

## · 癌症筛查指南及共识质量评价 ·

## 全球乳腺癌筛查指南质量评价

李江<sup>1</sup> 杨珂璐<sup>2</sup> 蔡依彤<sup>2</sup> 田金徽<sup>2</sup> 郑亚迪<sup>1</sup> 温艳<sup>1</sup> 杨卓煜<sup>1</sup> 李霓<sup>1</sup> 陈万青<sup>1</sup>  
赫捷<sup>3</sup>

<sup>1</sup>国家癌症中心/国家肿瘤临床医学研究中心/中国医学科学院北京协和医学院肿瘤医院  
癌症早诊早治办公室, 北京 100021; <sup>2</sup>兰州大学循证医学中心 730000; <sup>3</sup>国家癌症中心/  
国家肿瘤临床医学研究中心/中国医学科学院北京协和医学院肿瘤医院胸外科, 北京 100021  
李江和杨珂璐对本文有同等贡献

通信作者: 李霓, Email: nil@cicams.ac.cn

**【摘要】目的** 对已发表的乳腺癌筛查指南进行质量评价, 为我国乳腺癌筛查指南的制订提供借鉴和参考。**方法** 检索 PubMed、Embase、Cochrane Library、Web of Science、中国知网、中国生物医学文献服务系统、维普网和万方数据知识服务平台中乳腺癌筛查指南, 检索时间为建库至 2020 年 8 月。由 2 名研究人员独立筛选文献和提取资料, 采用开发指南研究和评估工具 II (AGREE II) 和国际实践指南报告标准 (RIGHT) 对纳入的指南进行质量评价。**结果** 共纳入 15 部乳腺癌筛查指南, 其中 7 部由美国发布, 发表年份主要集中在 2015–2019 年, 11 部指南有更新版本。AGREE II 评价结果显示, 严谨性 (47.0%±22.1%) 和应用性 (44.0%±15.1%) 得分较低, RIGHT 评价结果显示, 评审和质量保证 (46.7%±39.9%) 及资金资助和利益冲突 (41.7%±24.4%) 报告较差。根据 AGREE II 整体得分推荐使用的指南共有 6 部, 修改后推荐使用的指南有 9 部; RIGHT 报告水平良好的指南有 4 部, 报告水平一般的指南共 11 部。综合质量最好的是美国国立综合癌症网络 2018 年发布的指南 (AGREE II : 83.3%, RIGHT : 80.0%) 和美国癌症学会 2015 年发布的指南 (AGREE II : 83.3%, RIGHT : 85.7%)。**结论** 乳腺癌筛查指南的质量以中等质量为主, 应加强对指南的制订过程和质量控制等内容的关注。

**【关键词】** 乳腺肿瘤; 筛查; 指南; 共识; 质量评价

**基金项目:** 北京市优秀人才培养资助-青年拔尖团队项目 (2017000021223TD05)

### Quality assessment of global breast cancer screening guidelines

Li Jiang<sup>1</sup>, Yang Kelu<sup>2</sup>, Cai Yitong<sup>2</sup>, Tian Jinhui<sup>2</sup>, Zheng Yadi<sup>1</sup>, Wen Yan<sup>1</sup>, Yang Zhuoyu<sup>1</sup>, Li Ni<sup>1</sup>,  
Chen Wanqing<sup>1</sup>, He Jie<sup>3</sup>

<sup>1</sup>Office of Cancer Screening, National Cancer Center/National Clinical Research Center for Cancer/  
Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing  
100021, China; <sup>2</sup>Evidence-based Medicine Center, Lanzhou University, Lanzhou 730000, China;  
<sup>3</sup>Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for  
Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College,  
Beijing 100021, China

Li Jiang and Yang Kelu contributed equally to the article

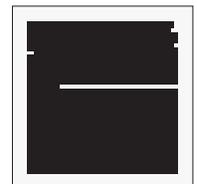
Corresponding author: Li Ni, Email: nil@cicams.ac.cn

**【Abstract】Objective** To evaluate the quality of the published breast cancer screening guidelines to provide a reference for domestic studies in the future. **Methods** PubMed, Embase, Cochrane Library, Web of Science, SinoMed, China National Knowledge Infrastructure, VIP, and Wanfang Data were searched to identify breast cancer screening guidelines on until August 2020.

DOI: 10.3760/cma.j.cn112338-20200806-01032

收稿日期 2020-08-06 本文编辑 万玉立

引用本文: 李江, 杨珂璐, 蔡依彤, 等. 全球乳腺癌筛查指南质量评价[J]. 中华流行病学杂志, 2021, 42(2): 219-226. DOI: 10.3760/cma.j.cn112338-20200806-01032.



Two reviewers screened literature and extracted data independently. The Appraisal of Guidelines for Research & Evaluation II (AGREE II) and Reporting Items for Practice Guidelines in Healthcare (RIGHT) tools were used to evaluate the quality of the included guidelines. **Results** A total of 15 breast cancer screening guidelines were included, of which seven were published in the United States, with publication years focusing on 2015 to 2019, and 11 guidelines had updated versions. "Rigour of development" (47.0%±22.1%) and "Applicability" (44.0%±15.1%) of AGREE II scored lower than other domains. "Review and quality assurance" (46.7%±39.9%) and "Funding, declaration, and management of interests" (41.7%±24.4%) of RIGHT were reported poorer than others. There were six guidelines recommended and another nine recommended with modifications based on the overall AGREE II score. There were four guidelines with a good level, and another 11 were with a moderate level of RIGHT. The National Comprehensive Cancer Network published the best overall quality guidelines in 2018 (AGREE II : 83.3%, RIGHT: 80.0%) and by the American Cancer Society in 2015 (AGREE II : 83.3%, RIGHT: 85.7%). **Conclusion** The quality of breast cancer screening guidelines was predominantly of moderate quality, and greater attention should be paid to the guideline development process and quality control of the guidelines.

