

药品临床综合评价工作流程指南[△]

李正翔^{1,2*}, 段蓉^{1,2}, 史录文³, 田金徽⁴, 左笑丛⁵, 张玉⁶, 张伶俐⁷, 张俊华⁸, 郑华林⁹, 赵荣生¹⁰, 郭武栋¹¹, 缪丽燕¹², 翟所迪¹⁰(1.天津医科大学总医院药学中心,天津 300052;2.天津市药品使用监测和临床综合评价中心,天津 300052;3.北京大学药学院药事管理与临床药学系,北京 100091;4.兰州大学基础医学院循证医学研究所,兰州 730099;5.中南大学湘雅三医院,长沙 422099;6.华中科技大学同济医学院附属协和医院,武汉 430022;7.四川大学华西第二医院,成都 620020;8.天津中医药大学循证医学中心,天津 301617;9.中国国际科技促进会,北京 100029;10.北京大学第三医院药学部,北京 100083;11.国家卫生健康委卫生发展研究中心,北京 100044;12.苏州大学附属第一医院,苏州 215006)

中图分类号 R98 文献标志码 A 文章编号 1001-0408(2025)19-2353-13
DOI 10.6039/j.issn.1001-0408.2025.19.01



摘要 目的 为规范药品临床综合评价工作的主要流程和相关技术环节,提升综合评价证据的质量和转化应用价值提供指导和参考。方法 以世界卫生组织标准指南制定方法为依据,严格遵循美国国家科学院医学研究所关于指南的最新定义,并符合指南研究与评价工具Ⅱ的六大领域要求制定《药品临床综合评价工作流程指南》。该指南采用德尔菲法构建研究问题;应用循证医学研究方法建立研究证据,采用中国循证医学中心的证据分级体系进行证据分级评价,采用牛津大学循证医学中心制定的推荐强度分级标准确定推荐强度,并通过专家共识法形成推荐意见。结果与结论 《药品临床综合评价工作流程指南》涵盖主题遴选、评价实施、证据评价和结果应用转化4类研究问题。该指南的制定规范了药品临床综合评价的全流程技术环节,可有效指导该项工作的高质、高效开展,提升评价证据的规范产出和转化应用价值,为卫生决策的科学化和临床用药合理化提供高质量的证据支持。
关键词 药品临床综合评价;工作流程;指南;团体标准;循证医学

Guideline for the workflow of clinical comprehensive evaluation of drugs

LI Zhengxiang^{1,2}, DUAN Rong^{1,2}, SHI Luwen³, TIAN Jinhui⁴, ZUO Xiacong⁵, ZHANG Yu⁶, ZHANG Lingli⁷, ZHANG Junhua⁸, ZHENG Hualin⁹, ZHAO Rongsheng¹⁰, GUO Wudong¹¹, MIAO Liyan¹², ZHAI Suodi¹⁰
(1. Pharmacy Center, Tianjin Medical University General Hospital, Tianjin 300052, China; 2. Tianjin Center for Drug Use Monitoring and Clinical Comprehensive Evaluation, Tianjin 300052, China; 3. Dept. of Pharmacy Administration and Clinical Pharmacy, School of Pharmaceutical Sciences, Peking University, Beijing 100091, China; 4. Institute of Evidence-Based Medicine, School of Basic Medical Sciences, Lanzhou University, Lanzhou 730099, China; 5. The Third Xiangya Hospital of Central South University, Changsha 422099, China; 6. Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China; 7. West China Second University Hospital, Sichuan University, Chengdu 620020, China; 8. Evidence-based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China; 9. China International Association for Promotion of Science and Technology, Beijing 100029, China; 10. Dept. of Pharmacy, Peking University Third Hospital, Beijing 100083, China; 11. Health Development Research Center, National Health Commission, Beijing 100044, China; 12. The First Affiliated Hospital of Soochow University, Jiangsu Suzhou 215006, China)

ABSTRACT OBJECTIVE To standardize the main processes and related technical links of the clinical comprehensive evaluation of drugs, and provide guidance and reference for improving the quality of comprehensive evaluation evidence and its transformation and application value. **METHODS** The construction of *Guideline for the Workflow of Clinical Comprehensive Evaluation of Drugs*

△基金项目 天津市卫生健康委药品临床综合评价项目(No.津卫食药[2025]67号)

*第一作者 主任药师。研究方向:医院药学、药品临床综合评价。
E-mail:13820893896@163.com

was based on the standard guideline formulation method of the World Health Organization (WHO), strictly followed the latest definition of guidelines by the Institute of Medicine of the National Academy of Sciences of the United States, and

conformed to the six major areas of the Guideline Research and Evaluation Tool II. Delphi method was adopted to construct the research questions; research evidence was established by applying the research methods of evidence-based medicine. The evidence quality classification system of the Chinese Evidence-Based Medicine Center was adopted for evidence classification and evaluation. The recommendation strength was determined by the recommendation strength classification standard formulated by the Oxford University Evidence-Based Medicine Center, and the recommendation opinions were formed through the expert consensus method.

RESULTS & CONCLUSIONS The *Guideline for the Workflow of Clinical Comprehensive Evaluation of Drugs* covers 4 major categories of research questions, including topic selection, evaluation implementation, evidence evaluation, and application and transformation of results. The formulation of this guideline has standardized the technical links of the entire process of clinical comprehensive evaluation of drugs, which can effectively guide the high-quality and high-efficient development of this work, enhance the standardized output and transformation application value of evaluation evidence, and provide high-quality evidence support for the scientific decision-making of health and the rationalization of clinical medication.

KEYWORDS clinical comprehensive evaluation of drugs; workflow; guideline; group standard; evidence-based medicine

药品临床综合评价是促进药品回归临床价值的基础性工作,是巩固和完善基本药物制度的重要措施,是健全药品供应保障制度的具体要求。近年来,国家卫生健康委员会相继发布《关于开展药品使用监测和临床综合评价工作的通知》(国卫药政函〔2019〕80号)、《关于规范开展药品临床综合评价工作的通知》(国卫办药政发〔2021〕16号)等多项政策,要求加强对各地和各类医疗卫生机构药品临床综合评价工作的重点部署,为药品临床综合评价工作提出了政策性要求和指导意见。而药品临床综合评价研究结果的稳定性和可靠性依赖于标准化、规范化的工作流程。为使相关文件的指导意见能够规范地落实到评价工作实践中,《药品临床综合评价工作流程指南》项目组制定了本指南,以供药品临床综合评价相关工作人员参考使用。

1 指南制定目的

本指南旨在阐明药品临床综合评价的定义范围,规范主要工作流程和相关技术环节,持续推动研究过程的标准化、科学化、同质化,提升综合评价证据的质量和转化应用价值,为药品临床综合评价工作的高质、高效开展提供参考和指导。通过指南的制定及应用,不仅可为药品临床综合评价相关研究者提供规范化的流程保障,还可提高广大卫生专业技术人员有效甄别、利用研究成果的能力,为卫生管理人员提供决策依据。

2 指南适用范围

本指南适用于开展药品临床综合评价工作的各类人员,包括但不限于政策决策人员和管理执行人员、医疗机构专业技术人员、大专院校科研人员、企事业单位相关从业人员等。

3 指南制定单位和项目组

本指南于2023年6月由中国药师协会药品临床评价工作委员会、中国药学会循证药学专业委员会、天津市药品使用监测和临床综合评价中心、天津医科大学总医院共同发起,由中国药师协会平台立项并形成团体标准;本指南项目组由26个省(区、市)67家单位(医疗机构、大专院校及研究所、行政管理部门)的75名专家组

成,除发起人外还包括指南指导专家组、指南制定专家组、指南外审专家组和指南工作组,各小组工作职责具体如下。

(1)指南指导专家组的工作职责包括:确定指南研究主题和应用范围;组建指南制定专家组、指南工作组和指南外审专家组,并管理所有成员的利益冲突;批准指南计划书;审核确定指南研究问题、证据质量及评审意见;监督指南全部制定流程;组织指南制定相关会议;审核指南推荐共识意见及指南全文,并批准指南发布。

(2)指南制定专家组的工作职责包括:根据指南研究主题和应用范围撰写指南计划书初稿;初步构建指南研究问题;指导指南证据检索及证据质量评估;制定德尔菲专家函询问卷;对推荐意见进行投票并达成共识,最终形成推荐意见决策表;撰写指南初稿,并根据指导意见进行修改和完善;规范推进指南的制定和发布。

(3)指南外审专家组的工作职责包括:审核和评价形成的指南推荐意见;评审指南初稿全文;形成专家反馈意见,提交指南指导专家组进行审定。

(4)指南工作组的工作职责包括:根据确定的指南研究问题进行系统性的证据检索、整理筛选和初步质量评估,将结果提交指南制定专家组;协调和记录工作会议、专家咨询等;协调其他指南制定相关事务性工作。

4 指南注册

《药品临床综合评价工作流程指南》在国际实践指南注册与透明化平台注册,注册号为PREPARE-2025CN052;注册分类为“标准指南”。

5 指南制定方法

《药品临床综合评价工作流程指南》的制定严格遵循美国国家科学院医学研究所关于指南的最新定义^[1],以世界卫生组织(World Health Organization, WHO)标准指南制定的方法学为依据^[2-3],并符合指南研究与评价工具II(Appraisal of Guidelines for Research and Evaluation II, AGREE II)的六大领域^[4]。本指南制定过程中的指南项目组组建、利益冲突声明、研究问题框架构建、证据检索和筛选、证据质量分级、推荐强度分级、指南外

