

团 体 标 准

T/CLPA XXXX—2025

药品临床综合评价工作流程指南
(2025 版)

Guideline for the workflow of comprehensive clinical evaluation of drugs

(征求意见稿)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

2025 - XX - XX 发布

2025 - XX - XX 实施

中国药师协会 发布

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 缩略语.....	2
5 总则	2
6 需考虑因素.....	3
6.1 药品临床综合评价工作流程内容的遴选和确定.....	3
6.2 证据分级与依据.....	3
6.3 推荐意见的形成.....	3
7 关键问题与推荐意见.....	4
7.1 主题遴选.....	4
7.2 评价实施.....	7
7.3 证据评价.....	17
7.4 结果应用转化.....	18
附录 A（规范性） 图 A.1 药品临床综合评价主题遴选评审流程.....	19
附录 A（规范性） 表 A.1 药品临床综合评价主题遴选评估指标体系.....	20
附录 A（规范性） 表 A.2 药品临床综合评价主题遴选评估指标体系.....	21
附录 B（资料性） 图 B.1 常见统计学方法.....	23
参考文献.....	24

前 言

本文件按照GB/T1.1-2020《标准化工作导则 第1部分：标准的结构和编写》、GB/T 20001.7-2017《标准编写规则 第7部分：指南标准》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国药师协会药品临床评价工作委员会提出。

本文件由中国药师协会归口。

本文件起草单位：中国药师协会药品临床评价工作委员会；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。

本文件主要起草人：李正翔（天津医科大学总医院）

（以下按姓名笔画排序）：于倩（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第一医院）、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（天津市卫生健康委员会）、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员会）、董梅（哈尔滨医科大学附属肿瘤医院）。

本文件指导专家（按姓名笔画排序）：

史录文（北京大学）、田金徽（兰州大学基础医学院循证医学研究所）、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）

本文件主要执笔人：段蓉（天津医科大学总医院）

文件工作组：张铭慧（天津医科大学总医院）、谢栋（天津医科大学总医院）、吴佳瑶（天津医科大学总医院）、孙蕾（天津医科大学总医院）、张子健（天津医科大学总医院）

本文件为首次发布。

引 言

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，是健全药品供应保障制度的具体要求^[1-2]。因此，药品临床综合评价工作流程的标准化、规范化、同质化、科学化，可极大提高研究结果的稳定性和可靠性。科学规范开展药品临床综合评价工作，可助力提升药学服务质量和药事管理能力，有效保障临床基本用药的供应与合理使用，更好地满足临床诊疗需求，并为推进国家药物政策决策提供坚实的基础和依据。

近年来，我国相继发布多项政策文件，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。2019年4月发布了《国家卫生健康委关于开展药品使用监测和临床综合评价工作的通知》^[1]、2021年7月发布了《国家卫生健康委办公厅关于规范开展药品临床综合评价工作的通知》^[2]，同时配发《药品临床综合评价管理指南（2021年版 试行）》（以下简称《管理指南》）。

为使相关文件指导意见能够规范落实在工作实践中，特编制了《药品临床综合评价工作流程指南》标准，旨在进一步规范药品临床综合评价工作的流程及内容，明确评价维度，加强研究证据的评价及研究成果转化应用。

药品临床综合评价工作流程指南

(2025 版)

1 范围

本文件确立了药品临床综合评价工作的主要流程及相关技术环节。

本文件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。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

国家卫生健康委《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)

国家卫生健康委办公厅《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)

国家卫生健康委办公厅《药品临床综合评价管理指南》(2021年版 试行)

国家药物和卫生技术综合评估中心(“国家卫生健康委卫生发展研究中心”)《抗肿瘤药品临床综合评价技术指南》(2022年版 试行)《心血管病药品临床综合评价技术指南》(2022年版 试行)《儿童药品临床综合评价技术指南》(2022年版 试行)《药品临床综合评价质量控制指南》(2024年版 试行)》

3 术语和定义

下列术语和定义适用于本文件。

3.1

药品临床综合评价

对已上市药品开展非注册目的的研究,通过系统整合药品临床应用与供应保障全链条的多维度证据,形成以人群健康获益为核心的药品临床价值评估结果。

3.2

循证医学

遵循证据的医学,又称实证医学,是一种医学诊疗方法,强调应用完善设计与执行的研究(证据)将决策最佳化。

3.3

真实世界研究

在真实世界环境下收集与患者有关的数据,通过分析,获得医疗产品的使用价值及潜在获益或风险的临床证据。

4 缩略语

下列缩略语适用于本文件。

ADL	日常生活能力量表 (Activity of Daily Living Scale)
AGREE	指南研究与评价工具 (Appraisal of Guidelines for Research and Evaluation)
CARC	基于共识的中医病例报告规范 (Consensus-based recommendations for case report in Chinese medicine)
CARE	病例报告指南 (Case REport guideline)
CBA	成本-效益分析 (Cost-Benefit Analysis)
CEA	成本-效果分析 (Cost-Effectiveness Analysis)
CHEERS	综合健康经济评估报告标准 (consolidated health economic evaluation reporting standards)
CMA	最小成本分析 (Cost-minimization analysis)
CONSORT	实用性随机对照试验报告规范 (Consolidated standards of reporting trials)
COSMIN	选择健康测量工具的共识标准 (Consensus-Based Standards For The Selection of Health Measurement Instruments)
CUA	成本-效用分析 (Cost-Utility Analysis)
ICD	《国际疾病分类》 (International Classification of Diseases)
IOM	美国国家科学院医学研究所 (Institute of Medicine)
PRISMA	系统评价和 Meta 分析优先报告的条目 (Preferred Reporting Items for Systems Reviews and Meta-Analyses)
QALY	质量调整生命年 (Quality-adjusted life year)
RCTs	随机对照试验 (randomized controlled trials)
RECORD	常规收集医疗卫生数据开展观察性研究的报告规范 (reporting of studies conducted using observational routinely collected health data)
RIGHT	国际实践指南报告规范 (Reporting Items for Practice Guidelines in healthcare)
SPIRIT	临床试验方案规范指南 (Standard Protocol Items: Recommendations for Interventional Trials)
STROBE	观察性流行病学研究的报告指南 (STrengthening the Reporting of OBServational studies in Epidemiology)
TREND	非随机对照试验报告规范 (transparent reporting of evaluations with nonrandomized designs)
WHO	世界卫生组织 (World Health Organization)

5 总则

《药品临床综合评价工作流程指南》的制定遵循美国国家科学院医学研究所 (IOM) 关于指南的最新定义, 以世界卫生组织 (WHO) 标准指南制定的方法学为依据, 并符合指南研究与评价工具 (AGREE II) 的六大领域。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。

本指南涵盖药品临床综合评价主题遴选、评价实施、证据评价和结果应用转化 4 个方面工作内容，19 个关键技术环节，为相关技术人员规范开展药品临床综合评价提供科学依据。

6 需考虑因素

6.1 药品临床综合评价工作流程内容的遴选和确定

以《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，指南工作组初步拟定 4 大类关键问题框架，涵盖 19 个技术环节的指南研究内容；基于德尔菲法以在线问卷的形式对相关问题的重要性进行调研和评分。共收集汇总 34 位专家函评反馈，最终遴选出 4 个方面工作内容，19 个关键技术环节的指南研究问题，涵盖药品临床综合评价的整体流程和主要工作内容。

