

遗传关联研究 Meta 分析中文文献方法学质量评价

李帅 方凯 孙傲伊 孙可欣 胡永华

【摘要】 目的 评价遗传关联研究 Meta 分析中文文献的方法学质量。方法 检索中国生物医学数据库、中国期刊全文数据库(CNKI)、万方数据资源系统和维普中国科技期刊全文数据库 2012 年 12 月前的遗传关联研究 Meta 分析文章。2 名研究者独立按纳入排除标准进行文献筛选,意见不一致时咨询第三位研究者。运用系统综述评价工具(AMSTAR)量表对入选文献的方法学质量进行评价。结果 共纳入 440 篇文献。方法学质量得分平均为 5.77 分(满分 11 分)。没有文章符合 AMSTAR 量表全部 11 项条目的要求。入选文章中,89.5% 有前期设计方案,38.6% 在纳入文献和信息提取方面有可重复性,72.7% 实施了广泛文献检索,14.8% 考虑了文献发表状态,10.9% 提供了纳入和排除文献清单,92.5% 描述了纳入文献特征,32.0% 评估了文献质量,50.0% 将文献质量用于结论推导,93.2% 合并文献结果方法恰当,82.3% 评估了发表偏倚,0.5% 声明利益冲突。结论 遗传关联研究 Meta 分析中文文章方法学质量中等,在文献的筛选和信息提取、考虑文献的发表情况、提供纳入和排除文献清单、对文献进行质量评估、声明利益冲突等方面需要提高。

【关键词】 遗传关联研究; Meta 分析; 方法学; 质量评价

Methodological quality of Meta-analyses regarding studies related to genetic association on papers published in Chinese journals LI Shuai, FANG Kai, SUN Ao-yi, SUN Ke-xin, HU Yong-hua. Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China

Corresponding author: HU Yong-hua, Email: yhhu@bjmu.edu.cn

This work was supported by a grant from the National Natural Science Foundation of China (No. 81230066).

【Abstract】 Objective To assess the methodological quality of Meta-analyses on papers published in Chinese journals regarding studies on genetic association. **Methods** Meta-analyses of genetic association study published in Chinese journals up to December 2012 had been searched through on 4 Chinese electronic databases (China biomedicine database, CNKI, Wanfang database and VIP Information). Articles independently selected by both two researchers under definite inclusion and exclusion criteria were included in this study (with consultation on a third researcher if inconsistent opinions existed). A Measurement Tool for the Assessment of Multiple Systematic Reviews (AMSTAR) checklist was used to evaluate the methodological quality together with the methodological quality score being calculated. **Results** Totally, 440 articles were included in this study, with mean score of the methodological quality score (maximum score is 11) as 5.77. Among the 440 articles, none had met the requirement of total 11 items in AMSTAR, with 89.5% and 38.6% of them had a priori design or duplicate study selection and data extraction. 72.7% of them performed a comprehensive literature search and 14.8% included studies regardless of their publication status. 10.9% and 92.5% of them provided a list of including or excluding studies or stated characteristics of the included ones. 32.0% of them assessed the scientific quality of those included studies, with 50.0% of them mentioned their qualities when formulating the conclusions. 93.2% of them used appropriate methods to combine data while 82.3% assessed the likelihood of publication bias. 0.5% of them declared the conflict of interests. **Conclusion** Our results from Meta-analyses showed an acceptable quality regarding their methodology related to genetic association on papers being published in Chinese journals. Improvement is needed especially on aspects as: selection of studies, data extraction, publication status with list of inclusion and exclusion, quality assessment and declaration on conflict of interests, etc.