审及报告撰写等技术流程均严格遵循标准指南制定方法进行^[5-8]。

5.1 研究内容及研究框架的确定

以国家药物和卫生技术综合评估中心以及国家卫生健康委员会药具管理中心发布的《药品临床综合评价管理指南(2021年版试行)》^[9](以下简称《管理指南》)为基础,根据现有研究文献及药品临床综合评价实际工作现状,初步整合指南研究内容;基于德尔菲法进行专家意见调研,构建指南研究问题框架,主要包括4类研究问题。

5.2 证据检索和筛选方法

针对已确定的研究问题,参照人群、干预、对照和结局(PICO)原则分解为主题词/自由词,并制定相应的检索策略,限定语言为中文和英文,检索时限自建库起至2024年12月。在指南制定专家组指导下由指南工作组研究人员统一检索和筛选。由2名研究人员同时独立进行证据的检索和筛选,存在争议时征求第3位研究人员的意见或召开指南工作组会议协商解决。

5.3 证据来源

本指南证据检索来源包括:(1)美国国立实践技术指南库、国际指南协作网、Trip循证医学数据库、英国国家卫生与临床优化研究所(National Institute for Health and Care Excellence, NICE)等指南数据库;(2)系统评价、卫生经济学评价研究数据库、the Cochrane Library、卫生技术评估数据库等;(3)PubMed、Embase、中国生物医学文献服务系统、万方、中国知网和维普等文献数据库。(4)国际和各国卫生行政主管部门、药监部门、医疗保障部门或相关行业协会的官方网站;(5)公共搜索引擎、AI检索工具(作为补充检索)。

5.4 证据纳入与排除标准

证据的纳入标准为:(1)内容涵盖卫生技术评估/药品临床综合评价工作流程相关研究内容;(2)指南、规范性文件及理论框架等研究性文献。

证据的排除标准为:(1)证据类型不符,包括摘要、讲座、评论、新闻、资讯等;(2)研究内容不符,包括具体药品或医疗器械的卫生技术评估/药品临床综合评价报告/临床研究;(3)其他内容不相关的文献或报告;(4)重复文献;(5)非中英文文献。

5.5 证据分级及推荐强度分级方法

本指南为管理型指南,采用中国循证医学中心的证据分级体系,将证据分为5级,证据分级标准见表1;采用牛津大学循证医学中心制定的推荐强度分级标准,将推荐强度分为4级,推荐强度分级标准见表2。

表1 证据分级标准

证据分级	描述
I级	系统评价、卫生技术评估、Meta分析
II级	政府及相关机构报告
III级	有确切研究方法的文献
IV级	综述
V级	专家意见

表2 推荐强度分级标准

推荐强度分级	描述
A	推荐,证据极有效(证据等级为I级)
B	可推荐,证据有效(证据等级为II级或III级),可能会在将来出现更高质量的新证据而后改变
C	证据在一定条件下有效(证据等级为IV级),应谨慎应用研究成果
D	证据的有效性具有局限性(证据等级为V级),只在较窄的范围内有效

5.6 推荐意见的形成

指南制定专家组根据证据质量、利弊权衡、各方意愿等因素形成初步的推荐意见决策表。按照欧洲临床营养与代谢学会指南手册,通过德尔菲法进行专家投票形成共识意见。共识意见形成规则为:(1)同意率>75%的建议被定义为最终推荐意见;(2)同意率<50%的建议被定义为未达成共识;(3)同意率在50%~75%的建议,代表大多数人同意,但会导致共识强度降级。

6 指南的研究内容

本指南涵盖药品临床综合评价的主题遴选、评价实施、证据评价、结果应用转化4类研究问题。

6.1 主题遴选

主题遴选在规范开展药品临床综合评价工作中是基础,更是前提和保障,对提升综合评价证据质量和转化应用具有重要意义。本指南主题遴选部分主要包括主题遴选组织管理、主题遴选评审流程、主题遴选评估方法3个方面,推荐内容已发表于《药品临床综合评价主题遴选指南》^[7],本文不再赘述。药品临床综合评价主题遴选的评估指标体系及权重见表3。

表3 药品临床综合评价主题遴选评估指标体系及权重表

一级指标	二级指标	三级指标	权重/分
必要性	卫生决策需求	对政策/医疗保健体系的重要性/影响	8.75
		评估结果的政策转化潜力	4.64
重要性	临床实践需求	评估的创新性和实践价值	12.16
		疾病负担	10.51
	健康需求	弱势群体影响	5.41
		降低患者死亡率/发病率潜力	3.23
	健康影响	患者生存质量/疾病预后的改善	2.91
		与现有治疗相比的潜在临床获益	1.71
可行性	经济性	安全性获益	1.69
		对社会和患者个人的经济负担	3.44
		对成本-效果/效用/效益的影响	2.34
社会效益	证据需求	证据可获得性和充分程度	7.75
		证据质量	4.58
		资源需求	当前是否可满足所需时间、资源和技术要求
公平性	社会、舆论关注程度	对伦理、法律、社会心理方面的影响	7.78
		个人/公共利益	5.34
	社会公平、伦理影响	药品使用对伦理和公平可及性的影响	11.42

6.2 评价实施

6.2.1 研究背景

研究背景的核心是阐述立题依据,陈述研究人员进行该项研究的意图以及开展此类工作的必要性,应重点介绍学科领域的基本情况,总结与研究主题直接相关的最新进展,展示不同研究的主要结论及其局限性,明确现有研究的不足或尚未解决的问题,并凸显本研究的必

要性及创新点,为本研究的内容确定及方法设计铺垫基础^[9-11](证据等级: I 级;推荐强度:A)。

6.2.2 研究目的

研究目的应清晰阐述研究的总体意图,明确本次药品临床综合评价的目的和待解决的问题。研究目的应简明扼要地阐述“运用何种理论和方法,解决何种主要问题,达到何种主要目的”。研究目的按重要程度可分为主要研究目的和次要研究目的^[10](证据等级: I 级;推荐强度:A)。

6.2.3 研究设计

研究设计应基于拟解决的问题、可获取的资源以及项目时间安排等因素,制定研究设计方案。按照认识论基础,研究设计可分为3大类:定性研究设计、定量研究设计和混合研究设计。药品临床综合评价工作由于其评价目的的多样化和评价维度的全面性,以混合研究设计的方法更为常见和适用,通常包括文献分析、临床研究、定性访谈和模型研究等。建议根据流行病学、卫生统计学、卫生技术评估的标准评价流程及方法进行药品临床综合评价研究设计^[10,12](证据等级: I 级;推荐强度:A)。

6.2.3.1 评价内容

(1)疾病类型及诊断标准:开展药品临床综合评价工作,需首先确立研究药品所针对的疾病类型,并明确该项研究所采纳的诊断标准,以便于明确研究对象的纳排标准,设计合理的研究结局指标。疾病类型方面,建议采用WHO发布的《国际疾病分类》作为分类标准和方法,对研究疾病进行描述和归类。疾病概况描述包括但不限于病因、临床表现、国内外发/患病率、国内外疾病经济负担、疾病诊疗方案与路径等。在药品临床综合评价工作中,应尽可能选择工具书(包括教材)及高质量指南作为疾病诊断标准的依据,并在研究设计中予以明确^[10,13-15](证据等级: I 级;推荐强度:A)。

(2)研究对象与纳排标准:研究对象方面,应明确研究对象的判断标准、源人群和选择方法;对于配对研究还应明确配对标准和对照例数,如需随访,则需明确随访方法。纳排标准方面,其制定主要涉及研究对象(患者)年龄、性别、疾病诊断、相关共病、是否为特殊人群等;同时也应尽可能明确患者的来源及招募方式^[10-11,16-17](证据等级: I 级;推荐强度:A)。

(3)干预措施:应明确研究药物相关干预措施实施的详细方案,包括药品通用名称、剂型、规格、生产企业、给药剂量、频率、给药方式、合并用药及疗程等。如研究药物为固定处方/协定方、经典方、院内制剂、名老中医经验方等,需明确药物的组成及剂量、剂型、制备方式及工艺。如有需要,还应明确药物的管理属性,如国家基本药物、国家基本医保属性等^[18-20](证据等级: I 级;推荐强度:A)。

(4)对照设置:包括对照方法和对照药物2个方面。

①对照方法:分为内对照和外对照,其划分取决于组间研究对象在时间、空间和人群间的相似性^[19-25],详见表4。值得注意的是,如选用外对照的方法,需评估其适用性和可比性、干预信息和重要混杂因素记录的全面性和完整性、结局指标的准确性等。②对照药物:分为安慰剂对照和阳性药物对照,应根据研究目的、研究类型、干预措施及评价维度选择对照药物^[19-25]。详见表5。

表4 对照方法的选择

对照方法	具体对照方式	证据等级	推荐强度
内对照	对照组与干预组来自同一目标总体、同一研究周期内,接受干预和对照措施。内对照包括安慰剂对照、阳性对照、不同剂量对照以及空白对照等	I 级	A
外对照	对照组与干预组并非来自同一目标总体,或非同期接受干预和对照措施,需从其他试验(研究)或历史病例中选取特征相似的研究对象作为对照。外对照可来源于既往已完成的随机对照试验,也可来源于回顾性/前瞻性的真实世界研究	I 级	A