6.2 证据分级与依据

指南制定工作组针对最终纳入的关键问题，参考人群、干预、对照和结局原则进行检索，检索数据库包括指南数据库（National Guideline Clearinghouse、Guidelines International Network、Trip database、The National Institute for Health and Care Excellence 等），系统评价、卫生技术评估、卫生经济学评价研究数据库（NHS Economic Evaluation Database、The Cochrane Library、HTA 数据库等），文献数据库（PubMed、Embase、中国生物医学文献数据库、万方、中国期刊全文数据库和维普数据库），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官方网站）；并将百度等公共搜索引擎及 AI 检索工具作为补充检索。本指南为管理型指南，采用中国循证医学中心中国循证医学中心的证据质量分级体系，采用牛津大学循证医学中心制定的推荐强度分级标准，见表 1。

表 1 证据分级与推荐强度分级

证据质量级别	描述
I 级	系统评价、HTA、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
推荐强度分级	描述
A	推荐，证据及有效（证据等级为 1 级）
B	可推荐，证据有效（证据等级为 2 级和 3 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 4 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 5 级），只在较窄的范围内有效

6.3 推荐意见的形成

指南制订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（ESPEN）指南手册，通过德尔菲法进行专家投票形成共识意见。共识形成规则为：（1）同意率超过 75% 的建议被定义为最终推荐意见；（2）同意率低于 50% 的建议为未达成共识；（3）同意率 50%-75% 的建议，代表大多数人同意，但会导

致共识强度降级。

7 关键问题与推荐意见

7.1 主题遴选

药品临床综合评价工作需要投入大量的人力、时间、资源成本，而研究资源是相对有限的，亟待科学规范的主题遴选评估流程和方法，有效选取在临床用药实践中存在的迫切需要解决的关键问题，进行高效规范评价。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、主题遴选评估方法三个方面。

7.1.1 主题遴选组织管理

为保障药品临床综合评价主题遴选工作的高质开展，需组建具有代表性强、权威度高、专业覆盖全面的主题遴选专家委员会。合理组建专家委员会有利于提高主题遴选的科学性、规范性和透明度，同时有利于减少因专业角度差异导致的评估偏倚以及专业知识能力不匹配的情况。

7.1.1.1 主题遴选专家委员会组成

药品临床综合评价主题遴选专家委员会建议由6个部分人群组成：

- ①卫生行政管理/决策部门代表^[3-9, 11-12, 16-18, 26-27, 29]（证据等级：Ⅱ级；推荐强度：B）；
- ②卫生专业技术人员及学者（包括临床医学、药学、护理学、卫生管理、循证医学和卫生经济学、伦理学等）^[3-6, 8-9, 12, 17-20, 23, 26-30]（证据等级：Ⅰ级；推荐强度：A）；
- ③患者方^[4, 9, 12]（证据等级：Ⅲ级；推荐强度：B）；
- ④公众代表^[9, 13]（证据等级：Ⅲ级；推荐强度：B）；
- ⑤企业方及商业团体代表^[9, 12]（证据等级：Ⅲ级；推荐强度：B）；
- ⑥监管机构代表^[18, 26-27, 29-30]（证据等级：Ⅰ级；推荐强度：A）。

在特定情况下，以上各方人员如与研究内容和研究目的存在利益冲突，不推荐参加主题遴选专家委员会^[9]（证据等级：Ⅲ级；推荐强度：C）

7.1.1.2 主题遴选专家委员会工作职责

药品临床综合评价主题遴选专家委员会应具备以下四个方面工作职责：

- ①主题遴选评估指标的建立与确认^[4-7, 9, 11, 13, 17]（证据等级：Ⅲ级；推荐强度：B）；
- ②临床综合评价研究主题技术审查^[3-9, 11, 13, 17-30]（证据等级：Ⅰ级；推荐强度：A）；
- ③待遴选主题评估及排序^[3, 5, 8-9, 11-12, 16, 18-30]（证据等级：Ⅰ级；推荐强度：A）；
- ④对主题清单遴选结果进行专家论证^[3-9, 11-13, 16-30]（证据等级：Ⅰ级；推荐强度：A）。

7.1.2 主题遴选评审流程

药品临床综合评价主题遴选评审流程可分为以下四个步骤：主题征集、确定参与主题遴选专家委员会、确定主题遴选评估指标体系、主题排序与决策。详见附录图 A.1

7.1.2.1 主题征集^[3, 10-12, 14-23, 25-30]

可根据政策制定需求、药品使用监测结果、临床用药问题、专家意见、文献综述、药品生产供应需求等多种途径收集待遴选主题。待遴选主题来源渠道包括但不限于来自国家/省市医疗、医保、医药等管理部门以及涉及相关专业的机构和研究者、患者/公众等社会群体。（证据等级：Ⅰ级；推荐强度：A）

7.1.2.2 确定参与主题遴选的专家委员会^[3, 11-12, 16, 21, 26, 28-30]

专家委员会组成及职责参见“7.1.1.1 主题遴选专家委员会组成”“7.1.1.2 主题遴选专家委员会工作职责”。（证据等级：Ⅰ级；推荐强度：A）

7.1.2.3 确定主题遴选评估指标体系^[3, 10-12, 14-21, 26-28]

主题遴选评估指标体系参见“7.1.3.1 主题遴选评估指标体系”；权重赋值方法参见“7.1.3.2 主题遴选评估指标权重赋值”。根据评价目的，选择全部或部分评估指标构建本机构药品临床综合评价主题遴选评估指标体系，并进行量化，以对遴选主题进行定量评分和优先排序。（证据等级：I级；推荐强度：A）

7.1.2.4 主题排序与决策^[3, 10-12, 14-23, 25-30]

专家委员会根据评估分数和优先排序，对各待遴选主题进行讨论审议，并做出最终决策，确定主题。（证据等级：I级；推荐强度：A）

7.1.3 主题遴选评估方法

7.1.3.1 主题遴选评估指标体系

本指南基于德尔菲法，构建了包括5个一级指标、9个二级指标、17个三级指标的药品临床综合评价主题遴选评估指标体系，并采用循证医学方法，明确各指标的评估内容，详见附录表A.1。

7.1.3.1.1 必要性

(1) 卫生决策需求

① 对政策/医疗保健体系的重要性/影响：评估目标药品有潜力解决医疗保健决策方面的困境，或是否可以促进卫生改革^[3, 12, 15, 21]。（证据等级：I级；推荐强度：A）

② 评估结果的政策转化潜力：评估药品临床综合评价结果支持或转化为卫生政策的潜力^[3]。（证据等级：III级；推荐强度：C）

(2) 临床实践需求

评估的创新性和实践价值：评估结果是否可以利益相关方（医疗机构、卫生决策机构、制药行业等）的卫生技术临床实践的改变提供依据^[20]。（证据等级：I级；推荐强度：B）

7.1.3.1.2 重要性

(1) 健康需求

① 疾病负担：疾病负担定义为某一特定疾病造成的残疾程度、死亡率及发病率对患者及其家庭和整个社会的影响。主要通过目标人群规模、疾病死亡率/发病率及伤残调整寿命年进行整体评估^[3-4, 6, 10-12, 15, 17-18, 21, 27]。（证据等级：I级；推荐强度：A）

② 弱势群体影响：弱势群体定义为儿童（新生儿）、妊娠期或哺乳期妇女、65岁以上老人、患有特殊疾病或患有神经心理疾病患者^[3-4, 6, 10-12, 15, 17-18, 21, 27]。（证据等级：I级；推荐强度：A）

(2) 健康影响

① 降低患者死亡率/发病率潜力：降低患者死亡率/发病率潜力是指在医疗干预、治疗方案或预防措施的实施下，能够有效减少患者因疾病导致的死亡或发病的风险。这种潜力通常通过比较干预前后的死亡率或发病率来衡量^[8, 16, 27]。（证据等级：I级；推荐强度：A）

