

乳腺癌筛查领域的系统评价再评价

李江,唐威,王昕,孙鑫,王勇,魏铎佩,陈宏达,李霓,陈万青,代敏
(国家癌症中心/中国医学科学院北京协和医学院肿瘤医院,北京 100021)

摘要: [目的] 评价乳腺癌筛查系统评价的方法学质量及其结论的可靠程度。 [方法] 计算机检索 PubMed、EMBASE、The Cochrane Library (2018 年第 2 期)、Web of Knowledge、WanFang、CNKI 和 CBM 数据库,纳入乳腺癌筛查相关的系统评价,检索时限截止为 2018 年 2 月 28 日。采用 ROBIS 工具评价纳入研究的方法学质量,并基于 GRADE 方法对系统评价纳入的结局指标进行证据质量分级。 [结果] 共纳入 13 个系统评价。ROBIS 工具评价结果显示,阶段一“目标问题与系统评价问题”的四个问题,有 8 个研究是完全符合。阶段二,系统评价制作过程中偏倚风险程度的评价结果为:研究的纳入排除标准领域中,8 个研究偏倚风险程度为“低”,2 个为“高”,3 个“不确定”;研究的检索和筛选领域中,7 个研究偏倚风险程度为“低”,4 个研究为“高”,2 个“不确定”;数据提取和质量评价领域中,6 个研究偏倚风险程度为“低”,7 个研究为“高”;数据合成和结果呈现领域中,6 个研究偏倚风险程度为“低”,6 个研究偏倚风险程度为“高”,1 个“不确定”。阶段三,偏倚风险的评价结果为:7 个系统评价为“低”风险偏倚,6 个系统评价为“高”风险偏倚。运用 GRADE 分级结果显示,乳腺癌筛查降低死亡率结局指标中 2 个研究的证据质量为 A 级,2 个为 B 级,3 个由于不精确性和发表偏倚而为 C 级(低质量)证据。乳腺癌检出率中 1 个研究为 A 级证据,1 个为 B 级,其余 2 个研究为 C 级。假阳性的 2 个研究的证据质量为 C 级。过度筛查中 2 个系统评价均为 C 级。筛查潜在危害中 1 个研究的证据级别为 B 级。 [结论] 当前针对乳腺癌筛查系统评价的方法学质量总体尚可,但结论的证据水平较低,对筛查终点结局的数据收集不足。因此,临床医生在使用这些证据进行临床决策时尚需依据当地实际情况讨论后应用。

关键词: 乳腺肿瘤;筛查;系统评价再评价;质量评估;证据分级

中图分类号:R737.9 文献标识码:A 文章编号:1004-0242(2018)06-0401-08

doi:10.11735/j.issn.1004-0242.2018.06.A001

Breast Cancer Screening: An Overview of Systematic Reviews

LI Jiang, TANG Wei, WANG Xin, et al.

(National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China)

Abstract: [Purpose] To evaluate the quality of methodology and evidence in systematic reviews (SRs) of breast cancer screening. [Methods] Electronic searches were carried out in the PubMed, The Cochrane Library (Issue 2, 2018), EMBASE, Web of Knowledge, CBM, WanFang Data and CNKI to collect SRs of breast cancer screening from inception to 28 February 2018. The ROBIS tool was applied for grading quality of methodology and the GRADE system was applied for grading quality of evidence of included SRs. [Results] A total of 13 SRs including 5 outcomes (mortality, detection rate, false-positive, over-diagnosis and potential benefits and harms) were included. The ROBIS evaluation showed that 8 studies were completely matched the 4 questions of phase I; in the phase II, 8 studies were in low risk of bias in the including criteria field, 7 studies were low risk of bias in the literature search and screening field, 6 studies were in low risk of bias in the data abstraction and quality assessment field, and 6 studies were in low risk of bias in the data synthesis field; in the phase III, 7 studies were in low risk. The GRADE classification showed that 2 studies had level A evidence on the outcome of mortality; 1 study had level A evidence on detection; others had level B and C on evidence. [Conclusion] The quality of methodology of SRs is modest; while the quality of evidence is low. The outcomes of screening have not been collected sufficiently. Therefore, when clinicians make clinical decisions they have to use these evidence with the local context.