【Key words】 Genetic association study; Meta-analysis; Methodology; Quality evaluation

DOI: 10.3760/cma.j.issn.0254-6450.2013.09.014

基金项目:国家自然科学基金(81230066)

作者单位:100191 北京大学公共卫生学院流行病与卫生统计学系

通信作者:胡永华, Email: yhhu@bjmu.edu.cn

遗传关联研究 Meta 分析已成为探索疾病与基因关联的常用研究方法^[1]。此方法不仅可以综合不同的研究以获得较多信息,还可以发现不同研究之间潜在的异质性^[2,3]。遗传关联研究 Meta 分析方法学质量的好坏,直接影响其结果真实性和可信度。已有研究发现,此类分析存在许多方法学质量问题^[4-6],但这些研究均针对英文文献进行评价,中文文献的方法学质量尚有待分析。

已有一些规范针对遗传关联研究 Meta 分析进行指导^[7,8]。但是运用这些规范进行评价存在一定困难,如没有成形量表,选取什么条目进行评价没有统一标准,操作不如量表简便。本研究采用系统综述评价工具 (AMSTAR, A Measurement Tool for the Assessment of Multiple Systematic Reviews)^[9,10]对遗传关联研究 Meta 分析中文文献方法学质量进行评价。AMSTAR 是针对 Meta 分析方法学质量的评价量表,由 Shea 等^[10,11]于 2007 年在综合几种标准的基础上得到。目前已有不同研究者验证了此量表的科学性^[12,13]。

资料与方法

1. 检索:检索中国生物医学数据库(CBM)、中国期刊全文数据库(CNKI)、万方数据资源系统和维普中国科技期刊全文数据库中收录的公开发表的中文文献,检索式为(“基因”或“遗传”或“变异”或“突变”或“位点”或“多态性”)和(“系统评价”或“Meta 分析”或“系统综述”或“荟萃分析”),检索项目为题名或关键词,时间为数据库起始时间至 2012 年 12 月。

2. 纳入和排除标准:纳入标准:①公开发表于中文期刊;②遗传关联研究 Meta 分析;③研究对象为人类;④文献研究设计为病例对照研究、队列研究或全基因组遗传关联研究。排除标准:①外文文章或外文文章翻译;②没有明确的基因;③研究的暴露因素为基因表达或其产物;④重复发表的文章(仅选择最近发表的 1 篇)。

3. 筛选:由 2 名研究者独立对检索得到的文献进行筛选,意见不一致时,咨询第三位研究者。搜集纳入文章的全文,难以直接获得全文者,向原作者联系以取得原文。

4. 方法学质量评价工具:采用 AMSTAR 量表评价纳入的文章。此量表含有 11 个条目,每个条目的评价为“是”、“否”、“不清楚”和“不适用”4 项。

5. 信息提取及评价:运用 EpiData 3.1 软件建立数据库,包括文章一般信息和 AMSTAR 量表各条目

内容。由 2 名研究者独立进行信息提取及评价。所有研究者均熟知遗传流行病学和系统评价的基本原理和方法。对提取的信息和评价进行交叉核对,不一致之处咨询第三位研究者。

6. 统计学分析:采用 SPSS 19.0 软件进行统计学分析。描述纳入文章符合 AMSTAR 量表各条目的比例以及方法学质量得分分布情况。分类资料的差异检查采用 χ^2 检验,定量资料的差异采用独立样本 t 检验。 $P < 0.05$ 为差异有统计学意义。方法学得分计算方法:每个条目评价结果为“是”赋 1 分,其他结果赋 0 分,11 个条目得分相加为总方法学质量得分,范围为 0~11 分^[11]。得分越高,说明方法学质量越好。

结 果

1. 文章筛选:共检索到 2003 篇文献,去除重复文献 914 篇,通过阅读题目和摘要初步纳入 488 篇;通过阅读全文,最终纳入 440 篇。

2. 纳入文章一般情况:440 篇文章中,作者单位为大专院校的占 34.8% (153/440),医院占 61.6% (271/440),疾病预防控制中心占 1.8% (8/440),研究所占 1.4% (6/440),其他单位占 0.5% (2/440)。最早的中文遗传关联性研究 Meta 分析文章发表于 1997 年^[14];1998—2000 年没有此类文章发表;2008 年之后,此类文章迅速增多。纳入参考文献数目中位数为 10 (3~61) 篇。150 篇文章的关注人群为中国人,占 34.1%;363 篇文献有阳性结果,占 82.5%。