② 患者生存质量/疾病预后的改善：干预措施的临床效益及生活质量的改善分为5个级别，分别为：治愈，延长寿命和生活质量主要改善，延长寿命和生活质量轻微改善，生活质量主要改善，生活质量轻微改善^[7, 9, 12-13, 16, 26]。（证据等级：II级；推荐强度：B）

③ 与现有治疗相比的潜在临床获益：优于现有替代方案的潜力或是否可以替代临床现有的卫生技术方案。可应用“可替代技术的数量”或“缺乏替代技术”作为评估内容^[3, 5-7, 11, 13, 16-18, 27]。（证据等级：I级；推荐强度：A）

④ 安全性获益：安全指标按照严重程度可分为5级（1级：轻微；无症状或症状极轻；仅检验可发现；无需干预 2级：轻度的；需要局部的小的，或非侵入性的干预；在工具性日常生活能力量表（instrument ADL）上评分下降 3级：严重的或临床上明显的，但非立即致命的；需入院干预或延长住院时间；残障；在躯体生活自理量表（self careADL）上评

分下降 4 级：威胁生命的；需要紧急干预 5 级：由于不良反应死亡)^[4-7, 18, 27]。（证据等级：I 级；推荐强度：A）

（3）经济性

① 对社会和患者个人的经济负担：社会 and 患者个人的经济负担可描述为直接成本及政府财政保障。

直接成本包括直接医疗成本和直接非医疗成本。

从医疗机构角度来看，直接医疗成本包括与卫生服务采购相关的一次性成本，直接非医疗成本包括经常性成本（如维护和运营成本）等。

从患者角度看，主要为家庭每年因疾病或健康问题而产生的直接医疗成本及与就医相关但并不具有医疗性质的成本，如异地就医的交通和食宿成本等非医疗支出。

政府财政保障：目标卫生技术纳入政府卫生支付范畴时，与目前的已纳入的卫生技术相比，这项技术的政策变化对政府预算的潜在增量影响是否在可承受范围内或是否具有可持续性^[4, 9, 11-12, 15-18, 27]。（证据等级：I 级；推荐强度：A）

② 对成本-效果/效用/效益的影响：主要通过增量效果比评估药品等卫生服务获得每单位健康所需花费的比较成本，使决策者能够考虑投资于一种卫生干预措施与其他选择相比，可以得到多少增量健康效益。如国内尚未有相关的经济性证据，可以考虑利用其他国家针对目标技术的成本效果/效用/效益评价结果进行初步评估^[4, 7, 16, 27]。（证据等级：I 级；推荐强度：A）

7.1.3.1.3 可行性

（1）证据需求

① 证据可获得性和充分程度：所需证据的可获得性和充分性，应综合已发表的文献证据及原始临床资料证据，重点关注新技术/药物的新证据^[3, 7-8, 10-14, 19]。（证据等级：II 级；推荐强度：B）

② 证据质量：证据质量指试验设计产生无偏倚结果的可能性，反映在研究过程中结论真实性的一系列因素^[5, 7]。（证据等级：III 级；推荐强度：B）

（2）资源需求

当前是否可满足所需时间、资源和技术的要求：研究团队资质、人员组成、专业分布、工作经验及当前是否可满足所需时间、资源和技术的要求^[8, 24, 26]。（证据等级：I 级；推荐强度：A）

7.1.3.1.4 社会效益

社会、舆论关注程度：

① 对伦理、法律、社会心理方面的影响：影响医疗保健决策、结果或成本的大多数群体及弱势群体的兴趣或利益诉求；低收入群体、罕见病患者群体等在道德意义上是否比其他群体更应该得到卫生资源分配的倾斜^[3, 6, 8-9, 14, 17, 26]。（证据等级：II 级；推荐强度：A）

② 个人/公共利益：目标主题应对个人/公益利益产生影响，具有社会争议度或引起社会关注^[4, 10, 12, 14]。（证据等级：III 级；推荐强度：C）

7.1.3.1.5 公平性

社会公平伦理影响：

药品使用对伦理和公平可及性的影响：社会公平伦理影响，定义为药品使用对伦理和公平可及性的影响，包括利益相关性冲突管理。旨在解决健康资源方面的不公平不平等问题，以确保每个人都有公平的机会从卫生系统中收益。主要通过评估干预措施是否以及如何促进公平或加剧不平等来解释卫生技术的公平性^[4, 9, 12]。（证据等级：III 级；推荐强度：C）

7.1.3.2 主题遴选评估指标权重赋值

药品临床综合评价主题遴选指标体系权重赋值方法推荐如下：

- ① 层次分析法^[4, 12, 17, 24]（证据等级：Ⅱ级；推荐强度：B）
- ② 多属性决策法^[5, 9, 11]（证据等级：Ⅲ级；推荐强度：B）
- ③ 德尔菲法^[8, 16]（证据等级：Ⅲ级；推荐强度：B）
- ④ 专家研讨会/小组讨论^[6, 17]（证据等级：Ⅲ级；推荐强度：B）
- ⑤ 配对比较法^[7]（证据等级：Ⅲ级；推荐强度：C）
- ⑥ 简单加性加权法^[15]（证据等级：Ⅲ级；推荐强度：C）
- ⑦ 等权重算法^[15]（证据等级：Ⅲ级；推荐强度：C）

因此,本指南采用证据等级最高和推荐强度最强的层次分析法对药品临床综合评价主题遴选评估指标体系进行权重赋值,详见附录表 A.1。

7.2 评价实施

7.2.1 研究背景^[31, 32, 33, 34, 35, 36]

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必要性和基础。在研究背景部分宜介绍学科领域的基本情况(如待评药品相关适应证的疾病流行病学概况、医疗服务利用情况、经济负担情况);总结与研究主题直接相关的最新进展(主要干预措施及其有效性、安全性和经济性等的评价现状),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;突显本研究的必要性,提出本研究的价值(必要性和重要性)或创新点,为本研究的研究内容确定及研究方法设计铺垫基础。(证据等级:Ⅰ级;推荐强度:A)

7.2.2 研究目的^[32, 33, 34, 35]

研究目的是研究实施的具体目标和意义,可以是特定假说,也可以是待解决的问题。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。

研究目的可依据前期主题遴选过程中收集的相关证据及评价结果提出,应清晰、具体、简洁、可测。研究目的越明确具体,研究的针对性越强,越可保障其实施效果;同时也需考虑其可行性,即具有相应的资源和条件,保障研究目的的达成。

研究目的中应简明扼要的阐述“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。(证据等级:Ⅰ级;推荐强度:A)

7.2.3 研究设计^[32, 33, 34, 35, 37, 38]

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。

研究方法及分类是多样化的,一般情况下,按照认识论基础,研究设计可分为三大类:定性研究设计、定量研究设计和混合研究设计。

(1) 定性研究设计:了解现象或问题;包括扎根理论、现象学研究、行为研究、个案研究等。

(2) 定量研究设计:探索各变量之间的关系或检测模型理论;包括实验设计和非实验设计(如问卷调查等)。

(3) 混合研究设计:综合运用定性和定量的研究设计方法;包括顺序法、并行法和转换法。

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,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四个方面。

在具体研究设计的选择和构建过程中,需重点考虑三个方面:

- (1) 与研究问题/目的的匹配程度;
- (2) 研究设计的可实施性/可执行性;

(3) 研究结果的实际应用意义（应用方和应用场景）。

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临床综合评价研究设计。（证据等级：I 级；推荐强度：A）

7.2.3.1 评价内容

7.2.3.1.1 疾病类型及诊断标准^[32, 33, 34, 35, 39-43]

开展药品临床综合评价工作，需首先确立研究药物所针对的疾病类型，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。

疾病类型建议采用世界卫生组织发布的《国际疾病分类》（International Classification of Diseases, ICD）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。