Key words: breast neoplasms; screening; overview of review; quality assessment; evidence classified

乳腺癌严重威胁全球女性健康,其造成的疾病

负担日益加重。近年来,全球疾病负担数据库(global burden of disease, GBD)中 188 个国家的相关数据资料显示,近 1500 万新发癌症患者中,约 180 万为乳腺癌患者,大约占女性癌症新发总数的 1/4,为绝大多数国家女性癌症发病率首位;同年,共

收稿日期:2018-03-24;修回日期:2018-04-13

基金项目:北京市优秀人才培养资助青年拔尖团队项目(2017000021223TD05);国家自然科学基金委员会项目(81602930)

通讯作者:李霓, E-mail: nli@cicams.ac.cn;

陈万青, E-mail: chenwq@cicams.ac.cn

有 46.4 万患者死于乳腺癌,是女性罹患癌症中导致死亡人数最多的一种,为近一半国家中女性癌症导致死亡中最多的一种^[1,2]。多个国家已陆续开展乳腺癌筛查项目,但其对死亡率的降低、过度诊断等问题尚存争议,同时不同国家间在疾病负担和卫生资源等方面存在差异,何种乳腺癌筛查项目最具有有效性和安全性也不得而知^[3]。随着循证医学的快速发展,循证肿瘤学的研究逐步深入。已有多个系统评价的结果阐述乳腺癌筛查的有效性和安全性。通常认为,基于严格设计的随机对照试验或观察性研究的系统评价,能够提供科学性较高的证据用于卫生决策以制定操作性强的推荐意见和指南,但由于系统评价方法学及结果质量等局限性,所得的证据是否可以被判定为高质量的证据收录于指南中尚待讨论。

本研究通过开展乳腺癌筛查系统评价的再评价^[4],评估其纳入研究的方法学质量及其结论的可靠程度,探索乳腺癌筛查有效性和安全性的特点和规律,为临床更好地进行乳腺癌筛查提供参考依据。

1 资料与方法

1.1 纳入与排除标准

纳入标准:①基于随机对照试验(randomized controlled trial, RCT)和/或观察性研究(包括队列研究和病例对照研究)的干预性系统评价。②研究对象为接受乳腺癌筛查方案的受试者,性别、种族和病例来源不限。年龄依据乳腺癌筛查正常或高危人群,限定年龄 40 岁以上的女性^[3]。③采用每年一次或每两年一次的 X 线钼靶或钼靶联合超声筛查其他检查等。④结局指标包括死亡率、检出率、假阳性率、过度诊断和潜在危害。

排除标准:①会议收录摘要;②重复发表文献;③评述类文献;④无法提取数据的文献;⑤非中、英文文献。

1.2 检索策略

计算机检索 PubMed、The Cochrane Library (2018 年第 2 期)、EMBASE、Web of Knowledge (SCI)、CBM、WanFang 和 CNKI 数据库,时间为从建库至 2018 年 2 月 28 日,纳入有关于乳腺癌筛查的系统评价。检索策略采用主题词和自由词相结合的方式。中文检索词:乳腺癌、乳腺肿瘤、筛查、早期诊

断、早期检测、系统评价、Meta 分析、荟萃分析、元分析;英文检索词:breast cancer, breast neoplasms, screening, early detection, Cochrane database of systematic reviews, meta-analysis 等。采用主题词联合关键词的检索方法。

1.3 文献筛选和资料提取

由 2 位评价人员独立进行文献筛选和资料提取,如遇分歧,则咨询第三方协助判断,缺乏的资料尽量与作者联系予以补充。文献筛选时首先阅读标题和摘要,在排除明显不相关的文献后,进一步阅读全文,以确定最终是否纳入。

采用自制的资料提取表提取资料,提取内容主要包括:①基本信息,包括题目、作者、发表年份、检索策略、纳入标准、质量评价方法、结局指标和结论等;②纳入研究的方法学质量和证据质量情况。

1.4 纳入研究的质量评价

该内容包括两个部分,第一部分为采用 ROBIS (risk of bias in systematic review)量表评价纳入研究的方法学质量^[5],第二部分为采用 GRADE 证据分级工具评价证据质量^[6]。分别由 2 位评价人员进行评价。

方法学质量评价采用 ROBIS 量表,包括三个阶段^[5],阶段一:评估相关性(视情况选择):系统评价类型确定(包括干预性、病因性、诊断准确性试验和预后性);阶段二:确定系统评价制定过程中的偏倚风险程度(包括研究的纳入排除标准、研究检索和/或筛选所使用方法、数据提取和质量评价所使用方法和数据合成和结果呈现这四个领域的 21 个标志性问题);阶段三:判断系统评价偏倚风险的 3 个标志性问题。

证据质量评价采用 GRADE 证据分级工具,包括 5 个降级因素和 3 个升级因素^[6]。对于原始是高质量的随机对照试验,可采用 5 个降级因素包括:①纳入的研究在随机、分配隐藏和盲法方面存在较大偏倚;②漏斗图不对称;③不同研究可信区间的重叠程度较差,且合并结果的 P 值较大;④纳入研究样本量太小,可信区间较宽;⑤纳入研究数量少,存在较大发表偏倚的可能。对于初始是低质量的观察性研究,可采用 3 个升级因素包括:①效应量大;②剂量反应;③所有可能的混杂因素。将纳入的证据质量分为 A(高质量)、B(中等质量)、C(低质量)和 D(极