**【 Key words 】** Breast neoplasms; Screening; Guidelines; Consensus; Quality assessment

**Fund program:** Training Programme Foundation for the Talents in Beijing City Grant (2017000021223TD05)

2018年全球乳腺癌的新发病例高达2 088 849例, 占有新发癌症病例的11.6%, 位居全球第二<sup>[1]</sup>。在美国, 乳腺癌是最常见的癌症和死亡率第二的癌症, 2012-2016年美国乳腺癌发病率每年都略有上升, 2020年美国预估新发病例为279 100例, 预计死亡病例为42 690例<sup>[2-3]</sup>。同样, 乳腺癌在我国女性恶性肿瘤中发病率也是第一, 每年发病人数约30.4万人, 是女性健康的巨大威胁<sup>[4]</sup>。乳腺癌筛查尚无统一的标准。高质量的乳腺癌筛查指南是缩小最佳证据与实践差距的重要工具<sup>[5-6]</sup>。近年临床实践指南数量不断增长, 但参与指南制订的团队不同、证据基础不同、制定流程不同, 导致指南质量参差不齐, 给临床实践造成一定的困扰<sup>[7]</sup>。目前已有多个国家和机构发布了乳腺癌筛查指南/共识, 指南的质量有待评价。本研究在已发表文章《乳腺癌筛查指南方法学质量和报告质量的系统评价》基础上<sup>[8]</sup>, 采用开发指南研究和评估工具II (The Appraisal of Guidelines for Research & Evaluation II, AGREE II)<sup>[9]</sup>和国际实践指南报告标准 (Reporting Items for Practice Guidelines in Healthcare, RIGHT)<sup>[10]</sup>评价相关指南的质量, 以期为我国制订乳腺癌筛查指南提供借鉴, 并为我国乳腺癌筛查工作提供参考。

## 资料与方法

1. 文献检索: 检索PubMed、Embase、Cochrane Library、Web of Science、中国知网、中国生物医学文

献服务系统、维普网和万方数据知识服务平台和相关部门网站, 包括医脉通网站 (<http://www.medlive.cn/>)、中国临床肿瘤学会 (<http://www.cSCO.org.cn/>)、美国国立综合癌症网络 (National Comprehensive Cancer Network, <https://www.cancer.gov/>)、美国临床肿瘤学会 (American Society of Clinical Oncology, <https://www.asco.org/>)、美国预防服务工作组 (U. S. Preventive Services Task Force, USPSTF, <https://www.uspreventiveservicestaskforce.org/uspstf/>)、英国国家临床诊治服务优化研究院 (National Institute of Clinical Excellence, <https://www.nice.org.uk/>) 和国际指南协作网 (Guideline International Network, <https://g-i-n.net/>) 等, 检索时间限定为建库至2020年8月。中文检索词包括乳腺癌、乳腺肿瘤、乳腺瘤、乳癌、乳房癌、指南、建议、筛查、早期诊断等, 英文检索词包括breast cancer、breast neoplasms、screening、early detection、guideline、recommendation等。同时追溯纳入指南的参考文献以补充检索结果。

2. 纳入与排除标准: 纳入公开发表的乳腺癌筛查指南, 同一部指南只纳入最新版本。排除指南摘要、指南翻译、指南解读、指南草案、会议摘要、同时包括乳腺癌及其他疾病的指南, 以及乳腺癌筛查内容非独立呈现的指南。

3. 文献筛选与资料提取: 2名研究人员根据纳入排除标准分别进行文献筛选, 再使用预先设计的信息提取表独立提取资料, 并进行核对, 如有分歧由第3名研究人员决定。提取内容包括指南名称、发表年份、国家、版本、资金、分级系统、证据等级、

推荐意见等级等信息。

4. 质量评价: 评价者分别采用 AGREE II<sup>[9]</sup> 和 RIGHT<sup>[10]</sup> 对纳入指南进行独立评价。AGREE II 工具是评估指南开发方法严谨性和透明度的工具, 共包含 6 个领域(范围 and 目的、参与人员、严谨性、清晰性、应用性、独立性) 23 个主要条目以及 2 个总体评估条目, 每个领域针对指南质量评价的一个特定问题。所有条目均按 1~7 分进行打分, 其中 1 分代表很不同意, 7 分代表非常同意, 2~6 分根据报道情况进行打分。评价完成后分别计算各领域的质量分值, 各领域分值=(实际得分-最小可能得分)/(最大可能得分-最小可能得分), 其中最大可能得分=7(非常同意)×条目数×评价者人数, 最小可能得分=1(很不同意)×条目数×评价者人数。由 4 位评价者独立进行评价, 在完成 23 个主要条目的评价之后, 对指南的质量做出一个综合判断, 并且根据得分回答是否推荐适用该指南, 当整体得分>60% 时推荐使用该指南, 整体得分为 30%~60% 时建议修改后推荐使用该指南, 整体得分<30% 时不推荐使用该指南<sup>[11]</sup>。RIGHT 工具是帮助指南制定者规范报告指南, 也为期刊编辑和同行评审提供了指南评审标准, 并帮助医务工作者更好地理解并执行指南。RIGHT 工具共包含 7 个领域(基本信息、背景、证据、推荐意见、评审和质量保证、资金资助和利益冲突、其他方面) 22 个条目, 所有条目以“是”和“否”进行评价, “是”指该条目的相关信息部分或完整呈现, “否”指该条目的相关信息完全缺失。由 2 名评价者独立进行评价并进行核对, 若有不一致通过讨论解决。总报告率≥70% 认为指南报告水平良好, 40%~70% 认为指南报告水平中等, ≤40% 认为指南报告水平较差。

5. 统计学分析: 采用 Excel 2018 软件进行资料整理和分析, AGREE II 评分和 RIGHT 的报告率采用百分比(%)、 $\bar{x} \pm s$  来描述, 以图表形式呈现结果, 并对相关的基本信息(指南的版本和国家)进行亚组分析, 使用 Stata 14 软件进行统计学比较, 采用平均差(mean difference, MD)、OR 值及其 95%CI 来描述。