疾病诊断的方法较多，其分类方法大致可分三种：

- (1) 根据获得临床资料的方法分类，包括症状诊断、体检诊断、实验诊断、超声诊断、心电图诊断、内窥镜诊断、放射性核素诊断、手术探查诊断、治疗诊断等；
- (2) 根据诊断的确切程度分类，包括疑似诊断、临时诊断、暂定诊断、临床诊断等；
- (3) 根据诊断内容分类，包括病因诊断、病理形态诊断、病理生理诊断等。

由于疾病诊断及分类较多，以及中西医对相同病症诊断方法可能存在巨大差异，对同一疾病的诊断标准往往并不能完全统一。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。

（证据等级：I 级；推荐强度：A）

7.2.3.1.2 研究对象及纳排标准^[23, 27, 32-36, 44-45]

无论是何种研究，均应明确研究对象（目标人群），并制定科学的纳入标准和排除标准。

研究对象：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数目，如需随访，则需明确随访方法。

纳排标准：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。

当研究对象（目标人群）存在较大的异质性时，可以根据研究需要开展亚组分析。

（证据等级：I 级；推荐强度：A）

7.2.3.1.3 干预措施^[32, 33, 34, 46-48]

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，还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。

（证据等级：I 级；推荐强度：A）

7.2.3.1.4 对照设置^[32-34, 47-54]

对比是流行病学研究方法的核心思想之一，可通过设置客观的对照来实现，而对照设置的合适与否直接决定了研究质量以及结果的可靠性。

(1) 对照方法的选择:对照方法可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。

① 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。包括安慰剂对照、阳性对照、不同剂量对照以及空白对照等方法。

② 外对照：对照组与干预组并非来自同一目标总体，或并非同期接受干预和对照措施（如单臂研究等），需要从其他试验（研究）或历史病例中选取特征相似的研究对象作为对照，即在研究之外选取接受不同干预措施的对比组。外对照可来源于既往已完成的 RCT 研究，

也可来源于回顾性/前瞻性的真实世界研究。

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) 对照药物的选择：应根据研究目的、研究类型及干预措施（主要研究药物）及评价维度选择对照药物（可选择一个或若干个对照药品）。本指南推荐以下两种对照药物选择方式。

① 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为自限性疾病。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。

② 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药品。

（证据等级：I级；推荐强度：A）

7.2.3.2 评价维度^[31, 35, 55]

药品临床综合评价是评价主体应用多种评价方法和工具开展的多维度、多层次证据的综合评判。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定量相结合的数据整合分析与综合研判。

药品临床综合评价的安全性、有效性、经济性是药品的技术评价属性；创新性、适宜性、可及性是药品的政策评价属性。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。

（证据等级：I级；推荐强度：A）

7.2.3.2.1 安全性评价^[31-34, 55-58]

(1) 评价维度定义：综合分析药品上市前后安全性信息结果。

(2) 评价内容：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。

(3) 评价报告依次内容为：

- 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 药学质量评价
- ④ 非临床安全性研究
- ⑤ 临床安全性研究结果及质量评价
- ⑥ 上市后药品不良反应/事件
- ⑦ 产品召回措施和实施情况
- ⑧ 预防用药差错/事故采取的警示措施

（证据等级：I级；推荐强度：A）

7.2.3.2.2 有效性评价^[31-34, 55-56, 59-61]

(1) 评价维度定义：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。

(2) 评价内容：核心指标主要包括生存时长和生命质量两大类。

① 生存时长相关指标包括：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 生命质量相关指标包括：健康相关生命质量和健康效用值，可用质量调整生命年（QALY）进行评价；

③ 在核心指标评价的基础上，还可以针对疾病主客观指标，包括但不限于临床有效

率等临床疗效指标，疾病实验室检查、影像检查等特异性指标以及患者主观感觉等指标进行评价。

(3) 评价报告依次内容为：

- ① 药理与药学特性
- 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价
- 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、共识、临床路径收录情况

(证据等级：I 级；推荐强度：A)

7.2.3.2.3 经济性评价^[18, 20, 23, 27, 31-34, 55-56, 62-65]

(1) 评价维度定义：综合运用流行病与卫生统计学、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。同时，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。

(2) 评价内容：根据药品决策的具体需求，可选择开展成本-效果分析（CEA）、成本-效用分析（CUA）、成本-效益分析（CBA）、最小成本分析（CMA）等。

(3) 评价报告依次内容为：

- 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价，包括临床研究及模型研究
- ③ 医疗保险目录收录情况

(证据等级：I 级；推荐强度：A)

7.2.3.2.4 创新性评价^[31-34, 55-56, 66-68]

(1) 评价维度定义：通过分析判断药品与参比药品满足临床需求程度、鼓励国产原研创新等情况，进行药品的创新性评价。

(2) 评价内容：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和推进国内自主研发等创新价值判断。

(3) 评价报告依次内容为：

- 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效率。
- 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

(证据等级：I 级；推荐强度：A)

7.2.3.2.5 适宜性评价^[31-34, 55-56, 69-70]

(1) 评价维度定义：评价药品技术特点适宜性和药品使用适宜性。

(2) 评价内容：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；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。

(3) 评价报告依次内容为：

- ① 药品技术适宜性

标签标注完整性
 药品说明书标注明确性
 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

② 药品使用适宜性

给药适宜性：药物用法、给药途径是否符合患者情况，用药是否对适应证，联合用药是否合理，用药是否准确排除禁忌证等。

用药适宜性：治疗方案、适用人群、给药间隔、给药周期等方面是否适宜。

（证据等级：I 级；推荐强度：A）

7.2.3.2.6 可及性评价^[32-34, 55-56, 61-65]

（1）评价维度定义：评价药品可获得性和可负担性。

（2）评价内容：

- ① 可获得性可由医疗机构药品配备使用情况或有无短缺情况等反映。
- 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。

（3）评价报告依次内容为：

- 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

（证据等级：I 级；推荐强度：A）

7.2.3.3 信息来源^[18, 23, 27-28, 32-36, 76-78]

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。全面的证据基础是评估过程的基础。来自多个来源的各种类型的证据可能会为评估提供信息。为确保综合评价结果的适当且可靠，部分资料信息获取来源及可获得的信息内容如下：

（1）文献数据库（包括指南数据库，系统评价、卫生技术评估、卫生经济学评价研究数据库，期刊文献数据库等）：各种研究文献，如卫生技术评估报告、指南/共识、循证研究、RCT 研究、真实世界研究、药物经济学评价、病例报告、基础试验研究、质量评价等。

（2）官方网站（包括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官方网站）：药品公开信息，如通报记录、药物警戒、政策属性等。

（3）书籍专著：该领域的基础知识、定义概念及研究进展等

（4）临床试验注册数据库：临床试验研究注册情况，包括试验基本信息，如试验名称、编号、注册日期及状态；试验设计信息，像研究类型、方法、样本量计算和分组情况；研究对象信息，涉及纳入标准、排除标准和退出标准；干预措施信息，有试验用药品或医疗器械详情及对照措施；还有主要研究者姓名、单位、联系方式和研究机构名称、地址等。

（5）生产企业：药品基础信息，如生产工艺、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。

（6）问卷调查：可根据待评价药品的特点和维度定向开展调查问卷，主要调研主观感受或认知程度等，可汇总专家或公众意见。

（7）临床资料：患者诊疗信息，如人口学信息、患者症状及诊断、疾病及治疗史、家族史、干预方式及随访结果等。

（8）数据库：卫生信息系统数据、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。

（9）灰色文献：未公开发表研究信息，可能包含了最新的研究成果和观点，能够研究提供新的研究思路 and 方向。

(10) 搜索引擎：公众/媒体关注度等补充数据。

(证据等级：I 级；推荐强度：A)