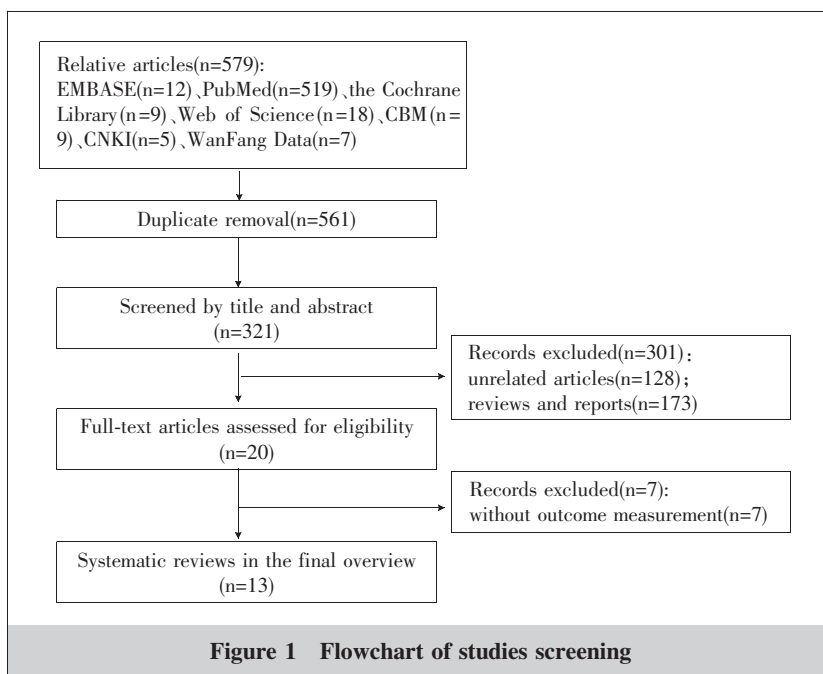
低质量)四个等级。

此外,为保证评价结果的可靠性,随机抽取2个纳入研究进行独立评估,共同衡量两名评价人员间对指南评价结果的一致性。

2 结果

2.1 文献筛选流程及结果

根据检索流程共检出579篇文献,经阅读题目和摘要逐层筛选获取全文后,最终纳入13个系统评价/Meta分析^[7-19](Figure 1)。



2.2 纳入研究的基本特征

文献发表年份在2007~2016年之间。纳入的系统评价中,5个^[7-9,11,19](39%)是美国的研究,2个^[16,17](15%)是澳大利亚的研究,2个^[10,15](15%)是丹麦的研究,瑞典、荷兰、意大利和日本各1个^[12-14,18](31%)。

纳入研究中的研究年龄为42岁(35~71岁)。4个研究^[10,12,17,18]报告了筛查随访时间;8个研究^[7,9-12,16,17,19]报告使用了质量评估工具,其中1个研究^[9]采用的是Cochrane 5.1.0推荐的偏倚风险评估工具,2个研究^[12,16]使用了QUADAS (quality assessment of diagnostic accuracy studies) 工具评判偏倚风险,1个研究^[9]使用其研究组自行修正的QUADAS标准,1个研究^[17]

使用Newcastle-Ottawa评分量表,2个研究^[11,19]使用的是其研究组如USPSTF研制的风险评估标准,1个研究^[7]使用的是其所属循证医学研究中心自制的评估量表(Table 1)。

2.3 纳入系统评价的方法学质量评价

2.3.1 阶段一:评估相关性

11个系统评价^[7-15,17,19]细化了纳入受试者人群的年龄和地域;12个系统评价^[7-12,14-19]中要解决的干预措施问题与目标干预措施完全匹配;12个系统评价^[7-12,14-19]要解决的对照措施问题与目标对照措施匹配;9个系统评价^[7,9-12,14-16,19]要解决的结局指标全面包括了目标结局指标,其余为部分匹配(Table 2)。

2.3.2 阶段二:偏倚风险程度

领域1:3个系统评价^[10,11,14]完全遵循了预先确定的目标和纳入标准,3个^[7,8,15]可能是遵循了预先确定的目的和纳入标准,7个研究^[9,12,13,16-19]无相关信息;7个系统评价^[9,10,15-19]的纳入标准适合系统评价的问题,5个^[7,8,11-13]可能适合,1个研究^[14]无相关信息;8个系统评价^[9,10,12,13,15-17,19]纳入标准明确且详细,3个^[7,8,18]较详细,2个研究^[11,14]无相关信息;7个系统评价^[10-12,16-19]纳入标准中所有基于研究特征的限制合适,5个^[7-9,13,15]可能合适,1个研究^[14]无相关信息。

因此,该领域中,纳入标准相关的绝大多数标志性问题的回答为“是”或“可能是”,8个研究^[9,10,12,13,16-19]的偏倚风险程度为“低”,2个^[11,14]为“高”,3个^[7,8,15]为“不确定”(Table 3)。