3. AMSTAR 各条目符合情况:对纳入的 440 篇文章,按 AMSTAR 量表进行评价,结果见表 1。

表 1 440 篇文章符合 AMSTAR 量表各条目情况

条 目	符合文章 数目	百分比 (%)
①提供前期设计方案	394	89.5
②文献纳入排除和信息提取有可重复性	170	38.6
③实施广泛文献检索	320	72.7
④考虑文献发表情况	65	14.8
⑤提供纳入和排除文献清单	48	10.9
⑥描述纳入文献特征	407	92.5
⑦评估纳入文献质量	141	32.0
⑧将文献质量用于结论推导	220	50.0
⑨合并文献结果方法恰当	410	93.2
⑩评估发表偏倚	362	82.3
⑪声明利益冲突	2	0.5

入选文章均有明确研究问题,89.5% (394/440) 有明确纳入标准,58.6% (258/440) 有明确排除标准,

10.0%(44/440)既无明确纳入标准也无明确排除标准。47.3%(208/440)有至少 2 名研究人员进行文献筛选,38.6%(170/440)提供解决不一致意见的方法。

入选文章提供的检索电子数据库数目中位数为 4(0~17),99.3%(437/440)报告了所用的数据库,90.5%(398/440)报告了检索时间范围,80.2%(353/440)报告了检索策略或关键词,45.7%(201/440)运用文献追溯进行文献补充,19.1%(84/440)运用手工检索进行文献补充。15.0%(66/440)纳入灰色文献,90.2%(397/440)纳入不同语言(汉语和英语)文献。97.5%(429/440)提供纳入文献清单,11.6%(51/440)提供排除文献清单,1.8%(8/440)两者均未提供。

407 篇运用一览表描述入选文献特征的文章中,91.2%(371/407)报告了作者,78.1%(318/407)报告了文章发表时间,96.3%(392/407)报告了样本量,74.2%(302/407)报告了基因型频率。141 篇进行了质量评估的文章中,36.2%(51/141)采用自定标准;63.8%(90/141)采用已有标准。68.8%(97/141)说明了评价标准内容;47.5%(67/141)说明了质量评估操作方法。220 篇将文献质量用于结论推导的文章中,48.2%(106/220)没有明确表明对文献进行过质量评估,却在讨论中出现了质量评估结果。

410 篇有异质性检验的文章中,72.4%(297/410)存在异质性。其中,11.1%(33/297)分析了产生异质性的原因,98.0%(291/297)采用随机效应模型进行结果合并,2.7%(8/297)进行 Meta 分析。共有 198 篇文章纳入文献数目 < 10 篇,154 篇文章评估了发表偏倚,占 77.8%。1.7%(6/362)分析了发表偏倚的原因,81.2%(294/362)采用漏斗图进行评估,36.2%(131/362)只采用漏斗图进行评估。声明利益冲突的 2 篇文章均声明自身文章无利益冲突,但是均未声明入选文献的潜在资助来源。

4. 方法学质量得分:440 篇文章中,方法学质量得分平均为(5.77 ± 1.68)分,95% 置信区间(CI)为(5.61 ~ 5.93)分。入选文章得分分布情况见图 1。可见,入选文章得分主要为 4~8 分,占 86.1%(379/440)。

不同发表时间文章的方法学质量得分见图 2。可见,方法学质量得分平均分有随着年份增加而增大的趋势。2012 年发表的文章平均分最高。

5. 不同单位发表的文章质量:按文章的第一作者单位情况,将作者单位分为两类:医院和非医院。不同单位发表的文章符合 AMSTAR 条目的情况见表 2。由表可知,作者单位为医院的文章,在提供前

期设计方案、文献纳入排除和信息提取具备可重复性和考虑文献发表情况方面,均好于单位为非医院的文章。两类单位文章在其余条目上的差异无统计学意义($P > 0.05$)。