为保证评价者之间的一致性, 对评价者评分结果进行组内相关系数(intra-class correlation coefficient, ICC) 检验<sup>[12]</sup>。若 ICC 值为 0.10~0.20, 说明评价者间一致性差; 0.21~0.40 说明一

致性较差; 0.41~0.60 说明一致性中等; 0.61~0.80 说明一致性良好; 0.81~1.00 说明一致性非常好<sup>[13]</sup>。

## 结 果

1. 文献筛选: 初步检索共获得 4 494 篇文献, 排除重复文献 1 465 篇和不符合纳入排除标准文献 2 872 篇后, 对 157 篇文献进行全文阅读, 最终纳入 15 部指南(图 1)。

2. 基本信息: 纳入的 15 部指南中美国发布的指南共有 7 部<sup>[14-20]</sup>、中国 3 部<sup>[21-23]</sup>、加拿大 2 部<sup>[24-25]</sup>、国际癌症研究署(International Agency for Research on Cancer, IARC)<sup>[26]</sup>、德国<sup>[27]</sup>和日本<sup>[28]</sup>各 1 部。发表时间主要集中在 2015~2019 年, 有更新版本的指南有 11 部<sup>[14-15, 17, 20-21, 23-29]</sup>, 有 4 部指南未明确报告资金来源<sup>[15, 18, 21, 23]</sup>, 6 部指南未报告证据等级<sup>[16, 18, 21-24]</sup>, 7 部指南未报告推荐意见等级<sup>[14, 16, 18, 21, 23-24, 26]</sup>(表 1)。

共 10 部指南使用了分级系统, 每部指南的分级系统不尽相同, 其中使用较多的分级系统是证据推荐分级的评估、制订与评价(The Grading of Recommendations Assessment, Development and Evaluation, GRADE)系统和基于 GRADE 系统改编的分级系统(3 部)<sup>[17, 19, 25]</sup>, 其次为 USPSTF 的分级系统和基于 USPSTF 改变的分级系统(2 部)<sup>[15, 20]</sup>。IARC<sup>[26]</sup>和美国国立综合癌症网络<sup>[14]</sup>的指南仅有证据等级而没有清楚的推荐意见等级, 而中国抗癌协会(China Anti-Cancer Association, CACA)的指南仅有推荐意见等级<sup>[22]</sup>(表 2)。

3. 质量评价:

(1) AGREE II: 6 个领域的 ICC 值为 0.86、0.96、0.98、0.87、0.97、0.96, 总 ICC=0.99, 95%CI: 0.977~0.991, 评价者之间的一致性很好。纳入指南的 AGREE II 整体得分最低得分为 33.3%, 最高得分为

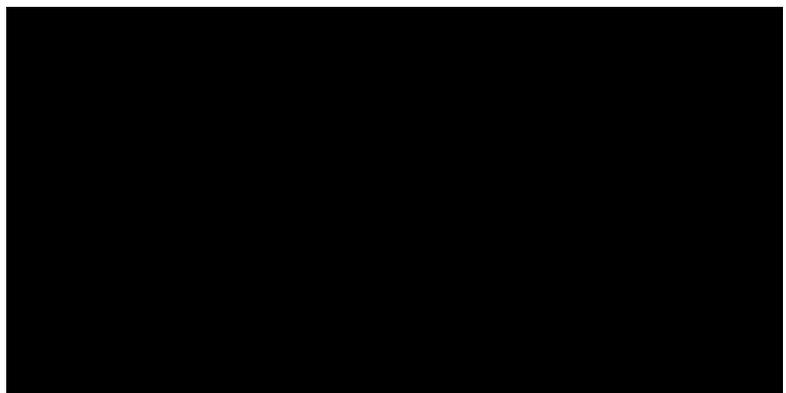


图 1 乳腺癌筛查文献筛选流程

表 1 乳腺癌筛查指南基本信息

文献	发布机构	年份	国家	版本更新	基金明确	证据等级明确	推荐意见等级明确
[14]	美国国立综合癌症网络(National Comprehensive Cancer Network)	2018	美国	是	是	是	未报告
[28]	日本国家癌症中心(National Cancer Center)	2016	日本	是	是	是	是
[25]	加拿大预防保健工作组(Canadian Task Force on Preventive Health Care)	2018	加拿大	是	是	是	是
[20]	美国预防服务工作组(U.S. Preventive Services Task Force)	2016	美国	是	是	是	是
[19]	美国放射学会(American College of Radiology)	2017	美国	否	是	是	是
[23]	中国香港地区癌症预防及普查专家工作组(Cancer Expert Working Group on Cancer Prevention and Screening, CEWG)	2018	中国	是	未报告	未报告	未报告
[24]	加拿大安大略癌症治疗中心(Cancer Care Ontario)	2008	加拿大	是	是	未报告	未报告
[18]	美国癌症学会(American Cancer Society)	2007	美国	否	未报告	未报告	未报告
[17]	美国癌症学会(American Cancer Society)	2015	美国	是	是	是	是
[26]	国际癌症研究署(International Agency for Research on Cancer)	2015	国际组织	是	是	是	未报告
[16]	美国医师协会(American college of physicians)	2007	美国	否	是	未报告	未报告
[27]	德国医学科学联合会(Association of the Scientific Medical Societies)	2009	德国	是	是	是	是
[22]	中国抗癌协会(China Anti-Cancer Association)	2019	中国	否	是	未报告	是
[15]	美国妇产科医师学会(American Congress of Obstetricians and Gynecologists)	2017	美国	是	未报告	是	是
[21]	中国抗癌协会(China Anti-Cancer Association)	2019	中国	是	未报告	未报告	未报告