领域2:7个系统评价^[8-10,12,16,17,19]的检索已发表和未发表的研究时所包含的数据库或电子资源的范围合适;10个系统评价^[7-12,16-19]使用了除数据库检索以外的其他方法来确定相关研究;6个系统评价^[9-12,16,19]纳入检索策略的检索词和结构能尽可能多地检索到符合的研究;7个系统评价^[10-12,15-17,19]基于时间、发表形式、语言的限制合适;6个系统评价^[11,12,16-19]研究的筛选中尽可能地减小了误差。

因此,该领域中,7个研究^[8-10,12,16,17,19]其偏倚风险程度为“低”,4个研究^[13-15,18]为“高”,2个研究^[7,11]

Table 1 Characteristics of included systematic reviews

Included studies	Region	Type of included studies	Age (years)	Interventions	Outcomes	Follow-up periods (T/C)	Quality assessment tool
Armstrong 2007 ^[7]	USA	117 RCTs+OSs	40~49	Mammography	①③	Unreported	Self-made scale
Brewer 2007 ^[8]	USA	23 OSs	≥40	Mammography, ultrasonography	③	Unreported	None
Noble 2009 ^[9]	USA	3 OSs+4 CSs	23~95	Mammography	②③	Unreported	QUADAS revised edition
Gøtzsche PC 2009 ^[10]	Denmark	4 RCTs+3 CSs	50~69	Mammography, ultrasonography	①	6.5 years	Cochrane handbook
Harris 2011 ^[11]	USA	17 CSs	50~69	Mammography	①	Unreported	USPSTF criteria
Azavedo 2012 ^[12]	Sweden	4 CSs	40~74	Mammography	②	3 years	QUADAS
Broeders 2012 ^[13]	Netherlands	12 OSs	40~74	Mammography	①	Unreported	None
Puliti 2012 ^[14]	Italy	13 CSs	20~84	Mammography	④	Unreported	None
Njor 2012 ^[15]	Denmark	20 OSs	50~69	Ultrasonography with mammography, mammography	①	Unreported	None
Vreugdenburg 2013 ^[16]	Australia	DTE:8 CSs; EIS: 11 CSs	36~58	Radiography, electrical impedance scanning	②	Unreported	QUADAS
Leung 2013 ^[17]	Australia	28 CSs	≥40	Mammography	②	3 years	Newcastle-Ottawa Scale
Hamashima 2015 ^[18]	Japan	8 RCTs	40~74	Mammography, palpation	①	5 years	None
Nelson 2016 ^[19]	USA	9 RCTs	39~74	Mammography, ultrasonography	①③④⑤	Unreported	USPSTF criteria

Note: RCT: randomized controlled trial; OS: observation study; CS: cohort study; CCS: case-control study; ①: mortality; ②: detection rate; ③: false positive; ④: over-diagnosis; ⑤: potential harm.

为“不确定”(Table 3)。

领域 3: 5 个系统评价^[8,10,16,17,19]数据资料提取过程中尽可能地减小了误差; 6 个系统评价^[10,12,14,15,18,19]为作者和读者提供足够的信息以获取研究特征来解读结果; 9 个系统评价^[7,10-12,14,15,17-19]提取了所有相关的研究结果来进行数据合成; 6 个系统评价^[9-12,16,17]使用了合适的工具来评价原始研究的偏倚风险或/和方法学质量; 6 个系统评价^[10-12,16,17,19]偏倚风险评价中尽可能地减小了误差。

因此, 该领域中 6 个研究^[10-12,16,17,19]偏倚风险程度为“低”, 7 个研究^[7-9,13-15,18]为“高”偏倚风险(Table 3)。

领域 4: 12 个系统评价^[7-15,17-19]数据分析合成纳入了所有符合要求的研究; 5 个系统评价^[9-11,14,17]遵循了所有预先设置的数据分析方法; 9 个系统评价^[8-10,12,13,16-19]基于纳入研究问题、设计和结局指标的性质和相似性, 数据合成方法恰当; 7 个系统评价^[8,9,13,16-19]数据合成中研究之间的异质性最小; 4 个系统评价^[9,10,16,17]做了敏感性分析以保证结果稳定; 4 个系统评价^[10,16,17,19]原始研究的偏倚最小。

因此, 该领域中 6 个研究^[8-10,16,17,19]偏倚风险程度为“低”, 6 个研究^[7,12-15,18]偏倚风险程度“高”, 1 个^[11]“不确定”(Table 3)。

2.3.3 阶段三: 偏倚风险判定

7 个系统评价^[8,10-12,16,17,19]遵循了基于阶段二中 4 个领域偏倚风险的结果解释; 9 个^[7-11,16-19]较合理地考虑到了纳入研究与系统评价研究问题的相关性; 12 个研究^[7-13,15-19]的评价者避免强调有统计学意义的结果。

