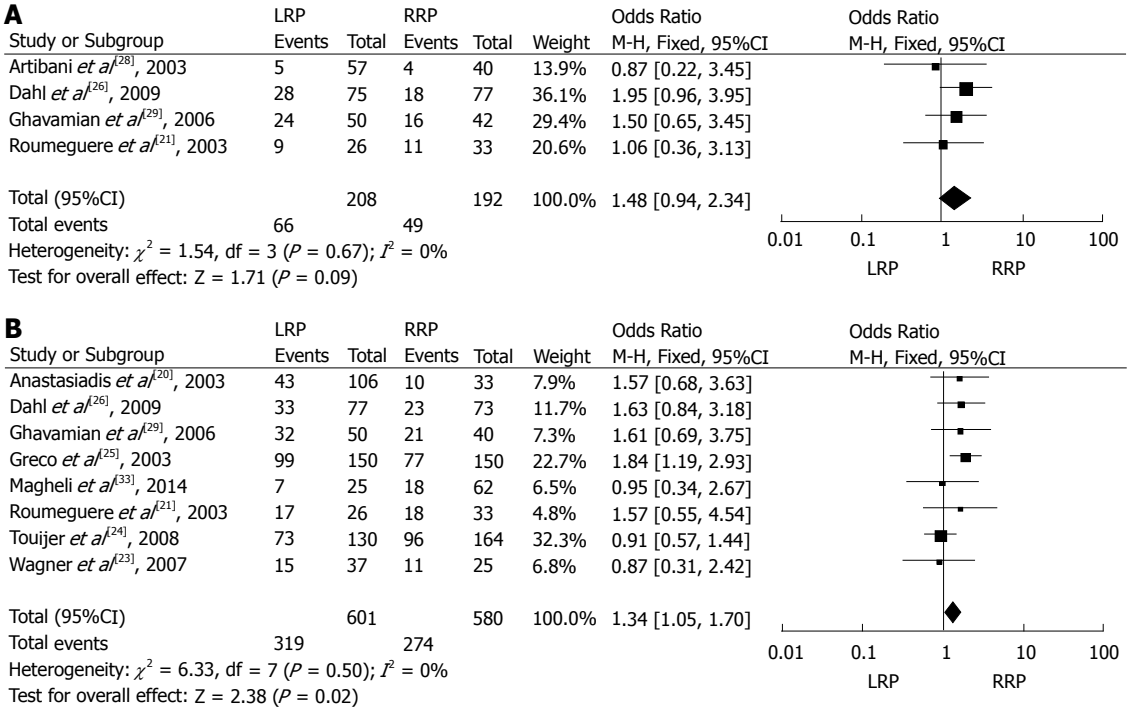



**Figure 5 Forest plots and meta-analyses of laparoscopic radical prostatectomy and retropubic radical prostatectomy.** A: 6-mo potency recovery; B: 12-mo potency recovery. RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy.

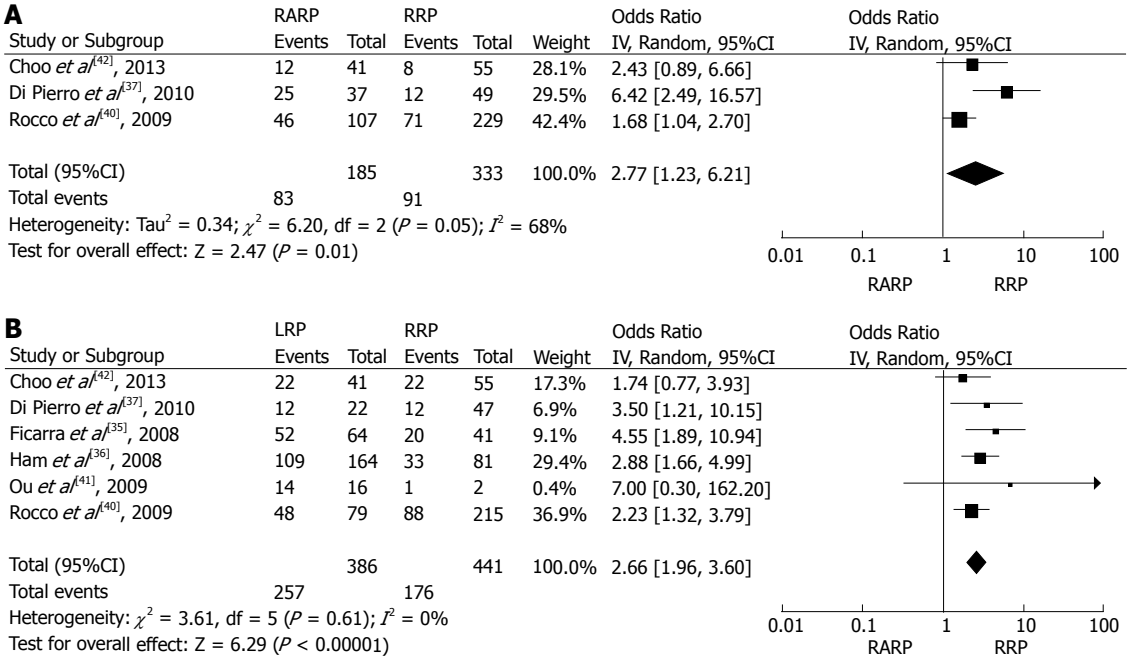



**Figure 6 Forest plots and meta-analyses of robot-assisted radical prostatectomy and retropubic radical prostatectomy.** A: 6-mo potency recovery; B: 12-mo potency recovery. RRP: Retropubic radical prostatectomy; RARP: Robot-assisted radical prostatectomy.

confirmed that the NS measures were advantageous factors to potency recovery ($P < 0.05$). All of the other remaining outcomes were proved to be stable and reliable by either using model conversion or exclusion of the study with the largest proportion.

Regrettably, in our meta-regression analyses, none of the adjustments such as age, BMI, prostate volume, Gleason score or PSA, achieved a statistical significance ($P < 0.05$) (Tables 11 and 12); however, the l'Abbé graphs showed an overall trend either as a positive correlation or a negative correlation between those potential factors and different surgical techniques. The older age, lower BMI and lower PSA level were associated with lower odds of different technical groups (Figure 8). The prostate vol-

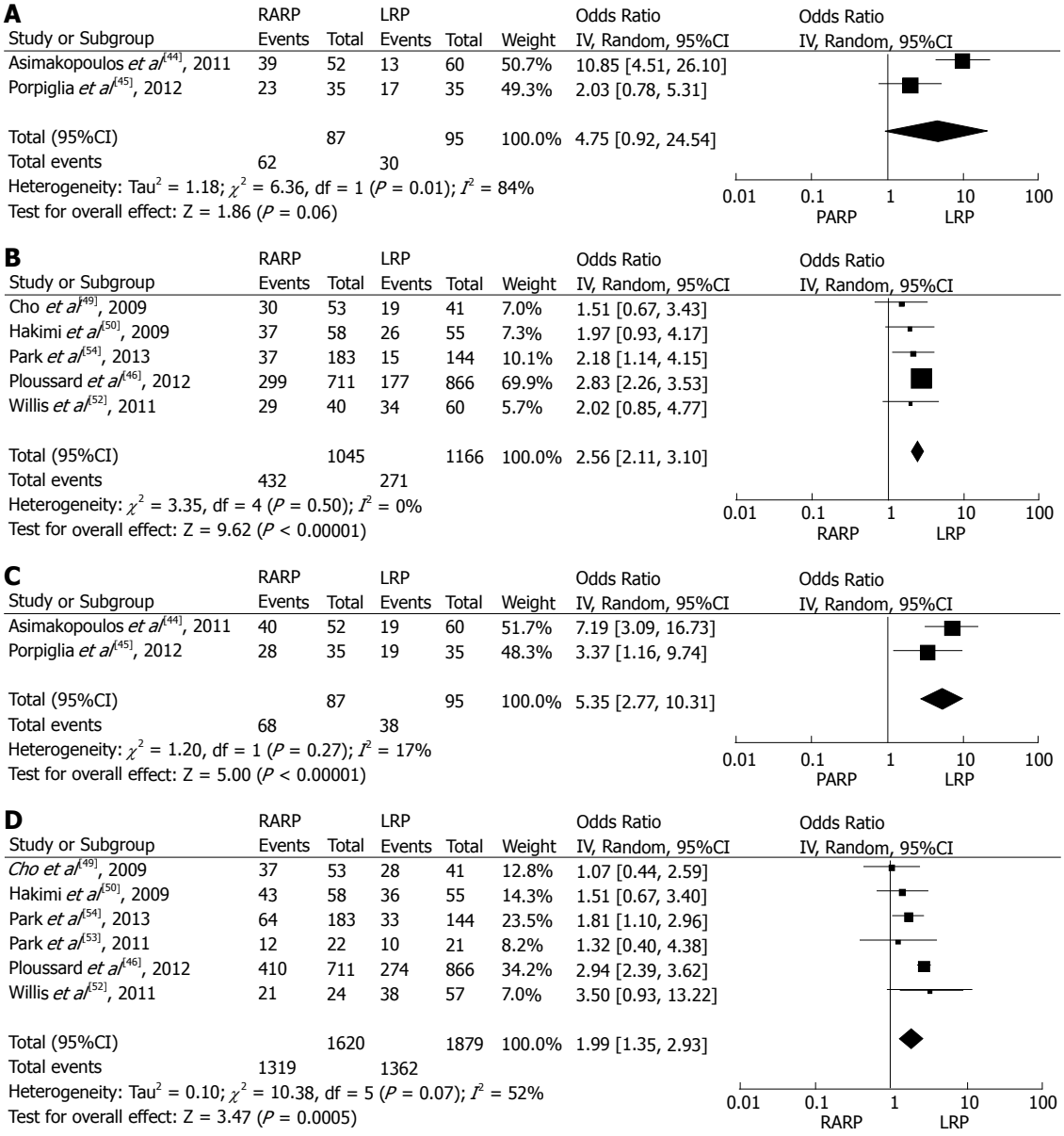



**Figure 7 Forest plots and meta-analyses of robot-assisted radical prostatectomy and laparoscopic radical prostatectomy.** A: 6-mo potency recovery based on randomized control trials (RCTs); B: 6-mo potency recovery based on non-randomized control trials (NRCTs); C: 12-mo potency recovery based on RCTs; D: 12-mo potency recovery based on NRCTs. LRP: Laparoscopic radical prostatectomy; RARP: Robot-assisted radical prostatectomy.

ume and Gleason score did not demonstrate any trend between the different methods of surgery (Figure 8).

### *Publication bias*

The funnel plots of two comparative results (6-mo potency recovery after LRP/RRP and after RARP/LRP) were asymmetrical (Figure 9), indicating the existence of publication bias; this was also confirmed by Egger linear regression test ($P = 0.024$ and $P = 0.013$, respectively). All the other comparisons demonstrated symmetrical funnel plots and found no statistical significance ($P > 0.05$) by using the Egger's test, indicating no publication bias.

## DISCUSSION

This meta-analysis was designed in accordance with the MOOSE reporting guidelines[10]. In 2011, Ficarra *et al*[8,9] had performed two meta-analyses which tried to compare the superiority of techniques concerning RARP *vs* RRP and RARP *vs* LRP. However, a deep investigation focusing on the deficiencies of these two studies made them possibly inconvincible: (1) limited number of studies included; (2) the lack of credible quality assessment tool for the included studies; (3) as for the comparison between RARP and LRP, it did not correspond with the methodological rules of a meta-analysis to integrate the RCT with the NRCT studies to analyze the outcomes, as they were totally two different level of evidences; (4) in the few included studies, Kim *et al*[10] and Krambeck *et al*[11] results were not available for the comparison between RARP and RRP; and (5) though all the outcomes of these two studies were apparently heterogeneous, the

**Table 5 Comparative studies evaluating potency recovery after retropubic radical prostatectomy or robot-assisted radical prostatectomy**

| Quality | Case, *n* | Author, yr | Country | Age (y) | BMI (kg/m$^2$) | Prostate volume (mL, g) | Gleason score (biopsy) | PSA (ng/mL) | Study design | Potency definition | Data collection | Loss of follow-up (N/Y,%) | Potency recovery(UNS/BNS), %(*n*) | | Potency recovery (unclear NS), %(*n*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | 6 mo | 12 mo | 6 mo | 12 mo |
| 3/2/3(H) | RRP, 100 | Tewari *et al*[34], 2003 | United States | 63.1 (42.8-72) | 27.6 (17-41) | 48.4 (24.2-70) | - | 7.3 (1.9-35) | Prospective | Presence of erection | Interview | - | Median: 440 d | | Median:440 d | |
| | RARP, 200 | | | 59.9 (40-72) | 27.7 (19-38) | 58.8 (18-140) | | 6.4 (0.6-41) | | | | | Median: 180 d[a] | | Median:180 d[a] | |
| 3/2/2(H) | RRP, 41 | Ficarra *et al*[35], 2008 | Italy | 65 (61-69) | 26 (24-28) | 40 (30-47) | - | 6 (5-10) | Prospective | IIEF-5 > 17 | IIEF-5 | N | - | 49.0 (20/41) | - | - |
| | RARP, 64 | | | 61 (57-67) | 26 (24-28) | 37.5 (30-48) | | 6.4 (4.6-9) | | | | | | 81.0[a] (52/64) | | |
| 3/2/3(H) | RRP, 81 | Ham *et al*[36], 2008 | South Korea | 66.9 ± 6.0 | 23.6 ± 1.8 | - | - | 55.2 ± 23.7 | Prospective | ESI | IIEF-5 | N | - | 40.7 (33/81) | - | - |
| | RARP, 164 | | | 67.3 ± 6.2 | 23.6 ± 2.3 | | | 22.3 ± 34.3 | | | | | | 66.5 (109/164) | | |
| 3/1/2(M) | RRP, 49 | Di Pierro *et al*[37], 2010 | Switzerland | 64.3 (59.1-68.0) | - | - | - | 7.57 (5.1-10.4) | Prospective | ESI | Institutional questionnaire | Y, > 20% | - | - | 25.0 (12/49) | 26.0 (12/47) |
| | RARP, 37 | | | 62.8 (58.4-67.0) | | | | 7.72 (5.6-12.1) | | | | | | | 68.0 (25/37) | 55.0 (12/22) |
| 1/1/1(L) | RRP, 122 | Kim *et al*[10], 2011 | South Korea | 66.5 ± 5.7 | - | 18.2 ± 23.4 | - | 14.6 ± 22.1 | Prospective | ESI | Validated questionnaire | N | - | - | 6.7 (8/122) | 28.1 (34/122) |
| | RARP, 373 | | | 64.2 ± 7.3 | | 15.2 ± 20.2 | | 10.4 ± 16.0 | | | | | | | 33.0 (123/373) | 57.1 (213/373) |
| 2/0/1(L) | RRP, 588 | Krambeck *et al*[11], 2008 | United States | 61.0 (41.0-77.0) | - | - | - | 5.0 (0.6-39.7) | Retrospective | ESI | Institutional questionnaire | Y, > 20% | - | - | - | 62.8 (262/417) |
| | RARP, 294 | | | 61.0 (38.0-76.0) | | | | 4.9 (0.5-33.5) | | | | | | | | 70.0 (142/203) |
| 3/1/2(M) | RRP, 240 | Rocco *et al*[40], 2009 | Italy | 63 (46-77) | - | - | 6 (4-10) | 6.7 (0.7-22.0) | Retrospective | ESI | Interview | Y, > 20% | - | - | 31.0 (71/229) | 41.0 (88/215) |
| | RARP, 120 | | | 63 (47-76) | | | 6 (4-9) | 6.9 (0.4-23.0) | | | | | | | 43.0 (46/107) | 61.0 (48/79) |
| 3/1/3(H) | RRP, 2 | Ou *et al*[41], 2009 | United States | 70.03 ± 6.10 | 24.09 ± 3.28 | 15.89 ± 14.15 | 6.22 ± 1.62 | | Retrospective | Presence of erection | Unspecified | N | - | 50.0 (1/2) | - | - |
| | RARP, 16 | | | 67.27 ± 6.21 | 24.22 ± 3.16 | 16.45 ± 18.80 | 6.13 ± 0.9 | | | | | | | 87.5.0 (14/16) | | |
| 3/2/3(H) | RRP, 55 | Choo *et al*[42], 2013 | South Korea | 67 ± 6.25 | 24 ± 2.73 | 42 ± 18.82 | | 7.6 ± 19.33 | Retrospective | ESI | IIEF-5 | N | 15.0 (8/55) | 40.0 (22/55) | - | - |
| | RARP, 41 | | | 66 ± 7.75 | 24 ± 2.55 | 41 ± 15.77 | | 7.2 ± 13.19 | | | | | 29.0 (12/41) | 54.0 (22/41) | | |

[a]$P < 0.05$. RRP: Retropubic radical prostatectomy; RARP: Robot-assisted radical prostatectomy; ESI: Erection sufficient for intercourse; PSA: Prostate-specific antigen; IIEF-5: International index of erectile function; UNS: Unilateral nerve sparing; BNS: Bilateral nerve sparing.

authors did not use any sensitivity analysis or subgroup analysis to explain the source of heterogeneity.

In contrast, our meta-analysis directly compared these three surgical approaches for the 6- and the 12-mo functional outcomes following radical prostatectomy (RP). In 2009, Ficarra *et al*[6] conducted a meta-analysis including only 6 studies and reported the 12-mo continence recovery following LRP and RRP, whose result was consistent with ours (OR = 0.87; 95%CI: 0.54-1.39; $P = 0.56$ and OR = 0.92; 95%CI: 0.57-1.51; $P = 0.75$, respectively). However, in our meta-analysis, the study by Rassweiler *et al*[55] was excluded because of its preoperative surgery and neo-adjuvant therapy and 13 eligible studies were included. Moreover, we evaluated the 6-mo continence recovery (OR, 0.84; 95%CI: 0.50-1.41; $P = 0.52$), so this result would be more convincible and complete. Compared with the previous meta-analysis by Ficarra *et al*[8], whose results for 12-mo urinary continence recovery based on a pooled analysis of 5 studies comparing RARP *vs* RRP, and 5 studies comparing RARP *vs* LRP identified the advantage in favor of RARP (OR = 1.53; 95%CI: 1.04-2.25; $P < 0.05$ and OR = 2.39; 95%CI: 1.29-4.45; $P < 0.01$, respectively), our meta-analyses identified the similar advantage in favor of RARP both at 6-mo and 12-mo follow-up. A critical review by Coelho *et al*[7] also indicated better outcomes after RARP compared with RRP (92% *vs* 80%) or with LRP (92% *vs* 84%). Obviously,

**Table 6 Comparative studies evaluating potency recovery after laparoscopic radical prostatectomy or robot-assisted radical prostatectomy**

| Quality | Case, *n* | Author, yr | Country | Age (yr) | BMI (kg/m$^2$) | Prostate Volume (mL,g) | Gleason score (biopsy) | PSA (ng/mL) | Study design | Potency definition | Data collection | Loss of follow-up (N/Y, %) | Potency recovery (UNS/BNS), %(*n*) | | Potency recovery (unclear NS), %(*n*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | 6 mo | 12 mo | 6 mo | 12 mo |
| High | LRP, 60 | Asimakopoulos *et al*[44], 2011 | Italy | 61.1 ± 5.1 | 26.3 ± 2.2 | - | - | 7.37 (1.5-9.15) | RCT | ESI | IIEF-6 | N | - | - | 22.0 (13/60) | 32.0 (19/60) |
| | RARP, 52 | | | 59.6 ± 5.4 | 25.8 ± 2.6 | | | 8.9 (5.8-9.2) | | | | | | | 75.0[a] (39/52) | 77.0[a] (40/52) |
| High | LRP, 35 | Porpiglia *et al*[45], 2012 | Italy | 64.7 ± 5.9 | 26.8 ± 2.9 | 37.7 ± 14.1 | - | 8.3 ± 6.5 | RCT | IIEF-5 > 17 | IIEF-5 | N | 48.5 (17/35) | 54.2 (19/35) | - | - |
| | RARP, 35 | | | 63.9 ± 6.7 | 26.2 ± 2.5 | 36.2 ± 12.6 | | 6.9 ± 4.2 | | | | | 65.7 (23/35) | 80.0[a] (28/35) | | |
| 3/1/3(H) | LRP, 866 | Ploussard *et al*[46], 2012 | France | 62.7 | 26.6 | - | - | 9.8 | Prospective | ESI | IIEF-5 | N | 20.4 (177/866) | 31.6 (274/866) | - | - |
| | RARP, 711 | | | 62.7 | 26.5 | | | 9.2 | | | | | 42.1 (299/711) | 57.7 (410/711) | | |
| 3/1/2(M) | LRP, 41 | Cho *et al*[49], | South Korea | 66.5 (57-75) | 23.65 (18.1-28.4) | 39.7 (19-72) | 6.81 (5-9) | 11.04 (2.72-36.6) | Retrospective | ESI | Interview | N | 46.3 (19/41) | 68.3 (28/41) | - | - |
| | RARP, 53 | | | 66.3 (50-77) | 24.61 (19.9-26.3) | 36.6 (22-92.8) | 6.83 (5-8) | 9.98 (2.91-26.3) | | | | | 56.6 (30/53) | 69.8 (37/53) | | |
| 4/2/3(H) | LRP, 55 | Hakimi *et al*[50], 2009 | United States | 59.6 (43-72) | - | - | - | 7.5 | Retrospective | Presence of | IIEF-5 | N | 47.3 (26/55) | 65.5 (36/55) | - | - |
| | RARP, 58 | | | 59.8 (42-71) | | | | 8.4 | | Erection | | | 63.8 (37/58) | 74.1 (43/58) | | |
| 3/2/2(H) | LRP, 86 | Willis *et al*[52], 2011 | United States | 58.0 ± 6.7 | 27.0 ± 3.4 | 35.2 ± 10.1 | - | 5.7 ± 2.9 | Retrospective | ESI | Validated | Y, > 20% | 57.0 (34/60) | 67.0 (38/57) | - | - |
| | RARP, 74 | | | 58.1 ± 6.3 | 26.7 ± 3.3 | 41.5 ± 15.2 | | 5.0 ± 2.2 | | | questionnaire | | 73.0 (29/40) | 88.0 (21/24) | | |
| 3/1/2(M) | LRP, 35 | Park *et al*[53], 2011 | South Korea | 65.7 (38-77) | 24.6 (19.4-31.4) | 30.1 (12.0-56.0) | - | 9.14 (2.65-30.77) | Retrospective | ESI | Interview | Y, > 20% | - | 47.6 (10/21) | - | - |
| | RARP, 37 | | | 62.7 (46-71) | 26.0 (19.7-39.4) | 32.9 (15.5-66.8) | | 6.32 (1.86-29.5) | | | | | | 54.5 (12/22) | | |
| 3/2/3(H) | LRP, 144 | Park *et al*[54], 2013 | South Korea | 67 (38-77) | 24.2 (17.2-31.4) | 28.8 (12.0-74.0) | - | 5.84 (0.08-41.26) | Retrospective | ESI | Interview | N | 30.8 (26/83) | 32.7 (27/83) | 10.2 (15/144) | 22.9 (33/144) |
| | RARP, 183 | | | 63 (44-75) | 24.7 (16.4-39.4) | 30.3 (15.5-82.8) | | 4.98 (0.05-51.46) | | | | | 31.1 (49/156) | 36.5 (57/156) | 20.1 (37/183) | 35.0 (64/183) |

[a]$P < 0.05$. LRP: Laparoscopic radical prostatectomy; RARP: Robot-assisted radical prostatectomy; RCT: Randomized controlled trial; ESI: Erection sufficient for intercourse; IIEF: International index of erectile function. EPIC: Expanded prostate cancer index composite; SHIM: Sexual health inventory for men; UNS: Unilateral nerve sparing; PSA: Prostate-specific antigen; BNS: Bilateral nerve sparing.

except the inclusion of more studies and the exclusion of two studies[10,11], our meta-analyses separated the RCT from the NRCT studies to analyze the outcomes, therefore, the result was subjected to fewer confounding and biases of study design.

In terms of potency recovery, for the first time, with 8 studies included, our meta-analysis supported the superiority of LRP than RRP at 12-mo follow-up (OR = 1.34; 95%CI: 1.05-1.70; $P < 0.05$). Compared with the previous meta-analysis by Ficarra *et al*[9], whose results for 12-mo potency recovery based on pooled 6 studies comparing RARP *vs* RRP, and 4 studies comparing RARP *vs* LRP demonstrated a better outcome in favor of RARP against RRP (OR = 2.84; 95%CI: 1.48-5.43; $P < 0.01$) and a non-statistically significant trend between RARP and LRP (OR = 1.89; 95%CI: 0.70-5.05; $P = 0.21$), our meta-analyses showed a statistically significant advantage in favor of RARP *vs* RRP (6-mo: $P < 0.05$ and 12-mo: $P < 0.01$, respectively) and also showed a better result in favor of RARP *vs* LRP (6-mo: $P < 0.01$ and 12-mo: $P < 0.01$, respectively). In ad-

**Table 7 Subgroup analyses of 6-mo urinary continence recovery after laparoscopic radical prostatectomy or retropubic radical prostatectomy**

| Subgroup | Study | Sample size | Heterogeneity $I^2$(%) | *P*-value | Meta-analysis | |
|---|---|---|---|---|---|---|
| | | | | | OR | 95%CI |
| Country | Asia | 553 | 63 | 0.06 | 0.45 | 0.20-1.04 |
| | America | 346 | 0 | 0.45 | 0.83 | 0.51-1.34 |
| | Europe | 763 | 80 | 0.40 | 1.46 | 0.60-3.55 |
| Continence definition | 0 pad | 1662 | 74 | 0.52 | 0.84 | 0.50-1.41 |
| | 0-1 pad | 0 | - | - | - | - |
| Study design | prospective | 968 | 77 | 0.55 | 1.24 | 0.61-2.50 |
| | retrospective | 694 | 59 | 0.08 | 0.56 | 0.29-1.07 |
| Loss of follow-up | ≤ 20% | 911 | 71 | 0.87 | 1.06 | 0.53-2.09 |
| | > 20% | 751 | 78 | 0.32 | 0.66 | 0.29-1.51 |

OR: Odds ratio; CI: Confidence

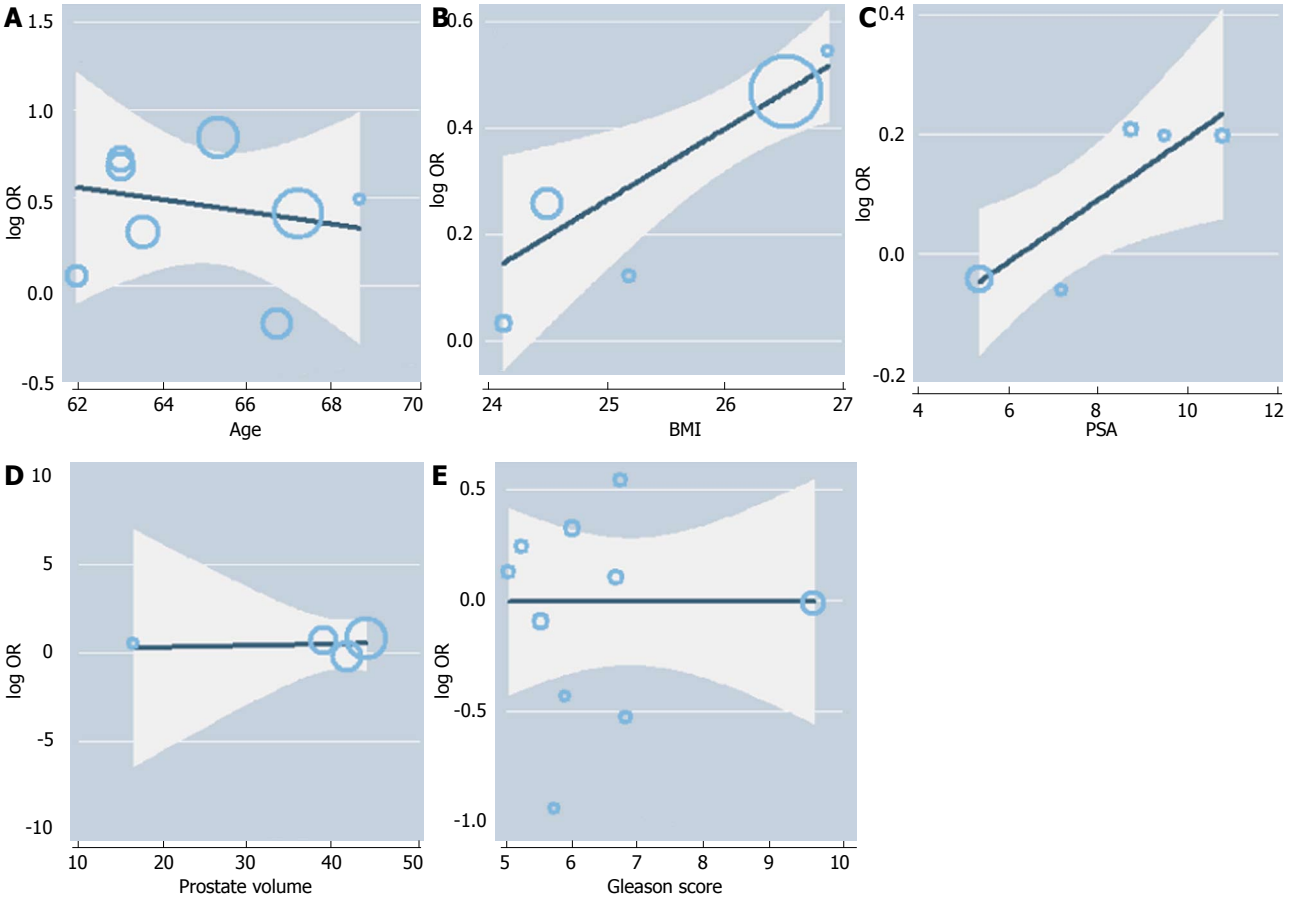



**Figure 8 Representative l'Abbé plots show the overall trend.** A: 12-mo continence of robot-assisted radical prostatectomy (RARP) and retropubic radical prostatectomy (RRP); B: 12-mo potency of RARP and laparoscopic radical prostatectomy (LRP); C: 12-mo potency of LRP and RRP; D: 12-mo continence of RARP and RRP; E: 12-mo continence of LRP and RRP. PSA: Prostate-specific antigen; BMI: Body mass index.

dition, there were some potential biases in Ficarra *et al*[9] meta-analysis which included only 6 studies, and two of them[10,11] were considered ineligible. While in our meta-analyses, the increased study number and the separation of the RCT and the NRCT studies, would be helpful to minimize the confounding of study design. Briefly, we supported a dramatic grading by superiority level for different comparisons of potency: RARP > LRP > RRP.

In this review, statistically significant heterogeneity was observed for several comparisons. So the subgroup analyses were added according to adjustment for country, continence or potency definition, study design and the NS procedures. We found that Western country and strict definition indicated better outcomes in favor of RARP against RRP ($P < 0.01$) for 6-mo urinary continence recovery. This difference may be explained by the popularity of robotic technique in Western countries. As the classic NS technique was repeatedly proved to be a sig-

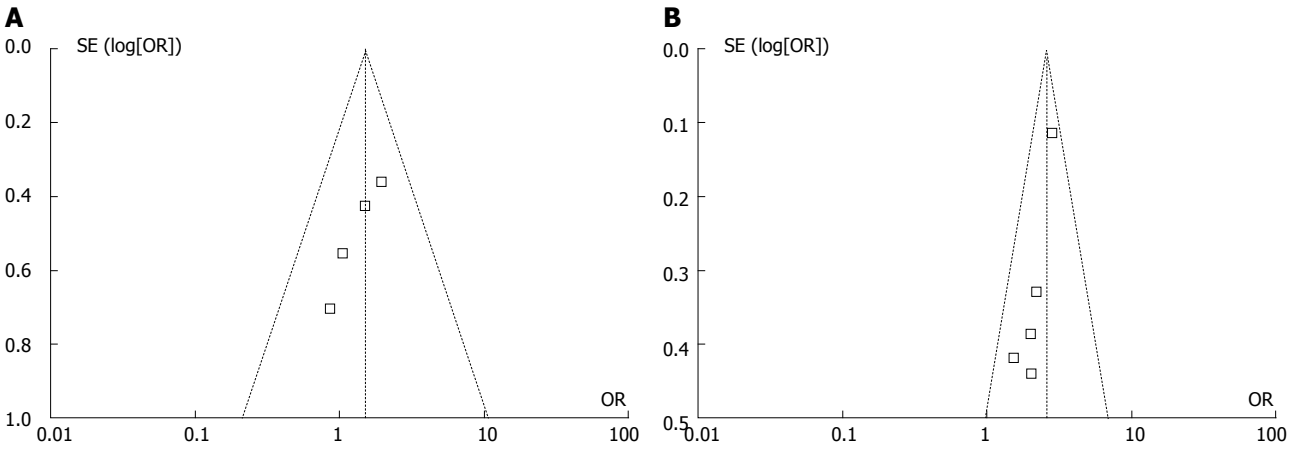



**Figure 9 Funnel plots for 6-mo potency recovery.** A: Comparison of laparoscopic radical prostatectomy (LRP) and retropubic radical prostatectomy (RRP); B: Comparison of robot-assisted radical prostatectomy (RARP) and LRP based on non-randomized control trials (NRCTs).

**Table 8 Subgroup analyses of 12-mo urinary continence recovery after laparoscopic radical prostatectomy or retropubic radical prostatectomy**

| Subgroup | Study | Sample size | Heterogeneity $I^2$ (%) | *P*-value | Meta-analysis | |
|---|---|---|---|---|---|---|
| | | | | | OR | 95%CI |
| Country | Asia | 553 | 72 | 0.18 | 0.38 | 0.09-1.54 |
| | America | 911 | 89 | 0.91 | 0.95 | 0.35-2.55 |
| | Europe | 1343 | 29 | 0.33 | 1.26 | 0.79-2.02 |
| Continence | 0 pad | 908 | 55 | 0.75 | 1.08 | 0.68-1.69 |
| definition | 0-1 pad | 754 | 88 | 0.27 | 0.53 | 0.17-1.63 |
| Study design | prospective | 509 | 83 | 0.51 | 1.26 | 0.63-2.53 |
| | retrospective | 1153 | 57 | 0.15 | 0.60 | 0.30-1.20 |
| Loss of follow-up | ≤ 20% | 451 | 82 | 0.82 | 1.09 | 0.51-2.33 |
| | > 20% | 1211 | 59 | 0.45 | 0.79 | 0.43-1.46 |

OR: Odds ratio; CI: Confidence interval.

**Table 9 Subgroup analyses of 6-mo urinary continence recovery after robot-assisted radical prostatectomy or retropubic radical prostatectomy**

| Subgroup | Study | Sample size | Heterogeneity $I^2$ (%) | *P*-value | Meta-analysis | |
|---|---|---|---|---|---|---|
| | | | | | OR | 95%CI |
| Country | Asia | 809 | 92 | 0.35 | 1.93 | 0.48-7.70 |
| | Europe/America | 862 | 0 | < 0.01 | 2.32 | 1.47-3.67 |
| Continence | 0 pad | 828 | 63 | < 0.01 | 3.09 | 1.65-5.80 |
| definition | 0-1 pad | 673 | 82 | 0.52 | 1.62 | 0.37-7.06 |
| Study design | prospective | 448 | 0 | < 0.01 | 2.48 | 1.44-4.26 |
| | retrospective | 1223 | 80 | 0.1 | 2.07 | 0.87-4.95 |
| Loss of follow-up | ≤ 20% | 1161 | 80 | 0.1 | 2.00 | 0.88-4.53 |
| | > 20% | 510 | 0 | < 0.01 | 2.99 | 1.55-5.77 |

OR: Odds ratio; CI: Confidence interval.

nificant predictor of return of potency by Coelho *et al*[7], by Ayyathurai *et al*[56] and by Briganti *et al*[57], this review independently evaluated it for 12-mo potency recovery between different techniques, and we confirmed again that the NS measures were advantageous factors to potency recovery ($P < 0.05$). Furthermore, the other factors such as age, BMI, prostate volume, Gleason score or PSA could also be a source of heterogeneity. Standford *et al*[58] found that urinary function varied with age and sexual function with age and race. Shikanov *et al*[59] emphasized other factors influencing continence and potency, such as baseline status, surgical technique, extent of NS and adjuvant therapy. In this review, we performed meta-regression analyses to explore the correlation between these factors and different techniques. Though no obviously statistical significance was found, the l'Abbé graphs predicted the trends that better functional outcomes were more easily achieved in patients with younger age, larger

**Table 10 Subgroup analyses of 12-mo potency recovery after nerve sparing procedures**

| Techniques | Subgroup | Sample size | Heterogeneity $I^2$ (%) | *P*-value | Meta-analysis | |
|---|---|---|---|---|---|---|
| | | | | | OR | 95%CI |
| LRP *vs* RRP | uni/bilateral NS | 735 | 0 | < 0.05 | 1.52 | 1.09-2.13 |
| | unclear NS | 802 | 22 | 0.37 | 1.17 | 0.83-1.65 |
| RARP *vs* RRP | uni/bilateral NS | 464 | 0 | < 0.01 | 2.83 | 1.90-4.22 |
| | unclear NS | 446 | 0 | < 0.01 | 2.43 | 1.52-3.90 |

RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy; RARP: Robot-assisted radical prostatectomy; OR: Odds ratio; CI: Confidence interval; NS: Nerve sparing.

**Table 11 Meta-regression of 12-mo continence recovery**

| Techniques | Factors | Sample, *n* | Coefficient | *P* value | 95%CI | |
|---|---|---|---|---|---|---|
| | | | | | Lower CI | Upper CI |
| LRP *vs* RRP | Age | 14 | -0.0422414 | 0.480 | -0.1685084 | 0.0840256 |
| | Prostate Volume | 7 | 0.0004602 | 0.976 | -0.0367033 | 0.0376237 |
| | Gleason Score | 10 | -0.0002758 | 0.998 | -0.2325786 | 0.2320269 |
| | PSA | 11 | 0.0381884 | 0.508 | -0.0871645 | 0.1635414 |
| RARP *vs* RRP | Age | 8 | -0.0347693 | 0.763 | -0.3038441 | 0.2343054 |
| | BMI | 5 | 0.178217 | 0.604 | -0.8030416 | 1.159476 |
| | Prostate Volume | 4 | 0.0076432 | 0.912 | -0.2556839 | 0.2709703 |
| | PSA | 5 | 0.0028508 | 0.882 | -0.053367 | 0.0590685 |
| RARP *vs* LRP | Age | 6 | -0.0026949 | 0.968 | -0.1735327 | 0.1789224 |
| | BMI | 4 | 0.0709043 | 0.680 | -0.7088789 | 0.5670703 |
| | PSA | 6 | 0.0275948 | 0.661 | -0.1898594 | 0.1346698 |

RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy; RARP: Robot-assisted radical prostatectomy; CI: Confidence interval; PSA: Prostate-specific antigen; BMI: Body mass index.

**Table 12 Meta-regression of 12-mo potency recovery**

| Techniques | Factors | Sample, *n* | Coefficient | *P* value | 95%CI | |
|---|---|---|---|---|---|---|
| | | | | | Lower CI | Upper CI |
| LRP *vs* RRP | Age | 8 | -0.0334222 | 0.682 | -0.156947 | 0.2237914 |
| | Gleason Score | 5 | -0.0059256 | 0.732 | -0.5614423 | 0.4429304 |
| | PSA | 5 | 0.0509797 | 0.558 | -0.1961242 | 0.2980837 |
| RARP *vs* RRP | Age | 6 | -0.006352 | 0.939 | -0.2221039 | 0.2093999 |
| | PSA | 5 | 0.0018209 | 0.892 | -0.0373331 | 0.0409749 |
| RARP *vs* LRP | Age | 6 | -0.0437647 | 0.535 | -0.2229024 | 0.1353731 |
| | BMI | 5 | 0.1340739 | 0.315 | -0.220684 | 0.4888318 |
| | Prostate Volume | 4 | -0.0080152 | 0.894 | -0.2365214 | 0.2204911 |
| | PSA | 6 | 0.0350044 | 0.588 | -0.1301063 | 0.2001150 |

RRP: Retropubic radical prostatectomy; LRP: Laparoscopic radical prostatectomy; RARP: Robot-assisted radical prostatectomy; CI: Confidence interval; PSA: Prostate-specific antigen; BMI: Body mass index.

BMI or higher PSA level in the RARP group than the other two groups (LRP or RRP), while it was difficult to judge the superiority of any technique in patients with different prostate volumes and Gleason scores.

Some potential limitations should be noted. First, moderate heterogeneity was found in several comparisons. Except the potential confounding factors controlled by the inclusion criteria and analyzed with subgroup stratification as described above, surgeon's experience and the means of modification varied from one to another, which could also influence the functional outcomes and were difficult to control. Second, contrary to expectation, due to the inclusion of few eligible studies for each comparison and the lack of data in available studies, all the meta-regression analyses presented non-statistically significant differences, which limited us to reach an exact correlation between those potential factors and the three techniques, this result still needs to be identified by further research. Third, the quality of eligible studies could potentially be another confounding factor. RCTs are powerful tools, which provide the highest level of evidence; however, because many patients refuse to participate in the randomization and the blinding degree is less, surgical RCTs are difficult to conduct. Only two RCTs were included for the comparison between RARP and LRP, and the remaining studies were all observational

comparative studies. In addition, the NOS tool itself has imperfections[60]. Finally, publication bias still existed. The failed acquisition of gray literature may contribute to this publication bias.

The superiority of a certain surgical approach in terms of functional outcomes is always a pivotal controversy. These outcomes were influenced by multiple factors including patient characteristics, surgical techniques and methodology used for data collection. In summary, concerning the urinary continence recovery, only RARP showed an advantage when compared with LRP or with RRP, and the result was comparable between LRP and RRP. In terms of potency recovery, for the first time, we ranked the three surgical approaches into a superiority level: RARP > LRP > RRP, which showed a statistically significant advantage both at 6 and 12 mo postoperatively. However, the limitation of this meta-analysis and potential factors should be taken into consideration and our results also need to be validated by further high quality multi-center RCTs with strict design and large sample size.

## COMMENTS

### Background

Radical prostatectomy (RP) is one of the recommended standard treatments for clinically localized prostate cancer (cT1-cT2) patients. The retropubic radical prostatectomy (RRP) was considered as the gold standard and the most widely used treatment for patients with localized prostate cancer (PCa). Recently, the authors have witnessed the emergence of laparoscopic radical prostatectomy (LRP) and robot-assisted laparoscopic prostatectomy (RARP).

### Research frontiers

Several experts have demonstrated that when compared with RRP, LRP and RARP have obvious advantages such as lower blood loss, less need for transfusion and shorter hospital-stay, but the lack of high-quality evidence and RCTs available precluded us from proving the superiority of any surgical option in terms of postoperative functional outcomes.

### Innovations and breakthroughs

In terms of potency recovery, for the first time, we ranked the three surgical approaches into a superiority level: RARP > LRP > RRP, which showed a statistically significant advantage both at 6 and 12 mo postoperatively.

### Applications

Current evidence suggests that minimally invasive approaches (RARP or LRP) are effective procedures for functional recovery. However, more high-quality randomized, controlled trials investigating the long-term functional outcomes are required to determine the advantages of RARP.

### Terminology

RARP means robot-assisted laparoscopic prostatectomy. LRP means laparoscopic radical prostatectomy. RRP means retropubic radical prostatectomy. The principal postoperative functional outcomes for patients with prostatectomy are urinary continence and potency recovery.

### Peer review

This manuscript compared the functional outcomes among three radical prostatectomy procedures. The project design and analyses of the data are acceptable. The figures and tables well summarize the existing data. Overall the manuscript is well written.

## REFERENCES


1 **Siegel R**, Naishadham D, Jemal A. Cancer Statistics, 2013. *CA Cancer J Clin* 2013; **63**: 11-30

2 **Ferlay J**, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 2013; **49**: 1374-1403 [PMID: 23485231 DOI: 10.1016/j.ejca.2012.12.027]

3 **Heidenreich A**, Bastian PJ, Bellmunt J, Bolla M, Joniau S, van der Kwast T, Mason M, Matveev V, Wiegel T, Zattoni F, Mottet N. EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *Eur Urol* 2014; **65**: 124-137 [PMID: 24207135]

4 **Walsh PC**, Donker PJ. Impotence following radical prostatectomy: Insight into etiology and prevention. *J Urol* 1982; **128**: 492-497

5 **Bill-Axelson A**, Holmberg L, Ruutu M, Garmo H, Stark JR, Busch C, Nordling S, Häggman M, Andersson SO, Bratell S, Spångberg A, Palmgren J, Steineck G, Adami HO, Johansson JE. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med* 2011; **364**: 1708-1717 [PMID: 21542742 DOI: 10.1056/NEJMoa1011967]

6 **Ficarra V**, Novara G, Artibani W, Cestari A, Galfano A, Graefen M, Guazzoni G, Guillonneau B, Menon M, Montorsi F, Patel V, Rassweiler J, Van Poppel H. Retropubic, laparoscopic, and robot-assisted radical prostatectomy: a systematic review and cumulative analysis of comparative studies. *Eur Urol* 2009; **55**: 1037-1063 [PMID: 19185977 DOI: 10.1016/j.eururo.2009.01.036]

7 **Coelho RF**, Rocco B, Patel MB, Orvieto MA, Chauhan S, Ficarra V, Melegari S, Palmer KJ, Patel VR. Retropubic, laparoscopic, and robot-assisted radical prostatectomy: a critical review of outcomes reported by high-volume centers. *J Endourol* 2010; **24**: 2003-2015 [DOI: 10.1089/end.2010.0295]

8 **Ficarra V**, Novara G, Rosen RC, Artibani W, Carroll PR, Costello A, Menon M, Montorsi F, Patel VR, Stolzenburg JU, Van der Poel H, Wilson TG, Zattoni F, Mottrie A. Systematic review and meta-analysis of studies reporting urinary continence recovery after robot-assisted radical prostatectomy. *Eur Urol* 2012; **62**: 405-417 [PMID: 22749852 DOI: 10.1016/j.eururo.2012.05.045]

9 **Ficarra V**, Novara G, Ahlering TE, Costello A, Eastham JA, Graefen M, Guazzoni G, Menon M, Mottrie A, Patel VR, Van der Poel H, Rosen RC, Tewari AK, Wilson TG, Zattoni F, Montorsi F. Systematic review and meta-analysis of studies reporting potency rates after robot-assisted radical prostatectomy. *Eur Urol* 2012; **62**: 418-430 [DOI: 10.1016/j.eururo.2012.05.046]

10 **Kim SC**, Song C, Kim W, Kang T, Park J, Jeong IG, Lee S, Cho YM, Ahn H. Factors determining functional outcomes after radical prostatectomy: robot-assisted versus retropubic. *Eur Urol* 2011; **60**: 413-419 [PMID: 21612859 DOI: 10.1016/j.eururo.2011.05.011]

11 **Krambeck AE**, DiMarco DS, Rangel LJ, Bergstralh EJ, Myers RP, Blute ML, Gettman MT. Radical prostatectomy for prostatic adenocarcinoma: a matched comparison of open retropubic and robot-assisted techniques. *BJU Int* 2009; **103**: 448-453 [DOI: 10.1111/j.1464-410X.2008.08012.x]

12 **Stroup DF**, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008-2012 [PMID: 10789670]

13 **Wells GA**, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. Cited: 2013-10-10. Available from: URL: http: //www.ohri.ca/programs/clinical_epidemiology/oxford.htm

14 **Higgins JPT**, Altman DG, Sterne JAC, editors. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of InterventionsVersion 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from: URL: http: //www.cochrane-handbook.org

15 **Higgins JP**, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539
16 **Mantel N**, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; **22**: 719-748 [PMID: 13655060]
17 **Fidler V**, Nagelkerke N. The Mantel-Haenszel procedure revisited: models and generalizations. *PLoS One* 2013; **8**: e58327 [PMID: 23516463 DOI: 10.1371/journal.pone.0058327]
18 **Deeks JJ**, Higgins JPT, Altman DG, editors. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from: URL: http: //www.cochrane-handbook.org
19 **Egger M**, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634 [PMID: 9310563]
20 **Anastasiadis AG**, Salomon L, Katz R, Hoznek A, Chopin D, Abbou CC. Radical retropubic versus laparoscopic prostatectomy: a prospective comparison of functional outcome. *Urology* 2003; **62**: 292-297 [PMID: 12893338]
21 **Roumeguere T**, Bollens R, Vanden Bossche M, Rochet D, Bialek D, Hoffman P, Quackels T, Damoun A, Wespes E, Schulman CC, Zlotta AR. Radical prostatectomy: a prospective comparison of oncological and functional results between open and laparoscopic approaches. *World J Urol* 2003; **20**: 360-366 [PMID: 12682770]
22 **Remzi M**, Klingler HC, Tinzl MV, Fong YK, Lodde M, Kiss B, Marberger M. Morbidity of laparoscopic extraperitoneal versus transperitoneal radical prostatectomy verus open retropubic radical prostatectomy. *Eur Urol* 2005; **48**: 83-89; discussion 89 [PMID: 15967256]
23 **Wagner AA**, Link RE, Trock BJ, Sullivan W, Pavlovich CP. Comparison of open and laparoscopic radical prostatectomy outcomes from a surgeon's early experience. *Urology* 2007; **70**: 667-671 [PMID: 17991534]
24 **Touijer K**, Eastham JA, Secin FP, Romero Otero J, Serio A, Stasi J, Sanchez-Salas R, Vickers A, Reuter VE, Scardino PT, Guillonneau B. Comprehensive prospective comparative analysis of outcomes between open and laparoscopic radical prostatectomy conducted in 2003 to 2005. *J Urol* 2008; **179**: 1811-1817; discussion 1817 [PMID: 18353387 DOI: 10.1016/j.juro.2008.01.026]
25 **Greco F**, Wagner S, Hoda MR, Kawan F, Inferrera A, Lupo A, Reichelt O, Jurczok A, Hamza A, Fornara P. Laparoscopic vs open retropubic intrafascial nerve-sparing radical prostatectomy: surgical and functional outcomes in 300 patients. *BJU Int* 2010; **106**: 543-547 [PMID: 20067455 DOI: 10.1111/j.1464-410X.2009.09157.x]
26 **Dahl DM**, Barry MJ, McGovern FJ, Chang Y, Walker-Corkery E, McDougal WS. A prospective study of symptom distress and return to baseline function after open versus laparoscopic radical prostatectomy. *J Urol* 2009; **182**: 956-965 [PMID: 19616252 DOI: 10.1016/j.juro.2009.05.044]
27 **Egawa S**, Kuruma H, Suyama K, Iwamura M, Baba S. Delayed recovery of urinary continence after laparoscopic radical prostatectomy. *Int J Urol* 2003; **10**: 207-212 [PMID: 12657100]
28 **Artibani W**, Grosso G, Novara G, Pecoraro G, Sidoti O, Sarti A, Ficarra V. Is laparoscopic radical prostatectomy better than traditional retropubic radical prostatectomy? An analysis of peri-operative morbidity in two contemporary series in Italy. *Eur Urol* 2003; **44**: 401-406 [PMID: 14499672]
29 **Ghavamian R**, Knoll A, Boczko J, Melman A. Comparison of operative and functional outcomes of laparoscopic radical prostatectomy and radical retropubic prostatectomy: single surgeon experience. *Urology* 2006; **67**: 1241-1246 [PMID: 16678887]
30 **Takenaka A**, Soga H, Kurahashi T, Miyake H, Tanaka K, Fujisawa M. Early recovery of urinary continence after laparoscopic versus retropubic radical prostatectomy: evaluation of preoperative erectile function and nerve-sparing procedure as predictors. *Int Urol Nephrol* 2009; **41**: 587-593 [PMID: 18810650 DOI: 10.1007/s]
31 **Simforoosh N**, Javaherforooshzadeh A, Aminsharifi A, Tabibi A. Early continence after open and laparoscopic radical prostatectomy with sutureless vesicourethral alignment: an alternative technique, 8 years' experience. *Urol J* 2009; **6**: 163-169 [PMID: 19711268]
32 **Springer C**, Inferrera A, Pini G, Mohammed N, Fornara P, Greco F. Laparoscopic versus open bilateral intrafascial nerve-sparing radical prostatectomy after TUR-P for incidental prostate cancer: surgical outcomes and effect on postoperative urinary continence and sexual potency. *World J Urol* 2013; **31**: 1505-1510 [PMID: 23400788 DOI: 10.1007/s00345-013-1036-0]
33 **Magheli A**, Busch J, Leva N, Schrader M, Deger S, Miller K, Lein M. Comparison of surgical technique (open vs. laparoscopic) on pathological and long term functional outcomes following radical prostatectomy. *BMC Urol* 2014; **14**: 18 [PMID: 24506815 DOI: 10.1186/1471-2490-14-18]
34 **Tewari A**, Srivasatava A, Menon M. A prospective comparison of radical retropubic and robot-assisted prostatectomy: experience in one institution. *BJU Int* 2003; **92**: 205-210 [PMID: 12887468]
35 **Ficarra V**, Novara G, Fracalanza S, D'Elia C, Secco S, Iafrate M, Cavalleri S, Artibani W. A prospective, non-randomized trial comparing robot-assisted laparoscopic and retropubic radical prostatectomy in one European institution. *BJU Int* 2009; **104**: 534-539 [PMID: 19281468 DOI: 10.1111/j.1464-410X.2009.08419.x]
36 **Ham WS**, Park SY, Kim WT, Koo KC, Lee YS, Choi YD. Open versus robotic radical prostatectomy: a prospective analysis based on a single surgeon's experience. *J Robotic Surg* 2008; **2**: 235-241[DOI: 10.1007/s11701-008-0111-9]
37 **Di Pierro GB**, Baumeister P, Stucki P, Beatrice J, Danuser H, Mattei A. A prospective trial comparing consecutive series of open retropubic and robot-assisted laparoscopic radical prostatectomy in a centre with a limited caseload. *Eur Urol* 2011; **59**: 1-6 [PMID: 21035248 DOI: 10.1016/j.eururo.2010.10.026]
38 **Geraerts I**, Van Poppel H, Devoogdt N, Van Cleynenbreugel B, Joniau S, Van Kampen M. Prospective evaluation of urinary incontinence, voiding symptoms and quality of life after open and robot-assisted radical prostatectomy. *BJU Int* 2013; **112**: 936-943 [PMID: 23937206 DOI: 10.1111/bju.12258]
39 **Caballero Romeu JP**, Palacios Ramos J, Pereira Arias JG, Gamarra Quintanilla M, Astobieta Odriozola A, Ibarluzea González G. [Radical prostatectomy: evaluation of learning curve outcomes laparoscopic and robotic-assisted laparoscopic techniques with radical retropubic prostatectomy]. *Actas Urol Esp* 2008; **32**: 968-975 [PMID: 19143287]
40 **Rocco B**, Matei DV, Melegari S, Ospina JC, Mazzoleni F, Errico G, Mastropasqua M, Santoro L, Detti S, de Cobelli O. Robotic vs open prostatectomy in a laparoscopically naive centre: a matched-pair analysis. *BJU Int* 2009; **104**: 991-995 [PMID: 19426191 DOI: 10.1111/j.1464-410X.2009.08532.x]
41 **Ou YC**, Yang CR, Wang J, Cheng CL, Patel VR. Comparison of robotic-assisted versus retropubic radical prostatectomy performed by a single surgeon. *Anticancer Res* 2009; **29**: 1637-1642 [PMID: 19443379]
42 **Choo MS**, Choi WS, Cho SY, Ku JH, Kim HH, Kwak C. Impact of prostate volume on oncological and functional outcomes after radical prostatectomy: robot-assisted laparoscopic versus open retropubic. *Korean J Urol* 2013; **54**: 15-21 [PMID: 23362442 DOI: 10.4111/kju.2013.54.1.15]
43 **Son SJ**, Lee SC, Jeong CW, Jeong SJ, Byun SS, Lee SE. Comparison of Continence Recovery Between Robot-Assisted Laparoscopic Prostatectomy and Open Radical Retropubic Prostatectomy: A Single Surgeon Experience. *Korean J Urol*

2013; **54**: 598-602 [DOI: 10.4111/kju.2013.54.9.598]
44 **Asimakopoulos AD**, Pereira Fraga CT, Annino F, Pasqualetti P, Calado AA, Mugnier C. Randomized comparison between laparoscopic and robot-assisted nerve-sparing radical prostatectomy. *J Sex Med* 2011; **8**: 1503-1512 [PMID: 21324093 DOI: 10.1111/j.1743-6109.2011.02215.x]
45 **Porpiglia F**, Morra I, Lucci Chiarissi M, Manfredi M, Mele F, Grande S, Ragni F, Poggio M, Fiori C. Randomised controlled trial comparing laparoscopic and robot-assisted radical prostatectomy. *Eur Urol* 2013; **63**: 606-614 [PMID: 22840353 DOI: 10.1016/j.eururo.2012.07.007]
46 **Ploussard G,** de la Taille A, Moulin M, Vordos D, Hoznek A, Abbou CC, Salomon L. Comparisons of the perioperative, functional, and oncologic outcomes after robot-assisted versus pure extraperitoneal laparoscopic radical prostatectomy. *Eur Urol* 2014; **65**: 610-619 [PMID: 23245815 DOI: 10.1016/j.eururo.2012.11.049]
47 **Joseph JV**, Vicente I, Madeb R, Erturk E, Patel HR. Robot-assisted vs pure laparoscopic radical prostatectomy: are there any differences? *BJU Int* 2005; **96**: 39-42 [PMID: 15963117]
48 **Lee HW**, Lee HM, Seo SI. Comparison of initial surgical outcomes between laparoscopic radical prostatectomy and robot-assisted laparoscopic radical prostatectomy performed by a single surgeon. *Korean J Urol* 2009; **50**: 468-474 [DOI: 10.4111/kju.2009.50.5.468]
49 **Cho JW**, Kim TH, Sung GT. Laparoscopic radical prostatectomy versus robot-assisted laparoscopic radical prostatectomy: a single surgeon's experience. *Korean J Urol* 2009; **50**: 1198-202 [DOI: 10.4111/kju.2009. 50.12.1198]
50 **Hakimi AA**, Blitstein J, Feder M, Shapiro E, Ghavamian R. Direct comparison of surgical and functional outcomes of robotic-assisted versus pure laparoscopic radical prostatectomy: single-surgeon experience. *Urology* 2009; **73**: 119-123 [PMID: 18952268 DOI: 10.1016/j.urology.2008.08.491]
51 **Trabulsi EJ**, Zola JC, Gomella LG, Lallas CD. Transition from pure laparoscopic to robotic-assisted radical prostatectomy: a single surgeon institutional evolution. *Urol Oncol* 2010; **28**: 81-85 [PMID: 20123354 DOI: 10.1016/j.urolonc.2009.07.002]
52 **Willis DL**, Gonzalgo ML, Brotzman M, Feng Z, Trock B, Su LM. Comparison of outcomes between pure laparoscopic vs robot-assisted laparoscopic radical prostatectomy: a study of comparative effectiveness based upon validated quality of life outcomes. *BJU Int* 2012; **109**: 898-905 [PMID: 21933328 DOI: 10.1111/j.1464-410X]
53 **Park JW**, Won Lee H, Kim W, Jeong BC, Jeon SS, Lee HM, Choi HY, Seo SI. Comparative assessment of a single surgeon's series of laparoscopic radical prostatectomy: conventional versus robot-assisted. *J Endourol* 2011; **25**: 597-602 [PMID: 21438677 DOI: 10.1089/end.2010.0229]
54 **Park B**, Kim W, Jeong BC, Jeon SS, Lee HM, Choi HY, Seo SI. Comparison of oncological and functional outcomes of pure versus robotic-assisted laparoscopic radical prostatectomy performed by a single surgeon. *Scand J Urol* 2013; **47**: 10-18 [PMID: 22835035 DOI: 10.3109/00365599.2012.696137]
55 **Rassweiler J**, Seemann O, Schulze M, Teber D, Hatzinger M, Frede T. Laparoscopic versus open radical prostatectomy: a comparative study at a single institution. *J Urol* 2003; **169**: 1689-1693 [PMID: 12686809]
56 **Ayyathurai R**, Manoharan M, Nieder AM, Kava B, Soloway MS. Factors affecting erectile function after radical retropubic prostatectomy: results from 1620 consecutive patients. *BJU Int* 2008; **101**: 833-836 [PMID: 18190627 DOI: 10.1111/j.1464-410X.2007.07409.x]
57 **Briganti A**, Gallina A, Suardi N, Capitanio U, Tutolo M, Bianchi M, Passoni N, Salonia A, Colombo R, Di Girolamo V, Guazzoni G, Rigatti P, Montorsi F. Predicting erectile function recovery after bilateral nerve sparing radical prostatectomy: a proposal of a novel preoperative risk stratification. *J Sex Med* 2010; **7**: 2521-2531 [PMID: 20487236 DOI: 10.1111/j.1743-6109.2010.01845.x]
58 **Stanford JL**, Feng Z, Hamilton AS, Gilliland FD, Stephenson RA, Eley JW, Albertsen PC, Harlan LC, Potosky AL. Urinary and sexual function after radical prostatectomy for clinically localized prostate cancer: the Prostate Cancer Outcomes Study. *JAMA* 2000; **283**: 354-360 [PMID: 10647798]
59 **Shikanov S**, Desai V, Razmaria A, Zagaja GP, Shalhav AL. Robotic radical prostatectomy for elderly patients: probability of achieving continence and potency 1 year after surgery. *J Urol* 2010; **183**: 1803-1807 [DOI: 10.1016/j.juro.2010.01.016]
60 **Stang A**. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol* 2010; **25**: 603-605

**P- Reviewer:** Liu YY, Papatsoris AG **S- Editor:** Ji FF
**L- Editor:** Wang TQ **E- Editor:** Lu YJ



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i3.127

*World J Meta-Anal* 2014 August 26; 2(3): 127-134
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Decision-tree analysis for cost-effective management of solitary pulmonary nodules in China

Bei Lu, Li-Xin Sun, Xi Yan, Zhen-Zhong Ai, Jin-Zhi Xu

Bei Lu, Li-Xin Sun, Xi Yan, Zhen-Zhong Ai, Jin-Zhi Xu, Department of Thoracic Surgery, the Fourth Affiliated Hospital of Harbin Medical University, Harbin 150001, Heilongjiang Province, China

Author contributions: This research paper was written by Lu B as the first author under supervision of Xu JZ; Sun LX retrieved and analyzed all medical literatures; Yan X and Ai ZZ prepared the illustrations and the statistical analysis.

Correspondence to: Dr. Jin-Zhi Xu, Department of Thoracic Surgery, the Fourth Affiliated Hospital of Harbin Medical University, No.37, Yi yuan Street, Nan gang District, Harbin 150001, Heilongjiang Province, China. 13836027105@139.com
Telephone: +86-451-82576963
Received: February 25, 2014 Revised: May 17, 2014
Accepted: July 18, 2014
Published online: August 26, 2014

## Abstract

**AIM:** To analyze the cost-effectiveness of the diagnosis of solitary pulmonary nodule (SPN) in China.

**METHODS:** Decision analysis models were constructed to assess the cost-effectiveness of four strategies for the management of SPN: computed tomography (CT) alone, CT plus CT-guided automated cutting needle biopsy (ACNB), CT plus positron emission tomography/computed tomography (PET/CT), CT plus diffusion-weighted magnetic resonance imaging (DWI) plus PET/CT.

**RESULTS:** The prevalence of lung cancer among SPN discovered in the clinical setting was approximately 50%. The CT plus ACNB strategy had higher diagnostic accuracies (87% *vs* 81%), with a cost saving of ￥1945 RMB per patient, and reducing unnecessary thoracotomy by 16.5%; this was associated with a 4.5% missed diagnosis rate. CT plus DWI plus PET/CT strategy also had higher accuracies (95% *vs* 81%), with a cost saving of ￥590 RMB per patient, and reducing unnecessary thoracotomy by 13.5%; this was accompanied by 0.3% missed diagnosis rate. CT plus PET strategy is cost effective at a prevalence rate of 0-34%, but there was a larger prevalence range of lung cancer for CT plus ACNB strategy (from 0 to 0.6) and CT plus DWI plus PET/CT strategy(from 0 to 0.64).

**CONCLUSION:** CT plus DWI plus PET/CT strategy was cost-effective, and had a higher accuracy accompanied by a lower missed diagnosis rate than CT plus ACNB strategy.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Solitary pulmonary nodules; Diffusion-weighted Magnetic resonance imaging; Computed tomography-guided automated cutting needle biopsy; Positron emission tomography/computed tomography; Cost effectiveness

**Core tip:** It has become a major concern how to manage solitary pulmonary nodule (SPN) discovered in the clinical setting with low cost and high accuracy and reduce the morbidity and mortality associated with radiation, biopsy, and surgical procedures. However, up to now there has not been an analysis of cost-effectiveness of various strategies for the diagnosis and management of SPN in China. Recent studies in other countries may not hold true in a Chinese hospital because of differences in health-care systems and diagnostic strategy. We are the first to perform such an analysis for the cost-effective management of solitary pulmonary nodules in China.

Lu B, Sun LX, Yan X, Ai ZZ, Xu JZ. Decision-tree analysis for cost-effective management of solitary pulmonary nodules in China. *World J Meta-Anal* 2014; 2(3): 127-134 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i3/127.htm
DOI: http://dx.doi.org/10.13105/wjma.v2.i3.127

## INTRODUCTION

Lung cancer has become the leading cause of cancer deaths in China because the incidence and mortality are rapidly increasing. Although earlier detection of peripheral lung cancer may reduce the mortality rate, population-based screening using plain chest radiography has been not carried out in China. Therefore, most solitary pulmonary nodules (SPN), which are usually discovered in the clinical setting on hospital in China, have a higher malignant prevalence[1,2]. It has become a major concern how to manage SPN discovered in the clinical setting with low cost and high accuracy and reduce the morbidity and mortality associated with radiation, biopsy, and surgical procedures.

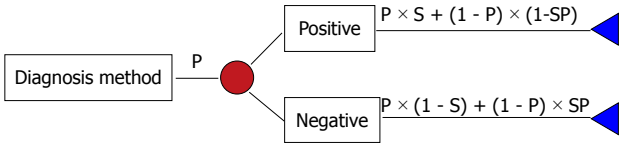
However, up to now there has not been an analysis of cost-effectiveness of various strategies for the diagnosis and management of SPN in China. Recent studies in other countries may not hold true in a Chinese hospital because of differences in health-care systems and diagnostic strategy[3-7]. The cost of surgery is not obviously higher than that of automated cutting needle biopsy (ACNB, six times) and positron emission tomography/computed tomography (PET/CT, three times) in China. But in the United States, Japan and European countries the cost of surgery was twenty times as much as that of biopsy and PET/CT. Because of high cost, PET/CT has not gained widespread popularity in most areas of China. A recent study showed that diffusion-weighted magnetic resonance imaging (DWI) may be able to be used in place of PET/CT for discriminating malignant from benign pulmonary nodules and is associated with significantly less false positives and lower costs than PET/CT[8]. Furthermore, in general, physicians in Chinese hospitals do not perform thoracoscopy with local anesthesia in patients who have or are suspected of having lung cancer.

According to the upper analysis, it is necessary to borrow previous study methods to evaluate the role of CT, ACNB, DWI, and PET/CT in the investigation of SPN which are discovered in the clinical setting in hospital in China. We used decision-tree analysis models and compared four strategies to understand under what conditions various strategies should be cost-effective and have high diagnostic accuracy.

## MATERIALS AND METHODS

In the current study, we have defined SPN as spherical intrapulmonary x-ray densities less than 4.0 cm in diameter with no calcium visible on a standard chest X-ray. There is no associated atelectasis, hilar enlargement, or pleural effusion. We do not include patients with any evidence of metastasis or a recent primary malignancy outside of the chest. We also exclude cases that have previous radiographs that have already established the stability of the rate of growth of the nodule. Multiple nodules are also not considered in the current analysis.

The analysis for cost-effectiveness was performed using quantitative methods of decision analysis. Decision-tree models were constructed with multiple competing strategies, and reported values of prior probability of cancer, and sensitivity and specificity of each diagnostic modality were applied to them using Chinese health care costs. The average cost per patient for each strategy, including all diagnostic tests and surgery when undertaken, was calculated. The medical literature was surveyed to obtain the performance of each diagnostic test. For the calculations, Decision Analysis Add-In for Microsoft Excel (Version1.0.6, Palisade Corporation) was used. The decision tree was constructed using the choices and potential outcomes of the choices. All conditional probabilities of each outcome in the tree were calculated and obtained as a function of the variables listed in Tables 1 and 2 by using Bayesian analysis, Calculation Methods:




Where P = prevalence, S = sensitivity, and Sp = specificity.

Calculations of overall cost of competing strategies were calculated by summing the products of the probabilities and values of the outcome of each strategy. Overall costs per patient in each strategy can be calculated automatically by the software and listed under "chance" in Figure 1. Accuracies (A) of each diagnosis method in each strategy were calculated by formula:

$A = (S - Sp)P + Sp$

False positive (FP), false negatives (FN), true positive (TP) and true negatives (TN) of each diagnosis method in each strategy were calculated by the formula:

$FP = 1/\{[(Sp-A)S/(1-Sp)(A-S)]+1\}$

$FP = 1/\{[(S-A)Sp/(1-S)(A-Sp)]+1\}$

Overall accuracies of each strategy were calculated by summing the products of the probabilities of the outcome of each strategy and TP plus TN in each strategy; The overall misdiagnosis rate of each strategy was calculated by summing the products of the probabilities of the outcome of each strategy and FP in each strategy; the overall missed diagnosis rate of each strategy was calculated by summing the products of the probabilities of the outcome of each strategy and FN in each strategy.

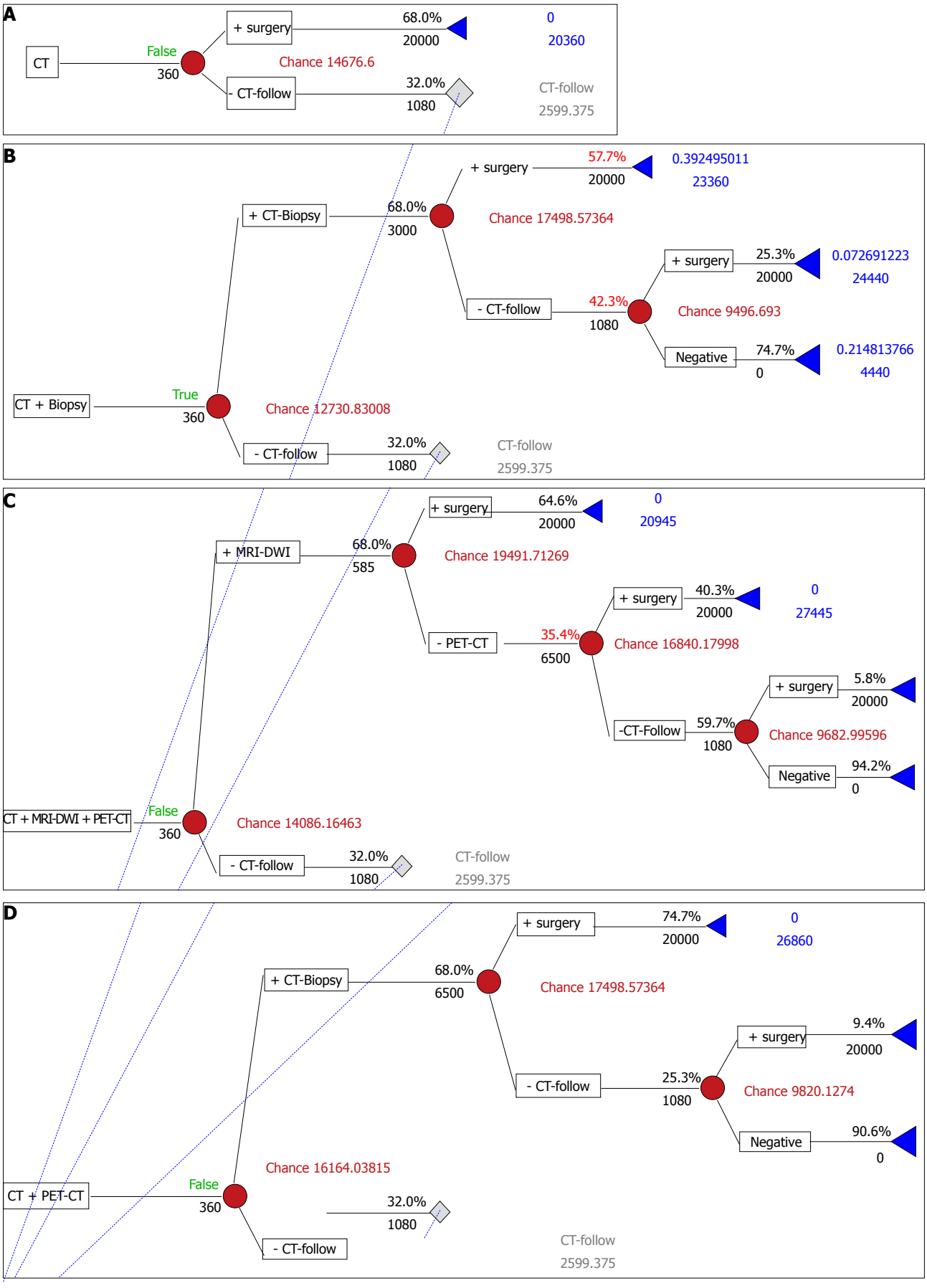
Cost-effectiveness was expressed as the incremental cost-accuracy ratio (ICAR) where:

$ICAR = (Cost_{strat}-Cost_{bl}/(Accuracy_{strat}-Accuracy_{bl})$ (RMB/%)[9].

$Cost_{strat}$ and $Accuracy_{strat}$ are the average cost per patient and accuracy of the strategy being compared, and $Cost_{bl}$ and $Accuracy_{bl}$ are the cost per patient and accuracy of a baseline strategy, which was the CT alone strategy in this study. A negative ICAR resulting from a negative numerator and positive denominator indicates that the strategy being compared is clearly preferred. In the case of a positive ICAR, a criterion must be chosen that gives the acceptable cost to get a higher accuracy.

### *Decision-tree analysis models*

We compared four strategies for the management of


A
CT
False
360
+ surgery
68.0%
20000
0
20360
Chance 14676.6
- CT-follow
32.0%
1080
CT-follow
2599.375
B
CT + Biopsy
True
360
+ CT-Biopsy
68.0%
3000
+ surgery
57.7%
20000
0.392495011
23360
Chance 17498.57364
- CT-follow
42.3%
1080
+ surgery
25.3%
20000
0.072691223
24440
Chance 9496.693
Negative
74.7%
0
0.214813766
4440
Chance 12730.83008
- CT-follow
32.0%
1080
CT-follow
2599.375
C
CT + MRI-DWI + PET-CT
False
360
+ MRI-DWI
68.0%
585
+ surgery
64.6%
20000
0
20945
Chance 19491.71269
- PET-CT
35.4%
6500
+ surgery
40.3%
20000
0
27445
Chance 16840.17998
-CT-Follow
59.7%
1080
+ surgery
5.8%
20000
Chance 9682.99596
Negative
94.2%
0
Chance 14086.16463
- CT-follow
32.0%
1080
CT-follow
2599.375
D
CT + PET-CT
False
360
+ CT-Biopsy
68.0%
6500
+ surgery
74.7%
20000
0
26860
Chance 17498.57364
- CT-follow
25.3%
1080
+ surgery
9.4%
20000
Chance 9820.1274
Negative
90.6%
0
Chance 16164.03815
- CT-follow
32.0%
1080
CT-follow
2599.375

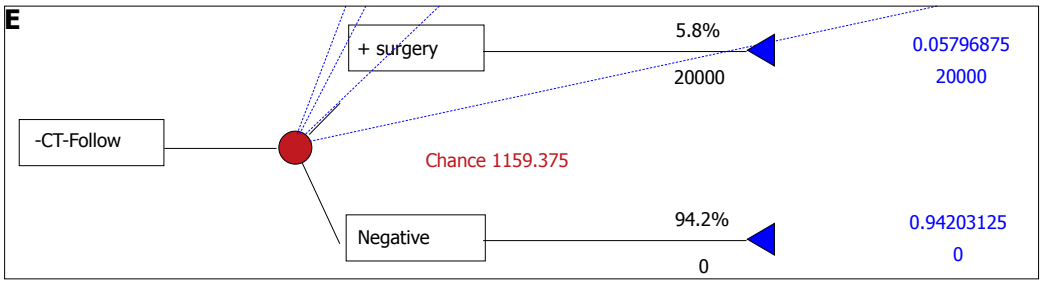



**Figure 1 Strategies for the management of solitary pulmonary nodule.** A: CT alone strategy (base line); B: CT plus CT-guided needle biopsy strategy; C: CT plus DWI plus PET strategy; D: CT plus PET strategy. PET: Positron emission tomography; CT: Computed tomography; DWI: Diffusion-weighted magnetic resonance imaging.