表5 对照药物的选择

对照药物	适用范围	证据等级	推荐强度
安慰剂对照	适用于轻症或功能性疾病或目标疾病尚无有效治疗措施或自限性疾病;可选择完全空白对照或在干预组和对照组给予相同的对症治疗措施基础上进行空白对照	I 级	A
阳性药物对照	适用于目标疾病已有明确或经典的治疗药物/治疗策略;采用具有相同适应证的经典治疗方案,并尽可能采用药理作用机制相同以及指南一线推荐的治疗方案中的药品	I 级	A

6.2.3.2 评价维度

药品临床综合评价主要聚焦药品临床使用实践中的重大技术问题和政策问题,围绕技术评价与政策评价两条主线,从安全性、有效性、经济性、创新性、适宜性、可及性6个维度,开展科学规范的定性和定量相结合的数据整合分析与综合研判。评价实施机构应基于评价目的及评价结果的应用转化场景,通过专家论证的方式确定合适的评价维度,并在研究方案中详细阐述评价维度确定的理由^[9-10,26],评价维度详见表6。

6.2.3.3 信息来源

药品临床综合评价工作可根据评价维度及评价方法,从多种渠道获得相关信息资料。为确保综合评价结果适当、可靠,信息来源相关情况^[10-11,31,39-44]详见表7。

6.2.4 研究类型

根据研究目的、研究设计、评价维度、数据来源等多种因素,选择适宜的研究类型。本指南仅列出药品临床综合评价中常用的各种研究类型^[9-10,45-46](证据等级: I 级;推荐强度:A)。

6.2.4.1 原始研究

原始研究以一手(原始)资料为研究基础,进行综合整理、分析研究,可根据是否有干预因素分为观察性研究和实验性研究,详见表8。

6.2.4.2 二次研究

二次研究是在已有研究基础上,应用统计技巧分析初始研究的数据资料,以回答初始研究提出的问题或试图回答新问题。二次研究方法主要包括系统评价、指南/专家共识、决策分析^[10-11,45,60-65],详见表9。

表6 药品临床综合评价的评价维度表

评价维度	维度定义	评价内容	评价要点	证据等级	推荐强度
安全性	综合分析药品上市前后安全性信息结果 ^[9, 26-28]	药物临床试验数据、不良反应、不良事件、禁忌证、注意事项、药物过量、相互作用、相对安全性(与同类产品比较)、药品质量、药品疗效稳定性等信息	(1)说明书安全性信息;(2)说明书以外的各国药监部门披露的药物警戒信息;(3)药理学评估;(4)非临床安全性研究;(5)临床安全性研究结果及质量评估;(6)上市后药品不良反应/事件;(7)产品召回措施和落实情况;(8)预防用药差错事故采取的警示措施	I级	A
有效性	通过定量分析,对拟评价药品及参比药品的临床效果进行人群测量,判断是否获得重要的健康收益 ^[9, 26-27, 29-30]	(1)核心指标:①生存时长,包括生存率、疾病控制率以及其他能够反映疾病进展的可测量指标;②生命质量,包括健康相关生命质量和健康效用值。(2)其他指标:①临床疗效指标;②实验室检查指标;③影像检查等特异性指标;④患者主观感受等	(1)药理与药理学特性;(2)药品上市前临床研究数据;(3)药品上市后临床研究数据及质量评估;(4)典型有效性病例报告/病例系列报告(主要针对上市时间较短的创新药物及罕见病用药);(5)指南、指导原则、专家共识、临床路径收录情况	I级	A
经济性	综合运用流行病学、卫生统计学、决策学、经济学等多学科理论及方法,分析测算药品的成本、效果、效用和效益等 ^[9, 26-27, 31-34]	(1)根据药品决策的具体需求,可选择开展成本-效果分析、成本-效用分析、成本-效益分析、最小成本分析等;(2)强化增量分析及不确定性分析,必要时进行卫生相关预算影响分析,全面研判药品临床应用的经济价值及影响	(1)药品价格、日费用、疗程费用;(2)药物经济学研究数据及质量评估,包括临床研究及模型研究;(3)医疗保险目录收录情况	I级	A
创新性	通过分析判断药品与参比药品满足临床需求的程度,鼓励国产原研创新等情况,评价药品的创新性 ^[9, 26-27, 35-36]	开展创新性评价,应当突出填补临床治疗空白,解决临床未满足的需求,满足患者急需诊疗需求,推进国内自主研发等创新价值判断	(1)临床创新性:满足临床尚未满足的诊疗需求,与同类药品比较的优势特色,药物作用机制创新、剂型创新。(2)企业服务体系创新性:优化诊疗服务流程,提升服务效率,提高诊疗服务资源利用率。(3)产业创新性:是否为自主研发的原研药或全球首个仿制药,是否获得国内药品结构或制剂工艺方面的专利,是否获得国际通用药品结构或制剂工艺方面的专利,是否为儿童专用药品(专用剂型规格)等	I级	A
适宜性	评价药品技术特点适宜性和药品使用适宜性 ^[9, 26-27, 37]	(1)药品技术特点适宜性可从药品标签标注、药品说明书、储存条件等方面进行评价;(2)药品使用适宜性主要包括患者给药方法、给药频率、药品规格、装量及包装是否恰当等	(1)药品技术适宜性:标签标注完整性、药品说明书标注准确性、起效时间、药物剂型等患者使用适宜性:储存条件适宜性、是否需要特殊装置、用药后是否需要监测或随访等;(2)药品使用适宜性:给药适宜性(如药物用法、给药途径是否符合患者情况,用药是否针对适应证,联合用药是否合理,用药是否准确排除禁忌证等)和用约适宜性(如治疗方案、适用人群、给药间隔、给药周期等是否适宜)	I级	A
可及性	评价药品可获得性和可负担性 ^[9, 26-27, 33-34, 38]	(1)可获得性可由医疗机构药品配备使用情况或有短缺情况等反映;(2)可负担性可由人均年用药治疗费用占城乡居民家庭年可支配收入比重(%)体现	(1)可获得性报告:药品销售范围、配备使用情况及可持续性;(2)可负担性报告:药品疗程费用/可支配收入比例等	I级	A

表7 药品临床综合评价的信息来源情况

信息来源	信息内容	证据等级	推荐强度
文献数据库	指南数据库;系统评价、卫生技术评估、卫生经济学评价研究数据库,以及期刊文献等数据库中的研究文献,包括卫生技术评估报告、指南/专家共识、循证研究、随机对照试验、真实世界研究、药物经济学评价、病例报告、基础试验研究、质量评估等	I级	A
官方网站	国际和各国卫生行政主管部门、药监部门、医疗保障部门或相关行业协会的官方网站;包括通报记录、药物警戒、政策属性等药品公开信息	I级	A
书籍专著	该领域的基础知识、定义概念及研究进展等	I级	A
临床试验注册数据库	临床试验研究注册情况;试验基本信息、设计方案以及研究者信息等	I级	A
生产企业	药品基本信息、生产工艺、专利、药品说明书、质量控制、药品产能/销量、药品定期安全性更新报告、未发表或正在开展的研究等	I级	A
问卷调查	可根据待评价药品的特点和维度定向开展问卷调查,主要调研主观感受或认知程度等,可汇总专家或公众意见	I级	A
临床资料	患者诊疗信息、人口学信息、疾病症状及诊断、疾病及治疗史、家族史、干预方式及随访结果等	I级	A
数据库	卫生信息系统数据库、医疗保险数据库、药品供应销售及采购数据库、药品临床有效性数据库、不良反应监测数据库、其他主要人口健康相关数据库等	I级	A
灰色文献	未公开发表的研究信息,可能包含了最新的研究成果和观点	I级	A
搜索引擎	公众/媒体关注度等补充数据	I级	A

表8 原始研究的研究类型、研究方法及其内容

研究类型	分类	研究方法及其内容	证据等级	推荐强度
观察性研究	描述性研究	主要用于产生假设,从而为进一步的分析性研究奠定基础。常用的描述性研究包括横断面研究、病例报告、病例系列及生态学等 ^[45, 47-57]	I级	A
	分析性研究	对结局的可能暴露因素进行检验的方法,按照暴露与结局的时间序列不同,可分为病例-对照研究、队列研究、巢式病例对照研究 ^[45, 47-57]	I级	A
实验性研究	随机对照试验	将研究对象随机分组,对不同组实施不同的干预,以对比效果的不同 ^[10, 45, 58-59]	I级	A
	非随机对照试验	不适宜或不必要开展随机对照试验时,可考虑采用非随机对照试验。常用的非随机对照试验包括自身前后对照研究、交叉对照研究、序贯试验等 ^[10, 45, 58-59]	I级	A

表9 二次研究的研究类型、研究方法及其内容

研究类型	研究方法及其内容	证据等级	推荐强度
系统评价	系统评价又称系统综述,是指对研究问题进行结构化的清楚定义,运用系统和明确的方法对相关研究进行检索、选择以及严格的评价。对所纳入研究的数据进行收集和分析,并进一步讨论总结形成的综述,可分为定性评价和定量评价(Meta分析),主要用于干预措施的有效性/安全性评价	I级	A
指南/专家共识	临床指南是基于系统评价在内的当前最新、最佳证据,总结不同于干预措施的利弊之后形成能够为临床问题提供最佳推荐意见,可作为指导临床实践的文件,是确保医疗质量的重要手段。专家共识是指多个学科专家代表组成团队针对具体临床问题达成推荐意见	I级	A
决策分析	决策分析是针对临床问题的风险与收益的不确定性,通过决策模型,利用概率和结局估计值等,确定最佳方案的过程。按照决策的可靠程度可分为3类:确定型决策(如成本分析法)、风险型决策(如期望值决策、贝叶斯决策)及不确定型决策(如最大最小准则)	I级	A