7.2.4 研究类型^[31, 35, 79-82]

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。（证据等级：I 级；推荐强度：A）

7.2.4.1 原始研究

原始研究以一手资料（原始资料）为研究基础，进行各项综合整理、分析研究，可以是否具有干预因素分为观察性研究和实验性研究。

7.2.4.1.1 观察性研究^[32-34, 80, 83-100]

观察性研究又称非实验性研究或对比研究，是在自然状态下对研究对象的特征进行观察、记录，并对结果进行描述和对比分析。该类研究不可设置人为干预处理，同时研究对象接受处理因素也不是随机化的。观察性研究根据有无对照组可分为描述性研究和分析性研究。

(1) 描述性研究

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、病例系列及生态学研究等。

① 横断面研究：又称现况调查，是指在一个特定的时间点或期间内对一个特定人群的各类情况进行调查研究，所收集的资料反应该时间断面的状态。主要用于了解药物使用的特点及与药物有关的事件分布特征，为制定合理的药物使用策略和进行效果考核提供依据。

② 病例报告：通过单个或少数病例的记录和描述，提供疾病表现、机制、诊断及治疗等方面的证据资料，主要用于描述药物干预与特殊事件（不良反应事件）相关性。病例报告由于研究对象的高选择性，一般只能形成警示信号和病因学假设，但不能用以因果关系的确定。

③ 病例系列：形式与病例报告类型，但多在 10 例以上，常为连续性病例的描述和总结。病例系列研究可收集多个单一暴露量因素病例，并对其结局指标进行描述评价；或收集相同结局的病例，再追溯其暴露因素；在一定程度上可进行因果关系验证；但因其无对照组，因果关系确证力度较弱，仍作为警示信号和病因学假设提出方法。

④ 生态学研究：又称相关性研究，是在群体水平上通过描述和比较不同群体中某因素的暴露情况与研究结局分布的一致性和差异性。生态学研究可分为生态比较研究和生态趋势研究两种类型，主要用于分析暴露因素与研究结局之间的关系（但因无法避免的混杂因素，可能会出现“生态学谬误”，因此只为病因学分析提供线索），而因果关系的确定仍需采用分析性研究或实验性研究方法。

(2) 分析性研究

分析性研究是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。

① 病例-对照研究：是指以现在患有某疾病的病患者为病例组，以未患该疾病但具有可比性的人群为对照组，通过询问、体检、实验室检查或复查病史等，搜集既往各种可能的危险因素的暴露史，测量各因素的暴露比例，通过统计分析，推断暴露因素与疾病之间的关系。主要用于检验病因假设和探讨疾病的危险因素、预后因素以及评价防治效果等。

② 队列研究：是指将人群按照是否暴露于某种因素及其暴露程度分为不同的亚组，追踪其各自的结局，比较不同亚组之间结局频率的差异，从而判定暴露因素与结局之间有无因果关联及关联程度大小的研究方法。主要用于确证暴露因素与结局指标的因果关系。

③ 巢式病例-对照研究：又称套叠式病例对照研究或队列内病例对照研究，是将病例对照研究和队列研究进行组合后形成的一种新的研究方法，即在对一个事先确定好的队列进行随访观察的基础上，再应用该病例对照研究（主要是匹配病例对照研究）的涉及思路进行

研究分析。相比队列研究，其因果推断与暴露因素和结局关联的时间顺序相符合，推断力度更强；同时，其统计效率和检验效率也高于病例对照研究，还可以计算临床结局发生率等。

（证据等级：I 级；推荐强度：A）

7.2.4.1.2 实验性研究^[32-35, 80, 101-104]

实验性研究是指研究者能够认为给予干预措施的研究，根据是否采用随机方法，分为随机对照研究和非随机对照研究。

（1）随机对照研究

随机对照研究（RCT）是指将研究对象随机分组，对不同组实施不同的干预，以对照效果的不同。随机对照研究具有能够最大程度避免研究设计及实施过程中可能出现的各种偏倚，平衡混杂因素，提高统计检验的有效性，被公认为是评价干预措施的金标准。

（2）非随机对照研究

虽然随机对照研究证据可靠性、真实性最佳，证据级别较高，但因其实施条件较为严格等因素限制，在药品临床综合评价工作中，很多研究不适宜或不必要开展随机对照研究，可考虑采用非随机对照研究。常用的非随机对照研究包括自身前后对照研究、交叉对照研究、序贯试验等。

① 自身前后对照研究：是指同一组病人先后接受两种不同的干预措施，以其中一种作为对照，比较两种疗效差别，以确定所考核干预措施的疗效。适用于慢性稳定或复发性疾病。自身前后对照研究的研究对象为同一组患者，以干预先后作为治疗组和对照组，具有良好的可比性，结果的可靠性亦高于不同病例组的前后对照研究。

② 交叉对照研究：是指对两组受试者使用两种不同的干预措施，然后相互交换处理措施，最后比较结果的试验方法。优点是每例病人先后接受试验组或对照组的的治疗，消除了不同个体间的差异。但不可应用于各种急性重症疾患、不能恢复到第一阶段治疗前状况的疾病以及不许可停止治疗让病情回到第一阶段的疾病等。

③ 序贯试验：试验样本数事先不固定，而是每试验一对研究对象后，立即分析，再决定下一步试验，直到可以判断出结果时即停止试验。序贯设计研究可避免盲目加大样本而造成浪费，适合临床工作的特点，计算亦较简便。缺点是仅适用于单指标的试验。

（证据等级：I 级；推荐强度：A）

7.2.4.2 二次研究^[28, 32-36, 80, 105-112]

二次研究是在已有研究基础（文献/资料）的基础上，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、经济学分析等。

（1）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，该研究方法是检索、获取、评价、综合证据的最佳方法，可分为定性评价和定量评价（Meta 分析），主要用于干预措施有效性/安全性的评价。

（2）指南/专家共识

临床实践指南是指与临床疾病的诊断和治疗有关的指南，是规范临床各种活动的依据。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的诊疗方案达成共识。临床指南的制定是基于系统评价在内的当前最新、最佳证据，总结不同干预措施的利弊之后形成能够为患者提供最佳诊疗服务的推荐意见，可作为指导临床实践的文件、更是确保医疗质量的重要手段。

（3）决策分析

决策分析是针对疾病和防治过程中风险与获益的不确定性，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三类：

- ① 确定型决策：如成本分析法
- ② 风险型决策：如期望值决策、贝叶斯决策等
- ③ 不确定型决策：如最大最小准则等

（证据等级：I级；推荐强度：A）

7.2.4.3 其他研究^[23, 28, 32-36, 113-115]

（1）卫生经济学研究

卫生经济学研究是指将经济学原理和方法应用于评价临床药物治疗过程，并指导制定合理的诊疗策略的研究方法。卫生经济学研究是对“可选措施的成本和结局”进行比较分析，是医疗决策和卫生技术评估中的重要成分。卫生经济学研究的常用方法包括：成本效果分析（CEA）、成本效用分析（CUA）、成本效益分析（CBA）和最小成本分析（CMA）。

① CEA：评价一定量的卫生资源（成本）后的个人健康产出，健康结果用非货币单位表示，多用于单一疾病、某一药品、某一干预措施、同一健康结果指标的评价。

② CUA：评价项目投入成本量和经质量调整的健康效益产出量，主要用于衡量卫生项目或治疗措施效率。

③ CBA：通过比较不同备选方案的全部预期成本和全部预期收益来评价各备选方案，追求效益大于成本的方案，可用于评价互相独立的方案、互相排斥的方案以及互相依赖的方案。