**Table 1 Performance parameters used in the decision-tree analyses**

| | Sensitivity | Specificity | Ref. |
|---|---|---|---|
| Chest CT | 0.99 | 0.63 | [5] |
| PET-CT | 0.97 | 0.85 | [5] |
| ACNB | 0.769 | 0.936 | [5] |
| CT-Follow | 0.56 | 0.95 | [3] |
| DWI | 0.7 | 0.97 | [3] |

PET: Positron emission tomography; CT: Computed tomography; DWI: Diffusion-weighted magnetic resonance imaging; ACNB: Automated cutting needle biopsy.

**Table 2 Medical costs in China used in the decision-tree analyses**

| | Cost (RMB) |
|---|---|
| Chest CT without contrast enhancement | ￥360 |
| PET-CT | ￥6500 |
| ACNB[1] | ￥3000 |
| CT-Follow: continuous three Chest CT | ￥1080 |
| DWI | ￥585 |
| Surgical resection of SPN | ￥20000 |

[1]Including the costs of possible treatment of pneumothorax. PET: Positron emission tomography; CT: Computed tomography; DWI: Diffusion-weighted magnetic resonance imaging; SPN: Solitary pulmonary nodule; ACNB: Automated cutting needle biopsy.

SPN (Figure 1): (1) CT alone strategy (baseline), (2) CT plus PET/CT strategy, (3) CT plus ACNB strategy, and (4) a CT-plus-DWI plus PET/CT strategy.

In these strategies, all patients initially undergo a chest CT examination without contrast enhancement. If an initial CT diagnosis of benign pulmonary nodule is made or no SPN is demonstrated on CT, the patients are followed up using an unenhanced chest CT. When follow-up CT shows the growth of the SPN, the diagnosis of lung cancer is made: (1) CT alone strategy: The CT alone strategy simulates a simple diagnostic approach to the management of patients with SPN. In this strategy, all patients, in whom SPN is diagnosed as lung cancer on the initial chest CT, proceed to surgical resection without pathological confirmation; (2) CT plus PET strategy: In this strategy, all patients who are CT positive (*i.e.*, an SPN is diagnosed as lung cancer) undergo chest PET-CT. If the PET/CT is then also positive, the patients proceed to surgical resection. If a chest PET/CT examination is negative after a positive CT examination, the patients are followed up by an unenhanced chest CT; (3) CT plus ACNB strategy: If the initial CT is positive, it is followed directly by ACNB. If the biopsy is then also positive, patient undergo surgical resection. If ACNB is negative, the patients are followed up by an unenhanced chest CT; (4) CT plus DWI plus PET/CT strategy: In this strategy, all patients who are CT positive undergo a chest CT plus DWI strategy. If the CT plus DWI strategy is then also positive, the patients proceed to surgical resection. If a chest a CT plus DWI strategy examination is negative after a positive CT examination, the patients are given an extra CT-plus-PET strategy above-mentioned.

### *Prevalence of cancer and diagnostic performance (sensitivity and specificity)*

Although there were no large sample statistics for epidemiology, some small sample studies and our experience indicated the prevalence of lung cancer among SPN was approximately 50% in China[1,2]. Therefore, we applied this value (50%) to the decision-tree analyses.

The diagnostic performance (sensitivity and specificity) of each diagnostic test are gleaned from the literature and entered into the models (Table 1). Although multi-slice spiral CT is taking the place of conventional thin slice CT gradually, the diagnostic performance of multi-slice spiral CT on SPN still lack of large sample statistics. Tsubamoto *et al*[10] reported that accuracy of the final diagnosis based on coronal multi-planar reconstruction of with a multi-detector-row CT scanner (74%) was almost equal to that based on transverse thin-section CT (71%) ($P = 0.3$) for the evaluation of solitary pulmonary nodules. As a result, we still used CT sensitivity of 99% and specificity of 63% in the largest series published by Siegelman *et al*[11], their study included 634 nodules, using conventional thin slice CT. From many published reports of PET/CT imaging of SPN, we adopted the results of Yi *et al*[12] and Kim *et al*[13]. We applied PET/CT sensitivity of 97% and specificity of 85% to the decision-tree analyses. The diagnostic accuracy of CT-guided transthoracic ACNB has been reported based on comparisons with

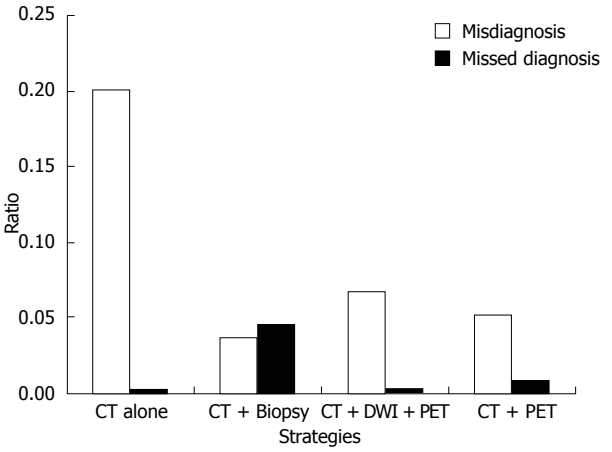



**Figure 2 Ratio of the patients who undergo unnecessary surgery for a benign solitary pulmonary nodule and in whom lung cancer is misdiagnosed as a benign nodule.** The strategies using CT + DWI + PET and CT + ACNB dramatically decrease the ratio of the patients who undergo unnecessary surgery for a benign SPN, although the ratio of cancer which is missed at diagnosis is increased. PET: Positron emission tomography; CT: Computed tomography; DWI: Diffusion-weighted magnetic resonance imaging; SPN: Solitary pulmonary nodule; ACNB: Automated cutting needle biopsy.

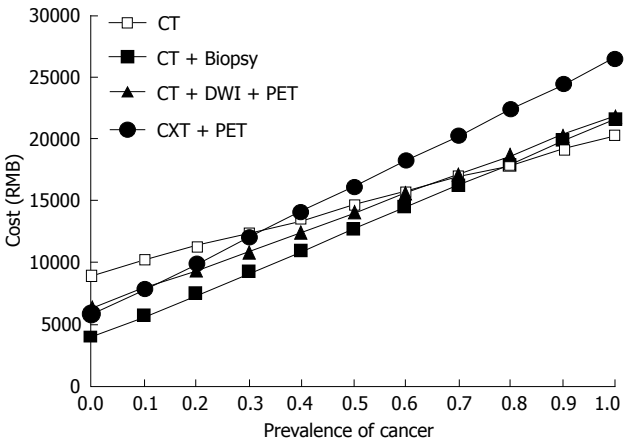



**Figure 3 Correlation between the prevalence of cancer and total costs per patient.** The total costs per patients were increased regardless of the strategies and the difference among the strategies became small, although the CT+PET alone strategy had the highest cost at the prevalence of 50%. PET: Positron emission tomography; CT: Computed tomography.

fine-needle aspiration biopsy. In China, ACNB has been often used by physicians, usually without combining the use of a long-throw biopsy needle, higher mean number of needle passes, and tandem system like Satoh *et al*[14]; as a result we selected ACNB sensitivity of 76.9% and specificity of 93.6% by Tsukada *et al*[15]. Although serial CT is probably the most common follow-up methodology to be adopted in the practice, there is little data on the utility of follow-up chest CT for the diagnosis of SPN. Takashima *et al*[16] reported the sensitivity of 0.56 and specificity of 0.95, when the nodule showing an increase of 0.5 mm or more in the maximum or perpendicular diameter or both on high-resolution CT images was diagnosed as a lung cancer. We applied these values to our analysis. There is a few data on the utility of DWI for the diagnosis of SPN. The study by Mori *et al*[8] reported the sensitivity of 0.70 and specificity of 0.97, which is similar to our unpublished results, we applied these values to our analysis.

### Cost of each medical procedure

The mean costs of diagnostic tests and thoracotomy in China are shown in Table 2. The cost of thoracotomy in RMB was based on the bills in our hospital during 2009 ($n$ = 20). The costs of diagnostic examinations include the costs of diagnostic procedure and radiological or/and pathological interpretations based on the data of the Harbin Price Bureau (These costs are uniform in China according to the Ministry of Public Health of the People's Republic of China). In China, patients undergoing ACNB do not stay overnight in a hospital. Although major complications are rare, pneumothorax is the most common complication after ACNB with a reported rate of 22%-54%[14,17,18]. According to our study, chest tube placement was not necessary for the incidence of pneumothorax. A central venous catheter attached to a negative pressure aspirator was used for the treatment of pneumothorax in our hospital. The costs of possible treatment of pneumothorax were included in the cost of CT-guided needle biopsy.

### Statistical analysis

The decision tree methodology can be extended by performing sensitivity analysis to determine the conditions under which the new test remains cost-effective. This was achieved by entering different values for the disease prevalence (0-100%) into the decision-tree models. Although the costs of medical procedures will clearly vary among countries, within China these costs are fixed and therefore need not constitute a variable in the sensitivity analysis.

## RESULTS

The prevalence of lung cancer among SPN discovered in hospital in the clinical setting was approximately 50%. In this prevalence, the strategies using CT plus ACNB and CT plus DWI plus PET/CT were the cost-effective alternatives to the CT alone strategy. The CT plus ACNB strategy had higher accuracies (87% *vs* 81%), with a cost saving of ￥1945 RMB per patient, and reduced the number of candidates who underwent unnecessary thoracotomy for a benign SPN by 16.5%; this was accompanied by a rate of missed diagnosis of 4.5%. The strategies using CT plus DWI plus PET/CT had also higher accuracies (95% *vs* 81%), with a cost saving of ￥590 RMB per patient, and reduced unnecessary thoracotomy by 13.5%; this was accompanied by a rate of missed diagnosis of 0.3% (Table 3, Figure 2).

At a prevalence of cancer up to 34%, CT plus PET strategy is cost effective with a higher accuracy than CT alone strategy. But CT plus ACNB strategy and CT plus DWI plus PET/CT strategy had a better cost-effective-

**Table 3 Overall preoperative diagnostic accuracy, costs per patient and incremental cost-accuracy ratio in each strategy, when the prevalence of cancer is 50%**

| Strategy | Accuracy | Cost per patient | ICAR (RMB/%) |
|---|---|---|---|
| CT alone | 0.81 | 14676 | Baseline |
| CT + PET-CT | 0.952 | 16164 | 8196 |
| CT + ACNB | 0.867 | 12731 | -17583 |
| CT + DWI + PET-CT | 0.949 | 14086 | -4247 |

PET: Positron emission tomography; CT: Computed tomography; DWI: Diffusion-weighted magnetic resonance imaging; ICAR: Incremental cost-accuracy ratio; ACNB: Automated cutting needle biopsy.

ness over a larger prevalence of lung cancer, ranging from 0 to 0.60 (CT plus ACNB) and from 0 to 0.63 (CT plus DWI plus PET/CT).

The total costs per patient were increased regardless of the strategies, and the difference among the strategies became small (Figure 3). By reducing the number of the candidates who undergo unnecessary thoracotomy for a benign SPN (Figure 2), CT plus ACNB strategy and CT plus DWI plus PET/CT strategy decreased the total costs per patients. The total costs per patient of CT plus DWI plus PET/CT strategy were reduced not only by reducing unnecessary thoracotomy by 13.5%, but also by giving 44% of the candidates a chance to avoid undergoing expensive PET/CT, meanwhile, the overall accuracy of the CT plus DWI plus PET/CT strategy was not lower than that of CT plus PET strategy (Figure 4).

ICAR of the strategies using CT plus ACNB and CT plus DWI plus PET/CT were negative values with higher accuracies, clearly representing cost-effectiveness compared to the CT alone strategy (Figure 5). Although CT plus ACNB strategy is more cost saving, CT plus DWI plus PET/CT strategy has a lower rate of missed diagnosis (Figure 2) and a higher overall accuracy (Figure 4).

Figure 5 plots the prevalence of cancer versus ICAR. Since the CT alone strategy had a higher cost at the prevalence of cancer between 0 and 64% (Figure 3), in our analysis, a negative ICAR means that CT plus ACNB strategy and CT plus DWI plus PET/CT strategy have lower costs with higher accuracies. At the prevalence of cancer up to 34%, all strategies were cost-effective.

## DISCUSSION

From the analyses performed here, adding ACNB, DWI, or PET/CT is advantageous in terms of greater accuracy and cost-effectiveness. The results of our sensitivity analyses showed that CT plus ACNB strategy and CT plus DWI plus PET/CT strategy have a wide range of prevalence of cancer, and CT alone strategy becomes more cost-effective only when the prevalence of cancer exceeds approximately 65%.

In the study, we assumed that SPN evaluated were discovered in the clinical setting in hospital. In our daily clinical setting, the prevalence of lung cancer among SPN was approximately 50%, far higher than those detected

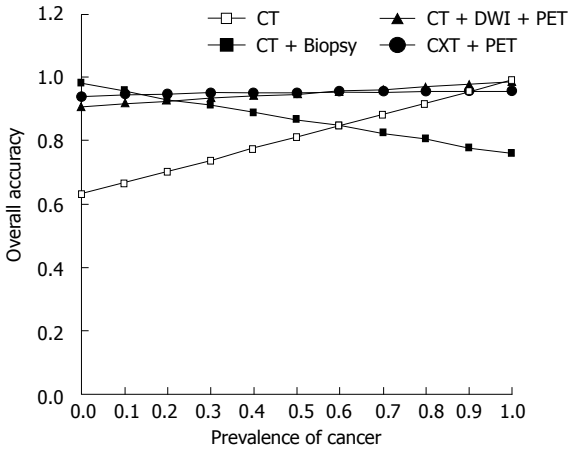



**Figure 4 Correlation between the prevalence of cancer and the overall accuracy.** The use of CT+PET or CT+DWI+PET resulted in a higher accuracy between the prevalence of 20%-90% compared to CT alone strategy and CT+Biopsy. PET: Positron emission tomography; CT: Computed tomography; DWI: Diffusion-weighted magnetic resonance imaging.

on lung cancer screening. For SPN discovered in the lung cancer screening, for a lower prevalence (approximately 10%), reducing false-positives to avoid more morbidity, mortality and increased costs for unnecessary thoracotomy is of importance; but for SPN discovered in the clinical setting in hospital, with a very high pretest probability, avoiding a missed diagnosis for false-negatives is as important as decreasing the misdiagnosis rate and reducing the cost. In the current study, CT plus ACNB strategy had a more cost saving and lower misdiagnosis rate, but also accompanied by higher rate of missed diagnosis up to 4.5%. The strategies using CT plus DWI plus PET/CT had higher accuracies (95% *vs* 81%), accompanied by a lower rate of missed diagnosis of 0.3%. This means there is an increase in life expectancy for decreasing rate of missed diagnosis.

The current analysis implies that the use of CT plus ACNB strategy or CT plus DWI plus PET/CT strategy can provide an advantage of a lower cost with higher accuracy. In the prevalence of cancer of 50%, the CT plus DWI plus PET/CT strategy is the optimal choice. In the United States, Japan and European countries the cost of surgery was twenty times as much as that of PET/CT. But in China the cost of surgery is only three times as much as that of PET/CT. Therefore, by reducing the number of the candidates who undergo unnecessary thoracotomy for a benign SPN, CT plus PET/CT strategy only decreased limited total costs per patients, which was hard to compensate for expensive cost of PET/CT, especially when the prevalence of lung cancer was up to 34%. A recent study showed that DWI may be able to be used in place of PET/CT for discriminating malignant from benign pulmonary nodules and is associated with significantly less false positive and lower cost than PET/CT[8]. In our unpublished study, a false-positive of DWI for SPN was only 3.5%. We assumed that most malignant SPN first could be identified by undergoing DWI previous to PET/CT, then negative nodules were added to

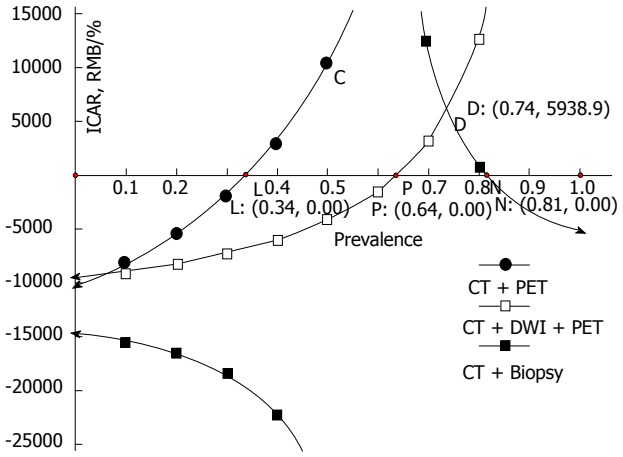



**Figure 5 Incremental cost-accuracy ratio compared with computed tomography alone strategy as the baseline.** A negative ICAR means cost-effectiveness. At the prevalence of cancer up to 34%, all strategies being compared are cost effective; in particular, the CT plus PET plus DWI strategy is cost-effective at the prevalence up to 63%. ICAR: Incremental cost-accuracy ratio; PET: Positron emission tomography; CT: Computed tomography; DWI: Diffusion-weighted magnetic resonance imaging.

PET/CT examination at a relatively low prevalence of lung cancer. The current study showed that the CT plus DWI plus PET/CT strategy decreased the total costs per patients not only by reducing unnecessary thoracotomy by 13.5%, but also by giving 44% of the candidates a chance to avoid undergoing expensive PET/CT; meanwhile, overall accuracies of CT plus DWI plus PET/CT strategy was not significantly lower than that of CT plus PET strategy.

ICAR of CT plus ACNB strategy are lower than that of the CT plus DWI plus PET/CT strategy, suggesting better cost-effectiveness of ACNB for SPN diagnosis, but we considered CT plus ACNB strategy is not an optimal option for the evaluation of SPN according to the analysis of current results. Most importantly, overall accuracies of CT plus ACNB strategy decreased gradually with the prevalence of lung cancer increasing, and was similar to that of CT alone strategy at a prevalence of lung cancer up to 50%. With the accompaniment of this, CT plus ACNB strategy had a rate of missed diagnosis of 4.5% for the evaluation of SPN, which was significantly higher than 0.3% of CT plus DWI plus PET/CT strategy. The high rate of missed diagnosis will inevitably reduce life expectancy. Secondly, unlike those discovered by screening chest radiography, most SPN discovered in the clinical setting in hospital are usually larger than 1 cm in diameter and with a high pretest probability. For such a large SPN, PET/CT has a higher negative predictive value and a lower rate of missed diagnosis; if this is accompanied with DWI with a higher positive predictive value and lower misdiagnosis rate, there will be an increase in cost-effectiveness. Our results underscored the point. In addition, although major complications are rare, pneumothorax is the most common complication with a reported rate of 22%-54% with ACNB[14,17,18]. A central venous catheter attached to negative pressure aspirator was used for the treatment of pneumothorax in our hospital. These invasions not only increased medical costs, but also caused some physical or mental injuries.

The current study is entirely a statistical simulation, and individual variation should and will inevitably occur in real medical practice. Therefore, the utility of a decision tree analysis must always be limited. Patient preferences and concerns must be considered when determining how to manage an individual patient. Although, according to this analysis, the CT plus DWI plus PET strategy was the optimal choice, the selection of strategy depends not only on the cost-effectiveness, but also on variable patient factors and accessibility to the modalities. Risk-taking attitudes of the patient and physician will also influence the choice of testing strategies. Because PET/CT is not yet widely available in China, with a difficulty in accessing it for many patients, CT alone strategy often is used for discrimination between malignant and benign SPN. Because of a high prevalence of lung cancer among SPN discovered in hospital in China, especially with a pretest probability up to 64%, in terms of ICAR in Figure 5 CT alone strategy is cost-effective. Between a prevalence of 80%-100%, CT plus ACNB is also cost-effective (Figure 5), but is often not used because of a rather low accuracy (Figure 4).

There were some limits to our study. First, SPN which are diagnosed as lung cancer were assumed to require surgery regardless of cancer staging. However, many factors interact to determine the needs of further examinations of the lung lesions and its resectability, including not only tumor staging, but also performance status and cardiopulmonary status. These factors were not traced in the models. Secondly, because of the lack of large sample study in China, the diagnostic performance (sensitivity and specificity) of each diagnostic test was gleaned from the literature which were published in the United States, Japan and European countries. We used those that were similar to results of a small sample study in China.

In conclusion, the introduction of CT plus DWI plus PET/CT strategy for the evaluation of SPN, which are discovered on chest radiography in the clinical setting, is potentially cost-effective in China with higher accuracy, over a large prevalence of cancer. When the prevalence of cancer rises up to 65%, the introduction of CT alone strategy for the evaluation of SPN is potentially cost-effective.

## COMMENTS

### *Background*

It has become a major concern about how to reduce the mortality rate of lung cancer, for it has become the leading cause of cancer deaths in China because the incidence and mortality are rapidly increasing. Although the authors have noted earlier detection of peripheral lung cancer, population-based screening using plain chest radiograph has been not carried out in China. Therefore, most solitary pulmonary nodules, which are usually discovered in the clinical setting in hospital in China, have a higher malignant prevalence.

### *Research frontiers*

Recent studies in other countries may not work for the Chinese hospital setting because of differences in health-care systems and diagnostic strategy. The cost

of surgery is not obviously higher than that of automated cutting needle biopsy (ACNB, six times) and positron emission tomography/computed tomography (PET/CT, three times) in China. But in the United States, Japan and European countries the cost of surgery was twenty times as much as that of biopsy and PET/CT. Because of expensive cost, PET/CT has not gained widespread popularity in most areas of China. A recent study showed that diffusion-weighted magnetic resonance imaging (DWI) may be able to be used in place of PET/CT for discriminating malignant from benign pulmonary nodules and is associated with significantly less false positives and lower cost than PET/CT.

### Innovations and breakthroughs

Up to now there has not been an analysis of cost-effectiveness of various strategies for the diagnosis of solitary pulmonary nodule (SPN) in China. Through the four strategies for the management of SPN: CT alone, CT plus CT-guided ACNB, CT plus PET/CT, CT plus DWI plus PET/CT, we analyzed the cost-effectiveness and chose CT plus DWI plus PET/CT strategy as an optimal option for the evaluation of SPN in China, because this strategy was not only cost-effective, but also had a higher accuracy accompanied by a lower missed diagnosis rate than CT plus ACNB strategy.

### Applications

According to this analysis, the CT plus DWI plus PET strategy was the optimal choice, which can be applied on chest radiograph in the clinical setting, having potential cost-effectiveness in China with higher accuracy, over a large prevalence of cancer.

### Terminology

SPN is the round or oval opaque areas of solitary pulmonary with the diameter < 3 cm performed on the X-line. The standards of the diagnosis of lung cancer by CT: there is a blood supply in the malignant tumor, while the benign tumor has lower or no blood supply; the malignant tumor has an irregular shape and sublobe with burrs on the edge.

### Peer review

In the manuscript Lu *et al* present the results of an analysis of cost-effectiveness of various strategies for the diagnosis of SPN in China. They have assessed the cost-effectiveness of four strategies for the management of SPN: CT alone, CT plus CT-guided ACNB, CT plus PET/CT, CT plus DWI plus PET/CT. The English of the text is of high quality, clear and easy to follow. The tables and the figures are clear and self explanatory. The statistical approach meets the current standards.

## REFERENCES


1 **Zhang L**, Wang M, Wang Y, Li L. [Clinico-pathological study of 98 patients with pulmonary solitary nodule]. *Zhonghua Zhong Liu Za Zhi* 2002; **24**: 491-493 [PMID: 12485507]

2 **Chen W**, Liu J, Chen Q, Li W, Xiong Z, Long X. [Bayes analysis in clinical decision-making for solitary pulmonary nodules]. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* 2009; **34**: 401-405 [PMID: 19483287 DOI: 10.3736/jcim20090501]

3 **Tsushima Y**, Endo K. Analysis models to assess cost effectiveness of the four strategies for the work-up of solitary pulmonary nodules. *Med Sci Monit* 2004; **10**: MT65-MT72 [PMID: 15114278]

4 **Keith CJ**, Miles KA, Griffiths MR, Wong D, Pitman AG, Hicks RJ. Solitary pulmonary nodules: accuracy and cost-effectiveness of sodium iodide FDG-PET using Australian data. *Eur J Nucl Med Mol Imaging* 2002; **29**: 1016-1023 [PMID: 12173015 DOI: 10.1007/s00259-002-0833-2]

5 **Gambhir SS**, Shepherd JE, Shah BD, Hart E, Hoh CK, Valk PE, Emi T, Phelps ME. Analytical decision model for the cost-effective management of solitary pulmonary nodules. *J Clin Oncol* 1998; **16**: 2113-2125 [PMID: 9626211]

6 **Kosuda S**, Ichihara K, Watanabe M, Kobayashi H, Kusano S. Decision-tree sensitivity analysis for cost-effectiveness of chest 2-fluoro-2-D-[(18)F]fluorodeoxyglucose positron emission tomography in patients with pulmonary nodules (non-small cell lung carcinoma) in Japan. *Chest* 2000; **117**: 346-353 [PMID: 10669673 DOI: 10.1378/chest.117.2.346]

7 **Gould MK**, Sanders GD, Barnett PG, Rydzak CE, Maclean CC, McClellan MB, Owens DK. Cost-effectiveness of alternative management strategies for patients with solitary pulmonary nodules. *Ann Intern Med* 2003; **138**: 724-735 [PMID: 12729427]

8 **Mori T**, Nomori H, Ikeda K, Kawanaka K, Shiraishi S, Katahira K, Yamashita Y. Diffusion-weighted magnetic resonance imaging for diagnosing malignant pulmonary nodules/masses: comparison with positron emission tomography. *J Thorac Oncol* 2008; **3**: 358-364 [PMID: 18379353 DOI: 10.1097/JTO.0b013e318168d9ed]

9 **Dietlein M**, Weber K, Gandjour A, Moka D, Theissen P, Lauterbach KW, Schicha H. Cost-effectiveness of FDG-PET for the management of solitary pulmonary nodules: a decision analysis based on cost reimbursement in Germany. *Eur J Nucl Med* 2000; **27**: 1441-1456 [PMID: 11083532 DOI: 10.1007/s002590000324]

10 **Tsubamoto M**, Johkoh T, Kozuka T, Honda O, Koyama M, Murai S, Inoue A, Sumikawa H, Tomiyama N, Hamada S, Yamamoto S, Nakamura H, Kudo M. Coronal multiplanar reconstruction view from whole lung thin-section CT by multidetector-row CT: determination of malignant or benign lesions and differential diagnosis in 68 cases of solitary pulmonary nodule. *Radiat Med* 2003; **21**: 267-271 [PMID: 14743900]

11 **Siegelman SS**, Khouri NF, Leo FP, Fishman EK, Braverman RM, Zerhouni EA. Solitary pulmonary nodules: CT assessment. *Radiology* 1986; **160**: 307-312 [PMID: 3726105 DOI: 10.1148/radiology.160.2.3726105]

12 **Yi CA**, Lee KS, Kim BT, Choi JY, Kwon OJ, Kim H, Shim YM, Chung MJ. Tissue characterization of solitary pulmonary nodule: comparative study between helical dynamic CT and integrated PET/CT. *J Nucl Med* 2006; **47**: 443-450 [PMID: 16513614]

13 **Kim SK**, Allen-Auerbach M, Goldin J, Fueger BJ, Dahlbom M, Brown M, Czernin J, Schiepers C. Accuracy of PET/CT in characterization of solitary pulmonary lesions. *J Nucl Med* 2007; **48**: 214-220 [PMID: 17268017]

14 **Satoh S**, Ohdama S, Matsubara O, Okochi Y, Tanaka R, Kimula Y. CT-guided automated cutting needle biopsy by a combined method for accurate specific diagnosis of focal lung lesions. *Radiat Med* 2005; **23**: 30-36 [PMID: 15786749]

15 **Tsukada H**, Satou T, Iwashima A, Souma T. Diagnostic accuracy of CT-guided automated needle biopsy of lung nodules. *AJR Am J Roentgenol* 2000; **175**: 239-243 [PMID: 10882279 DOI: 10.2214/ajr.175.1.1750239]

16 **Takashima S**, Sone S, Li F, Maruyama Y, Hasegawa M, Kadoya M. Indeterminate solitary pulmonary nodules revealed at population-based CT screening of the lung: using first follow-up diagnostic CT to differentiate benign and malignant lesions. *AJR Am J Roentgenol* 2003; **180**: 1255-1263 [PMID: 12704034 DOI: 10.2214/ajr.180.5.1801255]

17 **Klein JS**, Salomon G, Stewart EA. Transthoracic needle biopsy with a coaxially placed 20-gauge automated cutting needle: results in 122 patients. *Radiology* 1996; **198**: 715-720 [PMID: 8628859 DOI: 10.1148/radiology.198.3.8628859]

18 **Li H**, Boiselle PM, Shepard JO, Trotman-Dickenson B, McLoud TC. Diagnostic accuracy and safety of CT-guided percutaneous needle aspiration biopsy of the lung: comparison of small and large pulmonary nodules. *AJR Am J Roentgenol* 1996; **167**: 105-109 [PMID: 8659351 DOI: 10.2214/ajr.167.1.8659351]


**P- Reviewer:** Komócsi A **S- Editor:** Song XX
**L- Editor:** O'Neill M **E- Editor:** Lu YJ

**ISSN** 2308-3840 (online)

# World Journal of *Meta-Analysis*

*World J Meta-Anal* 2014 November 26; 2(4): 135-227



Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

## Contents

Quarterly Volume 2 Number 4 November 26, 2014

**REVIEW**

135 Anti-obesity drugs currently used and new compounds in clinical development
*Martins A, Morgado S, Morgado M*

**EVIDENCE-BASED MEDICINE**

154 Russian clinical research policy does not guarantee results availability
*Shamliyan TA, Avanesova AA*

**META-ANALYSIS**

162 Diagnostic yield of third eye retroscope on adenoma detection during colonoscopy: A systematic review and meta-analysis
*Thosani N, Rao B, Batra S, Adeyefa B, Raju GS, Bresalier RS, Guha S, Banerjee S*

171 Hydrocephalus after subarachnoid hemorrhage: A meta-analytic comparison of aneurysm treatments
*Lang SS, Sanborn MR, Ju C, Premjee A, Stein SC, Smith MJ*

179 Meta-analysis of bivariate *P* values
*Kocak M*

186 Observed communication between oncologists and patients: A causal model of communication competence
*Turkiewicz KL, Allen M, Venetis MK, Robinson JD*

194 Systematic review and meta-analysis of Statins-Fibrates therapy in diabetic dyslipidemia patients
*Zheng S, Li YX, Han TT, Zhang Y, Jiang DD, Hu YM*

204 Alcohol consumption and the risk of Barrett's esophagus: A meta-analysis
*Dai C, Liu WX, Wang K, Jiang HK, Jiang M, Sun MJ*

212 Using fluorodeoxy-D-glucose-positron emission tomography to monitor neoadjuvant chemotherapy response in sarcoma: A meta-analysis
*Wang YT, Pu H, Yin LL, Chen JY*

221 Female-specific *vs* conventional knee prostheses after total knee arthroplasty: A meta-analysis
*Rong GX, Huang L, Gui BJ, Xu AM, Zhang JL, Wang SS*

*World Journal of Meta-Analysis*
**Volume 2 Number 4 November 26, 2014**

## Contents


| | | |
|---|---|---|
| **APPENDIX** | I-V | Instructions to authors |
| **ABOUT COVER** | | Editorial Board Member of *World Journal of Meta-Analysis*, Yu-Ting Wang, MD, Doctor, Department of Radiology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu 610072, Sichuan Province, China |
| **AIM AND SCOPE** | | *World Journal of Meta-Analysis* (*World J Meta-Anal*, *WJMA*, online ISSN 2308-3840, DOI: 10.13105) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians, with a specific focus on meta-analysis, systematic review, mixed-treatment comparison, meta-regression, overview of reviews.<br>*WJMA* covers a variety of clinical medical fields including allergy, anesthesiology, cardiac medicine, clinical genetics, clinical neurology, critical care, dentistry, dermatology, emergency medicine, endocrinology, family medicine, gastroenterology and hepatology, geriatrics and gerontology, hematology, immunology, infectious diseases, internal medicine, obstetrics and gynecology, oncology, ophthalmology, orthopedics, otolaryngology, pathology, pediatrics, peripheral vascular disease, psychiatry, radiology, rehabilitation, respiratory medicine, rheumatology, surgery, toxicology, transplantation, and urology and nephrology, while maintaining its unique dedication to systematic reviews and meta-analyses. |
| **INDEXING/ABSTRACTING** | | *World Journal of Meta-Analysis* is now indexed in Digital Object Identifier. |
| **FLYLEAF** | I-IV | Editorial Board |


**EDITORS FOR THIS ISSUE**

**Responsible Assistant Editor:** *Xiang Li*
**Responsible Science Editor:** *Fang-Fang Ji*
**Responsible Electronic Editor:** *Huan-Liang Wu*
**Proofing Editorial Office Director:** *Jin-Lei Wang*
**Proofing Editor-in-Chief:** *Lian-Sheng Ma*

NAME OF JOURNAL
*World Journal of Meta-Analysis*

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Quarterly

EDITORS-IN-CHIEF
**Giuseppe Biondi-Zoccai, MD, Assistant Professor,** Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
*World Journal of Meta-Analysis*
Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
November 26, 2014

COPYRIGHT
© 2014 Baishideng Publishing Group Inc. Articles published by this Open-Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

SPECIAL STATEMENT
All articles published in journals owned by the Baishideng Publishing Group (BPG) represent the views and opinions of their authors, and not the views, opinions or policies of the BPG, except where otherwise explicitly indicated.

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.135

*World J Meta-Anal* 2014 November 26; 2(4): 135-153
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

REVIEW

# Anti-obesity drugs currently used and new compounds in clinical development

Armindo Martins, Sandra Morgado, Manuel Morgado

Armindo Martins, Manuel Morgado, CICS-UBI-Health Sciences Research Center, University of Beira Interior, 6201-001 Covilhã, Portugal
Sandra Morgado, Manuel Morgado, Hospital Centre of Cova da Beira, 6200-251 Covilhã, Portugal
Author contributions: Martins A analysis and interpretation of data, drafting of manuscript, statistical analysis and technical support; Morgado S acquisition of data, drafting of manuscript and critical revision of manuscript for important intellectual content; Morgado M concept and study design, analysis and interpretation of data, drafting of manuscript and critical revision of manuscript for important intellectual content.
Correspondence to: Manuel Morgado, MD, PhD, Hospital Centre of Cova da Beira, E.P.E., Quinta do Alvito, 6200-251 Covilhã, Portugal. manuelaugustomorgado@gmail.com
Telephone: +351-275-329002 Fax: +351-275-329099
Received: February 22, 2014 Revised: June 30, 2014
Accepted: August 27, 2014
Published online: November 26, 2014

## Abstract

Obesity is a chronic disease which requires treatment. As lifestyle interventions alone hardly ever result in long-term weight loss, pharmacotherapy is an important adjunct to lifestyle measures to improve the induction and maintenance of weight loss. Owing to the limited options currently available for the pharmacological treatment of obesity, it is imperative to develop new safe compounds. This study aims to review the current medications approved by European Medicines Agency and United States Food and Drug Administration (FDA) for the treatment of obesity, focusing essentially on their benefits and risks, as well as on the new drugs which are presently under clinical trials. Moreover, it lists the anti-obesity agents that have been recently withdrawn from the market. A revision of the scientific literature was carried out, through a search on Pubmed for papers published from January 2010 to January 2013. Orlistat (Xenical®) is currently the only long-term pharmacotherapy for obesity available in the European market, as rimonabant and sibutramine were withdrawn in 2008 and 2010, respectively, due to serious psychiatric and cardiovascular adverse effects. Lorcaserin (Belviq®) and the association of phentermine and topiramate (Qsymia™) were recently approved by FDA. Orlistat suppresses appetite inhibiting gastrointestinal lipase, being its adverse effects mostly gastrointestinal. Lorcaserin activates 5-HT2C receptors, phentermine is a norepinephrine releasing drug, and topiramate is an anticonvulsivant drug with weight loss properties.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Obesity; Weight loss; Pharmacological treatment; European Medicines Agency; Food and Drug Administration; Clinical trials

**Core tip:** Obesity is a chronic disease which requires treatment. As lifestyle interventions alone hardly ever result in long-term weight loss, pharmacotherapy is an important adjunct to lifestyle measures to improve the induction and maintenance of weight loss. This study reviews the current medications approved by European Medicines Agency and United States Food and Drug Administration for the treatment of obesity, focusing essentially on their benefits and risks, as well on the new drugs under clinical trials. An original algorithm for assessment and treatment of overweight and obese patients in clinical practice was also developed.

Martins A, Morgado S, Morgado M. Anti-obesity drugs currently used and new compounds in clinical development. *World J Meta-Anal* 2014; 2(4): 135-153 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/135.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.135

## INTRODUCTION

Obesity is a global epidemic, which is increasing not only

in Western cultures, but also in Mexico, Brazil, China and parts of Africa, due to increasing urbanization[1,2], and underlying sedentary lifestyles and overconsumption of calories. According to the most recent estimates by the World Health Organization, more than 1 billion people in the world are overweight or obese, accounting for up to 7% of the health care cost in most developed countries[3]. It is estimated that obesity affects 10%-30% of adults in the European Union (EU) countries and 46.3% in the United States[4]. This chronic disease increases morbidity and mortality, largely due to associated comorbidities, including type 2 diabetes mellitus (T2DM), cardiovascular diseases, premature atherosclerosis, metabolic syndrome, liver disease and cancer[5-9]. About 90% of cases of T2DM are due to excess weight, so professional societies recommend weight loss as a key initial step in the management of overweight or obese T2DM patients[10]. Furthermore, there is a five-to-six times increase in hypertension among obese individuals compared to those with normal weight[11,12].

The relation between obesity and several of its comorbidities appears to be mediated by risk factors characterized by the metabolic syndrome, which is closely associated with abdominal obesity[13-15]. Weight loss achieved by changes in eating behavior, diet and physical activity are the cornerstones in the treatment of the metabolic syndrome. However, weight control methods often produce short-term success, while sustained weight maintenance is difficult to achieve[16,17]. Therefore, pharmacological therapy has been proposed as an adjunct to diet and lifestyle changes to improve maintenance of weight loss[18]. Drug therapy is considered for individuals with a body mass index (BMI) greater than 30 kg/m$^2$ or ranging from 25 to 30 kg/m$^2$ if they have concomitant obesity-related risk factors or diseases[19].

Improving diet (*i.e.*, reduction in caloric intake) and increasing physical activity can address obesity. Nevertheless, adherence to lifestyle changes can be challenging for many reasons, such as a lack of will for change on the part of the patient, physical restrictions that limit activity and the need for a multidisciplinary health care team for effective treatment. In severe cases of obesity, with possibly evident comorbidities, pharmacological treatment or bariatric surgery are considered in addition to lifestyle changes[20].

In contrast to orlistat, a gastric and pancreatic lipase inhibitor that inhibits the absorption of fat from the intestines, all of the other mentioned drugs act on the central nervous system (CNS) to suppress appetite, thus reducing food intake, and/or stimulate energy expenditure, thereby reducing body weight. Orlistat is the first prescription treatment for obesity that does not act as an appetite suppressant, incapacitating lipase to break triglycerides down into fatty acids and monoglycerides. It also improves cardiovascular risk factors, including serum triglycerides levels[21-23].

There are currently several drugs towards the pharmacological treatment of obesity under clinical trials. An anti-obesity drug that is already on an advanced stage of clinical trials is Contrave®, a sustained-release formulation of bupropion and naltrexone, which significantly reduced body weight in a phase 3 trial[24]. The FDA has required an additional clinical trial prior to the approval of the drug, because of concern over the cardiovascular safety profile[25].

## SHORT-TERM TREATMENT OF OBESITY

### *Phentermine*

Phentermine hydrochloride (Adipex-P®) is a sympathomimetic anorectic drug of the phenethylamine class, with pharmacology similar to amphetamine. It is thought to act on the hypothalamus stimulating the adrenal glands to release norepinephrine, thus reducing appetite[26], although it has not been established that the primary action of amphetamine anorectics in treating obesity is appetite suppression, since other CNS actions or metabolic effects may also be involved[27]. It is approved by FDA since 1959 for short-term ($\leqslant$ 12 wk) treatment of obesity as monotherapy (15-37.5 mg/d) in a regimen of weight reduction based on exercise, behavioral modification and caloric restriction, in the management of exogenous obesity in patients with a BMI $\geqslant$ 30 kg/m$^2$ or $\geqslant$ 27 kg/m$^2$ in the presence of other risk factors (*e.g.,* hypertension, diabetes, dyslipidemia)[28]. In 2005, it was still the most prescribed anti-obesity drug in the United States[29]. However, it was definitely withdrawn from the EU market in May 2001, after a recommendation by European Medicines Agency (EMA) due to an unfavourable risk to benefits ratio, including tolerance and possible abuse potential[30]. Because phentermine has sympathomimetic properties, some reported adverse effects were insomnia, tremor, elevation in blood pressure and pulse rate, headache, palpitation and constipation[31].

### *Diethylpropion*

Diethylpropion hydrochloride (Tenuate®), also known as amfepramone, is another amphetamine-like analogue that was also approved by FDA in 1959 as a short-term ($\leqslant$ 12 wk) adjunct in the treatment of exogenous obesity as monotherapy only (one 25 mg tablet three times daily or one 75 mg controlled-release tablet once a day), in a regimen of weight reduction based on caloric restriction in patients with a BMI $\geqslant$ 30 kg/m$^2$[32,33]. Diethylpropion, as an amphetamine-like analogue, stimulates neurons to release or maintain high levels of catecholamines, namely dopamine and norepinephrine. High levels of these catecholamines tend to suppress hunger signals and appetite. Diethylpropion, through catecholamine elevation, may also indirectly affect leptin levels in the brain. It is theorized that diethylpropion can raise levels of leptin, which signal satiety. It is also theorized that increased levels of the catecholamines are partially responsible for halting neuropeptide Y. This peptide initiates eating, decreases energy expenditure, and increases fat storage[34]. The most common side effects are dry mouth and insomnia[35]. In 2001, EMA also recommended its withdrawal from the

EU market, for the same reasons as phentermine[30].

## LONG-TERM TREATMENT OF OBESITY

### *Intestinal lipase inhibitors*

**Orlistat:** Orlistat was until recently the only pharmacological weight loss drug on the international market. Xenical® (orlistat 120 mg) was approved as a prescription drug by FDA in 1999 for obesity management and to reduce the risk of regaining weight after prior weight loss. In 2007, Alli (orlistat 60 mg) was approved for over-the-counter use for weight loss in overweight adults (≥ 18 years), along with a reduced-calorie and low-fat diet[36]. Orlistat reversibly blocks the gastrointestinal lipase activity and thus the absorption of about 25% of ingested fat. The 60 mg dose achieves a weight loss of about 5% of total body weight over a 16 wk period, improving insulin resistance and lipid profile[37]. Weight loss with orlistat is about 3% (placebo-subtracted), but meets FDA categorical standard (FDA categorical weight loss is the proportion of subjects who lose at least 5% of baseline body weight in the active-product *vs* placebo-treated group.)[38]. It is indicated for obese individuals aged 18 or over with an initial BMI ≥ 30 kg/m$^2$ or ≥ 27 kg/m$^2$ in the presence of other risk factors. The safety and efficacy of Xenical® were evaluated in obese adolescent patients aged 12 to 16 years old. It was concluded that the use of Xenical® in this age group is supported by evidence from adequate and well-controlled studies of Xenical® in adults, from a 54-wk efficacy and safety study and a 21-d mineral balance study in obese adolescent patients aged 12 to 16 years. However, unlike FDA, EMA did not approve the use of orlistat in patients below 18 years of age. With respect to pediatric patients below the age of 12, safety and effectiveness have not been established yet and its use is not approved in this population[39]. The Royal Pharmaceutical Society recommends that BMI of individuals requesting orlistat is always checked and that patients should be advised to take an additional multivitamin supplement, to counteract the loss of fat-soluble vitamins[40]. As it is a lipases inhibitor, orlistat has several potential adverse effects patients need to be aware of, most of them involving the gastrointestinal system.

When combined with a reduced-calorie diet, orlistat has proved to be more effective than diet alone[41]. However, another study, carried out in order to compare a low-carbohydrate, ketogenic diet with orlistat therapy combined with a low fat diet over 48 wk, showed a statistically significant mean weight loss, although similar in the two groups (*i.e.*, -9.5% *vs* -8.5%)[42]. This intervention also had similar beneficial effects on most measures of cardiovascular disease risk, including waist circumference (which decreased as well), fasting serum lipid profiles and C-reactive protein. Weight loss was significantly greater in participants who attended group sessions regularly, which may indicate the usefulness of these sessions in motivating patients. Efforts should hence be made to incorporate similarly intensive weight loss programs into medical practice[42].

Another study was conducted to compare the effects of one year treatment with orlistat plus L-carnitine and orlistat alone on body weight, glycemic and lipid control and insulin resistance in uncontrolled T2DM[43]. This study led to the conclusion that the association of orlistat plus L-carnitine was more effective than orlistat monotherapy in improving body weight, glycemic and lipid profile, insulin resistance and inflammatory parameters. A weight loss was observed from baseline of 9.5% in the orlistat group and of 11.3% in the orlistat + L-carnitine group. No significant adverse effects were reported. These positive effects can be due to synergism between these two drugs[43].

A Portuguese study, carried out by de Castro *et al*[44], compared the efficacy and safety of orlistat *vs* placebo in obese patients with mild to moderate hypercholesterolemia, concluding that treatment with orlistat, plus a reduced calorie diet for 6 mo, achieved significant reductions in weight, BMI and in total and low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol changes being similar in both treatment groups. The mean difference in weight from baseline was 5.9% in the orlistat group and 2.3% in the placebo group. By the end of the study, 49.1% of patients in the orlistat group had lost 5%-10% of their baseline weight, compared to 22.4% of the placebo group and 8.8% had lost over 10% (4.7% more than the placebo group). The frequency of gastrointestinal adverse effects was slightly higher in the orlistat than in the placebo group[44].

**Cetilistat:** Cetilistat is a new, highly lipophilic benzoxazinone inhibitor of gastrointestinal and pancreatic lipases, which is currently under development for the management of weight loss in obese patients. In a 12-wk, Phase 2 clinical trial in obese patients without medicated comorbidities, administration of cetilistat three times daily combined with a hypocaloric diet produced significantly greater weight loss than placebo[45]. Furthermore, it was well tolerated, the adverse effects being predominantly gastrointestinal and mild (*i.e.*, aware of signs or symptoms, easily tolerated) or moderate (*i.e.*, sign or symptom causes discomfort but does not interfere with normal activities) in severity[45].

A randomized, double-blind study was carried out to determine the efficacy and safety of cetilistat and orlistat relative to placebo in obese patients with T2DM, treated with metformin[46]. Twelve weeks of treatment with cetilistat 80 or 120 mg three times daily significantly reduced body weight and improved glycemic control in obese T2DM patients. Absolute weight loss was significantly greater in the cetilistat 80 and 120 mg dose groups than in the placebo group ($P$ = 0.01 and $P$ = 0.0002, respectively), but in the group treated with cetilistat 40 mg weight loss was similar to placebo. Cetilistat was well tolerated, having withdrawals due to adverse effects been similar to placebo and markedly less frequent than with orlistat. Discontinuation in the orlistat group was entirely due to gastrointestinal adverse effects. The high level of tolerability to cetilistat led to an increase in compliance

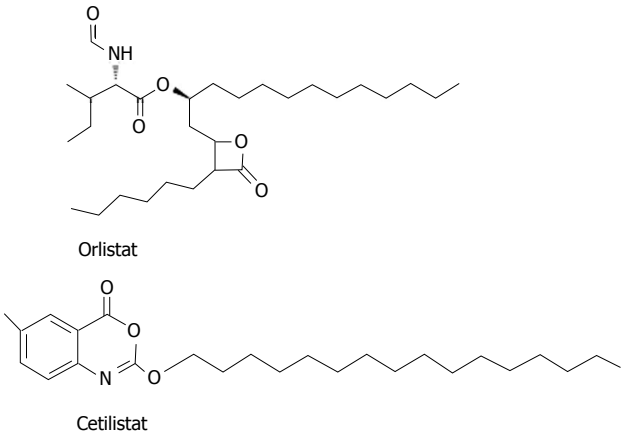



**Figure 1 Structures of orlistat and cetilistat.**

and could be clinically significant in the management of this patient population. The difference in adverse effects between cetilistat and orlistat could be attributable to structural differences between them (Figure 1). Although both molecules are lipase inhibitors, the chemical structural differences concerning hydrophilic and lipophilic components may influence the way they interact with fat micelles in the intestine. Orlistat, unlike cetilistat, may hence promote coalescence of micelles, leading to oils and increased gastrointestinal adverse effects[46].

A Bryson *et al*[47] article, published in 2009, described three phase 1 studies, which demonstrated that cetilistat effectively inhibits gastrointestinal lipases, substantially increasing the amount of faecal fat excreted in healthy male volunteers receiving a controlled calorie diet. It also stated that faecal fat excretion in the cetilistat groups was at least comparable to the orlistat 120 mg three times daily group[47].

In December 2008, phase 2 clinical trials data showing that cetilistat 80 mg and 120 mg promoted significant weight loss compared with placebo (3.85 kg and 4.32 kg *vs* 2.86 kg, respectively) and was well tolerated in clinically obese patients encouraged Takeda Pharmaceutical Company Limited to start a phase 3 clinical trial in Japan[48].

### *Agonists of the serotonin 5-HT2C receptors*

**Lorcaserin:** Lorcaserin is a potent and selective serotonin 2C receptor agonist, which has been recently (June 2012) approved by FDA for chronic weight management as Belviq® (lorcaserin hydrochloride), in addition to a hypocaloric diet and exercise. The drug is approved for use in obese (BMI $\geqslant$ 30 kg/m$^2$) or overweight (BMI $\geqslant$ 27 kg/m$^2$) adults who have at least one weight related condition, such as hypertension, T2DM or dyslipidemia[49].

Lorcaserin activates 5-hidroxytryptamine (5-HT or serotonin) receptors 5-HT2C, suppressing appetite through the pro-opiomelanocortin (POMC) neurons. This anti-obesity drug reduces body weight both in rodents and humans, although it is not clear whether weight loss is a result of reduced energy intake only or also of enhanced energy expenditure. When gauged by the standards of the Division's 2007 draft guidance for Developing Products for Weight Management, the mean weight loss associated with the lorcaserin 10 mg once daily and twice daily dose was 3%-4% greater than the mean weight loss with placebo. Therefore lorcaserin did not satisfy the guidance's mean efficacy criterion. However, the lorcaserin 10 mg twice a day dose did, by a slim margin, satisfy the categorical efficacy criterion[49,50].

Martin *et al*[51] investigated the effect of lorcaserin on energy intake and energy expenditure, having concluded, after 7 d of weight maintenance (consisting of a diet and exercise plan), that lorcaserin significantly reduced energy intake and appetite. However, it did not alter either energy expenditure or respiratory quotient (RQ) (The ratio of the volume of carbon dioxide produced to the volume of oxygen consumed per unit of time by the body; usually corresponding to the volumes given off and taken up by the lungs. It varies with the source of food energy used.). In combination with a healthy lifestyle program, lorcaserin treatment (10 mg twice daily) resulted in significantly larger weight loss than placebo (lorcaserin, -3.8 $\pm$ 0.4 kg; placebo, -2.2 $\pm$ 0.5 kg; $P < 0.01$) over 56 d[51]. Furthermore, treatment with lorcaserin improved the risk factors of cardiovascular disease, namely total and LDL cholesterol, which were significantly reduced, as well as diastolic blood pressure[51]. Concerning adverse effects, lorcaserin was generally well tolerated, headache being the most common adverse effect[51].

Another 2-year trial, carried out by Smith *et al*[52], and known as Behavioral Modification and Lorcaserin for Overweight and Obesity Management trial, led to the conclusion that lorcaserin, along with behavioral modification (consisting of a nutrition and physical exercise program), produced significant mean weight loss and improved maintenance of weight loss, when compared with placebo (5.8% *vs* 2.2%, $P < 0.001$). It was also observed that, at the end of year 1, 47.5% of patients who received lorcaserin lost, at least, 5% of their initial body weight, compared to 20.3% of patients who received placebo (*i.e.*, the proportion of patients in the lorcaserin group was more than twice that in the placebo group) ($P < 0.001$). The average weight loss achieved in the lorcaserin group after 1 year was associated with an improvement in plasma lipid levels, insulin resistance, and blood pressure. Furthermore, lorcaserin produced a reduction in waist circumference and in the levels of markers of inflammation, such as high-sensitivity C-reactive protein (hs-CRP), thus reducing the risk of future cardiovascular events. Total and LDL cholesterol, as well as triglyceride levels, at the end of year 1 were considerably lower in the lorcaserin group than in the placebo group, but it was increased at the end of year 2 in both groups. Concerning adverse effects, more patients in the lorcaserin group than in the placebo group withdrew from the study because of headache (2.0% *vs* 0.8%) and dizziness (0.8% *vs* 0.1%). At the end of year 1, 2.3% of patients in the placebo group and 2.7% of patients in the lorcaserin group developed serotonin-associated valvulopathy ($P = 0.70$) (relative risk with lorcaserin, 1.1; 95%CI: 0.69-1.85). On the other hand, in year 2, the rate of heart valvulopathy was 2.7%

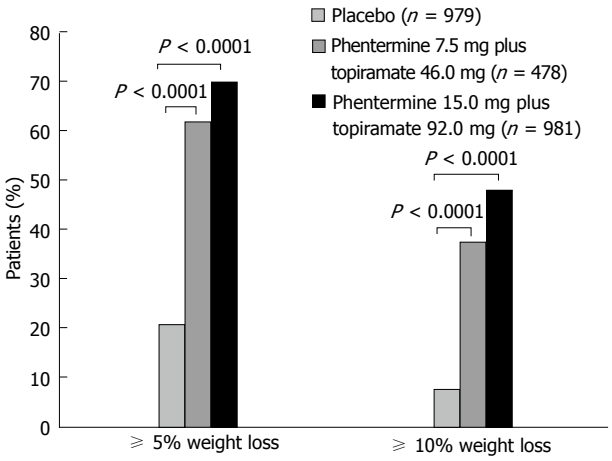



**Figure 2 Patients with at least 5% and at least 10% weight loss in the CONQUER study.** From[59], with permission from Elsevier.

in the placebo group and 2.6% in patients who received lorcaserin during year 1 and year 2[52].

Belviq® should not be used during pregnancy. Treatment may lead to serious adverse effects, including serotonin syndrome, particularly when administered with drugs commonly used to treat depression and migraine, and disturbances in attention or memory. However, the most common adverse effects of Belviq® are headache, dizziness, fatigue, nausea, dry mouth and constipation in non-diabetic patients; and hypoglycemia, headache, back pain, cough and fatigue in diabetic patients. Belviq® should be used with caution in patients with congestive heart failure, due to a possible increase in the number of 5-HT$_{2B}$ receptors in this pathology[49].

### *Combination therapies*

**Phentermine/topiramate:** Topiramate, used as anticonvulsivant and for prevention of migraine headaches, has shown weight loss properties but is not currently approved as a sole therapy for obesity[53]. The mechanism of action for topiramate is unknown, although it is known that it blocks repetitive firing by acting on sodium channels, may enhance GABA$_A$-mediated chloride ions flux, and seems to be an antagonist of the α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid and kainate receptors, blocking the effect of glutamate (an excitatory neurotransmitter)[54].

Qsymia™, an extended-release formulation of phentermine hydrochloride and topiramate, has been approved by FDA, in July 2012, as weight loss therapy, in combination with a reduced-calorie diet and exercise. The drug is approved for use in obese (BMI ⩾ 30 kg/m$^2$) or overweight (BMI ⩾ 27 kg/m$^2$) adults, who have at least one weight-related comorbidity, such as hypertension, T2DM or dyslipidemia. Qsymia™ is available in four different extended-release formulations: 3.75 mg of phentermine and 23 mg of topiramate; 7.5 mg of phentermine and 46 mg of topiramate; 11.25 mg of phentermine and 69 mg of topiramate; and a higher dose containing 15 mg of phentermine and 92 mg of topiramate for specific patients. Qsymia™ should be taken once a day in the morning, with or without food. The dose of Qsymia™ should be increased if the patient does not lose a certain amount of weight within the first 12 wk of treatment at the recommended dose. If, after an additional 12 wk period of treatment on a higher dose, the patient has not lost a certain amount of weight, treatment should be discontinued[55]. With respect to pharmacological interactions, Qsymia™ should not be used with: birth control (contraceptive) medications, diuretics (*e.g.*, hydrochlorotiazide), carbonic anhydrase inhibitors, anticonvulsivant medicines (*e.g.,* valproic acid) and medicines that impair or decrease thinking, concentration or muscle coordination[55].

Qsymia™ must not be used in patients with glaucoma or hyperthyroidism and can increase heart rate 0.6 to 1.6 beats per minute (bpm)[56]. Its use in patients with recent or unstable heart disease or stroke is, therefore, not recommended. Heart rate should be regularly monitored in these patients, especially if they are starting or increasing the dose. The FDA approved Qsymia™ with a Risk Evaluation and Mitigation Strategy, the aim of which is to educate prescribers and their patients about the increased risk of birth defects associated with first 3 mo of exposure to Qsymia™, the need for pregnancy prevention and the need to discontinue therapy in pregnancy[57]. The most common adverse effects of Qsymia™ are paresthesia of hands and feet, dizziness, altered taste sensation, insomnia, constipation and dry mouth[57].

Allison *et al*[58] conducted a 56-wk randomized controlled trial to evaluate the safety and efficacy of a controlled-release formulation of phentermine and topiramate (PHEN/TPM CR) for weight loss and metabolic improvements. In this study, men and women with class II and III obesity (BMI ⩾ 35 kg/m$^2$) were randomized to placebo, PHEN/TPM CR 3.75/23 mg or PHEN/TPM CR 15/92 mg, in addition to a low-energy diet. PHEN/TPM CR demonstrated dose-related beneficial effects on weight and metabolic variables, having not been reported serious adverse effects resulting from treatment. Concerning mean weight loss, patients in the placebo, PHEN/TPM 3.75/23, and PHEN/TPM 15/92 groups lost 1.6%, 5.1%, and 10.9% of baseline body weight (BW), respectively, at 56 wk ($P < 0.0001$). In categorical analysis, 17.3% of placebo patients, 44.9% of PHEN/TPM 3.75/23 patients, and 66.7% of PHEN/TPM 15/92 patients, lost at least 5% of baseline BW at 56 wk ($P < 0.0001$)[58].

The same authors developed a 56-wk, randomized, placebo-controlled, Phase 3 trial, known as the CONQUER study[59], in order to evaluate the efficacy and safety of two doses of Phentermine/Topiramate CR, in combination with diet and lifestyle modification in overweight and obese adults (Figure 2). These authors demonstrated that both doses of the controlled formulation of phentermine and topiramate, along with lifestyle interventions, are useful for the treatment of obesity, both in reducing weight and improving cardiometabolic variables (blood pressure, waist circumference, concentrations of lipids, glycemia, adiponectin)[59]. It was obtained a mean weight loss (placebo-subtracted) of 6.6% and 8.6% with

PHEN/TPM 7.5/46 mg and PHEN/TPM 15/92 mg, respectively, and a categorical weight loss of 62% and 70% with PHEN/TPM 7.5/46 mg and PHEN/TPM 15/92 mg, respectively. These results satisfy the FDA mean and categorical efficacy draft guidance criteria[38]. Furthermore, improvements in risk factors were most pronounced in patients with pre-existing comorbidities. Psychiatric and cognitive adverse dose-dependent effects occurred mainly during the early phase of treatment, and were resolved on drug discontinuation. It is important to mention that patients with clinically significant depression were excluded from this study[59].

The SEQUEL study, a double-blind placebo controlled, 52-wk extension of the CONQUER study[53], was designed to evaluate the longer-term efficacy and safety of lifestyle modification and two doses of PHEN/TPM CR for an additional 52-wk period (in patients who had completed the CONQUER study). Mean weight losses from baseline were -1.8%, -9.3% and -10.5% for placebo, PHEN/TPM 7.5/46 mg and PHEN/TPM 15/92 mg, respectively. A weight loss of 10% was achieved by more than 50% of subjects treated with PHEN/TPM CR. Phentermine/Topiramate CR improved cardiovascular and metabolic variables (*e.g.*, hyperglycemia, dyslipidemia, hypertension), despite reduced use of concomitant medications to control blood pressure, lipid variables, and glycemic variables and decreased the rate of progression to T2DM. The greatest benefic effects were seen in patients receiving PHEN/TPM CR 15 mg/92 mg[53]. However, a mean increase was reported in heart rate of 0.4 bpm in placebo subjects, 1.3 bpm in PHEN/TPM CR 7.5/46 subjects, and 1.7 bpm in PHEN/TPM CR 15/92 subjects. Despite this, PHEN/TPM CR was generally well tolerated, the most commonly reported adverse effects being upper respiratory infection, constipation, paraesthesia, sinusitis and dry mouth. During the extension study, discontinuation rates due to adverse effects were similar between placebo (3.1%) and patients treated with PHEN/TPM CR 7.5/46 (4.5%) and with PHEN/TPM CR 15/92 (4.4%). Therefore, PHEN/TPM CR constitutes a well-tolerated and effective option for the sustained treatment of obese individuals with cardiometabolic disease[53].

**Naltrexone/bupropion:** Naltrexone is a non-selective μ-opioid receptor antagonist, but is considered rather preferential for μ receptors. It is a synthetic analogue of oxymorphone and naloxone, a white crystalline compound soluble in water. Bupropion is an aminoketone antidepressant, inhibiting dopamine reuptake. Its structure is closely related to that of diethylpropion, an anorexigenic and sympathomimetic agent previously mentioned. Bupropion is highly soluble in water and is related to the phenylethylamines, which are known for their stimulant effects[60].

Combined treatment with sustained-release (SR) naltrexone and bupropion (Contrave®) was developed to produce complementary actions in CNS pathways regulating body weight: on the one hand, by stimulating hypothalamic pro-opiomelanocortin (POMC) neurons with bupropion; on the other hand, by simultaneously blocking opioid-mediated POMC autoinhibition with naltrexone[24]. POMC producing neurons in the hypothalamus release α-melanocyte-stimulating hormone (α-MSH) and β-endorphin. α-MSH mediates the anorexigenic effect of POMC, whereas β-endorphin performs an autoinhibitory feedback through activation of opioid receptors on POMC neurons, thus inactivating the anorexigenic effect[61-63].

In February 2011, the FDA declined the approval of Contrave® due to concerns related to the long-term cardiovascular safety profile in overweight and obese individuals. The FDA advised the company responsible for the development of Contrave® (Orexigen Therapeutics Inc.) that they would have to conduct a randomized, double-blind, placebo-controlled trial in order to demonstrate that the risk of major cardiovascular adverse effects in overweight and obese patients receiving the drug does not negatively affect the risk-benefit profile[60,64]. One of the conducted trials was the Contrave Obesity Research-I (COR-I) study[24], a randomized, double-blind, placebo-controlled, phase 3, 56-wk trial, which examined the effect of such treatment on body weight in overweight and obese subjects from 34 locations in the United States. The participants were men and women aged between 18 and 65 years with: (1) a BMI ranging from 30 to 45 kg/m$^2$ and uncomplicated obesity; or (2) a BMI ranging from 27 to 45 kg/m$^2$ and dyslipidemia or hypertension. Participants were submitted to a mild hypocaloric diet and exercise and were randomly assigned in a 1:1:1 ratio to receive: (1) naltrexone 32 mg plus bupropion 360 mg; (2) naltrexone 16 mg plus bupropion 360 mg; and (3) a placebo (Figure 3).

The combination treatment produced greater improvements in waist circumference, insulin resistance, and concentrations of HDL cholesterol, triglycerides and hs-CRP compared with placebo[24]. Moreover, naltrexone 32 mg plus bupropion 360 mg originated a better improvement in fasting insulin and glucose concentrations[24]. The groups treated with naltrexone plus bupropion revealed a better weight-related quality of life than the placebo group did, mainly as a result of increased physical function and self-esteem[24]. Treatment with naltrexone 32 mg plus bupropion 360 mg provided greater weight loss as well as greater improvements in cardiometabolic risk factors and quality of life than the naltrexone 16 mg plus bupropion 360 mg formulation. In participants who completed 56 wk of treatment, weight loss was greater in the combination treatment groups (naltrexone 32 mg plus bupropion, -8.1%; naltrexone 16 mg plus bupropion, -6.7%) than in the placebo group (-1.8%). Due to the effects of bupropion, combination treatment originated a small and transient increase in mean systolic and diastolic blood pressure. However, it was subsequently observed that greater weight loss correlated with larger decreases in systolic and diastolic blood pressure in all treatment groups. Although more adverse effects were reported in the groups receiving naltrexone plus bupropion than in the placebo group, those adverse effects were mostly mild

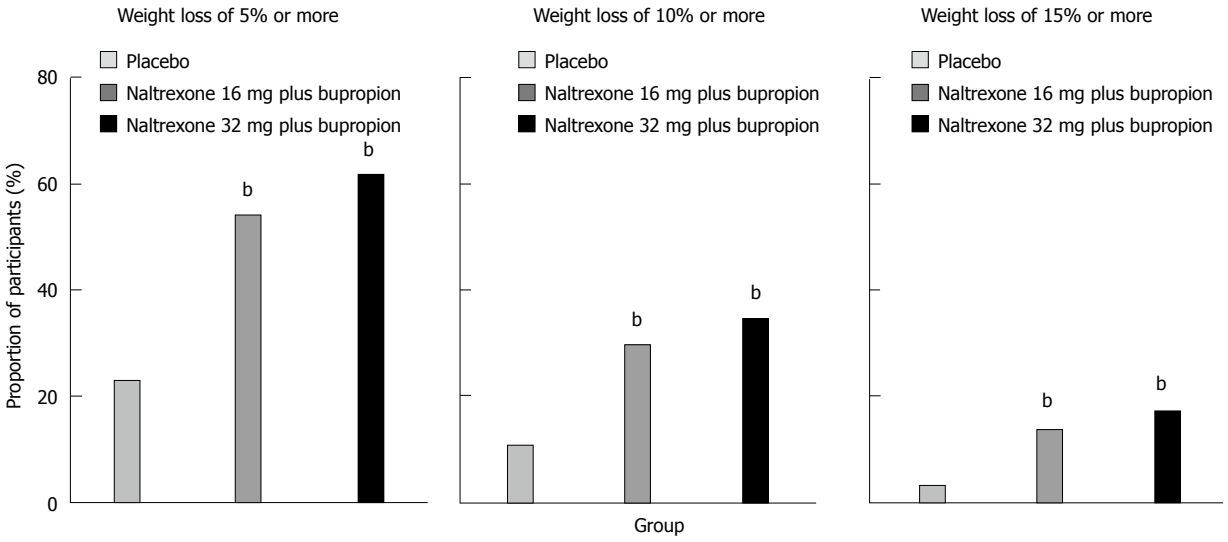



**Figure 3 Proportion of participants who lost at least 5%, 10% and 15% of baseline weight at week 56.** $^{b}P < 0.0001$ *vs* placebo. From[24], with permission from Elsevier.

to moderate and transient, having been consistent with the known safety and tolerability profiles of the individual drugs. It was then concluded that treatment with sustained-release naltrexone plus bupropion provides a new approach to manage obesity that might improve the ability to control eating behavior and response to food images[24].

The effects of bupropion and naltrexone in addictive disorders, along with evidence suggesting synergism of these drugs in reward pathways, suggest that this combination improves control of eating and response to food cravings. Since many overweight adults report food cravings to be an important obstacle to their ability to adhere to a diet, these actions could add to the utility of naltrexone plus bupropion in treatment of obesity[65].

Another study was the COR-behavior modification (BMOD)[66], a 56-wk, randomized, placebo-controlled trial, aiming to evaluate the efficacy and safety of naltrexone plus bupropion, in addition to an intensive BMOD. The participants were randomly assigned, in a 1:3 ratio, to receive either placebo or naltrexone SR 32 mg/d plus bupropion SR 360 mg/d. The main positive endpoint in the patients receiving treatment was change in body weight, a decrease in weight of 5% or more from baseline at week 56 having been achieved. The larger weight losses produced by naltrexone 32 mg plus bupropion 360 mg were also associated with greater reductions in markers of cardiometabolic risk, such as waist circumference, triglyceride and fasting insulin concentrations. Furthermore, this drug combination, along with BMOD, produced an increase in HDL cholesterol concentration, compared with placebo + BMOD[66]. Despite the reduction in systolic and diastolic blood pressure being higher in the placebo group, additional analyses revealed a relationship between greater blood pressure reductions and greater weight loss in patients treated with naltrexone 32 mg plus bupropion 360 mg + BMOD. The most common adverse effect in the groups treated with naltrexone SR/bupropion SR was mild-to-moderate transient nausea, a well-known side effect of naltrexone, which was dose-dependent. Other reported adverse effects were constipation, dizziness, dry mouth, urticaria, anxiety, disturbance in attention, increase in blood pressure and vomiting. In conclusion, the combination of naltrexone SR/bupropion SR, along with an intensive behavioral modification program, achieved a greater weight loss than BMOD alone, having generally been well tolerated.

**Bupropion/zonisamide:** The sustained-release combination of bupropion with the anti-epileptic agent zonisamide (Empatic™), developed by Orexigen Therapeutics Inc., has been evaluated in three Phase 2 trials, having produced more weight loss than treatment with either drug given alone[67-69]. The second Phase 2b clinical trial was designed to assess the safety and efficacy of Empatic™ in healthy, nondiabetic, obese patients[70].

Bupropion was approved by FDA in 1985 as antidepressant and in 1997 for smoking cessation. An immediate release formulation of zonisamide was approved in the United States in 2000 for the adjunctive treatment of partial seizures, a form of epilepsy. It is believed to exert its anticonvulsivant activity by blocking calcium and sodium channels[71]. The mechanism by which it decreases appetite is not known. Zonisamide increases dopamine and serotonin activities, possibly leading to a reduction in appetite and an increase in energy expenditure, and it is believed to suppress the neuropeptide Y/agouti-related protein (NPY/AgRP) neuron, further stimulating the POMC neurons[68]. It also inhibits carbonic anhydrase activity, which possibly alters the perception of taste[72].

A 24-wk Phase 2 study conducted by Orexigen Therapeutics Inc., demonstrated a weight loss of 7.5%, 6.1% and 1.4% from baseline weight with bupropion plus zonisamide 360 mg, bupropion plus zonisamide 120 mg and a placebo, respectively[73]. The most frequently reported adverse effects of zonisamide SR/bupropion SR combination were headache, nausea and insomnia[71].

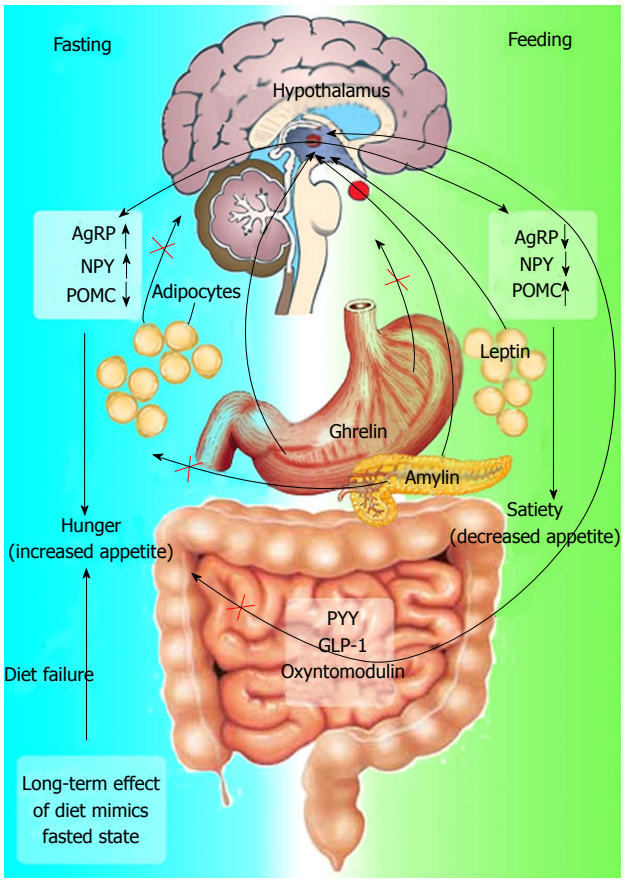



**Figure 4 Action of the orexigenic and anorexigenic hormones in the fasting and feeding states.** AgRP: Agouti-related protein; GLP-1: Glucagon-like peptide-1; NPY: Neuropeptide Y; PYY: Peptide tyrosine-tyrosine; POMC: Pro-opiomelanocortin.

### *Norepinephrine, dopamine and serotonin reuptake inhibitors*

**Tesofensine:** Tesofensine was initially developed for the treatment of Parkinson's disease or Alzheimer's disease, although it demonstrated a limited efficacy for this application[74]. However, it is a very promising development drug candidate for obesity management, having produced twice the weight loss seen with currently marketed drugs (-7.2 ± 6.7 kg, $P < 0.0001$)[75]. It simultaneously inhibits three monoaminergic neurotransmitters (dopamine, norepinephrine and serotonin) presynaptic reuptake, thereby increasing their neurotransmission. Dopamine acts in the nucleus accumbens of the forebrain, modulating reward and the "pleasure" feeling of food. Serotonin and norepinephrine act on the hypothalamus, increasing metabolism and reducing appetite. Tesofensine is under development by NeuroSearch for the potential treatment of obesity[76].

Sjödin *et al*[77] investigated the mechanisms by which tesofensine produces weight loss, measuring energy expenditure and appetite sensations in overweight and obese individuals. Tesofensine induced a 1.8 kg weight loss above placebo after 2 weeks' treatment ($P < 0.0001$). It was concluded that tesofensine has a significant effect on appetite sensations and a slight effect on energy expenditure at night and that both effects can contribute to the strong effect of tesofensine in weight reduction[77].

Phase 2 clinical trials with tesofensine were conducted in obese individuals and it was observed that tesofensine produced dose-related reductions in body weight, body fat and waist circumference, as well as improvements in other obesity-related endocrine factors. At the highest dose tested, tesofensine caused dose-dependent elevations in heart rate, with significant increases in blood pressure. However, the findings from these studies suggest that tesofensine may be a well-tolerated long-term treatment for obesity, with minimal cardiovascular effects. FDA appears to share this view, having authorized the Phase 3 trial program for this agent[78].