6.2.4.3 其他研究

在药品临床综合评价工作中,原始研究和二次研究以外,还有其他研究方法可被应用^[10-11, 66-67],详见表10。

表10 其他研究的研究类型、研究方法及其内容

研究类型	研究方法及其内容	证据等级	推荐强度
药物经济学研究	常用方法包括成本-效果分析、成本-效用分析、成本-效益分析和最小成本分析	I级	A
社会调查	社会调查是社会“调查”和“研究”的简称。药品临床综合评价工作中,社会调查通常通过问卷调查的方式进行,以考察对药品相关情况的主观认知	I级	A

6.2.5 样本量

样本量估算是为满足统计的准确性和可靠性(I类错误的控制和检验效能的保证)计算出所需的样本量,直接关系到研究结论的可靠性、可重复性及研究效率。样本量的确定既需保证有足够的研究样本验证研究目的,也需避免过多的受试者在非必要的情况下暴露于研究风险,减少人力、物力和时间的浪费,从而使研究

设计更符合科学性和伦理学原则,便于研究的具体实施^[11,68-74]。

(1)样本量估算的前提:①明确设计方法;②确定统计分析方法;③考察主要结局指标的特性。(2)样本量估算的影响因素:①统计检验水准(α ,即第一类错误的概率),一般采用0.05;②统计检验效能($1-\beta$, β 为第二类错误的概率), β 一般采用0.1或0.2,统计检验效能分别为90%或80%;③允许误差(δ): δ 表示所比较的2个总体参数间的差别, δ 越大,所需样本量越小;④总体相关信息,包括总体平均数、标准差或总体率等,可通过预试验、查阅文献、经验估计获得;⑤确定单侧/双侧检验:根据研究目的和研究设计确定。样本量估算的适用性需符合研究设计和研究方法(证据等级:Ⅲ级;推荐强度:B)。

6.2.6 随访时间

在医学随访中,应对各研究对象追踪观察,记录各时间点的具体情况以评价疗效或预后。对于随访时间的设置应明确其起止时间,设置的适用性需符合研究设计和研究方法。

6.2.6.1 明确开始随访的时间

明确开始治疗时间、入院时间、确诊时间等,一般为可明确记录的时间或开始研究的时间^[32,75-82](证据等级:I级;推荐强度:A)。

6.2.6.2 明确随访截止时间

(1)定时截尾:预设观察截止时间,即观察到该时间就停止随访;(2)定数截尾:预设观察人数/观察事件,即达到该人数或发生一定数量的事件就停止随访^[32,75-82](证据等级:I级;推荐强度:A)。

6.2.7 结局指标

药品临床综合评价工作中,可根据不同疾病或治疗领域设定针对性的评价指标作为研究的结局指标。结局指标的设置应与临床实践相一致,并考虑其可理解与可应用的程度。结局指标的选择可参考已发布的《管理指南》、技术指南、核心结局指标集、高质量指南/专家共识、系统评价及随机对照试验等。

结局指标可依据重要程度及应用需求分为主要结局指标和次要结局指标。建议采用国际公认的临床指标作为主要结局指标;采用生命质量、患者主观评价、健康总体评价等辅助性指标作为次要结局指标^[9-10,18,20,32,45-46](证据等级:I级;推荐强度:A)。

6.2.8 测量方法

在完成结局指标的构建后,应对每一个结局指标确定一个合适的结局指标测量工具,既需评价测量工具的质量,也应充分考虑临床医生和研究者对测量工具性能的反馈及选择测量工作的偏好。测量工具的选择适宜性可采用“基于共识标准选择健康测量工具”(consensus-based standards for the selection of health

measurement instruments, COSMIN)工作组制定的COSMIN清单进行评价^[10,83-86](证据等级:I级;推荐强度:A)。

6.2.9 偏倚控制

6.2.9.1 常见偏倚类型

研究偏倚是指观察到关联存在系统误差,需要控制偏倚,以保障研究结果的稳定性。根据偏倚的来源,可分为选择偏倚、信息偏倚和混杂偏倚^[87-90]。

6.2.9.2 偏倚控制方法

偏倚控制方法包括:(1)选择有恰当纳排标准的研究对象;(2)纳入适当的样本量(兼顾区域平衡性);(3)采用匹配或限制;(4)选择合适的对照药物;(5)验证主要协变量;(6)采用恰当的统计分析方法。任何有偏倚风险或证据不确定性的问题都需要与研究细节一起报告,或者作为结论(总结)的一部分进行报告^[87-90](证据等级:I级;推荐强度:A)。

6.2.10 质量控制

为保障综合评价质量,需对评价关键环节实施严谨、规范的质量控制。研究质量控制应贯穿方案设计、项目实施、数据管理、统计分析、报告撰写的整个过程,对研究中可能存在的质量问题预先进行把控。

6.2.10.1 质量控制措施

所有研究者均应采用标准操作规程,以保证研究的实施和质量。研究中所有的调研项目应确保其来源于原始数据,具有可靠性^[10-11,26,43-45](证据等级:I级;推荐强度:A)。

6.2.10.2 质量控制方法

质量控制方法主要包括现场质控、实验室质控、双人评价、争议处理、原始数据存档等^[10-11,26,43-45](证据等级:I级;推荐强度:A)。

6.2.11 统计分析

6.2.11.1 统计分析计划

为保障统计分析过程的规范完整,在药品临床综合评价工作开始实施前,需拟定研究统计分析计划,包括数据集的选择、缺失数据的处理、统计分析方法和统计模型的选择等^[9-11,91-93](证据等级:I级;推荐强度:A)。

6.2.11.2 统计分析方法

在药品临床综合评价工作基础数据收集完成后进行数据统计分析,包括缺失值处理、控制混杂、失访处理、统计描述、单因素统计分析和多因素统计分析、敏感度分析等^[9-11,91-93](证据等级:I级;推荐强度:A)。

6.2.11.3 统计分析报告的撰写

根据研究方案、统计分析计划书和统计分析结果,撰写统计分析报告,包括原始数据/资料的质量评估,统计分析方法的选择及分析结果等:(1)统计分析结果尽可能以图表形式展现,以提高可读性和条理性;(2)统计检验结果应包括检验水准、统计值和精确的P值;(3)应注明使用的统计学软件及版本号;(4)所有统计计算机

程序应以文件形式保存以备核查,应包括各评价维度证据合成的方法、模型、算法、软件、结果展示方式等^[9-11,91-93](证据等级: I 级;推荐强度:A)。

6.2.12 研究结果

6.2.12.1 适用性

适用性指综合评价结果的适用人群(评价角度)及应用场景^[10-11](证据等级: I 级;推荐强度:A)。

6.2.12.2 局限性

局限性指综合评价实施过程中存在的方法学等方面的局限^[10-11](证据等级: I 级;推荐强度:A)。

6.2.13 报告规范

对研究结果进行报告时,须按照相应的报告规范及

条目进行报告,以提高研究报告质量,使研究更为完整、透明和清晰^[9-11,94-97](证据等级: I 级;推荐强度:A),详见表 11。在药品临床综合评价工作实施完成后,须考察所有工作流程和工作环节的完成情况,并将所有相关资料留存备案。

6.3 证据评价

药品临床综合评价工作中,无论是自身完成的研究结果还是外部获得的直接证据,在应用转化之前,均需进行严格的证据评价,以保证其可靠性和科学性。本指南的证据评价主要指方法学质量评估。按照不同研究类型对质量评估工具进行分类,以便合理选用^[9-10,60,118-150],详见表 12。

表 11 研究结果的报告规范

报告内容	亚组	具体研究	报告规范	证据等级	推荐强度
临床研究报告			《临床试验方案规范指南》(SPIRIT) ^[98]	I 级	A
原始研究	观察性研究	队列研究、病例-对照研究、横断面研究	《观察性流行病学研究的报告指南》(STROBE) ^[99]	I 级	A
		真实世界队列研究、真实世界病例-对照研究	《常规收集医疗卫生数据开展观察性研究的报告规范》(RECORD、RECORD-PE) ^[100-102]		
		病例系列报告	尚无国际通用的报告标准,建议参考 STROBE 中的部分内容 ^[99]		
		病例报告	《病例报告指南》(CARE) ^[103-104]		
		中药病例报告	《中医病例报告建议条目》以及《基于共识的中医病例报告规范》(CARC) ^[105]		
	试验性/干预性研究	随机对照试验	《实用性随机对照试验报告规范》(CONSORT) ^[106]		
		真实世界随机交叉研究	《CONSORT 2010 声明:随机交叉试验扩展版》 ^[107]		
		真实世界单病例随机对照试验	《单病例随机对照试验 CONSORT 扩展版》(CENT) ^[108]		
		中药随机对照试验	《草药干预措施随机对照试验报告规范》以及《中药复方临床随机对照试验报告规范》(CONSORT-CHM Formulas) ^[109-110]		
		中药单病例随机对照试验	《中医药单病例随机对照试验 CONSORT 扩展版》(CENT for TCM) ^[111]		
二次研究	系统评价和 Meta 分析	使用队列或收集日常数据的随机对照试验	《使用队列或收集日常数据的随机对照试验的报告规范》(CONSORT-ROUTINE) ^[112]	I 级	A
		非随机对照试验	《非随机对照试验报告规范》(TREND) ^[113]		
		系统评价和 Meta 分析优先报告的条目》(PRISMA 2020) ^[114]			
其他研究	药物经济学研究	中草药系统评价和 Meta 分析	《中草药系统评价和 Meta 分析优先报告的条目》(PRISMA-CHM 2020) ^[115]	I 级	A
		临床实践指南报告规范	《国际实践指南报告规范》(RIGHT) ^[116]		
			综合健康经济评估报告标准(CHEERS 2022)声明 ^[117]		