表 2 乳腺癌筛查指南分级

文献	分级系统	证据等级	推荐意见等级
[14]	NCCN <sup>a</sup>	1, 2A, 2B, 3	
[28]	改编 SIGN <sup>b</sup>	1++, 1+, 1-, 2++, 2+, 2-, 3, 4	A, B, C, D, I
[17, 25]	GRADE <sup>c</sup>	高, 中, 低, 极低	强推荐, 弱推荐
[20]	USPSTF <sup>d</sup>	高, 中, 低	A, B, C, D, 专家意见
[19]	改编 GRADE <sup>c</sup>	1, 2, 3, 4, 高质量 Meta 分析, 低质量 Meta 分析	强推荐, 弱推荐, 有条件的推荐, 专家共识, 专家意见
[26]	IARC <sup>e</sup>	充分证据, 有限证据, 不充分证据	-
[27]	AWMF <sup>f</sup>	1a, 1b, 1c, 2a, 2b, 2c, 3a, 3b, 4, 5	A, B, C
[22]	CACA <sup>g</sup>		A, B, C
[15]	改编 USPSTF <sup>d</sup>	I, II-1, II-2, II-3, III	A, B, C

注:<sup>a</sup>NCCN:美国国立综合癌症网络;<sup>b</sup>SIGN:苏格兰校际指南网;<sup>c</sup>GRADE:证据推荐分级的评估、制订与评价;<sup>d</sup>USPSTF:美国预防服务工作组;<sup>e</sup>IARC:国际癌症研究署;<sup>f</sup>AWMF:德国医学科学联合会;<sup>g</sup>CACA:中国抗癌协会

83.3%(表3)。应用性最高得分为75.0%,最低得分为22.9%,平均得分为44.0%±15.1%,是所有领域中最低平均分;清晰性最高得分为97.2%,最低得分为58.3%,平均得分最高(77.9%±10.7%);整体得分中>60%共有6部,分别为美国国立综合癌症网络(2018年)<sup>[14]</sup>、加拿大预防保健工作组(2018年)<sup>[25]</sup>、美国癌症学会(2015年)<sup>[17]</sup>、日本国家癌症中心(2016年)<sup>[28]</sup>、USPSTF(2016年)<sup>[20]</sup>、美国放射学会(2017年)<sup>[19]</sup>颁布的指南,其余指南得分在30%~60%之间。

(2) RIGHT:基本信息的平均报告率最高(83.3%±12.6%),4部指南<sup>[14-15,19,22]</sup>报告率为100.0%;资金资助和利益冲突的平均报告率最低(41.7%±24.4%),其中有2部指南未报告此领域内容<sup>[20-21]</sup>;评审和质量保证的平均报告率(46.7%±39.9%)仅次于资金资助和利益冲突。总报告率≥

70%的指南共有4部,分别是美国国立综合癌症网络(2018年)<sup>[14]</sup>、加拿大预防保健工作组(2018年)<sup>[25]</sup>、美国癌症学会(2015年)<sup>[17]</sup>、IARC(2015年)<sup>[26]</sup>,其余指南的报告率为40%~70%(表4)。

(3)总体质量:其中有2部指南的AGREE II评分和RIGHT报告率均较好,分别为美国癌症学会(2015年)<sup>[17]</sup>和美国国立综合癌症网络(2018年)<sup>[14]</sup>制订的(图2)。

(4)亚组分析:根据指南国家所在地区和是否有最新版本进行亚组分析,美洲地区的指南和欧洲地区的相比RIGHT各领域报告率差异无统计学意义,总报告率美洲地区略逊于欧洲地区( $OR=0.5$ ,  $95\%CI:0.3\sim1.0$ ),在AGREE II的严谨性( $MD=-0.2$ ,  $95\%CI:-0.3\sim-0.1$ )和独立性( $MD=-0.3$ ,  $95\%CI:-0.5\sim-0.1$ )美洲地区指南也略逊于欧洲地区指南,

表 3 乳腺癌筛查指南 AGREE II 各领域评分(%)

文献	领域						整体得分
	范围和目的	参与人员	严谨性	清晰性	应用性	独立性	
[14]	95.8	69.4	70.3	80.6	66.7	87.5	83.3
[28]	77.8	41.7	61.5	91.7	25.0	75.0	75.0
[25]	75.0	86.1	61.5	97.2	66.7	50.0	83.3
[20]	61.1	66.7	30.2	83.3	45.8	83.3	75.0
[19]	61.1	77.8	52.1	77.8	41.7	75.0	66.7
[23]	61.1	44.4	19.8	58.3	75.0	41.7	58.3
[24]	77.8	2.8	71.9	75.0	43.8	33.3	58.3
[18]	52.8	66.7	35.4	66.7	37.5	0.0	33.3
[17]	75.0	75.0	65.6	86.1	43.8	58.3	83.3
[26]	80.6	47.2	72.9	68.1	35.4	77.1	50.0
[16]	83.3	55.6	52.1	80.6	29.2	41.7	66.7
[27]	84.7	50.0	69.3	66.7	47.9	77.1	50.0
[22]	75.0	36.1	8.3	69.4	22.9	16.7	16.7
[15]	80.6	63.9	22.9	88.9	37.5	12.5	58.3
[21]	80.6	29.2	26.0	77.8	41.7	79.2	58.3
平均得分( $\bar{x}\pm s$ )	74.8±11.3	54.2±21.6	47.9±22.1	77.9±10.7	44.0±15.1	53.9±28.5	61.1±18.8

表 4 乳腺癌筛查指南 RIGHT 各领域报告率(%)