### *Glucagon-like peptide-1 receptor agonists*

In the last few years, knowledge concerning the mechanisms involved in appetite control has increased. The immediate postprandial state is accompanied by hormonal changes that include: (1) a reduction in the concentrations of the orexigenic gut hormone ghrelin; and (2) an increase in the anorexigenic peptides, such as peptide tyrosine-tyrosine (PYY), glucagon-like peptide-1 (GLP-1) and oxyntomodulin[79].

All of these hormones act on the hypothalamus, playing an important regulatory role in the mediation of hunger, satiety and energy intake (Figure 4)[79]. GLP-1 also exerts an incretin effect on insulin synthesis and secretion, inhibiting glucagon secretion, which may be responsible for postprandial insulin response[80].

GLP-1 is an incretin hormone secreted by enteroendocrine L-cells, which increases insulin secretion after oral ingestion of a meal[80]. Carbohydrates and fat are the most potent secretagogues for GLP-1 release[81].

The incretin response after meals is impaired in many patients with T2DM, owing to resistance to glucose-dependent insulinotropic peptide (GIP) that cannot be compensated for by endogenous GLP-1. Two classes of incretin therapies have been developed, in order to overcome the native GLP-1 short half-life (1-2 min): dipeptidyl peptidase-4 (DPP-4) inhibitors (gliptins), which raise endogenous levels of plasma GLP-1 and GIP, and DPP-4-resistant GLP-1 receptor agonists (liraglutide, exenatide)[82]. In Phase 3 clinical trials, agents from both of these classes of incretin-based therapies produced improvements in glycemic control in patients with T2DM[82]; however, efficacy is greater with GLP-1 RAs, which additionally promote weight loss, reducing systolic blood pressure and improving β-cell function. These effects are particularly advantageous to T2DM patients, as most are affected by β-cell dysfunction, obesity and hypertension[82].

The glucagon-like peptide-1 receptor agonists (GLP-1 RAs) *liraglutide* and *exenatide* (or exendin-4) improve glycemic control by stimulating insulin release through pancreatic β-cells, depending on glucose. GLP-1 receptors do not exist only in the pancreas.

Liraglutide (Victoza®) is currently approved by FDA and EMA as an anti-hyperglycemic agent in T2DM adults. Treatment with the 1.8 mg dose resulted in weight

loss of 2.0 ± 0.3 kg from baseline after 26 wk, while weight change with placebo varied from a weight loss of 1.5 kg to a 0.6 kg weight gain over that time[83]. Nevertheless, liraglutide causes dose-dependent thyroid C-cell tumors in rodents, being thus contraindicated in patients with a personal or family history of medullary thyroid carcinoma, although it is unknown whether liraglutide causes C-cell tumors in humans[84]. Gastrointestinal adverse events, namely nausea, diarrhea and vomiting, were the most frequently reported events during treatment with liraglutide[84]. Other reported adverse effects were: pancreatitis; hypoglycemia, when liraglutide was used with an insulin secretagogue or insulin; renal impairment; hypersensitivity reactions; headache; and anti-liraglutide antibody formation[85].

At the present, the main adverse effects of GLP-1 and GLP-1 RAs seem to be transient gastrointestinal events. A cardiovascular protective role has also been suggested for GLP-1 in Phase 3 clinical studies, raising the possibility that GLP-1 RAs may lower the risk of cardiovascular events in T2DM patients. In those studies, T2DM patients treated with liraglutide exhibited reduced lipid levels as well as decreased levels of cardiovascular risk biomarkers, including plasminogen activator inhibitor-1 (PAI-1), brain natriuretic peptide and hs-CRP (-7.6%, -11.9% and -23.1%, respectively)[82].

Exenatide (Byetta®) is approved by both FDA and EMA as an adjunct to diet and exercise to improve glycemic control in adults with T2DM[86,87]. A study was conducted to assess the effects of exenatide on body weight and glucose tolerance in nondiabetic obese individuals with normal or impaired glucose tolerance or impaired fasting glucose, submitting obese patients to exenatide or placebo, with simultaneous lifestyle intervention, for 24 wk[88]. Patients treated with exenatide lost 5.1 ± 0.5 kg from baseline versus 1.6 ± 0.5 kg with placebo. Placebo-subtracted difference in percent weight reduction was -3.3% ± 0.5% ($P < 0.001$). It was observed that exenatide plus lifestyle modification reduced caloric intake, having resulted in greater weight loss than lifestyle modification alone (treatment difference, -3.3%), in nondiabetic obese subjects, with improved glucose tolerance in individuals with impaired glucose tolerance and impaired fasting glucose. Thus, some authors have speculated that exenatide, along with lifestyle modification, has potential as a treatment for obesity in individuals at high risk for developing T2DM[89]. However, the efficacy of weight loss for GLP-1 agonists does not meet FDA standards and they are not well-tolerated for long-term use because they result in nausea and diarrhea. Indeed, nausea was experienced by 25% and 4% and diarrhea by 14% and 3% of exenatide and placebo-treated subjects, respectively[88].

Oxyntomodulin is another gut hormone and GLP-1 receptor agonist, which is released from the small intestine after each meal in proportion to the calories intake. Early studies showed that administration of synthetic oxyntomodulin reduces appetite, suggesting that endogenous oxyntomodulin signals to the brain a feeling of "fullness" or "satiety" after meals[90]. Wynne *et al*[91] demonstrated that oxyntomodulin reduces energy intake and increases energy expenditure, resulting in negative energy balance. These data support its role as a potential anti-obesity therapy. Field *et al*[92] carried out a study aiming to investigate whether the anorectic effects of PYY, a $Y_2$ receptor agonist, and oxyntomodulin, a GLP-1 receptor agonist, can be additive. Likewise PYY, oxyntomodulin is secreted by intestinal L-cells after meals. It was observed that a combined infusion of PYY and oxyntomodulin produces an additive anorectic effect in overweight and obese humans, reducing food intake. Furthermore, this and other studies suggest that coadministration of a $Y_2$ receptor agonist and a GLP-1 receptor agonist may be an important strategy for treating obesity, although further studies are necessary[92].

### Amylinomimetic drugs

*Amylin* is cosecreted with insulin from pancreatic β-cells and, like GLP-1, it has been reported to inhibit glucagon secretion, delay gastric emptying, reduce appetite and, therefore, food intake in humans (Figure 4)[93,94]. Amylin may, thus, act as an important regulator of postprandial glucose homeostasis[95]. *In vitro*, it forms amyloid and undergoes a change in conformation, which leads to loss of biological activity[96,97]. However, the study of analogues of human amylin showed that substitution of alanine and serine at positions 25, 28 and 29, originating pramlintide, promotes stability in solution and allows preservation of biological activity[98]. Thus, responses to the new drug pramlintide provided knowledge of the biology of amylin, as it has the same mechanism of action. Pramlintide, a synthetic analogue of amylin, reduces appetite and food intake, thus producing body weight loss[99]. Pramlintide (Symlin®) is currently approved in the United States as an anti-hyperglycemic agent for patients with diabetes mellitus as an adjunctive therapy to mealtime insulin, being under investigation as a potential treatment for obesity[100].

A 2010 randomized, placebo-controlled trial investigated the efficacy and safety of pramlintide monotherapy or in association with either phentermine or sibutramine, in addition to lifestyle modification[101]. A greater weight loss was reported with both pramlintide plus phentermine (11.3% ± 0.9%) and pramlintide plus sibutramine (11.1% ± 1.1%) than with pramlintide (3.7% ± 0.7%) or placebo alone (2.2% ± 0.7%), confirming the potential of both combinations in treating obesity. With respect to categorical weight loss, ≥ 5% weight loss was achieved by 78% of pramlintide + sibutramine and 82% of pramlintide + phentermine evaluable subjects. These percentages for both combination treatments were significantly greater ($P < 0.001$) than the percent of subjects achieving ≥ 5% weight loss with placebo (28%) or pramlintide alone (36%). Nevertheless, a greater reduction in blood pressure was observed with both placebo and pramlintide than with combination treatments, which generally produced an elevation in diastolic blood pressure and heart

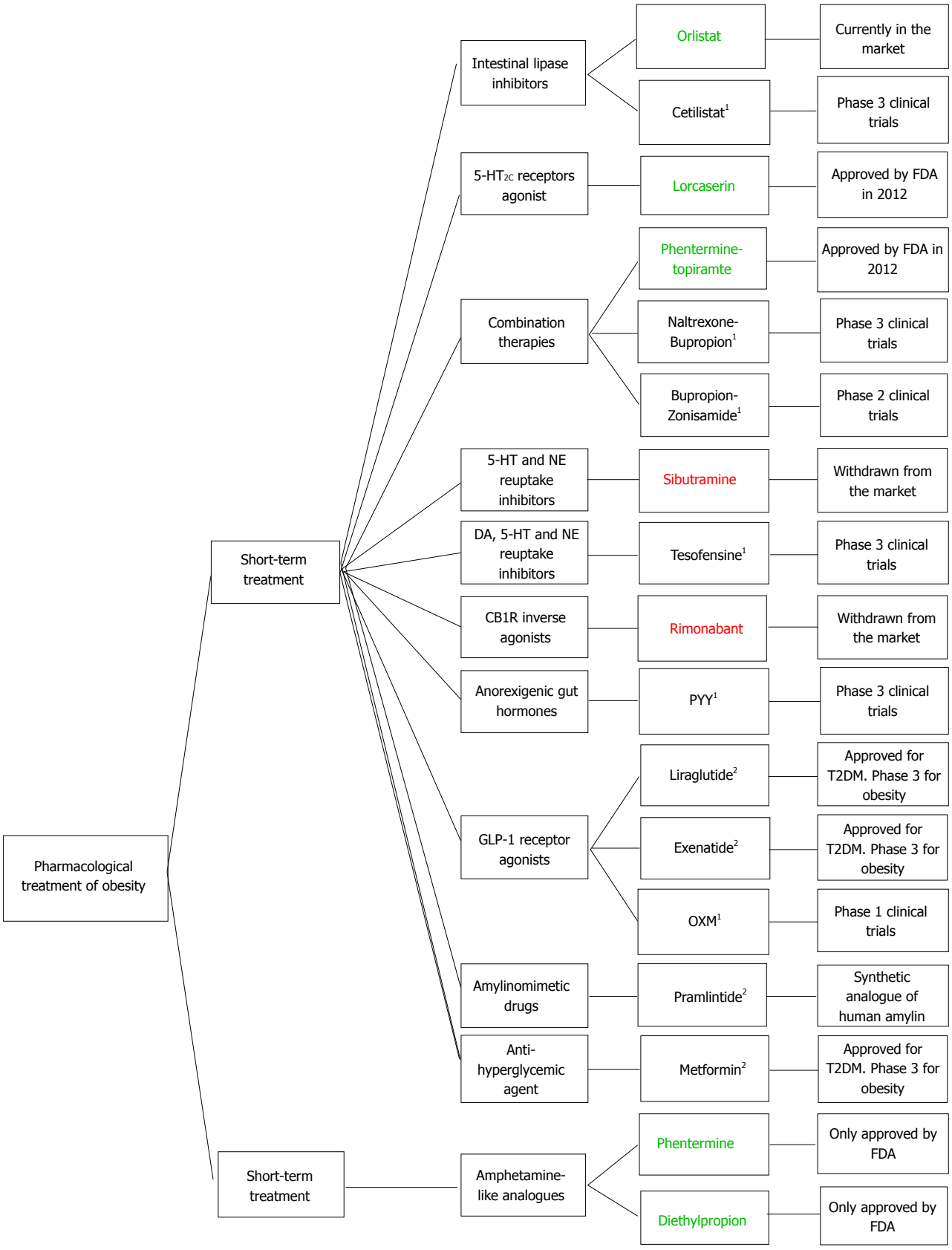



**Figure 5 Short-term and long-term pharmacological treatment of obesity.** [1]Under investigation (clinical trials); [2]Anti-obesity effects being investigated. 5-HT: 5-hidroxytryptamine; CB1R: Cannabinoid-1 receptors; DA: dopamine; FDA: United States Food and Drug Administration; GLP-1: Glucagon-like peptide 1; NE: Norepinephrine; OXM: Oxyntomodulin; PYY: Peptide tyrosine-tyrosine; POMC: Pro-opiomelanocortin; T2DM: Type 2 diabetes mellitus.

rate[101]. This is a matter of concern, owing to typically elevated cardiometabolic risk profile in obese individuals. An increase in systolic blood pressure was not observed in the group treated with pramlintide plus sibutramine,

in contrast to what had been previously reported with sibutramine monotherapy. Both combination treatments improved total and HDL cholesterol, as well as triglycerides levels. On the other hand, pramlintide plus phentermine improved LDL cholesterol levels. The most frequent adverse effect seen with pramlintide monotherapy was nausea, which generally resolved over time. The incidence of psychiatric adverse effects was low, the combination having produced a reduction in the Hospital Anxiety and Depression Scale depression subscore, a self-assessment mood questionnaire[101].

Smith *et al*[102] demonstrated that pramlintide treatment reduces daily food intake (including fast food intake) and meal sizes, besides improving perceived control of eating in obese individuals and reducing binge eating tendencies. Binge eating disorder is a psychopathology characterized by frequent and persistent episodes of binge eating with the feeling of losing the control due to the absence of regular compensatory behaviors[103].

Leptin is a neurohormone mainly secreted by adipocytes and plays an important role in regulating insulin action and long-term energy homeostasis (Figure 4). Obesity is associated with resistance to many of the metabolic effects of leptin, leading to the notion that resistance to leptin is a cause of insulin resistance associated with obesity[104]. Leptin deficiency produces an increase in body weight and insulin resistance, with consequences to glucose metabolism[105]. It is known that replacement treatment with leptin improves insulin sensitivity in people with leptin deficiency[106,107].

As the neurohormonal control of body weight involves an interaction between long-term adiposal signals (*e.g.,* leptin, insulin) and short-term satiation signals (*e.g.*, amylin, PYY, GLP-1, cholecystokinin), Ravussin *et al*[108] examined the effect on weight loss of the agonism amylin/leptin, using pramlintide and metreleptin. A 24-wk, randomized, double-blind, active-drug controlled study was developed in obese or overweight patients[108]. The combination pramlintide plus metreleptin produced a greater weight loss (-12.7% ± 0.9%) in the first 20 wk of the study than monotherapy with pramlintide (-8.4% ± 0.9%; $P < 0.001$) or metreleptin (-8.2% ± 1.3%; $P < 0.01$), demonstrating synergism between the two drugs. The most frequent adverse effects were mild to moderate nausea and reactions at the injection site. However, this study did not elucidate the mechanism(s) for the observed increase in weight loss, even though the results support further development of pramlintide/metreleptin, towards obesity pharmacotherapy[108].

### Metformin

There are currently no medications approved by FDA to treat obesity in children aged under 12 years. Metformin is approved by FDA for the treatment of T2DM in children aged 10 years or over, having received attention for its potential to be used with weight control purposes in children[109]. It suppresses hepatic glucose production and, at higher doses, improves peripheral insulin sensitivity[110,111]. It is proved that metformin decreases food consumption, stabilizes weight and can induce small weight losses in diabetic and nondiabetic adults[112-114]. Yanovski *et al*[115] studied the effects of metformin in obese insulin-resistant children aged between 6 and 12 years old, having concluded that treatment with metformin produces a modest reduction in body weight and adiposity, although it improves glucose homeostasis in obese children. Metformin may also be helpful in counteracting the weight gain observed in children taking atypical psychotropic medications. Nevertheless, cessation of metformin therapy may lead to rebound hyperinsulinemia and rapid weight gain[116]. Concerning side effects of metformin, the most common are nausea, flatulence, bloating and diarrhea, all of them usually resolved. Metformin should not be used in renal failure or with intravenous contrast. Multivitamin supplementation is strongly recommended[109].

Ghandi *et al*[117] carried out a study to compare the effects of metformin and orlistat in obese women with polycystic ovary syndrome and concluded that both metformin and orlistat produce similar reductions in weight, BMI and waist circumference. Furthermore, metformin caused a significant reduction in serum triglyceride and a higher ovulation rate than orlistat, although the difference was not statistically significant.

On the other hand, Wiegand *et al*[118] studied the effect of metformin therapy in obese insulin-resistant adolescents, having reached the conclusion that metformin and lifestyle modification improved insulin sensitivity, although it did not produce weight loss. Curiously, a higher prevalence of side effects was reported in the placebo group than in the group receiving metformin, showing how important the belief in therapy effects and self-efficacy may be, as it increases the motivation to adopt a healthy lifestyle[118].

### Other possible treatments under investigation

Baclofen, a γ-aminobutiric acid B ($GABA_B$) agonist, was demonstrated to reduce body weight more than 2% (baseline: 93.3 ± 9.8 kg; 12 wk: 91.7 ± 10.3 kg, $P < 0.05$) and waist circumference more than 2 cm in 5 out of 10 subjects, without observed adverse effects in obese patients[119]. Although the reported anti-obesity effects were mild, baclofen may constitute a new anti-obesity drug in humans.

Lactoferrin is a Fe-binding glycoprotein which is present at high concentrations in mammalian milk. It is multi-functional, having antibacterial, antiviral, immunostimulatory, antioxidant and cancer-preventive potential[120-123]. As lactoferrin is ingested by infants in breast milk, it is considered to be safe, having been approved as a food additive in Japan. Ono *et al*[124] demonstrated that the ingestion of enteric-coated lactoferrin during 8 wk can reduce visceral fat, body weight (-1.5 kg, + 1.0 kg than placebo), BMI and waist circumference both in men and women without the need for lifestyle modifications. Thus, it has potential to prevent obesity and decrease the risk of metabolic syndrome.

## DRUGS THAT HAVE BEEN WITHDRAWN FROM THE MARKET

### *Serotonin and norepinephrine reuptake inhibitors*

**Sibutramine:** Sibutramine is a norepinephrine and serotonin reuptake inhibitor that was approved in 1997 for weight management and was withdrawn from the market by FDA and EMA in 2010, based on the Sibutramine Cardiovascular OUTcome (SCOUT) trial, which demonstrated a 16% increase in risk of major adverse cardiovascular events (non-fatal heart attack, non-fatal stroke, resuscitation after cardiac arrest and cardiovascular death). The long-term effects of sibutramine treatment in combination with diet and exercise on the rates of cardiovascular events and cardiovascular death among individuals with high cardiovascular disease risk were evaluated in the SCOUT trial[125,126].

James *et al*[125] investigated the effect of sibutramine on cardiovascular outcomes in overweight and obese subjects, 55 years of age or older, with preexisting cardiovascular disease, T2DM, or both, to examine the cardiovascular consequences of weight management with or without sibutramine in individuals at high risk for cardiovascular events. In the SCOUT trial, event rates in the sibutramine group were lower than expected, and the sibutramine group lost significantly more weight than the placebo group ($P = 0.001$) and maintained the weight loss[125]. However, the risk of a primary outcome event was 16% higher in the sibutramine group than in the placebo group ($P = 0.02$). It was concluded that individuals with a history of cardiovascular conditions who were receiving long-term sibutramine treatment had an increased risk of nonfatal myocardial infarction ($P = 0.02$) and nonfatal stroke ($P = 0.03$), but not of cardiovascular death ($P = 0.90$) or death from any cause ($P = 0.54$). On the basis of these results, the authors of the SCOUT trial concluded that sibutramine should continue to be excluded from use in patients with preexisting cardiovascular disease[125]. These results also led to the withdrawal of sibutramine from the market.

### *Cannabinoid type 1 receptor inverse agonists*

The cannabinoid-1 receptors (CB1Rs) are highly expressed in the CNS, autonomic gastric endings of the peripheral nervous system, and other key cells involved in body energy metabolism, such as adipocytes, hepatocytes and myocytes[127-129]. It has been hypothesized that CB1R inverse agonists act directly on adipocytes, hepatocytes, pancreatic islets and skeletal muscle to mediate metabolic effects[130].

**Rimonabant:** Rimonabant was the first selective cannabinoid-1 receptor blocker tested in large clinical trials, having been used to reduce body weight and improve cardiometabolic risk factor in obese patients. Rimonabant (Acomplia® in Europe) received marketing approval from the EMA in June 2006. However, it was withdrawn from the market by EMA in 2008, due to safety concerns, namely psychiatric adverse effects. In June 2007, FDA unanimously voted against approval of rimonabant (Zimulti® in the United States) because of increased risk of neurological and psychiatric side effects, namely seizures, depression, anxiety, insomnia, aggressiveness and suicidal thoughts, having never approved its use[131,132].

A prospective, double-blind, placebo-controlled, randomized, two-arm parallel group trial (Atherosclerosis Underlying Development assessed by Intima-media Thickness in patients On Rimonabant/AUDITOR[133] was carried out to determine whether treatment with 20 mg rimonabant would reduce progression of carotid intima-media thickness over 30 mo of treatment in abdominally obese patients with metabolic syndrome. The results of this study demonstrate that rimonabant correlates with an increase in psychiatric symptoms, nervous system disorders (*e.g.*, depression, anxiety) and gastrointestinal disorders. Furthermore, the study suggests that a 5% loss of body weight over a 30-mo period with rimonabant is not sufficient to modify atherosclerosis progression in the carotid artery in obese individuals with the metabolic syndrome[133].

Motaghedi *et al*[134] observed that rimonabant administration may be efficacious for weight loss in adults with Prader Willi syndrome, characterized by increased appetite and decreased energy expenditure with lower insulin-like growth factor-1, leading to adiposity. Nevertheless, it causes high risk psychiatric adverse effects.

Horder *et al*[135] findings were the first to show that rimonabant inhibits the neural process of rewarding food stimuli in humans. Thus, it may not only promote weight loss, but it may also indicate a mechanism for inducing anhedonia (*i.e.*, loss of the capacity to experience pleasure), which could increase the risk of depressive symptomatology observed in clinical use.

Clinical trials did not support further development of both taranabant and CP 945598, because of psychiatric adverse effects, similar to those of rimonabant.

Figure 5 provides an overview of all drugs currently approved for both long-term and short-term treatment of obesity, the new drugs presently registered in ongoing clinical trials and those recently withdrawn from the market.

Figure 6 summarizes assessment of overweight (25 kg/m$^2$ $\leqslant$ BMI $<$ 30 kg/m$^2$) and obese (BMI $\geqslant$ 30 kg/m$^2$) patients and pharmacological and non-pharmacological treatment measures, taking into account currently approved anti-obesity drugs both in Europe (by EMA) and in the United States (by FDA).

Table 1 summarizes the subtracted drug/placebo mean weight loss (expressed in terms of percentages or kilograms, as available) of the currently available drugs for long-term treatment of obesity, and of the candidate drugs that have shown more efficacy in clinical trials. It also mentions the main beneficial and adverse effects of each anti-obesity drug, allowing to compare their efficacy and safety.

## CONCLUSION

Obesity is a chronic disease and an important risk factor for cardiovascular and metabolic diseases, such as T2DM.

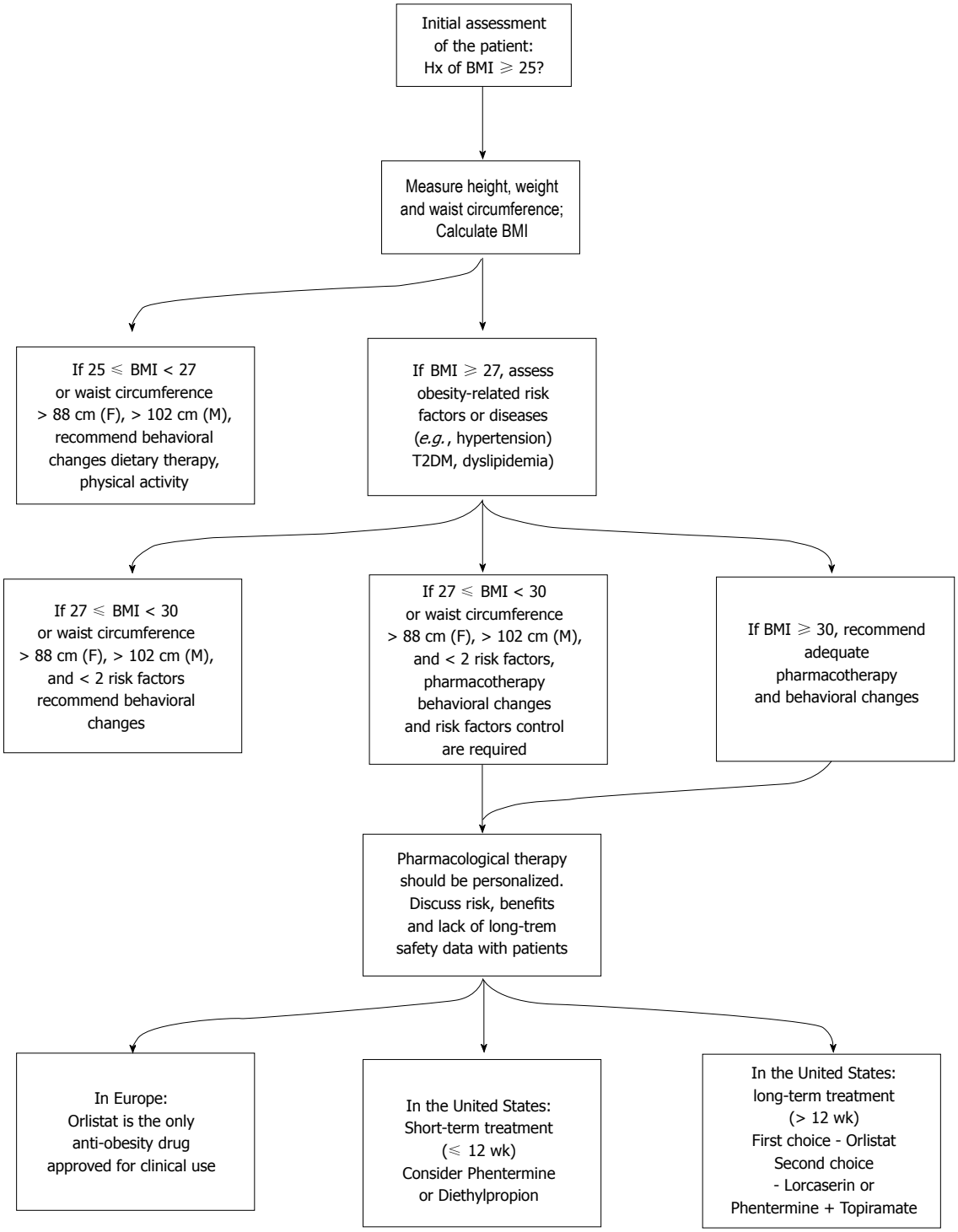



**Figure 6 Algorithm for assessment and treatment of overweight and obese patients.** BMI: Body mass index; F: Female; Hx: History; M: Male; T2DM: Type 2 diabetes mellitus.

As lifestyle interventions alone hardly ever result in long-term weight loss, pharmacotherapy is an important adjunct to lifestyle measures to improve the induction and maintenance of weight loss.

Orlistat was until recently the only pharmacological long-term treatment for obesity available on the market, as rimonabant and sibutramine had been withdrawn in 2008 and 2010, respectively, due to serious psychiatric and cardiovascular adverse effects. Lorcaserin (Belviq®) and phentermine plus topiramate (Qsymia™) were recently approved by FDA, in June and July 2012, respectively. Considering the drugs currently available on the market, orlistat is the only whose mechanism of action does not involve the CNS, but the inhibition of gastrointestinal lipase. Lorcaserin activates 5-HT$_{2C}$ receptors, thus suppressing appetite. Qsymia™ is an association of phentermine, a norepinephrine releasing drug, and topiramate, an antiepileptic drug which also exhibits weight loss

**Table 1 Benefits and risks of anti-obesity drugs**

| Drug | Drug/Placebo subtracted weight loss | Mechanism of action | Other positive effects | Adverse effects |
|---|---|---|---|---|
| Orlistat (Xenical®) | 3.6% | Decreases fat absorption | Improvement in most risk factors of cardiovascular disease and CRP | Increased flatulency Fecal incontinence Diarrhea |
| Lorcaserin (Belviq®) | 3.6% | 5-HT2C receptors agonist | Reduction in energy intake and appetite. Improvement in the markers of cardiovascular disease and inflammation | Headache, fatigue, dizziness, nausea, dry mouth, constipation. Disturbances in attention or memory. Risk of birth defects |
| Phentermine 7.5 + Topiramate 46 (Qsymia™) | 6.6% | Amphetamine analogue + Anticonvulsivant | Improvement of cardiometabolic variables | Psychiatric and cognitive adverse effects. Risk of birth defects |
| Phentermine 15 + Topiramate 92 | 8.6% | | | |
| Cetilistat | 1.46 kg | Decreases fat absorption | Improvement in glycemic control | Mild or moderate gastrointestinal effects |
| Naltrexone 16 + Bupropion (Contrave®) | 4.9% | μ-opioid antagonist + Stimulator of the hypothalamic POMC neurons | Improvement in cardiometabolic risk factors and in insulin resistance | Nausea, constipation, anxiety, increase in blood pressure |
| Naltrexone 32 + Bupropion | 6.3% | | | |
| Zonisamide 120 + Bupropion (Empatic™) | 4.7% | Increases monoamines (*e.g.*, DA) and POMC neurons activity | Reduction in appetite. Increase in energy expenditure | Headache, nausea, insomnia |
| Zonisamide 360 + Bupropion | 6.1% | | | |
| Tesofensine | 5.4 ± 6.7 kg | Inhibits reuptake of DA, 5-HT and NE | Reduction in appetite, body fat and waist circumference | Increase in heart rate and blood pressure |
| Liraglutide (Victoza®) | 0.5 ± 0.3 kg | GLP-1 agonist; increases insulin secretion after meals | Improvement in glycemic control and systolic blood pressure | Nausea, diarrhea, vomiting |
| Exenatide (Byetta®) | 3.5 ± 0.5 kg | GLP-1 agonist; increases insulin secretion after meals | Improvement in glycemic control and systolic blood pressure | Nausea, diarrhea |
| Pramlintide (Symlin®) | 1.5% ± 0.7% | Anti-hyperglycemic agent; synthetic analogue of human amylin | Reduction in binge eating tendencies and blood pressure | Nausea |

5-HT: 5-hidroxytryptamine; 5-HT2C: 5-hidroxytryptamine 2C receptors; CNS: Central nervous system; CRP: C-reactive protein; DA: Dopamine; GLP-1: Glucagon-like peptide 1; NE: Norepinephrine; POMC: Pro-opiomelanocortin.

properties. All these drugs require a combination with lifestyle changes (hypocaloric diet, physical activity) in order to show efficacy. Among the drugs under Phase 3 clinical trials, which have not been approved yet, the most promising ones appear to be cetilistat, which mechanism of action is identical to that of orlistat; Contrave®, an association of naltrexone and bupropion which acts on hypothalamic POMC neurons; tesofensine, a monoamine reuptake inhibitor; and GLP-1 agonists, such as liraglutide and exenatide, which exert an anorexigenic effect and may be helpful to T2DM obese patients. However, Contrave® and tesofensine have shown cardiovascular side effects in clinical trials, which led FDA not to approve them yet. Pramlintide, a synthetic analogue of human amylin currently approved for T2DM management, also produced significant weight loss in clinical trials, having shown a low incidence of psychiatric adverse effects. Enteric coated lactoferrin appears to be a promising treatment, as it does not require lifestyle changes. Baclofen may also constitute a new anti-obesity therapy, since it reduces body weight without reported adverse effects. Regarding the short-term treatment of obesity, phentermine and diethylpropion are currently the two available pharmacological options in the United States, having been withdrawn from the EU market.

In conclusion, owing to the limited options currently available for the pharmacological treatment of obesity, it is imperative to develop new safe compounds. The rapid growth of biomedical science in the last decade and the increasing understanding of the energy metabolism may provide new and better therapies for this pathologic condition.

## REFERENCES


1 **Cecchini M**, Sassi F, Lauer JA, Lee YY, Guajardo-Barron V, Chisholm D. Tackling of unhealthy diets, physical inactivity, and obesity: health effects and cost-effectiveness. *Lancet* 2010; **376**: 1775-1784 [PMID: 21074255 DOI: 10.1016/S0140-6736(10)61514-0]

2 **Ababakari AR**, Lauder W, Jones MC, Kirk A, Agyemang C, Bhopal RS. Prevalence and time trends in diabetes and physical inactivity among adult West African populations: the epidemic has arrived. *Public Health* 2009; **123**: 602-614 [PMID: 19748643 DOI: 10.1016/j.puhe.2009.07.009]

3 **WHO**. Facts about overweight and obesity (2012 July 2012). Available from: URL: http://www.who.int/mediacentre/factsheets/fs311/en/index.html

4 **WHO**. WHO Global Infobase, data for saving lives. (2011 August 2012). Available from: URL: http://apps.who.int/infobase/Index.aspx.
5 **Flegal KM**, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. *JAMA* 2010; **303**: 235-241 [PMID: 20071471 DOI: 10.1001/jama.2009.2014]
6 **Pi-Sunyer X**. The medical risks of obesity. *Postgrad Med* 2009; **121**: 21-33 [PMID: 19940414 DOI: 10.3810/pgm.2009.11.2074]
7 **Després JP**, Lemieux I. Abdominal obesity and metabolic syndrome. *Nature* 2006; **444**: 881-887 [PMID: 17167477 DOI: 10.1038/nature05488]
8 **Flegal KM**, Graubard BI, Williamson DF, Gail MH. Cause-specific excess deaths associated with underweight, overweight, and obesity. *JAMA* 2007; **298**: 2028-2037 [PMID: 17986696 DOI: 10.1001/jama.298.17.2028]
9 **Nguyen T**, Lau DC. The obesity epidemic and its impact on hypertension. *Can J Cardiol* 2012; **28**: 326-333 [PMID: 22595448 DOI: 10.1016/j.cjca.2012.01.001]
10 **American Diabetes Association.** Executive summary: Standards of medical care in diabetes--2012. *Diabetes Care* 2012; **35** Suppl 1: S4-S10 [PMID: 22187471 DOI: 10.2337/dc12-s004]
11 **Haslam DW**, James WP. Obesity. *Lancet* 2005; **366**: 1197-1209 [PMID: 16198769 DOI: 10.1016/S0140-6736(05)67483-1]
12 **Stamler R**, Stamler J, Riedlinger WF, Algera G, Roberts RH. Weight and blood pressure. Findings in hypertension screening of 1 million Americans. *JAMA* 1978; **240**: 1607-1610 [PMID: 691146]
13 **Laaksonen DE**, Lakka HM, Niskanen LK, Kaplan GA, Salonen JT, Lakka TA. Metabolic syndrome and development of diabetes mellitus: application and validation of recently suggested definitions of the metabolic syndrome in a prospective cohort study. *Am J Epidemiol* 2002; **156**: 1070-1077 [PMID: 12446265]
14 **Lakka HM**, Laaksonen DE, Lakka TA, Niskanen LK, Kumpusalo E, Tuomilehto J, Salonen JT. The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *JAMA* 2002; **288**: 2709-2716 [PMID: 12460094]
15 **Lemieux I**, Pascot A, Couillard C, Lamarche B, Tchernof A, Alméras N, Bergeron J, Gaudet D, Tremblay G, Prud'homme D, Nadeau A, Després JP. Hypertriglyceridemic waist: A marker of the atherogenic metabolic triad (hyperinsulinemia; hyperapolipoprotein B; small, dense LDL) in men? *Circulation* 2000; **102**: 179-184 [PMID: 10889128]
16 **Ayyad C**, Andersen T. Long-term efficacy of dietary treatment of obesity: a systematic review of studies published between 1931 and 1999. *Obes Rev* 2000; **1**: 113-119 [PMID: 12119984]
17 **Wadden TA**, Butryn ML, Byrne KJ. Efficacy of lifestyle modification for long-term weight control. *Obes Res* 2004; **12** Suppl: 151S-162S [PMID: 15687411 DOI: 10.1038/oby.2004.282]
18 **Cannon CP**, Kumar A. Treatment of overweight and obesity: lifestyle, pharmacologic, and surgical options. *Clin Cornerstone* 2009; **9**: 55-68; discussion 69-71 [PMID: 19789064]
19 **Khan A**, Raza S, Khan Y, Aksoy T, Khan M, Weinberger Y, Goldman J. Current updates in the medical management of obesity. *Recent Pat Endocr Metab Immune Drug Discov* 2012; **6**: 117-128 [PMID: 22435392]
20 **Colquitt JL**, Picot J, Loveman E, Clegg AJ. Surgery for obesity. *Cochrane Database Syst Rev* 2009; **(2)**: CD003641 [PMID: 19370590 DOI: 10.1002/14651858.CD003641.pub3]
21 **Kiortsis DN**, Filippatos TD, Elisaf MS. The effects of orlistat on metabolic parameters and other cardiovascular risk factors. *Diabetes Metab* 2005; **31**: 15-22 [PMID: 15803108]
22 **Tzotzas T**, Samara M, Constantinidis T, Tziomalos K, Krassas G. Short-term administration of orlistat reduced daytime triglyceridemia in obese women with the metabolic syndrome. *Angiology* 2007; **58**: 26-33 [PMID: 17351155 DOI: 10.1177/0003319706297915]
23 **Didangelos TP**, Thanopoulou AK, Bousboulas SH, Sambanis CL, Athyros VG, Spanou EA, Dimitriou KC, Pappas SI, Karamanos BG, Karamitsos DT. The ORLIstat and CArdiovascular risk profile in patients with metabolic syndrome and type 2 DIAbetes (ORLICARDIA) Study. *Curr Med Res Opin* 2004; **20**: 1393-1401 [PMID: 15383188 DOI: 10.1185/030079904125004466]
24 **Greenway FL**, Fujioka K, Plodkowski RA, Mudaliar S, Guttadauria M, Erickson J, Kim DD, Dunayevich E. Effect of naltrexone plus bupropion on weight loss in overweight and obese adults (COR-I): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 2010; **376**: 595-605 [PMID: 20673995 DOI: 10.1016/S0140-6736(10)60888-4]
25 **FDA**. CONTRAVE® (NALTREXONE SR/BUPROPION SR COMBINATION) Advisory Meeting Briefing Document (Endocrinologic and Metabolic Drugs Advisory Committee Meeting 2010 July 2012). Available from: URL: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM235671.pdf
26 **Ioannides-Demos LL**, Piccenna L, McNeil JJ. Pharmacotherapies for obesity: past, current, and future therapies. *J Obes* 2011; **2011**: 179674 [PMID: 21197148 DOI: 10.1155/2011/179674]
27 **FDA**. Reference ID: 3115694. 2012 October 2012]. Available from: URL: http://www.accessdata.fda.gov/drugsatfda_docs/label/2012/088023s037lbl.pdf
28 **Maher W**. Phentermine Labeling: MAO Inhibition. 2000 August 2012]. Available from: URL: http://www.fda.gov/ohrms/dockets/dailys/00/Sep00/091500/pdn0001.pdf
29 **Melnikova I**, Wages D. Anti-obesity therapies. *Nat Rev Drug Discov* 2006; **5**: 369-370 [PMID: 16802443 DOI: 10.1038/nrd2037]
30 **MHRA**. European withdrawal of anorectic agents/appetite suppressants: new legal developments, no new safety issues: Licences for phentermine and amfepramone being withdrawn May 2001. [2001 August 2012]. Available from: URL: http://www.mhra.gov.uk/Safetyinformation/Safetywarningsalertsandrecalls/Safetywarningsandmessagesformedicines/CON019540
31 **Kang JG**, Park CY. Anti-Obesity Drugs: A Review about Their Effects and Safety. *Diabetes Metab J* 2012; **36**: 13-25 [PMID: 22363917 DOI: 10.4093/dmj.2012.36.1.13]
32 **FDA**. FDA approved obesity drugs. [2011 July 2012]. Available from: URL: http://www.fda.gov/ohrms/dockets/ac/04/briefing/2004-4068B1_05_Approved-Drugs.pdf
33 **FDA**. Tenuate (R). 2003 August 2012. Available from: URL: http://www.accessdata.fda.gov/drugsatfda_docs/label/2004/11722s029,12546s032lbl.pdf
34 **Arias HR**, Santamaría A, Ali SF. Pharmacological and neurotoxicological actions mediated by bupropion and diethylpropion. *Int Rev Neurobiol* 2009; **88**: 223-255 [PMID: 19897080 DOI: 10.1016/S0074-7742(09)88009-4]
35 **Cercato C**, Roizenblatt VA, Leança CC, Segal A, Lopes Filho AP, Mancini MC, Halpern A. A randomized double-blind placebo-controlled study of the long-term efficacy and safety of diethylpropion in the treatment of obese subjects. *Int J Obes* (Lond) 2009; **33**: 857-865 [PMID: 19564877 DOI: 10.1038/ijo.2009.124]
36 **FDA**. Orlistat (marketed as Alli and Xenical) Information. 2010 [October 2012]. Available from: URL: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm180076.htm
37 **Anderson JW**. Orlistat for the management of overweight individuals and obesity: a review of potential for the 60-mg, over-the-counter dosage. *Expert Opin Pharmacother* 2007; **8**: 1733-1742 [PMID: 17685889 DOI: 10.1517/14656566.8.11.1733]
38 **FDA**. Guidance for Industry Developing Products for Weight Management. 2007 [October 2012]. Available from: URL: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071612.pdf
39 **FDA**. Reference ID: 3074639. [2012 October 2012]. Available

from: URL: http://www.accessdata.fda.gov/drugsatfda_docs/label/2012/020766s029lbl.pdf
40 **Royal_Pharmaceutical_Society_of_Great_Britain**. Practice guidance: OTC Orlistat. 2009 [July 2012]. Available from: URL: http://www.youngsure.com/images/html/practice.pdf
41 **Yanovski SZ**, Yanovski JA. Obesity. *N Engl J Med* 2002; **346**: 591-602 [PMID: 11856799 DOI: 10.1056/NEJMra012586]
42 **Yancy WS**, Westman EC, McDuffie JR, Grambow SC, Jeffreys AS, Bolton J, Chalecki A, Oddone EZ. A randomized trial of a low-carbohydrate diet vs orlistat plus a low-fat diet for weight loss. *Arch Intern Med* 2010; **170**: 136-145 [PMID: 20101008]
43 **Derosa G**, Maffioli P, Ferrari I, D'Angelo A, Fogari E, Palumbo I, Randazzo S, Cicero AF. Orlistat and L-carnitine compared to orlistat alone on insulin resistance in obese diabetic patients. *Endocr J* 2010; **57**: 777-786 [PMID: 20683173]
44 **de Castro JJ**, Dias T, Chambel P, Carvalheiro M, Correia LG, Guerreiro L, Marques O, Medina JL, Nobre E, Nunes JS, Pereiraa MC, Polónia J, Portugal J, Raimundo A, Ruas A, da Silva PM, Vasconcelos C, Reis JL, Teles AG. A randomized double-blind study comparing the efficacy and safety of orlistat versus placebo in obese patients with mild to moderate hypercholesterolemia. *Rev Port Cardiol* 2009; **28**: 1361-1374 [PMID: 20301983]
45 **Kopelman P**, Bryson A, Hickling R, Rissanen A, Rossner S, Toubro S, Valensi P. Cetilistat (ATL-962), a novel lipase inhibitor: a 12-week randomized, placebo-controlled study of weight reduction in obese patients. *Int J Obes* (Lond) 2007; **31**: 494-499 [PMID: 16953261 DOI: 10.1038/sj.ijo.0803446]
46 **Kopelman P**, Groot Gde H, Rissanen A, Rossner S, Toubro S, Palmer R, Hallam R, Bryson A, Hickling RI. Weight loss, HbA1c reduction, and tolerability of cetilistat in a randomized, placebo-controlled phase 2 trial in obese diabetics: comparison with orlistat (Xenical). *Obesity* (Silver Spring) 2010; **18**: 108-115 [PMID: 19461584]
47 **Bryson A**, de la Motte S, Dunk C. Reduction of dietary fat absorption by the novel gastrointestinal lipase inhibitor cetilistat in healthy volunteers. *Br J Clin Pharmacol* 2009; **67**: 309-315 [PMID: 19220279]
48 **Drugdevelopment-technology.com**. Cetilistat - Investigational Drug for Obesity (2011 August 2012). Available from: URL: http://www.drugdevelopment-technology.com/projects/cetilistat/#.UCaiod1u38o
49 **FDA**. FDA approves Belviq to treat some overweight or obese adults. 2012 July 2012]. Available from: URL: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm309993.htm
50 **FDA**. FDA Briefing Document NDA 22529 Lorqess (lorcaserin hydrochloride) Tablets, 10 mg. 2012 [October 2012]. Available from: URL: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM225631.pdf
51 **Martin CK**, Redman LM, Zhang J, Sanchez M, Anderson CM, Smith SR, Ravussin E. Lorcaserin, a 5-HT(2C) receptor agonist, reduces body weight by decreasing energy intake without influencing energy expenditure. *J Clin Endocrinol Metab* 2011; **96**: 837-845 [PMID: 21190985]
52 **Smith SR**, Weissman NJ, Anderson CM, Sanchez M, Chuang E, Stubbe S, Bays H, Shanahan WR. Multicenter, placebo-controlled trial of lorcaserin for weight management. *N Engl J Med* 2010; **363**: 245-256 [PMID: 20647200]
53 **Garvey WT**, Ryan DH, Look M, Gadde KM, Allison DB, Peterson CA, Schwiers M, Day WW, Bowden CH. Two-year sustained weight loss and metabolic benefits with controlled-release phentermine/topiramate in obese and overweight adults (SEQUEL): a randomized, placebo-controlled, phase 3 extension study. *Am J Clin Nutr* 2012; **95**: 297-308 [PMID: 22158731 DOI: 10.3945/ajcn.111.024927]
54 **Lemke TL**, Williams DA, Roche VF, Zito SW. Foye's Principles of Medicinal Chemistry. 6ed, USA: Lipipincott Williams & Wilkins, 2008
55 **FDA**. Medication Guide Qsymia™ (2012 October 2012). Available from: URL: http://www.fda.gov/downloads/Drugs/DrugSafety/UCM312590.pdf
56 Avvo. Qsymia- Phenteramine/ Topiramate for weight reduction (2012 October 2012). Available from: URL: http://www.avvo.com/health-guides/qsymia--phenteramine-topiramate-for-weight-reduction
57 **FDA**. FDA approves weight-management drug Qsymia (2012 July 2012). Available from: URL: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm312468.htm
58 **Allison DB**, Gadde KM, Garvey WT, Peterson CA, Schwiers ML, Najarian T, Tam PY, Troupin B, Day WW. Controlled-release phentermine/topiramate in severely obese adults: a randomized controlled trial (EQUIP). *Obesity* (Silver Spring) 2012; **20**: 330-342 [PMID: 22051941]
59 **Gadde KM**, Allison DB, Ryan DH, Peterson CA, Troupin B, Schwiers ML, Day WW. Effects of low-dose, controlled-release, phentermine plus topiramate combination on weight and associated comorbidities in overweight and obese adults (CONQUER): a randomised, placebo-controlled, phase 3 trial. *Lancet* 2011; **377**: 1341-1352 [PMID: 21481449]
60 **Ornellas T**, Chavez B. Naltrexone SR/Bupropion SR (Contrave): A New Approach to Weight Loss in Obese Adults. *P T* 2011; **36**: 255-262 [PMID: 21785538]
61 **Cone RD**, Cowley MA, Butler AA, Fan W, Marks DL, Low MJ. The arcuate nucleus as a conduit for diverse signals relevant to energy homeostasis. *Int J Obes Relat Metab Disord* 2001; **25** Suppl 5: S63-S67 [PMID: 11840218 DOI: 10.1038/sj.ijo.0801913]
62 **Cowley MA**, Smart JL, Rubinstein M, Cerdán MG, Diano S, Horvath TL, Cone RD, Low MJ. Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature* 2001; **411**: 480-484 [PMID: 11373681 DOI: 10.1038/35078085]
63 **Saper CB**, Chou TC, Elmquist JK. The need to feed: homeostatic and hedonic control of eating. *Neuron* 2002; **36**: 199-211 [PMID: 12383777]
64 Orexigen. FDA Issues Complete Response to New Drug Application for Contrave(R) for the Management of Obesity (2011 July 2012). Available from: URL: http://ir.orexigen.com/phoenix.zhtml?c=207034&p=irol-newsArticle&ID=1522207&highlight
65 **Pelchat ML**. Food cravings in young and elderly adults. *Appetite* 1997; **28**: 103-113 [PMID: 9158846 DOI: 10.1006/appe.1996.0063]
66 **Wadden TA**, Foreyt JP, Foster GD, Hill JO, Klein S, O'Neil PM, Perri MG, Pi-Sunyer FX, Rock CL, Erickson JS, Maier HN, Kim DD, Dunayevich E. Weight loss with naltrexone SR/bupropion SR combination therapy as an adjunct to behavior modification: the COR-BMOD trial. *Obesity* (Silver Spring) 2011; **19**: 110-120 [PMID: 20559296]
67 **Gadde KM**, Yonish GM, Foust MS, Wagner HR. Combination therapy of zonisamide and bupropion for weight reduction in obese women: a preliminary, randomized, open-label study. *J Clin Psychiatry* 2007; **68**: 1226-1229 [PMID: 17854247]
68 Orexigen. Product candidates - Empatic™ 2012 August 2012. Available from: URL: http://www.orexigen.com/candidates/candidates_empatic.php
69 **Greenway F**, Atkinson JAR. Bupropion and zonisamide for the treatment of obesity [abstract 52-OR]. Obesity Research, 2006; 14: A17
70 **Orexigen.** Clinical Trials. 2012 [cited Accessed August 2012]. Available from: URL: http://www.orexigen.com/trials/default.php
71 **Faria AM**, Mancini MC, Melo ME, Cercato C, Halpern A. Recent progress and novel perspectives on obesity pharmacotherapy. *Arq Bras Endocrinol Metabol* 2010; **54**: 516-529 [PMID:

20857056]
72 **Klonoff DC**, Greenway F. Drugs in the pipeline for the obesity market. *J Diabetes Sci Technol* 2008; **2**: 913-918 [PMID: 19885278]
73 **Fujioka K**, Apovian C, Hill J. The Evolution of Obesity Therapies: New Applications for Existing Drugs (MedscapeCME Diabetes Endocrinology, 2010). Available from: URL: http://cme.medscape.com/viewarticle/722366
74 **Astrup A**, Meier DH, Mikkelsen BO, Villumsen JS, Larsen TM. Weight loss produced by tesofensine in patients with Parkinson's or Alzheimer's disease. *Obesity* (Silver Spring) 2008; **16**: 1363-1369 [PMID: 18356831 DOI: 10.1038/oby.2008.56]
75 **Gilbert JA**, Gasteyger C, Raben A, Meier DH, Astrup A, Sjödin A. The effect of tesofensine on appetite sensations. *Obesity* (Silver Spring) 2012; **20**: 553-561 [PMID: 21720440 DOI: 10.1038/oby.2011.197]
76 NeuroSearch. Tesofensine - Mode of Action. August 2012]. Available from: URL: http://neurosearch.com/Default.aspx?ID=8174
77 **Sjödin A**, Gasteyger C, Nielsen AL, Raben A, Mikkelsen JD, Jensen JK, Meier D, Astrup A. The effect of the triple monoamine reuptake inhibitor tesofensine on energy metabolism and appetite in overweight and moderately obese men. *Int J Obes* (Lond) 2010; **34**: 1634-1643 [PMID: 20479765 DOI: 10.1038/ijo.2010.87]
78 **Bello NT**, Zahner MR. Tesofensine, a monoamine reuptake inhibitor for the treatment of obesity. *Curr Opin Investig Drugs* 2009; **10**: 1105-1116 [PMID: 19777399]
79 **Chaudhri O**, Small C, Bloom S. Gastrointestinal hormones regulating appetite. *Philos Trans R Soc Lond B Biol Sci* 2006; **361**: 1187-1209 [PMID: 16815798 DOI: 10.1098/rstb.2006.1856]
80 **Holst JJ**. The physiology of glucagon-like peptide 1. *Physiol Rev* 2007; **87**: 1409-1439 [PMID: 17928588 DOI: 10.1152/physrev.00034.2006]
81 **Pilichiewicz AN**, Chaikomin R, Brennan IM, Wishart JM, Rayner CK, Jones KL, Smout AJ, Horowitz M, Feinle-Bisset C. Load-dependent effects of duodenal glucose on glycemia, gastrointestinal hormones, antropyloroduodenal motility, and energy intake in healthy men. *Am J Physiol Endocrinol Metab* 2007; **293**: E743-E753 [PMID: 17609258 DOI: 10.1152/ajpendo.00159.2007]
82 **Vilsbøll T**, Garber AJ. Non-glycaemic effects mediated via GLP-1 receptor agonists and the potential for exploiting these for therapeutic benefit: focus on liraglutide. *Diabetes Obes Metab* 2012; **14** Suppl 2: 41-49 [PMID: 22405268 DOI: 10.1111/j.1463-1326.2012.01579.x]
83 **Zinman B**, Gerich J, Buse JB, Lewin A, Schwartz S, Raskin P, Hale PM, Zdravkovic M, Blonde L. Efficacy and safety of the human glucagon-like peptide-1 analog liraglutide in combination with metformin and thiazolidinedione in patients with type 2 diabetes (LEAD-4 Met+TZD). *Diabetes Care* 2009; **32**: 1224-1230 [PMID: 19289857 DOI: 10.2337/dc08-2124]
84 **FDA**. Nda 22-341 victoza risk evaluation and mitigation strategy (REMS). 2012 [cited Accessed July 2012]. Available from: URL: http://www.fda.gov/downloads/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/UCM202063.pdf
85 **FDA**. Highlights of prescribing information - Victoza (2012 July 2012). Available from: URL: http://www.accessdata.fda.gov/drugsatfda_docs/label/2012/022341s007s009s013lbl.pdf
86 **FDA**. Exenatide (marketed as Byetta) Information (2009 August 2012). Available from: URL: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm113705.htm
87 **EMA**. Byetta exenatide. 2012 August 2012. Available from: URL: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000698/human_med_000682.jsp&jsenabled=true
88 **Rosenstock J**, Klaff LJ, Schwartz S, Northrup J, Holcombe JH, Wilhelm K, Trautmann M. Effects of exenatide and lifestyle modification on body weight and glucose tolerance in obese subjects with and without pre-diabetes. *Diabetes Care* 2010; **33**: 1173-1175 [PMID: 20332357]
89 **Bradley DP**, Kulstad R, Racine N, Shenker Y, Meredith M, Schoeller DA. Alterations in energy balance following exenatide administration. *Appl Physiol Nutr Metab* 2012; **37**: 893-899 [PMID: 22735035 DOI: 10.1139/h2012-068]
90 **Wynne K**, Bloom S. Oxyntomodulin: a new therapy for obesity? *Science in School* 2007; **(6)**: 25-29
91 **Wynne K**, Park AJ, Small CJ, Meeran K, Ghatei MA, Frost GS, Bloom SR. Oxyntomodulin increases energy expenditure in addition to decreasing energy intake in overweight and obese humans: a randomised controlled trial. *Int J Obes* (Lond) 2006; **30**: 1729-1736 [PMID: 16619056 DOI: 10.1038/sj.ijo.0803344]
92 **Field BC**, Wren AM, Peters V, Baynes KC, Martin NM, Patterson M, Alsaraf S, Amber V, Wynne K, Ghatei MA, Bloom SR. PYY3-36 and oxyntomodulin can be additive in their effect on food intake in overweight and obese humans. *Diabetes* 2010; **59**: 1635-1639 [PMID: 20357366 DOI: 10.2337/db09-1859]
93 **Cummings DE**, Overduin J. Gastrointestinal regulation of food intake. *J Clin Invest* 2007; **117**: 13-23 [PMID: 17200702]
94 **Woods SC**, Lutz TA, Geary N, Langhans W. Pancreatic signals controlling food intake; insulin, glucagon and amylin. *Philos Trans R Soc Lond B Biol Sci* 2006; **361**: 1219-1235 [PMID: 16815800 DOI: 10.1098/rstb.2006.1858]
95 **Weyer C**, Maggs DG, Young AA, Kolterman OG. Amylin replacement with pramlintide as an adjunct to insulin therapy in type 1 and type 2 diabetes mellitus: a physiological approach toward improved metabolic control. *Curr Pharm Des* 2001; **7**: 1353-1373 [PMID: 11472273]
96 **Goldsbury C**, Goldie K, Pellaud J, Seelig J, Frey P, Müller SA, Kistler J, Cooper GJ, Aebi U. Amyloid fibril formation from full-length and fragments of amylin. *J Struct Biol* 2000; **130**: 352-362 [PMID: 10940238 DOI: 10.1006/jsbi.2000.4268]
97 **Higham CE**, Jaikaran ET, Fraser PE, Gross M, Clark A. Preparation of synthetic human islet amyloid polypeptide (IAPP) in a stable conformation to enable study of conversion to amyloid-like fibrils. *FEBS Lett* 2000; **470**: 55-60 [PMID: 10722845]
98 **Janes S**, GL Beaumont K, Beeley N, Rink T. The selection of pramlintide for clinical evaluation. *Diabetes* 1996; **45** (Suppl 2): p. 235A (Abstract)
99 **Lutz TA**. Amylinergic control of food intake. *Physiol Behav* 2006; **89**: 465-471 [PMID: 16697020 DOI: 10.1016/j.physbeh.2006.04.001]
100 **Aronne L**, Fujioka K, Aroda V, Chen K, Halseth A, Kesty NC, Burns C, Lush CW, Weyer C. Progressive reduction in body weight after treatment with the amylin analog pramlintide in obese subjects: a phase 2, randomized, placebo-controlled, dose-escalation study. *J Clin Endocrinol Metab* 2007; **92**: 2977-2983 [PMID: 17504894 DOI: 10.1210/jc.2006-2003]
101 **Aronne LJ**, Halseth AE, Burns CM, Miller S, Shen LZ. Enhanced weight loss following coadministration of pramlintide with sibutramine or phentermine in a multicenter trial. *Obesity* (Silver Spring) 2010; **18**: 1739-1746 [PMID: 20094043 DOI: 10.1038/oby.2009.478]
102 **Smith SR**, Blundell JE, Burns C, Ellero C, Schroeder BE, Kesty NC, Chen KS, Halseth AE, Lush CW, Weyer C. Pramlintide treatment reduces 24-h caloric intake and meal sizes and improves control of eating in obese subjects: a 6-wk translational research study. *Am J Physiol Endocrinol Metab* 2007; **293**: E620-E627 [PMID: 17505051 DOI: 10.1152/ajpendo.00217.2007]
103 **Wilson GT**, Wilfley DE, Agras WS, Bryson SW. Psychological treatments of binge eating disorder. *Arch Gen Psychiatry* 2010; **67**: 94-101 [PMID: 20048227 DOI: 10.1001/archgenpsy-

chiatry.2009.170]
104 **Friedman JM**. Leptin at 14 y of age: an ongoing story. *Am J Clin Nutr* 2009; **89**: 973S-979S [PMID: 19190071 DOI: 10.3945/ajcn.2008.26788B]
105 **Farooqi IS**, O'Rahilly S. Leptin: a pivotal regulator of human energy homeostasis. *Am J Clin Nutr* 2009; **89**: 980S-984S [PMID: 19211814 DOI: 10.3945/ajcn.2008.26788C]
106 **Oral EA**, Simha V, Ruiz E, Andewelt A, Premkumar A, Snell P, Wagner AJ, DePaoli AM, Reitman ML, Taylor SI, Gorden P, Garg A. Leptin-replacement therapy for lipodystrophy. *N Engl J Med* 2002; **346**: 570-578 [PMID: 11856796 DOI: 10.1056/NEJMoa012437]
107 **Lee JH**, Chan JL, Sourlas E, Raptopoulos V, Mantzoros CS. Recombinant methionyl human leptin therapy in replacement doses improves insulin resistance and metabolic profile in patients with lipoatrophy and metabolic syndrome induced by the highly active antiretroviral therapy. *J Clin Endocrinol Metab* 2006; **91**: 2605-2611 [PMID: 16636130 DOI: 10.1210/jc.2005-1545]
108 **Ravussin E**, Smith SR, Mitchell JA, Shringarpure R, Shan K, Maier H, Koda JE, Weyer C. Enhanced weight loss with pramlintide/metreleptin: an integrated neurohormonal approach to obesity pharmacotherapy. *Obesity* (Silver Spring) 2009; **17**: 1736-1743 [PMID: 19521351 DOI: 10.1038/oby.2009.184]
109 **August GP**, Caprio S, Fennoy I, Freemark M, Kaufman FR, Lustig RH, Silverstein JH, Speiser PW, Styne DM, Montori VM. Prevention and treatment of pediatric obesity: an endocrine society clinical practice guideline based on expert opinion. *J Clin Endocrinol Metab* 2008; **93**: 4576-4599 [PMID: 18782869 DOI: 10.1210/jc.2007-2458]
110 **DeFronzo RA**, Barzilai N, Simonson DC. Mechanism of metformin action in obese and lean noninsulin-dependent diabetic subjects. *J Clin Endocrinol Metab* 1991; **73**: 1294-1301 [PMID: 1955512]
111 **Cigolini M**, Bosello O, Zancanaro C, Orlandi PG, Fezzi O, Smith U. Influence of metformin on metabolic effect of insulin in human adipose tissue in vitro. *Diabete Metab* 1984; **10**: 311-315 [PMID: 6397366]
112 **Lee A**, Morley JE. Metformin decreases food consumption and induces weight loss in subjects with obesity with type II non-insulin-dependent diabetes. *Obes Res* 1998; **6**: 47-53 [PMID: 9526970]
113 **Munro JF**, MacCuish AC, Marshall A, Wilson EM, Duncan LJ. Weight-reducing effect of diguanides in obese non-diabetic women. *Br Med J* 1969; **2**: 13-15 [PMID: 4887385]
114 Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). UK Prospective Diabetes Study (UKPDS) Group. *Lancet* 1998; **352**: 854-865 [PMID: 9742977]
115 **Yanovski JA**, Krakoff J, Salaita CG, McDuffie JR, Kozlosky M, Sebring NG, Reynolds JC, Brady SM, Calis KA. Effects of metformin on body weight and body composition in obese insulin-resistant children: a randomized clinical trial. *Diabetes* 2011; **60**: 477-85 [DOI: 10.2337/db10-1185]
116 **Morrison JA**, Cottingham EM, Barton BA. Metformin for weight loss in pediatric patients taking psychotropic drugs. *Am J Psychiatry* 2002; **159**: 655-657 [PMID: 11925306]
117 **Ghandi S**, Aflatoonian A, Tabibnejad N, Moghaddam MH. The effects of metformin or orlistat on obese women with polycystic ovary syndrome: a prospective randomized open-label study. *J Assist Reprod Genet* 2011; **28**: 591-596 [PMID: 21484319 DOI: 10.1007/s10815-011-9564-2]
118 **Wiegand S**, l'Allemand D, Hübel H, Krude H, Bürmann M, Martus P, Grüters A, Holl RW. Metformin and placebo therapy both improve weight management and fasting insulin in obese insulin-resistant adolescents: a prospective, placebo-controlled, randomized study. *Eur J Endocrinol* 2010; **163**: 585-592 [PMID: 20639355 DOI: 10.1530/EJE-10-0570]
119 **Arima H**, Oiso Y. Positive effect of baclofen on body weight reduction in obese subjects: a pilot study. *Intern Med* 2010; **49**: 2043-2047 [PMID: 20930428 DOI: 10.2169/internalmedicine.49.3918]
120 **Tomita M**, Bellamy W, Takase M, Yamauchi K, Wakabayashi H, Kawase K. Potent antibacterial peptides generated by pepsin digestion of bovine lactoferrin. *J Dairy Sci* 1991; **74**: 4137-4142 [PMID: 1787185 DOI: 10.3168/jds.S0022-0302(91)78608-6]
121 **Harmsen MC**, Swart PJ, de Béthune MP, Pauwels R, De Clercq E, The TH, Meijer DK. Antiviral effects of plasma and milk proteins: lactoferrin shows potent activity against both human immunodeficiency virus and human cytomegalovirus replication in vitro. *J Infect Dis* 1995; **172**: 380-388 [PMID: 7622881]
122 **Zimecki M**, Właszczyk A, Cheneau P, Brunel AS, Mazurier J, Spik G, Kübler A. Immunoregulatory effects of a nutritional preparation containing bovine lactoferrin taken orally by healthy individuals. *Arch Immunol Ther Exp* (Warsz) 1998; **46**: 231-240 [PMID: 9779289]
123 **Shoji H**, Oguchi S, Shinohara K, Shimizu T, Yamashiro Y. Effects of iron-unsaturated human lactoferrin on hydrogen peroxide-induced oxidative damage in intestinal epithelial cells. *Pediatr Res* 2007; **61**: 89-92 [PMID: 17211147 DOI: 10.1203/01.pdr.0000250198.22735.20]
124 **Ono T**, Murakoshi M, Suzuki N, Iida N, Ohdera M, Iigo M, Yoshida T, Sugiyama K, Nishino H. Potent anti-obesity effect of enteric-coated lactoferrin: decrease in visceral fat accumulation in Japanese men and women with abdominal obesity after 8-week administration of enteric-coated lactoferrin tablets. *Br J Nutr* 2010; **104**: 1688-1695 [PMID: 20691130 DOI: 10.1017/S0007114510002734]
125 **James WP**, Caterson ID, Coutinho W, Finer N, Van Gaal LF, Maggioni AP, Torp-Pedersen C, Sharma AM, Shepherd GM, Rode RA, Renz CL. Effect of sibutramine on cardiovascular outcomes in overweight and obese subjects. *N Engl J Med* 2010; **363**: 905-917 [PMID: 20818901 DOI: 10.1056/NEJMoa1003114]
126 **Torp-Pedersen C**, Caterson I, Coutinho W, Finer N, Van Gaal L, Maggioni A, Sharma A, Brisco W, Deaton R, Shepherd G, James P. Cardiovascular responses to weight management and sibutramine in high-risk subjects: an analysis from the SCOUT trial. *Eur Heart J* 2007; **28**: 2915-2923 [PMID: 17595194]
127 **Liu YL**, Connoley IP, Wilson CA, Stock MJ. Effects of the cannabinoid CB1 receptor antagonist SR141716 on oxygen consumption and soleus muscle glucose uptake in Lep(ob)/Lep(ob) mice. *Int J Obes* (Lond) 2005; **29**: 183-187 [PMID: 15558076 DOI: 10.1038/sj.ijo.0802847]
128 **Bensaid M**, Gary-Bobo M, Esclangon A, Maffrand JP, Le Fur G, Oury-Donat F, Soubrié P. The cannabinoid CB1 receptor antagonist SR141716 increases Acrp30 mRNA expression in adipose tissue of obese fa/fa rats and in cultured adipocyte cells. *Mol Pharmacol* 2003; **63**: 908-914 [PMID: 12644592]
129 **Osei-Hyiaman D**, Liu J, Zhou L, Godlewski G, Harvey-White J, Jeong WI, Bátkai S, Marsicano G, Lutz B, Buettner C, Kunos G. Hepatic CB1 receptor is required for development of diet-induced steatosis, dyslipidemia, and insulin and leptin resistance in mice. *J Clin Invest* 2008; **118**: 3160-3169 [PMID: 18677409 DOI: 10.1172/JCI34827]
130 **Pagotto U**, Pasquali R. Fighting obesity and associated risk factors by antagonising cannabinoid type 1 receptors. *Lancet* 2005; **365**: 1363-1364 [PMID: 15836868 DOI: 10.1016/S0140-6736(05)66348-9]
131 **FDA**. What is rimonabant and what are the associated risks? 2008 [July 2012]. Available from: URL: http://www.fda.gov/Drugs/ResourcesForYou/Consumers/QuestionsAnswers/ucm136187.htm
132 **FDA**. FDA Briefing Document NDA 21-888 Zimulti (rimonabant) Tablets, 20 mg. [2007 July 2012]. Available from: URL: http://www.fda.gov/ohrms/dockets/ac/07/briefing/2007-4306b1-fda-

backgrounder.pdf
133 **O'Leary DH**, Reuwer AQ, Nissen SE, Després JP, Deanfield JE, Brown MW, Zhou R, Zabbatino SM, Job B, Kastelein JJ, Visseren FL. Effect of rimonabant on carotid intima-media thickness (CIMT) progression in patients with abdominal obesity and metabolic syndrome: the AUDITOR Trial. *Heart* 2011; **97**: 1143-1150 [PMID: 21610270]
134 **Motaghedi R**, Lipman EG, Hogg JE, Christos PJ, Vogiatzi MG, Angulo MA. Psychiatric adverse effects of rimonobant in adults with Prader Willi syndrome. *Eur J Med Genet* 2011; **54**: 14-18 [PMID: 20965292]
135 **Horder J**, Harmer CJ, Cowen PJ, McCabe C. Reduced neural response to reward following 7 days treatment with the cannabinoid CB1 antagonist rimonabant in healthy volunteers. *Int J Neuropsychopharmacol* 2010; **13**: 1103-1113 [PMID: 20426883]

**P- Reviewer:** Gómez-Sáez J, Masaki T, Papazoglou D
**S- Editor:** Song XX **L- Editor:** A **E- Editor:** Wu HL

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.154

*World J Meta-Anal* 2014 November 26; 2(4): 154-161
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.



# Russian clinical research policy does not guarantee results availability

Tatyana Arstakovna Shamliyan, Anna Artemovna Avanesova

Tatyana Arstakovna Shamliyan, Evidence-Based Medicine Center, Clinical Solutions, Philadelphia, PA 19103, United States
Anna Artemovna Avanesova, Stavropol State University, Stavropol 355009, Russian Federation
Author contributions: Shamliyan TA and Avanesova AA solely contributed to this paper.
Correspondence to: Tatyana Arstakovna Shamliyan, MD, MS, Senior Director, Quality Assurance, Evidence-Based Medicine Center, Clinical Solutions, Elsevier, 1600 JFK Blvd, Philadelphia, PA 9103, United States. t.shamliyan@elsevier.com
Telephone: +1-215-2393821-3821 Fax: +1-215-2396114
Received: May 14, 2014 Revised: August 11, 2014
Accepted: September 4, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To investigate results availability from clinical studies enrolling Russian subjects and Russian clinical research policy.

**METHODS:** We analyzed Russian legislation and ethical regulations about drug and devices approval, clinical research registration and the results availability. In August 2012, we searched International Clinical Trials Registry Platform and clinicaltrials.gov to find all registered studies that had an investigational site in the territory of the Russian Federation. To find publication status, we searched the PubMed and Scirus bibliographical databases with trial registration number to find journal publications of the registered studies.

**RESULTS:** We identified 2062 registered research protocols comprising 2017 international and 45 protocols sponsored by the Russian funding agencies. The number of the studies enrolling Russian subjects increased dramatically from three studies in 2002 to 252 studies in 2012. Most studies (92%) were funded exclusively by industry, were interventions (94.6%), examined drugs (87%) and enrolled exclusively adults (86%) of both genders (89%). Only 383 (19%) of multinational studies and two (4.4%) of exclusively Russian studies were published. Posting of patient outcomes was available for 16% of the trials that recruited trial participants in the Russian territory including one study funded exclusively by Russian sponsors. Investigators terminated 99 studies of 38111 participants and did not provide the results in clinicaltrials.gov or in published manuscripts. Federal laws require clinical study registration and conflict of interest disclosure. However, routine monitoring of compliance to clinical research policy is not available.

**CONCLUSION:** Russian legislation does not guarantee the availability of clinical research results. Russian legislation should mandate transparent evidence- based market approval of the drugs and devices.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Clinical Research; Meta-analysis; Research standards; Publication bias; Medicine; Legal

**Core tip:** We identified clinical studies that enrolled Russian subjects and found very low rate of the publication of the results in peer reviewed clinical journals or posting of the results in trial registry, clinicaltrials.gov. We concluded that Russian legislation does not guarantee the availability of clinical research results. The Russian legislation should be revised to mandate transparent evidence- based market approval of the drugs and devices based on high quality clinical evidence applicable to the Russian population.

Shamliyan TA, Avanesova AA. Russian clinical research policy does not guarantee results availability. *World J Meta-Anal* 2014; 2(4): 154-161 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/154.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.154

## INTRODUCTION

Scientific research improves global health care best when

international legal and ethical regulations guarantee stakeholders access to the complete and unbiased information it generates[1-4]. Similarly, complete and readily available information about the results from international studies is required for comprehensive syntheses of evidence[5]. Voluntary publication of only selected studies that show impressive results threatens the validity of research. Mandatory study registration and posting the results in trial registry as well as free public access to the publications[6-9] help address this problem[10-12].

Several American initiatives have contributed to transparent public control of clinical studies. First, the United States Congress mandated the registration in ClinicalTrials.gov of all drug trials "for serious or life-threatening diseases"[13-16]. Second, the Food and Drug Administration Amendments Act (FDAAA) mandated posting of applicable trial results on ClinicalTrials.gov within 12 mo of trial's completion[17]. FDAAA applicable trials include all interventions regulated by the FDA and phase Ⅱ to phase Ⅳ recruiting subjects in the United States[18-21].

Finally, the International Committee of Medical Journal Editors and the World Association of Medical Editors required a registration status for all submitted manuscripts reporting the results of clinical trials[22]. Since then many countries launched the national trial registries reporting the minimum protocol information defined by the 20-item World Health Organization Registration Data Set[23]. The World Health Organization International Clinical Trials Registry Platform (ICTRP) permits a single search across all national trial registries in nine languages[23].

The Russian legislation adopted the international standard in biomedical research and committed itself to meet the highest standards and the integrity of clinical research[24-26]. The actual legislation in relation to the trends in clinical research involving Russian subjects as well as study sponsorships, types, and results availability has not been examined yet.

A contract research organization (http://www.synrg-pharm.com) and the Association of Clinical Trials Organizations, a non-commercial organization of the companies that engage clinical trials (http://acto-russia.org) provide inconsistent information about ongoing and completed clinical studies in the territory of the Russian Federation. Our objectives were to review Russian legal and ethical regulations of clinical research and to examine from the reliable sources the trends in conducting and availability of the results from the clinical studies that enrolled Russian subjects.

## MATERIALS AND METHODS

We analyzed Russian legislation and ethical regulations about drug and devices approval, clinical research registration and the results availability.

We searched the ICTRP and clinicaltrials.gov to find all registered studies that had an investigational site in the territory of the Russian Federation. We retrieved all studies from trial registries in August 2012. We retrieved all available fields as defined at http://prsinfo.clinicaltrials.gov/definitions.html and reported by principal investigators. We further categorized interventions according to classification from clinicaltrials.gov. We marked and separately analyzed a subgroup of the studies that were sponsored exclusively by the Russian sponsors. We defined studies as randomized controlled clinical trials if the field with study designs described random allocation of the subjects into the treatment groups. We defined studies as pharmacokinetics, efficacy, safety, or efficacy and safety studies per investigators definitions of study goals. We analyzed the target enrollment in recruiting and actual number of the enrolled subjects in completed or terminated studies per the numbers provided by the principal investigators.

To find publication status, we searched the PubMed and Scirus bibliographical databases with trial registration number to find journal publications of the registered studies. Two researchers rechecked the publication status. We did not analyze the actual results of the studies reported in published articles. We did not abstract any data from the published articles, did not analyze reporting bias, and did not construct funnel plots to quantify publication bias.

We calculated descriptive frequency statistics and compared results availability (in clinical.trials.gov or journal papers) by subject demographics, conditions, examined treatments, study design, funding, and recruitment status. We calculated odds ratios with 95%CI for posting of the results and publication using logistic regression with STATA or SAS 9.1 software (SAS Institute Inc., of Cary, North Carolina, United States).

## RESULTS

Rapidly developing clinical research legislation in the Russian Federation consists of several federal laws[25-27] and Government Decrees[28-32] adopted during 2010-2012. Several regulatory documents address ethical approval of the clinical research involving human subjects according to the international standards including informed consent procedure[33-37]. The most recent decree from the Ministry of Health and Social Development at December 30, 2011 "On approval of the technical tests, toxicology studies, and clinical trials of medical devices in order of their registration" outlined requirements for the preclinical trials (Phase 0-2) emphasize examination of the safety and tolerance with investigative drugs[38]. Amendments B and G recommend protocols and tabulate safety assessment methods as a basis for design and execution of the clinical trials. The Ministry of Health and Social Development is responsible for the control over the clinical research and oversee clinical studies design and execution[35,39].