Table 2 Phase I assessing relevance of target question

Included studies	Target population	Intervention	Control	Outcome
Armstrong 2007 ^[7]	Yes	Yes	Yes	Yes
Brewer 2007 ^[8]	Yes	Yes	Yes	Partly
Noble 2009 ^[9]	Yes	Yes	Yes	Yes
Gøtzsche PC 2009 ^[10]	Yes	Yes	Yes	Yes
Harris 2011 ^[11]	Yes	Yes	Yes	Yes
Azavedo 2012 ^[12]	Yes	Yes	Yes	Yes
Broeders 2012 ^[13]	Yes	Partly	Yes	Partly
Puliti 2012 ^[14]	Yes	Yes	Yes	Yes
Njor 2012 ^[15]	Yes	Yes	Partly	Yes
Vreugdenburg 2013 ^[16]	Partly	Yes	Yes	Yes
Leung 2013 ^[17]	Yes	Yes	Yes	Partly
Hamashima 2015 ^[18]	Partly	Yes	Yes	Partly
Nelson 2016 ^[19]	Yes	Yes	Yes	Yes

Table 3 Risk of bias of phase II and phase III

Included studies	Phase II				Phase III: risk of Bias
	Domain 1: study eligibility criteria	Domain 2: identification and selection of studies	Domain 3: data collection and study appraisal	Domain 4: synthesis and finding	
Armstrong 2007 ^[7]	Unclear	Unclear	High	High	High
Brewer 2007 ^[8]	Unclear	Low	High	Low	Low
Noble 2009 ^[9]	Low	Low	High	Low	High
Gøtzsche PC 2009 ^[10]	Low	Low	Low	Low	Low
Harris 2011 ^[11]	High	Unclear	Low	Unclear	Low
Azavedo 2012 ^[12]	Low	Low	Low	High	Low
Broeders 2012 ^[13]	Low	High	High	High	High
Puliti 2012 ^[14]	Unclear	High	High	High	High
Njor 2012 ^[15]	High	High	High	High	High
Vreugdenburg 2013 ^[16]	Low	Low	Low	Low	Low
Leung 2013 ^[17]	Low	Low	Low	Low	Low
Hamashima 2015 ^[18]	Low	High	High	High	High
Nelson 2016 ^[19]	Low	Low	Low	Low	Low

综合阶段二评价中的偏倚风险,最终得7个系统评价^[8,10-12,16,17,19]为“低”风险偏倚,6个系统评价^[7,9,13-15,18]为“高”风险偏倚(Table 3)。

2.4 主要观察指标及证据质量分级

2.4.1 死亡率

7个研究^[7,10,11,13,15,18,19]报告了乳腺癌筛查长期随访后的死亡率情况。其中,2个研究^[13,19]显示每年一次的钼靶筛查方式可降低25%死亡率(RR=0.75, 95%CI:0.69~0.81;RR=0.75,95%CI:0.63~0.89),另1个研究^[18]显示在40~74岁女性中实施钼靶或联合方式筛查可显著降低乳腺癌死亡率(RR=0.75,95%CI:0.67~0.83),3个基于随机对照试验的系统评价^[10,11,15]并未发现筛查手段可降低10年死亡率,RR分别为1.02(95%CI:0.95~1.10)、0.99(95%CI:0.95~1.03)和0.99(95%CI:0.96~1.01)。另1个研究^[7]未报告详细信息。基于GRADE证据分级,2个研究^[10,19]的证据质量为A等级,2个研究^[11,18]的证据质量为B等级,3个研究^[7,13,15]的证据质量由于不精确性和存在发表偏倚而为C级(低质量)证据。

2.4.2 检出率

4个研究^[9,12,16,17]均报告了乳腺癌筛查的灵敏度和特异性。一个研究^[9]显示钼靶的灵敏度和特异性分别为86%和88%,另1个研究^[16]提供灵敏度和特异性的范围分别为25%~97%和12%~85%。其他研究^[12,17]无法获取具体效应量的详细信息。

1个研究^[9]所纳入的队列研究由于效应量大而

证据升级,最终研究证据为A级证据;1个研究^[16]由于结果的不一致性而为B级中等质量证据;其余2个研究^[12,17]由于结果的不一致性、不精确性与发表偏倚而最终证据级别为C级。

2.4.3 假阳性率

4个研究^[7-9,19]提到了筛查的假阳性率,2个研究^[7,9]未给出详细的信息,因其不一致性故证据质量为C级;2个研究^[8,19]证据质量为B1级。

2.4.4 过度筛查

2个系统评价^[14,19]讨论了过度筛查的危害,1个结果^[19]给予的范围值为11%~22%,另1个研究未提供详细信息。2个研究的证据质量由于结果的不一致性和不精确性而降级为C级证据。