表 12 药品临床综合评价的质量评估工具

研究类型	亚组	分类1	分类2	质量评估工具(量表、清单、条目)	证据等级	推荐强度
原始研究	观察性研究	描述性研究	横断面研究	AXIS 工具 ^a 、NIH 质量评估工具、JBI 分析性横断面研究评估清单、AHRQ 方法学量表、Crombie 评估条目	I 级	A
			病例系列/病例报告	JBI 清单 ^a		
		分析性研究	队列研究	CASP 队列研究清单、SIGN 队列研究清单、NIH 观察性队列研究和横断面研究质量评估工具、NOS ^a 、JBI 清单		
	病例-对照研究		CASP 病例-对照研究清单、SIGN 病例-对照研究清单、NIH 病例-对照研究质量评估工具、JBI 病例-对照研究清单、NOS 病例-对照研究量表 ^a			
	其他观察性研究		CASP 量表、Cochrane 偏倚风险工具、ArRoWs 工具 ^a			
	实验性研究	随机对照试验	非随机对照试验	Cochrane 偏倚风险工具 ^a 、CASP 清单、Jadad 量表		
ROBINS- I 评价工具 ^a 、JBI 清单、MINORS 评分条目、EPOCRoB 工具、IHE 质量评估工具						
AMSTAR-2 量表 ^a 、ROBIS 工具 ^a 、CASP 系统评价清单、JBI 系统评价与研究整合评价清单、NIH 系统评价与 Meta 分析质量评估工具、SIGN 系统评价与 Meta 分析评价清单						
二次研究	系统评价和 Meta 分析	网状 Meta 分析	网状 Meta 分析方法学清单			
		临床实践指南	AGREE II ^a			
		其他研究	SIGN 经济学评价清单、CASP 经济学评价清单、NICE 经济学评价方法清单			
其他研究	卫生技术评估报告	卫生技术评估报告	卫生技术评估 Checklist 量表	I 级	A	

a: 目前最常用或最推荐的质量评估方法;AXIS: Appraisal tool for Cross-Sectional Studies, 横断面研究评价工具;NIH: National Institutes of Health, 美国国立卫生研究院;JBI: Joanna Briggs Institute, 乔安娜·布里格斯学院;AHRQ: Agency for Healthcare Research and Quality, 美国卫生保健和质量机构;Crombie: Crombie's Critical Appraisal Checklist, 克伦比批判性评价清单;CASP: Critical Appraisal Skills Programme, 文献严格评价项目;SIGN: The Scottish Intercollegiate Guidelines Network, 苏格兰国家指南小组;NOS: Newcastle-Ottawa Scale, 纽卡斯尔-渥太华量表;ArRoWs: Assessment of Real World Observational Studies, 真实世界观察性研究评估工具;Jadad 量表: Jadad scale, 贾达德量表;ROBINS- I: Risk of Bias in Non-randomised Studies of Interventions, 干预性非随机研究偏倚风险;MINORS: Methodological Index for Non-randomized Studies, 非随机研究的方法学指标;EPOCRoB: Cochrane Effective Practice and Organisation of Care Group Risk of Bias Tool, Cochrane 有效实践与医疗组织小组偏倚风险评估工具;IHE: Canada Institute of Health Economics, 加拿大卫生经济学研究所;AMSTAR: A Measurement Tool to Assess Systematic Reviews, 一种评估系统评价的测量工具;ROBIS: Risk of Bias in Systematic Review, 系统评价偏倚风险。

6.4 结果应用转化

6.4.1 证据应用与反馈

6.4.1.1 应用范围

根据《管理指南》提出的药品临床综合评价结果证据的应用范围,推荐综合评价证据结果可广泛应用于以下场景,充分发挥其指导价值。(1)医疗卫生机构运营层面:支撑机构药品采购与供应保障工作,同时推动用药目录遴选、上下级机构用药目录衔接,助力提升药学服务质量与安全合理用药水平;(2)卫生健康资源优化层面:通过科学指引,控制不合理药品费用支出,提升卫生健康资源配置效率,优化药品使用结构;(3)政策与科技发展层面:为完善国家药物政策提供参考依据,丰富行业内药品临床综合评价实践经验、扩大文献证据储备,同时推动科研领域对药品临床综合评价理论及方法的深入探索^[9-10,151](证据等级: I 级;推荐强度: A)。

6.4.1.2 反馈要求

(1)在药品临床综合评价结果的转化应用过程中,需重点关注证据质量与应用效果,建立动态优化机制;(2)注重对证据的适用性、科学性、规范性进行全面评价与考察,确保应用方向准确、依据可靠;(3)建立及时反馈机制,针对应用中发现的问题、不足进行实时收集与分析,进而持续改进评价结果的质量与应用效能,形成“评价-应用-反馈-优化”闭环^[9-10,151](证据等级: I 级;推荐强度: A)。

6.4.2 评价结果的优化与完善

药品临床综合评价工作应根据应用反馈的结果,对研究方法进行优化与完善。按照《管理指南》要求:药品临床综合评价实施机构应持续跟踪已完成评价药品的实际供应与应用情况,不断累积相关数据验证评价结果。长期用药持续跟踪时间根据疾病或病情而定,通常不少于1年。针对有调整需求的国家及省级药品临床综合评价结果,委托项目的主要承担机构应结合特定领域政策需要及国内外评估机构证据更新情况,适时开展证据优化和结果更新。国家及省级药品临床综合评价的证据优化及结果更新,应结合药品临床应用监测等相关评价研究数据,提出证据核查或更新的书面请示,经论证后,确定是否启动有关工作^[9-10,151](证据等级: I 级;推荐强度: A)。

7 指南更新

本指南将根据国家管理政策的调整、药品临床综合评价内涵的变化、技术方法的更新、证据评价体系的迭代等情况进行定期更新,对于有争议领域出现的新证据,须优先评价,更新推荐意见。

指南项目组

发起人

李正翔(天津医科大学总医院)

指南撰稿人

段 蓉(天津医科大学总医院)

指南指导专家组专家

史录文(北京大学)

田金徽(兰州大学)

左笑丛(中南大学湘雅三医院)

张 玉(华中科技大学同济医学院附属协和医院)

张伶俐(四川大学华西第二医院)

张俊华(天津中医药大学)

郑华林(中国国际科技促进会)

赵荣生(北京大学第三医院)

郭武栋(国家卫生健康委卫生发展研究中心)

缪丽燕(苏州大学附属第一医院)

翟所迪(北京大学第三医院)

指南制定专家组专家

于 倩(吉林大学中日联谊医院)

王婧雯(空军军医大学西京医院)

王 舒(中国医科大学附属盛京医院)

卢晓阳(浙江大学医学院附属第一医院)

卢海儒(青海省人民医院)

史 琛(华中科技大学同济医学院附属协和医院)

刘玉国(山东第一医科大学附属肿瘤医院)

许东航(浙江大学医学院附属第二医院)

朱建国(苏州大学附属第一医院)

刘茂柏(福建医科大学附属协和医院)

刘国强(河北医科大学第三医院)

伍俊妍(中山大学孙逸仙纪念医院)

张四喜(吉林大学第一医院)

张毕奎(中南大学湘雅二医院)

陈 孝(中山大学附属第一医院)

何金汗(四川大学华西药学院/华西医院)

杨建华(新疆医科大学第一附属医院)

陈孟莉(解放军总医院)

陈 琦(贵州省人民医院)

李 莎(内蒙古医科大学附属医院)

邱 峰(重庆医科大学附属第一医院)

李晓宇(复旦大学附属中山医院)

李晋奇(四川省人民医院)

沈承武(山东第一医科大学附属省立医院)

沈爱宗(中国科学技术大学附属第一医院)

吴 晶(天津大学)

吴 斌(上海交通大学医学院附属胸科医院)

吴 璠(天津市卫生健康委)

周 欣(天津医科大学总医院)

金鹏飞(北京医院)

封卫毅(西安交通大学第一附属医院)

胡建新(江西省人民医院)

赵志刚(首都医科大学附属北京天坛医院)

赵振宇(天津医科大学朱宪彝纪念医院)

侯锐钢(山西医科大学第二医院)

段 蓉(天津医科大学总医院)

郭 澄(上海交通大学医学院附属第六人民医院)
海 鑫(哈尔滨医科大学附属第一医院)
崔 林(江苏省卫生健康发展研究中心)
菅凌燕(中国医科大学附属盛京医院)
董占军(河北省人民医院)
韩 晟(北京大学医药管理国际研究中心)
蒋艳秋(黑龙江省卫生健康委)
董 梅(哈尔滨医科大学附属肿瘤医院)

指南外审专家组专家

刘 耀(陆军特色医学中心)
孟丽华(中国药师协会)
杨宏昕(内蒙古自治区人民医院)
张志清(河北医科大学第二医院)
宋燕青(吉林大学白求恩第一医院)
郑晓俊(山西医科大学第一医院)
赵青威(浙江大学医学院附属第一医院)
赵 泉(烟台毓璜顶医院)
赵冠人(解放军总医院)
高 申(海军军医大学第一附属医院)
夏培元(陆军军医大学第一附属医院)
李国辉(中国医学科学院肿瘤医院)
陈万生(海军军医大学附属长征医院)
戴海斌(浙江大学医学院附属第二医院)
林 阳(首都医科大学附属北京安贞医院)
夏 泉(安徽医科大学第一附属医院)
张抗怀(西安交通大学第二附属医院)
丁玉峰(华中科技大学同济医学院附属同济医院)
刘洪涛(河北医科大学第一医院)

指南工作组专家

孙 蕾(天津医科大学总医院)
吴佳瑶(天津医科大学总医院)
张铭慧(天津医科大学总医院)
张子健(天津医科大学总医院)
谢 栋(天津医科大学总医院)

参考文献

[1] Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice Guidelines, GRAHAM R, MANCHER M, et al. Clinical practice guidelines we can trust[M]. Washington, D. C.: National Academies Press, 2011:57-104.