Clinical trials in the Russian Federation can be conducted by the foreign medical and pharmacist professionals after accreditation and certification according to the article 100 of the Federal Law N 323[26]. The sponsors must request the trial approval from the Council on Ethics of Ministry of Health and Social Development[34-37].

**Table 1 Posting the results in clinicaltrials.gov**

| Study characteristic | Has results | No results available | Total | Percent with the results |
|---|---|---|---|---|
| Subject age | | | | |
| Child | 21 | 78 | 99 | 21.21 |
| Child/adult | 6 | 37 | 43 | 13.95 |
| Child/adult/senior | 28 | 117 | 145 | 19.31 |
| Adult | 17 | 192 | 209 | 8.13 |
| Adult/senior | 260 | 1301 | 1561 | 16.66 |
| Senior | 3 | 2 | 5 | 60.00 |
| Subject sex | | | | |
| Both | 293 | 1536 | 1829 | 16.02 |
| Female | 22 | 131 | 153 | 14.38 |
| Male | 20 | 60 | 80 | 25.00 |
| Total | 335 | 1727 | 2062 | 16.25 |
| Sponsorship | | | | |
| Industry | 322 | 1566 | 1888 | 17.06 |
| Industry combined with non-industry sources | 12 | 63 | 75 | 16.00 |
| Non industry sponsors | 1 | 98 | 99 | 1.01 |
| Examined Interventions | | | | |
| Behavioral | 0 | 9 | 9 | 0.00 |
| Biological | 14 | 115 | 129 | 10.85 |
| Device | 0 | 17 | 17 | 0.00 |
| Dietary supplement | 1 | 3 | 4 | 25 |
| Drug | 315 | 1468 | 1783 | 17.67 |
| Genetic | 0 | 1 | 1 | 0.00 |
| Other | 0 | 22 | 22 | 0.00 |
| Procedure | 0 | 31 | 31 | 0.00 |
| Radiation | 0 | 1 | 1 | 0.00 |
| Phases of clinical trials | | | | |
| Phase 1 | 1 | 42 | 43 | 2.33 |
| Phase 1/phase 2 | 3 | 31 | 34 | 8.82 |
| Phase 2 | 58 | 429 | 487 | 11.91 |
| Phase 2/phase 3 | 8 | 43 | 51 | 15.69 |
| Phase 3 | 221 | 936 | 1157 | 19.10 |
| Phase 4 | 38 | 133 | 171 | 22.22 |
| Recruitment | | | | |
| Active, not recruiting | 25 | 339 | 364 | 6.87 |
| Approved for marketing | 0 | 2 | 2 | 0.00 |
| Available | 0 | 3 | 3 | 0.00 |
| Completed | 288 | 744 | 1032 | 27.91 |
| Enrolling by invitation | 0 | 20 | 20 | 0.00 |
| No longer available | 0 | 1 | 1 | 0.00 |
| Not yet recruiting | 0 | 25 | 25 | 0.00 |
| Recruiting | 0 | 481 | 481 | 0.00 |
| Suspended | 0 | 3 | 3 | 0.00 |
| Temporarily not available | 0 | 1 | 1 | 0.00 |
| Terminated | 22 | 100 | 122 | 18.03 |
| Interventional | 327 | 1624 | 1951 | 16.76 |
| Observational | 8 | 96 | 104 | 7.69 |

The principal investigators certified to conduct clinical research in Russia must request the trial approval from the local research ethics committee and obtain informed consent before enrollment of the subjects. Russian law mandates trial sponsors to provide life and health insurance to all enrolled human subjects[32]. The current legislation does not well articulate the transparent process of medical and statistical review of the new medicine and devices as well as availability of the results from clinical studies for the public.

The Director of the State Department for the regulation of drugs Marat Sakaev stated in the International Conference "Ethics Review of Clinical Research in Pharmaceuticals" that 677 accredited medical organizations could conduct clinical research in Russia (http://www.coe.int/t/dg3/healthbioethic/conferences_and_symposia/Speaker_comp_enx.pdf).

He emphasized the importance of further improvement in the legislation and ethical regulations of the clinical research: "Despite the fact that in Russia, clinical studies have not as common as in the United States and the European Union, every year more and more Russian citizens to receive and accept the invitation to participate in clinical trials. In this regard, the formation of a clear system of state regulation and control in this area, not only to protect the rights, safety and health of patients participating in clinical studies, but the guarantee of the reliability of information on the safety and efficacy of drugs, obtained in clinical trials".

We found 2693 registered protocols (2062 protocols in clinicaltrials.gov) including 2017 international studies recruiting Russian subjects and 45 studies sponsored exclusively by the Russian funding agencies. The number of the studies enrolling Russian subjects increased dramatically from three studies in 2002 to 252 studies in 2012 (Figure 1). The studies enrolled 2409570 subjects. Most studies (92%) were funded exclusively by industry, were interventions (94.6%), examined drugs (87%) and enrolled exclusively adults (86%) of both genders (89%). Half of the studies were completed while 6.5% were initiated but not completed. Few studies provided clear reasons for termination.

The studies funded exclusively by the Russian sponsors more often employed interventional design (84%) with random allocation of the subjects to the treatment groups (73%), and aimed to examine efficacy (16%) or efficacy and safety (53%) with the treatments. Eleven studies did not provide goals of the studies. Most studies were phase 3 (42%) or phase four (24%) clinical trials. The trials funded by the Russian agencies recruited 43163 subjects.

### *Results availability in trial registries*

Among all registered presented in the ICTRP, only clinicaltrial.gov provides postings of baseline participant's characteristics, flow, and post-treatment outcomes. The results were available only for 16% of the studies that recruited Russian participants including one study sponsored by Russian funding agency (Table 1). Odds of posting the results varied among individual sponsors and by study characteristics (Table 1). Industry sponsors posted the results more often than not for profit sponsors (OR = 20, 95%CI: 3-125) (Figure 2). Completed studies had posted results more often than initiated but not completed studies (OR = 1.8, 95%CI: 1.1-2.9). Interventional and drug studies posted the results more often than observational studies or studies of non-pharmacological interventions (Figure 2). The posted post-intervention outcomes were available only for 524214 (22%) enrolled participants.

### *Results availability in peer reviewed journals indexed in Medline*

The results were published only for 383 (19%) of in-

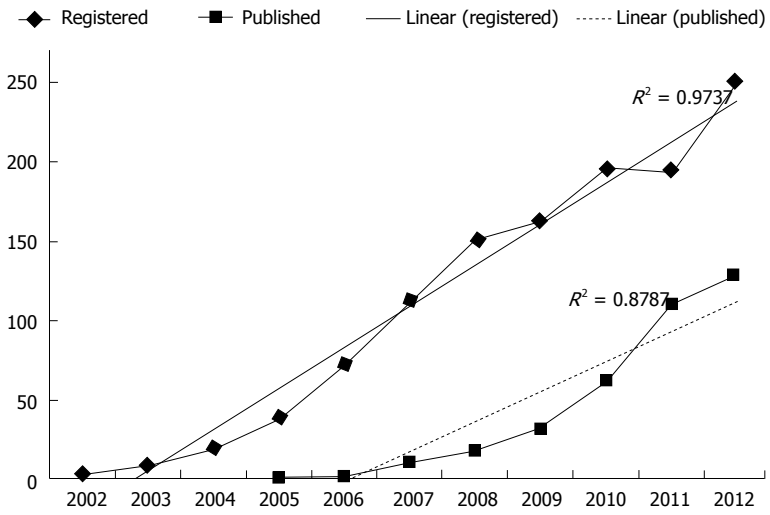



**Figure 1 Registration and publication of the clinical studies enrolling subjects in the Russian Federation.** Horizontal axis- years of the registration or publication; Vertical axis - the number of the studies.

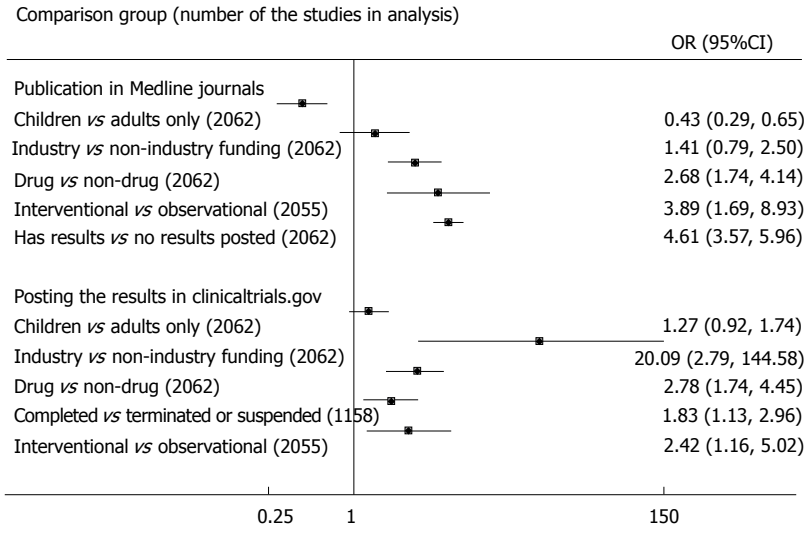



**Figure 2 Odds of results availability as posted in clinicaltrials.gov or published in journals among study characteristics.**

ternational studies and for two (4.4%) of the studies funding by the Russian agencies (Table 2). Journal manuscripts published the outcomes only for 902735 (38%) of the recruited participants (Table 3). Principal investigators terminated 99 studies of 38111 participants and did not provide the results in clinicaltrials.gov or in published manuscripts (Table 3). Principal investigators that posted the results in trial registry also published the results in peer reviewed journals (OR of journal publication of the studies with *vs* without results in clinicaltrials.gov 4.6, 95%CI: 3.6-5.9) (Figure 2). Pediatric studies were published less often than studies recruiting adults (OR = 0.43, 95%CI: 0.29-0.65) (Figure 2). Studies of pharmacological interventions (OR = 2.7, 95%CI: 1.7-4) published the results more often than studies examining non-drug treatments. Principal investigators conducting clinical trials published the results more often than principal investigators conducting observational studies (OR = 3.9, 95%CI: 1.7-8.9).

Principal investigators published manuscripts in 158 peer-reviewed journals available *via* PubMed search. The four journals including Lancet, New England Journal of Medicine, Current Medical Research and Opinions, and the Journal of Clinical Psychiatry published 24% of all studies enrolling Russian subjects. Only seven articles were led by Russian authors. Medline search found 821 published during last 10 years randomized controlled clinical trials with no registration information.

## DISCUSSION

Our study found dramatically increasing number of registered studies enrolling Russian subjects. However, patient outcomes were not available from the majority of the registered studies that recruited Russian participants. Many publications of the Russian randomized trials do not mention registration.

Our study demonstrated that existing international clinical research policy and Russian research regulations do not guarantee availability of the results from human studies. Comprehensive analyses of evidence are hindered by the substantial likelihood of publication bias. Publication of the results in peer-reviewed journals remains voluntary[6-9]. In contrast, posting of the results in Clinicaltrials.gov[40] provides public access to the subject flow, benefits with the treatments, and all adverse effects which

**Table 2 Publication of the registered studies that enrolled Russian subjects**

| Study characteristics | Published | Unpublished | Total | Percent published |
|---|---|---|---|---|
| Age | | | | |
| Adult | 26 | 183 | 209 | 12.4 |
| Adult/senior | 327 | 1234 | 1561 | 20.9 |
| Child | 7 | 92 | 99 | 7.1 |
| Child/adult | 3 | 40 | 43 | 7.0 |
| Child/adult/senior | 18 | 127 | 145 | 12.4 |
| Senior | 2 | 3 | 5 | 40.0 |
| Sex | | | | |
| Both | 352 | 1477 | 1829 | 19.2 |
| Female | 23 | 130 | 153 | 15.0 |
| Male | 8 | 72 | 80 | 10.0 |
| Sponsorship | | | | |
| Industry | 354 | 1534 | 1888 | 18.8 |
| Industry/National Institutes of Health | 0 | 2 | 2 | 0.0 |
| Industry/other | 14 | 42 | 56 | 25.0 |
| Industry/other/ National Institutes of Health | 0 | 1 | 1 | 0.0 |
| Industry/United States Fed | 1 | 1 | 2 | 50.0 |
| Other/industry | 0 | 13 | 13 | 0.0 |
| Other/National Institutes of Health/industry | 0 | 1 | 1 | 0.0 |
| National Institutes of Health | 4 | 6 | 10 | 40.0 |
| National Institutes of Health/other | 0 | 1 | 1 | 0.0 |
| Other | 10 | 70 | 80 | 12.5 |
| Other/National Institutes of Health | 0 | 8 | 8 | 0.0 |
| Interventions | | | | |
| Behavioral | 2 | 7 | 9 | 22.2 |
| Biological | 10 | 119 | 129 | 7.8 |
| Device | 1 | 16 | 17 | 5.9 |
| Dietary supplement | 1 | 3 | 4 | 25.0 |
| Drug | 359 | 1424 | 1783 | 20.1 |
| Genetic | 0 | 1 | 1 | 0.0 |
| Other | 2 | 20 | 22 | 9.1 |
| Procedure | 5 | 26 | 31 | 16.1 |
| Radiation | 0 | 1 | 1 | 0.0 |
| Phases of clinical trials | | | | |
| Phase 1 | 0 | 43 | 43 | 0.0 |
| Phase 1/phase 2 | 3 | 31 | 34 | 8.8 |
| Phase 2 | 68 | 419 | 487 | 14.0 |
| Phase 2/phase 3 | 5 | 46 | 51 | 9.8 |
| Phase 3 | 257 | 900 | 1157 | 22.2 |
| Phase 4 | 40 | 131 | 171 | 23.4 |
| Recruitment | | | | |
| Active, not recruiting | 51 | 313 | 364 | 14.0 |
| Approved for marketing | 1 | 1 | 2 | 50.0 |
| Completed | 285 | 747 | 1032 | 27.6 |
| No longer available | 1 | 0 | 1 | 100 |
| Suspended | 0 | 3 | 3 | 0.0 |
| Temporarily not available | 0 | 1 | 1 | 0.0 |
| Terminated | 23 | 99 | 122 | 18.9 |
| Withdrawn | 0 | 8 | 8 | 0.0 |
| Study type | | | | |
| Interventional | 375 | 1576 | 1951 | 19.2 |
| Observational | 6 | 98 | 104 | 5.8 |
| Posting the results in clinicaltrials.gov | | | | |
| Has results | 143 | 192 | 335 | 42.7 |
| No results available | 240 | 1487 | 1727 | 13.9 |

**Table 3 Enrollment of the subjects in the registered clinical studies by study characteristics**

| Study characteristic | Sum | Mean | Std Dev | Median |
|---|---|---|---|---|
| Subject age | | | | |
| Adult | 122622 | 592 | 1334 | 353 |
| Adult/senior | 1817056 | 1181 | 2836 | 462 |
| Child | 35069 | 358 | 891 | 172 |
| Child/adult | 9727 | 226 | 335 | 100 |
| Child/adult/senior | 415280 | 2925 | 9863 | 450 |
| Senior | 9816 | 1963 | 2158 | 450 |
| Subject sex | | | | |
| Both | 2224171 | 1232 | 3849 | 423 |
| Female | 132363 | 877 | 1388 | 419 |
| Male | 53036 | 689 | 1143 | 300 |
| Sponsorship | | | | |
| Industry | 2130061 | 1142 | 3628 | 432 |
| Industry/National Institutes of Health | 465 | 465 | | 465 |
| Industry/other | 123747 | 2250 | 4293 | 500 |
| Industry/other/National Institutes of Health | 8381 | 8381 | | 8381 |
| Industry/United States Fed | 750 | 375 | 318 | 375 |
| National Institutes of Health | 30401 | 3040 | 5516 | 860 |
| National Institutes of Health/other | 280 | 280 | | 280 |
| Other | 91308 | 1201 | 3831 | 145 |
| Other/industry | 17011 | 1309 | 2776 | 200 |
| Other/National Institutes of Health | 4166 | 521 | 477 | 335 |
| Other/National Institutes of Health/ industry | 3000 | 3000 | | 3000 |
| Interventions | | | | |
| Behavioral | 23514 | 2613 | 5846 | 400 |
| Biological | 116476 | 917 | 3958 | 288 |
| Device | 16116 | 948 | 2458 | 90 |
| Dietary supplement | 1712 | 428 | 605 | 150 |
| Drug | 2058994 | 1171 | 3623 | 448 |
| Genetic | 100 | 100 | | 100 |
| Other | 27135 | 1292 | 3539 | 160 |
| Procedure | 8566 | 276 | 381 | 150 |
| Phases of clinical trials | | | | |
| Phases = phase 1 | 2412 | 56 | 63 | 36 |
| Phases = phase 1/phase 2 | 3690 | 109 | 104 | 98 |
| Phases = phase 2 | 149238 | 310 | 374 | 200 |
| Phases = phase 2/phase 3 | 31469 | 629 | 960 | 295 |
| Phases = phase 3 | 1443064 | 1267 | 2773 | 587 |
| Phases = phase 4 | 234401 | 1379 | 3495 | 391 |
| Recruitment | | | | |
| Completed | 1355817 | 1331 | 4549 | 460 |
| Suspended | 792 | 264 | 216 | 210 |
| Terminated | 110736 | 915 | 2501 | 284 |
| Withdrawn | 540 | 77 | 133 | 0 |
| Study type | | | | |
| Interventional | 1845794 | 956 | 2415 | 404 |
| Observational | 563776 | 5421 | 11639 | 1069 |
| Posting of the results in clinicaltrials.gov | | | | |
| Study results = has results | 524214 | 1570 | 4804 | 484 |
| Study results = no results available | 1885356 | 1109 | 3381 | 402 |
| Terminated studies with no results available | | 102335 | 1034 | 2740 |
| Terminated studies with posted results | | 8401 | 382 | 573 |
| Terminated unpublished studies with posted results | | 3554 | 209 | 217 |
| Terminated unpublished studies without posted results | | 38111 | 471 | 760 |
| Publication | | | | |
| Published | 902735 | 2401 | 5662 | 670 |
| Unpublished | 1506835 | 909 | 2954 | 383 |

participating subjects experienced. However, no Russian policy addressed mandatory posting the results neither in

clinicaltrials.gov, nor in Russian trial registry[28]. The Russian Government Decree mandated the registration of all clinical studies in the official web-based Russian trial registry that has not been opened for the public yet[28]. The Russian trial registry and clinicaltrials.gov should include the protocols and patient outcomes for all clinical studies recruiting human subjects on the territory of the Russian Federation.

Further efforts should be made to ensure consistent public access to results of international clinical research[41]. Principal investigators are obligated to provide accurate and complete information about patient outcomes from all funded and conducted studies[24]. Posting results of registered studies on Clinicaltrials.gov has improved public access to the evidence somewhat, but not nearly enough[18-21]. Thus far, systematic reviews of evidence have not raised the issue of how lack of access to results from incomplete studies compromises the validity of review conclusions and decisions in health care in the United States, Russia, and around the world[42]. All multinational studies, complete, terminated, or suspended, and regardless of country specific market approval, should report participant flow and treatment outcomes on ClinicalTrials.gov. Incomplete multinational studies should always post results along with reasons for suspension or termination. Public data should include information about who initiated termination or suspension of the studies, and why. Early discontinuation of trials for commercial reasons has been determined unethical[43,44]. Trials discontinued for safety reasons should be reported in detail as a source of valuable information about treatment harms[45,46]. Evidence-based decisions in health care in the United States, Russia, and other countries can only be possible with complete and accessible information about the benefits and harms of treatment[47,48].

Our study had several limitations. We did not contact the Russian regulatory ministry to obtain more accurate information about all approved clinical studies although no registration information in Russian publications of randomized clinical trials indicated low registration rates. We relied on information provided in trial registries by sponsors and did not contact sponsors or principal investigators regarding missing data or exact reasons for termination. We could not know reasons for poor results availability since trial registry do not have a variable indicating compliance with the United States federal law regarding posting the results.

Nevertheless, our findings demonstrate that existing international, and specifically Russian clinical research, regulations and ethical policy does not guarantee public access to the results from all clinical studies enrolling Russian subjects. There is a growing number of clinical studies in Russia, multinational and national, with a shortage of internationally certified clinical research professionals[49]. Russia has internal training courses on Research Ethics for members of the Council on Ethics of Ministry of Health and Social Development and short training opportunities for the principal investigators in the local universities[50]. Based on our analysis of the Russian legislation, policy, and trends in clinical research, we propose Russian policy changes that can enhance integrity of human research and safety and quality of evidence based health care in accordance with the international ethical principles[44].

First, ethical approval and national and multinational studies conduct should be done by the clinical research professionals with internationally recognized training and certification in clinical research. Second, compliance with the Russian regulation to register all approved clinical studies should be routinely monitored and available to the public[28]. Third, the revised Russian legislation should mandate reporting the patient outcomes from all studies recruiting Russian participants. Forth, Russian scientific peer reviewed journals should adopt the international standards in publishing only registered clinical studies. Finally, transparent evidence based market approval based on high quality clinical evidence applicable to the Russian population should be introduced and routinely monitored including conflict of interest by policy and coverage decision makers[51].

## ACKNOWLEDGMENTS

We would like to thank Michael Berglin, B.S. for his help in editing this paper.

## COMMENTS

### *Background*

Scientific research improves global health care best when international legal and ethical regulations guarantee stakeholders access to the complete and unbiased information it generates. The Russian legislation adopted the international standard in biomedical research and committed itself to meet the highest standards and the integrity of clinical research. The actual legislation in relation to the trends in clinical research involving Russian subjects as well as study sponsorships, types, and results availability has not been examined yet.

### *Research frontiers*

International efforts should be made to ensure consistent public access to results of multinational clinical research. Principal investigators are obligated to provide accurate and complete information about patient outcomes from all funded and conducted studies. Registration of clinical research in The World Health Organization International Clinical Trials Registry Platform and posting results of registered studies on Clinicaltrials.gov has improved public access to the evidence somewhat, but not nearly enough. All multinational studies, complete, terminated, or suspended, and regardless of country specific market approval, should report participant flow and treatment outcomes on ClinicalTrials.gov. Evidence-based decisions in health care in the United Sttaes, Russia, and other countries can only be possible with complete and accessible information about the benefits and harms of available healthcare interventions.

### *Innovations and breakthroughs*

This study found dramatically increasing number of registered studies enrolling Russian subjects. However, patient outcomes were not available from the majority of the registered studies that recruited Russian participants. Many publications of the Russian randomized trials do not mention registration. This study demonstrated that existing international clinical research policy and Russian research regulations do not guarantee availability of the results from human studies. The Russian trial registry and clinicaltrials.gov should include the protocols and patient outcomes for all clinical studies recruiting human subjects on the territory of the Russian Federation.

### *Applications*

These findings demonstrate that existing international, and specifically Russian clinical research, regulations and ethical policy does not guarantee public access to the results from all clinical studies enrolling Russian subjects and therefore should be revised. Based on the analysis of the Russian legislation, policy, and

trends in clinical research, the authors propose Russian policy changes that can enhance integrity of human research and safety and quality of evidence based health care in accordance with the international ethical principles. First, ethical approval and national and multinational studies conduct should be done by the clinical research professionals with internationally recognized training and certification in clinical research. Second, compliance with the Russian regulation to register all approved clinical studies should be routinely monitored and available to the public. Third, the revised Russian legislation should mandate reporting the patient outcomes from all studies recruiting Russian participants. Forth, Russian scientific peer reviewed journals should adopt the international standards in publishing only registered clinical studies. Finally, transparent evidence based market approval based on high quality clinical evidence applicable to the Russian population should be introduced and routinely monitored including conflict of interest by policy and coverage decision makers.

### Peer review

This is a very valuable study about the trend of publication of clinical research. The conclusion is supported by solid data analysis.

## REFERENCES


1 **Clancy CM**, Slutsky JR, Patton LT. Evidence-based health care 2004: AHRQ moves research to translation and implementation. *Health Serv Res* 2004; **39**: xv-xxiii [PMID: 15333128 DOI: 10.1111/j.1475-6773.2004.00286.x]

2 **Kaiser J**. U.S. budget 2009. NIH hopes for more mileage from roadmap. *Science* 2008; **319**: 716 [PMID: 18258872 DOI: 10.1126/science.319.5864.716b]

3 **Loscalzo J**. The NIH budget and the future of biomedical research. *N Engl J Med* 2006; **354**: 1665-1667 [PMID: 16625006]

4 **Steinbrook R**. The NIH stimulus--the recovery act and biomedical research. *N Engl J Med* 2009; **360**: 1479-1481 [PMID: 19357402 DOI: 10.1056/NEJMp0901819]

5 **Slutsky JR**, Clancy CM. The Agency for Healthcare Research and Quality's Effective Health Care Program: creating a dynamic system for discovering and reporting what works in health care. *Am J Med Qual* 2005; **20**: 358-360 [PMID: 16280400 DOI: 10.1177/1062860605281608]

6 **Gass A**, Doyle H. PLOS position on NIH public access policy. *Science* 2005; **308**: 356 [PMID: 15831740 DOI: 10.1126/science.308.5720.356a]

7 **Rossner M**. The NIH policy on enhancing public access to publications resulting from NIH-funded research: can we streamline the process for our authors? *J Cell Biol* 2005; **168**: 991-998 [PMID: 15795313 DOI: 10.1083/jcb.200503015]

8 **Steinbrook R**. Public access to NIH-funded research. *N Engl J Med* 2005; **352**: 1739-1741 [PMID: 15858180 DOI: 10.1056/NEJMp058088]

9 **Kaiser J**. Scientific publishing. NIH wants public access to papers 'as soon as possible'. *Science* 2005; **307**: 825 [PMID: 15705812 DOI: 10.1126/science.307.5711.825]

10 **Dwan K**, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, Decullier E, Easterbrook PJ, Von Elm E, Gamble C, Ghersi D, Ioannidis JP, Simes J, Williamson PR. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One* 2008; **3**: e3081 [PMID: 18769481 DOI: 10.1371/journal.pone.0003081]

11 **Sterne JA**, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000; **53**: 1119-1129 [PMID: 11106885 DOI: 10.1016/S0895-4356(00)00242-0]

12 **Mathieu S**, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009; **302**: 977-984 [PMID: 19724045 DOI: 10.1001/jama.2009.1242]

13 **Food and Drug Administration**. Modernization Act (FDAMA) Pub L No. 105-115, 111 1997; Stat 2296, 2310(113). Available from: URL: http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/SignificantAmendmentstotheFDCAct/FDAMA/FullTextofFDAMAlaw/default.htm

14 **Zarin DA**, Keselman A. Registering a clinical trial in ClinicalTrials.gov. *Chest* 2007; **131**: 909-912 [PMID: 17303677 DOI: 10.1378/chest.06-2450]

15 **Zarin DA**, Ide NC, Tse T, Harlan WR, West JC, Lindberg DA. Issues in the registration of clinical trials. *JAMA* 2007; **297**: 2112-2120 [PMID: 17507347 DOI: 10.1001/jama.297.19.2112]

16 **Zarin DA**, Tse T, Ide NC. Trial Registration at ClinicalTrials.gov between May and October 2005. *N Engl J Med* 2005; **353**: 2779-2787 [PMID: 16382064 DOI: 10.1056/NEJMsa053234]

17 **Food and Drug Administration**. Amendments Act of 2007 (Pub L No. 110-85, 121 Stat 823, 904, Title VIII [September 27, 2007), 42 USC 282(j) (Supp 2009); 2010]. Available from: URL: http://www.gpo.gov/fdsys/pkg/PLAW-110publ85/pdf/PLAW-110publ85.pdf

18 **Chan AW**, Laupacis A, Moher D. Registering results from clinical trials. *JAMA* 2010; **303**: 2138-219; author reply 2139 [PMID: 20516411 DOI: 10.1001/jama.2010.702]

19 **Mancini J**, Reynier CJ. Registering results from clinical trials. *JAMA* 2010; **303**: 2139; author reply 2139 [PMID: 20516413]

20 **Miller JD**. Registering clinical trial results: the next step. *JAMA* 2010; **303**: 773-774 [PMID: 20179288 DOI: 10.1001/jama.2010.207]

21 **Tse T**, Williams RJ, Zarin DA. Update on Registration of Clinical Trials in ClinicalTrials.gov. *Chest* 2009; **136**: 304-305 [PMID: 19584213 DOI: 10.1378/chest.09-1219]

22 **World Association of Medical Editors (WAME) Editorial Policy Committee.** The Registration of Clinical Trials. Policy Statements (2005 Updated 2013 March 12). Available from: URL: http://www.wame.org/policy-statements

23 **The World Health Organization.** International Clinical Trials Registry Platform. 2010

24 **The Office of Research Integrity (ORI)**. Responsible Conduct of Research (RCR). 2nd ed. New York: Oxford University Press, 2009

25 **Russian Federation FL**. On Circulation of Medicines. Adopted by the State Duma. 2010; March 24 (as amended by Federal Laws of July, 27 2010, No.192-FZ; of October 11, 2010, No. 271-FZ, of November 29, No. 313-FZ, of December 06, 2011 No. 409-FZ, of June 25, 2012 No. 93-FZ). Available from: URL: http://www.pravo.gov.ru/index.html

26 **Russian Federation FL**. Public Health Protection in the Russian Federation adopted by the State Duma on November 1, 2011 and approved by the Federation Council on November 9, 2011 (November 21 2012 No. 323-FZ). Available from: URL: http://www.pravo.gov.ru/index.html

27 **Russian Federation FL**. Amendments to Part Second of the Tax Code of the Russian Federation and to Certain Legislative Acts of the Russian Federation (Passed by the State Duma on the 26th of March, 2010. 2010). Available from: URL: http://www.pravo.gov.ru/index.html

28 Russian Federation Government Decree No. 754H of August 26. Approval of the management, publication and posting of the clinical trials of drugs and other productd for medical use on the official web based trial register (2010). Available from: URL: http://www.pravo.gov.ru/index.html

29 Russian Federation Government Decree No. 771 of September 29. On Procedure of Import of medicines for medical use into the Russian Federation (2010). Available from: URL: http://www.pravo.gov.ru/index.html

30 Russian Federation Government Decree No. 683 of September 3. Approval of Rules for Accreditation of Medical Institutions for the Right to Conduct Clinical Trials of Pharmaceutical Drugs for Medical Application (2010). Available from: URL: http://www.pravo.gov.ru/index.html

31 Russian Federation Government Decree No. 673 of September 3. Approval of Rules for Import and Export of Biological Materials Obtained in Clinical Trials of a medicinal product for medical use into and from the Russian Federation (2010). Available from: URL: http://www.pravo.gov.ru/index.html

32 Russian Federation Government Decree No. 714 of September 13. On approval of Typical Rules for Compulsatory Insurance of Life and Health of a Patient Invovled in Clinical Trials of a Medicinal Product Amended by Government Decree of May 18, 2011, No.393. 2010. Available from: URL: http://www.pravo.gov.ru/index.html

33 **Minzdravsocrazvitie Russian Federation**. Conducting good trial of the bioequivalent (BEC) pharmaceutical products. Guidelines (2004). Available from: URL: http://acto-russia.org/index.php?option=com_content&task=view&id=7

34 **Russian Federation Government Decree**. About the Council on ethics. Ministry of Health Order 2010; from 31.08. 2010 N 774 N. Available from: URL: http://www.pravo.gov.ru/index.html

35 **Russian Federation Government Decree.** The order for approval of the clinical trials on the medicines for medical use. Ministry of Health Order. 2010; from 26 .08. 2010 No. 748. Available from: URL: http://www.pravo.gov.ru/index.html

36 **Russian Federation Government Decree**. The order for organization ethical review process for approval clinical trials on the medicines for medical use and the official forms Ministry of Health Order. 2010; from 26 .08. 2010 г. N 753N. Available from: URL: http://www.pravo.gov.ru/index.html

37 Good clinical practice GOST 52379-2005. N 232-art 2005; put into operation 01.04.2006

38 **Russian Federation Government Decree**. On approval of the administrative regulations of the Federal Service on Surveillance in Healthcare and Social Development for the state function to monitor the conduct of pre-clinical trials of drugs and clinical trials of drugs for medical use 2011; Registered with the Ministry of Justice on January 26, 2012, No. 23023. Available from: URL: http://www.pravo.gov.ru/index.html

39 **Russian Federation Government Decree.** Establishment the departments for clinical trials the medicines on volunteers. Ministry of Health Order. 1992; from 25.08.1992. Available from: URL: http://www.pravo.gov.ru/index.html

40 **Moja LP**, Moschetti I, Nurbhai M, Compagnoni A, Liberati A, Grimshaw JM, Chan AW, Dickersin K, Krleza-Jeric K, Moher D, Sim I, Volmink J. Compliance of clinical trial registries with the World Health Organization minimum data set: a survey. *Trials* 2009; **10**: 56 [PMID: 19624821 DOI: 10.1186/1745-6215-10-56]

41 **Drazen JM**. Transparency for clinical trials--the TEST Act. *N Engl J Med* 2012; **367**: 863-864 [PMID: 22873430 DOI: 10.1056/NEJMe1209433]

42 **Spence D**. Why evidence is bad for your health. *BMJ* 2010; **341**: c6368 [DOI: 10.1136/bmj.c6368]

43 Psaty BM, Rennie D. Stopping medical research to save money: a broken pact with researchers and patients. *JAMA* 2003; **289**: 2128-2131 [PMID: 12709471]

44 World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2000; **284**: 3043-3045 [PMID: 11122593 DOI: 10.1001/jama.284.23.3043]

45 **Bassler D**, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, Heels-Ansdell D, Walter SD, Guyatt GH, Flynn DN, Elamin MB, Murad MH, Abu Elnour NO, Lampropulos JF, Sood A, Mullan RJ, Erwin PJ, Bankhead CR, Perera R, Ruiz Culebro C, You JJ, Mulla SM, Kaur J, Nerenberg KA, Schünemann H, Cook DJ, Lutz K, Ribic CM, Vale N, Malaga G, Akl EA, Ferreira-Gonzalez I, Alonso-Coello P, Urrutia G, Kunz R, Bucher HC, Nordmann AJ, Raatz H, da Silva SA, Tuche F, Strahm B, Djulbegovic B, Adhikari NK, Mills EJ, Gwadry-Sridhar F, Kirpalani H, Soares HP, Karanicolas PJ, Burns KE, Vandvik PO, Coto-Yglesias F, Chrispim PP, Ramsay T. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA* 2010; **303**: 1180-1187 [PMID: 20332404 DOI: 10.1001/jama.2010.310]

46 **Sica DA**. Pharmacotherapy in congestive heart failure. Prematurely terminated clinical trials and their application to cardiovascular medicine. *Congest Heart Fail* 2001; **7**: 265-271 [PMID: 11832666 DOI: 10.1111/j.1527-5299.2001.00264.x]

47 **Dickersin K**. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997; **9**: 15-21 [PMID: 9083596]

48 **Egger M**, Smith GD. Bias in location and selection of studies. *BMJ* 1998; **316**: 61-66 [PMID: 9451274 DOI: 10.1136/bmj.316.7124.61]

49 **Tunnicliffe A**. Eastern Europe comes in from the cold. *Clinical Trials Insigh* 2012; **June**: 14-16. Available from: http: //www.worldpharmaceuticals.net

50 The European Forum for Good Clinical Practice. The Procedure for the Ethical Review of Protocols for Clinical Research Projects in Europe (Russia) (The EFGCP Report. 2011). Available from: http: //www.efgcp.be/

51 **Lichterman B**. Conflict or harmony? Clinical research and the medical press in Russia. *Sci Eng Ethics* 2002; **8**: 383-386 [PMID: 12353365 DOI: 10.1007/s11948-002-0057-7]


**P- Reviewer:** Floriani I, Shi Z **S- Editor:** Song XX
**L- Editor:** A **E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.162

*World J Meta-Anal* 2014 November 26; 2(4): 162-170
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Diagnostic yield of third eye retroscope on adenoma detection during colonoscopy: A systematic review and meta-analysis

Nirav Thosani, Bhavana Rao, Sachin Batra, Babatunde Adeyefa, Gottumukkala S Raju, Robert S Bresalier, Subhas Banerjee, Sushovan Guha

Nirav Thosani, Bhavana Rao, Sachin Batra, Babatunde Adeyefa, Sushovan Guha, Division of Gastroenterology, Hepatology and Nutrition, The University of Texas Medical School, Houston, TX 77030, United States
Nirav Thosani, Gottumukkala S Raju, Robert S Bresalier, Department of Gastroenterology, Hepatology and Nutrition, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, United States
Nirav Thosani, Subhas Banerjee, Department of Gastroenterology, Stanford University, Stanford, California, CA 94305, United States
Author contributions: Thosani N contributed to concept and study design, acquisition of data, analysis and interpretation of data, drafting of manuscript, critical revision of manuscript for important intellectual content, statistical analysis and technical support; Rao B contributed to acquisition of data and drafting of manuscript; Batra S contributed to acquisition of data, analysis and interpretation of data, revision of manuscript, statistical analysis, and drafting of manuscript; Adeyefa B contributed to acquisition of data and drafting of manuscript; Raju GS, Bresalier RS and Banerjee S contributed to analysis and interpretation of data and critical revision of manuscript for important intellectual content; Guha S contributed to concept and study design, analysis and interpretation of data, drafting of manuscript, critical revision of manuscript for important intellectual content, statistical analysis, administrative, technical, or material support and supervision.
Correspondence to: Nirav Thosani, MD, MHA, Division of Gastroenterology, Hepatology and Nutrition, The University of Texas Medical School, MSB 4.234, 6431 Fannin Street, Houston, TX 77030, United States. ncthosani@gmail.com
Telephone: +1-713-5006677 Fax: +1-713-5006699
Received: June 1, 2014 Revised: September 20, 2014
Accepted: October 14, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To determine the diagnostic yield of the "third eye retroscope", on adenoma detection rate during screening colonoscopy.

**METHODS:** The "third eye retroscope" when used with standard colonoscopy provides an additional retrograde view to visualize lesions on the proximal aspects of folds and flexures. We searched MEDLINE (PubMed and Ovid), SCOPUS (including MEDLINE and EMBASE databases), Cochrane Database of Systemic Reviews, Google Scholar, and CINAHL Plus databases to identify studies that evaluated diagnostic yield of "third eye retroscope" during screening colonoscopy. DerSimonian Laird random effects model was used to generate the overall effect for each outcome. We evaluated statistical heterogeneity among the studies by using the Cochran Q statistic and quantified by $I^2$ statistics.

**RESULTS:** Four distinct studies with a total of 920 patients, mean age 59.83 (95%CI: 56.77-62.83) years, were included in the review. The additional adenoma detection rate (AADR) defined as the number of additional adenomas identified due to "third eye retroscope" device in comparison to standard colonoscopy alone was 19.9% (95%CI: 7.3-43.9). AADR for right and left colon were 13.9% (95%CI: 9.4-20) and 10.7 (95%CI: 1.9-42), respectively. AADR for polyps $\geqslant$ 6 mm and $\geqslant$ 10 mm were 24.6% (95%CI: 16.6-34.9) and 24.2% (95%CI: 12.9-40.8), respectively. The additional polyp detection rate defined as the number of additional polyps identified due to "third eye retroscope" device in comparison to standard colonoscopy alone was 19.8% (95%CI: 7.9-41.8). There were no complications reported with use of "third eye retroscope" device.

**CONCLUSION:** The "third eye retroscope" device when used with standard colonoscopy is safe and detects 19.9% additional adenomas, compared to standard colonoscopy alone.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Third Eye Retroscope; Screening colonoscopy; Adenoma detection rate; Polyp detection rate; Additional adenoma detection rate; Additional polyp detection rate

**Core tip:** Missed lesions remain one of the main causes for development of interval colorectal cancer after screening colonoscopy. The third eye retroscope (TER) provides an additional retrograde view to visualize lesions on the proximal aspects of folds. In this meta-analysis of four studies including 920 patients undergoing screening colonoscopy, we observed that the additional adenoma detection rate, defined as the number of additional adenomas identified due to TER device in comparison to standard colonoscopy alone, was 19.9% (95%CI: 7.3-43.9). We found that TER device was safe and identified significant more numbers of adenoma which would have been missed by standard colonoscopy alone.

Thosani N, Rao B, Batra S, Adeyefa B, Raju GS, Bresalier RS, Banerjee S, Guha S. Diagnostic yield of third eye retroscope on adenoma detection during colonoscopy: A systematic review and meta-analysis. *World J Meta-Anal* 2014; 2(4): 162-170 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/162.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.162

## INTRODUCTION

Colorectal cancer is the third most frequent cancer diagnosis and the second leading cause of cancer death in the United States[1]. As the disease exhibits a characteristic sequence of progression from precancerous adenomas to carcinoma stage[2], colonoscopy has been recommended by the United States Preventive Services Task Force as one of the criterion standard for the detection as well as the removal of adenomas to halt the progression to malignancy[1,3,4]. However, despite meticulous techniques, adenoma miss rates as high as 5% to 24% have been reported[5-7], thus leading to the development of colorectal cancer within short intervals of completion of complete colonoscopy[8-11]. These lesions maybe missed due to suboptimal withdrawal technique, inadequate bowel cleansing or in some cases, impaired visualization due to flattened and depressed contour of the lesions. Another technical factor identified is the likelihood of missing lesions located on "blind spots" such as the proximal aspect of flexures or haustral folds, rectal and ileocecal valves and behind semilunar folds.

To circumvent this, Triadafilopoulos *et al*[12] evaluated an ancillary imaging device called the third eye retroscope (TER) (Avantis Medical Systems, Inc, Sunnyvale, CA) in anatomic models of colon with simulated polyps in 2007. This device provides a continuous backward or retrograde view that complements the forward view of the standard colonoscope as they are withdrawn together through the colon and thus enhances the visualization of hidden areas, which was supported by the increased adenoma detection rates with its use[13]. Subsequent trials have undertaken the use of TER in humans and have unanimously reported higher adenoma detection rates, especially for smaller sized lesions with moderate increase in procedure times and no predilection for occurrence of complications[14-16]. This finding is clinically significant as adenoma detection rate has been proven as an independent predictor for development of interval colorectal cancer after screening colonoscopy[17]. The aim of the study was to systematically review and perform a meta-analysis regarding the clinical impact of TER on adenoma detection.

## MATERIALS AND METHODS

### Search strategy

The systemic review was performed according to recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement[18]. We queried MEDLINE (PubMed and Ovid from 1990 to April 2014), SCOPUS, Cochrane Database of Systemic Reviews, Google scholar, and CINAHL Plus databases for eligible studies using the search term "Third Eye Retroscope" or "Retroscope" or "Retrograde colonoscopy" or "Retrograde viewing device". In addition, references of eligible articles retrieved by online search were reviewed for the eligible studies by the investigators. Only articles describing findings on human subjects were included. The screening and data abstraction were performed by two independent investigators (N.T. and B.R.) according to inclusion and exclusion criteria, which were pre-defined in this study. We resolved differences by discussion with third investigator (S.G.).

### Study characteristics

**Inclusion criteria:** The study population consisted of patients undergoing colonoscopy for screening or surveillance after previous polypectomy.

**Intervention:** The intervention was standard colonoscopy or standard colonoscopy with the use of TER.

**Study design:** Studies where the findings of standard colonoscopy were compared against the results of standard colonoscopy with TER, performed in the same patient, were included in the analysis.

**Outcome:** The main outcome of the study was to evaluate TER on the following parameters: Additional polyp detection rate (APDR) and Additional adenoma detection rate (AADR). The additional adenoma/polyps were defined as lesions which can be detected only by the TER. Any lesion which was visible on both TER and colonoscope was attributed to standard colonoscope and was not counted among the TER associated additional lesions.

**Exclusion criteria:** We excluded case reports, case series and review articles.

**Table 1 Characteristics of included studies**

| Ref. | Location | Study design | Mean patient age (yr) | Sample size | Intervention | Quality score | ΔTER/SC | | ΔTER / SC mean size (range) mm | | Mean time (min) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Polyps | Adenoma | Polyps | Adenoma | Withdrawal | Total |
| Triadafilopoulos *et al*[16] 2008 | United States | Prospective, Single center study | 64 ± 9.6 | (T): 24 (M): 16 (F): 8 | SC + TER | 23 | 4/34 | 1/8 | 7.5 (2-7)/ 4 (1-15) | 7/8 (4-15) | NA | NA |
| Waye *et al*[14] 2010 | United States | Prospective, Multicenter study | 63.1 ± 6.5 | (T): 249 (M): 132 (F): 117 | SC + TER | 22 | 34/257 | 15/136 | 5.2 (2-12)/ 4.4 (1-15) | NA | 10.9 | 26 |
| DeMarco *et al*[15] 2010 | United States | Prospective, Multicenter study | 56.8 ± 11.3 | (T): 298 (M): 144 (F): 154 | SC + TER | 22 | 27/182 | 16/100 | 6.5 (2-13)/ 5.5 (1-40) | 6.8 (2-13)/ 6.5 (1-40) | 9.9 ± 4.1 | 24.7 ± 9.7 |
| Leufkens *et al*[13] 2011[1] | Europe and United States | Randomized controlled trial, Multicenter study | Group A: 57.9 ± 9.7 | (T): 173 (M): NA (F): NA | SC → SC + TER | 23 | 78/160 | 49/107 | NA | NA | 9.52/7.58 | 20.87/16.97 |
| | | | Group B: 58.2 ± 9.7 | (T): 176 (M): NA (F): NA | SC + TER → SC | | 31/163 | 52/121 | NA | NA | | |

[1]Time TER/SC. SC: Standard colonoscopy; TER: Third eye retroscope; T: Total; F: Female; M: Male; NA: Not available.

### Data abstraction

Data abstraction was performed by two independent investigators (N.T. and B.R.) using MS-EXCEL based standardized data forms (Microsoft Excel, Microsoft Corporation, Redmond, Wash). The following information was recorded from individual studies: (1) study characteristics: country, publication year, study methodology including design, sample size, standardization of measurement and the clinical context; (2) summary measures of Demographic features such as age, race and gender distribution of the study population; and (3) interventions: manufacture and type of the TER used in each study.

### Assessment of risk of bias (Quality criteria)

Study eligibility and quality assessment were critically performed by three independent investigators (N.T., B.R., and S.B.). Any disagreements were adjudicated by a third investigator (S.G.) through discussion. We evaluated each study based on the inclusion and exclusion criteria as previously delineated. We used the Standards for the Reporting of Diagnostic Studies checklist to accurately report quality of our included studies[19].

### Data synthesis and statistical analysis

For the purpose of this meta-analysis, APDR was defined as the number of additional polyps identified due to TER device in comparison with standard colonoscopy. The additional AADR was defined as the number of additional adenomas identified due to TER device in comparison with standard colonoscopy. Technical error was defined as percentage of the patients in which TER failed to provide sufficient quality retrograde images either related issues with the cap size, failure to retroflex after coming out of the biopsy channel or due to malfunction of the image processor. DerSimonian Laird random effects model was used to generate the overall effect for each outcome. The random-effects model is more infallible than the fixed effect model as it incorporates into the weighing scheme both intra-study and inter-study variations[20].

Inter-Study heterogeneity was assessed based on Cochran Q statistic and $I^2$ statistics[21], with $I^2$ values of 25%, 50% and 75%, being classified as low, moderate and high heterogeneity, respectively[22]. We assessed the impact of removing one study at a time and determined the change in magnitude, direction, and the standard error of the pooled effect size. Meta-regression was applied to further explore heterogeneity in the pooled studies. In addition, we assessed the robustness of our meta-analysis with regards to the publication bias by using several tests including Egger regression asymmetry test[23], and Fail-safe N tests[24]. Further, Funnel plot was constructed to evaluate the publication bias using standard error and event rate[25,26]. Finally, Trim and Fill method was used to test and adjust for any significant publication bias in our meta-analysis. All statistical tests were performed with the Comprehensive Meta-analysis version 2.0 (Biostat, Englewood, NJ). We considered $P < 0.05$ as statistically significant for this meta-analysis.

## RESULTS

### Literature search

A total of 12 articles were identified and reviewed in detail by using the search terms of "Third Eye Retroscope". Of these 12 articles; 1 study was on animal model, and 6 were either review articles or letters to the editor, and were excluded from the analysis. In total, 4 studies, 1 randomized controlled trial and 3 prospective studies were selected for this meta-analysis[13-16]. One additional study was a post hoc analysis of the included randomized trial and these study was also excluded from the analysis[27]. The PRISMA checklist for systematic review was reported on Supplementary Table 1. The characteristics along with the quality of the included studies are shown in Table 1.

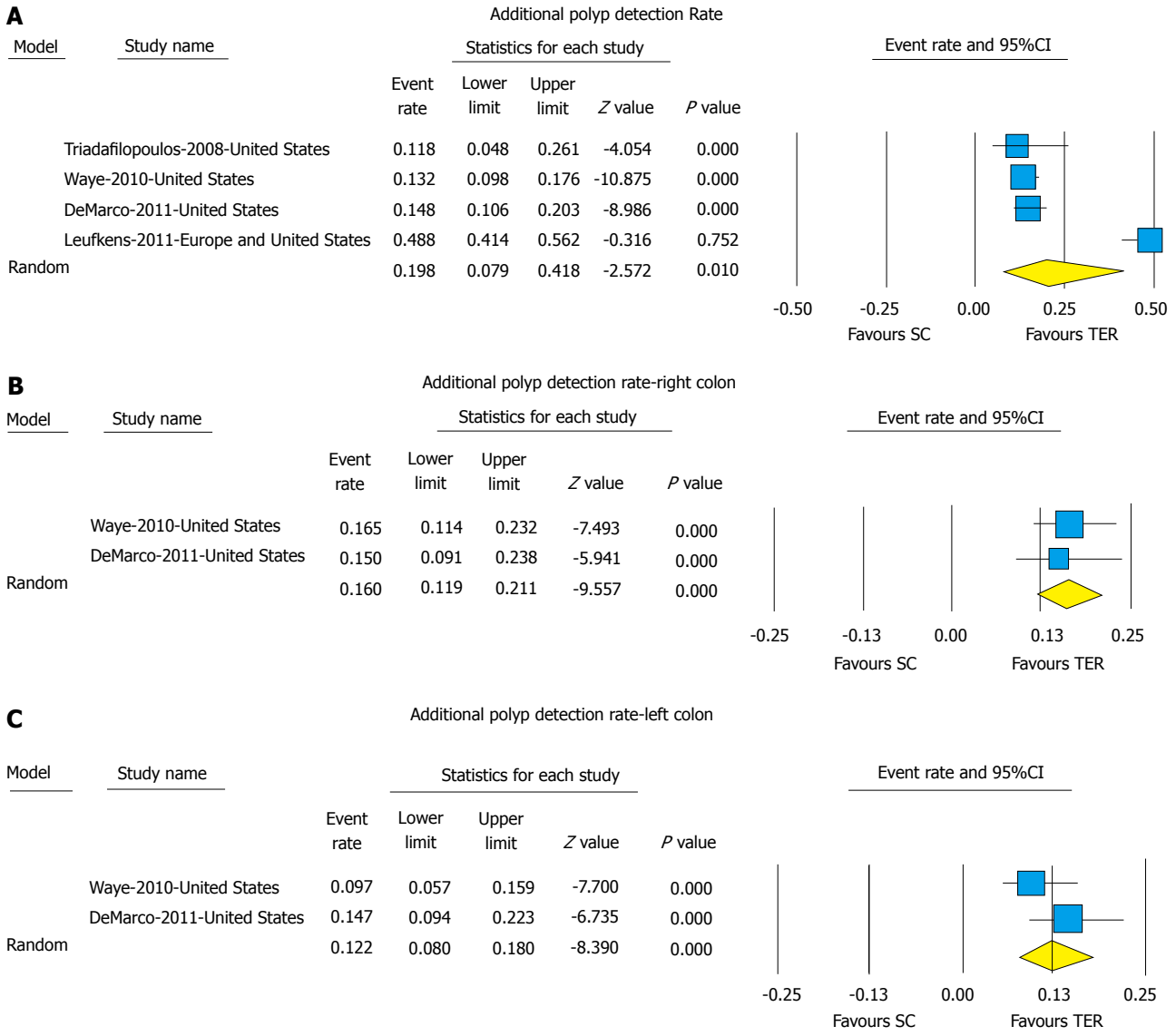



**Figure 1 (A) Pooled additional polyp detection rate, (B) additional polyp detection rate-right colon, and (C) additional polyp detection rate-left colon with use of third eye retroscope compared to standard colonoscopy alone.** The size of the each square is proportional to the sample size for each study, and the horizontal lines through the square indicate the 95%CI for that study. For the pooled analysis, the diamond indicated the pooled value and the right and left ends of the diamond indicate the 95%CI for the analysis. TER: Third eye retroscope; SC: Standard colonoscope.

Of the four studies included in the meta-analysis, Triadafilopoulos *et al*[12] found a total of 38 polyps with mean size of 4 ± 3 mm. Resected polyps consisted of 9 adenomatous polyps and 29 hyperplastic polyps. Of the 38 polyps, 30 polyps were detected in antegrade view, 4 in both antegrade-retrograde views, and TER helped in detection of 4 polyps that were located in the proximal aspect of the haustrae and visible only in retrograde view. Thus use of TER led to incremental yield of standard colonoscope (SC) by 11.8% through detection of additional polyps: 3 hyperplastic polyps (sizes: 2 mm, 3 mm, and 3 mm) and one tubullovillous adenoma (size: 7 mm). DeMarco *et al*[15] in demonstrated that TER detected 27 (14.8%) polyps in addition to 182 polyps detected with SC. While, 100 adenomas were detected on SC, TER facilitated detection of 16 (16%) more adenomas ($P < 0.001$). In addition, the mean sizes of the polyps or adenomas detected with SC and TER were similar ($P = 0.37$ and $P = 0.87$, respectively) and subgroup analyses revealed 23% increment in detection of polyps 6 mm or larger. Similar to above findings, Waye *et al*[14] also found that polyps and adenomas detected by the TER were comparable in terms of size to those detected by SC. However, in a two arm comparison of SC with TER, Leufkens *et al*[13] demonstrated TER detecting higher number of advanced polyps and adenoacarcinomas.

### Meta-analysis

The four studies recruited a total of 920 patients with a mean age of 59.83 (95%CI: 56.77-62.83) years. A total of 46 patients were excluded from the studies due to technical error/device malfunction, with overall technical error rate of 4.6% (95%CI: 2.2-9.3) in intention to treat analysis.

The results of the meta-analysis are shown in detail in Table 2. The pooled APDR due to TER device in comparison to standard colonoscopy was 19.8 % (95%CI:

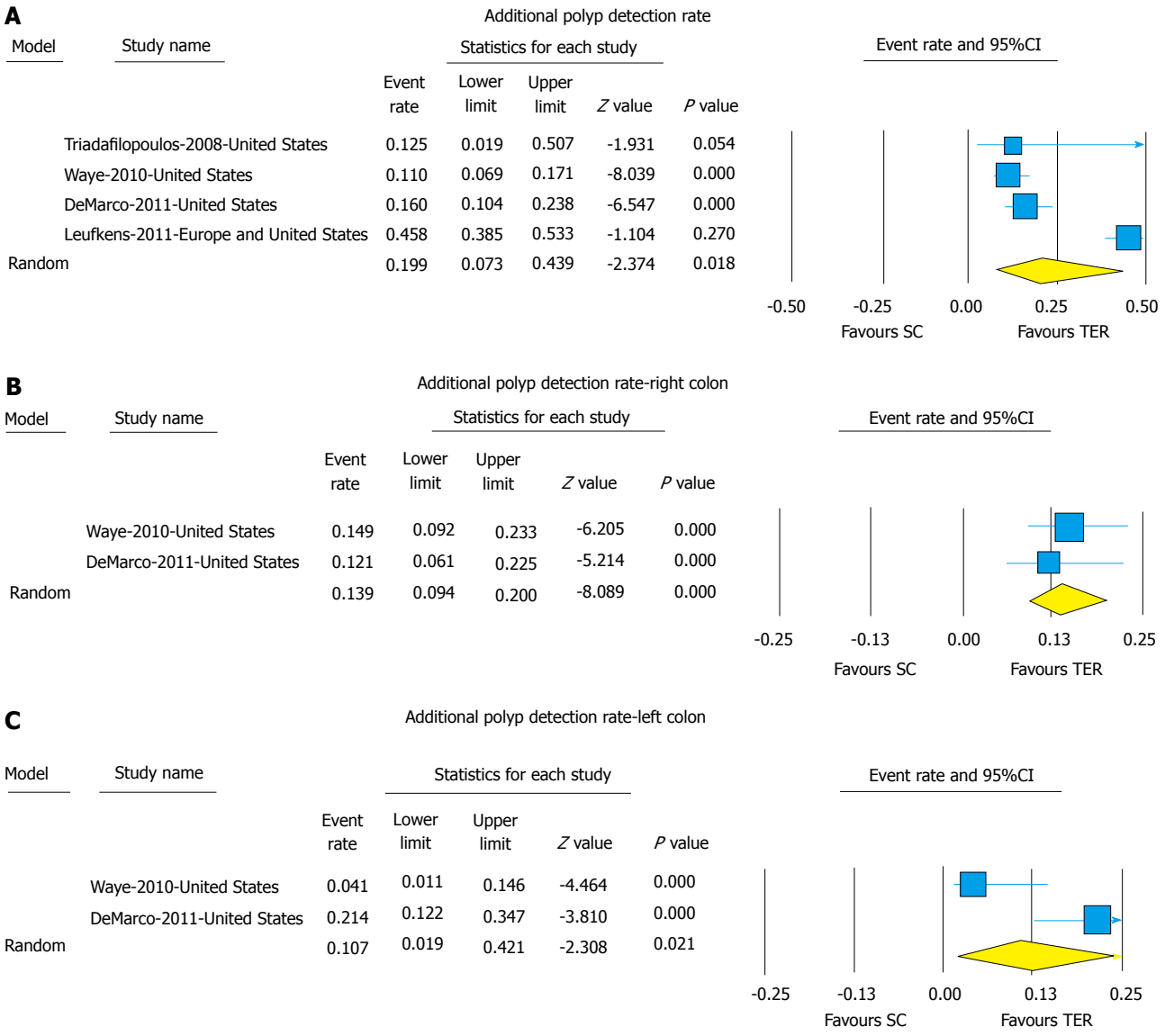



**Figure 2 (A) Pooled additional adenoma detection rate, (B) additional adenoma detection rate-right colon, and (C) additional adenoma detection rate-left colon with use of third eye retroscope compared to standard colonoscopy alone.** The size of the each square is proportional to the sample size for each study, and the horizontal lines through the square indicate the 95%CI for that study. For the pooled analysis, the diamond indicated the pooled value and the right and left ends of the diamond indicate the 95%CI for the analysis. TER: Third eye retroscope; SC: Standard colonoscope.

7.9-41.8) (Figure 1A). APDR for right colon and left colon was 16% (95%CI: 11.9-21.2) and 12.2% (95%CI: 8-18), respectively (Figures 1B and C). APDR for polyp ⩾ 6 mm in size was 20.9% (95%CI: 14.8-28.6) and for polyp ⩾ 10 mm in size was 25.3% (95%CI: 15.5-38.4).

The pooled AADR due to TER device in comparison to standard colonoscopy was 19.9% (95%CI: 7.3-43.9) (Figure 2A). AADR for right colon and left colon was 13.9% (95%CI: 9.4-20) and 10.7% (95%CI: 1.9-42), respectively (Figures 2B and C). AADR for polyp ⩾ 6 mm in size was 24.6% (95%CI: 16.6-34.9) and for polyp ⩾ 10 mm in size was 24.2% (95%CI: 12.9-40.8). There were no complications related to TER during the procedure.

The Funnel plots evaluating publication bias for both APDR and AADR are shown in Figure 3A and B. The Fail-safe N test showed that for the combined 2-tailed *P* value to be no longer significant ($P > 0.05$), it would need an additional 149 studies with "null" results. By using the random-effect model, the APDR and AADR were noted as 19.8% (95%CI: 7.9-41.8) and 19.9% (95%CI: 7.3-43.9), respectively and these values did not change using the Trim and Fill method.

## DISCUSSION

Colonoscopy has so far been considered the best available method for colorectal cancer screening as it is sensitive and specific for detection of precancerous lesions. Yet, when the performance of colonoscopy have been evaluated with back to back colonoscopic examinations[5-7,28] or when compared with the results of CT colonography[29,30], colonoscopy has often been found to have missed a significant number of lesions. Two tandem colonoscopy studies reported adenoma miss rates of

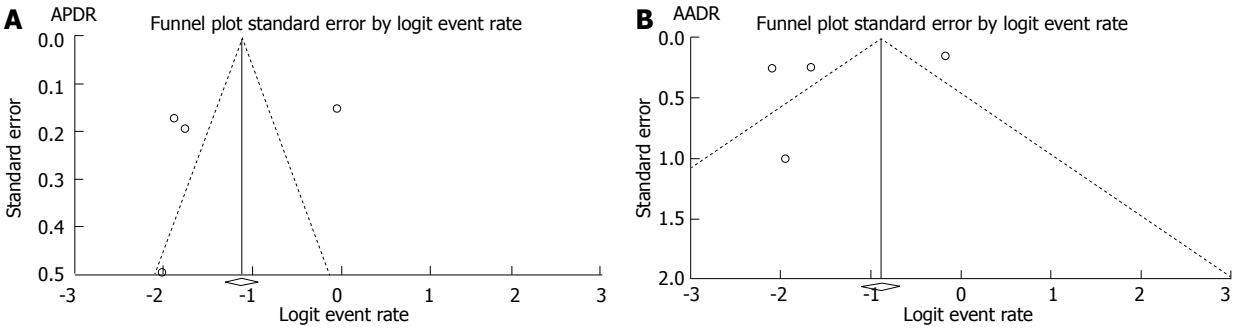



**Figure 3 Funnel Plot for publication bias for (A) additional polyp detection rate and (B) additional adenoma detection rate.**

**Table 2 Results of meta-analysis: Additional polyp detection rate and additional adenoma detection rate**

| Category | No. of studies | Event rate (%) | 95%CI | Heterogeneity | |
|---|---|---|---|---|---|
| | | | | *P* | $I^2$ (%) |
| APDR | 4 | 19.8 | 7.9-41.8 | 0 | 96.38 |
| APDR right colon | 2 | 16 | 11.9-21.2 | 0.76 | 0 |
| APDR left colon | 2 | 12.2 | 8-18 | 0.22 | 31.9 |
| APDR 6 mm or more | 2 | 20.9 | 14.8-28.6 | 0.47 | 0 |
| PDR 10 mm or more | 2 | 25.3 | 15.5-38.4 | 0.51 | 0 |
| AADR | 4 | 19.9 | 7.3-43.9 | 0 | 94.44 |
| AADR right colon | 2 | 13.9 | 9.4-20 | 0.61 | 0 |
| AADR left colon | 2 | 10.7 | 1.9-42 | 0.02 | 82 |
| AADR 6 mm or more | 2 | 24.6 | 16.6-34.9 | 0.94 | 0 |
| AADR 10 mm or more | 2 | 24.2 | 12.9-40.8 | 0.34 | 0 |

APDR: Additional polyp detection rate; AADR: Additional adenoma detection rate.

20%-24%[6,28]. A study comparing colonoscopy with CT colonography reported a colonoscopy miss rate of 12% for polyps 1 cm or larger[29]. In addition, colon cancer had been diagnosed within a short interval after a meticulous colonoscopy[31,32]. These studies shed light on missed lesions after standard colonoscopy and raise the issue that colonoscopy is a less sensitive screening tool than previously perceived, with profound clinical and medico-legal implications[33-35].

Factors contributing to the colonoscopy miss rates include operator related variables such as lower level of expertise, sub optimal withdrawal technique[36], decreased observation time during extubation[36-40] and other technical factors such as poor bowel preparation[36,41-43] and inadequate visualization of all areas[29,36]. Pickhardt *et al*[29] identified that 67% of the adenomas that were missed by colonoscopy but had been picked up by CT colonography were located on the proximal aspect of folds, which account for 7.8% of colon surface area[44]. Another study concurred that lesions located on the proximal aspect of colonic mucosal or haustral folds, behind the rectal valves, beneath the ileocecal valve, or on the inner curve of flexures are difficult to detect with a forward-viewing colonoscope, even with various manipulations[45]. To enhance exposure of mucosa in these "hidden areas", TER has been evaluated with promising results[10-14].

The TER system consists of 3 parts: the retroscope, a cap with a polarizing filter, and a dedicated video-image processor. After the standard colonoscope is advanced into the cecum, the retroscope is inserted through the "working channel" of the colonoscope until it extends beyond its other end. Since the distal tip of the TER naturally assumes a bent "J" shape after emerging from channel, it provides 180-degree retrograde illumination and visualization of the colon when the standard colonoscope is withdrawn. In comparison to TER, standard colonoscope only provides 140-degree field of view during examination. During its initial evaluation in an anatomic model with simulated polyps, the use of TER contributed to greater visualization of simulated polyps placed on obscure locations including behind the folds when compared to the standard colonoscope alone (81% *vs* 12%, $P = 0.00001$)[12]. Also TER improved visualization of colonic mucosal surface area from 87% to 99% when used in conjunction with a standard colonoscope, thus facilitating enhanced detection of lesions[46].

This enhanced detection is gauged by two parameters - APDR and AADR. Our analysis has showed APDR of 19.8 % (95%CI: 7.9-41.8) and AADR of 19.9% (95%CI: 7.3-43.9) with the use of TER device in comparison to standard colonoscopy alone. These additional polyps were detected in both right and left colon (16% *vs* 12.2%) and a similar observation was also made for adenoma detection in both right and left colon (13.9% *vs* 10.7%). This diagnostic benefit provided by the TER is of clinical importance, since standard colonoscopy has been reported to be less protective for right sided colorectal cancer[8,47]. With the use of TER, the APDR and AADR for polyps greater than $\geqslant$ 10 mm in size was 25.3% and 24.2 % respectively. These lesions have been shown to have a greater likelihood for malignant transformation[48,49] and hence their detection and removal would enhance the efficacy of screening colonoscopy.

Some technical difficulties with the use of the device have been noted including the inability to advance the catheter through the "working channel", inadequate clasping onto the cap and kinking of the retroscope. These device related reasons led to the exclusion of 20 patients in the 4 studies. However no procedural complications or device related safety issues were reported, although the use of TER was avoided in patients with potential risk factors including ulcerative colitis or Crohn's disease, familial adenomatous polyposis, and severe di-

verticulosis[14,15].

Some of the disadvantages of using a TER include the inferior image quality[50], the impaired ability to aspirate luminal contents[50] and the increase in the withdrawal and procedure time, which may be attributed to the time required for removal and reintroduction of the auxiliary device when a polyp is encountered[51] and partly due to the increased number of polypectomies[13]. However with increasing endoscopist's experience with the device, significant decreases in withdrawal times and total procedure times have been noted[15]. Although the polarizing filter minimizes the blinding of the retroscope's image by the colonoscope's light, residual glare has been reported as a minor distraction[12]. In order to diminish this effect, a short period of adjustment is necessary to adapt to the side-by-side viewing of images with distinctive directional orientations[50].

In the TER system, the cap and catheter are single use only and a large initial investment is necessary for acquiring the video processor. The post hoc analysis performed on the TERRACE trial showed that the AADR with TER for screening colonoscopy was only 4.4%, whereas for surveillance and diagnostic colonoscopy it was 35.7% and 55.4% respectively[27]. When this modest improvement in the diagnostic yield of screening colonoscopy is weighed against the cost of performing the procedure, the benefit of using the TER in routine clinical practice is brought to question. Results of this meta-analysis will help future cost-effectiveness analysis to determine routine use of TER.

In summary, the results from this systematic review suggest that the TER is a safe and promising tool for enhancing the efficacy of the standard colonoscope for detection of colonic polyps and adenomas. This device allows retrograde views of the proximal parts of colonic mucosal or haustral folds, back of the flexures and thereby increased diagnostic yield of identifying adenomatous polyps by almost 20%.

## COMMENTS

### Background

Colorectal cancer is the third most frequently diagnosed cancer and the second leading cause of cancer death in the United States. Colonoscopy facilitates early detection of adenomatous polyps in colon and therefore has been recommended by the United States Preventive Services Task Force as a screening tool for prevention of colorectal cancer. However, despite several improvements in the colonoscope design and operator performance, up-to one fourth of adenomas are missed and could lead to increase in the risk of colorectal cancer. These lesions maybe missed due to various reasons including being located on 'blind spots' such as the proximal aspect of flexures or haustral folds, rectal and ileocecal valves and behind semilunar folds. Third eye retroscope (TER) is an attachment to standard colonoscope which allows endoscopist to inspect regions of the colon not visible to standard colonoscope. The published evidence for the benefit of this device has been equivocal and limited mostly to small studies. The current study undertakes a systematic review of the existing evidence regarding the benefit of using TER attachment with standard colonoscopy and performs a meta-analysis of adenoma and polyp miss-rate to summarize the results.

### Research frontiers

The meta-analysis evaluates the hypothesis that third eye retroscope reduces the polyp and adenoma miss-rate when used in conjunction with the standard colonoscope. In addition, the study determines the safety of the device.

### Innovations and breakthroughs

The key findings of the study indicate that use of third eye retroscope attachment significantly improves the adenoma detection and reduces the risk of colorectal cancer over the standard colonoscope. The results show that this device when used with standard colonoscopy is safe and detects 19.9% additional adenomas, compared to standard colonoscopy alone.

### Applications

TER is a safe device and its use in addition to standard colonoscopy can make latter a more effective screening tool to prevent colorectal cancer.

### Terminology

Third eye retroscope is a linear device that can be easily passed through the working channel of standard colonoscope. Since the distal tip of the TER naturally assumes a bent "J" shape it provides 180-degree retrograde illumination and visualization of the colon when the standard colonoscope is withdrawn. It has a video processor to capture images. Randomized clinical trial is a clinical trial which assigns treatments to patients randomly to compare effectiveness of two treatments. It is a gold standard of comparing treatments. Meta-analysis is a method of combining treatment effects from several studies in order to derive more conclusive and robust result on efficacy of treatments.

### Peer review

This was a well-written article.

## REFERENCES


1 **United States Preventive Services Task Force.** Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2008; **149**: 627-637 [PMID: 18838716 DOI: 10.7326/0003-4819-149-9-200811040-00243]

2 **Baretton GB**, Aust DE. Colorectal adenoma-carcinoma-sequence an update. Neues von der Adenom-Karzinom-Sequenz. *Endo heute* 2011; **24**: 164-170 [DOI: 10.1055/s-0031-1283721]

3 **Levin B**, Lieberman DA, McFarland B, Smith RA, Brooks D, Andrews KS, Dash C, Giardiello FM, Glick S, Levin TR, Pickhardt P, Rex DK, Thorson A, Winawer SJ. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *CA Cancer J Clin* 2008; **58**: 130-160 [PMID: 18322143 DOI: 10.3322/CA.2007.0018]

4 **Winawer SJ**, Zauber AG, Fletcher RH, Stillman JS, O'brien MJ, Levin B, Smith RA, Lieberman DA, Burt RW, Levin TR, Bond JH, Brooks D, Byers T, Hyman N, Kirk L, Thorson A, Simmang C, Johnson D, Rex DK. Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. *CA Cancer J Clin* 2006; **56**: 143-159; quiz 184-185 [PMID: 16737947 DOI: 10.3322/canjclin.56.3.143]

5 **Ahn SB**, Han DS, Bae JH, Byun TJ, Kim JP, Eun CS. The Miss Rate for Colorectal Adenoma Determined by Quality-Adjusted, Back-to-Back Colonoscopies. *Gut Liver* 2012; **6**: 64-70 [PMID: 22375173]

6 **Rex DK**, Cutler CS, Lemmel GT, Rahmani EY, Clark DW, Helper DJ, Lehman GA, Mark DG. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology* 1997; **112**: 24-28 [PMID: 8978338 DOI: 10.1016/S0016-5085(97)70214-2]

7 **Hixson LJ**, Fennerty MB, Sampliner RE, Garewal HS. Prospective blinded trial of the colonoscopic miss-rate of large colorectal polyps. *Gastrointest Endosc* 1991; **37**: 125-127 [PMID: 2032595 DOI: 10.1016/S0016-5107(91)70668-8]

8 **Baxter NN**, Goldwasser MA, Paszat LF, Saskin R, Urbach DR, Rabeneck L. Association of colonoscopy and death from colorectal cancer. *Ann Intern Med* 2009; **150**: 1-8 [PMID: 19075198 DOI: 10.7326/0003-4819-150-1-200901060-00306]

9 **Citarda F**, Tomaselli G, Capocaccia R, Barcherini S, Crespi M.