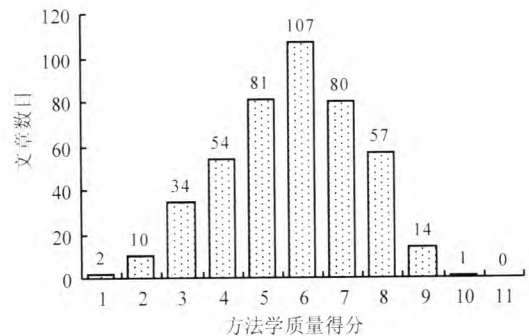
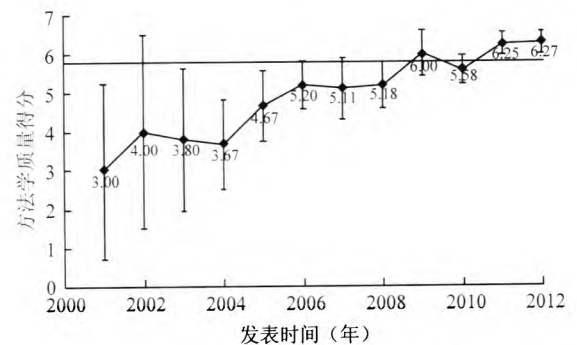


图 1 440 篇文章方法学质量得分分布



注:因 1997 年只有 1 篇文章,绘图时去除,其得分为 3 分;440 篇文章方法学质量平均分为 5.77 分;误差条为 95% CI

图 2 不同发表时间文章方法学质量得分

表 2 不同单位发表的文章 AMSTAR 条目符合情况

条 目	医院 (n=271)	非医院 (n=169)	χ^2 值	P 值
①提供前期设计方案	24(91.9)	14(85.8)	4.114	0.043
②文献纳入排除和信息提取有可重复性	125(46.1)	45(26.6)	16.691	<0.001
③实施广泛文献检索	201(74.2)	119(70.4)	0.740	0.390
④考虑文献发表情况	49(18.1)	16(9.5)	6.134	0.013
⑤提供纳入和排除文献清单	26(9.1)	22(13.0)	1.255	0.263
⑥描述纳入文献特征	253(93.4)	154(91.1)	0.749	0.387
⑦评估纳入文献质量	93(34.3)	48(28.4)	1.672	0.196
⑧将文献质量用于结论推导	143(52.8)	77(45.6)	2.162	0.141
⑨合并文献结果方法恰当	252(93.0)	158(93.5)	0.041	0.839
⑩评估发表偏倚	277(83.8)	135(79.9)	1.076	0.300
⑪声明利益冲突	1(0.4)	1(0.6)	<0.001	1.000

注:括号外数据为频数,括号内数据为百分比(%)

单位为医院的文章方法学质量得分平均为(5.97 ± 1.58)分;非医院的文章方法学质量得分平均为(5.44 ± 1.79)分,两者差异有统计学意义($t = 3.255, P = 0.001$)。

讨 论

Meta 分析在检索和收集文献、筛选文献、资料提取和利益冲突等方面容易产生偏倚。任何方面操作不科学均会影响最终结果的可信性。遗传关联研究 Meta 分析中文文章数目逐年增多, 这些文章方法学质量好坏, 直接关系其偏倚有无和大小, 进而影响其结果真实性。

检索和收集文献过程的偏倚主要是文献检索过程偏倚、发表偏倚和限制语种偏倚等。本研究发现, 72.7% 入选文章实施了广泛文献检索, 在数据库名称、检索策略和时间范围三要素中, 有明确检索策略的文章所占比例最少。只有不足 1/2 文章运用了其他检索方法。15.0% 入选文章纳入灰色文献, 接近 10% 文章进行了语种限制, 提示遗传关联中文文章在文献检索方面仍有待提高。广泛文献检索是防止文献选择偏倚的一个重要步骤, 应检索尽可能多的数据库, 以便广泛搜集文献。电子数据库检索方便, 但并不能保证包括所有文献, 如 Medline 数据库, 其覆盖范围因领域不同而不同^[15], 故应在电子检索同时, 辅以其他检索方法进行补充。

文献发表偏倚和限制语种同样会造成文献选择偏倚。只选择发表的文献可能会造成阳性结果选择偏倚。由于观察性研究没有像随机对照实验那样的记录, 故查找观察性研究 Meta 分析灰色文献更加不易。同时检索各种会议汇编和硕/博士论文库有助于解决这一问题。研究发现, 中国研究者更倾向于将有统计学意义的结果发表于国际期刊^[16], 提示若限制语种则可能会造成严重偏倚。因此, 进行此类研究时, 应至少同时检索中文和英文文献; 若有条件, 还应该检索其他语种文献。