[2] 杨克虎. 世界卫生组织指南制定手册[M]. 兰州: 兰州大学出版社, 2013:9-72.

[3] SINCLAIR D, ISBA R, KREDO T, et al. World Health Organization guideline development: an evaluation[J]. PLoS One, 2013, 8(5):e63715.

[4] BROUWERS M C, KERKVLIIET K, SPITHOFF K. The AGREE reporting checklist: a tool to improve reporting of clinical practice guidelines[J]. BMJ, 2016, 352:i1152.

[5] CHEN Y L, YANG K H, 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.

[6] 李正翔, 张玉, 张伶俐, 等. 医疗机构药品遴选指南[J]. 中国药房, 2022, 33(7):769-776.

[7] 李正翔, 张伶俐, 段蓉, 等. 药品临床综合评价主题遴选指南[J]. 中国药房, 2025, 36(8):897-902.

[8] 段蓉, 李正翔. 基于德尔菲法和层次分析法构建《药品临床综合评价主题遴选指南》研究问题及主题遴选评估指标体系[J]. 中国药房, 2025, 36(7):789-793.

[9] 国家药物和卫生技术综合评估中心, 国家卫生健康委药具管理中心. 药品临床综合评价管理指南: 2021年版试行[EB/OL]. [2025-07-10]. <http://www.nhc.gov.cn/cms-search/downFiles/863665c55fd74f29a656990d1a4ea5ce.pdf>.

[10] 国家药物和卫生技术综合评估中心. 药品临床综合评价质量控制指南: 2024年版试行[EB/OL]. [2025-07-10]. <http://nhei.cn/nhei/znfb/202412/d33db3fcddf94daeabe49-cfa95fa24f9.shtml>.

[11] CUSCHIERI S. The STROBE guidelines[J]. Saudi J Anaesth, 2019, 13(5):31.

[12] COHEN L, MANION L, MORRISON K. Research methods in education[M]. London: Routledge, 2017:1-15.

[13] 万学红, 卢雪峰. 诊断学[M]. 9版. 北京: 人民卫生出版社, 2018:1, 89, 237.

[14] 尚红, 王兰兰. 实验诊断学[M]. 3版. 北京: 人民卫生出版社, 2015:2-8.

[15] 李永, 陈启亮, 赵文, 等. 中西医诊断术语的差异性比较研究[J]. 天津中医药, 2020, 37(9):972-975.

[16] 詹思延. 第三讲: 如何报告观察性流行病学研究: 国际报告规范 STROBE 解读[J]. 中国循证儿科杂志, 2010, 5(3):223-227.

[17] CHENG A, KESSLER D, MACKINNON R, et al. Reporting guidelines for health care simulation research: extensions to the CONSORT and STROBE statements[J]. Adv Simul, 2016, 1:25.

[18] 中华中医药学会《中医药真实世界研究技术规范》制订组. 中医药真实世界研究技术规范: 证据质量评价与报告[J]. 中医杂志, 2022, 63(3):293-300.

[19] GRAY C M, GRIMSON F, LAYTON D, et al. A framework for methodological choice and evidence assessment for studies using external comparators from real-world data[J]. Drug Saf, 2020, 43(7):623-633.

[20] 彭晓霞, 舒啸尘, 谭婧, 等. 基于真实世界数据评价治疗结局的观察性研究设计技术规范[J]. 中国循证医学杂志, 2019, 19(7):779-786.

[21] 李戈, 杨智荣, 赵厚宇, 等. 真实世界研究外对照设置的方法学进展[J]. 医药导报, 2022, 41(1):38-43.

[22] JAHANSHAHI M, GREGG K, DAVIS G, et al. The use of external controls in FDA regulatory decision making[J]. Ther Innov Regul Sci, 2021, 55(5):1019-1035.

- [23] 王晓敏,田勇泉. 安慰剂对照试验的伦理辩护[J]. 伦理学研究,2013(2):124-127.
- [24] 雷翔,刘智,翟静波,等. 试论中药临床研究阳性对照药选择的原则及方法[J]. 中国新药杂志,2014,23(19):2276-2279.
- [25] 翁鸿,尹庆锋,王朝阳,等. 临床研究方案设计要点之对照药物的选择[J]. 中国循证心血管医学杂志,2017,9(4):385-387.
- [26] 李正翔. 医疗机构药品遴选指南[M]. 北京:化学工业出版社,2022:119.
- [27] 李正翔,段蓉. 基于德尔菲法构建《医疗机构药品遴选指南》研究问题及药品遴选指标体系[J]. 中国医院药学杂志,2020,40(22):2372-2376.
- [28] 徐小微. 中国药品综合评价指南参考大纲:第一章 药品安全性评价指南[J]. 药品评价,2011,8(18):6-7.
- [29] 刘芳,翟所迪. 中国药品综合评价指南参考大纲:第二版:第二章 药品有效性评价指南[J]. 药品评价,2015,12(8):8-11.
- [30] 朱珠. 中国药品综合评价指南参考大纲:第二版:第三章 体内药理学特性评价指南[J]. 药品评价,2015,12(8):11-12.
- [31] Agency for Care Effectiveness(ACE). Process & methods [EB/OL]. [2025-07-08]. <https://www.ace-hta.gov.sg/resources/process-methods>.
- [32] The Canadian Agency for Drugs and Technologies in Health (CADTH). Topic identification and prioritization process[EB/OL]. [2025-07-13]. https://www.cadth.ca/sites/default/files/pdf/HTA_OU_Topic_ID_and_Prioritization_Process.pdf.
- [33] 石丰豪,李薇,韩晟. 药物经济学评价中的风险决策与信息价值分析[J]. 世界临床药物,2021,42(12):1066-1072.
- [34] 陈景,杨磊,张琴,等. 中药药物经济学国内外评价现状及问题[J]. 中国药物经济学,2021,16(11):5-9.
- [35] 苏子仁,陈建南,沈小玲,等. 试论中成药制备过程的药性改变与中成药创新[J]. 中药新药与临床药理,2002,13(1):1-3,64.
- [36] 彭龙希,李蒙,杨圆圆,等. 药品生产企业核心竞争力的定义与其评价体系构建[J]. 中国药房,2021,32(7):782-787.
- [37] 孙春华. 中国药品综合评价指南参考大纲:第二版:第五章 药品顺应性评价指南[J]. 药品评价,2015,12(8):15-16.
- [38] 管晓东,史录文. 基于WHO/HAI标准调查法的我国基本药物可及性评价方法研究[J]. 中国药房,2013,24(24):2212-2215.
- [39] 徐陆欣怡,王妮,尹适成,等. 我国罕见病药品的可及性及采购使用现状研究[J]. 中国卫生政策研究,2022,15(2):60-64.
- [40] 徐倩,徐斌. 上海市口服抗帕金森病药品可负担性评价[J]. 中国药业,2021,30(23):1-4.
- [41] 刘思雨. 中国药品专利期补偿制度下的药品可及性问题及应对[C]//《上海法学研究》集刊:上海市法学会知识产权法研究会文集,2021:56-62.
- [42] 曹瑾,张铭,赵俊贤,等. 甘肃省急性病患者家庭用药合理性和药品可及性调查研究[J]. 中国全科医学,2022,25(4):438-444,452.
- [43] 董名扬,王舒,菅凌燕. 药品临床综合评价方法的应用现状[J]. 中国药房,2021,32(22):2813-2816.
- [44] 张强,王志飞,谢雁鸣,等. 中成药临床综合评价报告规范[J]. 中国中药杂志,2021,46(23):6062-6067.
- [45] 曾繁典,郑荣远,詹思延,等. 药物流行病学[M]. 2版. 北京:中国医药科技出版社,2016:1-16.
- [46] 钟婕,周英凤. 实施性研究的方法学及应用进展[J]. 中华护理杂志,2018,53(7):862-866.
- [47] 张鹤巍,郭晓晶,叶小飞,等. 观察性研究中混杂因素控制的敏感性分析方法[J]. 药物流行病学杂志,2022,31(4):240-247.
- [48] 郑轶,郭晓晶,许金芳,等. 观察性健康数据科学和信息学组织实用工具简介[J]. 药物流行病学杂志,2021,30(12):837-840,846.
- [49] 黎国威,李立康,刘颖欣,等. 观察性药物流行病学研究的P值相关问题探讨[J]. 中国药物警戒,2022,19(1):23-26.
- [50] 覃青连,李峤,颜星星,等. 四种因果图模型在观察性研究因果推断中的比较研究[J]. 中国卫生统计,2020,37(4):496-500.
- [51] 张惠文,邹强,张拥军. 罕见病病例报告写作规范专家共识[J]. 临床儿科杂志,2022,40(3):229-234.
- [52] 王小琴,陈耀龙,渠清源,等. 病例系列研究方法学质量评价工具解读[J]. 中国循证儿科杂志,2015,10(5):381-385.
- [53] 李业棉,赵芄,杨崧惠,等. 队列研究中纵向缺失数据填补方法的模拟研究[J]. 中华流行病学杂志,2021,42(10):1889-1894.
- [54] 覃青连,李峤,颜星星,等. 空间病例对照研究理论方法进展与趋势展望[J]. 中国卫生统计,2021,38(1):155-160.
- [55] 刘澳林,周莉,韩如雪,等. 中医领域病例对照研究方法学质量与报告质量的相关性分析[J]. 中医杂志,2020,61(7):594-600.
- [56] 翟静波,商洪才,李江,等. 单病例随机对照试验的统计分析方法[J]. 中国循证医学杂志,2017,17(4):494-496.
- [57] 王梅,王建华,张抗,等. 中医疗疗效评价队列研究的方法学质量评价[J]. 中医杂志,2016,57(16):1379-1383.
- [58] 王亮,吴晓枫,宋薇,等. 真实世界研究和RCT随机对照研究方法评价中西医结合干预慢性乙型肝炎疗效[J]. 浙江中医药大学学报,2020,44(10):980-985.
- [59] 焦明旭,张晓,刘迪,等. 倾向性评分匹配在非随机对照研究中的应用[J]. 中国卫生统计,2016,33(2):350-352.
- [60] HIGGINS J, GREEN S. Cochrane handbook for systematic reviews of interventions version 5.1.0: the cochrane collaboration[EB/OL]. [2025-07-10]. <http://www.cochrane-handbook.org>.
- [61] 杭州市药事管理质控中心药品评价专家组,浙江大学医学院附属杭州市第一人民医院. 心血管慢病药品临床综合评价杭州专家共识[J]. 中国药房,2022,33(9):1025-1030.