文献	领域							总报告率
	基本信息	背景	证据	推荐意见	评审和质量保证	资金资助和利益冲突	其他方面	
[14]	100.0	87.5	100.0	85.7	50.0	50.0	33.3	80.0
[28]	83.3	50.0	60.0	57.1	100.0	50.0	66.7	62.9
[25]	66.7	75.0	100.0	100.0	0.0	50.0	66.7	74.3
[20]	66.7	75.0	60.0	57.1	50.0	75.0	33.3	62.9
[19]	83.3	62.5	40.0	42.9	50.0	50.0	33.3	54.3
[23]	83.3	50.0	0.0	42.9	0.0	25.0	33.3	40.0
[24]	66.7	75.0	60.0	57.1	100.0	50.0	66.7	65.7
[18]	83.3	62.5	0.0	42.9	50.0	0.0	66.7	45.7
[17]	100.0	75.0	80.0	100.0	100.0	75.0	66.7	85.7
[26]	100.0	87.5	60.0	85.7	100.0	75.0	100.0	85.7
[16]	83.3	62.5	40.0	28.6	50.0	50.0	33.3	51.4
[27]	100.0	75.0	80.0	57.1	0.0	25.0	100.0	68.6
[22]	83.3	75.0	20.0	42.9	0.0	25.0	33.3	48.6
[15]	66.7	75.0	60.0	57.1	50.0	25.0	66.7	60.0
[21]	83.3	75.0	20.0	42.9	0.0	0.0	66.7	48.6
平均报告率( $\bar{x}\pm s$ )	83.3±12.6	70.8±11.2	52.0±31.9	60.0±22.4	46.7±39.9	41.7±24.4	57.8±23.5	62.3±14.5

而清晰性( $MD=0.1, 95\%CI: 0.1\sim 0.2$ )和整体得分( $MD=0.2, 95\%CI: 0.1\sim 0.3$ )美洲地区优于欧洲地区;美洲地区与亚洲地区的指南相比在 AGREE II 的参与人员( $MD=0.3, 95\%CI: 0.1\sim 0.4$ )和 RIGHT 的证据( $OR=4.5, 95\%CI: 1.4\sim 14.6$ )方面优于亚洲地区, AGREE II 整体得分差异无统计学意义,而 RIGHT 的总报告率美洲地区优于亚洲地区( $OR=1.8, 95\%CI: 1.2\sim 2.7$ );亚洲地区与欧洲地区相比,在 AGREE II 的参与人员( $MD=-0.1, 95\%CI: -0.2\sim 0.0$ )和严谨性( $MD=-0.4, 95\%CI: -0.7\sim -0.2$ ),以及 RIGHT 的证据( $OR=0.1, 95\%CI: 0.0\sim 0.8$ )和总报告率( $OR=0.3, 95\%CI: 0.2\sim 0.6$ )方面均不如欧洲地区指南;有更新版的指南与无更新版的指南相比,在

AGREE II 评分的应用性( $MD=0.2, 95\%CI: 0.0\sim 0.3$ )以及在 RIGHT 的证据( $OR=4.9, 95\%CI: 1.5\sim 15.3$ )、推荐意见( $OR=3.2, 95\%CI: 1.3\sim 7.9$ )和总报告率( $OR=2.0, 95\%CI: 1.4\sim 3.0$ )方面有更新版的指南报告较好(表 5)。

## 讨 论

癌症的三级预防包括病因预防、发病学预防和康复预防,其中发病学预防的主要措施为普查,通过早发现、早诊断和早治疗有效降低死亡率并提高患者预后,也可降低社会经济负担<sup>[30]</sup>。通过分析和评价目前已发表的乳腺癌筛查指南,可以为临床工

注:橙色代表 AGREE II 评为推荐使用的指南,蓝色代表 AGREE II 评为修改后推荐使用的指南,气泡大小为指南报告 RIGHT 的条目数,红圈为最优指南;NCCN:美国国立综合癌症网络;NCC:日本国家癌症中心;CTFPHC:加拿大预防保健工作组;USPSTF:美国预防服务工作组;ACR:美国放射学会;CEWG:中国香港地区癌症预防及普查专家工作组;CCO:加拿大安大略癌症治疗中心;ACS:美国癌症学会;IARC:国际癌症研究署;ACP:美国医师协会;AWMF:德国医学科学联合会;CACA:中国抗癌协会;ACOG:美国妇产科医师学会

图 2 乳腺癌筛查指南总体质量气泡图

作者选择合适的指南提供参考和依据。

本研究共纳入 15 部乳腺癌筛查指南,其中多数指南由美国制订,且发表时间主要在 2014 年以后。不同机构制订的指南质量多有不同,但不同地区间和不同版本间的差异较小。通过对目前已发表的乳腺癌筛查指南的质量进行评估发现,多数协会/学会在发布关于乳腺癌筛查指南后都会进行更新,例如美国国立综合癌症网络每年都会对乳腺癌筛查指南进行更新,并且同一年中也会有多版本的

指南发布和更新,同样 USPSTF 自 2010 年起每年都会对其发布的乳腺癌筛查指南进行更新。但美国癌症学会 2007 年单独发布了关于使用磁共振成像(MRI)作为钼靶的辅助筛查的指南后,再未对其进行更新,而美国医师协会在 2008 年发布的乳腺癌筛查后,2019 年发布了中风险女性乳腺癌筛查的指南声明<sup>[31]</sup>,由于不符合纳入标准故未纳入本研究。有更新版本的指南和无更新版本的指南相比,报告质量无明显差异,在 AGREE II 评价的应用性领域中有更新版本的指南得分较高,说明有更新版本的指南更加注重指南在实际临床工作中的应用,例如提供指南应用的工具、应用时可能遇到的障碍、如何评判指南实施的标准等,从而可以在实际工作中更好地实施指南的内容<sup>[9,32]</sup>。