④ CMA：衡量和比较不同医疗干预措施成本的方法，只能用于比较具有相同收益且可以准确确定成本的干预措施。

（2）社会调查

社会调查是社会“调查”和“研究”的简称。社会调查是指人们为达到一定目的，有意识地通过对社会现象的考察、了解和分析、研究，来了解社会真实情况的一种自觉认识活动。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主观认知情况。

（证据等级：I级；推荐强度：A）

7.2.5 样本量^[36, 116-122]

样本量估算是指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。样本量的确定需既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，并便于研究具体实施。

（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 统计检验的水准（ α ，即第一类错误的概率），一般采用 0.05；

② 统计检验的效能（ $1-\beta$ ， β 为第二类错误的概率），通常 β 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 ：表示所比较的两个总体参数间的差别，误差 δ 越大，所需样本越小；

④ 总体相关信息：包括总体平均数、标准差或总体率等，可通过预实验、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/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）

7.2.6 随访时间^[20, 23, 123-131]

在医学随访中，应对各研究对象追踪观察，记录各时间点上时间的发生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。

（1）明确开始随访的时间：如开始治疗、入院时间、确诊时间等，一般应为可明确记录的时间或开始研究的时间。

（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。

③ 随访截止时间的设定应综合考虑所研究的疾病特征、主要结局指标类型及可执行性。

④ 用于分析的时间范围应足够长（至少满足能够判断药物治疗取得阶段性指征结果的周期且符合临床诊疗思维），以反映被比较干预措施之间成本、结局等指标的所有重要差异。

随访时间设置的适用性需符合研究设计和研究方法。

（证据等级：I级；推荐强度：A）

7.2.7 结局指标^[20, 23, 32-36, 48, 80-82, 132]

药品临床综合评价工作中，可根据研究针对的不同疾病或治疗领域设定针对性的评价核心指标，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相应核心结局指标集、高质量指南/专家共识、系统评价及随机对照研究等。

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。主要结局指标和次要结局指标的选用确认是辩证的，研究者也可根据研究目的、研究类型、研究过程合理确定。不同评价维度应优选公认的最佳评价指标，建议参考已发布的药品临床综合评价管理指南和技术指南，采用适宜的评价方法和恰当的数据来实施评价。

（证据等级：I级；推荐强度：A）

7.2.8 测量方法^[35, 133-136]

药品临床综合评价研究中，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COSMIN工作组制定的测量工具质量评价清单（COSMIN清单）进行评价。（证据等级：I级；推荐强度：A）

7.2.9 偏倚控制^[23, 32-34, 137-144]

（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，可分为三类，即：选择偏倚、信息偏倚和混杂偏倚。

① 选择偏倚：是指由于入选研究和未入院研究的对象特征差异而导致研究结果偏离真实情况所产生的偏倚。常见的选择偏倚包括转介偏倚、自我选择偏倚、渠道偏倚、存活者治疗偏倚、新发-现患病例偏倚、易感者损耗偏倚、检出症候偏倚等。

② 信息偏倚：又称观察偏倚，是指研究时再进行信息收集过程中所产生的系统误差，从而导致研究记过与真实结果不符。常见信息偏倚包括错分偏倚和非死亡时间偏倚。

③ 混杂偏倚：是指由于一个或多个外来因素的存在，歪曲了研究因素与研究结局之间的关联。根据混杂因素与暴露因素、结局的关联方向一致性，可分为正混杂和负混杂。

(2) 偏倚控制方法

- ① 选择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 主要协变量的验证
- 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，或者作为结论和总结的一部分进行报告。

（证据等级：I 级；推荐强度：A）

7.2.10 质量控制^[23, 32-36, 55, 76-80]

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。

(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。在数据处理的每一阶段必须采用质量控制且处理正确，以保证所有数据可靠。

(2) 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。

（证据等级：I 级；推荐强度：A）

7.2.11 统计分析^[23, 31, 35-36, 145-148]

(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。

(2) 统计分析：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。

(3) 根据研究方案、统计分析计划书和统计分析结果，应撰写统计分析报告。包括原始数据/资料的质量评价，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使用图表形式展现，以提高可读性和条理性。
- ② 统计检验结果应包括统计意义的水平、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。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。

常见的统计学方法，详见附录图 B.1

（证据等级：I 级；推荐强度：A）

7.2.12 研究结果^[23, 35-36]

明确研究结果的适用性和局限性。

- (1) 适用性：综合评价结果的适用人群（评价角度）及应用场景。
- (2) 局限性：综合评价实施过程中存在的方法学等方面的局限性。

（证据等级：I 级；推荐强度：A）

7.2.13 报告规范^[23, 35-36, 149-153]

对研究结果进行报告时，需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。（证据等级：I级；推荐强度：A）

（1）临床研究计划书的报告规范^[154]

临床研究计划书报告规范参照《临床试验方案规范指南》（SPIRIT）（证据等级：I级；推荐强度：A）

（2）观察性研究的报告规范^[155-161]

① 队列研究、病例-对照研究、横断面研究：参照《观察性流行病学研究的报告指南》（STROBE）

② 真实世界队列研究、真实世界病例-对照研究：参照《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）

③ 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

④ 病例报告：参照《病例报告指南》（CARE）

⑤ 中药病例报告：参照《中医病例报告建议条目》、《基于共识的中医病例报告规范》（CARC）

（证据等级：I级；推荐强度：A）

（3）干预性研究的报告规范^[162-169]

① 随机对照研究：参照《实用性随机对照试验报告规范》（CONSORT）

② 真实世界随机交叉研究：参照《CONSORT2010 声明：随机交叉试验扩展版》

③ 真实世界单病例随机对照研究：参照《单病例随机对照试验 CONSORT 扩展版》（CENT）

④ 中药随机对照研究：参照《草药干预措施随机对照试验报告规范》、《中药复方临床随机对照试验报告规范》（CONSORT-CHMFormulas）

⑤ 中药单病例随机对照研究：参照《中医药单病例随机对照试验 CONSORT 扩展版》（CENTforTCM）

⑥ 使用队列或收集日常数据的随机对照研究：参照《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）

⑦ 非随机对照研究：参照《非随机对照试验报告规范》（TREND）

（证据等级：I级；推荐强度：A）

（4）系统评价研究的报告规范^[170-171]

① 系统评价和 Meta 分析：参照《系统评价和 Meta 分析优先报告的条目》（PRISMA2020）

② 中草药系统评价和 Meta 分析：参照《中草药系统评价和 Meta 分析优先报告的条目》（PRISMA-CHM2020）

（证据等级：I级；推荐强度：A）

（5）卫生经济学研究报告规范^[172]

卫生经济学研究报告规范参照综合健康经济评估报告标准（CHEERS）声明。（证据等级：I级；推荐强度：A）

（6）临床实践指南报告规范^[173]

临床实践指南报告规范参照《国际实践指南报告规范》（RIGHT）。（证据等级：I级；推荐强度：A）

在药品临床综合评价工作评价实施完成后，需考察所有工作流程和工作环节的完成情况，并对所有相关资料留存备案。

7.3 证据评价^[31-35, 105, 174-207]

药品临床综合评价工作中，无论是自身完成的研究结果或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。

方法学质量评估工具根据不同研究类型特征进行研发，各自适用范围有所不同，在证据

的产生和应用时应选择最佳工具。

值得注意的是：所有的评价工具都是主观的，工具的应用会受到应用者的技能和知识水平的影响。因此，质量评估人员需具备必需的专业知识、严谨的学术态度，并至少由两名独立的评估人员独立评估、交叉检查，以最大程度的避免实施偏倚。

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。详见附录表 A.2（证据等级：I 级；推荐强度：A）

7.4 结果应用转化

7.4.1 证据应用与反馈^[1-2, 31-35, 208-209]