文献筛选过程的偏倚主要是文献纳入标准和排除标准偏倚, 筛选者偏倚等。明确的纳入和排除标准可以保证对相似文献进行结果合并, 从而降低纳入文献间异质性; 至少有 2 名研究者进行筛选可以避免单个研究者主观作用的影响。本研究发现, 47.3% 文章有至少 2 名研究人员进行筛选, 但只有 38.6% 文章有解决不一致意见的方法。提示遗传关联研究中文文章在文献筛选过程存在许多问题, 后续研究应重视这个问题。

本研究发现, 只有 32.0% 文章对文献进行了质量评估。目前, 已有 80 多种观察性研究质量标准^[17], 有些专门针对遗传关联研究^[18, 19]。但本研究发现, 在进行文献质量评估的文章中, 多于 1/3 文章采用自己

制定的标准, 这种小规模研究组制定的标准, 科学性难以验证, 故其评定的文献质量难以保证。虽然文献质量对 Meta 分析的影响存在一定争议^[20], 但它会影响研究偏倚大小和结果真实性, 故应对纳入文献进行质量评估, 对不同质量文献进行敏感性分析, 观察其对结果的影响。

由于不同种族间遗传特异性, 研究对象为不同种族的遗传关联研究之间难免存在异质性, 故遗传关联研究 Meta 分析需要更加注意异质性问题。此类 Meta 分析在异质性处理方面, 探寻异质性原因可能比盲目合并不同文献结果更有意义^[21, 22], 单纯进行文献结果合并可能因为种族混杂的存在而造成结果不准确。Thakkinstian 等^[21]建议, 对所有基因型间的两两比较进行异质性检验, 除非均不存在异质性, 否则不进行数据合并。

发表偏倚评估意指判断纳入文献是否存在发表偏倚。漏斗图是常用方法之一, 但是由于对称性的判断存在很大主观性^[23, 24], 故并不提倡单纯使用漏斗图进行评估。本研究发现 36.2% 文章只使用漏斗图, 此问题需引起重视。有研究发现, 单纯一种评估方法并不足以进行发表偏倚评估^[25], 故建议同时采用多种方法。此外, 若 Meta 分析纳入文献少于 10 篇, 由于统计效力低等问题并不适合进行发表偏倚评估^[26], 但有 77.8% 文章进行了评估。

纳入文章方法学质量得分平均为 5.77 分, 分数至少为 8 分的文章仅有 72 篇, 占 16.3%。虽然文章得分随发表时间有增高趋势, 但 2012 年发表的文章平均分也仅为 6.27 分。提示纳入文章整体方法学质量中等, 有待进一步提高。

在提供前期设计方案, 操作的可重复性和考虑文献发表情况方面, 作者单位为医院的 Meta 分析文章要好于为非医院的文章。

本研究存在局限: 第一, 未涉及遗传关联研究特有内容。本研究只关注传统 Meta 分析方法学内容。第二, 方法学质量得分可能不合理。有的研究者认为, AMSTAR 可以提供质量评价, 但是难以提供定量的质量评价, 故将 AMSTAR 量表各条目量化, 衍生出了 R-AMSTAR^[27], 但其实用性仍待验证。第三, AMSTAR 量表主要用于针对随机对照实验 Meta 分析, 而遗传关联研究为观察性研究, 为非随机研究。研究小组认为, 利用 AMSTAR 量表评价非随机研究 Meta 分析可能会导致发表偏倚; 目前, 此小组正在考虑针对观察性研究对 AMSTAR 量表进行修订^[28]。