Efficacy in standard clinical practice of colonoscopic polypectomy in reducing colorectal cancer incidence. *Gut* 2001; **48**: 812-815 [PMID: 11358901 DOI: 10.1136/gut.48.6.812]
10 **Robertson DJ**, Greenberg ER, Beach M, Sandler RS, Ahnen D, Haile RW, Burke CA, Snover DC, Bresalier RS, McKeown-Eyssen G, Mandel JS, Bond JH, Van Stolk RU, Summers RW, Rothstein R, Church TR, Cole BF, Byers T, Mott L, Baron JA. Colorectal cancer in patients under close colonoscopic surveillance. *Gastroenterology* 2005; **129**: 34-41 [PMID: 16012932 DOI: 10.1053/j.gastro.2005.05.012]
11 **Brady AP**, Stevenson GW, Stevenson I. Colorectal cancer overlooked at barium enema examination and colonoscopy: a continuing perceptual problem. *Radiology* 1994; **192**: 373-378 [PMID: 8029400 DOI: 10.1148/radiology.192.2.8029400]
12 **Triadafilopoulos G**, Watts HD, Higgins J, Van Dam J. A novel retrograde-viewing auxiliary imaging device (Third Eye Retroscope) improves the detection of simulated polyps in anatomic models of the colon. *Gastrointest Endosc* 2007; **65**: 139-144 [PMID: 17185094]
13 **Leufkens AM**, DeMarco DC, Rastogi A, Akerman PA, Azzouzi K, Rothstein RI, Vleggaar FP, Repici A, Rando G, Okolo PI, Dewit O, Ignjatovic A, Odstrcil E, East J, Deprez PH, Saunders BP, Kalloo AN, Creel B, Singh V, Lennon AM, Siersema PD. Effect of a retrograde-viewing device on adenoma detection rate during colonoscopy: the TERRACE study. *Gastrointest Endosc* 2011; **73**: 480-489 [PMID: 21067735 DOI: 10.1016/j.gie.2010.09.004]
14 **Waye JD**, Heigh RI, Fleischer DE, Leighton JA, Gurudu S, Aldrich LB, Li J, Ramrakhiani S, Edmundowicz SA, Early DS, Jonnalagadda S, Bresalier RS, Kessler WR, Rex DK. A retrograde-viewing device improves detection of adenomas in the colon: a prospective efficacy evaluation (with videos). *Gastrointest Endosc* 2010; **71**: 551-556 [PMID: 20018280 DOI: 10.1016/j.gie.2009.09.043]
15 **DeMarco DC**, Odstrcil E, Lara LF, Bass D, Herdman C, Kinney T, Gupta K, Wolf L, Dewar T, Deas TM, Mehta MK, Anwer MB, Pellish R, Hamilton JK, Polter D, Reddy KG, Hanan I. Impact of experience with a retrograde-viewing device on adenoma detection rates and withdrawal times during colonoscopy: the Third Eye Retroscope study group. *Gastrointest Endosc* 2010; **71**: 542-550 [PMID: 20189513 DOI: 10.1016/j.gie.2009.12.021]
16 **Triadafilopoulos G**, Li J. A pilot study to assess the safety and efficacy of the Third Eye retrograde auxiliary imaging system during colonoscopy. *Endoscopy* 2008; **40**: 478-482 [PMID: 18543136 DOI: 10.1055/s-2007-995811]
17 **Kaminski MF**, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, Zwierko M, Rupinski M, Nowacki MP, Butruk E. Quality indicators for colonoscopy and the risk of interval cancer. *N Engl J Med* 2010; **362**: 1795-1803 [PMID: 20463339 DOI: 10.1056/NEJMoa0907667]
18 **Moher D**, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535 [PMID: 19622551]
19 **Bossuyt PM**, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Standards for Reporting of Diagnostic Accuracy. *Clin Chem* 2003; **49**: 1-6 [PMID: 12507953 DOI: 10.1373/49.1.1]
20 **DerSimonian R**, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177-188 [PMID: 3802833 DOI: 10.1016/0197-2456(86)90046-2]
21 **Higgins JP**, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539-1558 [PMID: 12111919 DOI: 10.1002/sim.1186]
22 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557]
23 **Egger M**, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634 [PMID: 9310563 DOI: 10.1136/bmj.315.7109.629]
24 **Duval S**, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; **56**: 455-463 [PMID: 10877304 DOI: 10.1111/j.0006-341X.2000.00455.x]
25 **Sterne JA**, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol* 2001; **54**: 1046-1055 [PMID: 11576817 DOI: 10.1016/S0895-4356(01)00377-8]
26 **Sterne JA**, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001; **323**: 101-105 [PMID: 11451790 DOI: 10.1136/bmj.323.7304.101]
27 **Siersema PD**, Rastogi A, Leufkens AM, Akerman PA, Azzouzi K, Rothstein RI, Vleggaar FP, Repici A, Rando G, Okolo PI, Dewit O, Ignjatovic A, Odstrcil E, East J, Deprez PH, Saunders BP, Kalloo AN, Creel B, Singh V, Lennon AM, DeMarco DC. Retrograde-viewing device improves adenoma detection rate in colonoscopies for surveillance and diagnostic workup. *World J Gastroenterol* 2012; **18**: 3400-3408 [PMID: 22807609 DOI: 10.3748/wjg.v18.i26.3400]
28 **Heresbach D**, Barrioz T, Lapalus MG, Coumaros D, Bauret P, Potier P, Sautereau D, Boustière C, Grimaud JC, Barthélémy C, Sée J, Serraj I, D'Halluin PN, Branger B, Ponchon T. Miss rate for colorectal neoplastic polyps: a prospective multicenter study of back-to-back video colonoscopies. *Endoscopy* 2008; **40**: 284-290 [PMID: 18389446 DOI: 10.1055/s-2007-995618]
29 **Pickhardt PJ**, Nugent PA, Mysliwiec PA, Choi JR, Schindler WR. Location of adenomas missed by optical colonoscopy. *Ann Intern Med* 2004; **141**: 352-359 [PMID: 15353426 DOI: 10.7326/0003-4819-141-5-200409070-00009]
30 **Pickhardt PJ**, Choi JR, Hwang I, Butler JA, Puckett ML, Hildebrandt HA, Wong RK, Nugent PA, Mysliwiec PA, Schindler WR. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N Engl J Med* 2003; **349**: 2191-2200 [PMID: 14657426 DOI: 10.1056/NEJMoa031618]
31 **Pabby A**, Schoen RE, Weissfeld JL, Burt R, Kikendall JW, Lance P, Shike M, Lanza E, Schatzkin A. Analysis of colorectal cancer occurrence during surveillance colonoscopy in the dietary Polyp Prevention Trial. *Gastrointest Endosc* 2005; **61**: 385-391 [PMID: 15758908 DOI: 10.1016/S0016-5107(04)02765-8]
32 **Postic G**, Lewin D, Bickerstaff C, Wallace MB. Colonoscopic miss rates determined by direct comparison of colonoscopy with colon resection specimens. *Am J Gastroenterol* 2002; **97**: 3182-3185 [PMID: 12492208 DOI: 10.1111/j.1572-0241.2002.07128.x]
33 **Rex DK**, Bond JH, Feld AD. Medical-legal risks of incident cancers after clearing colonoscopy. *Am J Gastroenterol* 2001; **96**: 952-957 [PMID: 11316211 DOI: 10.1111/j.1572-0241.2001.03677.x]
34 **Plumeri PA**. Colonoscopic gamble. *Am J Gastroenterol* 2001; **96**: 946-947 [PMID: 11316210 DOI: 10.1111/j.1572-0241.2001.03716.x]
35 **Feld AD**. Malpractice risks associated with colon cancer and inflammatory bowel disease. *Am J Gastroenterol* 2004; **99**: 1641-1644 [PMID: 15330895 DOI: 10.1111/j.1572-0241.2004.40943.x]
36 **Rex DK**. Maximizing detection of adenomas and cancers during colonoscopy. *Am J Gastroenterol* 2006; **101**: 2866-2877 [PMID: 17227527 DOI: 10.1111/j.1572-0241.2006.00905.x]
37 **Rex DK**. Colonoscopic withdrawal technique is associated with adenoma miss rates. *Gastrointest Endosc* 2000; **51**: 33-36 [PMID: 10625792 DOI: 10.1016/S0016-5107(00)70383-X]
38 **Barclay RL**, Vicari JJ, Doughty AS, Johanson JF, Greenlaw RL. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. *N Engl J Med* 2006; **355**: 2533-2541 [PMID: 17167136 DOI: 10.1056/NEJMoa055498]
39 **Sanchez W**, Harewood GC, Petersen BT. Evaluation of polyp detection in relation to procedure time of screening or surveillance colonoscopy. *Am J Gastroenterol* 2004; **99**: 1941-1945 [PMID: 15447753 DOI: 10.1111/j.1572-0241.2004.40569.x]
40 **Simmons DT**, Harewood GC, Baron TH, Petersen BT, Wang KK, Boyd-Enders F, Ott BJ. Impact of endoscopist withdraw-

al speed on polyp yield: implications for optimal colonoscopy withdrawal time. *Aliment Pharmacol Ther* 2006; **24**: 965-971 [PMID: 16948808 DOI: 10.1111/j.1365-2036.2006.03080.x]
41 **Rostom A**, Jolicoeur E. Validation of a new scale for the assessment of bowel preparation quality. *Gastrointest Endosc* 2004; **59**: 482-486 [PMID: 15044882 DOI: 10.1016/S0016-5107(03)02875-X]
42 **Harewood GC**, Sharma VK, de Garmo P. Impact of colonoscopy preparation quality on detection of suspected colonic neoplasia. *Gastrointest Endosc* 2003; **58**: 76-79 [PMID: 12838225 DOI: 10.1067/mge.2003.294]
43 **Froehlich F**, Wietlisbach V, Gonvers JJ, Burnand B, Vader JP. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. *Gastrointest Endosc* 2005; **61**: 378-384 [PMID: 15758907 DOI: 10.1016/S0016-5107(04)02776-2]
44 **Haidry R**, Butt MA, Lovat LB. Advances in diagnostic endoscopy. *Medicine* 2011; **39**: 279-283 [DOI: 10.1016/j.mpmed.2011.02.006]
45 **Lieberman D**. Quality and colonoscopy: a new imperative. *Gastrointest Endosc* 2005; **61**: 392-394 [PMID: 15758909 DOI: 10.1016/S0016-5107(05)00133-1]
46 **East JE**, Saunders BP, Burling D, Boone D, Halligan S, Taylor SA. Surface visualization at CT colonography simulated colonoscopy: effect of varying field of view and retrograde view. *Am J Gastroenterol* 2007; **102**: 2529-2535 [PMID: 17640320 DOI: 10.1111/j.1572-0241.2007.01429.x]
47 **Brenner H**, Chang-Claude J, Seiler CM, Stürmer T, Hoffmeister M. Does a negative screening colonoscopy ever need to be repeated? *Gut* 2006; **55**: 1145-1150 [PMID: 16469791 DOI: 10.1136/gut.2005.087130]
48 **Le Bodic L**, Cerbelaud C, Bouchand S, Auffret N, Clément A, Le Bodic MF. Follow-up of a cohort of 2604 colorectal adenomas treated in 1991 and 1992. Search for parameters related to adenomatous recurrence. *Gastroenterol Clin Biol* 2003; **27**: 466-470 [PMID: 12843910]
49 **Rocha Ramírez JL**, Peña JP, Franco Gutiérrez JR, Villanueva Sáenz E. Colonic adenoma: risk factors for their malignant transformation. *Rev Gastroenterol Mex* 1996; **61**: 178-183 [PMID: 9102738]
50 **Mamula P**, Tierney WM, Banerjee S, Desilets D, Diehl DL, Farraye FA, Kaul V, Kethu SR, Kwon RS, Pedrosa MC, Rodriguez SA, Wong Kee Song LM. Devices to improve colon polyp detection. *Gastrointest Endosc* 2011; **73**: 1092-1097 [PMID: 21628010 DOI: 10.1016/j.gie.2011.01.062]
51 **Vemulapalli KC**, Rex DK. Evolving techniques in colonoscopy. *Curr Opin Gastroenterol* 2011; **27**: 430-438 [PMID: 21785352 DOI: 10.1097/MOG.0b013e328349cfc0]

**P- Reviewer:** Costa F, Deutsch JC, Kobayashi N
**S- Editor:** Ji FF **L- Editor:** A **E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.171

*World J Meta-Anal* 2014 November 26; 2(4): 171-178
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.



# Hydrocephalus after subarachnoid hemorrhage: A meta-analytic comparison of aneurysm treatments

Shih-Shan Lang, Matthew R Sanborn, Connie Ju, Akiff Premjee, Sherman C Stein, Michelle J Smith

Shih-Shan Lang, Matthew R Sanborn, Connie Ju, Akiff Premjee, Sherman C Stein, Michelle J Smith, Department of Neurosurgery, University of Pennsylvania Medical Center, Philadelphia, PA 19104, United States

Author contributions: Lang SS, Stein SC and Smith MJ designed the research; Ju C and Premjee A collected the data and assisted in the analysis; Stein SC performed analyses and assisted in the writing; Lang SS, Sanborn MR and Smith MJ wrote most of the paper; all authors participated in editing the manuscript.

Correspondence to: Sherman C Stein, MD, Clinical Professor of Neurosurgery, Department of Neurosurgery, University of Pennsylvania Medical Center, Philadelphia, 3400 Spruce Street, Philadelphia, PA 19104, United States. sherman.stein@uphs.upenn.edu

Telephone: +1-215-4400919 Fax: +1-215-4403798
Received: April 23, 2014 Revised: August 7, 2014
Accepted: August 27, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To compare two treatments for ruptured cerebral aneurysm with reference to the relative risk of developing hydrocephalus.

**METHODS:** We reviewed the English language literature on the risk of developing hydrocephalus after aneurysm treatment. Data were divided by type of study (randomized controlled trial, cohort trial, nonrandomized comparison, prospectively- and retrospectively-collected observational study). They were also divided by type of aneurysm treatment (microvascular - clipping, or endovascular - coiling). Additional predictive variables collected for each publication were average age, gender distribution, measures of hemorrhage volume and subarachnoid hemorrhage severity, aneurysm locations, time to treatment, duration of follow-up and date of publication. We employed meta-analysis to calculate pooled risk ratios of developing hydrocephalus in cases receiving aneurysm clipping *vs* those receiving coiling. Meta-regression was used to correct pooled results for covariates.

**RESULTS:** Because indications for the two treatments are different, there is little clinical equipoise for treating most cases. The single randomized, controlled trial dealt with a small subset of ruptured aneurysms. Neither this nor pooled values from other studies which compared the two treatments had the power to demonstrate significant differences between the two treatments. Nor was there an apparent difference when observational data were meta-analytically pooled. However, when meta-regression was used to correct for predictive variables known to differ between the two treatment groups, a highly-significant difference appeared. Coiling is used more commonly in older, sicker patients with aneurysms in certain locations. These cases are more likely to develop hydrocephalus. When corrected for these covariates, the risk of hydrocephalus was found to be significantly lower in coiled *vs* clipped cases ($P$ = 0.014).

**CONCLUSION:** Pooled observational data were necessary to demonstrate that coiling ruptured cerebral aneurysms is associated with a lower risk of developing hydrocephalus than is clipping.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Subarachnoid hemorrhage; Cerebral aneurysm; Hydrocephalus; Meta-analysis; Meta-regression; Observational data

**Core tip:** Several treatment comparisons in clinical medicine are not amenable to randomized controlled trials. The conditions may be too rare for trials to obtain adequate statistical power. There may be a lack of clinical equipoise on the part of patients or clinicians. Comparisons of different treatments are nevertheless still important and can only be addressed by pooling observational data.

Lang SS, Sanborn MR, Ju C, Premjee A, Stein SC, Smith

MJ. Hydrocephalus after subarachnoid hemorrhage: A meta-analytic comparison of aneurysm treatments. *World J Meta-Anal* 2014; 2(4): 171-178 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/171.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.171

## INTRODUCTION

Cerebral aneurysms are small spherical outpouchings of the arteries that run along the surface of the brain. Affecting approximately 3.2% of adults[1], they are usually clinically silent and cause no symptoms. However, as its weakened wall stretches, the aneurysm is prone to rupture, bleeding into the subarachnoid space that surrounds the brain[2]. This subarachnoid hemorrhage occurs annually in about 9 per 100000 population worldwide, many countries experiencing much higher rates[3]. The most common cause of nontraumatic subarachnoid hemorrhage (SAH), ruptured cerebral aneurysm, is associated with high mortality and morbidity, and is considered an important cause of disabling stroke. One common complication of SAH is hydrocephalus, enlarging of the brain's fluid spaces[4]. This is thought to be caused by the blood and the inflammation it incites clogging the pathways and preventing circulation and absorption of cerebrospinal fluid (CSF). Hydrocephalus is extremely common in the first days after SAH; as many as 30% of hydrocephalus cases are permanent and require long-term treatment.

Ruptured aneurysms are commonly treated acutely, in order to prevent repeat hemorrhage. One standard treatment is craniotomy with application of a spring-loaded clip across the neck of the aneurysm. Usually this procedure is performed with the aid of an operative microscope and is termed "microvascular" or "clipping". The other technique involves threading a catheter within the artery from the groin to the aneurysm and filling the aneurysm sac with tiny platinum coils. Termed "endovascular" or "coiling", this approach induces the blood within the aneurysm to clot. Both approaches are usually successful in preventing aneurysms from bleeding again. The choice of which procedure is most appropriate often varies according to the characteristics of the aneurysm and the patient[5,6]. Recent guidelines have been published to establish the relative strengths and weaknesses of both approaches[7].

One unresolved issue is whether one approach is more successful in preventing chronic hydrocephalus. Each has its strong points. Microvascular surgery promotes dissipation of the blood surrounding the aneurysm. On the other hand, the brain exposure and manipulation involved in microvascular surgery may itself predispose to hydrocephalus, a problem the endovascular approach avoids.

Clinical trials reporting incidence of hydrocephalus after each aneurysm treatment have reported wide variations in the incidence of chronic hydrocephalus. Studies reporting outcomes from both approaches have measured conflicting relative risks. Since there are few controlled clinical trials, it is apparent that meta-analytic pooling of results is necessary to predict the relative risks associated with the two aneurysm treatments. Point estimates of the actual probability of hydrocephalus are often required for decision analyses and cost-effectiveness determinations. Since these values may be difficult to determine from meta-analytically pooled risk ratios, Einarson proposed meta-analysis of observational data[8]. However, this and similar approaches to meta-analysis have been criticized as precise but not necessarily valid[9,10]. Our study compares analysis of trials, in which both treatments are compared, with observational data to predict relation between treatment and risk of chronic hydrocephalus

## MATERIALS AND METHODS

We performed a structured search in Medline, EMBASE and the Cochrane database of articles published in English before March 2013 containing a combination of the terms "subarachnoid", "hemorrhage/haemorrhage", "intracranial aneurysm", "treatment" and "hydrocephalus," either as keywords or in the text or title. We supplemented our search using the "Related Citations" feature of PubMed and manual reviews of the bibliographies of selected articles. Abstracts of returned articles were examined for a likelihood of yielding useful data. Unlikely abstracts were excluded; the others were reviewed by at least two authors. We excluded series reporting fewer than 10 treated cases. We also excluded publications that reviewed or duplicated data from other reports or dealt only with special cases, such as giant or pediatric aneurysms. For purposes of this study, we defined "hydrocephalus" as the placement of a permanent ventricular shunt at any time in the acute or follow-up period. The articles were divided into two groups, those treated by microvascular clip occlusion and those with endovascular coil occlusion. If an article reported on both approaches, it was included in the analysis only if results of the two treatment groups were reported separately. We included only cases in which a permanent CSF shunt was placed; acute but transient hydrocephalus was not included.

In addition to treatment and hydrocephalus incidence, the following categories of data were also extracted: (1) number of cases surviving acute care; (2) median Fisher grade (FG) a measure of hemorrhage volume, median Hunt-Hess grade (HHG) a measure of the hemorrhage's clinical severity; (3) mean age; (4) gender distribution; (5) ruptured aneurysm location and diameter; (6) proportion of microvascular cases with attempted clot removal; (7) mean time to treatment; (8) duration of follow-up; and (9) date of publication.

We used meta-analytic technique to calculate pooled risk ratios from randomized controlled trials (RCTs), non-randomized cohort studies, and retrospective reviews of case series that reported results from both treatment approaches. In addition we pooled observational data meta-analytically, creating pooled means and distributions for each variable {Einarson, 1997 #36}, calculating risk ratios from the pooled results.

**Table 1 Series used for analysis**

| Series | Trial type | Treatment | No. of cases | Hydrocephalus incidence |
|---|---|---|---|---|
| Akyuz *et al*[21] | Ret | Clip | 145 | 6.9% |
| Andaluz *et al*[22] | Pro | Clip | 106 | 6.6% |
| Auer *et al*[23] | Ret | Clip | 138 | 6.5% |
| Bailes *et al*[24] | Ret | Clip | 27 | 96.3% |
| Chan *et al*[25] | Ret | Clip | 89 | 57.3% |
| de Oliveira *et al*[26] | Ret | Clip | 212 | 17.5% |
| Dehdashti *et al*[12] | Cohort | Both | 245 | 13.9% |
| Diaz *et al*[27] | Ret | Coil | 31 | 15.5% |
| Dorai *et al*[28] | Ret | Clip | 651 | 23.3% |
| Fanning *et al*[29] | Ret | Coil | 178 | 1.7% |
| Ferch *et al*[30] | Ret | Clip | 73 | 53.4% |
| Fujimura *et al*[31] | Ret | Clip | 39 | 17.9% |
| Fukuhara *et al*[32] | Pro | Clip | 62 | 9.7% |
| Gao *et al*[33] | Ret | Coil | 221 | 4.1% |
| Gilsbach *et al*[34] | Ret | Clip | 138 | 23.2% |
| Graff-Radford *et al*[35] | Ret | Clip | 3521 | 8.9% |
| Gruber *et al*[36] | Ret | Both | 187 | 21.4% |
| Harrigan *et al*[37] | Ret | Both | 349 | 41.2% |
| Hernesniemi *et al*[38] | Ret | Clip | 57 | 31.6% |
| Hirashima *et al*[39] | Ret | Clip | 114 | 34.2% |
| Hoh *et al*[40] | Ret | Both | 20585 | 5.4% |
| Hornyak *et al*[41] | Ret | Clip | 92 | 20.7% |
| Jartti *et al*[42] | Ret | Both | 209 | 30.1% |
| Joakimsen *et al*[43] | Pro | Clip | 47 | 10.6% |
| Kaku *et al*[44] | Ret | Both | 100 | 18.0% |
| Kang *et al*[45] | Pro | Coil | 76 | 5.3% |
| Kasuya *et al*[46] | Ret | Clip | 108 | 29.6% |
| Kolluri *et al*[47] | Ret | Clip | 500 | 4.2% |
| Komotar *et al*[48] | Ret | Clip | 582 | 10.3% |
| Krayenbühll *et al*[49] | Ret | Clip | 218 | 10.1% |
| Kung *et al*[50] | Ret | Coil | 102 | 30.4% |
| Kwon *et al*[51] | Ret | Coil | 107 | 12.1% |
| Liew *et al*[52] | Ret | Clip | 13 | 69.2% |
| Lin *et al*[53] | Ret | Clip | 147 | 15.6% |
| Milhorat *et al*[54] | Ret | Clip | 29 | 20.7% |
| Moriyama *et al*[55] | Ret | Clip | 58 | 32.8% |
| Nam *et al*[56] | Ret | Both | 736 | 19.7% |
| Natarajan *et al*[57] | Ret | Both | 192 | 14.3% |
| Ohwaki *et al*[58] | Ret | Clip | 74 | 33.8% |
| O'Kelly *et al*[59] | Ret | Both | 3120 | 18.7% |
| Park *et al*[60] | Ret | Clip | 94 | 13.8% |
| Park *et al*[61] | Ret | Coil | 118 | 5.9% |
| Pinsker *et al*[62] | Ret | Clip | 31 | 48.4% |
| Raimondi *et al*[63] | Ret | Clip | 30 | 13.3% |
| Säveland *et al*[64] | Pro | Clip | 247 | 1.2% |
| Schmieder *et al*[65] | Ret | Clip | 138 | 10.9% |
| Sethi *et al*[66] | Ret | Both | 100 | 6.0% |
| Sindou *et al*[67] | Ret | Clip | 197 | 12.7% |
| Spallone *et al*[68] | Ret | Clip | 125 | 7.2% |
| Stenhouse *et al*[69] | Ret | Clip | 27 | 11.1% |
| Sundt *et al*[70] | Ret | Clip | 616 | 9.4% |
| Taha *et al*[71] | Ret | Both | 133 | 22.5% |
| Tapaninaho *et al*[72] | Ret | Clip | 835 | 9.7% |
| Theander *et al*[73] | Ret | Clip | 16 | 6.3% |
| Tomasello *et al*[74] | Pro | Clip | 47 | 4.3% |
| Tsementzis *et al*[75] | Pro | Clip | 61 | 6.6% |
| Vale *et al*[76] | Ret | Clip | 90 | 20.0% |
| Vanninen *et al*[11] | RCT | Both | 106 | 13.2% |
| Varelas *et al*[77] | Ret | Both | 163 | 7.3% |
| Vassilouthis *et al*[78] | Ret | Clip | 181 | 6.1% |
| Widenka *et al*[79] | Ret | Clip | 283 | 18.4% |
| Williams *et al*[80] | Ret | Clip | 19 | 42.1% |
| Yang *et al*[81] | Ret | Both | 88 | 25.0% |
| Yasui *et al*[82] | Ret | Clip | 82 | 46.3% |
| Yoshioka *et al*[83] | Ret | Clip | 576 | 37.3% |

RCT: Randomized controlled trial; Pro: Prospective data collection; Ret: Retrospective; Clip: Microvascularapproach; Coil: Endovascular approach; Both: Clipping and coiling procedures reported in series.

To assess the effects of follow-up length, date of publication and other covariates on results, we employed meta-regression {Thompson, 2002 #39}. Statistical associations between a predictor variable such as type of aneurysm treatment and hydrocephalus incidence assume that treatment was independent of other influences. If there were other predictor variables that might interact with treatment (covariates), their contributions were analyzed using multiple regression analysis. In the case of pooled data, this required meta-regression. Meta-analytic pooling, meta-regressions and statistical analyses were done using Stata 12 (StataCorp LP; College Station, Texas).

## RESULTS

The search returned 857 abstracts of which 303 were relevant to our study. Some published series were excluded from our analysis since they reported only on elderly, pediatric or other specialized groups of patients. Figure 1 illustrates our search results, and Table 1 summarizes the publications used in our analysis, along with the study type, treatments used and incidence of hydrocephalus. In all, there were 65 publications reporting 38081 cases surviving acute care. Note that 14 papers (termed "comparative studies") reported outcomes for both clipped and coiled patients separately. There was only one RCT[11] and one prospectively collected cohort study[12]. The other 12 papers involved retrospective data review of cases having one procedure or the other. This suggests that the bulk of the evidence was Level 4[13]. The incidence of hydrocephalus varied from 1 to 96% in individual series.

We compared the risk ratios (endovascular/microvascular) reported by the single RCT, the cohort study, the pooled values of the 14 comparative studies and pooled value of all 65 publications treated as observational studies. The results are shown in Table 2. The pooled mean incidence of chronic hydrocephalus, calculated from all sources, is 0.191 (± 0.023). The probability of hydrocephalus was lower after coiling than clipping (0.161 ± 0.199 and 0.209 ± 0.115, respectively). This corresponds to a risk ratio of 0.770 (relative risk = 0.726, absolute difference 0.048), the difference being non-significant ($P$ = 0.289) on the Student's $t$ test.

For the observational analysis we also investigated the influence of variables other than treatment on hydrocephalus incidence. Table 3 reports the results of simple linear meta-regression for the associations between hydrocephalus incidence and the other measured variables, demonstrating several other significant differences between the two groups. These include HHG, time to treatment, and patient age, the last of which is

**Table 2 Risk ratios by study type**

| Study type | Number of studies | Risk ratio (coil/clip) | 95%CI |
|---|---|---|---|
| RCT | 1 | 0.286 | 0.008-1.011 |
| Cohort study | 1 | 1.512 | 0.956-2.068 |
| All comparative studies | 14 | 1.150 | 0.982-1.318 |
| Observational studies | 65 | 0.770 | 0.539-1.005 |

RCT: Randomized controlled trial; Clip: Microvascularapproach; Coil: Endovascular approach.

**Table 3 Association between variables and incidence of hydrocephalus**

| Variable | B coefficient | 95%CI | *P* value |
|---|---|---|---|
| Treatment[1] | -0.048 | -0.131-0.034 | 0.249 |
| Mean hunt-hess grade | 0.146 | 0.064-0.228 | 0.001 |
| Mean fisher grade | 0.015 | -0.106-0.136 | 0.805 |
| Mean age (yr) | 0.010 | 0.002-0.018 | 0.019 |
| % Female patients | -0.047 | -0.544-0.451 | 0.852 |
| % Vertebrobasilar aneurysms | 0.001 | -0.002-0.004 | 0.434 |
| Time to treatment | -0.011 | -0.021-0.0004 | 0.042 |
| Duration of follow-up | -0.000 | -0.004-0.003 | 0.833 |
| Year of publication | 0.001 | -0.002-0.005 | 0.455 |

[1]Clipping scored 1, coiling 2.

shown in Figure 2. Of note, there was no significant association between date of publication and hydrocephalus incidence, suggesting no progress has been made over time to lower its risk. In addition, neither sex distribution nor the mean FG appeared to influence hydrocephalus occurrence. Other variables, such as the presence of vasospasm, intraventricular hemorrhage (IVH), removing the subarachnoid clot surgically or with antifibrinolytics, acute hydrocephalus and aneurysm diameter lacked sufficient pooled data to permit meaningful analysis.

We cannot assume that all the predictor variables for hydrocephalus are independent. Since the treatment was not chosen randomly, we must consider whether characteristics of the patients or the hemorrhages influenced choice of treatment. Statistical comparisons revealed significant differences between clipped and coiled cases regarding patient age ($P = 0.045$), % females ($P = 0.038$), and % vertebrobasilar aneurysms ($P = 0.018$). In all cases, the higher values were in the coiled group. Although the lower risk of hydrocephalus in coiled patients was not significant in single regression, adjusting with multivariable meta-regression for these covariates (age, sex, aneurysm location) revealed that treatment became a highly significant independent predictor of hydrocephalus ($P = 0.014$). HHG ($P = 0.001$) and age ($P = 0.032$) remained the only other independent significant predicting variables. Correcting the comparative series for covariates did not alter the non-significant associated between treatment type and hydrocephalus incidence.

## DISCUSSION

Meta-analytic techniques to pool observational data and

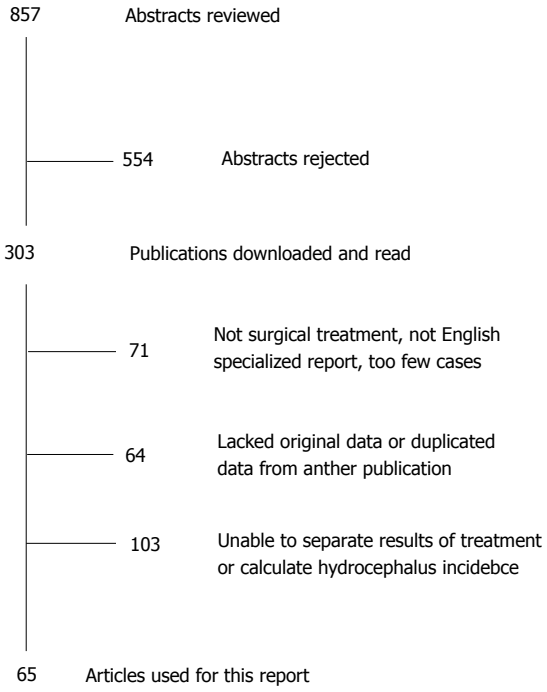



**Figure 1 Articles reviewed for this study**

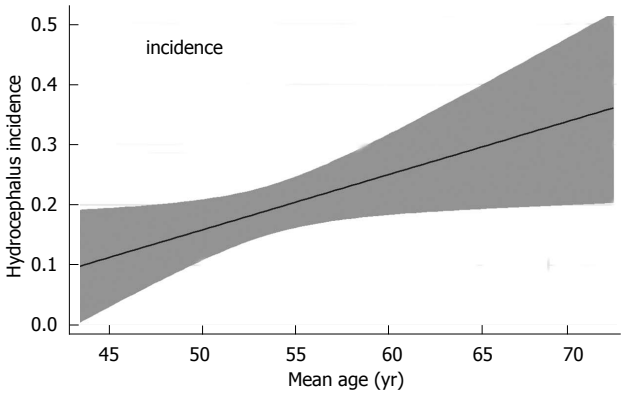



**Figure 2 Influence of age on hydrocephalus incidence.** In the meta-regression, the line represents the fitted mean and the gray are the 95%CI.

correct for covariates enabled us to demonstrate a lower risk of hydrocephalus following endovascular aneurysm repair compared to microvascular treatment. This knowledge may aid aneurysm surgeons in selecting the ideal approach to individual patients after subarachnoid hemorrhage.

The task is complicated by a strong bias in selecting patients for each surgical approach. For example, older and sicker patients and those harboring aneurysms in certain locations are more likely to be selected for endovascular over microvascular surgery. The single RCT[11] was limited to the small subset of cases in which clinical equipoise was thought to exist. Not only is this report of limited generalizability to the ruptured aneurysm population, it is too small to demonstrate significance. Nor is a significant difference obtained by pooling results from non-randomized trials in which the two approaches are compared. Only the larger numbers obtained by treating

all studies as observational allowed the underlying pattern to emerge. Surgeons are more likely to employ coiling to treat aneurysms in older, sicker patients and those whose aneurysms are in the vertebrobasilar circulation, all of which have been shown to predispose to hydrocephalus. Correcting for these covariates allowed us to demonstrate that the risk of hydrocephalus was significantly lower after endovascular coiling than microvascular clipping. Among attempts to extend the utility of meta-analysis in clinical medicine[14,15], employing observational data has been reported as a valuable addition[8,16-18]. Not only can it be used to generate pooled probabilities, but, as here, to explore possible causal relationships in situations in which RCTs have not or cannot be done.

This study has several important limitations. It could be argued that possible selection bias creates two overlapping but distinct populations, for which any statistical comparison is invalid. The role of meta-analysis for observational data has been criticized as introducing possible confounding and heterogeneity[9], and there are situations in which it creates misleading conclusions[10]. Nevertheless, guidelines have been established in reporting observational meta-analyses[19] along the lines of the QUOROM specifications[20].

In a conclusion, we propose that analysis of observational data is necessary to display a significant association between treatment choice for subarachnoid hemorrhage and subsequent hydrocephalus. We used meta-analysis to create pooled estimates of frequency and meta-regression to correct for covariates.

## COMMENTS

### Background

Cerebral aneurysms are common causes of hemorrhagic stroke, and hydrocephalus is a common complication of a ruptured aneurysm. Aneurysm treatment usually consists of occlusion by intracranial clipping or intravascular coiling. There is lack of consensus about which treatment has the lower risk of hydrocephalus. However, because of the different indications for the two treatments, different patient ages and other discordances, a suitable trial has not been done.

### Research frontiers

Authors' present an approach to the problem which employs observational meta-analysis and meta-regression to compare the two procedures in reference to the risk of hydrocephalus and to correct for covariates, such as subject age and aneurysm location.

### Innovations and breakthroughs

Previous comparisons of relative risks and treatment effects were generally limited to randomized clinical trials.

### Applications

This approach permits comparison of risks and effectiveness of treatment strategies from observational data in the absence of randomized clinical trials.

### Terminology

Cerebral aneurysms: small spherical outpouchings of the arteries that run along the surface of the brain. Aneurysm rupture: tear in the thinned wall of the aneurysm. Subarachnoid hemorrhage: bleeding resulting from aneurysm rupture. Hydrocephalus: enlargement of the fluid spaces of the brain secondary to impairment of fluid flow or absorption.

### Peer review

This paper presented a very well-written meta-analysis of published literature aiming to determine whether endovascular coiling or (open) surgical clipping of (cerebral) aneurysm affects incidence of chronic hydrocephalus.

## REFERENCES


1 **Vlak MH**, Algra A, Brandenburg R, Rinkel GJ. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. *Lancet Neurol* 2011; **10**: 626-636 [PMID: 21641282 DOI: 10.1016/S1474-4422(11)]

2 **Chalouhi N**, Hoh BL, Hasan D. Review of cerebral aneurysm formation, growth, and rupture. *Stroke* 2013; **44**: 3613-3622 [PMID: 24130141 DOI: 10.1161/STROKEAHA.113.002390]

3 **de Rooij NK**, Linn FH, van der Plas JA, Algra A, Rinkel GJ. Incidence of subarachnoid haemorrhage: a systematic review with emphasis on region, age, gender and time trends. *J Neurol Neurosurg Psychiatry* 2007; **78**: 1365-1372 [PMID: 17470467 DOI: 10.1136/jnnp.2007.117655]

4 **Germanwala AV**, Huang J, Tamargo RJ. Hydrocephalus after aneurysmal subarachnoid hemorrhage. *Neurosurg Clin N Am* 2010; **21**: 263-270 [PMID: 20380968 DOI: 10.1016/j.nec.2009.10.013]

5 **Li H**, Pan R, Wang H, Rong X, Yin Z, Milgrom DP, Shi X, Tang Y, Peng Y. Clipping versus coiling for ruptured intracranial aneurysms: a systematic review and meta-analysis. *Stroke* 2013; **44**: 29-37 [PMID: 23238862 DOI: 10.1161/STROKEAHA.112.663559]

6 **van der Schaaf I**, Algra A, Wermer M, Molyneux A, Clarke M, van Gijn J, Rinkel G. Endovascular coiling versus neurosurgical clipping for patients with aneurysmal subarachnoid haemorrhage. *Cochrane Database Syst Rev* 2005; **(4)**: CD003085 [PMID: 16235314 DOI: 10.1002/14651858.CD003085.pub2]

7 **Connolly ES**, Rabinstein AA, Carhuapoma JR, Derdeyn CP, Dion J, Higashida RT, Hoh BL, Kirkness CJ, Naidech AM, Ogilvy CS, Patel AB, Thompson BG, Vespa P. Guidelines for the management of aneurysmal subarachnoid hemorrhage: a guideline for healthcare professionals from the American Heart Association/american Stroke Association. *Stroke* 2012; **43**: 1711-1737 [PMID: 22556195 DOI: 10.1161/STR.0b013e3182587839]

8 **Einarson TR**. Pharmacoeconomic applications of meta-analysis for single groups using antifungal onychomycosis lacquers as an example. *Clin Ther* 1997; **19**: 559-569; discussion 538-539 [PMID: 9220219]

9 **Egger M**, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998; **316**: 140-144 [PMID: 9462324]

10 **Smith GD**, Egger M. Meta-analyses of observational data should be done with due care. *BMJ* 1999; **318**: 56 [PMID: 9872898]

11 **Vanninen R**, Koivisto T, Saari T, Hernesniemi J, Vapalahti M. Ruptured intracranial aneurysms: acute endovascular treatment with electrolytically detachable coils--a prospective randomized study. *Radiology* 1999; **211**: 325-336 [PMID: 10228510]

12 **Dehdashti AR**, Rilliet B, Rufenacht DA, de Tribolet N. Shunt-dependent hydrocephalus after rupture of intracranial aneurysms: a prospective study of the influence of treatment modality. *J Neurosurg* 2004; **101**: 402-407 [PMID: 15352596 DOI: 10.3171/jns.2004.101.3.0402]

13 **Oxford Centre for Evidence-Based Medicine** (Levels of evidence 2012). Available from: URL: http://www.cebm.net/index.aspx?o=1025. Accessed March 30, 2014

14 **Follmann D**, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 1992; **45**: 769-773 [PMID: 1619456]

15 **Goodman SN**. Meta-analysis and evidence. *Control Clin Trials* 1989; **10**: 188-204 [PMID: 2666026]

16 **Binder DA**. Methodological Issues in the Meta-Analysis of Observational Studies: Discussion (Statistical Society of Canada – Joint Statistical Meeting, 2010. Accessed April 11, 2014). Available from: URL: http://www.amstat.org/sections/srms/proceedings/y2010/.../205375_56394.pdf

17 **Lewington S**, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002; **360**: 1903-1913 [PMID: 12493255]

18 **Monroe J**. Meta-Analysis for Observational Studies: Statistical Methods for Heterogeneity, Publication Bias and Combining Studies. Los Angeles: Statistics, UCLA, 2007

19 **Stroup DF**, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008-2012 [PMID: 10789670]

20 **Moher D**, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; **354**: 1896-1900 [PMID: 10584742]

21 **Akyuz M**, Tuncer R. The effects of fenestration of the interpeduncular cistern membrane arousted to the opening of lamina terminalis in patients with ruptured ACoA aneurysms: a prospective, comparative study. *Acta Neurochir* (Wien) 2006; **148**: 725-733; discussion 731-732 [PMID: 16489503 DOI: 10.1007/s00701-006-0738-0]

22 **Andaluz N**, Zuccarello M. Fenestration of the lamina terminalis as a valuable adjunct in aneurysm surgery. *Neurosurgery* 2004; **55**: 1050-1059 [PMID: 15509311]

23 **Auer LM**, Mokry M. Disturbed cerebrospinal fluid circulation after subarachnoid hemorrhage and acute aneurysm surgery. *Neurosurgery* 1990; **26**: 804-808; discussion 808-809 [PMID: 2352599]

24 **Bailes JE**, Spetzler RF, Hadley MN, Baldwin HZ. Management morbidity and mortality of poor-grade aneurysm patients. *J Neurosurg* 1990; **72**: 559-566 [PMID: 2319314 DOI: 10.3171/jns.1990.72.4.0559]

25 **Chan M**, Alaraj A, Calderon M, Herrera SR, Gao W, Ruland S, Roitberg BZ. Prediction of ventriculoperitoneal shunt dependency in patients with aneurysmal subarachnoid hemorrhage. *J Neurosurg* 2009; **110**: 44-49 [PMID: 18950263 DOI: 10.3171/2008.5.17560]

26 **de Oliveira JG**, Beck J, Setzer M, Gerlach R, Vatter H, Seifert V, Raabe A. Risk of shunt-dependent hydrocephalus after occlusion of ruptured intracranial aneurysms by surgical clipping or endovascular coiling: a single-institution series and meta-analysis. *Neurosurgery* 2007; **61**: 924-933; discussion 933-934 [PMID: 18091269 DOI: 10.1227/01.neu.0000303188.72425.24]

27 **Diaz RJ**, Wong JH. Clinical outcomes after endovascular coiling in high-grade aneurysmal hemorrhage. *Can J Neurol Sci* 2011; **38**: 30-35 [PMID: 21156426]

28 **Dorai Z**, Hynan LS, Kopitnik TA, Samson D. Factors related to hydrocephalus after aneurysmal subarachnoid hemorrhage. *Neurosurgery* 2003; **52**: 763-769; discussion 769-771 [PMID: 12657171]

29 **Fanning NF**, Willinsky RA, ter Brugge KG. Wall enhancement, edema, and hydrocephalus after endovascular coil occlusion of intradural cerebral aneurysms. *J Neurosurg* 2008; **108**: 1074-1086 [PMID: 18518706 DOI: 10.3171/JNS/2008/108/6/1074]

30 **Ferch R**, Pasqualin A, Barone G, Pinna G, Bricolo A. Surgical management of ruptured aneurysms in the eighth and ninth decades. *Acta Neurochir* (Wien) 2003; **145**: 439-345; discussion 445 [PMID: 12836067 DOI: 10.1007/s00701-003-0037-y]

31 **Fujimura M**, Sugawara T, Seki H, Oku T, Niimura K, Otawara Y, Higuchi H. Perivascular coating with fibrin glue of cerebral arteries in patients with aneurysmal subarachnoid hemorrhage; incidence of chronic hydrocephalus. *Tohoku J Exp Med* 1996; **179**: 267-272 [PMID: 8944428]

32 **Fukuhara T**, Shimizu T, Namba Y. Limited efficacy of endoscopic third ventriculostomy for hydrocephalus following aneurysmal subarachnoid hemorrhage. *Neurol Med Chir* (Tokyo) 2009; **49**: 449-455 [PMID: 19855140]

33 **Gao X**, Liang G, Li Z, Wei X, Hong Q. Complications and adverse events associated with Neuroform stent-assisted coiling of wide-neck intracranial aneurysms. *Neurol Res* 2011; **33**: 841-852 [PMID: 22004708 DOI: 10.1179/1743132811Y.0000000012]

34 **Gilsbach JM**, Harders AG, Eggert HR, Hornyak ME. Early aneurysm surgery: a 7 year clinical practice report. *Acta Neurochir* (Wien) 1988; **90**: 91-102 [PMID: 3354369]

35 **Graff-Radford NR**, Torner J, Adams HP, Kassell NF. Factors associated with hydrocephalus after subarachnoid hemorrhage. A report of the Cooperative Aneurysm Study. *Arch Neurol* 1989; **46**: 744-752 [PMID: 2742543]

36 **Gruber A**, Reinprecht A, Bavinzski G, Czech T, Richling B. Chronic shunt-dependent hydrocephalus after early surgical and early endovascular treatment of ruptured intracranial aneurysms. *Neurosurgery* 1999; **44**: 503-509; discussion 509-512 [PMID: 10069587]

37 **Harrigan MR**, Rajneesh KF, Ardelt AA, Fisher WS. Short-term antifibrinolytic therapy before early aneurysm treatment in subarachnoid hemorrhage: effects on rehemorrhage, cerebral ischemia, and hydrocephalus. *Neurosurgery* 2010; **67**: 935-939; discussion 939-940 [PMID: 20881558]

38 **Hernesniemi J**, Vapalahti M, Niskanen M, Kari A. Management outcome for vertebrobasilar artery aneurysms by early surgery. *Neurosurgery* 1992; **31**: 857-861; discussion 861-862 [PMID: 1279451]

39 **Hirashima Y**, Hamada H, Hayashi N, Kuwayama N, Origasa H, Endo S. Independent predictors of late hydrocephalus in patients with aneurysmal subarachnoid hemorrhage--analysis by multivariate logistic regression model. *Cerebrovasc Dis* 2003; **16**: 205-210 [PMID: 12865606]

40 **Hoh BL**, Kleinhenz DT, Chi YY, Mocco J, Barker FG. Incidence of ventricular shunt placement for hydrocephalus with clipping versus coiling for ruptured and unruptured cerebral aneurysms in the Nationwide Inpatient Sample database: 2002 to 2007. *World Neurosurg* 2011; **76**: 548-554 [PMID: 22251503 DOI: 10.1016/j.wneu.2011.05.054]

41 **Hornyak M**, Gilsbach J, Harders A. Clinical significance of computed tomography in early aneurysm surgery. *Neurochirurgia* (Stuttg) 1991; **34**: 135-140 [PMID: 1745317 DOI: 10.1055/s-2008-1052073]

42 **Jartti P**, Karttunen A, Isokangas JM, Jartti A, Koskelainen T, Tervonen O. Chronic hydrocephalus after neurosurgical and endovascular treatment of ruptured intracranial aneurysms. *Acta Radiol* 2008; **49**: 680-686 [PMID: 18568561 DOI: 10.1080/02841850802050754]

43 **Joakimsen O**, Mathiesen EB, Monstad P, Selseth B. CSF hydrodynamics after subarachnoid hemorrhage. *Acta Neurol Scand* 1987; **75**: 319-327 [PMID: 3618109]

44 **Kaku Y**, Yamashita K, Kokuzawa J, Hatsuda N, Andoh T. Treatment of ruptured cerebral aneurysms - clip and coil, not clip versus coil. *Acta Neurochir Suppl* 2010; **107**: 9-13 [PMID: 19953365 DOI: 10.1007/978-3-211-99373-6_2]

45 **Kang HS**, Han MH, Lee TH, Shin YS, Roh HG, Kwon OK, Kwon BJ, Kim SY, Kim SH, Byun HS. Embolization of intracranial aneurysms with hydrogel-coated coils: result of a Korean multicenter trial. *Neurosurgery* 2007; **61**: 51-58; discussion 58-59 [PMID: 17621018 DOI: 10.1227/01.neu.0000279723.67779.f0]

46 **Kasuya H**, Shimizu T, Kagawa M. The effect of continuous drainage of cerebrospinal fluid in patients with subarachnoid hemorrhage: a retrospective analysis of 108 patients. *Neurosurgery* 1991; **28**: 56-59 [PMID: 1994282]

47 **Kolluri VR**, Sengupta RP. Symptomatic hydrocephalus following aneurysmal subarachnoid hemorrhage. *Surg Neurol* 1984; **21**: 402-404 [PMID: 6701774]

48 **Komotar RJ**, Olivi A, Rigamonti D, Tamargo RJ. Microsurgical fenestration of the lamina terminalis reduces the inci-

dence of shunt-dependent hydrocephalus after aneurysmal subarachnoid hemorrhage. *Neurosurgery* 2002; **51**: 1403-112; discussion 1403-112; [PMID: 12445345]
49 **Krayenbühl HA**, Yaşargil MG, Flamm ES, Tew JM. Microsurgical treatment of intracranial saccular aneurysms. *J Neurosurg* 1972; **37**: 678-686 [PMID: 4654697 DOI: 10.3171/jns.1972.37.6.0678]
50 **Kung DK**, Policeni BA, Capuano AW, Rossen JD, Jabbour PM, Torner JC, Howard MA, Hasan D. Risk of ventriculostomy-related hemorrhage in patients with acutely ruptured aneurysms treated using stent-assisted coiling. *J Neurosurg* 2011; **114**: 1021-1027 [PMID: 20950080 DOI: 10.3171/2010.9.JNS10445]
51 **Kwon BJ**, Chang HW, Youn SW, Kim JE, Han MH. Intracranial aneurysm perforation during endosaccular coiling: impact on clinical outcome, initial occlusion, and recanalization rates. *Neurosurgery* 2008; **63**: 676-672; discussion 682-683 [PMID: 18981878 DOI: 10.1227/01.NEU.0000325500.73330.C2]
52 **Liew D**, Ng PY, Ng I. Surgical management of ruptured and unruptured symptomatic posterior inferior cerebellar artery aneurysms. *Br J Neurosurg* 2004; **18**: 608-612 [PMID: 15799193]
53 **Lin CL**, Kwan AL, Howng SL. Acute hydrocephalus and chronic hydrocephalus with the need of postoperative shunting after aneurysmal subarachnoid hemorrhage. *Kaohsiung J Med Sci* 1999; **15**: 137-145 [PMID: 10224837]
54 **Milhorat TH**. Acute hydrocephalus after aneurysmal subarachnoid hemorrhage. *Neurosurgery* 1987; **20**: 15-20 [PMID: 3808257]
55 **Moriyama E**, Matsumoto Y, Meguro T, Kawada S, Mandai S, Gohda Y, Sakurai M. Combined cisternal drainage and intrathecal urokinase injection therapy for prevention of vasospasm in patients with aneurysmal subarachnoid hemorrhage. *Neurol Med Chir* (Tokyo) 1995; **35**: 732-736 [PMID: 8532128]
56 **Nam KH**, Hamm IS, Kang DH, Park J, Kim YS. Risk of Shunt Dependent Hydrocephalus after Treatment of Ruptured Intracranial Aneurysms : Surgical Clipping versus Endovascular Coiling According to Fisher Grading System. *J Korean Neurosurg Soc* 2010; **48**: 313-318 [PMID: 21113357 DOI: 10.3340/jkns.2010.48.4.313]
57 **Natarajan SK**, Sekhar LN, Ghodke B, Britz GW, Bhagawati D, Temkin N. Outcomes of ruptured intracranial aneurysms treated by microsurgical clipping and endovascular coiling in a high-volume center. *AJNR Am J Neuroradiol* 2008; **29**: 753-759 [PMID: 18184845 DOI: 10.3174/ajnr.A0895]
58 **Ohwaki K**, Yano E, Nakagomi T, Tamura A. Relationship between shunt-dependent hydrocephalus after subarachnoid haemorrhage and duration of cerebrospinal fluid drainage. *Br J Neurosurg* 2004; **18**: 130-134 [PMID: 15176553 DOI: 10.1080/02688690410001680975]
59 **O'Kelly CJ**, Kulkarni AV, Austin PC, Urbach D, Wallace MC. Shunt-dependent hydrocephalus after aneurysmal subarachnoid hemorrhage: incidence, predictors, and revision rates. Clinical article. *J Neurosurg* 2009; **111**: 1029-1035 [PMID: 19361256 DOI: 10.3171/2008.9.JNS08881]
60 **Park BE**. Spontaneous subarachnoid hemorrhage complicated by communicating hydrocephalus: epsilon amino caproic acid as a possible predisposing factor. *Surg Neurol* 1979; **11**: 73-80 [PMID: 451853]
61 **Park HK**, Horowitz M, Jungreis C, Genevro J, Koebbe C, Levy E, Kassam A. Periprocedural morbidity and mortality associated with endovascular treatment of intracranial aneurysms. *AJNR Am J Neuroradiol* 2005; **26**: 506-514 [PMID: 15760857]
62 **Pinsker MO**, Gerstner W, Wolf S, Trost HA, Lumenta CB. Surgery and outcome for aneurysmal subarachnoid hemorrhage in elderly patients. *Acta Neurochir Suppl* 2002; **82**: 61-64 [PMID: 12378993]
63 **Raimondi AJ**, Torres H. Acute hydrocephalus as a complication of subarachnoid hemorrhage. *Surg Neurol* 1973; **1**: 23-26 [PMID: 4784573]
64 **Säveland H**, Hillman J, Brandt L, Edner G, Jakobsson KE, Algers G. Overall outcome in aneurysmal subarachnoid hemorrhage. A prospective study from neurosurgical units in Sweden during a 1-year period. *J Neurosurg* 1992; **76**: 729-734 [PMID: 1564533 DOI: 10.3171/jns.1992.76.5.0729]
65 **Schmieder K**, Koch R, Lücke S, Harders A. Factors influencing shunt dependency after aneurysmal subarachnoid haemorrhage. *Zentralbl Neurochir* 1999; **60**: 133-140 [PMID: 10726336]
66 **Sethi H**, Moore A, Dervin J, Clifton A, MacSweeney JE. Hydrocephalus: comparison of clipping and embolization in aneurysm treatment. *J Neurosurg* 2000; **92**: 991-994 [PMID: 10839260 DOI: 10.3171/jns.2000.92.6.0991]
67 **Sindou M**. Favourable influence of opening the lamina terminalis and Lilliequist's membrane on the outcome of ruptured intracranial aneurysms. A study of 197 consecutive cases. *Acta Neurochir* (Wien) 1994; **127**: 15-16 [PMID: 7942174]
68 **Spallone A**, Pastore FS, Rizzo A, Guidetti B. Low-dose tranexamic acid combined with aprotinin in the pre-operative management of ruptured intracranial aneurysms. *Neurochirurgia* (Stuttg) 1987; **30**: 172-176 [PMID: 2447515 DOI: 10.1055/s-2008-1054089]
69 **Stenhouse LM**, Knight RG, Longmore BE, Bishara SN. Long-term cognitive deficits in patients after surgery on aneurysms of the anterior communicating artery. *J Neurol Neurosurg Psychiatry* 1991; **54**: 909-914 [PMID: 1744646]
70 **Sundt TM**, Kobayashi S, Fode NC, Whisnant JP. Results and complications of surgical management of 809 intracranial aneurysms in 722 cases. Related and unrelated to grade of patient, type of aneurysm, and timing of surgery. *J Neurosurg* 1982; **56**: 753-765 [PMID: 6804608 DOI: 10.3171/jns.1982.56.6.0753]
71 **Taha MM**, Nakahara I, Higashi T, Iwamuro Y, Iwaasa M, Watanabe Y, Tsunetoshi K, Munemitsu T. Endovascular embolization vs surgical clipping in treatment of cerebral aneurysms: morbidity and mortality with short-term outcome. *Surg Neurol* 2006; **66**: 277-284; discussion 284 [PMID: 16935636 DOI: 10.1016/j.surneu.2005.12.031]
72 **Tapaninaho A**, Hernesniemi J, Vapalahti M, Niskanen M, Kari A, Luukkonen M, Puranen M. Shunt-dependent hydrocephalus after subarachnoid haemorrhage and aneurysm surgery: timing of surgery is not a risk factor. *Acta Neurochir* (Wien) 1993; **123**: 118-124 [PMID: 8237488]
73 **Theander S**, Granholm L. Sequelae after spontaneous subarachnoid haemorrhage, with special reference to hydrocephalus and Korsakoff's syndrome. *Acta Neurol Scand* 1967; **43**: 479-488 [PMID: 6081211]
74 **Tomasello F**, d'Avella D, de Divitiis O. Does lamina terminalis fenestration reduce the incidence of chronic hydrocephalus after subarachnoid hemorrhage? *Neurosurgery* 1999; **45**: 827-831; discussion 831-832 [PMID: 10515477]
75 **Tsementzis SA**, Hitchcock ER, Meyer CH. Benefits and risks of antifibrinolytic therapy in the management of ruptured intracranial aneurysms. A double-blind placebo-controlled study. *Acta Neurochir* (Wien) 1990; **102**: 1-10 [PMID: 2407050]
76 **Vale FL**, Bradley EL, Fisher WS. The relationship of subarachnoid hemorrhage and the need for postoperative shunting. *J Neurosurg* 1997; **86**: 462-466 [PMID: 9046303 DOI: 10.3171/jns.1997.86.3.0462]
77 **Varelas P**, Helms A, Sinson G, Spanaki M, Hacein-Bey L. Clipping or coiling of ruptured cerebral aneurysms and shunt-dependent hydrocephalus. *Neurocrit Care* 2006; **4**: 223-228 [PMID: 16757827]
78 **Vassilouthis J**, Richardson AE. Ventricular dilatation and communicating hydrocephalus following spontaneous subarachnoid hemorrhage. *J Neurosurg* 1979; **51**: 341-351 [PMID: 313979 DOI: 10.3171/jns.1979.51.3.0341]
79 **Widenka DC**, Wolf S, Schürer L, Plev DV, Lumenta CB. Fac-

tors leading to hydrocephalus after aneurysmal subarachnoid hemorrhage. *Neurol Neurochir Pol* 2000; **34**: 56-60 [PMID: 11452856]

80 **Williams JP**, Pribram HF, Lynde RH, Sharpe AR. Isotope cisternography in the evaluation of patients with subarachnoid hemorrhage. *J Nucl Med* 1970; **11**: 592-596 [PMID: 5506342]

81 **Yang TC**, Chang CH, Liu YT, Chen YL, Tu PH, Chen HC. Predictors of shunt-dependent chronic hydrocephalus after aneurysmal subarachnoid haemorrhage. *Eur Neurol* 2013; **69**: 296-303 [PMID: 23445755 DOI: 10.1159/000346119]

82 **Yasui N**, Nathal E, Fujiwara H, Suzuki A. The basal interhemispheric approach for acute anterior communicating aneurysms. *Acta Neurochir* (Wien) 1992; **118**: 91-97 [PMID: 1456108]

83 **Yoshioka H**, Inagawa T, Tokuda Y, Inokuchi F. Chronic hydrocephalus in elderly patients following subarachnoid hemorrhage. *Surg Neurol* 2000; **53**: 119-24; discussion 124-125 [PMID: 10713188]

**P- Reviewer:** Algin O, Chu D, Leonardi M, Lesley WS
**S- Editor:** Wen LL **L- Editor:** A **E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.179

*World J Meta-Anal* 2014 November 26; 2(4): 179-185
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.



# Meta-analysis of bivariate *P* values

Mehmet Kocak

Mehmet Kocak, Department of Preventive Medicine, University of Tennessee Health Sciences Center, Memphis, TN 38105, United States
Author contributions: Kocak M solely contributed to this paper.
Correspondence to: Mehmet Kocak, PhD, Assistant Professor, Department of Preventive Medicine, University of Tennessee Health Sciences Center, 66 N. Pauline Street Office 626, Memphis, TN 38105, United States. mkocak1@uhtsc.edu
Telephone: +1-901-4482947 Fax: +1-901-4487041
Received: February 25, 2014 Revised: October 16, 2014
Accepted: October 28, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To propose a new meta-analysis method for bivariate *P* value which account for the paired structure.

**METHODS:** Studies that look to test two different features from the same sample gives rise to bivariate *P* value. A relevant example of this is testing for periodicity as well expression from time-course gene expression studies. Kocak *et al* (2010) uses George and Mudholkar' (1983) "Difference of Two Logit-Sums" method to pool bivariate *P* value across independent experiments, assuming independence within a pair. As bivariate *P* value need not to be independent within a given study, we propose a new meta-analysis approach for pooling bivariate *P* value across independent experiments, which accounts for potential correlation between paired P-values. We compare the "Difference of Two Logit Sums" method with our novel approach in terms of their sensitivity and specificity through extensive simulations by generating *P* value samples from most commonly used tests namely, *Z* test, *t* test, chi-square test, and *F* test, with varying sample sizes and correlation structure.

**RESULTS:** The simulations results showed that our new meta-analysis approach for correlated and uncorrelated bivariate *P* value has much more desirable sensitivity and specificity features compared to the existing method, which treats each member of the paired *P* value as independent. We also compare these meta-analysis approaches on bivariate *P* value from periodicity and expression tests of 4936 *S.Pombe* genes from 10 independent time-course experiments and we showed that our new approach ranks the periodic, conserved, and cycling genes significantly higher, and detects many more periodic, "conserved" and "cycling" genes among the top 100 genes, compared to the 'Difference of Two Logit-Sums' method. Finally, we used our meta-analytic approach to compare the relative evidence in the association of pre-term birth with preschool wheezing versus pre-school asthma.

**CONCLUSION:** The new meta-analysis method has much better sensitivity and specific characteristics compared to the "Difference of Two-Logit Sums" method and it is not computationally more expensive.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Meta-analysis; Bivariate *P* value; Independent experiments; Cell cycle data

**Core tip:** In meta-analysis of bivariate *P* value, keeping the inherent paired structure and thus reserving the correlation between the each member of the paired P-values is critical. In this work, we propose a novel meta-analysis technique which does keep this paired structure intact and thus results in much more favorable sensitivity and specificity characteristics compared to the existing method by George and Mudholkar' (1983), which treats the *P* value as independent.

Kocak M. Meta-analysis of bivariate *P* values. *World J Meta-Anal* 2014; 2(4): 179-185 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/179.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.179

## INTRODUCTION

It is common to combine independent *P* values when pooling data from independent experiments. For meta-

analysis of univariate *P* values, Fisher[1] proposed a simple transformation of *P* values to reach a chi-square distribution. Similarly, Stouffer *et al*[2] used a probit transformation and George *et al*[3] proposed the use of sum of logit transformations of the *P* values. All these approaches exploit the fact that the *P* values has a uniform distribution under the null hypothesis. Each of these approaches also has a weighted version.

Bivariate (or multivariate) *P* values arise when two (or more) different hypotheses are tested on the same data to summarize the evidence for two (or more) features. For example, in cell-cycle gene expression setting, two running hypotheses can be the test for periodicity (FEATURE- Ⅰ) and test for expression (FEATURE- Ⅱ). The main interest here is to show which of the two features has relatively more evidence of significance. This "relative evidence" can provide practical advantages when the researcher wants to pick one factor over the other. For example, in clinical trials, a researcher may want to stratify the patients based on a single diagnostic or prognostic factor, while other significant factors may also be present. In such a case, comparing the "relative evidence" for a given factor (feature) over the others may be quite practical as the researcher may want to choose the factor that has the highest relative evidence of significance for stratification.

As the tests are applied to the same data, some types of correlation are expected between the resulting *P* values. For example, it can be argued that genes that follow a cyclic pattern overtime are more likely to be highly expressed than genes that do not follow a cyclic pattern. Due to this nature of these types of concurrent hypothesis tests, a meta-analytic method that takes the correlation structure between the two *P* values testing (*i.e.*, two different features) using the same dataset is highly needed.

In the following sections, we briefly describe a meta-analysis method for bivariate *P* values by George *et al*[4], which is based on logit transformation of *P* values, followed by a new proposal for meta-analysis of bivariate *P* values. We compare the two approaches using extensive simulations and through an application to the time-course cell cycle gene expression data from 10 independent S.Pombe experiments. We also utilize our meta-analytic approach to compare the relative evidence in the association of pre-term birth with preschool wheezing versus pre-school asthma, and we finish with discussions.

## MATERIALS AND METHODS

### Different of logit sums method

In Kocak *et al*[5], the difference of two logit sums method by George *et al*[4] was used as described in Section 2.3. The test statistic for this approach is a simple difference of the sums of logits of *P* values as follows:

For a sample of pairs of *P* values,

$$\left\{\begin{bmatrix} P_{1j} \\ P_{2j} \end{bmatrix}\right\}_{j=1}^{n},$$

$$T = -(T_1 - T_2) = -\left[\sum_{j=1}^{n} \log\left(\frac{P_{1j}}{1-P_{1j}}\right) - \sum_{j=1}^{n} \log\left(\frac{P_{2j}}{1-P_{2j}}\right)\right]$$

which can be rewritten as

$$T = \sum_{j=1}^{n}\left[\log\left(\frac{P_{1j}}{1-P_{1j}}\right) - \log\left(\frac{P_{2j}}{1-P_{2j}}\right)\right] = -\sum_{j=1}^{n} \log\left(\frac{P_{1j}/(1-P_{1j})}{P_{2j}/(1-P_{2j})}\right).$$

Then under the intersection of null hypotheses,

$$\bigcap_{j=1}^{n} H_{1j}^{0} \quad \text{and} \quad \bigcap_{j=1}^{n} H_{2j}^{0},$$

$$T_1 = \sum_{j=1}^{n} \log\left(\frac{P_{1j}}{1-P_{1j}}\right) \quad \text{and}$$

$$T_2 = \sum_{j=1}^{n} \log\left(\frac{P_{2j}}{1-P_{2j}}\right)$$

are convolutions of logistic random variables. George *et al*[4] computed the exact distribution of such convolutions and showed that this distribution can be accurately approximated using a *t* distribution. Specifically, $T_1$ and $T_2$ are both distributed as

$$\pi\sqrt{\frac{n(5n+2)}{3(5n+4)}}\,(t_{5n+4}).$$

Consequently, under the intersection of null hypotheses,

$$\bigcap_{j=1}^{n} H_{1j}^{0} \quad \text{and} \quad \bigcap_{j=1}^{n} H_{2j}^{0},$$

$T_1 - T_2$ is a convolution of $2n$ logistic random variables. Thus,

$$T = -\sum_{j=1}^{n} \log\left(\frac{P_{1j}/(1-P_{1j})}{P_{2j}/(1-P_{2j})}\right)$$

is accurately approximated by

$$\pi\sqrt{\frac{2n(5n+1)}{3(5n+2)}}\,(t_{10n+4}).$$

Although the above test statistic is in fact the sum of the log of the odds ratios, it is clear that the "paired" nature of the *P* values is broken and irrelevant as the test statistic is invariant to swapping the member of a pair with another member from another independent experiment. We believe that the "paired" structure should not be broken as it preserves the correlation structure between the *P* values intact.

### BiPMeta: A new proposal for meta-analysis of bivariate P values

To keep the "paired" structure of the pairs of *P* values, we employ the following approach: For pairs of *P* values,

$$\left\{\begin{bmatrix} P_{1j} \\ P_{2j} \end{bmatrix}\right\}_{j=1}^{n},$$

representing the evidence for Feature- Ⅰ and Feature- Ⅱ respectively from n independent experiments, we keep the "paired" structure through the test statistic,

$$T_{Meta} = \frac{1}{n}\sum_{j=1}^{n} \frac{\log(P_{1j}/P_{2j})}{P_{1j}+P_{2j}}$$

**Table 1 The estimates of α0, α1, α2, α3 in the approximating cumulative distribution function function**

| Sample size | α0 | α1 | α2 | α3 |
|---|---|---|---|---|
| 3 | -0.7842 | 0.0822 | 4.5194 | 1.4353 |
| 4 | -1.2092 | 0.071 | 6.9087 | 1.5018 |
| 5 | -1.5419 | 0.0639 | 9.6286 | 1.5474 |
| 6 | -1.8161 | 0.0596 | 12.6637 | 1.5761 |
| 7 | -2.0475 | 0.0557 | 15.9471 | 1.6005 |
| 8 | -2.2507 | 0.0528 | 19.555 | 1.6202 |
| 9 | -2.4347 | 0.0511 | 23.4612 | 1.6324 |
| 10 | -2.5823 | 0.0493 | 27.1803 | 1.6403 |

This test statistic has a symmetric probability density function about zero ($\mu_0 = 0$) and has sharply reducing left and right tails as shown in Figure 1 below, where the *P*-values for two features come from two independent uniform distributions [*i.e.*, Beta(1,1) *vs* Beta(1,1)] :Then, we test the following hypotheses: H1:$\mu_0 \leqslant 0$ *vs* H1:$\mu_0 > 0$, where the null hypothesis states that the meta-evidence for Feature- II is at least as much as the meta-evidence for Feature- I whereas the alternative hypothesis states the exact opposite.

After numerous attempts to find an approximating cumulative distribution function (CDF) for $T_{meta}$, we observed that a CDF of the form

$$\begin{cases} 1 - \dfrac{e^{-(a_0 + a_1 t)}}{a_2 + t^{a_3}} & \text{if } t \geqslant 0 \\ \dfrac{e^{-(a_0 + a_1 |t|)}}{a_2 + |t|^{a_3}} & \text{if } t < 0 \end{cases}$$

provides very close approximation to the CDF of our meta-test statistic ($T_{meta}$) under the null hypothesis. For a sample of bivariate *P*-values, we estimate the four parameters involved in this approximating CDF (namely, $\alpha_0$, $\alpha_1$, $\alpha_2$, $\alpha_3$) by fitting a non-linear regression model to the value of the test statistic for 10000000 *P* value samples generated from independent standard uniform distributions as Beta(1,1) and Beta(1,1) with varying sizes. For sample size $n = 5$, Figure 2 illustrates the closeness of such an approximation to the empirical CDF for $T_{meta}$, and we provide the estimates of $\alpha_0$, $\alpha_1$, $\alpha_2$, $\alpha_3$ in Table 1.

### Simulation design

The null case was generated from a pair of beta distributions, [Beta(1,1),Beta(1,1)]. To generate the alternative cases, we used the following strategy: (1) Generate samples of P-values from a pair of *Z* tests, *t* Tests, $\chi^2$ tests, and F-tests. At this point, we have a sample of pairs of *P* values,

$$\left\{\begin{bmatrix} P_{1j} \\ P_{2j} \end{bmatrix}\right\}_{j=1}^{n},$$

where $n = 100000$; (2) Apply the probit transformation

$$\left\{\begin{bmatrix} \Phi^{-1}(P_{1j}) \\ \Phi^{-1}(P_{2j}) \end{bmatrix}\right\}_{j=1}^{n} = \left\{\begin{bmatrix} X_{1j} \\ X_{2j} \end{bmatrix}\right\}_{j=1}^{n}$$

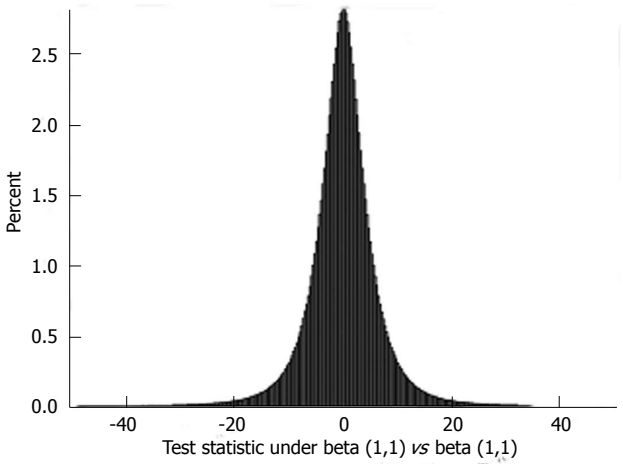



**Figure 1 Probability density function of the test statistics for Bivariate meta-analysis under Beta(1,1) *vs* Beta(1,1) (*i.e.*, two independent uniform distributions) (*n* = 5).**

and compute the corresponding mean and variance,

$[(\tilde{\mu}_1, \tilde{\sigma}_1^2), (\tilde{\mu}_2, \tilde{\sigma}_2^2)]^T$, of $\left\{\begin{bmatrix} X_{1j} \\ X_{2j} \end{bmatrix}\right\}_{j=1}^{n}$;

and (3) Generate two samples of desired size from standard normal distributions,

$$\left\{\begin{bmatrix} Z_{1j} \\ Z_{2j} \end{bmatrix}\right\}_{j=1}^{n},$$

and for each pair,

$$\begin{bmatrix} Z_{1j} \\ Z_{2j} \end{bmatrix},$$

obtain the corresponding correlated pair via

$$\begin{bmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \end{bmatrix} + \left[Chol\begin{bmatrix} \tilde{\sigma}_1^2 & \rho\,\tilde{\sigma}_1\tilde{\sigma}_2 \\ \rho\,\tilde{\sigma}_1\tilde{\sigma}_2 & \tilde{\sigma}_1^2 \end{bmatrix}\right] \times \begin{bmatrix} Z_{1j} \\ Z_{2j} \end{bmatrix}^T$$

for a given ρ.

In our simulations, we set the *r* to be -0.8, -0.5, 0, 0.5, and 0.8, and generated 1000 paired samples of *P* values with sample sizes 5 and 10 based on the bivariate tests listed in Table 2.

## RESULTS

We present the results of the simulations below in Figure 3 for Bivariate *P* value samples of size 3, 5, and 10, respectively, and for the remaining sample sizes ($n = 4, 6, 7, 8, 9$) where the first row of the graph shows the degree of correlation between the *P* values in a given pair.

Clearly, our new meta-analysis approach has much more desirable sensitivity and specificity, which is more pronounced when the correlation between the *P* value pairs gets stronger towards the positive end of the correlation spectrum. It is worth noting that the sensitivity and specificity get weaker for the Difference of Two Logit-sums method as the correlation gets stronger towards positive correlations, while the sensitivity and specificity of our new meta-analysis method increases as the corre-

**Table 2 Simulation design for Bivariate meta-analysis**

| Scenario | Distribution-1 | Distribution-2 | Comment |
|---|---|---|---|
| 1 | Beta(1,1) | Beta(1,1) | Null Case |
| 2 | *Z* test, Δ = 0.5 | *Z* test, Δ = 0.25 | One-sample *Z* test with sample size = 10 |
| 3 | *T* test, Δ = 0.5 | *T* test, Δ = 0.25 | One-sample *t* test with sample size = 10 |
| 4 | $\chi^2$ test, Δ = 1.5 | $\chi^2$ test, Δ = 1.25 | One-sample $\chi^2$ test of variance with sample size = 10 |
| 5 | *F* test, Δ = 0.5 | *F* test, Δ = 0.25 | One-way *ANOVA* with three class-levels of size = 10 |

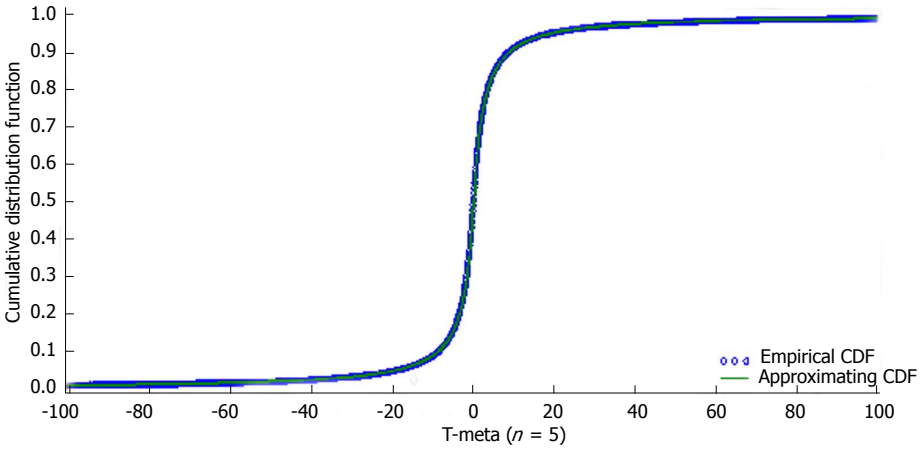



**Figure 2 Empirical cumulative distribution function *vs* approximating cumulative distribution function for Tmeta when *n* = 5.** CDF: Cumulative distribution function.

**Table 3 Performance of the new meta-analysis method compared to the "Difference of Two Logit Sums" method on detecting periodic, conserved and cycling genes**

| Method | Periodic genes by Marguerat *et al*[10] (2006) *n* = 40 | | Conserved genes by Lu *et al*[11] (2007) *n* = 52 | | Cycling genes by Lu *et al*[11] (2007) *n* = 235 | |
|---|---|---|---|---|---|---|
| | Median rank | No. of genes in top 100 genes | Median rank | No. of genes in top 100 genes | Median rank | No. of genes in top 100 genes |
| Difference of two logit sums | 4072 | 0 | 3738 | 2 | 3405 | 5 |
| New meta-analysis method | 102 | 20 | 697 | 13 | 745 | 37 |

lation between the *P* values increases.

### Application to cell-cycle gene expression experiments

Kocak *et al*[6] investigated the cyclic behavior of 4936 genes from 10 independent time-course experiments conducted on Schizosaccharomyces Pombe (*S. Pombe*) yeast cells (Rustici *et al*[7], Oliva *et al*[8], Peng *et al*[9]. We have utilized the P-values from their Empirical Bayes Periodicity test as FEATURE-1 of the bivariate *P* values. Kocak *et al*[5] analyzed data from the same set of experiments in terms of whether or not a given gene is expressed, which served as FEATURE-2 of the bivariate *P* values. In short, we have a pair of *P* values for a given *S. Pombe* gene: one for the testing of periodicity, and the other for the testing for expression. Our aim in this analysis is to identify genes that are "relatively more periodic and being expressed" and we will compare the two meta-analytic approaches in terms of their ability to detect truly periodic genes. To do that, we used a benchmark set of 40 periodic genes reported by Marguerat *et al*[10], and sets of 52 "conserved" genes 235 "cycling" genes reported by Lu *et al*[11].

From Table 3, it is clear that our new meta-analysis approach ranked the periodic genes much higher and detected 20 of the periodic genes in Top 100 while the "Difference of Two Logit Sums" method did not detect any of them in Top 100.

Similarly, the new Meta-analysis method ranked the conserved and cycling genes significantly higher compared to the other three methods, and among the top 100 genes, it detected 13 "Conserved" genes and 37 "Cycling" genes while the "Difference of two Logit Sums" method detected only 2 and 5, respectively.