虽然目前已有许多证据等级和推荐意见等级的分级系统,但仍有指南未能明确报告所使用的分级系统或未对证据和推荐意见进行等级的划分。此外,不同分级系统中等级的级数和含义不尽相同,对临床使用造成一定的困扰,甚至同一内容的推荐意见在不同指南中的分级都会有所不同,例如日本国家癌症中心(National Cancer Center, NCC)2016 年的指南<sup>[28]</sup>和 CACA 2019 年的指南<sup>[21]</sup>都不建议单独采用乳腺触诊进行乳腺癌筛查,但 NCC 给出的推荐意见等级为 I 级(没有足够证据证明该筛查方法与降低死亡率相关,不推荐作为普查方法),而 CACA 的推荐等级为 B 级(基于有限或结论相对一致的研究证据,中等推荐)。NCC 的推荐意见基于 6 个随机对照试验和 4 个病例对照研究,而

表 5 乳腺癌筛查指南质量评价亚组分析

评分/报告率	美洲地区 vs. 欧洲地区	美洲地区 vs. 亚洲地区	亚洲地区 vs. 欧洲地区	有更新版 vs. 无更新版
AGREE II [MD 值(95%CI)]				
范围和目的	-0.1(-0.2~0.1)	0.0(-0.1~0.1)	0.0(-0.2~0.0)	0.1(-0.1~0.2)
参与人员	0.1(0.0~0.3)	0.3(0.1~0.4)	-0.1(-0.2~0.0)	-0.1(-0.3~0.2)
严谨性	-0.2(-0.3~-0.1)	0.2(0.0~0.5)	-0.4(-0.7~-0.2)	0.2(-0.1~0.4)
清晰性	0.1(0.1~0.2)	0.1(-0.1~0.2)	0.1(-0.1~0.2)	0.1(0.0~0.2)
应用性	0.0(-0.1~0.2)	0.1(-0.2~0.3)	0.0(-0.3~0.3)	0.2(0.0~0.3)
独立性	-0.3(-0.5~-0.1)	0.0(-0.4~0.3)	-0.2(-0.5~0.1)	0.3(-0.1~0.6)
整体得分	0.2(0.1~0.3)	0.2(-0.1~0.4)	0.0(-0.2~0.3)	0.2(-0.1~0.4)
RIGHT [OR 值(95%CI)]				
基本信息	0.2(0.0~2.8)	0.8(0.2~2.8)	0.2(0.0~3.7)	1.0(0.3~3.5)
背景	0.6(0.2~2.3)	1.6(0.7~3.8)	0.4(0.1~1.6)	1.4(0.6~3.3)
证据	0.6(0.2~2.8)	4.5(1.4~14.6)	0.1(0.0~0.8)	4.9(1.5~15.3)
推荐意见	0.7(0.2~2.5)	2.0(0.8~5.0)	0.4(0.1~1.4)	3.2(1.3~7.9)
评审和质量保证	1.3(0.1~10.9)	3.8(0.6~23.9)	0.3(0.0~4.2)	1.7(0.3~8.7)
资金资助和利益冲突	0.9(0.2~4.1)	2.7(0.7~9.9)	0.3(0.1~2.0)	1.8(0.6~6.2)
其他方面	0.1(0.0~1.6)	1.1(0.3~4.2)	0.1(0.0~1.7)	2.5(0.6~9.4)
总报告率	0.5(0.3~1.0)	1.8(1.2~2.7)	0.3(0.2~0.6)	2.0(1.4~3.0)

CACA 的推荐意见的证据基础不详。推荐意见等级的不同除了与证据基础有关外,还与参与指南制定的成员组成有很大关系,日本和中国的指南制订成员主要为本国人员,并且指南的目标人群为本国人,所以在推荐意见等级评定时更多的是结合本国实际国情进行评定,所以相同的推荐意见的等级有所不同。我们建议提供更多的报告形成推荐意见及评定等级的过程和结果,给予足够多的证据和参考以供指南使用者选择。

AGREE II 评定中多数指南为修改后推荐使用,其中严谨性和应用性的平均分最低。严谨性分别从证据的准备过程、推荐意见的形成过程、指南全文的质量把控和后续更新等角度进行评估,多数指南只着重于推荐意见部分的呈现,而对于指南形成过程描述较少。对于指南制作过程的详细描述,可以给指南阅读者更加明确清晰地呈现每一个推荐意见的形成过程和证据基础,使其在使用指南时有足够的信息去判断推荐意见的合理性以及是否适合临床情景。指南的最终目的是服务于临床工作,所以指南应该具有足够的应用性,指南应详细描述推荐意见应用过程中促进和阻碍因素;还应提供配套的总结文件、快速参考指南、培训工具、患者说明等以帮助指南实施和推广;在应用推荐意见时应考虑潜在的资源投入的影响,例如成本效益、成本效用、预算等信息;在指南的主要推荐意见中还应包括可以监控和审计的标准,以促进指南的实用性<sup>[9,32]</sup>。不同地区所制订的指南差异不大,美国和加拿大发表的指南 AGREE II 整体得分略优于欧洲地区,可能与纳入欧洲地区国家的指南较少有关,结果可能存在一定偏倚。

RIGHT 判定中多数指南为中等,其中评审和质量保证及资金资助和利益冲突平均报告率最低。指南的评审和质量保证是指南获得认可的重要步骤,外部评审促进了指南的准确性、清晰性和实用性,而质量保证侧重于指南制定的方法和程序得到官方的认可、批准和支持。资金资助和利益冲突是指南制定过程中一个重要的潜在偏倚来源,指南应该充分报告何时收到何种资助及其影响,也同时应报告参与指南制订成员是否有利益冲突,以及如何管理这些利益冲突,从而提高指南的可信度<sup>[10,33]</sup>。从不同地区制订的指南的报告规范来看,欧洲地区的指南报告最为规范,而且有更新版本的指南的报告情况也优于无更新版本的指南。

本研究共纳入了 3 部中国制订的指南,遗憾的

是这 3 部指南的 AGREE II 评分和 RIGHT 报告率都不高,普遍存在报告内容过于简单、证据基础不详、指南形成过程报告不全或未报告、资金资助和利益冲突声明过于简单或未报告等问题。