2.4.5 筛查潜在危害

1个研究^[19]报告了筛查潜在危害,未报告具体的效应量分析,但由于其报告了可能混杂的原因,证据级别为B级中等质量。

3 讨论

本研究首次使用ROBIS工具评估乳腺癌筛查领域系统评价的方法学质量及采用GRADE评估系统评价证据质量。研究结果显示,方法学质量方面,7个系统评价风险程度低;6个系统评价风险程度高。分析偏倚来源因素,则发现偏倚主要存在于研究的检索和筛选、数据提取和质量评价、数据合成和结果

呈现部分;证据质量方面,显示有关于乳腺癌筛查对于死亡率和检出率的结局中证据质量较高,其他结局指标如过度筛查和潜在危害中证据级别较低。

系统评价再评价(overview of review,OR)是基于Cochrane系统评价而来,是更加科学严谨的一种评价方法。它是全面搜集针对同一疾病或同一健康问题的病因、诊断、治疗或预后等方面的相关系统评价/Meta分析进行再评价的一种综合研究方法,能够鉴别系统评价存在的方法学偏倚及其结论证据质量的高低,为决策者提供更为集中的高质量证据^[20]。自2010年以来,系统评价/Meta分析的数量成倍数增长^[21],众多评估方法学质量的工具应运而生,如OQAQ(overview quality assessment questionnaire),CASP(critical appraisal skill program),AMSTAR(a measurement tool to assess systematic reviews)和ROBIS(risk of bias in systematic review)等。ROBIS工具是2014年英国布里斯托尔大学(University of Bristol)社会医学部制定发布的一种全新的评价工具,也是目前值得推荐的一个系统评价偏倚风险评价工具,在近三届Cochrane Colloquium年会均对其进行了专题讨论和相关培训。本研究组在乳腺癌筛查领域的系统评价再评价中探索使用了ROBIS工具进行评估。GRADE证据质量分级2000年由McMaster大学成立的国际证据分级与推荐工作组开始创立,2004年面试,自2008年起,就已经被世界卫生组织(WHO)、美国医疗保健研究与质量局(agency for healthcare research and quality,AHRQ)等相关组织和机构认可,并用于医学研究所获证据的分级评判中。

本研究严格按照ROBIS工具要求进行,在目标问题、研究的纳入排除标准、研究的检索和筛选、数据提取和质量评价以及数据合成和结果呈现上对每个纳入的系统评价均做出了较高的要求,结果显示ROBIS评估后所纳入的系统评价高低偏倚风险程度各占一半,而乳腺癌筛查领域的系统评价多是由于原始研究的试验设计、数据分析和质量评价、数据合成和结果呈现部分存在有相对较大的风险。本研究纳入系统评价的终点结局最主要是死亡率,4个系统评价纳入了三个设计严谨的大样本随机对照试验,其他均基于队列研究。单从系统评价结论看,乳腺癌筛查具有确定的有效性,但安全性尚不得知。癌

症的筛查,需要的随访时间较长,实施随机分配隐藏难度较大,随访率等问题堪忧,如在大规模依从性较好的高风险人群队列中进行筛查,检出率较高,能有效地降低死亡率,但是安全性需要更加较长的时间去验证,现有的系统评价中有探讨安全性的问题,但是由于原始研究的数据较少而未能给出确切的证据支持。虽然队列研究结果在一定程度上证据质量能够支持综合证据的质量,但如果原始研究就为高质量的随机对照试验,那么所得的结果在一定程度上就具备推广意义^[22]。运用GRADE证据质量等级分级工具降级最多的因素是研究的不一致性和发表偏倚,说明系统评价纳入的随机对照试验和队列研究等在时间以及人群综合数据分析方面存在较大偏倚。基于此,我们建议筛查领域系统评价制作时,就应当采用GRADE证据分级工具对所要纳入的原始研究进行证据质量的划分,以得到高质量的、与真实情况具备相关性的结论,形成丰富的高质量证据资源,为相关筛查推荐意见的构建提供了可靠的证据支持。

基于此,我们建议今后乳腺癌筛查领域,乃至癌症筛查领域的系统评价研究人员在制定系统评价前,应做到以下几点:①应该制定计划书,同时在系统评价的官方地址进行注册或发表以避免出现类似研究相关性的偏倚^[23];②研究人员应该接受医学文献检索和资料提取等相关培训以避免研究检索和筛选过程中产生的偏倚;③在方法学质量评价、证据质量评级等步骤时需严格遵循独立性和一致性的实施原则;④在制定系统评价之前,对ROBIS工具和GRADE工具进行学习和研讨,以降低系统评价的偏倚风险程度,提高系统评价所得证据的可信度。