《药品临床综合评价管理指南》（2021 年版 试行）明确提出药品临床综合评价结果证据的应用范围。

- （1）医疗卫生机构药品采购与供应保障等；
- （2）推动医疗卫生机构用药目录遴选和上下级医疗卫生机构用药目录衔接，提高药学服务和安全合理用药水平；
- （3）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；
- （4）为完善国家药物政策提供参考；
- （5）丰富行业药品临床综合评价的实践，扩大文献证据储备；
- （6）推动科研领域对于药品临床综合评价理论及方法的深入探索。

药品临床综合评价结果的转化应用过程中，应注重对证据适用性、科学性、规范性等方面的评价和考察，及时反馈，不断改善。

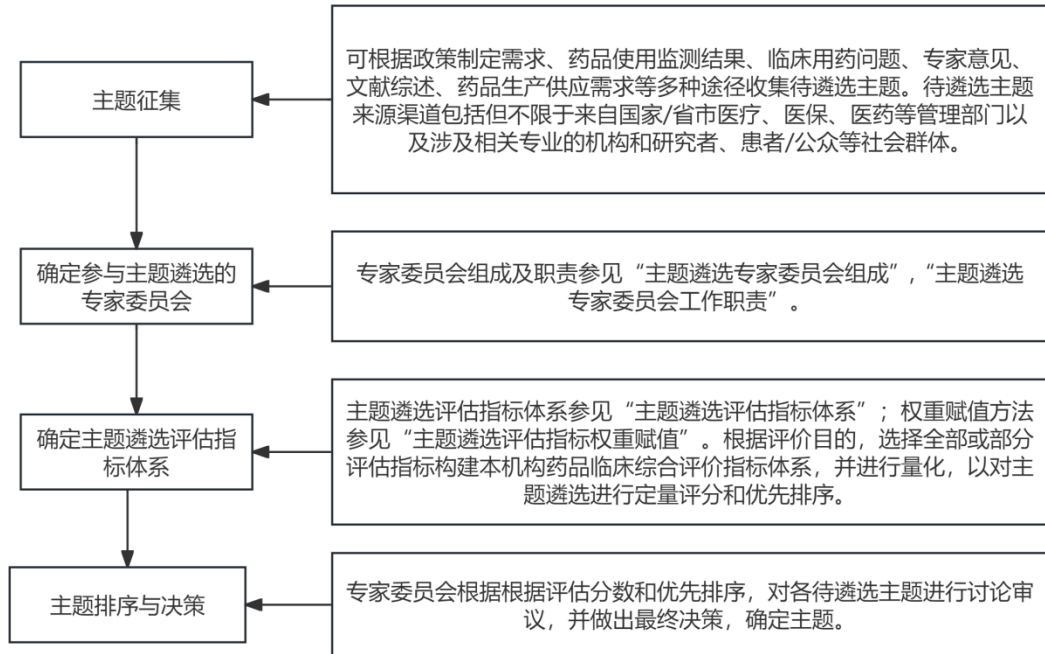
（证据等级：I 级；推荐强度：A）

7.4.2 评价结果优化与完善^[1-2, 31-35, 208-209]

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《药品临床综合评价管理指南》（2021 年版 试行）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，委托项目主要承担机构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，应当由有关部门及医疗卫生机构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，经牵头组织机构委托开展咨询论证后，确定是否启动有关工作。

（证据等级：I 级；推荐强度：A）

附录 A
(规范性)
药品临床综合评价主题遴选评审流程



附录图 A.1 药品临床综合评价主题遴选评审流程

附录 A

(规范性)

药品临床综合评价主题遴选评估指标体系

表 A.1 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
	临床实践需求	评估的创新性和实践价值	12.16
重要性	健康需求	疾病负担	10.51
		弱势群体影响	5.41
	健康影响	降低患者死亡率/发病率潜力	3.23
		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
		安全性获益	1.69
	经济性	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	
可行性	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、资源和技术的要求	6.34
社会效益	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
公平性	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

附录表 A.2 药品临床综合评价质量评估工具

研究类型				质量评价工具（量表、清单、条目）
原始研究	观察性研究	描述性研究	横断面研究	横断面研究评价工具（AXIS 工具）* NIH 质量评价工具 JBI 分析性横断面研究评估清单 AHRQ 方法学量表 Crombie 评估条目
			病例系列/病例报告	JBI 清单*
		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价工具 纽卡斯尔-渥太华量表（NOS）* 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-对照研究质量评价工具 JBI 病例-对照研究清单 NOS 病例-对照研究量表*
	其他观察性研究		CASP 量表 Cochrane 偏倚风险工具	
	实验性研究	随机对照研究		Cochrane 偏倚风险工具（RoB）* CASP 清单 Jadad 量表
		非随机对照研究		ROBINS-I 评价工具* JBI 清单

			MINORS 工具 EPOCRoB 工具 IHE 质量评价工具
		真实世界研究	ArRoWs 评价工具*
二次研究		系统评价/Meta 分析	AMSTAR-2 量表* 系统评价偏倚风险工具 (ROBIS) 工具* 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 分析质量评价工具 SIGN 系统评价与 Meta 分析评价清单
		网状 Meta 分析	网状 Meta 分析 (NMA) 方法学清单
		临床实践指南	指南研究与评价工具 (AGREEII) *
		卫生经济学评价	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
		HTA 报告	HTAchecklist 量表