由于不同机构制订指南流程和方法学上不同,以及基于的证据基础不同,所以对于乳腺癌筛查的推荐意见有所不同。对于乳腺癌筛查的起始年龄各个指南的推荐略有不同,加拿大安大略癌症治疗中心(Cancer Care Ontario, CCO)的指南基于专家意见推荐从 30 岁起女性就进行每年规律筛查<sup>[24]</sup>,而 CACA 的指南建议从 45 岁开始进行规律筛查,但证据基础不详<sup>[22]</sup>。同样,对于停止筛查的年龄各个指南也有所出入,CCO 建议 69 岁前都应进行规律筛查,与 CACA 的推荐意见相同,且 CACA 还建议对于 >69 岁的一般风险女性,身体健康且预期寿命 > 10 年应接受机会性筛查,而美国妇产科医师学会(American Congress of Obstetricians and Gynecologists)建议应截止在 75 岁,对于 >75 岁的妇女也应根据健康状况、预期寿命和本人意愿综合做出决定<sup>[15,22,24]</sup>。钼靶筛查因其能有效降低乳腺癌死亡率,是各个指南首推的筛查方法,部分指南还对钼靶筛查的时间间隔也进行了推荐,认为一般风险女性应 1~2 年进行一次钼靶筛查<sup>[15,22]</sup>,但对于不同的钼靶类型、拍摄时的体位选择和读片规范等内容的推荐意见较少<sup>[21]</sup>。对于高危女性,多部指南推荐可以采用 MRI 联合钼靶作为主要的筛查方法,当无法使用或不耐受 MRI 时可以采用超声作为代替<sup>[18,22]</sup>。超声和乳腺临床检查不建议作为单独的筛查工具,有指南推荐对致密性乳腺女性可以在钼靶筛查阴性后补充超声筛查,也可以将临床检查和钼靶相结合用以评估高危女性<sup>[15,18,21-22]</sup>。乳腺自我检查因其可能造成的假阳性结果有危害的风险被多个指南所不推荐<sup>[15,21]</sup>。

本研究此次纳入的指南为中文和英文,对于其他语言的指南暂未纳入,且未对推荐意见进行深入分析和总结。在之后的研究中,我们会尽可能纳入乳腺癌筛查领域的所有指南,除了指南质量的评价外,对于指南推荐意见的一致性和证据基础也会进行进一步探索。

乳腺癌筛查指南的总体质量中等,应加强指南的严谨性、应用性、评审和质量保证、资金资助和利益冲突等方面的内容。我国现有的乳腺癌筛查指南质量不高,应遵循科学的指南制定方法和规范的报告方法制订出符合我国国情的乳腺癌筛查指南,