### Application to pre-term birth data

Sonnenschein-van der Voort *et al*[12] conduct a meta-analysis of 28 independent European studies investigating the association between pre-term birth and pre-school wheezing, and 18 European studies investigating pre-term birth and pre-school asthma. There were 16 studies reporting the results for both pre-school wheezing and asthma. We applied our bivariate meta-analysis approach to assess the relative evidence for the pre-school wheezing compared to the pre-school asthma. The resulting meta-pvalue from those 16 studies is 0.0000054, suggesting that there is significantly more evidence for an association of pre-school wheezing with pre-term birth compared to that of pre-school asthma.

The above authors similarly investigated the association of pre-school wheezing and asthma in relation with low birth weight and there were 14 studies with results

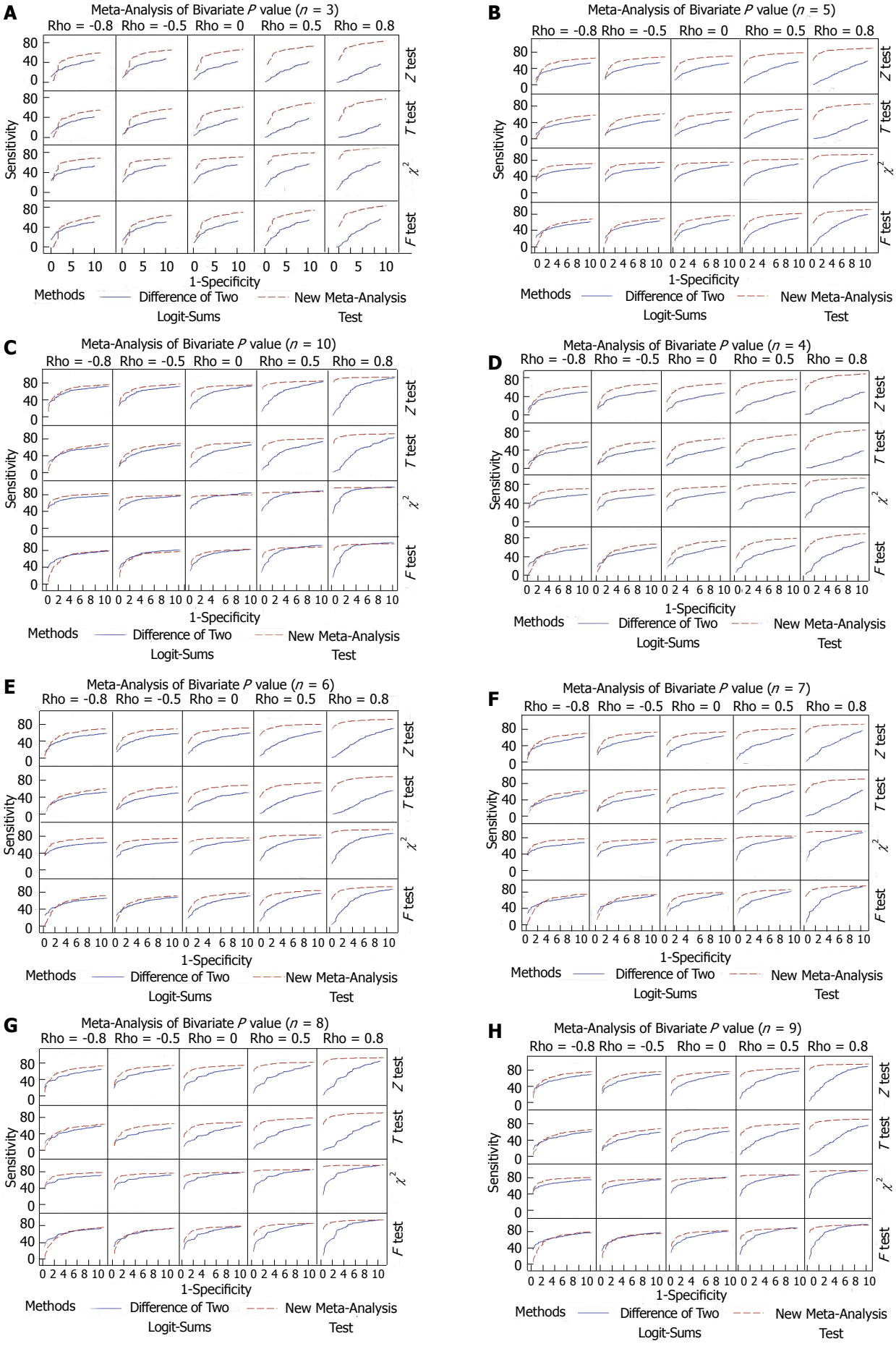



**Figure 3 Meta-analysis of Bivariate *P* values.** A: *n* = 3; B: *n* = 5; C: *n* = 10; D: *n* = 4; E: *n* = 6; F: *n* = 7; G: *n* = 8; H: *n* = 9.

for both pre-school wheezing and asthma. Meta-pvalue based on our meta-analysis method for these 14 studies turned out to be 0.016, suggesting that there exists slightly more evidence for an association between pre-school asthma and low birth weight compared to that of pre-school wheezing.

## DISCUSSION

We proposed a new approach for pooling bivariate *P*-values and we have shown through simulations that the proposed method for bivariate *P* values has much better sensitivity and specificity under varying degree of correlation between the *P* values in a given pair of *P* values. Our new approach for meta-analysis of bivariate *P* values preserves the "paired" structure between the *P* values in a given *P* values pair, which in term keeps the possible correlation within a pair intact.

In application to bivariate *P* values testing for periodicity and expression of *S.Pombe* genes, we have shown that our new meta-analysis method ranked the periodic, "conserved" and "cycling" genes much higher compared to the "Difference of Two Logit Sums" method, which ignores the "paired" nature of the *P* value pairs while this "paired" nature is especially relevant in periodicity and expression testing since it is expected that the periodic genes are expected to be more expressed as the gene expression oscillates as the cell division process moves from a cell cycle phase to another phase.

We also applied our meta-analysis method for a collection of European studies investigating the association between pre-school wheezing and asthma with pre-term birth and as well as low birth weight. We showed that there is relatively more evidence for an association between pre-school wheezing and pre-term birth compared to that of pre-school asthma. This finding supports the results of the meta-analysis conducted by Sonnenschein-van der Voort *et al*[12] who reported a meta odds ratios (95%CI) of 1.34 (1.25, 1.43) for the former association and 1.40 (1.18, 1.67) for the latter association, where it is clear that the former meta-odds ratio has much lower standard error.

Currently, we are using the empirical approximating cumulative distribution function of our meta-analysis test statistic ($T_{meta}$) and an analytic version of the probability density and distribution functions of our test statistic is needed to be identified. We also plan to approach the meta-analysis of bivariate (and multivariate) *P* values on a Bayesian setting without breaking the "paired" (or "joint" in the multivariate case) nature of the *P* values.

## ACKNOWLEDGMENTS

I like to thank Dr. Ebenezer O George of the University of Memphis, and Dr. Stanley Pounds of St. Jude Children's Research Hospital for the helpful discussions on the new meta-analysis method on the results of the simulations.

## COMMENTS

### Background

Meta-analysis of p-values is an efficient way of utilizing and summarizing study results coming from similar experimentation conditions. Meta-analysis of univariate *P* values was of interest for almost a century now and several competing methods were introduced. Meta-analysis of bivariate *P* values did not attract much interest as most research focused on single endpoints and carried out multiple parallel univariable meta-analysis instead of meta-analysis with bivariate or multivariate *P* values.

### Research frontiers

There is no readily available meta-analysis method for bivariate p-values coming from two concurrent hypotheses on the same data, which take into account the correlation structure between the two hypothesis tests.

### Innovations and breakthroughs

The only existing method of the meta-analysis of bivariate p-values assumes that the pairs of *P* values are independent of each other, which is limitation to be statistically addressed as pairs of *P* values stemming from the same study are expected to be correlated. In this manuscript, the authors introduce a new meta-analytic technique that takes into account the correlation structure between the p-values in a given pair.

### Applications

The authors have applied their new method on cell-cycle gene expression data from ten independent studies as well as the pre-term birth outcome data.

### Terminology

Meta-Analysis: Combining statistical results from independent yet similar studies to produce an overall finding in a specific area of research. Bivariate *P* values: A pair of *P* values produced from concurrent hypothesis tests for two associations in the same dataset.

### Peer review

The author showed that the meta-analysis approach has much more desirable sensitivity and specificity, which is more pronounced when the correlation between the *P* values pairs gets stronger towards the positive end of the correlation spectrum (Figure 2).

## REFERENCES

1 **Fisher RA**. Statistical Methods for Research Workers. 14th ed. Edinburg and London: Oliver and Boyd, 1970
2 **Stouffer SA**, Suchman EA, Devinney LC, Star SA, Williams RM. The American soldier: Adjustment ring army life. Princeton, NJ: Princeton University Press, 1949; 1: 43-53
3 **George EO**, Mudholkar GS. (1977). The logit method for combining independent tests. *Institute of Mathematical Statistics Bulletin* 1977; **6**: 212
4 **George EO**, Mudholkar GS. On the Convolution of Logistic Random Variables. *Metrika* 1983; **30**: 1-14 [DOI: 10.1007/BF02056895]
5 **Kocak M**, Zheng G, Narasimhan G, George EO, Pyne S. Differential meta-analysis for testing the relative importance of two competing null hypotheses over multiple experiments. *J Indian Soc Agr Stat* 2010; **64**: 1-10
6 **Kocak M**, George EO, Pyne S, Pounds S. An empirical Bayes approach for analysis of diverse periodic trends in time-course gene expression data. *Bioinformatics* 2013; **29**: 182-188 [PMID: 23172863 DOI: 10.1093/bioinformatics/bts672]
7 **Rustici G**, Mata J, Kivinen K, Lió P, Penkett CJ, Burns G, Hayles J, Brazma A, Nurse P, Bähler J. Periodic gene expression program of the fission yeast cell cycle. *Nat Genet* 2004; **36**: 809-817 [PMID: 15195092 DOI: 10.1038/ng1377]
8 **Oliva A**, Rosebrock A, Ferrezuelo F, Pyne S, Chen H, Skiena S, Futcher B, Leatherwood J. The cell cycle-regulated genes of Schizosaccharomyces pombe. *PLoS Biol* 2005; **3**: e225 [PMID: 15966770 DOI: 10.1371/journal.pbio.0030225]
9 **Peng X**, Karuturi RK, Miller LD, Lin K, Jia Y, Kondu P, Wang L, Wong LS, Liu ET, Balasubramanian MK, Liu J. Identification of cell cycle-regulated genes in fission yeast. *Mol Biol Cell* 2005; **16**: 1026-1042 [PMID: 15616197 DOI: 10.1091/mbc.

E04-04-0299]
10 **Marguerat S**, Jensen TS, de Lichtenberg U, Wilhelm BT, Jensen LJ, Bähler J. The more the merrier: comparative analysis of microarray studies on cell cycle-regulated genes in fission yeast. *Yeast* 2006; **23**: 261-277 [PMID: 16544289 DOI: 10.1002/yea.1351]
11 **Lu Y**, Mahony S, Benos PV, Rosenfeld R, Simon I, Breeden LL, Bar-Joseph Z. Combined analysis reveals a core set of cycling genes. *Genome Biol* 2007; **8**: R146 [PMID: 17650318 DOI: 10.1186/gb-2007-8-7-r146]
12 **Sonnenschein-van der Voort AM**, Arends LR, de Jongste JC, Annesi-Maesano I, Arshad SH, Barros H, Basterrechea M, Bisgaard H, Chatzi L, Corpeleijn E, Correia S, Craig LC, Devereux G, Dogaru C, Dostal M, Duchen K, Eggesbø M, van der Ent CK, Fantini MP, Forastiere F, Frey U, Gehring U, Gori D, van der Gugten AC, Hanke W, Henderson AJ, Heude B, Iñiguez C, Inskip HM, Keil T, Kelleher CC, Kogevinas M, Kreiner-Møller E, Kuehni CE, Küpers LK, Lancz K, Larsen PS, Lau S, Ludvigsson J, Mommers M, Nybo Andersen AM, Palkovicova L, Pike KC, Pizzi C, Polanska K, Porta D, Richiardi L, Roberts G, Schmidt A, Sram RJ, Sunyer J, Thijs C, Torrent M, Viljoen K, Wijga AH, Vrijheid M, Jaddoe VW, Duijts L. Preterm birth, infant weight gain, and childhood asthma risk: a meta-analysis of 147,000 European children. *J Allergy Clin Immunol* 2014; **133**: 1317-1329 [PMID: 24529685 DOI: 10.1016/j.jaci.2013.12.1082]

**P- Reviewer:** Puddu PE **S- Editor:** Ji FF **L- Editor:** A
**E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.186

*World J Meta-Anal* 2014 November 26; 2(4): 186-193
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Observed communication between oncologists and patients: A causal model of communication competence

Katie LaPlant Turkiewicz, Mike Allen, Maria K Venetis, Jeffrey D Robinson

Katie LaPlant Turkiewicz, Department of Communication and Theater Arts, University of Wisconsin - Waukesha, Waukesha, WI 53188, United States
Mike Allen, Department of Communication, University of Wisconsin - Milwaukee, Milwaukee, WI 53201, United States
Maria K Venetis, Brian Lamb School of Communication, Purdue University, West Lafayette, IN 47907, United States
Jeffrey D Robinson, Department of Communication, Portland State University, Portland, OR 97201, United States
Author contributions: All authors contributed to this paper.
Correspondence to: Mike Allen, PhD, Department of Communication, University of Wisconsin - Milwaukee, PO Box 413, Milwaukee, WI 53201, United States. mikealle@uwm.edu
Telephone: +1-414-2294510 Fax: +1-414-2293859
Received: July 29, 2014 Revised: September 16, 2014
Accepted: October 1, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To investigate and test a causal model derived from previous meta-analytic data of health provider behaviors and patient satisfaction.

**METHODS:** A literature search was conducted for relevant manuscripts that met the following criteria: Reported an analysis of provider-patient interaction in the context of an oncology interview; the study had to measure at least two of the variables of interest to the model (provider activity, provider patient-centered communication, provider facilitative communication, patient activity, patient involvement, and patient satisfaction or reduced anxiety); and the information had to be reported in a manner that permitted the calculation of a zero-order correlation between at least two of the variables under consideration. Data were transformed into correlation coefficients and compiled to produce the correlation matrix used for data analysis. The test of the causal model is a comparison of the expected correlation matrix generated using an Ordinary Least Squares method of estimation. The expected matrix is compared to the actual matrix of zero order correlation coefficients. A model is considered a possible fit if the level of deviation is less than expected due to random sampling error as measured by a chi-square statistic. The significance of the path coefficients was tested using a $z$ test. Lastly, the Sobel test provides a test of the level of mediation provided by a variable and provides an estimate of the level of mediation for each connection. Such a test is warranted in models with multiple paths.

**RESULTS:** A test of the original model indicated a lack of fit with the summary data. The largest discrepancy in the model was between the patient satisfaction and the provider patient-centered utterances. The observed correlation was far larger than expected given a mediated relationship. The test of a modified model was undertaken to determine possible fit. The corrected model provides a fit to within tolerance as evaluated by the test statistic, $\chi^2$ (8, average $n$ = 342) = 10.22. Each of the path coefficients for the model reveals that each one can be considered significant, $P < 0.05$. The Sobel test examining the impact of the mediating variables demonstrated that patient involvement is a significant mediator in the model, Sobel statistic = 3.56, $P < 0.05$. Patient active was also demonstrated to be a significant mediator in the model, Sobel statistic = 4.21, $P < 0.05$. The statistics indicate that patient behavior mediates the relationship between provider behavior and patient satisfaction with the interaction.

**CONCLUSION:** The results demonstrate empirical support for the importance of patient-centered care and satisfy the need for empirical casual support of provider-patient behaviors on health outcomes.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Provider-patient communication; Communication competence; Oncologist; Cancer; Causal model; Meta-analysis

**Core tip:** The meta-analysis provides advice about how to deliver the diagnosis of cancer to a patient that promotes more acceptance. The more constructive reaction a patient has to negative news increases adherence and speed of treatment. The focus on communication that is patient-centered creates the basis for improved clinical practice.

Turkiewicz KL, Allen M, Venetis MK, Robinson JD. Observed communication between oncologists and patients: A causal model of communication competence. *World J Meta-Anal* 2014; 2(4): 186-193 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/186.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.186

## INTRODUCTION

A substantial amount of research has been dedicated to illuminating the provider-patient communication dynamic during primary care medical interviews[1]. Effective/competent provider-patient communication has been empirically linked to both positive relational and health outcomes[2-4]. Demonstrable outcomes of competent provider-patient communication include: improved understanding and recall of medical information, improved health outcomes, higher levels of satisfaction, decreased anxiety and depression, improved coping, improved adherence/compliance, and improved perceptions of quality of life and well-being[5,6]. When the context is shifted from primary care medical interviews to consultations between oncologists and cancer patients, a variety of dimensions are affected including: standards of care, consultation length, patient's health status, patient's anxiety level, and the actual health facility[3]. Therefore, the typical provider-patient communication script of discussing medical problems, general patient education, treatment plans, and relationship building is fundamentally altered and confounded by the issues related to cancer treatment[7]. "Physicians are often challenged to communicate bad news, discuss prognosis, and often switch from curative treatment to supportive care and give appropriate and realistic information all along[7]".

The increasing global cancer diagnosis rate[8] combined with the inherent issues related to effective/competent provider-patient communication creates both a research challenge and an opportunity. Similar to research efforts focused on the primary care medical interview, data is becoming increasingly available on communication during a cancer consultation. While not as substantial as primary care medical interactions, there has been enough published data to conduct a meta-analysis of observed (*i.e.*, taped and coded) communication between oncologists and patients[4]. In brief, the results provide preliminary evidence to support meta-analytically derived data for patient-centered oncology care. The authors conclude by calling for more advanced analysis to demonstrate actual effects-pathways between provider behaviors, patient behaviors, and patient satisfaction levels[4]. This requires the formulation of a causal model to determine specific associations and sequential relationships among provider-patient behavioral variables, provider-patient communication variables, and patient satisfaction levels. A causal model with detailed statistical support for which specific provider behaviors promote the best patient outcomes would serve to strengthen the evidence base for the efficacy of patient-centered oncology care and provide physicians with direct guidance on maximizing particular communication behaviors that enhance positive patient outcomes.

The following review of the literature will examine provider-patient oncology communication and the issues and challenges associated with this specific context. A brief review of communication competence theory[9] and its relationship to provider-patient communication will also be provided. Next, each of the components (input, throughput, and output) of the proposed causal model will be examined: (1) input - provider active, provider facilitative, and provider patient-centered behaviors; (2) throughput - patient involvement and patient active behaviors; and (3) output - patient satisfaction. Finally, a related research question will be advanced regarding provider communication competence within an oncology context.

### *Review of the literature*

**Provider-patient oncology communication:** The general features of provider-patient communication in oncology contexts share many of the same features as primary care provider-patient interactions[10]. Some of the key areas of departure are due to the emotionally charged and uncertain nature of cancer diagnoses[10,11]. The relevant literature on cancer communication highlights the problematic features of the oncology context. The amount of information conveyed during cancer consultations is more substantial and complex than primary care visits[10,11]. Additionally, the emotional component of the situation is salient and covers a range of largely negative reactions including anxiety, fear, frustration, and stigma[10-13]. These general informational and emotional barriers provide the foundation for the more specific issues that contribute to the "unique ecology" of cancer communication contexts[4].

In their summary of the literature, Venetis *et al*[4]. identify three specific ways that oncology consultations impact the provider-patient communication dynamic: (1) evidence suggests that oncologists are perceived as less competent communicators than their primary care counterparts and demonstrate more resistance to improving their communication skills; (2) the agenda for oncology visits is quite narrow and focused (*i.e.*, the development of a treatment plan), and does not follow the traditional script of history taking, exam, and diagnosis; and (3) the psychosocial (affective, emotional) *vs* biomedical component of cancer consultations is heightened, corresponding with the emotional issues the patient is facing. Each of these components presents substantial challenges, but they are not insurmountable. Effective and targeted skills

training, based on basic principles of communication competence, can improve the interaction for both provider and patient.

**Communication competence theory:** Communication competence is the ability to choose a communication behavior that is appropriate and effective for a given situation, allowing a person to achieve their communication goals. The model most often used to describe competence is the component model, which includes knowledge, skill, and motivation[9]. Knowledge means knowing what behavior is best suited for a given situation, skill is having the ability to apply that behavior, and motivation is having the desire to communicate in a competent manner. This theory and its components are salient to provider-patient interactions within the oncology context and provide a relevant framework through which to examine communication skills and shortcomings.

The issue of substandard communication competence has been established as a barrier to effective provider-patient communication in the primary care setting[1]. As discussed previously, characteristics of the oncology context exacerbate this issue. According to Cegala *et al*[1], communication competence in a medical setting can be operationalized as, "Communicative moves that effectively advance a participant's goals and, at the same time, reflect understanding and appropriate accommodation of the other's goals" (p.3). Although there may be several goals working in tandem during a medical exchange, the primary goal for both provider and patient is information exchange[1]. Competent information exchange facilitates discussion regarding medical history, the medical problem/issue, diagnosis, treatment plan, and procedures[1].

In a recent study of a communication skills training curriculum based on communication competence theory and specifically designed for oncologists, Bylund *et al*[14] provided a detailed framework divided into six different skill set modules: (1) establishing the consultation framework skills; (2) checking skills; (3) questioning skills; (4) empathic communication skills; (5) information organization skills; and (6) shared decision-making skills. Based upon these categories, oncologists were assessed for their use of these skill sets both prior to and after attending a communication-training program. The results indicated marked increases in the first two skill sets, establishing the consultation framework and checking. Questioning, empathic communication, and information organization skills demonstrated limited changes. The final skill set, shared decision-making, did not demonstrate any change at all. These findings are informative, promising for ongoing communication skills training efforts, and provide clear indications of where improvements need to be made. Oncologists in this sample clearly demonstrated the ability to absorb and put into practice critical communication competence skills. However, the more affective and socioemotional communicative skills demonstrated the least amount of change.

Communication competence theory[9] is a vital component to informed provider communication skills training. Regardless of the framework or specific training approach, the goal of communication competence is the unifying foundation for all physician skills training programs. As recently demonstrated by Bylund *et al*[14], the oncology context is no exception. The communication competence theoretical framework provided a tangible set of skill modules for training to be designed around, a clear indication of what skills were successfully adopted by oncologists, and a reasonable understanding of the skills that require additional training.

**Proposed causal model components:** The explicit and implicit communication challenges associated with the oncology context combined with study results demonstrating both promise and challenges for oncologists in adopting competent communication behaviors, indicate the importance of an enhanced understanding of specific ways oncologists can adopt behaviors with a high likelihood of positively affecting patient outcomes. Recent meta-analysis results confirm the importance of patient-centered oncology care and call for more advanced statistical support of effects pathways among provider behaviors, patient behaviors, and patient outcomes[4]. In an effort to respond to this need, the following investigation advances a proposed causal model with three general components: (1) input; (2) throughput; (3) output. Each of these components includes provider behaviors (*i.e.*, input), patient behaviors (*i.e.*, throughput); and (4) patient outcomes (*i.e.*, output). The following expands on each of the three components and provides detailed explanation of the classification system employed for the casual model.

**Provider communication:** Three distinct types of provider communication behaviors comprise the input component of the casual model: (1) provider active; (2) provider facilitative; and (3) provider patient-centered behaviors. The proposed casual model positions these three behaviors, when employed effectively, as catalysts for positive patient responses and satisfaction outcomes. Provider active is the most straightforward element and represents measureable behaviors that include frequencies and time allotments (*i.e.*, the amount of information provided and duration of time spent with the patient). Provider facilitative is focused on the emotional tone of the provider. Provider patient-centered is less straightforward, as it also centers on emotional elements. However, the key differences between facilitative and patient-centered are the direction of the emotion (*i.e.*, provider to patient) and the element of inclusion the provider extends to the patient. According to Dimoska *et al*[15], "A patient-centered approach is one in which the doctor listens to patients attentively and sympathetically, talks about psychosocial and non-medical issues, appears warm and caring towards the patient rather than hurried, and allows the patient to have input into the consultation" (p. 1508). It is this definition, which attends to the affective, psychosocial, and empowerment needs of the patient that will be employed for the proposed model[4].

Based on the 14 eligible studies[2,3,5,7,10,11,13,15-21] used to contribute to the proposed model, the following measured communication behaviors were categorized accordingly. Provider active behaviors include: (1) amount of provider communication; (2) amount of information given by the provider; (3) duration of the provider's talk; and (4) time spent with the patient. Provider facilitative communication and behaviors are conceptualized to encompass provider: (1) affect/emotionality; (2) friendliness/warmth; (3) depersonalization; (4) emotional responsiveness; (5) empathy; and (6) hostile *vs* friendly behaviors. The third element, provider patient-centered behaviors, encompasses the largest range of communication and behaviors: (1) partnership building; (2) supportive communication; (3) communicative responsiveness; (4) interest/engagement; (5) use of open-ended questions; (6) framing; (7) health-related quality of life concerns; (8) authoritative *vs* affiliative behaviors; (9) psychological exchange; and (10) psychosocial partnership issues.

**Patient communication:** The intermediaries of the proposed causal model are patient-focused and include two elements, patient involvement and patient active communication and behaviors. The relevant literature generally combines these two categories and discusses them more broadly as "patient participation" and more narrowly as "patient question-asking"[16,17,22]. "Patient participation has been operationalized as patients' relative contribution to the conversation, their discussing of experiences with the illness, the number of questions they ask, their expressions of concerns and worries, and their agenda setting talk[22]". Street *et al*[17] provide a more general definition of patient participation, "as the extent to which patients produce verbal responses that have the potential to significantly influence the content and structure of the interaction as well as the health care provider's beliefs and behaviors" (p. 62).

For the purpose of this model, it is critical to be more specific and to parse out patient participation behaviors into the two categories. Similar to the approach taken with the input component of the model, patient involvement is conceptualized from the 14 eligible studies and includes: (1) length of visit/consultation; (2) preference for information; (3) decision style; (4) question-asking; and (5) treatment decisions. Patient active includes: (1) preference for participation; (2) increases in question asking; and (3) patient information-giving.

**Patient satisfaction:** As the least complex and most straightforward component of the proposed model, the output represents positive patient outcomes, namely satisfaction with the interaction. As discussed previously, patient outcomes in oncology contexts are substantial and could include physical and emotional dimensions (*i.e.*, improved health, better treatment adherence, improved perceived quality of life, *etc.*). According to Eide *et al*[7], patient satisfaction is an outcome measure that usually represents a combination of a patients' assessment of their relationship with the provider and the medical support they receive. As a driving and central element to almost all provider-patient communication research, patient satisfaction represents an important outcome, particularly in oncology interviews.

Cancer represents a frightening and uncertainty-producing diagnosis. The belief is that more satisfied patients are more likely to complete treatment as well as experience more successful outcomes from treatment. Based on the studies that contributed to the proposed model, patient satisfaction was conceptualized by patient self-reported anxiety levels post-consultation or expressions of satisfaction with the consultation.

### Summary and research question

It is no longer necessary to argue the importance of effective/competent provider-patient communication, as that has been clearly established in the literature. Similarly, the communication challenges unique to oncology interactions are largely above dispute. The results of a recent meta-analysis establish the importance of patient-centered oncology communication and demonstrate the measurable potential it could have on positive patient outcomes[4]. The remaining piece to this line of research is to provide measurable relationships among these behaviors. These statistics will promote a refined understanding of patient-centered oncology care, identify and isolate the behaviors with the most impact, and inform more directed communication competence skills training programs for oncologists. The driving question is how do all of these components fit together to form an ideal model of physician communication competence?

## MATERIALS AND METHODS

### Data acquisition procedure

A search was made of electronic databases and existing reviews of the literature for relevant manuscripts that met the following criteria: (1) reported an analysis of provider-patient interaction in the context of an oncology interview; (2) the study had to measure at least two of the variables of interest to the model (provider activity-measured by total amount of communication/information, provider patient-centered communication-supportive/interest/engagement/framing/asking for open-ended responses, provider facilitative communication-emotional/personalized, patient activity-amount of participation, patient involvement-question asking/length of visit/amount of discussion, and patient satisfaction or reduced anxiety); and (3) the information had to be reported in a manner that permitted the calculation of a zero-order correlation between at least two of the variables under consideration.

Manuscripts that failed to report adequate information for a statistical estimate were not included. A complete list is available from the first author. In some cases the data report only included enough information to provide estimates for only some of the relevant relationships and was included. The lack of complete statistical information reduced the contribution of that data set to

**Table 1 Description of studies**

| |
|---|
| Butow *et al*[11] (1995). Sydney, Australia. 142 patients. Interactions coded using the CN-LOGIT Interaction Analysis System. Patients completed the psychological adjustment to cancer scale, the functional living index: cancer, profile of mood states, and satisfaction with the consultation scale |
| Detmar *et al*[18] (2001). Amsterdam, Netherlands. 71 patients. Interactions coded using the RIAS. Patients completed health-related quality of life survey |
| Dimoska *et al*[15] (2008). Sydney, Australia. 155 patients. Interactions coded using CANCODE interaction analysis. Patients completed Spielberger State Anxiety Scale, patient satisfaction, and mental adjustment to cancer scale |
| Eggly *et al*[16] (2006). United States. 28 patients. Interactions coded for content by the Neuendorf system. Communication was evaluated using the Karmanos Accrual Assessment System |
| Eide *et al*[7] (2003). Norway. 36 patients. Interactions coded using the Roter Interaction Analysis System. Patients completed a satisfaction instrument |
| Ishikawa *et al*[2] in *Patient Education & Counseling* (2002). Tokyo, Japan. 128 patients from National Cancer Center Hospital. Interactions coded using the RIAS (Roter Interaction Analysis System). Patients evaluated quality of provider communication |
| Ishikawa *et al*[5] in *Social Science & Medicine* (2002). Uses same data as above study |
| Koedoot *et al*[19] (2004). Amsterdam, Netherlands. 140 patients. Interactions coded on the basis of information provided by the provider. Patients completed surveys involving anxiety, depression, preference for participation in decision-making style, and communicative responsiveness of the provider |
| Leighl *et al*[20] (2001). Australia. 101 patients. Interactions coded using coded content analysis system developed by the authors. Patients completed adapted form of the Cassileth Information Styles Questionnaire, patient satisfaction, and Spielberger State Anxiety Scale |
| Ong *et al*[13] (2000). Amsterdam, Netherlands. 96 patients. Interactions coded using the Roter Interaction Analysis System. Patients completed the Rotterdam Symptoms Checklist, and patient satisfaction questionnaire |
| Siminoff *et al*[10] (2000). Cleveland, OH and San Antonio, TX, United States. 50 patients. Interactions coded using the RIAS. Patients completed knowledge comprehension and satisfaction items (VAS - Visual Analog Scales), decisional conflict scale, and a satisfaction with decision scale |
| Street *et al*[3] (2006). United States. 62 patients. Interactions coded using system developed by first author |
| Street *et al*[17] (2001). United States. 9 patients. Interactions coded using system developed by first author |
| Street *et al*[21] (1995). College Station, TX, United States. 60 patients. Interactions coded using a system developed by the first author. Patients completed the perceived involvement in healthcare scale, the perceived decision control instrument, a knowledge test about cancer treatment, and an optimism scale |

**Table 2 Data compilation reported**

| | | | | | |
|---|---|---|---|---|---|
| Provider active | | | | | |
| Provider facilitative | 0.15 (128) | | | | |
| Provider patient-center | 0.10 (328) | 0.18 (71) | | | |
| Patient involvement | 0.31 (531) | 0.37 (458) | 0.16 (426) | | |
| Patient active | 0.39 (128) | 0.37 (220) | 0.20 (208) | 0.34 (323) | |
| Patient satisfaction | 0.14 (450) | 0.18 (670) | 0.23 (732) | 0.23 (158) | 0.16 (251) |

this analysis. A simple editorial and publication solution requires the reporting of a complete zero order matrix among all variables and the data could have been included in this report.

### Data compilation procedure

Data from 14 studies[2,3,5,7,10,11,13,15-21] were transformed into correlation coefficients and corrected (see Table 1 for details of included studies), where applicable, for various measurement artifacts (for procedures see Hunter *et al*[23]). Each effect also has the accompanying sample size that was used to estimate the observed relationship. The effects for each separate correlation were averaged and compiled to produce the correlation matrix used for data analysis reported in Table 2. The correlation matrix provides the average estimate for each separate effect and the corresponding sample size estimate for the combined average effect. The averaging process for the estimation is weighted by sample size to reflect the varying contribution and accuracy of the individual effect.

### Test of the causal model

The test of the causal model is a comparison of the expected correlation matrix generated using an Ordinary Least Squares (OLS) method of estimation. A good model considered an adequate explanation or fit to the available data will meet a variety of tests: (1) overall fit of the model between actual and expected correlations; (2) each path coefficient is significant; (3) each predicted and actual zero order correlation is within sampling error; and (4) each intervening or mediating variable provides a significant level of mediation. A good or sufficient model will meet each of these four tests described in detail below.

The expected matrix is compared to the actual matrix of zero order correlation coefficients. A model is considered a possible fit if the level of deviation is less than expected due to random sampling error as measured by a $\chi^2$ statistic. A nonsignificant chi-square indicates that the level of deviation between the expected and actual matrix is within sampling error, indicating that the model is not inconsistent with the observed data.

The significance of the path coefficients was tested using a *Z* test. The first test only describes whether the proposed model is consistent with the observed data. This test indicates that the sizes of the relationships are considered significant. The importance of this test is that the first test would still work if the observed path coefficients were all zero. The fit of the overall data, while important, does not indicate the value of the proposed causal connections.

The comparison of the actual and expected correla-

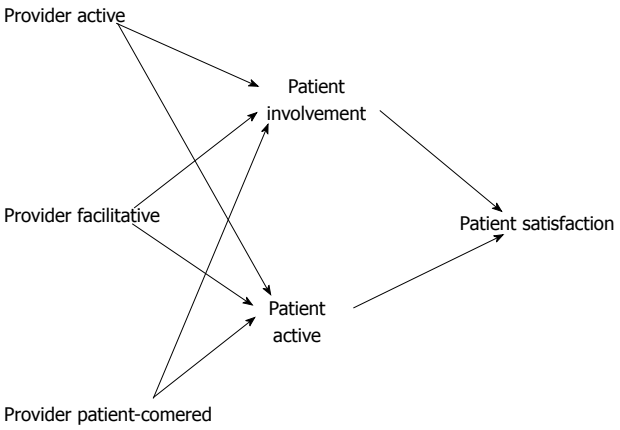



**Figure 1 Proposed causal model.** The proposed model hypothesized the following casual relationship: A patient's perception of provider active, facilitative, and patient-centered behavior (*i.e.*, input) generates an intermediate outcome of patient involvement and activity (*i.e.*, throughput), which in turn increases overall patient satisfaction (*i.e.*, output).

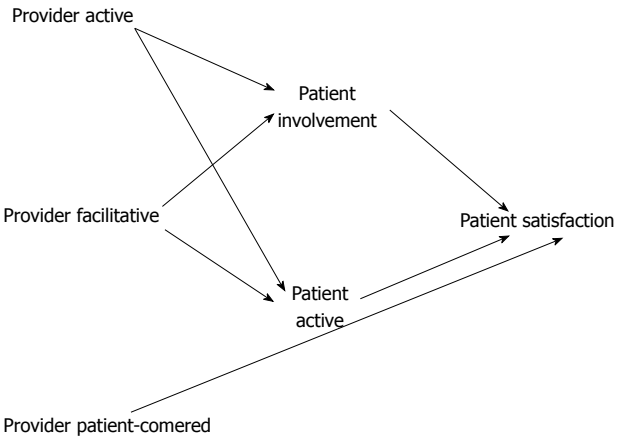



**Figure 2 Final causal model.** The modified model that emerged after statistical analysis: a patient's perception of provider active and facilitative behavior (*i.e.*, input) generates an intermediate outcome of patient involvement and activity (*i.e.*, throughput), which in turn increases overall patient satisfaction (*i.e.*, output). Similarly, the model also demonstrates a direct (rather than mediated) relationship between a patient's perception of a provider's patient-centered behavior and a patient's overall satisfaction.

tions in the zero order matrix employs a $Z$ test. The overall test indicates that the combined errors are less than expected due to sampling error but does not indicate whether all connections are within sampling error. An adequate model would demonstrate fit, within sampling error for all observed correlations compared to the expected correlations given the model.

The Sobel test provides a test of the level of mediation provided by a variable and provides an estimate of the level of mediation for each connection. For example, suppose variable A causes variable B causes variable C. The Sobel test considers the connection between A and C and the impact or strength of the mediating effect of variable B on that relationship. The test provides an estimate of that impact and should demonstrate significant reduction. Such a test is warranted in models with multiple paths.

## RESULTS

### *Test of the original model*

A test of the original model, appearing in Figure 1, indicated a lack of fit with the summary data as reported in Table 2. The first test of the model generated a comparison of the expected and actual matrix that was significant, $\chi^2$ (7, average $n$ = 342) = 32.17. The results indicate that the model is not consistent with the data. Examination of the expected and observed zero order matrix was examined to determine if any identifiable changes could be made to the model to arrive at a more consistent model.

The largest discrepancy in the model was between the patient satisfaction and the provider patient-centered utterances. The observed correlation was far larger than expected given a mediated relationship. The decision was made to test the model without the mediated relationship (*i.e.*, removing patient involvement and patient active) to determine if provider patient-centered utterances would have a direct (rather than mediated) impact on patient satisfaction.

### *Test of a modified model*

The test of a modified model was undertaken to determine possible fit. The corrected model provides a fit to within tolerance as evaluated by the test statistic, $\chi^2$ (8, average $n$ = 342) = 10.22.

Each of the path coefficients for the model displayed in Figure 2 reveals that each one can be considered significant, $P < 0.05$. The path for provider active to patient involvement (0.39) and provider active to patient active (0.31) are significant. The path coefficients from provider facilitative to patient involvement (0.37) and to patient active (0.37) are significant. All three paths from patient involvement (0.16), patient active (0.23), and provider patient-centered utterances (0.23) to patient satisfaction are significant.

A test of the individual discrepancies between expected and observed correlations demonstrates no value-achieved significance, $P < 0.05$. What this indicates is that each of the values between the expected and observed correlation matrix are within sampling error.

The Sobel test examines the impact of the mediating variables (patient involvement and patient active) on the relationship between provider active and provider facilitative communication and patient satisfaction. Patient involvement was demonstrated to be a significant mediator in the model, Sobel statistic = 3.56, $P < 0.05$. Patient active was also demonstrated to be a significant mediator in the model, Sobel statistic = 4.21, $P < 0.05$. The statistics indicate that patient behavior mediates the relationship between provider behavior and patient satisfaction with the interaction.

## DISCUSSION

The proposed model hypothesized the following casual relationship: a patient's perception of provider active, facilitative, and patient-centered behavior (*i.e.*, input) gener-

ates an intermediate outcome of patient involvement and activity (*i.e.*, throughput), which in turn increases overall patient satisfaction (*i.e.*, output). The modified model that emerged after statistical analysis is: a patient's perception of provider active and facilitative behavior (*i.e.*, input) generates an intermediate outcome of patient involvement and activity (*i.e.*, throughput), which in turn increases overall patient satisfaction (*i.e.*, output). Similarly, the model also demonstrates a direct (rather than mediated) relationship between a patient's perception of a provider's patient-centered behavior and a patient's overall satisfaction.

The initial proposed model required slight modification after advanced statistical analysis, indicating that provider patient-centered behaviors were not mediated by any type of patient behavior or involvement. Instead, they simply directly impacted patient satisfaction outcomes. This finding is salient to provider communication competence training efforts, as it demonstrates that providers can learn to adopt specific behaviors that will have direct demonstrable outcomes on their success with patients. Similarly, the model also provides substantial insight and support for specific provider behaviors that promote and support patient involvement.

The data remains limited to a relatively small set of studies. While the quality of the data is high - because it is based on actual recorded observations of interaction as opposed to self-reported recollections - the number of providers and patients in the sample is less than 1000. The types of cancers were limited as well and did not include many different cancers with a variety of lifestyle changes. Similarly, the advice is restricted only to oncology and does not consider a number of other important medical contexts where the delivery of information to a patient is important in influencing subsequent participation in treatment or other aspects of the healthcare system.

The most pressing need for future research is to address the questions of generalizability of the findings beyond the limited application to oncology. The theoretical expectations seem to suggest very strongly that the focus of providers on the patient to facilitate understanding and address concerns seems almost tautological in nature. What the findings clearly do indicate is that providers lacking in communication competence will enable a situation of nonparticipation by the patient. Providers should continue to keep in mind that what is offered is medical advice and advice is only valuable from an expert if it is followed.

The results of this investigation demonstrate additional empirical support for the importance of patient-centered care within the oncology context and begin to satisfy the need for empirical casual support of provider-patient behaviors on overall health outcomes. Furthermore, the results provide guidance and proof of tangible competent behavioral patterns that providers can train for and adopt to improve the emotional and physical outcomes for their patients. A statistical causal assessment of provider-patient communication behaviors within the oncology context through the lens of communication competence theory is a logical continuation of previous related research[4]. The findings of this investigation support and enlarge the existing evidence base for the importance of patient-centered care.

## COMMENTS

### Background

The article proposes a model based on actual observations of doctor/patient interactions involving the delivery of the presence of cancer. The challenge is to find a means of delivering this bad news in a constructive manner that creates a focus on patient participation in available treatments as soon as possible.

### Research frontiers

The challenge for oncologists is to generate a means of patient-centered communication when delivering the results of medical tests. The goal of the provider is to create a full understanding of the medical condition and provide a motivation for the patient to respond in a constructive emotional manner as well as to develop a sense of urgency about participation in treatment options.

### Innovations and breakthroughs

The use of meta-analysis to derive and test the model means that the successful model works consistently across the entire body of research and not simply the individual data set. Establishing a model that can explain an entire set of divergent data creates the basis for generalized scientific knowledge.

### Applications

The study provides advice about how to deliver the diagnosis of cancer to a patient that promotes more acceptance. The more constructive reaction a patient has to negative news increases adherence and speed of treatment. The focus on communication that is patient-centered creates the basis for improved clinical practice.

### Terminology

Meta-analysis refers to a variety of statistical methods used to compare the results from different studies. A causal model is a quantitative representation of real-world dynamics and attempts to describe the causal and other relationships, among a set of multiple variables.

### Peer review

This is an interesting study, indicating that oncologists can directly influence patient satisfaction, by training and adopting specific behavioral patterns.

## REFERENCES


1 **Cegala DJ**, Socha McGee D, McNeilis KS. Components of patients' and doctors' perceptions of communication competence during a primary care medical interview. *Health Commun* 1996; **8**: 1-27 [DOI: 10.1207/s15327027hc0801_1]

2 **Ishikawa H**, Takayama T, Yamazaki Y, Seki Y, Katsumata N. Physician-patient communication and patient satisfaction in Japanese cancer consultations. *Soc Sci Med* 2002; **55**: 301-311 [PMID: 12144143 DOI: 10.1016/S0277-9536(01)00173-3]

3 **Street RL**, Gordon HS. The clinical context and patient participation in post-diagnostic consultations. *Patient Educ Couns* 2006; **64**: 217-224 [PMID: 16859865 DOI: 10.1016/j.pec.2006.02.004]

4 **Venetis MK**, Robinson JD, Turkiewicz KL, Allen M. An evidence base for patient-centered cancer care: a meta-analysis of studies of observed communication between cancer specialists and their patients. *Patient Educ Couns* 2009; **77**: 379-383 [PMID: 19836920 DOI: 10.1016/j.pec.2009.09.015]

5 **Ishikawa H**, Takayama T, Yamazaki Y, Seki Y, Katsumata N, Aoki Y. The interaction between physician and patient communication behaviors in Japanese cancer consultations and the influence of personal and consultation characteristics. *Patient Educ Couns* 2002; **46**: 277-285 [PMID: 11932127 DOI: 10.1016/S0738-3991(01)00164-1]

6 **Siminoff LA**, Graham GC, Gordon NH. Cancer communication patterns and the influence of patient characteristics:

disparities in information-giving and affective behaviors. *Patient Educ Couns* 2006; **62**: 355-360 [PMID: 16860520 DOI: 10.1016/j.pec.2006.06.011]
7 **Eide H**, Graugaard P, Holgersen K, Finset A. Physician communication in different phases of a consultation at an oncology outpatient clinic related to patient satisfaction. *Patient Educ Couns* 2003; **51**: 259-266 [PMID: 14630382 DOI: 10.1016/S0738-3991(02)00225-2]
8 **American Cancer Society**. Cancer facts and figures (2014). Available from: URL: http://www.cancer.org
9 **Spitzberg BH**, Cupach WR. Interpersonal communication competence. 1st ed. Beverly Hills, CA: Sage, 1984: 26-29
10 **Siminoff LA**, Ravdin P, Colabianchi N, Sturm CM. Doctor-patient communication patterns in breast cancer adjuvant therapy discussions. *Health Expect* 2000; **3**: 26-36 [PMID: 11281909 DOI: 10.1046/j.1369-6513.2000.00074.x]
11 **Butow PN**, Dunn SM, Tattersall MH, Jones QJ. Computer-based interaction analysis of the cancer consultation. *Br J Cancer* 1995; **71**: 1115-1121 [PMID: 7734311 DOI: 10.1038/bjc.1995.216]
12 **Brown RF**, Butow PN, Dunn SM, Tattersall MH. Promoting patient participation and shortening cancer consultations: a randomised trial. *Br J Cancer* 2001; **85**: 1273-1279 [PMID: 11720460 DOI: 10.1054/bjoc.2001.2073]
13 **Ong LM**, Visser MR, Lammes FB, de Haes JC. Doctor-patient communication and cancer patients' quality of life and satisfaction. *Patient Educ Couns* 2000; **41**: 145-156 [PMID: 12024540 DOI: 10.1016/S0738-3991(99)00108-1]
14 **Bylund CL**, Brown R, Gueguen JA, Diamond C, Bianculli J, Kissane DW. The implementation and assessment of a comprehensive communication skills training curriculum for oncologists. *Psychooncology* 2010; **19**: 583-593 [PMID: 19484714 DOI: 10.1002/pon.1585]
15 **Dimoska A**, Butow PN, Dent E, Arnold B, Brown RF, Tattersall MH. An examination of the initial cancer consultation of medical and radiation oncologists using the Cancode interaction analysis system. *Br J Cancer* 2008; **98**: 1508-1514 [PMID: 18454160 DOI: 10.1038/sj.bjc.6604348]
16 **Eggly S**, Penner LA, Greene M, Harper FW, Ruckdeschel JC, Albrecht TL. Information seeking during "bad news" oncology interactions: Question asking by patients and their companions. *Soc Sci Med* 2006; **63**: 2974-2985 [PMID: 16962218 DOI: 10.1016/j.socscimed.2006.07.012]
17 **Street RL**, Millay B. Analyzing patient participation in medical encounters. *Health Commun* 2001; **13**: 61-73 [PMID: 11370924 DOI: 10.1207/S15327027HC1301_06]
18 **Detmar SB**, Muller MJ, Wever LD, Schornagel JH, Aaronson NK. The patient-physician relationship. Patient-physician communication during outpatient palliative treatment visits: an observational study. *JAMA* 2001; **285**: 1351-1357 [PMID: 11255393 DOI: 10.1001/jama.285.10.1351]
19 **Koedoot CG**, Oort FJ, de Haan RJ, Bakker PJ, de Graeff A, de Haes JC. The content and amount of information given by medical oncologists when telling patients with advanced cancer what their treatment options are. palliative chemotherapy and watchful-waiting. *Eur J Cancer* 2004; **40**: 225-235 [PMID: 14728937 DOI: 10.1016/j.ejca.2003.10.008]
20 **Leighl N**, Gattellari M, Butow P, Brown R, Tattersall MH. Discussing adjuvant cancer therapy. *J Clin Oncol* 2001; **19**: 1768-1778 [PMID: 11251008]
21 **Street RL**, Voigt B, Geyer C, Manning T, Swanson GP. Increasing patient involvement in choosing treatment for early breast cancer. *Cancer* 1995; **76**: 2275-2285 [PMID: 8635032 DOI: 10.1002/1097-0142(19951201)76: ]
22 **Timmermans LM**, van der Maazen RW, Verhaak CM, van Roosmalen MS, van Daal WA, Kraaimaat FW. Patient participation in discussing palliative radiotherapy. *Patient Educ Couns* 2005; **57**: 53-61 [PMID: 15797153 DOI: 10.1016/j.pec.2004.03.016]
23 **Hunter J**, Schmidt F. Methods of meta-analysis: Correcting for artifact and bias in research. 2nd ed. Beverly Hills, CA: Sage, 2004: 185-197 [DOI: 10.4135/9781412985031]

**P- Reviewer:** Al-Badawi I, Robotin MC, Sajid MS
**S- Editor:** Ji FF **L- Editor:** A **E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.194

*World J Meta-Anal* 2014 November 26; 2(4): 194-203
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Systematic review and meta-analysis of Statins-Fibrates therapy in diabetic dyslipidemia patients

Shuang Zheng, Yang-Xue Li, Ting-Ting Han, Yao Zhang, Dong-Dong Jiang, Yao-Min Hu

Shuang Zheng, Yang-Xue Li, Ting-Ting Han, Yao Zhang, Dong-Dong Jiang, Yao-Min Hu, Department of Endocrinology, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai 200127, China

Author contributions: Hu YM designed the review; Zheng S and Li YX collected the data; Han TT, Zhang Y and Jiang DD analyzed the data; Zheng S and Hu YM wrote the paper.

Supported by The National Scientific Foundation of China, No. 81270946, No. 81170758, No. 30670988

Correspondence to: Yao-Min Hu, PhD, Department of Endocrinology, Renji Hospital, School of Medicine, Shanghai Jiaotong University, NO.160 Pujian Road, Shanghai 200127, China. amin1031@hotmail.com
Telephone: +86-21-68383447 Fax: +86-21-58394262

Received: June 6, 2014 Revised: June 24, 2014
Accepted: September 6, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To evaluate the efficacy, effect of preventing cardiovascular diseases and safety of statins-fibrates combination therapy in diabetic dyslipidemia patients.

**METHODS:** We searched the databases of MEDLINE, EMBASE, web of knowledge and Cochrane central register of Controlled Trials for literatures about the coadministration of statins and fibrates as the treatment of patients with dyslipidemia and type 2 diabetes mellitus. We included related randomized controlled trials, controlled clinical trials and cross-sectional studies and excluded animal trials and clinical observations. The primary endpoints outcomes were the concentration of plasma total cholesterol (TC), triglyceride (TG), high density lipoprotein cholesterol (HDL-C) and low density lipoprotein cholesterol (LDL-C). The secondary outcomes were cardiovascular diseases (CVD) and adverse events.

**RESULTS:** Ten studies were included in this meta-analysis. For lipid modifying efficacy, the combination of statins and fibrates therapy had more significant effect on reducing TC [$P$ = 0.004, weighted mean difference (WMD) = -8.19, 95%CI: -13.82--2.56] and TG concentration ($P$ < 0.001, WMD = -47.29, 95%CI: -68.66--25.92) and increasing HDL-C concentration ($P$ < 0.00001, WMD = 3.79, 95%CI: 2.25-5.33) when compared with statins monotherapy, while the effect of reducing LDL-C concentration ($P$ = 0.50, WMD = -2.52, 95%CI: -9.76-4.72) was insignificant. To fibrates monotherapy, the combination therapy was more effective on reducing TC ($P$ < 0.00001, WMD = -48.51, 95%CI: -57.14--39.89), TG ($P$ < 0.00001, WMD = -26.07, 95%CI: -30.96--21.18), LDL-C concentration ($P$ < 0.00001, WMD = -45.74, 95%CI: -53.35--38.13) and increasing HDL-C concentration ($P$ = 0.04, WMD = 1.38, 95%CI: 0.04-2.73). For cardiovascular diseases, the coadministration therapy had no significant effect on reducing the incidence of these events when compared with monotherapy (For primary clinical endpoints, $P$ = 0.12, OR = 0.61, 95%CI: 0.33-1.14); for secondary clinical endpoints, $P$ = 0.13, OR = 0.66, 95%CI: 0.38-1.14). For adverse events happened during the follow-up, both the incidence of hepatic-related (alanine aminotransferase and/or aspartate aminotransferase of patients were $\geqslant$ 3 times of upper limit of normal) ($P$ = 0.38, OR = 0.55, 95%CI: 0.15-2.06) and muscular-related (myopathy and/or creatine phosphokinase $\geqslant$ 3 times of upper limit of normal) adverse events ($P$ = 0.10, OR = 1.62, 95%CI: 0.91-2.86) had no significant difference between these two therapies.

**CONCLUSION:** The results showed statins-fibrates combination therapy was more effective on lipid modification and well tolerated but there was no significant effect on preventing cardiovascular diseases.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Statin; Fibrate; Dyslipidemia; Type 2 diabetes; Combination therapy

**Core tip:** Both dyslipidemia and type 2 diabetes were established risk factors of cardiovascular diseases. Statins therapy was highly effective at lowering low density lipo-

protein cholesterol (LDL-C). However, despite the increasing use of statins as monotherapy for LDL-C reduction, a significant residual cardiovascular risk was still presented in patients with diabetic dyslipidemia. At the same time, Fenofibrate failed to alter the primary clinical endpoints significantly. How about the efficacy of statins-fibrates combination therapy in patients with diabetic dyslipidemia? The results of this meta-analysis showed the combination therapy was more effective on lipid modification and well tolerated but there was no significant effect on preventing cardiovascular diseases.

Zheng S, Li YX, Han TT, Zhang Y, Jiang DD, Hu YM. Systematic review and meta-analysis of Statins-Fibrates therapy in diabetic dyslipidemia patients. *World J Meta-Anal* 2014; 2(4): 194-203 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/194.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.194

## INTRODUCTION

Cardiovascular diseases (CVD) represent the leading cause of mortality worldwide. Many researches have confirmed that both dyslipidemia and type 2 diabetes have tight relationship with CVD[1,2]. The combination of dyslipidemia and diabetes could be called diabetic dyslipidemia, which is a well-recognized reason for atherosclerotic cardiovascular diseases[3]. Elevated low density lipoprotein cholesterol (LDL-C) is a major risk factor for CVD[4]. As a result, current guidelines recommend management of LDL-C as the primary goal of therapy for diabetic dyslipidemia[5]. Statins are the drug of first choice for aggressive lipid lowering actions and reducing risk of CVD in these patients[6]. However, current therapeutic use of statins as monotherapy is still leaving many patients with diabetic dyslipidemia at high risk for CVD[7]. Some studies have come up with the conclusion that the coadministration of statins and fibrates may be more effective with no more adverse events as the treatment to patients with mixed dyslipidemia than statins or fibrates monotherapy[8-12], which may reduce the incidence of CVD at the same time. So we are curious about the efficacy and safety of the combined statins-fibrates therapy and their benefits in reducing CVD incidence in patients with dyslipidemia and type 2 diabetes.

## MATERIALS AND METHODS

### Search strategy and selection criteria

The search was performed using database of MEDLINE from 1980 to March 2014, EMBASE from 1980 to March 2014, the forth quarter 2014 Cochrane central register of controlled trials and web of knowledge from 1980 to March 2014. We conducted a comprehensive and systematic search of the published literature for trials of coadministration of statins and fibrtaes as the treatment of dyslipidemia with type 2 diabetes. The initial search terms were "statin", "fibrate", "dyslipidemia", "diabetes" and their combination [statin AND fibrate /"combination therapy" AND dyslipidemia/dyslipidemia AND diabetes/(statin AND fibrate) AND dyslipidemia/(statin AND fibrate) AND diabetes/"diabetic dyslipidemia" /(statin AND fibrate) AND "diabetic dyslipidemia" /"combination therapy" AND "diabetic dyslipidemia"].

We included randomized controlled trials, controlled clinical trials and cross-sectional studies those investigated the efficacy and safety of coadministration of statins and fibrates in patients with diabetic dyslipidemia and excluded animal trials and clinical observations. The treatment group involved the coadministration of statins and fibrates. The control group used statins or fibrates monotherapy. The statins included Simvastatin, Fluvastatin, Atorvastatin, Pravastatin, Rosuvastatin and Cerivastatin. The fibrates included Fenofibrate, Bezafibrate and Fenofibric acid.

Two reviewers independently evaluated the articles and any disagreement was resolved by consensus.

### Statistical analysis

Study design data including design synopsis, duration of treatment and basic characteristics of patients. The primary endpoints outcomes were the concentration of total cholesterol (TC), triglyceride (TG), high density lipoprotein cholesterol (HDL-C) and LDL-C. The secondary outcomes were cardiovascular diseases and adverse events. For cardiovascular diseases, the primary clinical endpoints included death from cardiovascular causes, nonfatal myocardial infarction or nonfatal stroke. The secondary clinical endpoints included the combination of the primary outcome plus revascularization or hospitalization for congestive heart failure (termed the "expanded macrovascular outcome"); a combination of a fatal coronary event, nonfatal myocardial infarction, or unstable angina (termed "major coronary disease events"); nonfatal myocardial infarction; fatal or nonfatal stroke; nonfatal stroke; death from any cause; death from cardiovascular causes; and hospitalization or death due to heart failure. For adverse events, if the alanine aminotransferase and/or aspartate aminotransferase of patients were $\geqslant$ 3 times of upper limit of normal during follow-up, it would be counted as hepatic-related adverse events. While the muscular-related adverse events included myopathy and CK (creatine phosphokinase) $\geqslant$ 3 times of upper limit of normal during follow-up.

We combined the results and expressed them as odds radio (OR) or weighted mean difference (WMD) with corresponding 95%CIs, using a fixed effect (FE) or randomized effect (RE) model for the studies with sufficient data (in this article, if the homogeneity was no more than 50%, FE model was used, else RE model was used). The homogeneity was assessed with $I^2$ and $\chi^2$ test. The above statistical calculations were performed on Revman Manager 5.0 Software (Copenhagen, Denmark) for meta-analysis.

## RESULTS

We searched 107 potentially relevant studies and retrieved 62 full-text articles. 52 of the full-text articles were exclud-

**Table 1 Basic characteristics of included studies**

| Ref. | Type | Number (T/C1/C2) | Age (mean) (T/C1/C2) | Men (%) | Treatment | Contral 1 | Contral 2 | Duration (mo) |
|---|---|---|---|---|---|---|---|---|
| Athyros *et al*[19] 2002 | RCT | 40/40/40 | 58/57/58 | 56.7 | A + Fe | A | Fe | 6 |
| Durrington *et al*[22] 2004 | RCT | 115/53/48 | 60/60/60 | 53.0 | R + Fe | R | Fe | 6 |
| Rosenson *et al*[21] 2011 | RCT | 177/173/123 | 60/58/58 | 41.4 | R + FA | R | FA | 3 |
| Derosa *et al*[18] 2004 | RCT | 25/23/0 | 61/59/0 | 50.0 | F + Fe | F | | 6 |
| Farnier *et al*[16] 2011 | RCT | 145/146/0 | 56/57/0 | 48.1 | P + Fe | S | | 3 |
| Hamilton *et al*[17] 2010 | RCT | 15/0/15 | 63/0/63 | 86.7 | A/S/P/R + Fe | | Fe | 3 |
| Ginsberg *et al*[20] 2010 | RCT | 2765/2753/0 | 62/62/0 | 69.3 | S + Fe | S | | 56.4 |
| Gavish *et al*[13] 2000 | CCT | 146/100/48 | 59/58/60 | 57.4 | S + B | S | B | 21 |
| Constantinides *et al*[15] 2012 | CCT | 14/14/14 | > 18/> 18/> 18 | 100 | S + B | S | B | 2 |
| Klempfner *et al*[14] 2014 | CSS | 225/2838/0 | 60/65/0 | 71.0 | A/S/P/C/F + B | | B | 12 |

RCT: Randomized controlled trial; CCT: Controlled clinical trial; CSS: Cross-sectional study; T/C1/C2: Treatment/Control 1/Control 2; S: Simvastatin; B: Bezafibrate; F: Fluvastatin; Fe: Fenofibrate; A: Atorvastatin; P: Pravastatin; R: Rosuvastatin; C: Cerivastatin; FA: Fenofibric acid.

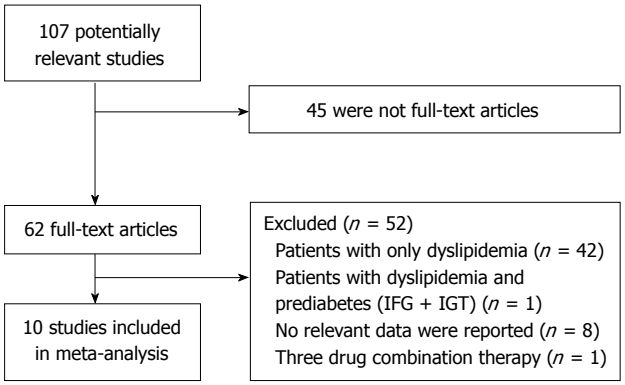



**Figure 1 Flow diagram of study screening process.** IFG: Impaired fasting glucose; IGT: Impaired glucose tolerance.

ed because the patients were not having both dyslipidemia and diabetes or the data were not complete (Figure 1). Finally only 10 studies[13-22] were included in our research and their basic characteristics were shown in Table 1.

### *Comparisons of lipid modifying efficacy between coadministration and monotherapy*

Figure 2 showed the comparisons of lipid modifying efficacy between coadministration of statins and fibrates and statins monotherapy. Figure 2A presented the results of plasma TC concentration in two different therapies; six trials investigated it with a total of 721 patients. From this figure we observed that compared with statins monotherapy, the coadministration of statins and fibrates had more strong function on reducing the concentration of plasma TC in patients with diabetic dyslipidemia ($P$ = 0.004, WMD = -8.19, 95%CI: -13.82--2.56). Figure 2B presented the results of plasma TG concentration in two different therapies, four trials with 663 patients were included, which had the similar results with TC ($P$ < 0.001, WMD = -47.29, 95%CI: -68.66--25.92). Both of them showed a beneficial and statistically significant effect of coadministration on diabetic dyslipidemia. Figure 2C presented the results of plasma HDL-C concentration with six trials and 721 patients. From this figure we observed the concentration of HDL-C was lower in statins monotherapy group ($P$ < 0.00001, WMD = 3.79, 95%CI: 2.25-5.33). Figure 2D presented the results of plasma LDL-C concentration with four trials and 691 patients, which showed the modification of LDL-C was not significant between coadministration and statins monotherapy ($P$ = 0.50, WMD = -2.52, 95%CI: -9.76-4.72). In conclusion, for lipid modifying efficacy, the combination of statins and fibrates therapy had more significant effect on reducing plasma TC and TG concentration and increasing HDL-C concentration when compared with statins monotherapy, while the effect of reducing LDL-C concentration was insignificant.

Figure 3 showed the comparisons of lipid modifying efficacy between coadministration of statins and fibrates and fibrates monotherapy. Figure 3A, B and D presented the results of plasma TC, TG and LDL-C concentration respectively. All of them showed a beneficial and statistically significant effect on reducing TC, TG and LDL-C with combination therapy ($P$ < 0.00001). Figure 3C presented the results of plasma HDL-C concentration with three trials and 302 patients. From this figure we observed the concentration of HDL-C was lower in fibrates monotherapy group ($P$ = 0.04, WMD = 1.38, 95%CI: 0.04-2.73). In conclusion, the combination therapy was more effective on reducing TC, TG, LDL-C concentration and increasing HDL-C concentration than fibrates monotherapy.

### *Comparisons of cardiovascular diseases between coadministration and monotherapy*

Figure 4 showed the incidence of cardiovascular diseases between coadministration and statins or fibrates monotherapy. Three identified studies were included in this analysis with 8875 patients. Figure 4A showed the primary clinical endpoints, Figure 4B showed the secondary clinical endpoints. From Figure 4 we observed that in patients with diabetic dyslipidemia, the coadministration therapy had insignificant effect on reducing the incidence of cardiovascular diseases when compared with statins or fibrates monotherapy. For primary clinical endpoints, $P$ = 0.12, OR = 0.61, 95%CI: 0.33-1.14); for secondary clinical endpoints, $P$ = 0.13, OR = 0.66, 95%CI: 0.38-1.14).

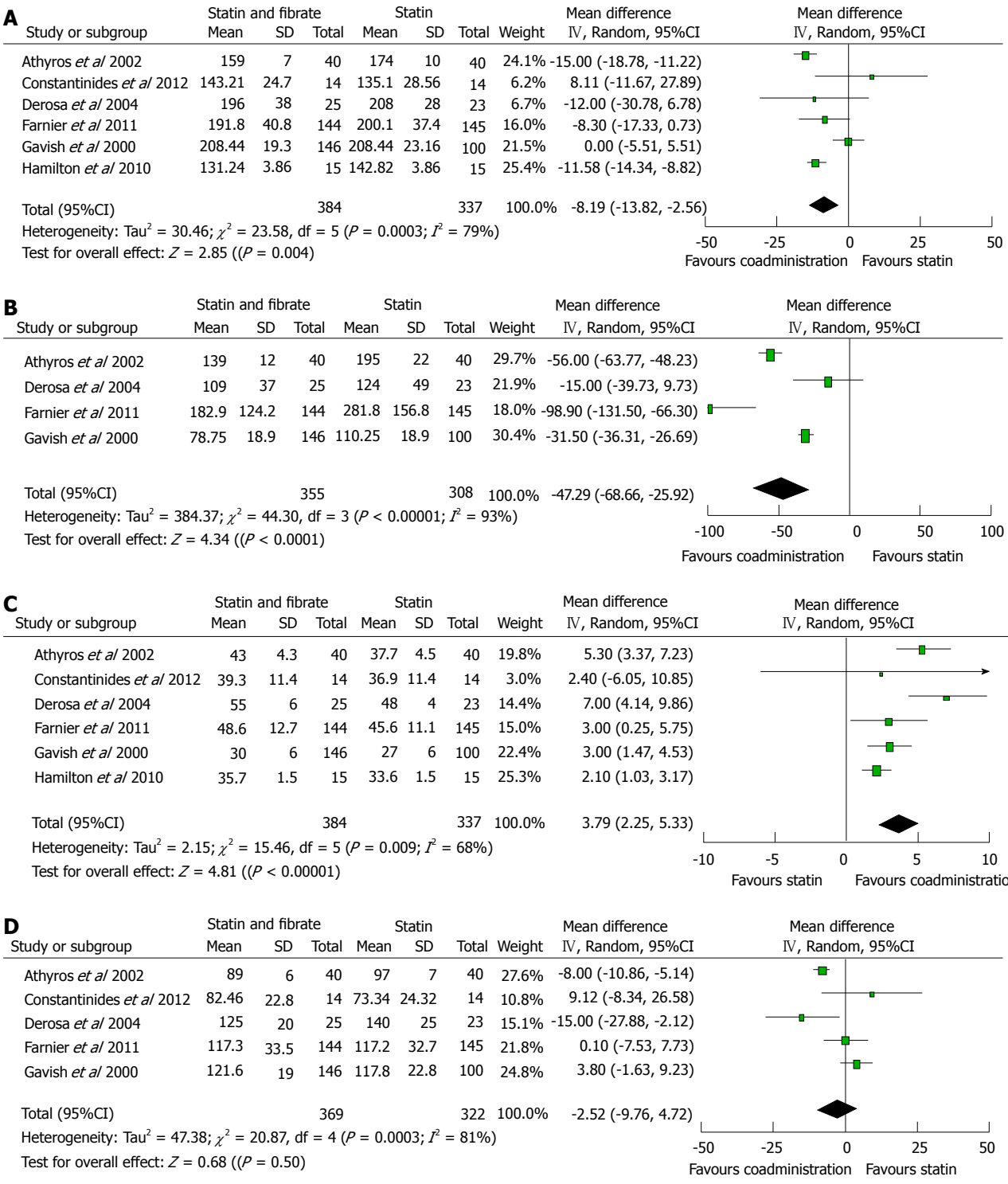



**Figure 2 Summary of results of meta-analysis on the lipid modifying function between coadministration of statins and fibrates and statins monotherapy.** A: Comparison of plasma TC concentration between coadministration and statins monotherapy; B: Comparison of plasma TG concentration between coadministration and statins monotherapy; C: Comparison of plasma HDL-C concentration between coadministration and statins monotherapy; D: Comparison of plasma LDL-C concentration between coadministration and statins monotherapy. LDL-C: Low density lipoprotein cholesterol; HDL-C: High density lipoprotein cholesterol; TG: Triglyceride.

In conclusion, the coadministration therapy had no significant effect on reducing the incidence of these events when compared with monotherapy.

### *Comparisons of adverse events between coadministration and monotherapy*

Figure 5 showed the adverse events happened in the period of two different therapies. Four related trials were included in meta-analysis with a total of 1274 patients. Figure 5A presented the hepatic-related adverse events. Figure 5B presented the muscular-related adverse events. There was no significant difference between the two therapies ($P = 0.38$ and 0.10 respectively). In conclusion, statins-fibrates combination therapy was tolerated as well

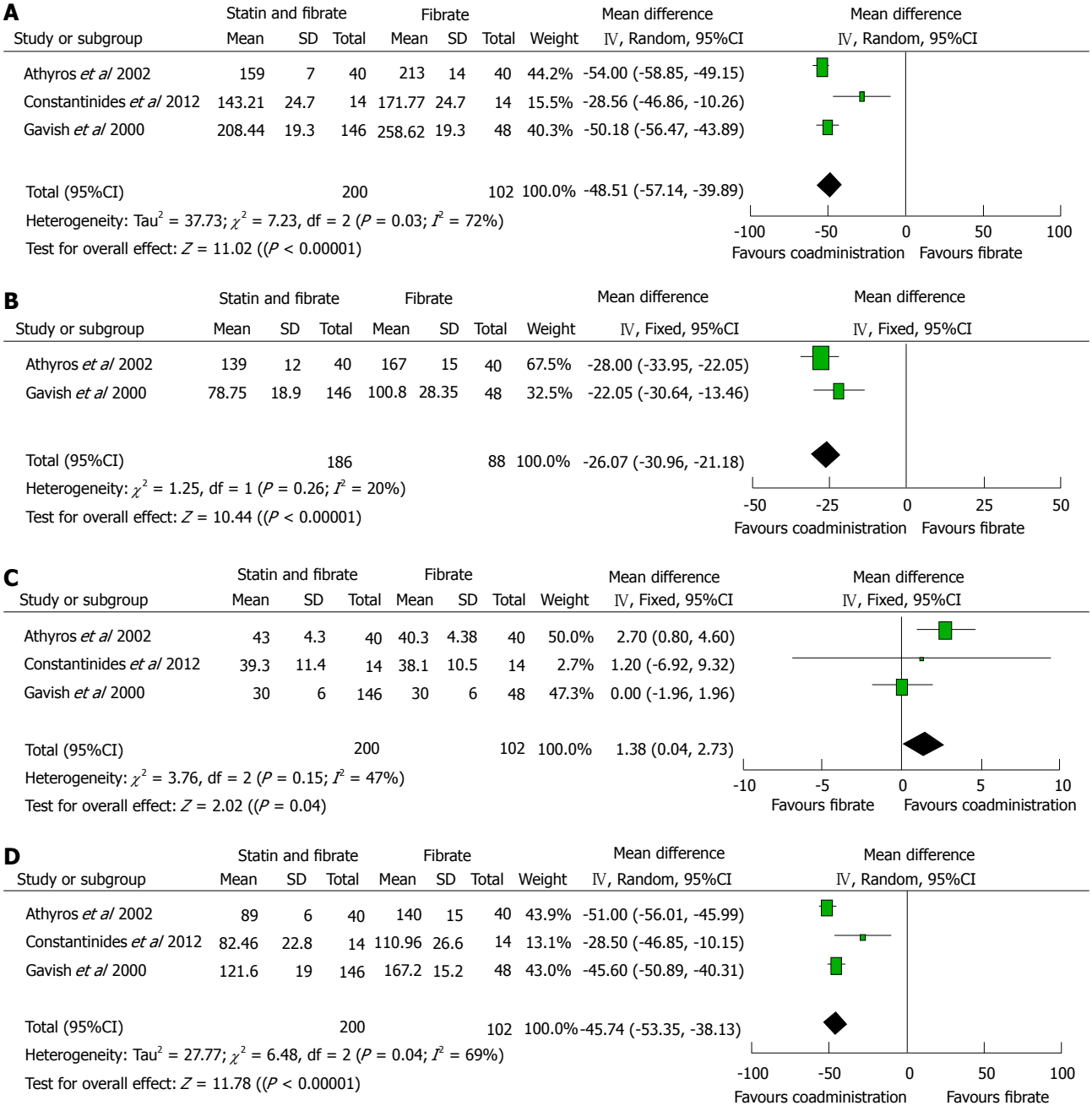



**Figure 3 Summary of results of meta-analysis on the lipid modifying function between coadministration of statins and fibrates and fibrates monotherapy.** A: Comparison of plasma TC concentration between coadministration and fibrates monotherapy; B: Comparison of plasma TG concentration between coadministration and fibrates monotherapy; C: Comparison of plasma HDL-C concentration between coadministration and fibrates monotherapy; D: Comparison of plasma LDL-C concentration between coadministration and fibrates monotherapy. LDL-C: Low density lipoprotein cholesterol; HDL-C: High density lipoprotein cholesterol; TG: Triglyceride.

as statins or fibrates monotherapy.

## DISCUSSION

CVD continues the leading cause of death worldwide. Both dyslipidemia and type 2 diabetes were established risk factors for CVD[1,2]. Moreover, dyslipidemia was strikingly common in patients with type 2 diabetes, affecting almost 50% of this population[2]. So it was easy to understand that CVD was more common in patients with diabetic dyslipidemia than in the general populations[23]. Several researches had come out with the idea that elevated LDL-C was a major risk factor for CVD[4,24,25]. As a result, management of LDL-C was the primary goal of therapy for patients with dyslipidemia and type 2 diabetes[5,26].

Statins therapy was highly effective at lowering LDL-C. Hydroxymethylglutaryl-coenzyme A reductase inhibitors (statins) have emerged as the cornerstone for LDL-C lowering since the first agent, lovastatin, was approved in 1987[27]. However, despite the increasing use of statins as monotherapy for LDL-C reduction, a

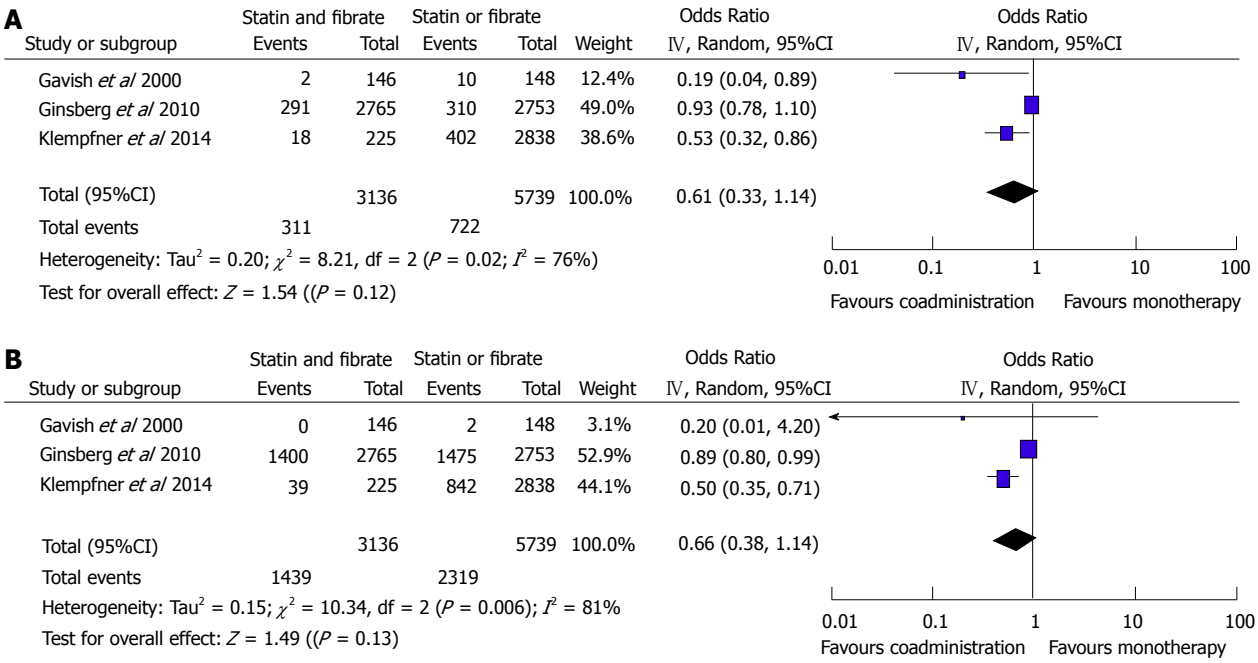



**Figure 4 Summary of results of meta-analysis on cardiovascular diseases.** A: Primary clinical endpoints events; B: Secondary clinical endpoints events.