- [62] 指南与共识[J]. 中华消化外科杂志, 2022, 21(1): C71-C76.
- [63] 梁昌昊, 夏如玉, 黄子玮, 等. 共识法在中医临床实践指南及专家共识中的应用及方法学报告特征分析[J]. 中医杂志, 2022, 63(1): 26-34.
- [64] 戴泽琦, 徐思敏, 吴雪, 等. 多准则决策分析用于中成药临床综合评价的建议[J]. 中国中药杂志, 2022, 47(12): 3155-3160.
- [65] 黄倩倩, 王涛, 康琪, 等. 多准则决策分析在药物效益风险评价领域的应用研究进展[J]. 中国新药杂志, 2021, 30(22): 2069-2074.
- [66] 黄丽花, 李维嫣, 吴家园, 等. 卫生经济学评价方法在临床医学研究中的应用概况[J]. 中国医药指南, 2019(14): 58-60.
- [67] 弓孟春, 刘莉, 吴琛, 等. 传染病爆发流行期间的临床研究方法与卫生经济学考量[J]. 南方医科大学学报, 2020, 40(3): 353-357.
- [68] 潘岳松, 金奥铭, 王梦星. 临床研究样本量的估计方法和常见错误[J]. 中国卒中杂志, 2022, 17(1): 31-35.
- [69] 谢爱磊, 陈嘉怡. 质性研究的样本量判断: 饱和的概念、操作与争议[J]. 华东师范大学学报(教育科学版), 2021, 39(12): 15-27.
- [70] 张颖, 王俊慧, 胡焯胤, 等. 中医药临床研究中预试验样本量的确定[J]. 中医杂志, 2021, 62(4): 307-311.
- [71] 夏万元, 钟晓妮, 陶浩, 等. 参考值范围样本量估计中参数的适度性研究[J]. 中国卫生统计, 2019, 36(5): 644-648.
- [72] 刘丹, 周吉银. 临床科研项目样本量的要求[J]. 中国医学伦理学, 2019, 32(6): 716-718, 723.
- [73] 黄悦勤. 医学科研中随机误差控制和样本量确定[J]. 中国心理卫生杂志, 2015, 29(11): 874-880.
- [74] 陈平雁. 临床试验中样本量确定的统计学考虑[J]. 中国卫生统计, 2015, 32(4): 727-731, 733.
- [75] 沈炎娟. 以追踪方法学为指导的随访管理体系构建[J]. 中医药管理杂志, 2021, 29(21): 245-246.
- [76] 杨锋, 陈欣, 尤东方, 等. 基于现实世界研究中临床随访数据的两种动态预测建模方法的实证研究[J]. 中国临床医学, 2021, 28(5): 751-756.
- [77] ROCHE J S, CLERY M J, CARTER P M, et al. Tracking assault-injured, drug-using youth in longitudinal research: follow-up methods[J]. Acad Emerg Med, 2018, 25(11): 1204-1215.
- [78] 刘文贵. 临床病例的远期随访方法[J]. 中国病案, 2011, 12(6): 5-7.
- [79] 邓伟胜. 医疗服务随访的方法与要求[J]. 医院管理论坛, 2010, 27(7): 36-38.
- [80] 李天资, 黄赞松, 潘兴寿, 等. 临床随访资料的方差统计分析方法[J]. 右江医学, 2006, 34(6): 618-620.
- [81] 胡良平, 李长平. 生存分析方法在随访研究中的应用: I [J]. 中华耳鼻咽喉头颈外科杂志, 2006, 41(1): 73-76.
- [82] 贺宪民, 贺佳, 范思昌. 临床随访资料的多元分析方法[J]. 第二军医大学学报, 2001, 22(1): 83-85.
- [83] 韩如雪, 李庚, 陈贤坤, 等. 中医临床研究中“核心结局指标集”测量工具的选择[J]. 中药新药与临床药理, 2019, 30(6): 745-749.
- [84] 邱瑞瑾, 孙杨, 胡嘉元, 等. 临床研究中选择结局指标测量工具的方法[J]. 中国循证医学杂志, 2018, 18(2): 238-243.
- [85] PRINSEN C A C, VOHRA S, ROSE M R, et al. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set”: a practical guideline[J]. Trials, 2016, 17(1): 449.
- [86] MOKKINK L B, TERWEE C B, PATRICK D L, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study[J]. Qual Life Res, 2010, 19(4): 539-549.
- [87] 王丽, 王晓墨, 杨晨露. 基于医疗大数据开展药品上市后监测与评价研究的偏倚及其控制[J]. 中国药物警戒, 2022, 19(3): 239-243.
- [88] 柏柳安宁, 夏结来, 王陵, 等. 真实世界研究中的常见偏倚及其控制[J]. 中国临床药理学与治疗学, 2020, 25(12): 1422-1428.
- [89] 刘新辉. 同时控制测量偏倚和混杂偏倚的工具变量方法研究[D]. 济南: 山东大学, 2020.
- [90] 张彤, 何文婷, 费宇彤, 等. 干预性队列研究的偏倚控制及其在中医领域的特殊性[J]. 中医杂志, 2019, 60(11): 923-927.
- [91] 游蓉丽, 李兵, 党海霞, 等. 药品上市价值评估方法研究文献综述[J]. 药物流行病学杂志, 2016, 25(10): 658-663.
- [92] 鲁路, 倪世豪, 黄育生, 等. 探讨中药人用经验的临床数据类型和统计分析方法[J]. 中国中药杂志, 2022, 47(13): 3681-3685.
- [93] 刘敏. 基于统计分析与张量分解的图像去噪算法研究[D]. 南京: 南京大学, 2021.
- [94] TURNER L, SHAMSEER L, ALTMAN D G, et al. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals: a cochrane review[J]. Syst Rev, 2012, 1: 60.
- [95] TAN Z W, TAN A C, LI T, et al. Has the reporting quality of published randomised controlled trial protocols improved since the SPIRIT statement: a methodological study[J]. BMJ Open, 2020, 10(8): e038283.
- [96] PANIC N, LEONCINI E, DE BELVIS G, et al. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses[J]. PLoS One, 2013, 8(12): e83138.
- [97] 陈耀龙. 医学研究报告规范的发展与简介[J]. 兰州大学学报(医学版), 2022, 48(1): 1-4.
- [98] CHAN A W, TETZLAFF J M, ALTMAN D G, et al. SPIRIT 2013: new guidance for content of clinical trial protocols[J]. Lancet, 2013, 381(9861): 91-92.