从而有效指导乳腺癌筛查工作,提高乳腺癌的早期诊断,降低乳腺癌死亡率。

利益冲突 所有作者均声明不存在利益冲突

### 参 考 文 献

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68(6): 394-424. DOI: 10.3322/caac.21492.
- [2] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020[J]. *CA Cancer J Clin*, 2020, 70(1): 7-30. DOI: 10.3322/caac.21590.
- [3] DeSantis CE, Ma JM, Gaudet MM, et al. Breast cancer statistics, 2019[J]. *CA Cancer J Clin*, 2019, 69(6): 438-451. DOI: 10.3322/caac.21583.
- [4] Chen WQ, Zheng RS, Baade PD, et al. Cancer statistics in China, 2015[J]. *CA Cancer J Clin*, 2016, 66(2): 115-132. DOI: 10.3322/caac.21338.
- [5] 王吉耀. 用临床实践指南的评价标准指导指南的临床应用[J]. *上海医学*, 2018, 41(6): 327-328. Wang JY. Use of AGREE II to guide the clinical application of clinical practice guidelines[J]. *Shanghai Med J*, 2018, 41(6): 327-328.
- [6] Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on[J]. *Lancet*, 2017, 390(10092): 415-423. DOI: 10.1016/S0140-6736(16)31592-6.
- [7] 陈耀龙, 杨克虎. 正确理解、制订和使用临床实践指南[J]. *协和医学杂志*, 2018, 9(4): 367-373. DOI: 10.3969/j.issn.1674-9081.2018.04.015. Chen YL, Yang KH. How to correctly understand, develop, and apply clinical practice guidelines[J]. *Med J Peking Union Med Coll Hosp*, 2018, 9(4): 367-373. DOI: 10.3969/j.issn.1674-9081.2018.04.015.
- [8] 李江, 唐威, 李霓, 等. 乳腺癌筛查指南方法学质量和报告质量的系统评价[J]. *中国循证医学杂志*, 2018, 18(6): 629-636. DOI: 10.7507/1672-2531.201803081. Li J, Tang W, Li N, et al. Methodological and reporting quality of breast cancer screening guidelines: a systematic review[J]. *Chin J Evid-Based Med*, 2018, 18(6): 629-636. DOI: 10.7507/1672-2531.201803081.
- [9] Brouwers MC, Kerkvliet K, Spithoff K, et al. The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines[J]. *BMJ*, 2016, 352: i1152. DOI: 10.1136/bmj.i1152.
- [10] Chen YL, Yang KH, Marušić A, et al. A reporting tool for practice guidelines in health care: The RIGHT statement[J]. *Ann Intern Med*, 2017, 166(2): 128-132. DOI: 10.7326/M16-1565.
- [11] Jiang M, Guan WJ, Fang ZF, et al. A critical review of the quality of cough clinical practice guidelines[J]. *Chest*, 2016, 150(4): 777-788. DOI: 10.1016/j.chest.2016.04.028.
- [12] Landis JR, Koch GG. The measurement of observer agreement for categorical data[J]. *Biometrics*, 1977, 33(1): 159-174. DOI: 10.2307/2529310.
- [13] Kramer MS, Feinstein AR. Clinical biostatistics: LIV. The biostatistics of concordance[J]. *Clin Pharmacol Ther*, 1981, 29(1): 111-123. DOI: 10.1038/clpt.1981.18.
- [14] Bevers TB, Helvie M, Bonaccio E, et al. NCCN clinical practice guidelines in Oncology: breast cancer screening and diagnosis[J]. *J Natl Compr Canc Netw*, 2018, 16(11): 1362-1389. DOI: 10.6004/jnccn.2018.0083.
- [15] American Congress of Obstetricians and Gynecologists. Practice bulletin number 179: breast cancer risk assessment and screening in average-risk women[J]. *Obstet Gynecol*, 2017, 130(1): e1-16. DOI: 10.1097/AOG.0000000000002158.
- [16] Qaseem A, Snow V, Sherif K, et al. Screening mammography for women 40 to 49 years of age: a clinical practice guideline from the American College of Physicians[J]. *Ann Intern Med*, 2007, 146(7): 511-515. DOI: 10.7326/0003-4819-146-7-200704030-00007.
- [17] Oeffinger KC, Fontham ETH, Etzioni R, et al. Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society[J]. *JAMA*, 2015, 314(15): 1599-1614. DOI: 10.1001/jama.2015.12783.
- [18] Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography[J]. *CA Cancer J Clin*, 2007, 57(2): 75-89. DOI: 10.3322/canjclin.57.2.75.
- [19] Mainiero MB, Moy L, Baron P, et al. ACR Appropriateness Criteria® Breast Cancer Screening[J]. *J Am Coll Radiol*, 2017, 14(11 Suppl 1): S383-390. DOI: 10.1016/j.jacr.2017.08.044.
- [20] Siu AL. Screening for breast cancer: U.S. Preventive services task force recommendation statement[J]. *Ann Intern Med*, 2016, 164(4): 279-296. DOI: 10.7326/M15-2886.
- [21] 中国抗癌协会乳腺癌专业委员会. 中国抗癌协会乳腺癌诊治指南与规范(2019年版)[J]. *中国癌症杂志*, 2019, 29(8): 609-680. DOI: 10.19401/j.cnki.1007-3639.2019.08.009. Breast Cancer Professional Committee of China Anti-Cancer Association. China anti-cancer society breast cancer diagnosis and treatment guidelines and protocols[J]. *Chin Oncol*, 2019, 29(8): 609-680. DOI: 10.19401/j.cnki.1007-3639.2019.08.009.
- [22] 黄育北. 中国女性乳腺癌筛查指南[J]. *中国肿瘤临床*, 2019, 46(9): 429-431. DOI: 10.3969/j.issn.1000-8179.2019.09.572. Huang YB. Guidelines for women's breast cancer screening in China[J]. *Chin J Clin Oncol*, 2019, 46(9): 429-431. DOI: 10.3969/j.issn.1000-8179.2019.09.572.
- [23] Lam TH, Wong KH, Chan KKL, et al. Recommendations on prevention and screening for breast cancer in Hong Kong[J]. *Hong Kong Med J*, 2018, 24(3): 298-306. DOI: 10.12809/hkmj177037.
- [24] Warner E, Messersmith H, Causer P, et al. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer[J]. *Ann Intern Med*, 2008, 148(9): 671-679. DOI: 10.7326/0003-4819-148-9-200805060-00007.
- [25] Klarenbach S, Sims-Jones N, Lewin G, et al. Recommendations on screening for breast cancer in women aged 40-74 years who are not at increased risk for breast cancer[J]. *CMAJ*, 2018, 190(49): E1441-1451. DOI: 10.1503/cmaj.180463.
- [26] International Agency for Research on Cancer. Breast-Cancer Screening—Viewpoint of the IARC Working Group [EB/OL]. [2020-08-28]. <https://publications.iarc.fr/>.
- [27] Albert US, Altland H, Duda V, et al. 2008 update of the guideline: early detection of breast cancer in Germany[J]. *J Cancer Res Clin Oncol*, 2009, 135(3): 339-354. DOI: 10.1007/s00432-008-0450-y.
- [28] Hamashima C, Hattori M, Honjo S, et al. The Japanese guidelines for breast cancer screening[J]. *Jpn J Clin Oncol*, 2016, 46(5): 482-492. DOI: 10.1093/jcco/hyw008.
- [29] Bevers TB, Anderson BO, Bonaccio E, et al. NCCN clinical practice guidelines in oncology: breast cancer screening and diagnosis[J]. *J Natl Compr Canc Netw*, 2009, 7(10): 1060-1096. DOI: 10.6004/jnccn.2009.0070.
- [30] 董超, 徐正丰. 乳腺癌三级预防的研究进展[J]. *长江大学学报: 自然科学版*, 2011, 8(4): 209-212. DOI: 10.3969/j.issn.1673-1409.2011.04.095. Dong C, Xu ZF. Research progress in tertiary prevention of breast cancer[J]. *J Yangtze University: Nat Sci Ed*, 2011, 8(4): 209-212. DOI: 10.3969/j.issn.1673-1409.2011.04.095.
- [31] Qaseem A, Lin JS, Mustafa RA, et al. Screening for breast cancer in average-risk women: a guidance statement from the American College of Physicians[J]. *Ann Intern Med*, 2019, 170(8): 547-560. DOI: 10.7326/M18-2147.
- [32] 谢利民, 王文岳. 《临床指南研究与评价系统 II》简介[J]. *中西医结合学报*, 2012, 10(2): 160-165. DOI: 10.3736/jcim20120206.
- [33] Xie LM, Wang WY. A brief introduction to appraisal of guidelines for research and Evaluation II [J]. *J Chin Integ Med*, 2012, 10(2): 160-165. DOI: 10.3736/jcim20120206. Chen YL, Yang KH, Marušić A, et al. RIGHT Explanation and Elaboration: guidance for reporting practice guidelines [EB/OL]. [2020-08-28]. <http://right-statement.org/uploads/e1c3506a2062dd4ccc7f36c0.pdf>.