世界卫生组织国际癌症研究署(international agency for research on cancer,IARC)推荐要适当地、合理地进行大范围乳腺癌筛查。我国亦有较大规模的乳腺癌筛查项目^[24-26],如何经济且有效地实施筛查项目是一个挑战。在癌症筛查领域引入循证医学的方法和理念,提供筛查有效性和安全性的可靠、科学的依据,并积极开展筛查的循证评价将有利于提高国家及各级别乳腺癌筛查项目的水平,可促进我国癌症筛查项目的国际化进程。WHO明确提出卫生健康实践指南/推荐意见等必须基于循证的方法论构建,并也明确系统评价的重要作用之一是为临床

指南/推荐意见提供高质量的证据支持,证据分级研究团队专家建议系统评价制作者在确定研究题目时,就应当考虑到所制作的系统评价将会作为实践指南的参考依据,应当严格按照方法学评估工具与证据质量分级工具细致研究,严谨地产生高质量证据,从而提高被引用的效率,避免研究资源的浪费^[27-29]。

本研究存在一定的局限性,仅检索中英文乳腺癌筛查系统评价,语种的限制可能会造成不完全的检索。我们会在今后的研究中,尽量纳入其他语种的系统评价,通过专业人员进行翻译。所以本研究仅限于探索性的使用 ROBIS 和 GRADE 工具进行系统评价再评价时的方法学评价和证据质量评价。

总之,当前针对乳腺癌筛查系统评价的方法学质量总体尚可,结论的证据水平中等,对筛查终点结局的数据收集不足。因此,建议临床医生在使用这些证据进行临床决策时尚需依据当地实际情况讨论后应用。

参考文献:

- [1] US Burden of Disease Collaborators, Mokdad AH, Ballestros K, et al. The State of US Health, 1990–2016: burden of diseases, injuries, and risk factors among US States [J]. JAMA, 2018, 319(14): 1444–1472.
- [2] Marmot MG, Altman DG, Cameron DA, et al. The benefits and harms of breast cancer screening: an independent review[J]. Br J Cancer, 2013, 108(11): 2205–2240.
- [3] Stout NK, Lee SJ, Schechter CB, et al. Benefits, harms, and costs for breast cancer screening after US implementation of digital mammography [J]. J Natl Cancer Inst, 2014, 106(6): 92.
- [4] Whiting P, Savović J, Higgins JPT, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed[J]. J Clin Epidemiol, 2016, 69: 225–234.
- [5] Bühn S, Mathes T, Prengel P, et al. The risk of bias in systematic reviews tool showed fair reliability and good construct validity[J]. J Clin Epidemiol, 2017, 91: 121–128.
- [6] Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations [J]. BMJ, 2008, 336(7650): 924–926.
- [7] Armstrong K, Moye E, Williams S, et al. Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians [J]. Ann Intern Med, 2007, 146(7): 516–526.
- [8] Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms[J]. Ann Intern Med, 2007, 146(7): 502–510.
- [9] Noble M, Bruening W, Uhl S, et al. Computer-aided detection mammography for breast cancer screening: systematic review and meta-analysis[J]. Arch Gynecol Obstet, 2009, 279(6): 881–890.
- [10] Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography [J]. Cochrane Database Syst Rev, 2009, 4: CD001877.
- [11] Harris R, Yeatts J, Kinsinger L. Breast cancer screening for women ages 50 to 69 years a systematic review of observational evidence[J]. Prev Med, 2011, 53(3): 108–114.
- [12] Azavedo E, Zackrisson S, Mejøre I, et al. Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review [J]. BMC Med Imaging, 2012, 12: 22.
- [13] Broeders M, Moss S, Nyström L, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies[J]. J Med Screen, 2012, 19 Suppl 1: 14–25.
- [14] Puliti D, Duffy SW, Miccinesi G, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review[J]. J Med Screen, 2012, 19 Suppl 1: 42–56.
- [15] Njor S, Nyström L, Moss S, et al. Breast cancer mortality in mammographic screening in Europe: a review of incidence-based mortality studies[J]. J Med Screen, 2012, 19 Suppl 1: 33–41.
- [16] Vreugdenburg TD, Willis CD, Mundy L, et al. A systematic review of electrography, electrical impedance scanning, and digital infrared thermography for breast cancer screening and diagnosis[J]. Breast Cancer Res Treat, 2013, 137(3): 665–676.
- [17] Leung J, McKenzie S, Martin J, et al. Effect of rurality on screening for breast cancer: a systematic review and meta-analysis comparing mammography[J]. Rural Remote Health, 2014, 14(2): 2730.
- [18] Hamashima C, Ohta K, Kasahara Y, et al. A meta-analysis of mammographic screening with and without clinical breast examination[J]. Cancer Sci, 2015, 106(7): 812–818.
- [19] Nelson HD, Cantor A, Humphrey L, et al. Screening for breast cancer: a systematic review to update the 2009 U.S. Preventive Services Task Force Recommendation. Rockville(MD): Agency for Healthcare Research and Quality(US)[M]. U.S. Preventive Services Task Force Evidence