- [99] VANDENBROUCKE J P, VON ELM E, ALTMAN D G, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration [J]. *Int J Surg*, 2014, 12(12):1500-1524.
- [100] BENCHIMOL E I, SMEETH L, GUTTMANN A, et al. The reporting of studies conducted using observational routinely-collected health data (RECORD) statement[J]. *PLoS Med*, 2015, 12(10):e1001885.
- [101] 聂晓璐, 武泽昊, 赵厚宇, 等. 使用常规收集医疗卫生数据开展观察性研究的报告规范(药物流行病学版) RECORD-PE 规范中文版: 上[J]. *药物流行病学杂志*, 2019, 28(3):190-198, 212.
- [102] 聂晓璐, 武泽昊, 赵厚宇, 等. 使用常规收集医疗卫生数据开展观察性研究的报告规范(药物流行病学版) RECORD-PE 规范中文版: 下[J]. *药物流行病学杂志*, 2019, 28(4):263-272.
- [103] GAGNIER J J, KIENLE G, ALTMAN D G, et al. The CARE guidelines: consensus-based clinical case reporting guideline development[J]. *BMJ Case Rep*, 2013, 2013:bcr2013201554.
- [104] 于河, 杨红, 刘建平. 专家临证验案与经验的报告方法: 病例系列研究的设计和评价[J]. *中医杂志*, 2008, 49(5):407-410.
- [105] FU S F, CHENG C W, ZHANG L, et al. Consensus-based recommendations for case report in Chinese medicine (CARC)[J]. *Chin J Integr Med*, 2016, 22(1):73-79.
- [106] ZWARENSTEIN M, TREWEEK S, GAGNIER J J, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement[J]. *BMJ*, 2008, 337:a2390.
- [107] DWAN K, LI T J, ALTMAN D G, et al. CONSORT 2010 statement: extension to randomised crossover trials[J]. *BMJ*, 2019, 366:14378.
- [108] VOHRA S, SHAMSEER L, SAMPSON M, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015 statement[J]. *BMJ*, 2015, 350:h1738.
- [109] GAGNIER J J, BOON H, ROCHON P, et al. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement[J]. *Ann Intern Med*, 2006, 144(5):364-367.
- [110] CHENG C W, WU T X, SHANG H C, et al. CONSORT extension for Chinese herbal medicine formulas 2017: recommendations, explanation, and elaboration: traditional Chinese version[J]. *Ann Intern Med*, 2017, 167(2):W7-W20.
- [111] LI J, HU J Y, ZHAI J B, et al. CONSORT extension for reporting N-of-1 trials for traditional Chinese medicine (CENT for TCM): recommendations, explanation and elaboration[J]. *Complement Ther Med*, 2019, 46:180-188.
- [112] KWAKKENBOS L, IMRAN M, MCCALL S J, et al. CONSORT extension for the reporting of randomised controlled trials conducted using cohorts and routinely collected data (CONSORT-ROUTINE): checklist with explanation and elaboration[J]. *BMJ*, 2021, 373:n857.
- [113] VLAHOV D. Transparent reporting of evaluations with nonrandomized designs (TREND) [J]. *J Urban Health*, 2004, 81(2):163-164.
- [114] PAGE M J, MCKENZIE J E, BOSSUYT P M, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews[J]. *BMJ*, 2021, 372:n71.
- [115] ZHANG X, TAN R, LAM W C, et al. PRISMA (preferred reporting items for systematic reviews and meta-analyses) extension for Chinese herbal medicines 2020 (PRISMA-CHM 2020)[J]. *Am J Chin Med*, 2020, 48(6):1279-1313.
- [116] WANG X Q, ZHOU Q, CHEN Y L, et al. Protocol of reporting items for public versions of guidelines: the reporting tool for practice guidelines in health care: public versions of guidelines[J]. *BMJ Open*, 2019, 9(3):e023147.
- [117] HUSEREAU D, DRUMMOND M, AUGUSTOVSKI F, et al. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations[J]. *Value Health*, 2022, 25(1):3-9.
- [118] MA L L, WANG Y Y, YANG Z H, et al. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better?[J]. *Mil Med Res*, 2020, 7(1):7.
- [119] MUCKART D J J. Evidence-based medicine: are we boiling the frog?[J]. *S Afr Med J*, 2013, 103(7):447-448.
- [120] MULROW C D. The medical review article: state of the science[J]. *Ann Intern Med*, 1987, 106(3):485-488.
- [121] HIGGINS J P T, ALTMAN D G, GÖTZSCHE P C, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials[J]. *BMJ*, 2011, 343:d5928.
- [122] STERNE J A C, SAVOVIĆ J, PAGE M J, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials [J]. *BMJ*, 2019, 366:14898.
- [123] IBBOTSON T, GRIMSHAW J, GRANT A. Evaluation of a programme of workshops for promoting the teaching of critical appraisal skills[J]. *Med Educ*, 1998, 32(5):486-491.
- [124] PURSELL E. Can the Critical Appraisal Skills Programme check-lists be used alongside grading of recommendations assessment, development and evaluation to improve transparency and decision-making? [J]. *J Adv Nurs*, 2020, 76(4):1082-1089.
- [125] TAYLOR R, REEVES B, EWINGS P, et al. A systematic review of the effectiveness of critical appraisal skills training for clinicians[J]. *Med Educ*, 2000, 34(2):120-125.
- [126] JADAD A R, MOORE R A, CARROLL D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary?[J]. *Control Clin Trials*, 1996, 17(1):1-12.

- [127] SCHULZ K F. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials[J]. *Jama*, 1995, 273(5):408.
- [128] HARTLING L, OSPINA M, LIANG Y Y, et al. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study[J]. *BMJ*, 2009, 339:b4012.
- [129] SLIM K, NINI E, FORESTIER D, et al. Methodological index for non-randomized studies (minors): development and validation of a new instrument[J]. *ANZ J Surg*, 2003, 73(9):712-716.
- [130] STERNE J A, HERNÁN M A, REEVES B C, et al. ROBINS- I : a tool for assessing risk of bias in non-randomised studies of interventions[J]. *BMJ*, 2016, 355:i4919.
- [131] WELLS G, SHEA B, O'CONNELL D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses[EB/OL]. [2025-07-13]. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
- [132] STANG A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses[J]. *Eur J Epidemiol*, 2010, 25(9):603-605.
- [133] DOWNES M J, BRENNAN M L, WILLIAMS H C, et al. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS) [J]. *BMJ Open*, 2016, 6(12):e011458.
- [134] DRUMMOND M F. Users' guides to the medical literature: XIII : how to use an article on economic analysis of clinical practice A: are the results of the study valid?[J]. *Jama*, 1997, 277(19):1552.
- [135] VARDELL E, MALLOY M. Joanna briggs institute: an evidence-based practice database[J]. *Med Ref Serv Q*, 2013, 32(4):434-442.
- [136] HANNES K, LOCKWOOD C. Pragmatism as the philosophical foundation for the Joanna Briggs meta-aggregative approach to qualitative evidence synthesis[J]. *J Adv Nurs*, 2011, 67(7):1632-1642.
- [137] SACKS H S, BERRIER J, REITMAN D, et al. Meta-analyses of randomized controlled trials[J]. *N Engl J Med*, 1987, 316(8):450-455.
- [138] SHEA B J, GRIMSHAW J M, WELLS G A, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews[J]. *BMC Med Res Methodol*, 2007, 7:10.
- [139] SHEA B J, REEVES B C, WELLS G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both[J]. *BMJ*, 2017, 358:j4008.
- [140] WHITING P, SAVOVIĆ J, HIGGINS J P T, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed[J]. *J Clin Epidemiol*, 2016, 69:225-234.
- [141] VLAYEN J, AERTGEERTS B, HANNES K, et al. A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit [J]. *Int J Qual Health Care*, 2005, 17(3):235-242.
- [142] CATTANI A, TEIXEIRA P P, DA CONCEIÇÃO ECKERT I, et al. Quality appraisal of clinical nutrition practice guidelines for critically ill adult patients: a systematic review using the advancing guideline development, reporting and evaluation in health care instrument II (AGREE-II) and AGREE-recommendation excellence (AGREE-REX)[J]. *Br J Nutr*, 2023, 129(1):66-76.
- [143] HAILEY D. Toward transparency in health technology assessment: a checklist for HTA reports[J]. *Int J Technol Assess Health Care*, 2003, 19(1):1-7.
- [144] Critical Appraisal Skills Programme. CASP (case control study) checklist[EB/OL]. (2019-01-01)[2025-07-14]. <https://casp-uk.net/casp-tools-checklists/>.
- [145] COLES B, TYRER F, HUSSEIN H, et al. Development, content validation, and reliability of the assessment of real-world observational studies (ArRoWS) critical appraisal tool[J]. *Ann Epidemiol*, 2021, 55:57-63.e15.
- [146] ZHANG J, WANG Y Y, WENG H, et al. Management of non-muscle-invasive bladder cancer: quality of clinical practice guidelines and variations in recommendations[J]. *BMC Cancer*, 2019, 19(1):1054.
- [147] HOOIJMANS C R, ROVERS M M, DE VRIES R B M, et al. SYRCLE's risk of bias tool for animal studies[J]. *BMC Med Res Methodol*, 2014, 14:43.
- [148] LI B H, YU Z J, WANG C Y, et al. A preliminary, multi-center, prospective and real world study on the hemostasis, coagulation, and safety of Hemocoagulase Bothrops Atrox in patients undergoing transurethral bipolar plasmakinetic prostatectomy[J]. *Front Pharmacol*, 2019, 10:1426.
- [149] MOONS K G M, WOLFF R F, RILEY R D, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration [J]. *Ann Intern Med*, 2019, 170(1):W1-W33.
- [150] GOLD C, ERKKILÄ J, CRAWFORD M J. Shifting effects in randomised controlled trials of complex interventions: a new kind of performance bias?[J]. *Acta Psychiatr Scand*, 2012, 126(5):307-314.
- [151] 石秀园,赵锐,李璠,等. 构建我国药品临床综合评价工作机制的思考[J]. *中国药房*, 2020, 31(23):2828-2833.
(收稿日期:2025-07-30 修回日期:2025-09-08)
(编辑:刘明伟)