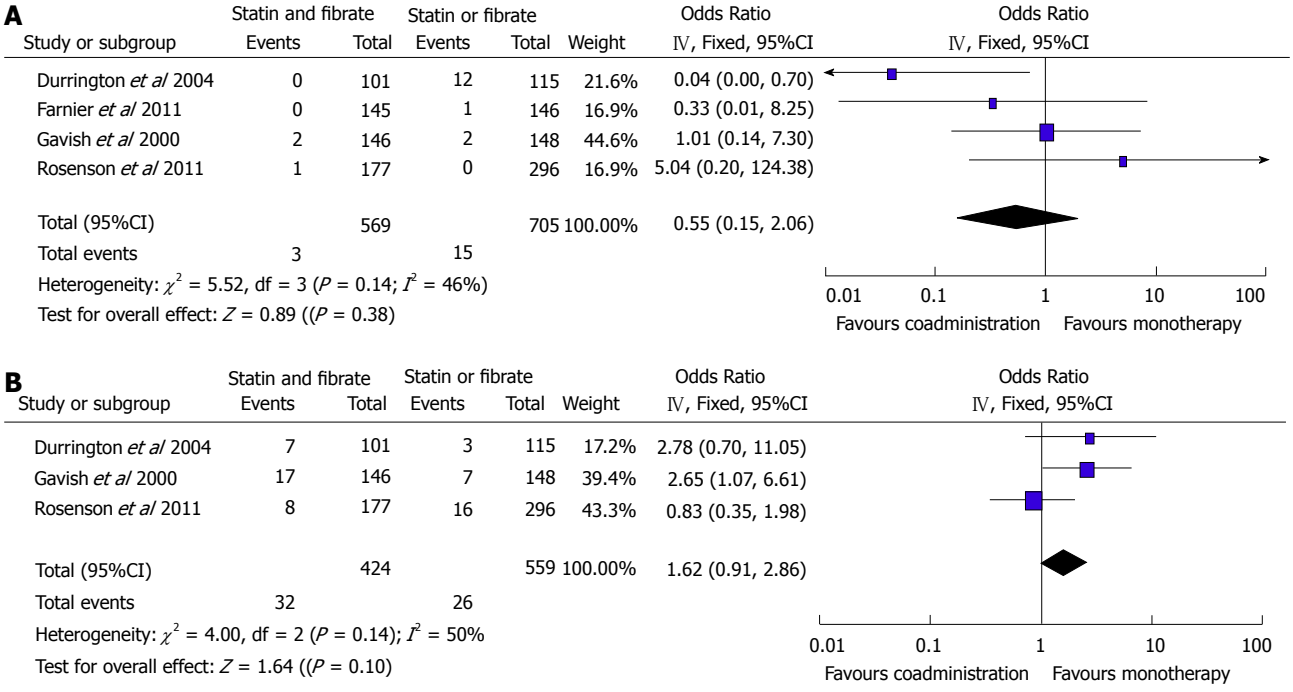



**Figure 5 Summary of results of meta-analysis on adverse events.** A: Hepatic-related adverse events; B: Muscular-related adverse events.

significant residual cardiovascular risk was still presented in patients with diabetic dyslipidemia[7]. The reason was that LDL-C could not reflect the classic "diabetic dyslipidemia", which consisted of hypertriglyceridemia and low levels of HDL-C[28,29]. In patients with type 2 diabetes, LDL particles were small and dense, carrying less cholesterol per particle; therefore, at any given LDL-C concentration, there were more LDL particles present in an individual with type 2 diabetes relative to an individual without the disease, which might make the LDL-C level a misleading measure of risk in patients with type 2 diabetes[5]. It was increasingly recognized that insulin resistance contributed to the characteristic dyslipidemia associated with type 2 diabetes[30]; and this dyslipidemia associated with insulin resistance was also referred to as atherogenic dyslipidemia.

Since diabetic dyslipidemia was characterized by moderately increased TG levels and reduced HDL-C, fibrates therapy should be considered as the option. Several large clinical and angiographic trials had evaluated the efficacy

of fibrates as monotherapy in halting the progression of atherosclerotic diseases[31-33]. The Fenofibrate Intervention and Event Lowering in Diabetes study was a 5-year, randomized, placebo-controlled trial testing the safety and efficacy of fenofibrate 200 mg in 9795 type 2 diabetic patients[34]. In this trial Fenofibrate failed to alter the primary clinical endpoints significantly.

Although LDL-C levels didn't differ substantially from individuals with or without type 2 diabetes, data had demonstrated that lowering LDL-C levels reduced the risk for major CVD in patients with type 2 diabetes[5]. It was well accepted that statins were the primary and more efficient method of reducing LDL-C levels even at low doses[35]. However, statins manifested minimal effects on raising HDL-C levels (5%-15%) and on decreasing TG levels (7%-30%)[36]. Fibrates had small or minimal effects on LDL-C levels, which depended on baseline TG levels[35]. These data implied that a combination of statins and fibrates may have additional benefits, especially in patients with dyslipidemia and type 2 diabetes.

Some researches have shown the efficacy of statins-fibrates combination therapy in patients with mixed dyslipidemia. Research by Goldberg *et al*[8] focused on the efficacy of fenofibric acid plus statins on multiple lipid parameters in women with mixed dyslipidemia and the results showed the coadministration could increase the HDL-C level and decrease TG level more effecctive than statin monotherapy. Similar results could observed from Research by Pepine *et al*[9] showed that in elderly patients with mixed dyslipidemia, rosuvastatin 5, 10, or 20 mg in combination with fenofibric acid 135 mg improved the overall lipid profile. Farnier *et al*[10] also found in high-risk patients with mixed hyperlipidemia not controlled by pravastatin 40 mg monotherapy, the fenofibrate 160 mg/pravastatin 40 mg fixed-dose combination therapy significantly improved the global atherogenic lipid profile. So it was easy for us to suppose that combination therapy was more effective on patients with mixed dyslipidemia than statins or fibrates monotherapy. How about patients with diabetic dyslipidemia?

From our meta-analysis, we observed that in patients with both dyslipidemia and type 2 diabetes, compared with statins or fibrates monotherapy, the coadministration of statins and fibrates had more significant effect on lowering TG concentration. For plasma LDL-C concentration, combination therapy had statistically significant effect on lowering it than fibrates monotherapy. These data gave us the implication that the coadministration may be more effective on lipid modification than statins or fibrates monotherapy in diabetic dyslipidemia patients.

Since the combination therapy had additional benefit than monotherapy on lipid modifying efficacy in diabetic dyslipidemia patients, it was reasonable for us to give the hypothesis that the combination therapy would result in an additional cardiovascular benefit, as compared with statins therapy alone. One study focused on cardiovascular events in patients received combined fibrates/statins treatment versus statin monotherapy had showed that a significantly lower risk of 30-d major adverse cardiovascular events rate was observed in patients receiving combined fibrates/statins treatment following acute coronary syndrome compared with statins monotherapy[37]. However, from the results of our meta-analysis, there was no significant difference on the incidence of cardiovascular diseases, no matter the primary or the secondary clinical endpoints, between the coadministration therapy and monotherapy in diabetic dyslipidemia patients. So our analysis did not support the use of combination therapy to reduce cardiovascular risk in the majority of patients with type 2 diabetes who were at high risk for cardiovascular diseases. But this evidence was not so robust since only three identified studies were included in this part of meta-analysis. Further studies should continue focus on the rate of CVD and maybe we could find different answers to this kind of questions.

In addition to lipid modifying efficacy, safety was an important issue influencing the selection of combination therapy or monotherapy. Common adverse events associated with statins use included gastrointestinal upset and muscle aches, although dose related hepatoxicity and myotoxicity were the most clinically significant adverse events[38]. Common adverse events associated with fibrates included gastrointestinal disturbance, rash, headache, pancreatitis, myalgia, and myotoxicity (in rare instances-and possibly more likely with gemfibrozil than with fenofibrate)[23]. Combination therapy with statins and gemfibrozil was more likely to be accompanied by severe myopathy[39,40]. This might due to the fact that gemfibrozil had significant pharmacokinetic interactions with statins that lead to increased plasma levels of statins[41]. This limitation was not observed with fenofibrate, bezafibrate, or ciprofibrate and no significant side effects had been observed with combination treatment with statins and fibrates[23]. Same conclusions could be driven from our meta-analysis, both incidence of hepatic-related adverse events and muscular-related adverse events had no obvious difference between coadministration and monotherapy. So the combination therapy could be well tolerated in diabetic dyslipidemia patients.

In conclusion, diabetic dyslipidemia was associated with elevated serum TG, low serum HDL-C levels, and a preponderance of small, dense LDL particles. Disturbance of lipid metabolism linked to insulin resistance may be the primary event in the development of type 2 diabetes, which had the tight relationship with CVD. The present meta-analysis had shown that the coadministration was more effective on lipid modification than monotherapy and it was well tolerated, though the rate of CVD had no significant difference when combination therapy was given.

However, our meta-analysis had several potential limitations. Firstly, most studies included in this mate-analysis were of small sample size and didn't describe withdrawals or dropouts. Secondly, few studies were included in the analysis of the incidence of CVD, for few related published articles were found, so our meta-analysis may

be affected by publication bias. Thus, more high quality studies were needed to evaluate the efficacy of coadministration therapy on reducing the rate of CVD in patients with diabetic dyslipidemia. Finally, our results showed there was no significant difference between combination therapy and statins monotherapy for reducing the incidence of CVD in patients with diabetic dyslipidemia. However, considering the limited literatures and most of the researches were not focus on dyslipidemia patients who need the combination therapy to prevent CVD and residual atherogenic risk after statins monotherapy, there may be some potential bias. As a result, further study aiming at these patients in need for combination therapy was necessary for a more precise and reliable comparison.

## COMMENTS

### Background

Statins-fibrates combination therapy has been suggested to be more effective with no more adverse events as the treatment to patients with mixed dyslipidemia than statins or fibrates monotherapy, which may reduce the incidence of cardiovascular diseases (CVD) at the same time. However, it is still unclear of the efficacy and safety of the combined statins-fibrates therapy and their benefits on reducing CVD incidence in patients with dyslipidemia and type 2 diabetes.

### Research frontiers

Statins therapy was highly effective at lowering low density lipoprotein cholesterol (LDL-C). However, despite the increasing use of statins as monotherapy for LDL-C reduction, a significant residual cardiovascular risk was still presented in patients with diabetic dyslipidemia. The Fenofibrate Intervention and Event Lowering in Diabetes study was a 5-year, randomized, placebo-controlled trial testing the safety and efficacy of fenofibrate 200 mg in 9795 type 2 diabetic patients. In this trial Fenofibrate failed to alter the primary clinical endpoints significantly. Some researches have shown the efficacy and safety of statins-fibrates combination therapy in patients with mixed dyslipidemia. But no consensus was reached among patients with diabetic dyslipidemia of using the combination therapy.

### Innovations and breakthroughs

This meta-analysis included 10 articles and more than 10 thousand patients were contained. Most of the included studies were randomized controlled trials. The authors found that the combination therapy was more effective on lipid modification and well tolerated but there was no significant effect on preventing cardiovascular diseases.

### Applications

Statins-fibrates combination therapy may potentially be used for the therapy of diabetic dyslipidemia; however, more high quality studies are required for further estimate of the efficacy of combination therapy on reducing the rate of CVD.

### Terminology

Statins are Hydroxymethylglutaryl-coenzyme A reductase inhibitors, which may inhibit HMG-CoA reductase, a precursor to the formation of cholesterol, and up-regulate the LDL-C-receptor. What's more, statins could reduce the hepatic cholesterol synthesis. Fibrates exert their primary effects on lipid metabolism *via* the activation of peroxisome proliferator activated receptor-alpha. It is able to reduce plasma triglyceride levels by inhibiting their synthesis and stimulating their clearance.

### Peer review

In this article, Zheng *et al* present evidence that the coadministration of statins and fibrates in the treatment of patients with dyslipidemia and type 2 diabetes was more effective on lipid modification than statins or fibrates monotherapy. They also present evidence that statins-fibrates combination therapy was tolerated as well as statins or fibrates monotherapy. This is an interesting report for the clinical practice. Overall the report appears to be carefully examined and data adequately discussed.

## REFERENCES


1 **Kannel WB**, McGee DL. Diabetes and cardiovascular disease. The Framingham study. *JAMA* 1979; **241**: 2035-2038 [PMID: 430798]

2 **Saydah SH**, Fradkin J, Cowie CC. Poor control of risk factors for vascular disease among adults with previously diagnosed diabetes. *JAMA* 2004; **291**: 335-342 [PMID: 14734596 DOI: 10.1001/jama.291.3.335291/3/335]

3 **Ng DS**. Diabetic dyslipidemia: from evolving pathophysiological insight to emerging therapeutic targets. *Can J Diabetes* 2013; **37**: 319-326 [PMID: 24500559 DOI: 10.1016/j.jcjd.2013.07.062S1499-2671]

4 **Turner RC**, Millns H, Neil HA, Stratton IM, Manley SE, Matthews DR, Holman RR. Risk factors for coronary artery disease in non-insulin dependent diabetes mellitus: United Kingdom Prospective Diabetes Study (UKPDS: 23) *BMJ* 1998; **316**: 823-828 [PMID: 9549452]

5 **Goff DC**, Gerstein HC, Ginsberg HN, Cushman WC, Margolis KL, Byington RP, Buse JB, Genuth S, Probstfield JL, Simons-Morton DG. Prevention of cardiovascular disease in persons with type 2 diabetes mellitus: current knowledge and rationale for the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial. *Am J Cardiol* 2007; **99**: 4i-20i [PMID: 17599424]

6 **Tenenbaum A**, Fisman EZ. Which is the best lipid-modifying strategy in metabolic syndrome and diabetes: fibrates, statins or both? *Cardiovasc Diabetol* 2004; **3**: 10 [PMID: 15574199]

7 **Tenenbaum A**, Fisman EZ. Balanced pan-PPAR activator bezafibrate in combination with statin: comprehensive lipids control and diabetes prevention? *Cardiovasc Diabetol* 2012; **11**: 140 [PMID: 23150952 DOI: 10.1186/1475-2840-11-1401475-2840-11-140]

8 **Goldberg AC**, Bittner V, Pepine CJ, Kelly MT, Thakker K, Setze CM, Lele A, Sleep DJ. Efficacy of fenofibric acid plus statins on multiple lipid parameters and its safety in women with mixed dyslipidemia. *Am J Cardiol* 2011; **107**: 898-905 [PMID: 21247520 DOI: 10.1016/j.amjcard.2010.11.009S0002-9149]

9 **Pepine CJ**, Jacobson TA, Carlson DM, Kelly MT, Setze CM, Gold A, Stolzenbach JC, Williams LA. Combination rosuvastatin plus fenofibric acid in a cohort of patients 65 years or older with mixed dyslipidemia: subanalysis of two randomized, controlled studies. *Clin Cardiol* 2010; **33**: 609-619 [PMID: 20960535 DOI: 10.1002/clc.20830]

10 **Farnier M**, Ducobu J, Bryniarski L. Efficacy and safety of adding fenofibrate 160 mg in high-risk patients with mixed hyperlipidemia not controlled by pravastatin 40 mg monotherapy. *Am J Cardiol* 2010; **106**: 787-792 [PMID: 20816118 DOI: 10.1016/j.amjcard.2010.05.005S0002-9149]

11 **Davidson MH**, Rooney MW, Drucker J, Eugene Griffin H, Oosman S, Beckert M. Efficacy and tolerability of atorvastatin/fenofibrate fixed-dose combination tablet compared with atorvastatin and fenofibrate monotherapies in patients with dyslipidemia: a 12-week, multicenter, double-blind, randomized, parallel-group study. *Clin Ther* 2009; **31**: 2824-2838 [PMID: 20110022 DOI: 10.1016/j.clinthera.2009.12.007S0149-2918]

12 **Goldberg AC**, Bays HE, Ballantyne CM, Kelly MT, Buttler SM, Setze CM, Sleep DJ, Stolzenbach JC. Efficacy and safety of ABT-335 (fenofibric acid) in combination with atorvastatin in patients with mixed dyslipidemia. *Am J Cardiol* 2009; **103**: 515-522 [PMID: 19195513 DOI: 10.1016/j.amjcard.2008.10.025S0002-9149]

13 **Gavish D**, Leibovitz E, Shapira I, Rubinstein A. Bezafibrate and simvastatin combination therapy for diabetic dyslipidaemia: efficacy and safety. *J Intern Med* 2000; **247**: 563-569 [PMID: 10809995]

14 **Klempfner R**, Goldenberg I, Fisman EZ, Matetzky S, Amit U,

Shemesh J, Tenenbaum A. Comparison of statin alone versus bezafibrate and statin combination in patients with diabetes mellitus and acute coronary syndrome. *Am J Cardiol* 2014; **113**: 12-16 [PMID: 24157192 DOI: 10.1016/j.amjcard.2013.08.033S0002-9149]

15 **Constantinides A**, de Vries R, van Leeuwen JJ, Gautier T, van Pelt LJ, Tselepis AD, Lagrost L, Dullaart RP. Simvastatin but not bezafibrate decreases plasma lipoprotein-associated phospholipase $A_2$ mass in type 2 diabetes mellitus: relevance of high sensitive C-reactive protein, lipoprotein profile and low-density lipoprotein (LDL) electronegativity. *Eur J Intern Med* 2012; **23**: 633-638 [PMID: 22902096 DOI: 10.1016/j.ejim.2012.05.008S0953-6205]

16 **Farnier M**, Steinmetz A, Retterstøl K, Császár A. Fixed-dose combination fenofibrate/pravastatin 160/40 mg versus simvastatin 20 mg monotherapy in adults with type 2 diabetes and mixed hyperlipidemia uncontrolled with simvastatin 20 mg: a double-blind, randomized comparative study. *Clin Ther* 2011; **33**: 1-12 [PMID: 21397769 DOI: 10.1016/j.clinthera.2011.02.006S0149-2918]

17 **Hamilton SJ**, Chew GT, Davis TM, Watts GF. Fenofibrate improves endothelial function in the brachial artery and forearm resistance arterioles of statin-treated Type 2 diabetic patients. *Clin Sci* (Lond) 2010; **118**: 607-615 [PMID: 20047560 DOI: 10.1042/CS20090568CS20090568]

18 **Derosa G**, Cicero AE, Bertone G, Piccinni MN, Ciccarelli L, Roggeri DE. Comparison of fluvastatin + fenofibrate combination therapy and fluvastatin monotherapy in the treatment of combined hyperlipidemia, type 2 diabetes mellitus, and coronary heart disease: a 12-month, randomized, double-blind, controlled trial. *Clin Ther* 2004; **26**: 1599-1607 [PMID: 15598476]

19 **Athyros VG**, Papageorgiou AA, Athyrou VV, Demitriadis DS, Kontopoulos AG. Atorvastatin and micronized fenofibrate alone and in combination in type 2 diabetes with combined hyperlipidemia. *Diabetes Care* 2002; **25**: 1198-1202 [PMID: 12087019]

20 **Ginsberg HN**, Elam MB, Lovato LC, Crouse JR, Leiter LA, Linz P, Friedewald WT, Buse JB, Gerstein HC, Probstfield J, Grimm RH, Ismail-Beigi F, Bigger JT, Goff DC, Cushman WC, Simons-Morton DG, Byington RP. Effects of combination lipid therapy in type 2 diabetes mellitus. *N Engl J Med* 2010; **362**: 1563-1574 [PMID: 20228404 DOI: 10.1056/NEJMoa1001282NEJMoa1001282]

21 **Rosenson RS**, Carlson DM, Kelly MT, Setze CM, Hirshberg B, Stolzenbach JC, Williams LA. Achievement of lipid targets with the combination of rosuvastatin and fenofibric Acid in patients with type 2 diabetes mellitus. *Cardiovasc Drugs Ther* 2011; **25**: 47-57 [PMID: 21174145 DOI: 10.1007/s10557-010-6273-5]

22 **Durrington PN**, Tuomilehto J, Hamann A, Kallend D, Smith K. Rosuvastatin and fenofibrate alone and in combination in type 2 diabetes patients with combined hyperlipidaemia. *Diabetes Res Clin Pract* 2004; **64**: 137-151 [PMID: 15063607 DOI: 10.1016/j.diabres.2003.11.012]

23 **Vijayaraghavan K**. Treatment of dyslipidemia in patients with type 2 diabetes. *Lipids Health Dis* 2010; **9**: 144 [PMID: 21172030 DOI: 10.1186/1476-511X-9-144]

24 **Buse JB**, Ginsberg HN, Bakris GL, Clark NG, Costa F, Eckel R, Fonseca V, Gerstein HC, Grundy S, Nesto RW, Pignone MP, Plutzky J, Porte D, Redberg R, Stitzel KF, Stone NJ. Primary prevention of cardiovascular diseases in people with diabetes mellitus: a scientific statement from the American Heart Association and the American Diabetes Association. *Diabetes Care* 2007; **30**: 162-172 [PMID: 17192355]

25 **Grundy SM**, Cleeman JI, Merz CN, Brewer HB, Clark LT, Hunninghake DB, Pasternak RC, Smith SC, Stone NJ. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. *Circulation* 2004; **110**: 227-239 [PMID: 15249516 DOI: 10.1161/01.CIR.0000133317.49796.0E110/2/227]

26 **Brunzell JD**, Davidson M, Furberg CD, Goldberg RB, Howard BV, Stein JH, Witztum JL. Lipoprotein management in patients with cardiometabolic risk: consensus statement from the American Diabetes Association and the American College of Cardiology Foundation. *Diabetes Care* 2008; **31**: 811-822 [PMID: 18375431 DOI: 10.2337/dc08-901831/4/811]

27 **Backes JM**, Gibson CA, Howard PA. Optimal lipid modification: the rationale for combination therapy. *Vasc Health Risk Manag* 2005; **1**: 317-331 [PMID: 17315604]

28 **Elam MB**, Hunninghake DB, Davis KB, Garg R, Johnson C, Egan D, Kostis JB, Sheps DS, Brinton EA. Effect of niacin on lipid and lipoprotein levels and glycemic control in patients with diabetes and peripheral arterial disease: the ADMIT study: A randomized trial. Arterial Disease Multiple Intervention Trial. *JAMA* 2000; **284**: 1263-1270 [PMID: 10979113]

29 **Oki JC**. Dyslipidemias in patients with diabetes mellitus: classification and risks and benefits of therapy. *Pharmacotherapy* 1995; **15**: 317-337 [PMID: 7667166]

30 **Adiels M**, Olofsson SO, Taskinen MR, Borén J. Diabetic dyslipidaemia. *Curr Opin Lipidol* 2006; **17**: 238-246 [PMID: 16680028 DOI: 10.1097/01.mol.0000226115.97436.c000041433-200606000-00007]

31 **Frick MH**, Syvänne M, Nieminen MS, Kauma H, Majahalme S, Virtanen V, Kesäniemi YA, Pasternack A, Taskinen MR. Prevention of the angiographic progression of coronary and vein-graft atherosclerosis by gemfibrozil after coronary bypass surgery in men with low levels of HDL cholesterol. Lopid Coronary Angiography Trial (LOCAT) Study Group. *Circulation* 1997; **96**: 2137-2143 [PMID: 9337181]

32 **de Faire U**, Ericsson CG, Grip L, Nilsson J, Svane B, Hamsten A. Secondary preventive potential of lipid-lowering drugs. The Bezafibrate Coronary Atherosclerosis Intervention Trial (BECAIT). *Eur Heart J* 1996; **17** Suppl F: 37-42 [PMID: 8960446]

33 **Bezafibrate Infarction Prevention (BIP) study**. Secondary prevention by raising HDL cholesterol and reducing triglycerides in patients with coronary artery disease. *Circulation* 2000; **102**: 21-27 [PMID: 10880410]

34 **Keech A**, Simes RJ, Barter P, Best J, Scott R, Taskinen MR, Forder P, Pillai A, Davis T, Glasziou P, Drury P, Kesäniemi YA, Sullivan D, Hunt D, Colman P, d'Emden M, Whiting M, Ehnholm C, Laakso M. Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): randomised controlled trial. *Lancet* 2005; **366**: 1849-1861 [PMID: 16310551]

35 **Bruckert E**, De Gennes JL, Malbecq W, Baigts F. Comparison of the efficacy of simvastatin and standard fibrate therapy in the treatment of primary hypercholesterolemia and combined hyperlipidemia. *Clin Cardiol* 1995; **18**: 621-629 [PMID: 8590530]

36 **Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults.** Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 2001; **285**: 2486-2497 [PMID: 11368702]

37 **Tenenbaum A**, Medvedofsky D, Fisman EZ, Bubyr L, Matetzky S, Tanne D, Klempfner R, Shemesh J, Goldenberg I. Cardiovascular events in patients received combined fibrate/statin treatment versus statin monotherapy: Acute Coronary Syndrome Israeli Surveys data. *PLoS One* 2012; **7**: e35298 [PMID: 22523582 DOI: 10.1371/journal.pone.0035298PONE-D-11-25899]

38 **Moon YS**, Kashyap ML. Pharmacologic treatment of type 2 diabetic dyslipidemia. *Pharmacotherapy* 2004; **24**: 1692-1713 [PMID: 15585439 DOI: 10.1592/phco.24.17.1692.52340]

39 **Taher TH**, Dzavik V, Reteff EM, Pearson GJ, Woloschuk BL, Francis GA. Tolerability of statin-fibrate and statin-niacin combination therapy in dyslipidemic patients at high risk for

cardiovascular events. *Am J Cardiol* 2002; **89**: 390-394 [PMID: 11835917]

40 **Shek A**, Ferrill MJ. Statin-fibrate combination therapy. *Ann Pharmacother* 2001; **35**: 908-917 [PMID: 11485144]

41 **Koh KK**, Quon MJ, Rosenson RS, Chung WJ, Han SH. Vascular and metabolic effects of treatment of combined hyperlipidemia: focus on statins and fibrates. *Int J Cardiol* 2008; **124**: 149-159 [PMID: 17658632]

**P- Reviewer:** Cicero AFG, Sakabe K **S- Editor:** Song XX
**L- Editor:** A **E- Editor:** Wu HL

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.204

*World J Meta-Anal* 2014 November 26; 2(4): 204-211
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.



# Alcohol consumption and the risk of Barrett's esophagus: A meta-analysis

Cong Dai, Wei-Xin Liu, Ke Wang, Hong-Kun Jiang, Min Jiang, Ming-Jun Sun

Cong Dai, Wei-Xin Liu, Min Jiang, Ming-Jun Sun, Department of Gastroenterology, First Affiliated Hospital, China Medical University, Shenyang 110001, Liaoning Province, China
Ke Wang, Department of Cadre Ward V, First Affiliated Hospital, China Medical University, Shenyang 110001, Liaoning Province, China
Hong-Kun Jiang, Department of Pediatrics, First Affiliated Hospital, China Medical University, Shenyang 110001, Liaoning Province, China
Author contributions: Dai C, Wang K and Liu WX carried out the literature search, selection, validity assessment, data abstraction and data analysis; Dai C, Liu WX and Jiang M wrote the paper and incorporated the comments from other authors and peer reviewers; Dai C and Sun MJ had the original idea for the paper, formulated the protocol and contributed to data abstraction and analysis; all authors reviewed and approved the final draft of the paper.
Supported by National Natural Science Foundation of China, No. 81300273 and 81300130
Correspondence to: Ming-Jun Sun, Professor, Department of Gastroenterology, First Affiliated Hospital, China Medical University, 92 Bei'er Road, Heping District, Shenyang 110001, Liaoning Province, China. 273159833@qq.com
Telephone: +86-24-23414369 Fax: +86-24-23414369
Received: July 14, 2014 Revised: November 1, 2014
Accepted: November 7, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To evaluate the possible association between alcohol consumption and Barrett's esophagus (BE).

**METHODS:** We performed a systematic literature search of multiple online electronic databases. Inclusion criteria entailed studies about alcohol and BE. Meta-analysis was conducted to evaluate odds ratio (OR) and 95%CIs for the association between alcohol consumption and BE.

**RESULTS:** Twenty studies comprising 4758 patients with BE were included in the meta-analysis. The risk of BE in patients with alcohol consumption was increased compared with control groups (OR = 1.01; 95%CI: 1.00-1.02), especially in case-control and cohort, European and Asian, and hospital studies, but there was a decreased risk of BE associated with alcohol consumption from American studies (OR = 0.86; 95%CI: 0.77-0.96). At the same time, there was no significant association between BE and alcohol consumption in community studies (OR = 0.97; 95%CI: 0.84-1.12) and the type of alcohol (wine, beer and liquor) studies.

**CONCLUSION:** Our meta-analysis found that alcohol consumption was associated with an increased risk of BE, especially for European and Asian drinkers.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Barrett's esophagus; Alcohol consumption; Risk factors; Meta-analysis

**Core tip:** Barrett's esophagus (BE) is an acquired condition in which specialized columnar epithelium replaces the usual stratified squamous epithelium lining the esophagus. Some studies have investigated the association between alcohol consumption and BE, with conflicting conclusions. We performed a meta-analysis of the possible association between alcohol consumption and BE and found that alcohol consumption was associated with an increased risk of BE, especially for European and Asian drinkers.

Dai C, Liu WX, Wang K, Jiang HK, Jiang M, Sun MJ. Alcohol consumption and the risk of Barrett's esophagus: A meta-analysis. *World J Meta-Anal* 2014; 2(4): 204-211 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/204.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.204

## INTRODUCTION

Barrett's esophagus (BE) is an acquired condition in

**Table 1 Characteristics of included studies**

| Ref. | Year | Country | Source | Study type | Number of BE patients |
|---|---|---|---|---|---|
| Levi *et al*[20] | 1990 | Switzerland | Hospital | Case-control | 140 |
| Gray *et al*[21] | 1993 | United Kingdom | Hospital | Case-control | 58 |
| Eloubeidi *et al*[22] | 2001 | United States | Hospital | Case-control | 88 |
| Gerson *et al*[25] | 2002 | United States | Hospital | Case-control | 24 |
| Conio *et al*[24] | 2002 | Italy | Hospital | Case-control | 147 |
| Avidan *et al*[23] | 2002 | United States | Hospital | Case-control | 1189 |
| Olliver *et al*[26] | 2005 | United Kingdom | Hospital | Case-control | 50 |
| de Jonge *et al*[27] | 2006 | The Netherlands | Hospital | Case-control | 232 |
| Kim *et al*[29] | 2007 | South Korea | Community | Cohort study | 101 |
| Anderson *et al*[28] | 2007 | Ireland | Community | Case-control | 223 |
| Akiyama *et al*[30] | 2008 | Japan | Hospital | Case-control | 211 |
| Anderson *et al*[17] | 2009 | Ireland | Hospital | Case-control | 224 |
| Fouad *et al*[32] | 2009 | Egypt | Hospital | Case-control | 73 |
| Kubo *et al*[18] | 2009 | United States | Community | Case-control | 320 |
| Akiyama *et al*[31] | 2009 | Japan | Hospital | Cohort study | 374 |
| Shi *et al*[39] | 2010 | China | Hospital | Case-control | 179 |
| Thrift *et al*[19] | 2011 | Australia | Community | Case-control | 393 |
| Steevens *et al*[16] | 2011 | The Netherlands | Community | Cohort study | 370 |
| Thrift *et al*[33] | 2014 | Australia | Hospital | Case-control | 258 |
| Yates *et al*[34] | 2014 | United Kingdom | Community | Cohort study | 104 |

BE: Barrett's esophagus.

which specialized columnar epithelium replaces the usual stratified squamous epithelium lining the esophagus[1,2]. BE affects approximately 2% of the Western population and is most prevalent in white men aged over 50 years[3-6]. The prevalence of BE in other populations, such as Asian and African, is uncertain because many patients with BE are asymptomatic and do not accept a diagnostic endoscopy. Among patients with gastroesophageal reflux disease (GERD), the prevalence of BE increases to between 3% and 15%[7,8]. BE is an important risk factor for the development of esophageal adenocarcinoma (EAC)[9] and the incidence of EAC has significantly increased in Western countries over recent decades[10,11]. Moreover, survival for patients with EAC remains poor, with median survival less than 12 mo[12]. Understanding the causes of BE is a necessary step towards preventing EAC.

Potential risk factors for BE in the general population include white race, older age, cigarette smoking, abdominal obesity and GERD[13,14], while use of aspirin and non-steroidal anti-inflammatory drugs (NSAIDs) may prevent the development of BE[15]. There is no meta-analysis about the association between alcohol consumption and BE. If the associations were shown to be causal, then alcohol would be a key target for early prevention of EAC. The results among these studies reporting the association between BE and alcohol consumption have been conflicting and the type of alcohol (wine, beer, liquor) can affect this association. For example, some studies have reported an inverse association with wine consumption, others have found lower risk associated with beer and some evidence for higher risk associated with liquor[16-19].

To confirm the relationship between alcohol consumption and BE, we conducted a meta-analysis of studies which calculated the prevalence of BE in an alcohol consumption population that also provided a quantitative estimate of the risk of BE in an alcohol consumption population. At the same time, we conducted further analyses to examine the association in different study types (case-control, cohort study), countries (Europe, America and Asia), sources (hospital, community) and alcohol type (wine, beer and liquor).

## MATERIALS AND METHODS

### Literature search

We conducted the meta-analysis with the standard protocol for study identification, inclusion, data abstraction and analysis. The PubMed, EMBASE, MEDLINE and Web of Science databases were used primarily to identify potentially relevant published studies. A search of studies in these databases from inception through to 30 October 2014 was performed using medical subject headings including "Barrett's esophagus", "BE", "risk factors", "alcohol", "wine", "beer" and "liquor". The search was not limited by language.

### Study selection

Inclusion criteria were as follows: (1) case-control, cross-sectional and cohort study designs; (2) BE was diagnosed by endoscopy and confirmed that specialized columnar epithelium replaced the usual stratified squamous epithelium lining the esophagus; (3) the prevalence of BE in alcohol consumption groups and control groups was examined; and (4) the odds ratio (OR) and the 95%CI was reported or these data could be calculated.

The information collected from the studies included the publication year, country, source, study design and number of BE patients (Table 1). The outcome of the analysis was the OR of the prevalence of BE in an alcohol consumption group versus a control group and subgroup analysis was conducted in different study types (case-control, cohort study), countries (Europe, America

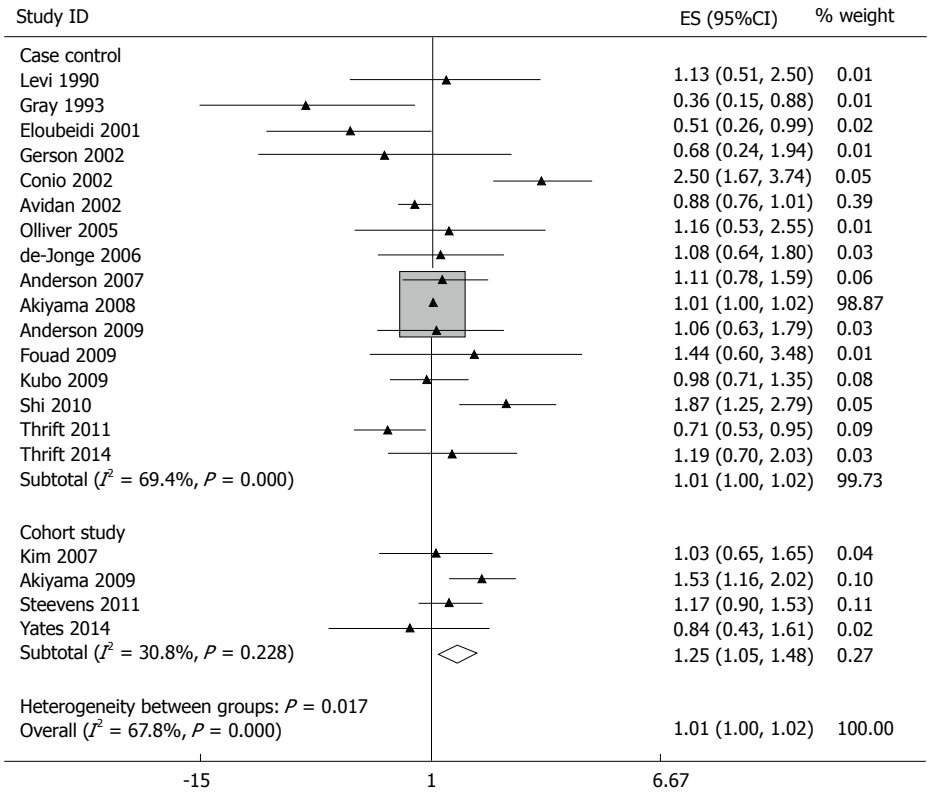



**Figure 1 Forest plot of the prevalence of Barrett's esophagus in alcohol consumption groups *vs* control groups and subgroup analysis by study type (case-control and cohort study).**

and Asia), sources (hospital, community) and alcohol type (wine, beer and liquor). Included studies were collected in EndNote X4 software. Two investigators browsed the titles and abstracts of the studies and studies that met the criteria were further analyzed for eligibility. For studies that met the inclusion criteria of the meta-analysis, we extracted the data using standard operating norms. Information about qualifying studies included the publication year, country, source, study design and number of BE patients. Inconsistent results were determined by the senior investigator. The consistency rate among the authors was 0.994. The methodological quality of each trial was assessed by the following criteria: study design, method of BE diagnosis and method of patient enrollment.

### *Data synthesis and analysis*

The inconsistency index ($I^2$) was calculated to assess the inter-study heterogeneity and significant difference was set at the $P < 0.10$ level. We chose the random-effects model to calculate the pooled OR if there was heterogeneity among these studies and the fixed-effects model to calculate the pooled OR if there was homogeneity. Summary OR and 95%CI were calculated to describe the relationship between alcohol consumption and the risk of BE. We chose a funnel plot analysis to estimate publication bias. The meta-analyses were done using STATA 12.0 software.

## RESULTS

Our systematic review identified 38 studies. After we carefully reviewed these studies, twenty met the inclusion criteria. Information regarding the author, year, country, source, study type and number of BE patients of the included studies can be found in Table 1. There are two types of studies, case-control and cohort. The studies come from different geographical regions, with 10 European, 6 American and 4 Asian studies. The number of the study population ranged from 97 to 4736 and BE patients ranged from 24 to 1189.

Twenty studies were included in our meta-analysis for the prevalence of BE in comparing alcohol consumption groups with control groups. BE patients were more likely to be in alcohol consumption groups than control groups (OR = 1.01; 95%CI: 1.00-1.02) (Figure 1). There was significant heterogeneity ($P$ = 0.017). The funnel plot (Figure 2) is generally symmetrical, showing no evidence of publication bias (Begg's Test, $Z$ = 0.36, $P$ = 0.721). The subgroup analysis of sixteen case-control studies showed an increased risk of BE associated with alcohol consumption (OR = 1.01; 95%CI: 1.00-1.02) (Figure 1) and the four cohort studies also showed an increased risk of BE associated with alcohol consumption (OR = 1.25; 95%CI: 1.05-1.48) (Figure 1).

The subgroup analysis of ten European studies showed an increased risk of BE associated with alcohol consumption (OR = 1.21; 95%CI: 1.04-1.41) (Figure 3), six American studies showed a decreased risk of BE associated with alcohol consumption (OR = 0.86; 95%CI: 0.77-0.96) (Figure 3) and the four Asian studies showed an increased risk of BE associated with alcohol con-

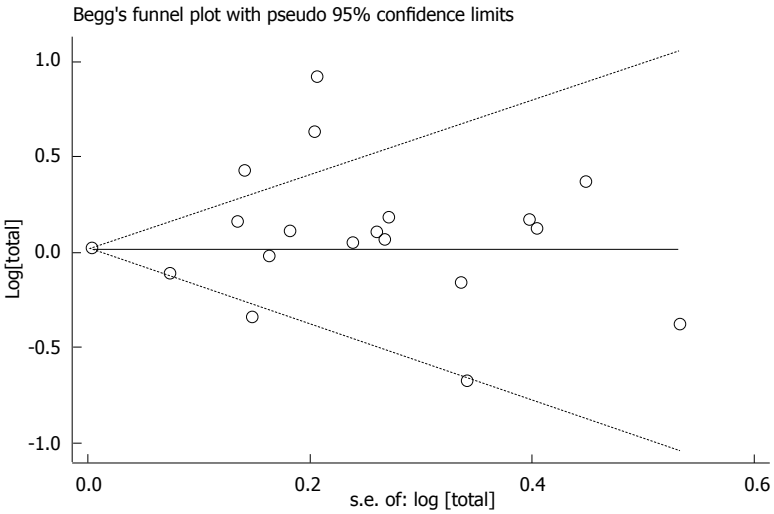



**Figure 2 Begg's funnel plot analysis of each study.**

sumption (OR = 1.01; 95%CI: 1.00-1.02) (Figure 3).

The subgroup analysis of fourteen hospital studies showed an increased risk of BE associated with alcohol consumption (OR = 1.01; 95%CI: 1.00-1.02) (Figure 4) and the six community studies did not show an increased risk of BE associated with alcohol consumption (OR = 0.97; 95%CI: 0.84-1.12) (Figure 4).

At the same time, the type of alcohol consumption is divided into wine, beer and liquor. The subgroup analysis of five studies about wine did not show an increased risk of BE associated with alcohol consumption (OR = 1.00; 95%CI: 0.85-1.17) (Figure 5), the five studies about beer did not showed an increased risk of BE associated with alcohol consumption (OR = 0.94; 95%CI: 0.79-1.12) (Figure 5) and the five studies about liquor did not show an increased risk of BE associated with alcohol consumption (OR = 1.10; 95%CI: 0.93-1.29) (Figure 5).

## DISCUSSION

Lifestyle factors, such as cigarette smoking, alcohol consumption and exercise, are increasingly important for many diseases. Alcohol consumption is considered to be an important risk factor for gastric cancer, cirrhosis and liver cancer. Although there are many studies about the association between alcohol consumption and BE[17,18,20-34], there is no exact definition. Our meta-analysis identified numerous studies examining the relationship between BE and alcohol consumption. We found that BE patients were more likely to be in alcohol consumption groups than control groups. At the same time, we conducted further analyses to examine the association in different study types (case-control, cohort study), countries (Europe, America and Asia), sources (hospital, community) and alcohol type (wine, beer and liquor). We found that there was an increased risk of BE associated with alcohol consumption from case-control and cohort, European and Asian, and hospital studies, but there was a decreased risk of BE associated with alcohol consumption from American studies. At the same time, there was no significant association between BE and alcohol consumption in community studies and the type of alcohol (wine, beer and liquor) studies.

There are several potential mechanisms through which alcohol consumption may be associated with BE, an early event in the development of esophageal adenocarcinoma. First, gastroesophageal reflux (GER) is the main predisposing risk factor for BE. Alcohol, which has been shown to reduce gastroesophageal sphincter pressure, can lead to an increase in GER and the risk of developing BE. This is reflected in the 1.6 to 3.0-fold increased risk of GER symptoms in people who drink alcohol. However, cautious interpretation of this finding is advised because of the lack of internal consistency. Alcohol consumption was significantly lower in population controls reporting GER symptoms, suggesting that patients with GER symptoms may avoid or reduce their alcohol intake after the development of GER symptoms. Second, there are different results in drinking patterns of alcohol. Wine drinkers are more likely than liquor drinkers to consume their alcoholic beverage with food. Consumption of alcohol with food may reduce the direct damage of the lining of esophagus, reducing the carcinogenesis process. At the same time, red wine contains compounds, such as polyphenol, that have a significant protective effect on oxidative stress[35]. Some clinical and animal studies have demonstrated that anti-oxidants may decrease the risk of BE and prevent the development of esophageal adenocarcinoma[36-38]. Drinking wine may reduce the oxidative damage caused by GERD, thereby decreasing the risk of esophagitis and BE among GERD patients, and the protective effects of alcohol consumption may arise from reduction in insulin resistance or increased levels of lipoproteins. Our results showed that there was an increased risk of BE associated with alcohol consumption from European and Asian studies, but there was a decreased risk of BE associated with alcohol consumption from American studies. This is because Americans are more likely than Europeans and Asians to drink wine. However, our meta-analysis showed that there was no significant association between BE and alcohol consumption in the type of alcohol (wine, beer and

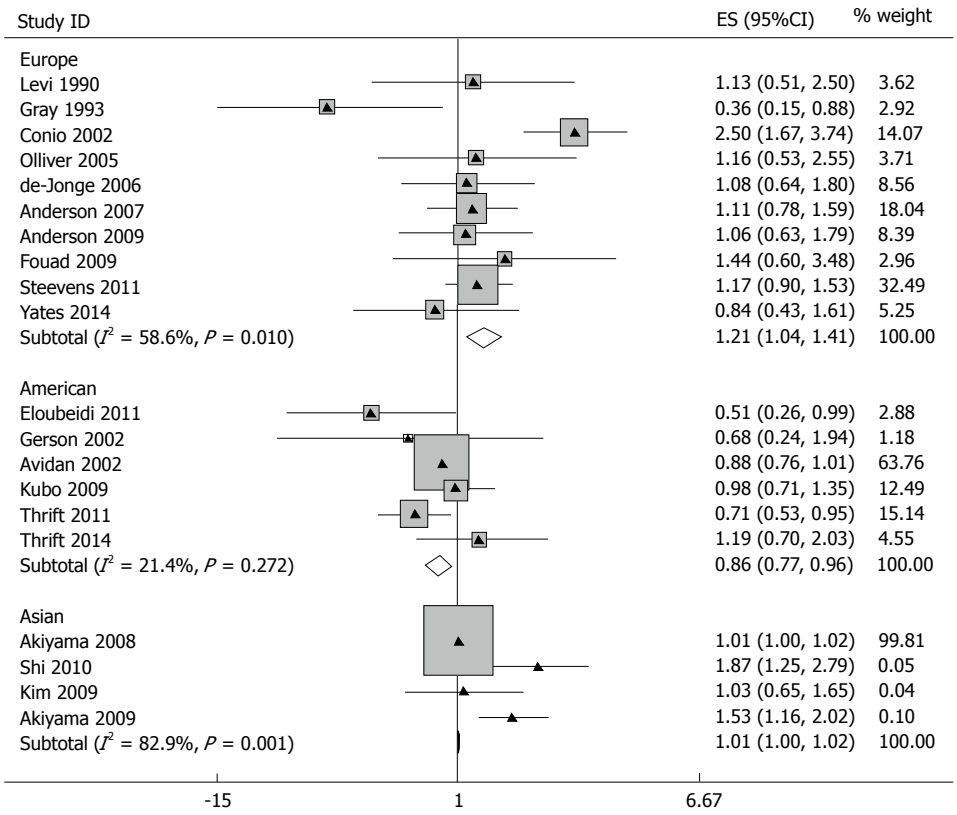



**Figure 3 Forest plot of the prevalence of Barrett's esophagus in alcohol consumption groups *vs* control groups in different geographical regions (Europe, America, Asia).**

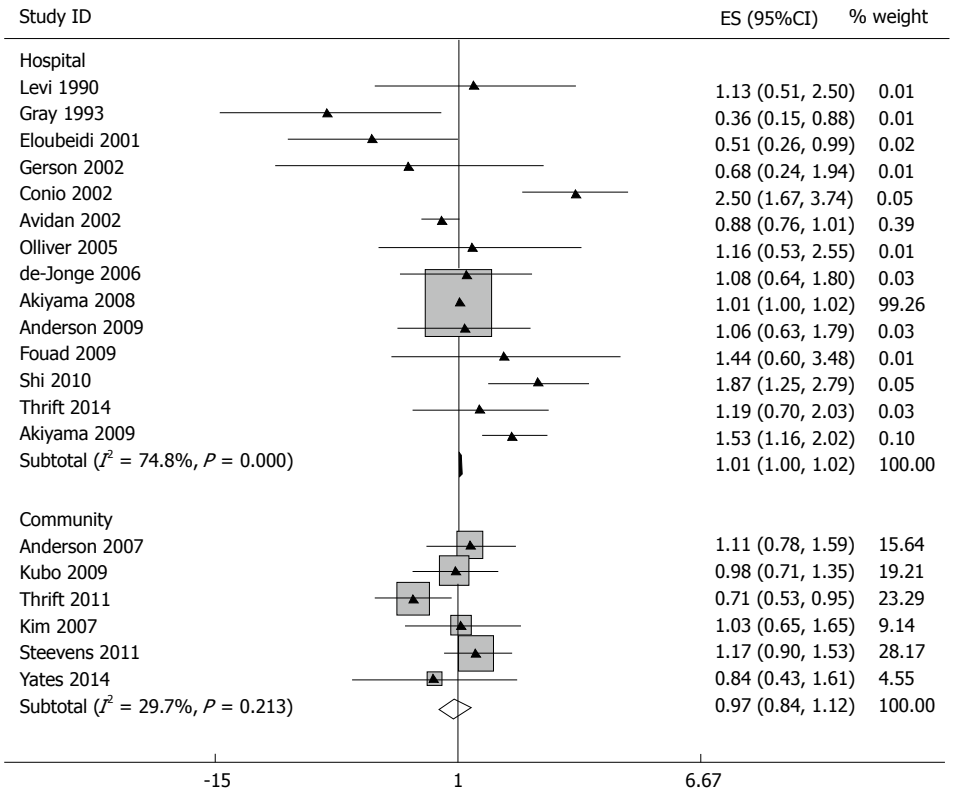



**Figure 4 Forest plot of the prevalence of Barrett's esophagus in alcohol consumption groups *vs* control groups from different sources (hospital, community).**

liquor) studies so we cannot use these theories to explain our results; perhaps this is due to the small sample size.

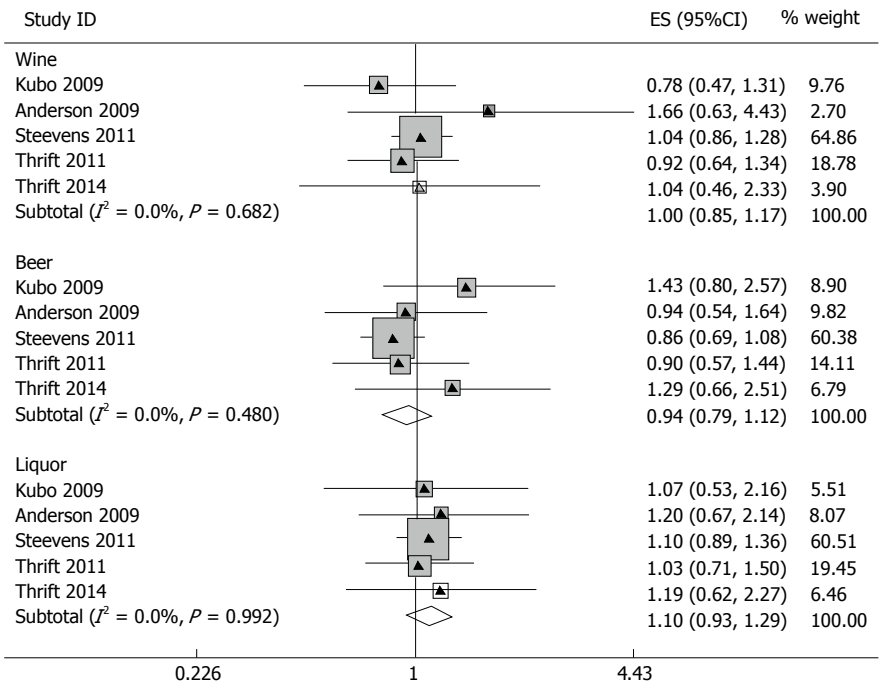



**Figure 5 Forest plot of the prevalence of Barrett's esophagus in alcohol consumption groups *vs* control groups in different type of alcohol (wine, beer and liquor).**

In addition, some unmeasured lifestyle factors are used to explain the association between alcohol consumption and BE. For example, wine drinkers were more likely to be educated and have other markers of a healthy lifestyle, such as a better diet and vitamin use, while beer and liquor drinkers were more likely to engage in unhealthier lifestyles, such as eating fewer servings of fruit and vegetables and having a higher body mass index (BMI).

Although demographic and lifestyle factors were associated with Barrett's esophagus, these factors did not appear to explain the association between alcohol consumption and Barrett's esophagus. The association between alcohol consumption and Barrett's esophagus should also consider other factors, such as income, education, smoking, BMI, Helicobacter pylori, fruit and vegetable intake, vitamin supplement intake, NSAIDs use and television watching. For example, *Helicobacter pylori* infection can increase the activity of alcohol dehydrogenase (ADH) in the gastric epithelium. Ethanol is metabolized in the esophageal mucosa to acetaldehyde mainly by the action of ADH, which has been shown to have strong carcinogenic potential. In addition, homozygosity for the $ADH_3$ gene has recently been reported in patients with BE and could increase the rate of conversion of ethanol to ADH in the esophageal mucosa.

At the same time, our meta-analysis has some limitations. First, there were some bias and confounders in the included studies. For example, some case-control studies were subject to recall and selection bias which are inherent to retrospective studies. However, the recall bias with regards to alcohol consumption in the included studies was not a significant bias as recall of alcohol consumption is reliable among study participants and our meta-analysis included four prospective cohort studies which can reduce the risk of recall and selection bias. The second is the lack of information of the presence of GERD and the use of medications for this disease. Consequently, we were not able to investigate the possible confounding effects of these factors or their possible intermediate role in the associations investigated. There is also the lack of information about the length of alcohol consumption time in the included studies.

In conclusion, our meta-analysis demonstrated that alcohol consumption is a risk factor for the development of BE. Since alcohol consumption is a modifiable risk factor, it is an important concept in terms of BE prevention. Future studies should investigate the risk factors for progression of BE to EAC.

## COMMENTS

### *Background*

Barrett's esophagus (BE) is an acquired condition in which specialized columnar epithelium replaces the usual stratified squamous epithelium lining the esophagus. The main clinical significance of BE is its association with an increased risk of developing esophageal adenocarcinoma (EAC). Survival for patients with EAC remains poor, with median survival less than 12 mo. Understanding the causes of BE is a necessary step toward preventing EAC.

### *Research frontiers*

Potential risk factors for BE include white race, male sex, older age, cigarette smoking, abdominal obesity and gastroesophageal reflux disease, while use of aspirin and non-steroidal anti-inflammatory drugs may prevent the development of BE. To date, no meta-analysis of the relationship between alcohol consumption and BE has been performed. The results among these studies reporting the association between BE and alcohol consumption have been conflicting and the type of alcohol (wine, beer, liquor) can affect this association.

### *Innovations and breakthroughs*

Alcohol consumption was associated with an increased risk of BE compared

with control groups (OR = 1.01; 95%CI: 1.00-1.02). Subgroup analysis showed an increased risk of BE associated with alcohol consumption from case-control and cohort, European and Asian, and hospital studies, but there was a decreased risk of BE associated with alcohol consumption from American studies (OR = 0.86; 95%CI: 0.77-0.96). At the same time, there was no significant association between BE and alcohol consumption in community studies (OR = 0.97; 95%CI: 0.84-1.12) and the type of alcohol (wine, beer and liquor) studies.

### Applications

Alcohol consumption was associated with an increased risk of BE, especially for European and Asian drinkers.

### Terminology

BE is an acquired condition in which specialized columnar epithelium replaces the usual stratified squamous epithelium lining the esophagus.

### Peer review

This manuscript about alcohol consumption and BE is interesting for the readership and should be published.

## REFERENCES


1 **Spechler SJ**. Clinical practice. Barrett's Esophagus. *N Engl J Med* 2002; **346**: 836-842 [PMID: 11893796 DOI: 10.1056/NEJMcp012118]

2 **Wang KK**, Sampliner RE. Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. *Am J Gastroenterol* 2008; **103**: 788-797 [PMID: 18341497 DOI: 10.1111/j.1572-0241.2008.01835.x]

3 **Cook MB**, Chow WH, Devesa SS. Oesophageal cancer incidence in the United States by race, sex, and histologic type, 1977-2005. *Br J Cancer* 2009; **101**: 855-859 [PMID: 19672254 DOI: 10.1038/sj.bjc.6605246]

4 **Devesa SS**, Blot WJ, Fraumeni JF. Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. *Cancer* 1998; **83**: 2049-2053 [PMID: 9827707 DOI: 10.1002/(SICI)1097-0142(19981115)83]

5 **Lepage C**, Rachet B, Jooste V, Faivre J, Coleman MP. Continuing rapid increase in esophageal adenocarcinoma in England and Wales. *Am J Gastroenterol* 2008; **103**: 2694-2699 [PMID: 18853967 DOI: 10.1111/j.1572-0241.2008.02191.x]

6 **Bosetti C**, Levi F, Ferlay J, Garavello W, Lucchini F, Bertuccio P, Negri E, La Vecchia C. Trends in oesophageal cancer incidence and mortality in Europe. *Int J Cancer* 2008; **122**: 1118-1129 [PMID: 17990321 DOI: 10.1002/ijc.23232]

7 **Cameron AJ**. Epidemiology of columnar-lined esophagus and adenocarcinoma. *Gastroenterol Clin North Am* 1997; **26**: 487-494 [PMID: 9309399]

8 **Hirota WK**, Loughney TM, Lazas DJ, Maydonovitch CL, Rholl V, Wong RK. Specialized intestinal metaplasia, dysplasia, and cancer of the esophagus and esophagogastric junction: prevalence and clinical data. *Gastroenterology* 1999; **116**: 277-285 [PMID: 9922307]

9 **Lagergren J**, Bergström R, Lindgren A, Nyrén O. Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. *N Engl J Med* 1999; **340**: 825-831 [PMID: 10080844 DOI: 10.1056/NEJM199903183401101]

10 **Pohl H**, Welch HG. The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. *J Natl Cancer Inst* 2005; **97**: 142-146 [PMID: 15657344 DOI: 10.1093/jnci/dji024]

11 **Thrift AP**, Whiteman DC. The incidence of esophageal adenocarcinoma continues to rise: analysis of period and birth cohort effects on recent trends. *Ann Oncol* 2012; **23**: 3155-3162 [PMID: 22847812 DOI: 10.1093/annonc/mds181]

12 **Polednak AP**. Trends in survival for both histologic types of esophageal cancer in US surveillance, epidemiology and end results areas. *Int J Cancer* 2003; **105**: 98-100 [PMID: 12672037 DOI: 10.1002/ijc.11029]

13 **Edelstein ZR**, Farrow DC, Bronner MP, Rosen SN, Vaughan TL. Central adiposity and risk of Barrett's esophagus. *Gastroenterology* 2007; **133**: 403-411 [PMID: 17681161 DOI: 10.1053/j.gastro.2007.05.026]

14 **Demeester SR**, Peters JH, Demeester TR. Barrett's esophagus. *Curr Probl Surg* 2001; **38**: 558-640 [PMID: 11468640 DOI: 10.1067/msg.2001.115514]

15 **Anderson LA**, Johnston BT, Watson RG, Murphy SJ, Ferguson HR, Comber H, McGuigan J, Reynolds JV, Murray LJ. Nonsteroidal anti-inflammatory drugs and the esophageal inflammation-metaplasia-adenocarcinoma sequence. *Cancer Res* 2006; **66**: 4975-4982 [PMID: 16651456 DOI: 10.1158/0008-5472.CAN-05-4253]

16 **Steevens J**, Schouten LJ, Driessen AL, Huysentruyt CJ, Keulemans YC, Goldbohm RA, van den Brandt PA. A prospective cohort study on overweight, smoking, alcohol consumption, and risk of Barrett's esophagus. *Cancer Epidemiol Biomarkers Prev* 2011; **20**: 345-358 [PMID: 21173169 DOI: 10.1158/1055-9965.EPI-10-0636]

17 **Anderson LA**, Cantwell MM, Watson RG, Johnston BT, Murphy SJ, Ferguson HR, McGuigan J, Comber H, Reynolds JV, Murray LJ. The association between alcohol and reflux esophagitis, Barrett's esophagus, and esophageal adenocarcinoma. *Gastroenterology* 2009; **136**: 799-805 [PMID: 19162028 DOI: 10.1053/j.gastro.2008.12.005]

18 **Kubo A**, Levin TR, Block G, Rumore GJ, Quesenberry CP, Buffler P, Corley DA. Alcohol types and sociodemographic characteristics as risk factors for Barrett's esophagus. *Gastroenterology* 2009; **136**: 806-815 [PMID: 19111726 DOI: 10.1053/j.gastro.2008.11.042]

19 **Thrift AP**, Pandeya N, Smith KJ, Mallitt KA, Green AC, Webb PM, Whiteman DC. Lifetime alcohol consumption and risk of Barrett's Esophagus. *Am J Gastroenterol* 2011; **106**: 1220-1230 [PMID: 21427711 DOI: 10.1038/ajg.2011.89]

20 **Levi F**, Ollyo JB, La Vecchia C, Boyle P, Monnier P, Savary M. The consumption of tobacco, alcohol and the risk of adenocarcinoma in Barrett's oesophagus. *Int J Cancer* 1990; **45**: 852-854 [PMID: 2335388]

21 **Gray MR**, Donnelly RJ, Kingsnorth AN. The role of smoking and alcohol in metaplasia and cancer risk in Barrett's columnar lined oesophagus. *Gut* 1993; **34**: 727-731 [PMID: 8314502]

22 **Eloubeidi MA**, Provenzale D. Clinical and demographic predictors of Barrett's esophagus among patients with gastroesophageal reflux disease: a multivariable analysis in veterans. *J Clin Gastroenterol* 2001; **33**: 306-309 [PMID: 11588545]

23 **Avidan B**, Sonnenberg A, Schnell TG, Chejfec G, Metz A, Sontag SJ. Hiatal hernia size, Barrett's length, and severity of acid reflux are all risk factors for esophageal adenocarcinoma. *Am J Gastroenterol* 2002; **97**: 1930-1936 [PMID: 12190156 DOI: 10.1111/j.1572-0241.2002.05902.x]

24 **Conio M**, Filiberti R, Blanchi S, Ferraris R, Marchi S, Ravelli P, Lapertosa G, Iaquinto G, Sablich R, Gusmaroli R, Aste H, Giacosa A. Risk factors for Barrett's esophagus: a case-control study. *Int J Cancer* 2002; **97**: 225-229 [PMID: 11774268 DOI: 10.1002/ijc.1583]

25 **Gerson LB**, Shetler K, Triadafilopoulos G. Prevalence of Barrett's esophagus in asymptomatic individuals. *Gastroenterology* 2002; **123**: 461-467 [PMID: 12145799]

26 **Olliver JR**, Hardie LJ, Gong Y, Dexter S, Chalmers D, Harris KM, Wild CP. Risk factors, DNA damage, and disease progression in Barrett's esophagus. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 620-625 [PMID: 15767340 DOI: 10.1158/1055-9965.EPI-04-0509]

27 **de Jonge PJ**, Steyerberg EW, Kuipers EJ, Honkoop P, Wolters LM, Kerkhof M, van Dekken H, Siersema PD. Risk factors for the development of esophageal adenocarcinoma in Barrett's esophagus. *Am J Gastroenterol* 2006; **101**: 1421-1429 [PMID: 16863542 DOI: 10.1111/j.1572-0241.2006.00626.x]

28 **Anderson LA**, Watson RG, Murphy SJ, Johnston BT, Comber H, Mc Guigan J, Reynolds JV, Murray LJ. Risk factors for Barrett's oesophagus and oesophageal adenocarcinoma: results from the FINBAR study. *World J Gastroenterol* 2007; **13**: 1585-1594 [PMID: 17461453]

29 **Kim JH**, Rhee PL, Lee JH, Lee H, Choi YS, Son HJ, Kim JJ, Rhee JC. Prevalence and risk factors of Barrett's esophagus in Korea. *J Gastroenterol Hepatol* 2007; **22**: 908-912 [PMID: 17565647 DOI: 10.1111/j.1440-1746.2006.04448.x]

30 **Akiyama T**, Inamori M, Iida H, Mawatari H, Endo H, Hosono K, Yoneda K, Fujita K, Yoneda M, Takahashi H, Goto A, Abe Y, Kobayashi N, Kubota K, Saito S, Nakajima A. Alcohol consumption is associated with an increased risk of erosive esophagitis and Barrett's epithelium in Japanese men. *BMC Gastroenterol* 2008; **8**: 58 [PMID: 19077221 DOI: 10.1186/1471-230X-8-58]

31 **Akiyama T**, Inamori M, Akimoto K, Iida H, Mawatari H, Endo H, Ikeda T, Nozaki Y, Yoneda K, Sakamoto Y, Fujita K, Yoneda M, Takahashi H, Hirokawa S, Goto A, Abe Y, Kirikoshi H, Kobayashi N, Kubota K, Saito S, Nakajima A. Risk factors for the progression of endoscopic Barrett's epithelium in Japan: a multivariate analysis based on the Prague C & amp; M Criteria. *Dig Dis Sci* 2009; **54**: 1702-1707 [PMID: 19003532 DOI: 10.1007/s10620-008-0537-y]

32 **Fouad YM**, Makhlouf MM, Tawfik HM, el-Amin H, Ghany WA, el-Khayat HR. Barrett's esophagus: prevalence and risk factors in patients with chronic GERD in Upper Egypt. *World J Gastroenterol* 2009; **15**: 3511-3515 [PMID: 19630106]

33 **Thrift AP**, Kramer JR, Richardson PA, El-Serag HB. No significant effects of smoking or alcohol consumption on risk of Barrett's esophagus. *Dig Dis Sci* 2014; **59**: 108-116 [PMID: 24114046 DOI: 10.1007/s10620-013-2892-6]

34 **Yates M**, Cheong E, Luben R, Igali L, Fitzgerald R, Khaw KT, Hart A. Body mass index, smoking, and alcohol and risks of Barrett's esophagus and esophageal adenocarcinoma: a UK prospective cohort study. *Dig Dis Sci* 2014; **59**: 1552-1559 [PMID: 24500448 DOI: 10.1007/s10620-013-3024-z]

35 **Rivero-Pérez MD**, Muñiz P, Gonzalez-Sanjosé ML. Antioxidant profile of red wines evaluated by total antioxidant capacity, scavenger activity, and biomarkers of oxidative stress methodologies. *J Agric Food Chem* 2007; **55**: 5476-5483 [PMID: 17579427 DOI: 10.1021/jf070306q]

36 **Chen X**, Mikhail SS, Ding YW, Yang Gy, Bondoc F, Yang CS. Effects of vitamin E and selenium supplementation on esophageal adenocarcinogenesis in a surgical model with rats. *Carcinogenesis* 2000; **21**: 1531-1536 [PMID: 10910955]

37 **Kubo A**, Corley DA. Meta-analysis of antioxidant intake and the risk of esophageal and gastric cardia adenocarcinoma. *Am J Gastroenterol* 2007; **102**: 2323-2330; quiz 2331 [PMID: 17581269 DOI: 10.1111/j.1572-0241.2007.01374.x]

38 **Kubo A**, Levin TR, Block G, Rumore GJ, Quesenberry CP, Buffler P, Corley DA. Dietary patterns and the risk of Barrett's esophagus. *Am J Epidemiol* 2008; **167**: 839-846 [PMID: 18218607 DOI: 10.1093/aje/kwm381]

39 **Shi XS**, Wang LJ, Hao HJ, Wang LJ, Zhang YL, Ma CY. Study on Reflux Esophagitis, Barrett's Esophagus and Esophageal Adenocarcinoma. *Chin J Gastroenterol* 2010; **15**: 233-236


**P- Reviewer:** Theisen J **S- Editor:** Ji FF **L- Editor:** Roemmele A **E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.212

*World J Meta-Anal* 2014 November 26; 2(4): 212-220
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Using fluorodeoxy-D-glucose-positron emission tomography to monitor neoadjuvant chemotherapy response in sarcoma: A meta-analysis

Yu-Ting Wang, Hong Pu, Long-Lin Yin, Jia-Yuan Chen

Yu-Ting Wang, Hong Pu, Long-Lin Yin, Jia-Yuan Chen, Department of Radiology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu 610072, Sichuan Province, China
Author contributions: Wang YT and Chen JY designed the research; Wang YT, Pu H and Yin LL performed the research; Wang YT and Pu H analyzed the data; Wang YT wrote the manuscript.
Supported by Scientific Research Subject of Health Department of Sichuan, China, No. 070045
Correspondence to: Yu-Ting Wang, MD, Department of Radiology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, No. 32, Section 2, 1st Ring Road (West), Chengdu 610072, Sichuan Province, China. wangyuting_330@163.com
Telephone: +86-28-87394280
Received: May 13, 2014 Revised: August 11, 2014
Accepted: September 4, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To systematically evaluate the accuracy of 18-fluorodeoxy-D-glucose-positron emission tomography (18-FDG PET) to assess response to neoadjuvant chemotherapy in bone and soft tissue sarcomas.

**METHODS:** Studies published in English language regarding the accuracy of F-18 FDG PET for the indication were retrieved from MEDLINE. The QUADAS tool was utilized for methodological quality appraisal. Relevant data were extracted, and quantitative data synthesis included pooled estimation and subgroup analysis.

**RESULTS:** A total of fifteen studies involving 420 patients with pathologically confirmed sarcoma were collected. Methodological quality was relatively high. The pooled sensitivity and specificity of PET to predict histopathological response were 87% (95%CI: 81%-91%) and 83% (95%CI: 77%-87%), respectively. Ten studies employed a lower standardized uptake value (SUV) after chemotherapies (mostly 2.5) and/or a higher SUV reduction rate (mostly around 50%) as PET criteria of good response. Subgroup analysis showed that PET exhibited a significantly better specificity in osteosarcoma (OS) and Ewing sarcoma (ES) than in soft-tissue sarcoma (STS) (91% *vs* 75%, $P < 0.05$), and a higher specificity in pediatric patients than in adults (90% *vs* 74%, $P < 0.01$). PET yielded a lower specificity in ifosfamide-contained chemotherapies than in the alternative regimen (70% *vs* 97%, $P < 0.01$).

**CONCLUSION:** F-18 FDG PET is promising to predict neoadjuvant therapy response in sarcoma, especially in pediatric patients with OS or ES. Certain chemotherapeutic agents could potentially cause false positives of PET.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** 18-fluorodeoxy-D-glucose-positron emission tomography; Sarcoma; Therapy monitoring; Meta-analysis

**Core tip:** Histopathological response to neoadjuvant chemotherapy is a significant prognostic factor and could guild the following treatment for patients with sarcoma. Novel molecular imaging such as 18-fluorodeoxy-D-glucose-positron emission tomography (PET) could offer non-invasive and accurate assessment of such response in sarcoma based on our relatively large-sample statistical analysis for this relatively rare tumor. More information that clinical physicians concerned with was generated, such as recommended subgroup of patients, potential candidates for quantitative criteria of response and possible causes of false positives of PET.

Wang YT, Pu H, Yin LL, Chen JY. Using fluorodeoxy-D-glucose-positron emission tomography to monitor neoadjuvant chemo-

therapy response in sarcoma: A meta-analysis. *World J Meta-Anal* 2014; 2(4): 212-220 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/212.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.212

## INTRODUCTION

Soft tissue and bone sarcomas account for about 1% of malignancies and remain a diagnostic and therapeutic challenge. Osteosarcoma (OS) and Ewing's sarcoma (ES) are the most common malignant primary bone tumors in children and adolescents, while soft tissue sarcomas (STS) are a heterogeneous group of malignancies that exhibit mesenchymal differentiation, and may occur in both children and the elderly[1].

While sarcoma is not a common tumor, it had a high mortality rate[2]. The rate of systemic spread is so high that surgical treatment alone was rarely enough for cure. The introduction of neoadjuvant chemotherapy (NCT) brought a significant improvement in the long-term survival of patients and was therefore widely used in clinical practice[3,4]. Although histopathological response to neoadjuvant chemotherapy (mostly defined by tumor necrosis of more than 90%) was showed to be associated with a significantly higher survival rate, it is known that only a subcollective of patients benefit from NCT (32%-88% for OS and ES, 19%-48% for STS)[5]. Therefore, there is an urgent need for non-invasive methods which allow accurate identification of responses to NCT, and accordingly determine whether to switch to a more intensified chemotherapy regimen, and also to determine the most appropriate surgical approach.

Radiologic imaging has some limitations for this utility; for instance, the OS does not suffer evident shrinkage after chemotherapy[6]. In contrast to conventional anatomic imaging such as MRI and CT, [F-18]-fluorodeoxy-D-glucose (FDG)-positron emission tomography (PET) has the unique ability to detect response at the molecular level, and is being increasingly used in patients with sarcoma for therapy monitoring[7].

Nonetheless, the sample sizes in conducted trials were relatively small and the results among them were remarkably discrepant. We performed this meta-analysis to systematically estimate the accuracy of F-18 FDG PET to assess histopathological response to neoadjuvant therapy in patients with bone and soft tissue sarcomas.

## MATERIALS AND METHODS

### Literature search

This meta-analysis was performed following guidelines for the systematic review and meta-analysis of diagnostic studies[8]. Ethical committee approval and patients' consents were not required for this work due to its statistic nature. MEDLINE (January 1980 to August 2012) were used for searching English-language articles with the following keywords: ("positron emission tomography OR fluorodeoxyglucose") AND ("prediction OR response OR response monitoring OR neoadjuvant OR chemotherapy") AND ("carcinoma OR cancer OR tumor OR tumour") AND ("sensitivity OR specificity OR accuracy") AND ("sarcoma"). Two reviewers independently checked the abstracts of retrieved publications and obtained the fulltext of each eligible article. References of the retrieved studies were screened to expand the scope of the computerized literature search. Disagreements were resolved by discussion between the two reviewers.

Studies were eligible for inclusion based on the following criteria: (1) using F-18 FDG PET to assess response to neoadjuvant therapies in patients with primary bone or soft tissue sarcoma; (2) employing pathological outcome as the "gold standard"; and (3) clinical raw data could be drawn from data in the original published study. A minimal sample size for inclusion was not required given the relative rareness of sarcomas in clinical scenarios. Conference abstracts, case reports, comments, and letters to the editors were excluded.

### Quality appraisal and data extraction

To identify the methodological quality of included article studies, two reviewers independently assessed each article using the QUADAS tool, which is specifically designed and widely used in systematic reviews of diagnostic accuracy studies[9]. It provides a checklist of items regarding the representativeness and methodological quality of investigated studies and presents an accumulated score as a result. Description of the execution of PET scans and the pathological interpretation were assessed. The optimal time interval between F-18 FDG PET and histopathological reading was defined as less than 30 d. Whether the pathological interpretation was blind to the PET scan results was also recorded. However, the PET readers were naturally blind to pathological information, since these scans were performed before the pathological results from following surgeries.

Two reviewers independently extracted the following data by using a standardized data extraction form: (1) characteristics of the study, including investigated tumor pathology (OS or STS), sample size, patient age, timing for PET scans, specific criteria of PET and pathology, *etc.*; (2) methodological details which needed to be examined with the QUADAS tool; and (3) clinical outcome. A 2 × 2 contingency table was constructed consisting of true-positive, false-positive, true-negative and false-negative for each study. We considered response to therapies as "positive", while therapy failures as "negative".