- Syntheses, 2016.
- [20] Yang KH, Liu YL, Yuan JQ, et al. The development of overviews of reviews[J]. Chinese Journal of Evidence-Based Pediatrics, 2011, 6(1): 54-57. [杨克虎, 刘雅莉, 袁金秋, 等. 发展和完善中的系统评价再评价[J]. 中国循证儿科杂志, 2011, 6(1): 54-57.]
- [21] Tian JH, Chen YL, Yang KH, et al. Progresses and challenges for meta-analysis or systematic review[J]. Journal of Lanzhou University(Medical Sciences), 2016, 42(1): 42-46. [田金徽, 陈耀龙, 杨克虎, 等. SR/MA 研究进展与挑战[J]. 兰州大学学报(医学版), 2016, 42(1): 42-46.]
- [22] Brouwers M, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in healthcare[J]. CMAJ, 2010, 182(18): e839-e842.
- [23] Dal-Ré R. Clinical trials transparency: where are we today?[J]. Trends Cancer, 2018, 4(1): 1-3.
- [24] Dai M, Shi JF, Li N. The design of cancer early detection and treatment in urban China[J]. Chinese Journal of Preventive Medicine, 2013, 47(2): 179-182. [代敏, 石菊芳, 李霓. 中国城市癌症早诊早治项目设计及预期目标[J]. 中华预防医学杂志, 2013, 47(2): 179-182.]
- [25] Wei WQ. Ideas on cancer screening in urban area[J]. China Cancer, 2010, 19(2): 97-98. [魏文强. 关于城市地区癌症筛查工作的几点思考[J]. 中国肿瘤, 2010, 19(2): 97-98.]
- [26] Wu F, Liu XY, Zhao GM, et al. Health economic evaluation of cancer screening [J]. China Cancer, 2016, 25(2): 81-87. [吴菲, 刘霄宇, 赵根明, 等. 癌症筛查成本效果评估的研究进展[J]. 中国肿瘤, 2016, 25(2): 81-87.]
- [27] Wu QF, Ding HF, Deng W, et al. ROBIS: new tool on assessment risk of bias of systematic review [J]. Chinese Journal of Evidence-Based Medicine, 2015, 15(12): 1454-1457. [吴琼芳, 丁泓帆, 邓围, 等. ROBIS: 评估系统评价偏倚风险的新工具[J]. 中国循证医学杂志, 2015, 15(12): 1454-1457.]
- [28] Ding HF, Wu QF, Yang N, et al. Interpretation of ROBIS tool in evaluating the risk of bias of a selected systematic review [J]. Chinese Journal of Evidence-Based Medicine, 2016, 16(1): 115-121. [丁泓帆, 吴琼芳, 杨楠, 等. 评估系统评价偏倚风险的 ROBIS 工具实例解读 [J]. 中国循证医学杂志, 2016, 16(1): 115-121.]
- [29] Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence[J]. J Clin Epidemiol, 2011, 64(4): 401-406.

《中国肿瘤》《肿瘤学杂志》联合征订征稿启事

《中国肿瘤》杂志由国家卫生和计划生育委员会主管,中国医学科学院、全国肿瘤防治研究办公室主办,中国肿瘤医学综合类科技月刊,中国科学引文数据库(CSCD)来源期刊、中国科技论文统计源期刊、中国科技核心期刊 (ISSN 1004-0242 CN11-2859/R),大 16 开,80 页,单价 8 元,全年 96 元,邮发代号:32-100。主编赫捷院士。以交流肿瘤防治经验,推广肿瘤科技成果,促进肿瘤控制事业的发展为宗旨。紧扣肿瘤控制主题,尤其在肿瘤预防、流行病学方面独树一帜。每期刊出相应专题报道,配有癌情监测、防治工作、医院管理、研究进展、论著等栏目,已成为社会各方了解我国肿瘤防控工作进展和动态的重要论坛。

《肿瘤学杂志》为面向全国肿瘤学术类科技月刊,中国科技论文统计源期刊、中国科技核心期刊 (ISSN 1671-170X CN 33-1266/R),大 16 开,80 页,单价 8 元,全年 96 元,邮发代号:32-37。由浙江省卫生和计划生育委员会主管,浙江省肿瘤医院和浙江省抗癌协会联合主办,报道我国肿瘤学术研究领域的新技术、新成果和新进展,刊登肿瘤临床与基础类学术论文,报道重点为常见恶性肿瘤诊治研究,指导临床实践和科研。公平、公正,择优录用稿件,好稿快发。当地邮局订阅,漏订者可向编辑部补订。

两刊地址:浙江省杭州市拱墅区半山东路 1 号(310022) 咨询电话和传真: 0571-88122280

两刊网址: <http://www.chinaoncology.cn>