参 考 文 献

- [1] Kavvoura FK, Ioannidis JP. Methods for Meta-analysis in genetic association studies: a review of their potential and pitfalls. *Hum Genet*, 2008, 123(1): 1-14.
- [2] Ioannidis JP. Genetic associations: false or true? *Trends Mol Med*, 2003, 9(4): 135-138.
- [3] Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet*, 2006, 38(1): 3-5.
- [4] Attia J, Thakkinstian A, D'Este C. Meta-analyses of molecular association studies: methodologic lessons for genetic epidemiology. *J Clin Epidemiol*, 2003, 56(4): 297-303.
- [5] Minelli C, Thompson JR, Abrams KR, et al. The quality of Meta-analyses of genetic association studies: a review with recommendations. *Am J Epidemiol*, 2009, 170(11): 1333-1343.
- [6] Boccia S, De Feo E, Galli P, et al. A systematic review evaluating the methodological aspects of Meta-analyses of genetic association studies in cancer research. *Eur J Epidemiol*, 2010, 25(11): 765-775.
- [7] Little J, Higgins J, Bray M, et al. The HuGENet™ HuGE review handbook, version 1.0 [EB/OL]. [2013-02-25]. http://www.medicine.uottawa.ca/public-health-genomics/web/assets/documents/HuGE_Review_Handbook_V1_0.pdf.
- [8] Sagoo GS, Little J, Higgins JP. Systematic reviews of genetic association studies. *PLoS Med*, 2009, 6(3): e1000028.
- [9] Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*, 2007, 7(1): 10.
- [10] Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*, 2009, 62(10): 1013-1020.
- [11] Shea BJ, Bouter LM, Peterson J, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS One*, 2007, 2(12): e1350.
- [12] Bessa-Nogueira RV, Vasconcelos BC, Niederman R. The methodological quality of systematic reviews comparing temporomandibular joint disorder surgical and non-surgical treatment. *BMC Oral Health*, 2008, 8: 27.
- [13] Mokkink LB, Terwee CB, Stratford PW, et al. Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual Life Res*, 2009, 18(3): 313-333.
- [14] Shen HB, Yu SZ. Meta analysis on the relationship between gene polymorphism of angiotensin converting enzyme and myocardial infarction. *Dis Contr J*, 1997, 1(4): 267-269. (in Chinese)
沈洪兵, 俞顺章. 血管紧张素转换酶基因多态性与心肌梗塞关系的 Meta 分析. *疾病控制杂志*, 1997, 1(4): 267-269.
- [15] Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and Meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*, 2009, 6(7): e1000100.
- [16] Pan Z, Trikalinos TA, Kavvoura FK, et al. Local literature bias in genetic epidemiology: an empirical evaluation of the Chinese literature. *PLoS Med*, 2005, 2(12): e334.
- [17] Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol*, 2007, 36(3): 666-676.
- [18] Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature*, 2007, 447(7145): 655-660.
- [19] Little J, Higgins J, Ioannidis JP, et al. Strengthening the Reporting of Genetic Association studies (STREGA)—an extension of the STROBE statement. *Eur J Clin Invest*, 2009, 39(4): 247-266.
- [20] Herbison P, Hay-Smith J, Gillespie WJ. Adjustment of Meta-analyses on the basis of quality scores should be abandoned. *J Clin Epidemiol*, 2006, 59(12): 1249-1256.
- [21] Thakkinstian A, McElduff P, D'Este C, et al. A method for Meta-analysis of molecular association studies. *Stat Med*, 2004, 24(9): 1291-1306.
- [22] Zintzaras E, Lau J. Synthesis of genetic association studies for pertinent gene-disease associations requires appropriate methodological and statistical approaches. *J Clin Epidemiol*, 2008, 61(7): 634-645.
- [23] Lau J, Ioannidis JP, Terrin N, et al. The case of the misleading funnel plot. *BMJ*, 2006, 333(7568): 597-600.
- [24] Terrin N, Schmid CH, Lau J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *J Clin Epidemiol*, 2005, 58(9): 894-901.
- [25] Munafò MR, Clark TG, Flint J. Assessing publication bias in genetic association studies: evidence from a recent Meta-analysis. *Psychiat Res*, 2004, 129(1): 39-44.
- [26] Ioannidis JP, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in Meta-analyses: a large survey. *Can Med Assoc J*, 2007, 176(8): 1091-1096.
- [27] Kung J, Chiappelli F, Cajulis OO, et al. From systematic reviews to clinical recommendations for evidence-based health care: validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *Open Dent J*, 2010, 4: 84-91.
- [28] AMSTAR Developments [EB/OL]. [2013-02-27]. <http://www.amstar.ca/Developments.php>.

(收稿日期: 2013-03-16)

(本文编辑: 万玉立)