### Statistical analysis

Based on the 2 × 2 contingency tables, sensitivity was calculated as the ratio of the number of true-positive findings showed in the index test to the number of all positive findings supported by the golden standard. Specificity was calculated as the ratio of the number of true-negative findings reported by the index test to the number of all negative findings showed in the reference standard. The diagnostic odds ratio (DOR) was calculated as [True positive (TP) × True negative (TN)]/[False posi-

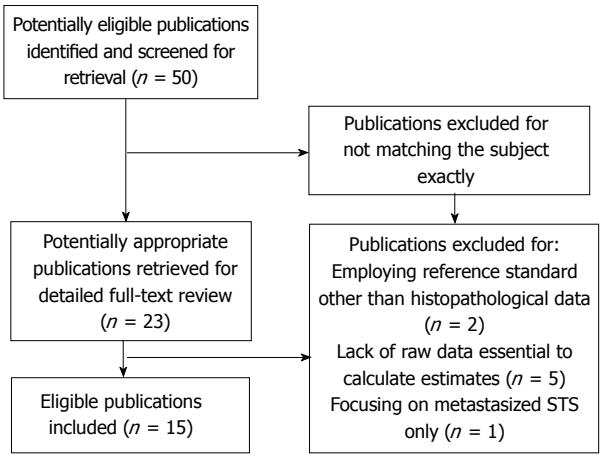



**Figure 1 Flow diagram of publication selection for this meta-analysis.** STS: Soft-tissue sarcoma.

tive (FP)× False negative (FN)], where TP, TN, FP and FN stand for true-positive, true-negative, false-positive and false-negative, respectively. If the DOR could not be calculated because one of the cells in the 2 × 2 table was empty, 0.5 was added to all cells[10]. We used the $\chi^2$ test to assess the heterogeneity among studies. The $I^2$ index was employed, with an $I^2$ value of over 50% indicating statistical heterogeneity (a random-effects model was accordingly utilized), and an $I^2$ value of less than 50% suggesting that homogeneity was reached (a fixed-effects model was then applied)[11,12].

If statistical heterogeneity was found, subgroup analyses were then performed to explore its possible sources (*e.g.*, different tumor pathology, patient age group and chemotherapy regimen). We employed the $Z$ test to examine the statistical differences between subgroups. Two-sided $P < 0.05$ was considered statistically significant.

Our summary receiver operating characteristic curves (SROC) was plotted in accordance with Mose's linear model, reflecting the summary trade-off between sensitivity and specificity across the included studies. The areas under the curves (AUC) and $Q^*$ value were also calculated. $Q^*$ value is defined by the point where sensitivity and specificity are equal, which is the point closest to the ideal top-left corner of the SROC space, reflecting the diagnostic value[13].

To analyze publication bias, we created inverted funnel plots of individual study log DORs plotted against the effective sample size. Both Begg's rank correlation test and Egger's regression test were utilized to assess the asymmetry of the funnel plot[14]. An asymmetrical funnel plot would suggest that additional small-sample studies may have been conducted but not published due to their unfavorable results.

## RESULTS

### Study evaluation and selection

After independent review, our research initially yielded a total of 50 articles in English-language from MEDLINE. Twenty-three articles were considered candidates after careful review of abstracts. After fulltext reading of these articles, eight were excluded because of employing reference standard other than pathological outcome ($n$ = 2); lack of raw data that was essential to built the 2 × 2 contingency tables ($n$ = 5); or focusing on metastasized STS only ($n$ = 1). Finally, 15 articles fulfilled all eligibility criteria and were included for data extraction and following analysis[3,6,7,15-26]. Flow diagram of publication selection is showed in Figure 1.

### Study characteristics and quality appraisal

Among the 15 studies, four were retrospectively designed[6,17,21,25], one had unclear design of whether prospective or retrospective[24], and the rest ten were prospective. All were in English. A total of 420 patients were included. Median sample size was 21 (range, 9 to 50), and mean age ranged from 10 to 54 years old. Three studies focused on pediatric patients only[3,6,24].

Detailed characteristics of the 15 included trials are outlined in Table 1. All articles assessed treatment of neoadjuvant chemotherapy, and in four studies some patients received neoadjuvant radiotherapy as well[19,20,25,25]. All trials had a sequential F-18 FDG PET both before and after the neoadjuvant therapies were performed.

All studies included in the meta-analysis fulfilled 9 or more of the 14 standards in the guidelines of the QUADAS tool. Drawbacks were commonly seen in description of how patients were selected and whether histopathological interpretation was blind to PET results. Moreover, the availability of clinical data in the diagnostic test reading was mostly unclear. The median methodological quality score (represented as the number of "Yes" to fourteen questions) was 12.

### Data analysis

Estimation of overall accuracy: Table 2 exhibited the absolute numbers from 2 × 2 contingency tables from each article. When more than one standard of PET was present in one trial, we took the one that generated higher accuracy into statistic calculation. The sensitivity of F-18 FDG PET to assess therapy response in sarcoma ranged from 40%-100%, and specificity from 44%-100%. The forest plot displayed the sensitivity, specificity and DOR with 95%CI for each individual study (Figure 2).

Based on the random-effects model, the pooled sensitivity and specificity were 87% (95%CI: 81%-91%) and 83% (95%CI: 77%-87%), respectively, and the pooled DOR was 40.5 (95%CI: -14.6-112.0). The SROC indicated an AUC of 0.930 and $Q^*$ value of 0.865 (Figure 3). Heterogeneity of sensitivity and specificity was detected, but no significant heterogeneity of DOR was found, as their $I^2$ exhibited in Figure 2.

### Subgroup analysis

Several subgroup analyses were then performed to examine the potential sources of heterogeneity displayed among the 15 studies. It was noted that these studies

**Table 1 Characteristics of the included studies**

| | Year | Tumor pathology | Patient No. (% of males) | Age, mean (range), yr | Therapy, regimen | No. (%) of responders | Pathologic criteria | Adopted PET criteria | Timing of PET | Quality score[b] |
|---|---|---|---|---|---|---|---|---|---|---|
| Im | 2012 | OS (resectable high-grade) | 14 (71) | 15 (10-25) | Neo-chemo: various combinations of D, C, M, I and E | 5 (36) | Necrosis fraction ⩾ 90% (by Salzer-Kuntschik *et al*) | Post-chemo SUVmax < 3 | Before, between the first and second cycles | 12 |
| Kim | 2011 | OS or ESFT | 23 (69) | 10 (3-19) | Neo-chemo: various combinations, I contained | 15 (65) | Viable tumor ⩽ 10% | SUV reduction rate ⩾ 50% and SUV2 ⩽ 2.5 | Before, after completion of chemo | 10 |
| Bajpai | 2011 | OS | 31 (81) | 17 (5-66) | Neo-chemo: C + D | 10 (32) | Necrosis fraction ⩾ 90% | V1 ⩽ 300 mL and SUV2:SUV1 ⩽ 0.48 | Before, after 3 cycle | 13 |
| Dimitrako | 2010 | STS (non-metastasized) | 24 (unclear) | Unclear | Neo-chemo: E + D | 14 (58) | Necrosis fraction ⩾ 90% | SUVmean+ Influx, value unclear | Before, after 2 cycles | 9 |
| Cheon | 2009 | OS (high-grade) | 70 (68) | 14[a] (5-59) | Neo-chemo: M, A and C | 33 (47) | Necrosis fraction ⩾ 90% | SUV2 < 2 or MVCR > 65% when SUV2 is between 2 to 5 | Before, after completion of chemo | 12 |
| Hamada | 2009 | OS | 9 | 51[a] (20-80) | Neo-chemo: D, C, I and M | 5 (56) | Necrosis fraction ⩾ 90% | SUV2 < 2.5 | Before, after completion of chemo | 10 |
| Benz | 2009 | STS (resectable high-grade) | 50 (52) | 51[a] (20-80) | Neo-chemo: mostly I contained | 8 (16) | Necrosis fraction ⩾ 90% | SUV reduction rate ⩾ 35% after the first cycle | Before, after the first cycle, after completion of chemo | 12 |
| Benz | 2008 | STS (high-grade) | 20 (50) | 49 (19-86) | Neo-chemo: I + D or G + D; some had radiotherapy | 6 (30) | Necrosis fraction ⩾ 95% | SUVmean changes, value unclear | Before, after completion of chemo | 12 |
| Evilevitch | 2008 | STS (high-grade) | 42 (57) | 54 (20-86) | Neo-chemo: I in 84%, G in 16% of patients, some had radiotherapy | 8 (19) | Necrosis fraction ⩾ 95% | SUV reduction rate ⩾ 60% | Before, after completion of chemo | 13 |
| Iagaru | 2008 | OS and STS | 14 (57) | 36 (18-56) | Neo-Chemo: I based | 6 (43) | Necrosis fraction ⩾ 90% | SUV reduction rate ⩾ 27% | Before, after completion of chemo | 11 |
| Ye | 2008 | OS | 15 (60) | 17[a] (7-33) | Neo-chemo: M, C, and A | 8 (53) | Necrosis fraction ⩾ 90% | TBR2/TBR1 < 0.46 | Before, after completion of chemo | 13 |
| Hawkins | 2005 | ESFT | 34 (unclear) | 19 (6-46) | Neo-chemo: V, D, Cy, I and E; some had radiotherapy | 25 (74 ) | Necrosis fraction ⩾ 90% | SUV2 < 2.5 | Before, after completion of chemo | 10 |
| Hawkins | 2002 | OS and ES | 31 (74) | 13 (6-19) | Neo-chemo: D, C, M, with or without I | 18 (58) | Necrosis fraction ⩾ 90% | SUV2 < 2 | Before, after completion of chemo | 10 |
| Franzius | 2000 | OS and ES | 17 (76) | 13[a] (5-36) | Neo-chemo: D, M, C and I; some had radiotherapy | 15 (88) | Necrosis fraction ⩾ 90% | TBR reduction rate > 30% | Before, after completion of chemo | 12 |
| Schulte | 1999 | OS | 27 (63) | 17[a] (5-36) | Neo-chemo: D, M, I and C | 17 (63) | Necrosis fraction ⩾ 90% | TBR2/TBR1< 0.6 | Before, after completion of chemo | 12 |

[a]: Median age; [b]: Evaluated by QUADAS tool, presented as the number of "Yes" to the fourteen questions of QUADAS criteria. A: Adriamycin; C: Cisplatin, Cy: Cyclophosphamide; D: Doxorubicin, E: Etoposide, G: Gemcitabine; M: Methotrexate, I: Ifosfamide; V: Vincristine; OS: Osteosarcoma; STS: Soft-tissue sarcomas; ES: Ewing sarcoma; ESFT: Ewing sarcoma family of tumors; Neo-chemo: Neoadjuvant chemotherapy; SUV: Standardized uptake value; SUV1: Pre-chemo SUV; SUV2: Post-chemo SUV; MVCR: Metabolic volume change ratio; TBR: Tumor to background ratio; TBR1: Pre-chemo TBR; TBR 2: post-chemo TBR.

investigated different tumor pathology, patient age group and chemotherapy regimen. Subgroup analysis results are showed in Table 3.

**Tumor pathology:** The $I^2$ values of sensitivity and specificity were close to 50% within both the group "OS or ES" and "STS", while within the group "STS" the

**Table 2 Absolute numbers from 2 × 2 contingency tables for the included studies**

| | TP | FN | FP | TN |
|---|---|---|---|---|
| Im | 5 | 0 | 1 | 8 |
| Kim | 14 | 1 | 1 | 7 |
| Bajpai[1] | 4 | 6 | 0 | 20 |
| Dimitrako | 11 | 3 | 1 | 9 |
| Cheon | 31 | 2 | 1 | 36 |
| Hamada | 5 | 0 | 0 | 4 |
| Benz-2009 | 8 | 0 | 14 | 28 |
| Benz-2008 | 6 | 0 | 0 | 14 |
| Evilevitch | 8 | 0 | 10 | 24 |
| Iagaru | 3 | 3 | 3 | 5 |
| Ye | 8 | 0 | 0 | 7 |
| Hawkins-2005 | 19 | 6 | 5 | 4 |
| Hawkins-2002 | 14 | 4 | 1 | 12 |
| Franzius | 15 | 0 | 0 | 2 |
| Schulte | 17 | 0 | 2 | 8 |

[1]One patient was not taken into calculation due to the unavailabilty of the relevant PET data. TP: True positive; FP: False positive; FN: False negative; TN: True negative.

heterogeneity of DOR was eliminated ($I^2$ = 0%). Statistic suggested that PET had better specificity and DOR when assessing the "OS or ES" group than the "STS" group.

**Patient age group:** Three studies focused on pediatric patients, and four studies investigated adult patients only. The heterogeneity of sensitivity, specificity and DOR was all resolved within the pediatric patient group ($I^2$: 0%-37%), and the $I^2$ of DOR within the adult patient group was 50%. Results showed that PET had significantly better specificity and DOR when used among pediatric patients than adults.

**Chemotherapy regimen:** Ifosfamide was reported to cause inflammatory reactions and consequently increase the chance of mismatch between PET scans and pathology[21]. To assess the effect of ifosfamide, we conducted this subgroup analysis. We found that within the group using chemotherapy agents without ifosfamide, the heterogeneity of specificity and DOR was resolved ($I^2$: 0% and 15.7%), and significantly better specificity and DOR than the group using ifosfamide were gained.

**PET response criteria:** Ten studies employed a lower standardized uptake value (SUV) after chemotherapies (mostly between 2 and 3) and/or a higher SUV reduction rate (27%-60%, mostly around 50%) as PET criteria of good response. Three studies used tumor to background ratio (TBR) related parameters to define good response. However, subgroup analysis did not resolve the heterogeneity of any parameters (sensitivity, specificity and DOR), hindering a reliable further comparison.

### *Publication bias*

The inverted funnel plot suggested some asymmetry of the distribution of the points (Figure 4). Begg's test indicated that publication bias was unlikely with a *P* value of 0.06. However, Egger's test showed that such bias might exist with a *P* value of 0.003.

## DISCUSSION

The diagnosis and response assessment of bone and soft tissue sarcomas by imaging remain a challenge, largely due to its high intratumoral heterogeneity. A previous meta-analysis on the value of F-18 FDG PET in detecting sarcoma has reported a relatively high accuracy (pooled sensitivity of 91%, specificity of 85%)[1]. Clinical evidence regarding the value of F-18 FDG PET to monitor response of sarcoma to neoadjuvant therapy is increasing, but a solid and quantitative evaluation is lacking. To our knowledge, this meta-analysis quantitatively summarized increased but different data from individual studies, and explored specific clinical uses of PET, such as the optimal patient groups and potential causes of false results, which could provide vital clues for the clinical applications of F-18 FDG PET in sarcoma. The statistical results supported the important role of F-18 FDG PET in assessing histopathological response to NCT in bone and soft tissue sarcomas, with relatively high pooled sensitivity (87%), specificity (83%), DOR (40.5), and AUC (0.930).

### *Optimal patient groups for F-18 FDG PET*

The difference of pathology and age predilection between bone and soft tissue sarcomas rationalize identification of optimal patient groups of using F-18 FDG PET to assess response to NCT. It is noted that all three studies that focused on pediatric patients investigated OS or ES. Further discrimination of OS and ES was hardly possible, since many studies did not report such data separately. Almost all studies investigating STS were comprised of merely adult patients (one study did not clearly report the patient age range[7]). One study included both bone and soft tissue sarcomas[21]. We conducted subgroup analysis between studies that included merely pediatric patients or adult patient to identify the effect of patient age, and so as to the effect of two pathological groups. Statistical results suggested that PET was especially promising among pediatric patients with OS and ES, with relatively high sensitivity and specificity, both around 90%.

### *Possible cause of misjudgment*

The pooled sensitivity of PET in response assessment of sarcoma was relatively high and stable among all subgroups in this meta-analysis, suggesting a relatively low chance of false negatives in clinical practice. By contrast, the pooled specificity was lower, and almost all the significant difference gained in subgroup analysis was attributed to specificity, calling for clinical caution of false positives. Certain chemotherapy agents have been reported to potentially cause misjudgments in PET scans after treatment of sarcoma, and ifosfamide was among the most widely known ones[21]. It is reported in bone tumors to theoretically increase SUVmax of adjacent inflammatory tissues,

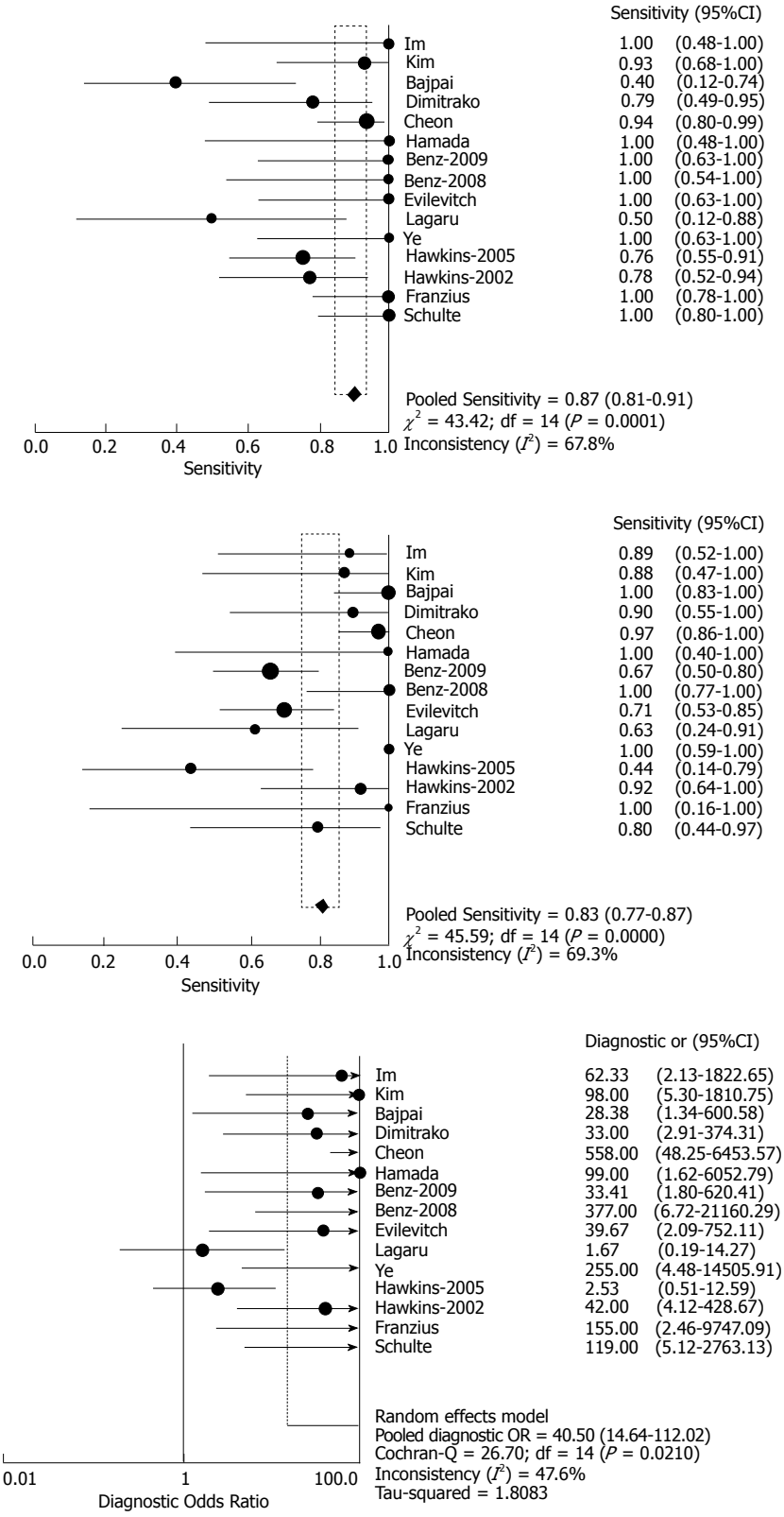



**Figure 2 Pooled sensitivity, specificity, and diagnostic odds ratio in forest plots for neoadjuvant therapy response prediction in bone and soft tissue sarcomas by F-18-fluorodeoxy-D-glucose-positron emission tomography.**

causing false negatives consequently. Differing from such possibility, our subgroup analysis found more chance of false positives. The most distinct and largest numbers of false positives were generated in studies by Benz *et al*[18] and Evilevitch *et al*[20], both investigating STS, suggesting that ifosfamide might have different mechanisms in STS from bone tumors. Since it was reported in clinical practice that addition of ifosfamide could improve patients'

**Table 3 Subgroup analysis based on tumor pathology, patient age groups and chemotherapy regimen**

| Subgroup criteria | Characteristics | No. of studies | No. of patients | Se % | *P* of Se | Sp % | *P* of Sp | DOR |
|---|---|---|---|---|---|---|---|---|
| Tumor pathology[a] | OS or ES (1-3, 5, 6, 11-15) | 10 | 270 | 87 (95%CI: 81-92) | > 0.05 | 91 (95%CI: 84-95) | < 0.05 | 57.8 |
| | STS (4, 7, 8, 9) | 4 | 55 | 92 (95%CI: 78-98) | | 75 (95%CI: 65-83) | | 47.9 |
| Patient age groups[b] | Pediatric patients (1, 2, 13) | 3 | 68 | 87 (95%CI: 72-96) | > 0.05 | 90 (95%CI: 73-98) | < 0.01 | 59.2 |
| | Adult patients (6, 7, 8, 9, 10) | 5 | 135 | 91 (95%CI: 76-98) | | 74 (95%CI: 64-82) | | 25.6 |
| Chemotherapy regimen[c] | With ifosfamide (2, 6, 7, 9, 10, 12, 14, 15) | 8 | 216 | 90 (95%CI: 82-95) | > 0.05 | 70 (95%CI: 61-78) | < 0.01 | 20.8 |
| | Without ifosfamide (3, 4, 5, 11) | 4 | 139 | 83 (95%CI: 72-91) | | 97 (95%CI: 91-100) | | 103.8 |

[a][22] not included due to covering both tumor pathology; [b]7 studies not included due to age ranges covering both pediatric and adult patients; [c]3 studies not included [3,20,25] for unclear portion of ifosfamide use. Se: Sensitivity; Sp: Specificity; OS: Osteosarcoma; ES: Ewing sarcoma; DOR: Diagnostic odds ratio.

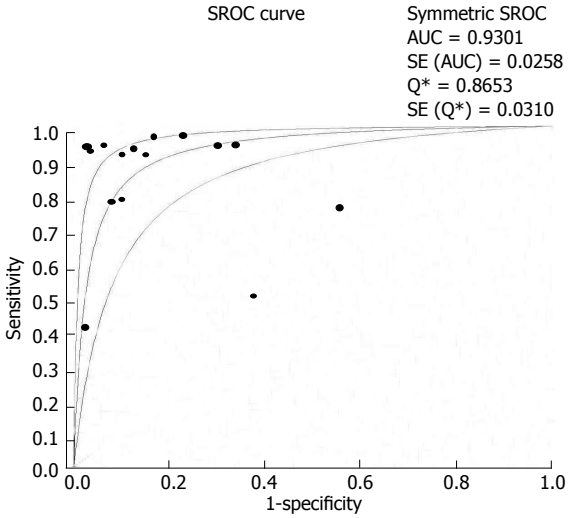



**Figure 3 Summary receiver operating characteristic curves for neoadjuvant therapy response prediction in bone and soft tissue sarcomas by F-18-fluorodeoxy-D-glucose-positron emission tomography.**

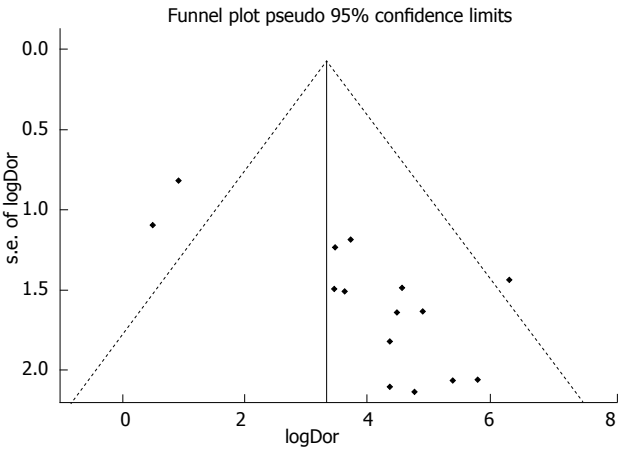



**Figure 4 Funnel plot to test publication bias shows some asymmetry.**

survival in both OS and ES, and ifosfamide was widely used in STS as well, it is noteworthy that this chemotherapeutic agent has the potential to cause misjudgment on PET scan results of sarcoma.

### *Candidate parameters for quantitative evaluation*

Different parameters have been employed to explore a better accuracy of F-18 FDG PET, among which maximum SUVs were considered quantitative and reproducible, therefore widely used in clinical trials[17]. Six studies included in this meta-analysis relied on post-treatment SUV (SUV2, mostly maximum SUV) and eight trials relied on SUV reduction rate to identify responders, and overlaps existed. It was pointed out that SUV2 may remain elevated because of inflammation or reactive changes. Moreover, tumors with an initial high pre-treatment SUV might still have a high SUV2, despite a significant decrease in the SUV value. SUV reduction rate is therefore proposed as a more reliable indicator[6]. In contrast to such knowledge, three included studies that compared the accuracy of SUV2 and SUV reduction rate all presented that SUV2 had a better correlation with histopathology than the latter[17,23,24]. TBR was employed in three studies, but was reported as nonreproducible and less practical due to inter-observer variability[3]. Since the subgroup analysis of different criteria still suffered unexplained heterogeneity, the optimal criteria remain to be defined in larger-sample, prospectively designed studies that include and compare directly different parameters.

### *Clinical roles of pre-operative F-18 FDG PET*

The insufficient accuracy of current imaging modalities raised reasonable concerns of clinicians regarding the roles of pre-operative PET. Potential clinical applications may include: (1) providing prognostic indicator of disease-free survival after multimodal treatment[5]; (2) guiding the selection of resection technique, because surgical oncologists may be more willing to perform limb-sparing surgical resections in the setting of a favorable response to therapy[27]; and (3) potentially identifying non-responders early (pooled NPV = 90%), since several studies explored the use of interim PET information and all generated positive results[3,7,15,18]. Benz *et al*[18] compared the performance of PET after the first cycle and after the

completion of NCT, reporting that the earlier scans had an even better accuracy. Because the currently available treatments are associated with significant toxicities[28], early identification of non-responders would allow clinicians to minimize the toxicity of ineffective chemotherapies and switch the treatment strategy to more aggressive regimen or addition of radiotherapy. However, further large-scale prospective studies are required to adequately position PET in sarcoma treatment management.

### Limitations of the study

Our meta-analysis had several drawbacks. First, the included studies had relatively small sample sizes in general, giving that sarcoma lacks abundant cases in clinical scenarios. Second, some extent of heterogeneity among studies was detected, while subgroup analysis could not eliminate all of the heterogeneity (although some was resolved). Third, some biases might be present such as publication bias, since *P* value yielded by Egger's test was 0.003. Spectrum biases were possible due to the failure of finding out comprehensive relevant articles or unpublished data, caused by the restriction of searching conditions.

In summary, this meta-analysis reveals a beneficial value of F-18 FDG PET to assess histopathological response after NCT in bone and soft tissue sarcomas. It is especially promising to apply to pediatric patients with OS or ES. Post-NCT SUV and SUV reduction rate on PET scans are potential candidates for quantitative criteria of response. It is noteworthy that certain chemotherapeutic agents, specifically ifosfamide, could contribute to false positives of PET results. Given the potential of PET to guide clinical treatment strategies and establish personalized medical plans, large-scale prospective studies are required to adequately position this modality in sarcoma management.

## COMMENTS

### Background

Sarcoma is a rare but highly malignant tumor that frequently requires neoadjuvant chemotherapy, but only a subcollective of patients benefit from such treatment. Therefore, there is an urgent need for non-invasive methods to accurately identify response of sarcoma to chemotherapy, and accordingly decide whether to switch to a more intensified schedule.

### Research frontiers

Radiologic imaging has some limitations for this utility due to the biological characteristics of sarcoma and the disadvantages of conventional anatomic imaging. Positron emission tomography (PET) using [F-18]-fluorodeoxy-D-glucose (F-18 FDG) is a newly applied imaging modality that measures tissue metabolism and has the unique ability to detect response at the molecular level, therefore is being increasingly used in chemotherapy response monitoring among patients with sarcoma. However, the actual accuracy of PET and some specific issues in clinical practice such as optimal patient groups, standards for response and possible causes for misjudgment remain to be explored.

### Innovations and breakthroughs

There have been multiple clinical trials assessing the accuracy of PET to monitor response to chemotherapy in sarcoma. Nonetheless, the sample sizes in conducted trials were relatively small and the results among them were remarkably discrepant. This meta-analysis provided a systematic and quantitative evaluation of the accuracy of PET for this utility. Moreover, it generated further information that clinical physicians concerned with, such as recommended subgroup of patients, potential candidates for quantitative criteria of response and possible causes of false positives of PET.

### Applications

The study results suggested that F-18 FDG PET is promising to predict neoadjuvant therapy response in sarcoma, especially in pediatric patients with osteosarcoma or Ewing sarcoma. A lower standardized uptake value (SUV) after chemotherapies and/or a higher SUV reduction rate are potential candidates for quantitative criteria of response. Certain chemotherapeutic agents such as ifosfamide could potentially cause false positives of PET.

### Terminology

Positron emission tomography is a functional imaging technique that produces a three-dimensional image of the body. When fludeoxyglucose (FDG), an analogue of glucose, was chosen as the tracer for PET, its concentrations will indicate tissue metabolic activity and therefore explore the possibility of viable tumor; Standardized uptake value is a frequently used quantitative parameter generated on PET scans, indicating concentrations of the tracer (*e.g.*, FDG) in targeted tissues and therefore reflecting their metabolic activity.

### Peer review

Interesting meta-analysis on the use of PET CT in the evaluation of patients with sarcomas of both soft tissue and bone after chemoradiation with a curative intent. The authors have used all the appropriate scientific methodologies used for the preparation of meta-analyses and their results are extremely interesting and supported by the literature. The authors conclude that PET scan is a promising diagnostic tool in the evaluation of neo-adjuvant treatment response in sarcomas.

## REFERENCES


1 **Bastiaannet E**, Groen H, Jager PL, Cobben DC, van der Graaf WT, Vaalburg W, Hoekstra HJ. The value of FDG-PET in the detection, grading and response to therapy of soft tissue and bone sarcomas; a systematic review and meta-analysis. *Cancer Treat Rev* 2004; **30**: 83-101 [PMID: 14766127 DOI: 10.1016/j.ctrv.2003.07.004]

2 **Ioannidis JP**, Lau J. 18F-FDG PET for the diagnosis and grading of soft-tissue sarcoma: a meta-analysis. *J Nucl Med* 2003; **44**: 717-724 [PMID: 12732672]

3 **Im HJ**, Kim TS, Park SY, Min HS, Kim JH, Kang HG, Park SE, Kwon MM, Yoon JH, Park HJ, Kim SK, Park BK. Prediction of tumour necrosis fractions using metabolic and volumetric 18F-FDG PET/CT indices, after one course and at the completion of neoadjuvant chemotherapy, in children and young adults with osteosarcoma. *Eur J Nucl Med Mol Imaging* 2012; **39**: 39-49 [PMID: 21953008 DOI: 10.1007/s00259-011-1936-4]

4 **Meyers PA**, Heller G, Healey J, Huvos A, Lane J, Marcove R, Applewhite A, Vlamis V, Rosen G. Chemotherapy for nonmetastatic osteogenic sarcoma: the Memorial Sloan-Kettering experience. *J Clin Oncol* 1992; **10**: 5-15 [PMID: 1370176]

5 **Bielack SS**, Kempf-Bielack B, Delling G, Exner GU, Flege S, Helmke K, Kotz R, Salzer-Kuntschik M, Werner M, Winkelmann W, Zoubek A, Jürgens H, Winkler K. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. *J Clin Oncol* 2002; **20**: 776-790 [PMID: 11821461 DOI: 10.1200/JCO.20.3.776]

6 **Kim DH**, Kim SY, Lee HJ, Song BS, Kim DH, Cho JB, Lim JS, Lee JA. Assessment of Chemotherapy Response Using FDG-PET in Pediatric Bone Tumors: A Single Institution Experience. *Cancer Res Treat* 2011; **43**: 170-175 [PMID: 22022294 DOI: 10.4143/crt.2011.43.3.170]

7 **Dimitrakopoulou-Strauss A**, Strauss LG, Egerer G, Vasamiliette J, Mechtersheimer G, Schmitt T, Lehner B, Haberkorn U, Stroebel P, Kasper B. Impact of dynamic 18F-FDG PET on the early prediction of therapy outcome in patients with high-risk soft-tissue sarcomas after neoadjuvant chemotherapy: a feasibility study. *J Nucl Med* 2010; **51**: 551-558 [PMID: 20351350 DOI: 10.2967/jnumed.109.070862]

8 **Barker FG**, Carter BS. Synthesizing medical evidence: systematic reviews and metaanalyses. *Neurosurg Focus* 2005; **19**:

E5 [PMID: 16241107]
9 **Whiting P**, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; **3**: 25 [PMID: 14606960 DOI: 10.1186/1471-2288-3-25]
10 **Sweeting MJ**, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004; **23**: 1351-1375 [PMID: 15116347 DOI: 10.1002/sim.1761]
11 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557]
12 **Reitsma JB**, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58**: 982-990 [PMID: 16168343 DOI: 10.1016/j.jclinepi.2005.02.022]
13 **Zamora J**, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006; **6**: 31 [PMID: 16836745 DOI: 10.1186/1471-2288-6-31]
14 **Song F**, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002; **31**: 88-95 [PMID: 11914301 DOI: 10.1093/ije/31.1.88]
15 **Bajpai J**, Kumar R, Sreenivas V, Sharma MC, Khan SA, Rastogi S, Malhotra A, Gamnagatti S, Kumar R, Safaya R, Bakhshi S. Prediction of chemotherapy response by PET-CT in osteosarcoma: correlation with histologic necrosis. *J Pediatr Hematol Oncol* 2011; **33**: e271-e278 [PMID: 22193290 DOI: 10.1097/MPH.0b013e31820ff78e]
16 **Cheon GJ**, Kim MS, Lee JA, Lee SY, Cho WH, Song WS, Koh JS, Yoo JY, Oh DH, Shin DS, Jeon DG. Prediction model of chemotherapy response in osteosarcoma by 18F-FDG PET and MRI. *J Nucl Med* 2009; **50**: 1435-1440 [PMID: 19690035 DOI: 10.2967/jnumed.109.063602]
17 **Hamada K**, Tomita Y, Inoue A, Fujimoto T, Hashimoto N, Myoui A, Yoshikawa H, Hatazawa J. Evaluation of chemotherapy response in osteosarcoma with FDG-PET. *Ann Nucl Med* 2009; **23**: 89-95 [PMID: 19205843 DOI: 10.1007/s12149-008-0213-5]
18 **Benz MR**, Czernin J, Allen-Auerbach MS, Tap WD, Dry SM, Elashoff D, Chow K, Evilevitch V, Eckardt JJ, Phelps ME, Weber WA, Eilber FC. FDG-PET/CT imaging predicts histopathologic treatment responses after the initial cycle of neoadjuvant chemotherapy in high-grade soft-tissue sarcomas. *Clin Cancer Res* 2009; **15**: 2856-2863 [PMID: 19351756 DOI: 10.1158/1078-0432.CCR-08-2537]
19 **Benz MR**, Allen-Auerbach MS, Eilber FC, Chen HJ, Dry S, Phelps ME, Czernin J, Weber WA. Combined assessment of metabolic and volumetric changes for assessment of tumor response in patients with soft-tissue sarcomas. *J Nucl Med* 2008; **49**: 1579-1584 [PMID: 18794268 DOI: 10.2967/jnumed.108.053694]
20 **Evilevitch V**, Weber WA, Tap WD, Allen-Auerbach M, Chow K, Nelson SD, Eilber FR, Eckardt JJ, Elashoff RM, Phelps ME, Czernin J, Eilber FC. Reduction of glucose metabolic activity is more accurate than change in size at predicting histopathologic response to neoadjuvant therapy in high-grade soft-tissue sarcomas. *Clin Cancer Res* 2008; **14**: 715-720 [PMID: 18245531 DOI: 10.1158/1078-0432.CCR-07-1762]
21 **Iagaru A**, Masamed R, Chawla SP, Menendez LR, Fedenko A, Conti PS. F-18 FDG PET and PET/CT evaluation of response to chemotherapy in bone and soft tissue sarcomas. *Clin Nucl Med* 2008; **33**: 8-13 [PMID: 18097248 DOI: 10.1097/RLU.0b013e31815c4fd4]
22 **Ye Z**, Zhu J, Tian M, Zhang H, Zhan H, Zhao C, Yang D, Li W, Lin N. Response of osteogenic sarcoma to neoadjuvant therapy: evaluated by 18F-FDG-PET. *Ann Nucl Med* 2008; **22**: 475-480 [PMID: 18670853 DOI: 10.1007/s12149-008-0147-y]
23 **Hawkins DS**, Schuetze SM, Butrynski JE, Rajendran JG, Vernon CB, Conrad EU, Eary JF. [18F]Fluorodeoxyglucose positron emission tomography predicts outcome for Ewing sarcoma family of tumors. *J Clin Oncol* 2005; **23**: 8828-8834 [PMID: 16314643 DOI: 10.1200/JCO.2005.01.7079]
24 **Hawkins DS**, Rajendran JG, Conrad EU, Bruckner JD, Eary JF. Evaluation of chemotherapy response in pediatric bone sarcomas by [F-18]-fluorodeoxy-D-glucose positron emission tomography. *Cancer* 2002; **94**: 3277-3284 [PMID: 12115361 DOI: 10.1002/cncr.10599]
25 **Franzius C**, Sciuk J, Brinkschmidt C, Jürgens H, Schober O. Evaluation of chemotherapy response in primary bone tumors with F-18 FDG positron emission tomography compared with histologically assessed tumor necrosis. *Clin Nucl Med* 2000; **25**: 874-881 [PMID: 11079583 DOI: 10.1097/00003072-200011000-00004]
26 **Schulte M**, Brecht-Krauss D, Werner M, Hartwig E, Sarkar MR, Keppler P, Kotzerke J, Guhlmann A, Delling G, Reske SN. Evaluation of neoadjuvant therapy response of osteogenic sarcoma using FDG PET. *J Nucl Med* 1999; **40**: 1637-1643 [PMID: 10520703]
27 **Picci P**, Sangiorgi L, Rougraff BT, Neff JR, Casadei R, Campanacci M. Relationship of chemotherapy-induced necrosis and surgical margins to local recurrence in osteosarcoma. *J Clin Oncol* 1994; **12**: 2699-2705 [PMID: 7989947]
28 **Lorigan P**, Verweij J, Papai Z, Rodenhuis S, Le Cesne A, Leahy MG, Radford JA, Van Glabbeke MM, Kirkpatrick A, Hogendoorn PC, Blay JY. Phase III trial of two investigational schedules of ifosfamide compared with standard-dose doxorubicin in advanced or metastatic soft tissue sarcoma: a European Organisation for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group Study. *J Clin Oncol* 2007; **25**: 3144-3150 [PMID: 17634494 DOI: 10.1200/JCO.2006.09.7717]

**P- Reviewer:** Rapidis AD, Reis JS **S- Editor:** Song XX
**L- Editor:** Wang TQ **E- Editor:** Wu HL



World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v2.i4.221

*World J Meta-Anal* 2014 November 26; 2(4): 221-227
ISSN 2308-3840 (online)
© 2014 Baishideng Publishing Group Inc. All rights reserved.

META-ANALYSIS

# Female-specific *vs* conventional knee prostheses after total knee arthroplasty: A meta-analysis

Gen-Xiang Rong, Lei Huang, Bin-Jie Gui, A-Man Xu, Jing-Ling Zhang, Si-Sheng Wang

Gen-Xiang Rong, Bin-Jie Gui, Jing-Ling Zhang, Si-Sheng Wang, Department of Orthopedics, the First Affiliated Hospital of Anhui Medical University, Hefei 230022, Anhui Province, China
Lei Huang, A-Man Xu, Department of Gastrointestinal Surgery, the First Affiliated Hospital of Anhui Medical University, Hefei 230022, Anhui Province, China
Author contributions: Rong GX conceived and designed the review, conducted the statistical analyses and drafted the manuscript; Huang L, Gui BJ and Xu AM critically revised the manuscript and approved the final version of the manuscript submitted for publication; Zhang JL and Wang SS identified and acquired reports of trials and analyzed the data.
Correspondence to: Lei Huang, MD, Department of Gastrointestinal Surgery, the First Affiliated Hospital of Anhui Medical University, 218 Jixi Road, Hefei 230022, Anhui Province, China. huangleizhenting@126.com
Telephone: +86-551-65334247 Fax: +86-551-63633742
Received: July 20, 2014 Revised: October 23, 2014
Accepted: October 28, 2014
Published online: November 26, 2014

## Abstract

**AIM:** To study whether female-specific prostheses are superior to conventional prostheses after total knee arthroplasty (TKA) by conducting this meta-analysis to evaluate the effects of the 2 different designs.

**METHODS:** A systematic electronic search was conducted in the databases of PubMed, EMBASE and the Cochrane Library for prospective and retrospective trials. Meta-analysis was performed for the outcomes, including range of motion (ROM), Knee Society score, Hospital for Special Surgery scores (HSS) and complications including deep infection, manipulation under anesthesia (MUA), revisions, anterior knee pain, deep vein thrombosis and overhang rate. Meta-analysis was conducted where applicable. Weighted mean difference (WMD) and odds ratio were calculated according to study type.

**RESULTS:** Seven studies including 1174 knees were eligible for data extraction and pooled analysis. The overhang rate of female-specific prostheses was significantly lower than the conventional ones (WMD, 3.25; 95%CI: 0.00-0.27; $P$ = 0.001). ROM in the female-specific prostheses group after TKA tended to be greater than the conventional prostheses group; however, with insignificant difference (WMD, 2.48; 95%CI: -0.83-5.78; $P$ = 0.14). HSS (WMD, 0.48; 95%CI: -1.45-0.88; $P$ = 0.63) and complications including deep infection (WMD, 0.39; 95%CI: 0.19-3.08; $P$ =0 .70), MUA (WMD, 1.53; 95%CI: 0.02-1.61; $P$ = 0.13) and revisions (WMD, 0.55; 95%CI: 0.07-4.34; $P$ = 0.13) were all comparable between the two groups with at least 1 year follow-up. Other indexes were revealed to be similar between the two treatments with a descriptive analytical method.

**CONCLUSION:** Although the overhang rate is lower with female-specific prostheses, the current evidence does not support that female-specific prostheses outweigh conventional ones after TKA.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Total knee arthroplasty; Female-specific; Conventional; Prosthesis; Meta-analysis

**Core tip:** In this meta-analysis, we systematically reviewed the literature and evaluated the clinical efficacy of female-specific *vs* conventional prostheses after total knee arthroplasty (TKA). We found that although the overhang rate is lower with female-specific prostheses and range of motion in the female-specific prostheses group tended to be greater than the conventional prostheses group, the current evidence does not support that female-specific prostheses are superior to conventional ones after TKA with all other parameters, including Knee Society score, Hospital for Special Surgery scores and complications including deep infection, manipulation under anesthesia, revisions, anterior knee pain and deep vein thrombosis comparable between the two treatments.

Rong GX, Huang L, Gui BJ, Xu AM, Zhang JL, Wang SS. Female-specific *vs* conventional knee prostheses after total knee arthroplasty: A meta-analysis. *World J Meta-Anal* 2014; 2(4): 221-227 Available from: URL: http://www.wjgnet.com/2308-3840/full/v2/i4/221.htm DOI: http://dx.doi.org/10.13105/wjma.v2.i4.221

## INTRODUCTION

Nowadays, total knee arthroplasty (TKA) has become an promising treatment for end-stage degenerative knee arthritis and nearly 60% of sufferers are women[1]. Viewed from the anatomical aspect, there are significant distinctions in arthrosis between men and women. Q angles of women are larger than men and the anterior heights of lateral and medial condyles are smaller in women compared with men. The ratio of anterior-posterior and medial-lateral of the anterior femoral condylar anatomy is larger than in men[2,3]. Conventional prostheses do not consider the gender difference so it may influence the clinical efficacy of TKA. Some research has shown that the clinical efficacy of females is inferior to men and the researchers think that this may because of the anatomical differences based on gender, which can result in patellofemoral joint tightness or overstuffing that eventually limits range of motion (ROM)[4-6]. In order to solve this problem, according to the knee joint anatomical structure of women, the following adjustments were performed based on the conventional prostheses: firstly, female-specific prostheses have 3 Q-angles increased to the distal femoral to solve the defective rate of postoperative patella trajectory; secondly, compared with conventional prostheses, the medial-lateral femoral condylar is shortened. Therefore, the aspect ratio of the medial-lateral to anterior–posterior femoral condylar become smaller and so the consequences of implant overhang is reduced; thirdly, female prostheses reduce the height of the anterior condyle to reduce the friction effectively[7]. Theoretically, female prostheses can obtain a better clinical effect compared with conventional prostheses. However, whether the female prostheses are better clinically is not clear.

Therefore, this meta-analysis was conducted to compare the clinical efficacy of female-specific prostheses versus conventional ones after primary TKA.

## MATERIALS AND METHODS

### Literature retrieval

We searched EMBASE and MEDLINE databases and the Cochrane Library with the search terms: total knee arthroplasty/total knee replacement and gender-specific/sex-specific/female-specific/NexGen. We limited our search to trials from January 1966 to June 2014. References of the retrieved papers were also screened for any additional relevant studies.

### Inclusion criteria

Selection criteria included: (1) clinical trials which compared female-specific prostheses with conventional ones after primary TKA; (2) outcomes including the postsurgical clinical data with a minimum of one year follow-up; (3) presurgical diagnosis of end-stage arthritis; and (4) manuscripts published in English before June 2014.

Exclusion criteria were: (1) patients with a history of previous upper tibial osteotomy, septic arthritis, patellectomy or fixed varus/valgus deformity with an angle larger than 20°; (2) trials with insufficient data, *e.g.*, without ROM, knee society score (KSS), Hospital for Special Surgery scores (HSS) or complications; and (3) comments, letters, guidelines, case reports, reviews, duplications or those that overlapped with other trials but with less information.

### Quality evaluation

The methodological quality of trials was evaluated according to the Cochrane Handbook (version 5.1.2) for systematic reviews of interventions[8]. The assessment criteria have five aspects, including: (1) generation of adequate sequence; (2) concealment of allocation sequence; (3) blinding; (4) incomplete outcome data; and (5) selective outcome report.

### Data extraction

Data were collected and extracted by two authors (Rong GX and Gui BJ) independently, including first author, publication date, study type, sample size, patient's age, gender, body mass index, intervention, follow-up, pre and post-surgical ROMs, Knee score and complications. Any discrepancy was resolved through discussion and consensus. The 95%CI or standard error was converted into standard deviation (SD).

KSS includes Knee Society clinical and function scores. HSS has a high degree of accuracy in evaluating postoperative knee function. Overhang rate was defined as the incidence of the cases where the medial-lateral diameter of prosthesis is larger than the femoral condyle, which will lead to soft tissue irritation and balancing. When patients develop stiffness or less knee flexion, they will undergo manipulation under anesthesia (MUA) to gain knee flexion.

### Statistical analysis

The effect measurements estimated were weighted mean difference (WMD) for continuous outcome and odds ratio for dichotomous outcome, both of which were reported with 95%CIs. Differences were considered to be statistically significant with a $P$ value less than 0.05. Pooled estimate parameters were firstly evaluated with the fixed effect model, while when there was heterogeneity, the random effect model was applied. The heterogeneity among included trials was evaluated with the $I^2$ statistic, representing the proportion of the total variation across trials ($I^2 \leqslant 75\%$, extreme heterogeneity; $50\% \leqslant I^2 \leqslant 75\%$, large heterogeneity; $25\% < I^2 < 50\%$, moderate heterogeneity; $I^2 \leqslant 25\%$, no heterogeneity)[9]. With $I^2 \leqslant 25\%$, a fixed effect model was selected and with $25\% < I^2 < 50\%$, a random effect model was applied. However,

**Table 1 Baseline characteristics of included trials**

| Study ID | Implant used | | Sample size (knees) | | Study design | Mean age (yr) | | Follow-up (mo) | BMI (kg/m$^2$) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | GSP | CP | GSP | CP | | GSP | CP | | GSP | CP |
| Kim *et al*[14] | Gender-specific CR- Flex | Standard CR- Flex | 138 | 138 | RCT | 71.2 | 71.2 | 39 | 27.3 | 27.3 |
| Kim *et al*[15] | Gender-specific LPS-Flex | Standard LPS-Flex | 85 | 85 | RCT | 69.7 | 69.7 | 25.6 | 27.1 | 27.1 |
| Tanavalee *et al*[16] | Gender-specific LPS-Flex | Standard LPS-Flex | 165 | 149 | CCT | 70 | 70.5 | 24 | 23.4 | 23.5 |
| Singh *et al*[17] | Gender-specific LPS High-Flex | Standard LPS-Flex | 100 | 100 | RCT | 64 | 68 | 24 | 30.7 | 31.1 |
| Song *et al*[18] | Gender Solutions | Standard CR- Flex | 46 | 46 | RCT | 68.8 | 68.8 | 31.5 | 26.8 | 26.8 |
| Lionberger *et al*[19] | Gender-specific CR or LPS-Flex | Standard CR- Flex | 39 | 39 | CCT | 67.9 | 68.3 | 12 | 29.9 | 30 |
| Thomsen *et al*[20] | Gender Solutions high-Flex | Standard LPS-Flex | 22 | 22 | RCT | 66 | 66 | 13 | 29.3 | 29.3 |

GSP: Gender-specific prostheses; CP: Conventional prostheses; BMI: Body mass index; CR: Posterior cruciate retaining; LPS: Legacy posterior stabilized; RCT: Randomized controlled trial; CCT: Case control trial.

**Table 2 Designs and methodological quality summary**

| Study | Sequence generation | Allocation concealment | Assessor blinding | Incomplete outcome | Selective reporting |
|---|---|---|---|---|---|
| Kim *et al*[14] | Yes | Yes | Unclear | Yes | Yes |
| Kim *et al*[15] | Yes | Yes | Yes | Yes | Yes |
| Tanavalee *et al*[16] | No | No | No | Yes | Yes |
| Singh *et al*[17] | Yes | Unclear | Unclear | Yes | Yes |
| Song *et al*[18] | Yes | Unclear | Yes | Yes | Yes |
| Lionberger *et al*[19] | No | No | No | Yes | Yes |
| Thomsen *et al*[20] | Yes | Unclear | Yes | Yes | Yes |

Yes: Low risk of bias; Unclear: Moderate risk of bias; No: High risk of bias.

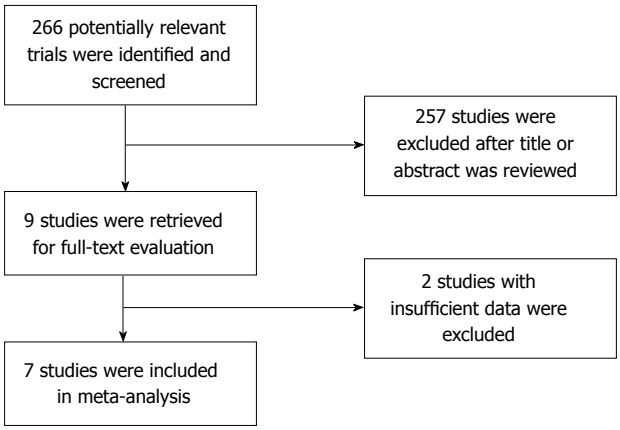



**Figure 1 Process of study selection.**

for a trial with $I^2 \geqslant 50\%$ indicating significant heterogeneity[10], we then tried to figure out the reasons and apply the method of sensitivity or subgroup analysis which was carried out via ruling out the studies possibly biasing the outcomes. The probable publication bias of the outcomes was evaluated using a funnel plot[11]. All statistical analyses were conducted with software RevMan 5.0.

## RESULTS

### Characteristics of included studies

266 potential papers were initially identified according to the search strategy. After the abstracts were read, 257 articles were excluded because they were not related to our research. Seven studies were read as full text and 2 studies[12,13] were excluded subsequently without sufficient data. Ultimately, 7 trials[14-20] with 1174 knees were included: 595 in the female-specific prosthesis arm and 579 in the conventional prosthesis arm. The selection process is shown in Figure 1.

Three trials compared posterior cruciate retaining prostheses and 4 studies compared legacy posterior stabilized prostheses. The follow-up period was 12 to 39 mo. The clinical and demographic features are summarized in Table 1.

Methodological quality of the eligible studies is shown in Table 2. In the 7 studies, 5 were randomized controlled studies and 2 were retrospective studies.

### HSS and KSS

Only 3 studies provided postoperative HSS and 1 does not offer SD (Table 3). Compared with preoperative HSS, both of the implants significantly enhanced postoperative HSS. However, no difference was found (WMD, 0.48; 95%CI: -1.45-0.88; $P = 0.63$). Four studies provided postoperative KSS but only one study provided SD; therefore, we also just applied a descriptive analysis (Table 4). The 4 trials all reported that the postoperative KSS clinical score was more than 90 points and that the postoperative KSS function score was more than 80 points in the female-specific and conventional groups. The KSS score was significantly improved but there was no difference between the two groups.

### Knee ROM

Statistical test with homogeneity based on all trials suggested significant heterogeneity ($I^2 = 50.6\%$, $P = 0.13$), so a random effects model was applied. We found that ROM in the female-specific prostheses group after TKA tended

**Table 3 Clinical results: preoperative and postoperative range of motion, Hospital for Special Surgery scores**

| Study | Preoperative ROM | | *P* | Postoperative ROM | | *P* | Preoperative HSS | | *P* | Postoperative HSS | | *P* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GSP | CP | | GSP | CP | | GSP | CP | | GSP | CP | |
| Kim *et al*[14] | 127 (100-150) | 123 (80-150) | NR | 124 (85-140) | 126 (85-140) | NR | NR | NR | NR | NR | NR | NR |
| Kim *et al*[15] | 123 (85-150) | 120 (58-150) | 0.139 | 126 (85-140) | 125 (80-140) | 0.739 | 59.7 (37-58) | 59.1 (37-68) | 0.120 | 91.2 (77-100) | 90.7 (84-100) | 0.252 |
| Tanavalee *et al*[16] | 126 | 125.6 | > 0.05 | 134.1 | 133.5 | > 0.05 | NR | NR | NR | NR | NR | NR |
| Singh *et al*[17] | 111.7 ± 13.5 | 110.5 ± 13.7 | 0.282 | 122.8 ± 8.0 | 119.9 ± 8.7 | 0.007 | 55.7 ± 9.1 | 56.7 ± 9.0 | 0.212 | 91.5 ± 4.8 | 91.9 ± 4.1 | 0.313 |
| Song *et al*[18] | 120.4 ± 16.6 | 121.8 ± 13.5 | 0.64 | 131.1 ± 9.2 | 133.7 ± 19.2 | 0.16 | 57.5 ± 11.5 | 57.1 ± 12.1 | 0.880 | 92.7 ± 8.0 | 92.1 ± 8.7 | 0.75 |
| Lionberger *et al*[19] | 102.7 ± 9.5 | 107.5 ± 11.5 | 0.048 | 123.4 ± 9.8 | 118.3 ± 9.7 | 0.023 | NR | NR | NR | NR | NR | NR |
| Thomsen *et al*[20] | NR | NR | NR | 125 (105-142) | 125 (105-139) | 0.82 | NR | NR | NR | NR | NR | NR |

ROM: Range of motion; HSS: Hospital for special surgery scores; GSP: Gender-specific prostheses; CP: Conventional prostheses; NR: Not reported.

**Table 4 Clinical results: preoperative and postoperative Knee Society score**

| Study | Preoperative KSS clinical score | | *P* | Postoperative KSS clinical score | | *P* | Preoperative KSS function score | | *P* | Postoperative KSS function score | | *P* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GSP | CP | | GSP | CP | | GSP | CP | | GSP | CP | |
| Kim *et al*[14] | 39.8 (10-70) | 35.3 (0-50) | 0.149 | 93 (70-100) | 94 (70-100) | 0.69 | 44.2 (20-70) | 44.2 (20-70) | 1.000 | 84 (60-100) | 83 (60-100) | 0.322 |
| Kim *et al*[15] | 34.4 (7-62) | 31.2 (0-55) | 0.145 | 96.5 (83-100) | 95.5 (81-100) | 0.424 | 48.3 (20-80) | 47.9 (20-80) | 0.320 | 84.8 (60-100) | 84.8 (60-100) | NR |
| Tanavalee *et al*[16] | 34.3 | 34.7 | > 0.05 | 92.1 | 92.9 | > 0.05 | 32.2 | 31.8 | > 0.05 | 89.7 | 89.5 | > 0.05 |
| Singh *et al*[17] | 34.2 ± 11.5 | 36.3 ± 13.4 | 0.118 | 94.9 ± 4.7 | 95.8 ± 3.6 | 0.061 | 37.6 ± 14.0 | 38.7 ± 10.2 | 0.254 | 80.2 ± 10.6 | 79.9 ± 13.0 | 0.429 |

Relevant data were not available for the other 3 studies. KSS: Knee Society score; GSP: Gender-specific prostheses; CP: Conventional prostheses.

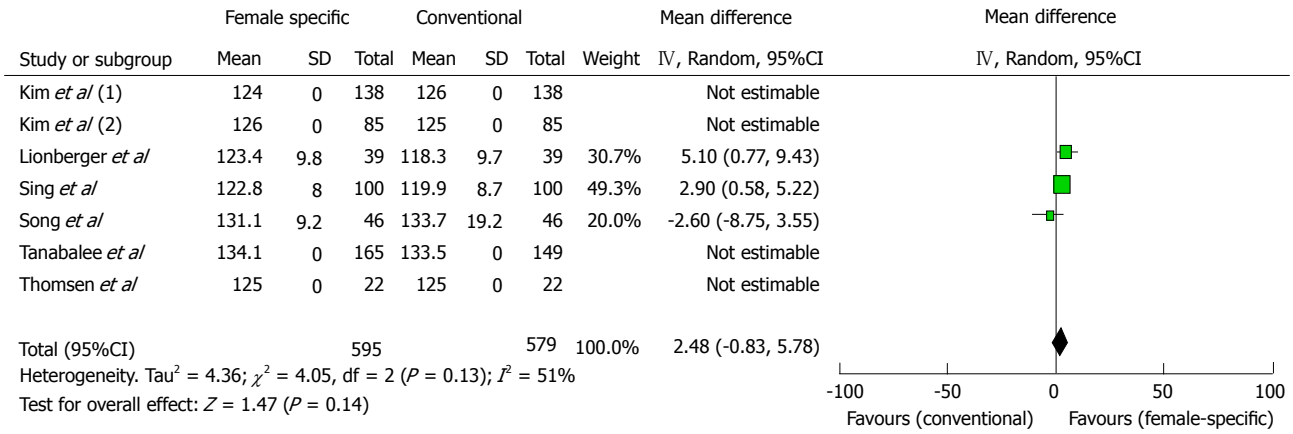



**Figure 2 Forest plot of range of motion with female-specific *vs* conventional prostheses after total knee arthroplasty.** ROM in the female-specific prostheses group after TKA tended to be greater than the conventional prostheses group; however, with insignificant difference between the 2 groups. ROM: Range of motion; TKA: Total knee arthroplasty; SD: Standard deviation; CI: Confidence interval.

to be greater than the conventional prostheses group; however, with insignificant difference between the 2 groups (WMD, 2.48; 95%CI: -0.83-5.78; $P = 0.14$; Figure 2).

### Complications

Seven trials recorded surgical complications. Information on deep infection was found in 4 trials[14-16,19] and the pooled WMD was 0.76 (95%CI: 0.19-3.08; $P = 0.76$), showing that there was no significant difference between the two groups. The $P$ value of heterogeneity test was 0.95 and $I^2$ was 0, indicating no heterogeneity (Figure 3).

Information on overhang was available in 3 trials[14,15,20]. Statistical heterogeneity was not found between the female-specific prosthesis group with conventional ones ($P = 0.88$), so we use a fixed effects model. The pooled WMD was 3.25 (95%CI: 0.00-0.27; $P = 0.001$), showing that the overhang rate of the female-specific prosthesis group was significantly lower than the conventional prosthesis group.

There was one patient with a superficial wound infection after gender-specific TKA, as reported by Thomsen *et al*[20]. MUA for patients was found in 2 trials[16,17]. Revisions performed were also reported in 2 trials[14,16]. Anterior knee pain was reported by Singh *et al*[17]. There was no supracondylar fracture or deep vein thrombosis reported. These are all shown in Table 5 and significant statistical differences were not found between the 2 groups (MUA: WMD, 1.53; 95%CI: 0.02-1.61; $P = 0.13$. Revisions: WMD, 0.55; 95%CI: 0.07-4.34; $P = 0.13$).

## DISCUSSION

The meta-analysis showed that there was no significant statistical difference in knee ROM, which is one of the most

**Table 5 Outcomes of complications comparing female-specific with conventional prostheses**

| Complications | Incidence | | Odds ratio (95%CI) | Overall effect (*P* value) | $I^2$ (%) | Heterogeneity *P* |
|---|---|---|---|---|---|---|
| | GSP | CP | | | | |
| Manipulation under anesthesia[16,17] | 0/265 | 4/249 | 0.19 (0.02-1.61) | 0.13 | 0 | 0.97 |
| Revisions[14,16] | 1/293 | 2/286 | 0.57 (0.07-4.34) | 0.55 | 0 | 0.64 |
| Anterior knee pain[17] | 5/50 | 4/50 | 1.28 (0.32-5.07) | 0.73 | NA | NA |

GSP: Gender-specific prostheses; CP: Conventional prostheses; CI: Confidence interval; NA: Not applicable.

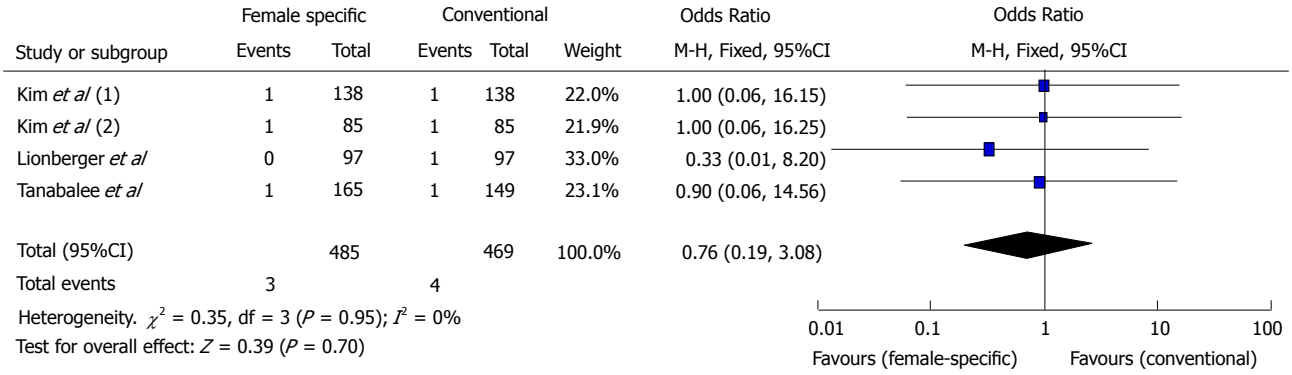



**Figure 3 Forest plot of deep infection with female-specific *vs* conventional prostheses after total knee arthroplasty.** There was no significant significance between the 2 groups. CI: Confidence interval.

significant measurements of knee arthroplasty but not the only measure of knee arthroplasty. Meneghini *et al*[21] reported that there was no difference between the patients who obtained flexion greater than 115° and those who obtained high flexion greater than 125° in Knee Society scores and function scores, which have been widely used in TKA since 1989. HSS was put forward by the American Special Surgery Hospital and has a high degree of accuracy in indicating postoperative knee function. We found no difference between the two groups on KSS and HSS.

Postoperative complications include deep infection which is one of the most serious postoperative complications that could lead to operational failure, MUA, revisions, anterior knee pain and deep vein thrombosis, which in our meta-analysis were not found to have significant statistical differences between the two groups. But, because the follow-up time is not long enough, the long-term complications are not shown. Female-specific prostheses were designed to reduce the ratio of mediolateral overhang in women. Clarke *et al*[12] found that the incidence of overhang was reduced in female patients when using a female-specific prosthesis. Our meta-analysis suggested that the overhang rate of female-specific prostheses was lower than conventional prostheses. So, the female specific prosthesis provides an easy way to manage the improper mediolateral matching between the femoral component and condyle, a phenomenon that frequently occurs in mid-sized to large-sized female patients[16].

However, some scholars question the design of female-specific prostheses. They report that when conventional prostheses are used, women achieve equal or even better results than men in aspects of prosthesis survival, pain, outcomes scores, risk of revision, ROM, stiffness, wear-related failures and satisfaction[22]. This finding does not support the theory that conventional prostheses place women at a disadvantage compared to men. Besides, Merchant *et al*[23] summarized 19 studies for full review and concluded that the Q angle and anterior heights of the medial and lateral condyles in women are not significantly different from men and that the smaller female medial-lateral/anterior-posterior aspect ratio has no measurable clinical effect. This finding refutes the theory that female-specific prostheses are needed to address anatomic differences between men and women. In addition, the clinical efficacy after TKA is not only concerned with prostheses but also with many other factors, such as body mass index, age, preoperative ROM, surgical technique, the removal of the posterior cruciate ligament and osteophyte, the management of postoperative pain and compliance of the patient[24].

Meta-analysis itself belongs to an observational study rather than controlled trial study so the result is likely to be affected by bias and confounding factors. Our meta-analysis has some limitations. Firstly, 2 retrospective studies were included and the inherent bias of nonrandomized studies may affect the reliability of conclusion. Secondly, not all the studies provided each clinical result in detail and some studies did not offer SD, which could not be included in the analysis using RevMan. Thirdly, the follow-up periods of the included studies were short so we could not learn about the long-term clinical efficacy. Fourthly, because small number of articles were incorporated in our study, we failed to assess the publication biases by a funnel plot analysis. So, the level of confidence in the estimates is relatively low. However, the inclusion

and exclusion criteria in the present meta-analysis are very strict and we evaluated the publication bias qualitatively so the conclusion is believable.

In conclusion, despite the limitations, the meta-analysis indicated that there were no differences between female-specific and conventional prostheses groups in knee ROM, KSS, HSS and complications, including deep infection, MUA, revisions, anterior knee pain and deep vein thrombosis with at least 1 year follow-up but the overhang rate of female-specific prosthesis group was lower than conventional prosthesis group. The absence of evidence does not necessarily mean evidence of absence and more studies of high quality are needed to demonstrate the results and long-term follow-up is needed to prove whether there are any differences in long-term clinical efficacy.

## COMMENTS

### Background

Nowadays, total knee arthroplasty (TKA) has become an effective therapy for end-stage degenerative arthritis of the knee and nearly 60% of sufferers are women. Female-specific prostheses consider the difference of knee joint anatomical structure between men and women. Theoretically, female-specific prostheses can obtain a better clinical effect compared to conventional prostheses. However, whether the female prostheses are superior to the conventional prostheses clinically is not clear.

### Research frontiers

In recent years, many studies have been performed to compare the clinical efficacy of female-specific *vs* conventional prostheses after TKA. Moreover, a systematic review was performed. However, these reviews were methodologically insufficient and thus could not achieve a comprehensive conclusion.

### Innovations and breakthroughs

Meta-analysis was performed for the outcomes, including range of motion, Knee Society Score (KSS), Hospital for Special Surgery scores (HSS) and complications, including manipulation under anesthesia (MUA), revisions, anterior knee pain, deep vein thrombosis and overhang rate. Meta-analysis was conducted where applicable. These findings were not presented clearly in previous systematic reviews.

### Applications

Although the overhang rate is lower with female-specific prostheses, the current evidence does not support that female-specific prostheses are superior to conventional ones after TKA. This can be used as a clinical reference when choosing prostheses for TKA.

### Terminology

KSS includes Knee Society clinical and function scores. HSS has a high degree of accuracy in evaluating postoperative knee function. Overhang rate was defined as the incidence of the cases where the medial-lateral diameter of prosthesis is larger than the femoral condyle, which will lead to soft tissue irritation and balancing. When patients develop stiffness or less knee flexion, they will be undergo MUA to gain knee flexion.

### Peer review

This is a well-performed meta-analysis of currently available studies on whether female-specific prostheses are better than conventional prostheses after TKA. The manuscript deals with a relevant topic; the style and structure are good.

## REFERENCES


1 **Ritter MA**, Wing JT, Berend ME, Davis KE, Meding JB. The clinical effect of gender on outcome of total knee arthroplasty. *J Arthroplasty* 2008; **23**: 331-336 [PMID: 18358368 DOI: 10.1016/j.arth.2007.10.031]

2 **Conley S**, Rosenberg A, Crowninshield R. The female knee: anatomic variations. *J Am Acad Orthop Surg* 2007; **15** Suppl 1: S31-S36 [PMID: 17766787]

3 **Cheng FB**, Ji XF, Lai Y, Feng JC, Zheng WX, Sun YF, Fu YW, Li YQ. Three dimensional morphometry of the knee to design the total knee arthroplasty for Chinese population. *Knee* 2009; **16**: 341-347 [PMID: 19230678 DOI: 10.1016/j.knee.2008.12.019]

4 **Mihalko W**, Fishkin Z, Krackow K. Patellofemoral overstuff and its relationship to flexion after total knee arthroplasty. *Clin Orthop Relat Res* 2006; **449**: 283-287 [PMID: 16888532 DOI: 10.1097/01.blo.0000218756.89439.06]

5 **SooHoo NF**, Lieberman JR, Ko CY, Zingmond DS. Factors predicting complication rates following total knee replacement. *J Bone Joint Surg Am* 2006; **88**: 480-485 [PMID: 16510811 DOI: 10.2106/JBJS.E.00629]

6 **Vincent KR**, Vincent HK, Lee LW, Alfano AP. Outcomes in total knee arthroplasty patients after inpatient rehabilitation: influence of age and gender. *Am J Phys Med Rehabil* 2006; **85**: 482-489 [PMID: 16715017 DOI: 10.1097/01.phm.0000219151.18749.50]

7 **Guy SP**, Farndon MA, Sidhom S, Al-Lami M, Bennett C, London NJ. Gender differences in distal femoral morphology and the role of gender specific implants in total knee replacement: a prospective clinical study. *Knee* 2012; **19**: 28-31 [PMID: 21277212 DOI: 10.1016/j.knee.2010.12.005]

8 **Xu AM**, Huang L, Liu W, Gao S, Han WX, Wei ZJ. Neoadjuvant chemotherapy followed by surgery versus surgery alone for gastric carcinoma: systematic review and meta-analysis of randomized controlled trials. *PLoS One* 2014; **9**: e86941 [PMID: 24497999 DOI: 10.1371/journal.pone.0086941]

9 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557]

10 **Higgins JP**, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539-1558 [PMID: 12111919 DOI: 10.1002/sim.1186]

11 **Schulz KF**, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273**: 408-412 [PMID: 7823387]

12 **Clarke HD**, Hentz JG. Restoration of femoral anatomy in TKA with unisex and gender-specific components. *Clin Orthop Relat Res* 2008; **466**: 2711-2716 [PMID: 18719972 DOI: 10.1007/s11999-008-0454-6]

13 **Song EK**, Park SJ, Yoon TR, Park KS, Seo HY, Seon JK. Hi-flexion and gender-specific designs fail to provide significant increases in range of motion during cruciate-retaining total knee arthroplasty. *J Arthroplasty* 2012; **27**: 1081-1084 [PMID: 22115763 DOI: 10.1016/j.arth.2011.10.011]

14 **Kim YH**, Choi Y, Kim JS. Comparison of standard and gender-specific posterior-cruciate-retaining high-flexion total knee replacements: a prospective, randomised study. *J Bone Joint Surg Br* 2010; **92**: 639-645 [PMID: 20435999 DOI: 10.1302/0301-620X.92B5.24129]

15 **Kim YH**, Choi Y, Kim JS. Comparison of a standard and a gender-specific posterior cruciate-substituting high-flexion knee prosthesis: a prospective, randomized, short-term outcome study. *J Bone Joint Surg Am* 2010; **92**: 1911-1920 [PMID: 20720134 DOI: 10.2106/JBJS.I.00910]

16 **Tanavalee A**, Rojpornpradit T, Khumrak S, Ngarmukos S. The early results of gender-specific total knee arthroplasty in Thai patients. *Knee* 2011; **18**: 483-487 [PMID: 21041089 DOI: 10.1016/j.knee.2010.10.005]

17 **Singh H**, Mittal V, Nadkarni B, Agarwal S, Gulati D. Gender-specific high-flexion knee prosthesis in Indian women: a prospective randomised study. *J Orthop Surg* (Hong Kong) 2012; **20**: 153-156 [PMID: 22933669]

18 **Song EK**, Jung WB, Yoon TR, Park KS, Seo HY, Seon JK. Comparison of outcomes after bilateral simultaneous total knee arthroplasty using gender-specific and unisex knees. *J Arthroplasty* 2012; **27**: 226-231 [PMID: 21704482 DOI: 10.1016/j.arth.2011.03.046]

19 **Lionberger DR**, Eggers MD, Brewer KE, Fang L. Improved

knee flexion following high-flexion total knee arthroplasty. *J Orthop Surg Res* 2012; **7**: 22 [PMID: 22672696 DOI: 10.1186/1749-799X-7-22]

20 **Thomsen MG**, Husted H, Bencke J, Curtis D, Holm G, Troelsen A. Do we need a gender-specific total knee replacement? A randomised controlled trial comparing a high-flex and a gender-specific posterior design. *J Bone Joint Surg Br* 2012; **94**: 787-792 [PMID: 22628593 DOI: 10.1302/0301-620X.94B6.28781]

21 **Meneghini RM**, Pierson JL, Bagsby D, Ziemba-Davis M, Berend ME, Ritter MA. Is there a functional benefit to obtaining high flexion after total knee arthroplasty? *J Arthroplasty* 2007; **22**: 43-46 [PMID: 17823014 DOI: 10.1016/j.arth.2007.03.011]

22 **Baker PN**, Khaw FM, Kirk LM, Esler CN, Gregg PJ. A randomised controlled trial of cemented versus cementless press-fit condylar total knee replacement: 15-year survival analysis. *J Bone Joint Surg Br* 2007; **89**: 1608-1614 [PMID: 18057361 DOI: 10.1302/0301-620X.89B12.19363]

23 **Merchant AC**, Arendt EA, Dye SF, Fredericson M, Grelsamer RP, Leadbetter WB, Post WR, Teitge RA. The female knee: anatomic variations and the female-specific total knee design. *Clin Orthop Relat Res* 2008; **466**: 3059-3065 [PMID: 18820981 DOI: 10.1007/s11999-008-0536-5]

24 **Rajgopal V**, Bourne RB, Chesworth BM, MacDonald SJ, McCalden RW, Rorabeck CH. The impact of morbid obesity on patient outcomes after total knee arthroplasty. *J Arthroplasty* 2008; **23**: 795-800 [PMID: 18534516 DOI: 10.1016/j.arth.2007.08.005]

**P- Reviewer:** Laudner K, Tunio MA, Trkulja V
**S- Editor:** Gong XM **L- Editor:** Roemmele A
**E- Editor:** Wu HL



Published by **Baishideng Publishing Group Inc**
8226 Regency Drive, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com



© 2014 Baishideng Publishing Group Inc. All rights reserved.