注：*为目前最常用或最推荐的质量评估方法

附录 B
(资料性)
常见统计学方法

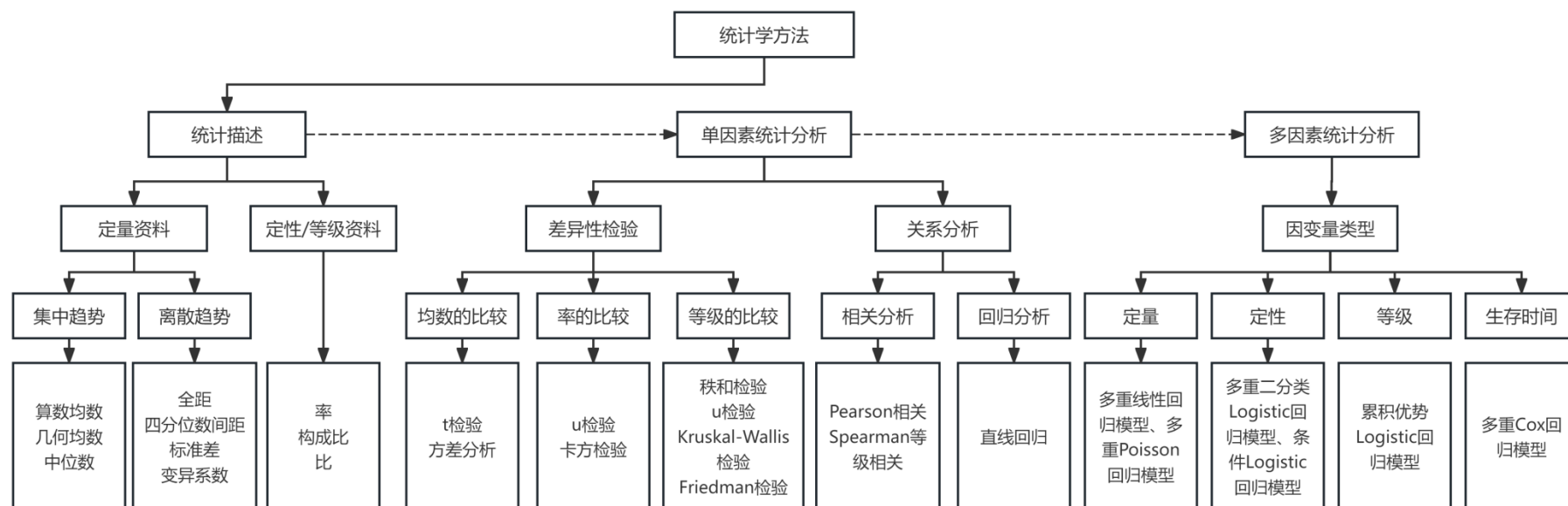


图 B.1 常见统计学方法

参 考 文 献

- [1] 国家卫生健康委办公厅. 国家卫生健康委办公厅关于规范开展药品临床综合评价工作的通知[EB/OL]. (2021-07-28) [2024-09-22].
<http://www.nhc.gov.cn/yaozs/s2908/202107/532e20800a47415d84adf3797b0f4869.shtml>
- [2] 国家卫生健康委. 国家卫生健康委关于开展药品使用监测和临床综合评价工作的通知[EB/OL]. (2019-04-09) [2024-09-22].
<http://www.nhc.gov.cn/yaozs/pqt/201904/31149bb1845e4c019a04f30c0d69c2c9.shtml>
- [3] 邱英鹏, 赵琨, 王晓玲, 等. 儿童药品临床综合评价主题的遴选方法初探[J]. 中国药物评价, 2021, 38(3):177-182.
- [4] Blaauw D, Chambers C, Chirwa T, et al. Introducing an Ethics Framework for health priority-setting in South Africa on the path to universal health coverage[J]. S Afr Med J, 2022, 112(3):240-244.
- [5] Mobinizadeh M, Mohamadi E, Arman H, et al. Topic selection for health technology assessment: An approach combining multiple attribute decision making and decision rules[J]. Med J Islam Repub Iran, 2021, 35:40.
- [6] Kosherbayeva L, Hailey D, Kurakbaev K, et al. A process of prioritizing topics for health technology assessment in Kazakhstan[J]. International Journal of Technology Assessment in Health Care, 2016, 32(3):147-151.
- [7] Mobinizadeh M, Raeissi P, Nasiripour A A, et al. A model for priority setting of health technology assessment: the experience of AHP-TOPSIS combination approach[J]. Daru, 2016, 24:10.
- [8] Jankauskiene D, Petronyte G. A model for hta priority setting: Experience in lithuania[J]. International Journal of Technology Assessment in Health Care, 2013, 29(4):450-455.
- [9] Youngkong S, Baltussen R, Tantivess S, et al. Multicriteria decision analysis for including health interventions in the universal health coverage benefit package in Thailand[J]. Value Health, 2012, 15(6):961-970.
- [10] Bastian H, Scheibler F, Knellingen M, et al. Choosing health technology assessment and systematic review topics: The development of priority-setting criteria for patients' and consumers' interests[J]. International Journal of Technology Assessment in Health Care, 2011, 27(4):348-356.
- [11] Husereau D, Boucher M, Noorani H. Priority setting for health technology assessment at CADTH[J]. International Journal of Technology Assessment in Health Care, 2010, 26(3):341-347.
- [12] Whitlock E P, Lopez S A, Chang S, et al. AHRQ series paper 3: identifying, selecting, and refining topics for comparative effectiveness systematic reviews: AHRQ and the effective health-care program[J]. J Clin Epidemiol, 2010, 63(5):491-501.
- [13] Menon D, Stafinski T. Engaging the public in priority-setting for health technology assessment: findings from a citizens' jury[J]. Health Expect, 2008, 11(3):282-293.
- [14] Carlsson P. Health technology assessment and priority setting for health policy

in Sweden[J]. INTERNATIONAL JOURNAL OF TECHNOLOGY ASSESSMENT IN HEALTH CARE, 2004, 20(1):44-54.

[15] Oortwijn W J, Vondeling H, Van Barneveld T, et al. Priority setting for health technology assessment in The Netherlands: Principles and practice[J]. Health Policy, 2002, 62(3):227-242.

[16] Shani S, Siebzehner M I, Luxemburg O, et al. Setting priorities for the adoption of health technologies on a national level the Israeli experience[J]. Health Policy, 2000, 54(3):169-185.

[17] Molla S, Donaldson And Harold C. Sox J. Setting Priorities for Health Technology Assessment: A Model Process[M]. Washington (DC): National Academies Press (US), 1992.

[18] Agency for Care Effectiveness(ACE). Process & Methods[EB/OL]. [2023-07-08]. <https://www.ace-hta.gov.sg/resources/process-methods>.

[19] Agency for Healthcare Research and Quality(AHRQ). EHC Program Topic Nomination and Selection | Effective Health Care (EHC) Program[EB/OL]. [2023-07-08]. <https://effectivehealthcare.ahrq.gov/about/epc/nomination>.

[20] The Canadian Agency for Drugs and Technologies in Health(CADTH). Topic Identification and Prioritization Process[EB/OL]. [2023-07-13]. https://www.cadth.ca/sites/default/files/pdf/HTA_OU_Topic_ID_and_Prioritization_Process.pdf.

[21] European Network for Health Technology Assessment(EUnetHTA). Topic Identification, Selection and Prioritisation - EUnetHTA[EB/OL]. [2023-07-07]. <https://www.eunethta.eu/ja3services/horizon-scanning/>.

[22] Health Intervention and Technology Assessment Program(HITAP). Health Technology Assessment Process Guidelines[EB/OL]. [2023-07-13]. <https://www.hitap.net/en/documents/21282>.

[23] Health Technology Wales(HTW). Our Appraisal Process - Health Technology Wales[EB/OL]. [2023-07-10]. <https://healthtechnology.wales/about/our-appraisal-process/>.

[24] Institute for Clinical & Economic Review(ICER). A Guide to ICER's Methods for Health Technology Assessment[EB/OL]. https://icer.org/wp-content/uploads/2021/01/ICER_HTA_Guide_102720.pdf

[25] German Drug Evaluation Agency(IQWiG). ThemenCheck Medizin: Thema vorschlagen | IQWiG.de[EB/OL]. [2023-07-10]. <https://www.iqwig.de/sich-einbringen/themencheck-medizin/formular-themenvorschlag/>.

[26] Belgian Health Care Knowledge Centre(KCE). Topic selection for the annual programme | KCE[EB/OL]. [2023-07-14]. <https://kce.fgov.be/en/submit-a-topic/topic-selection-for-the-annual-programme>.

[27] Health Technology Assessment Section(MAHTAS). Health Technology Assessment[EB/OL]. [2023-07-14]. https://www.moh.gov.my/moh/resources/HTA_MANUAL_MAHTAS.pdf.

[28] National Institute for Health and Care Excellence(NICE). About this guide | NICE health technology evaluation topic selection: the manual | Guidance |

- NICE[EB/OL]. [2023-07-10].
<https://www.nice.org.uk/process/pmg37/chapter/about-this-guide>.
- [29] Swedish Agency for Health Technology Assessment and Assessment of Social Services(SBU). Selection of topics[EB/OL]. [2023-07-10].
<https://www.sbu.se/en/about-sbu/selection-of-topics/>.
- [30] Swiss Federal Office of Public Health(SFOPH). HTA programme[EB/OL]. [2023-07-10].
<https://www.bag.admin.ch/bag/en/home/versicherungen/krankenversicherung/krankenversicherung-leistungen-tarife/hta/hta-programm.html>.
- [31] 国家药物和卫生技术综合评估中心（国家卫生健康委卫生发展研究中心）、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]. [2023-07-10]. <http://www.nhc.gov.cn/cms-search/downFiles/863665c55fd74f29a656990d1a4ea5ce.pdf>
- [32] 国家药物和卫生技术综合评估中心（国家卫生健康委卫生发展研究中心）、国家心血管病中心、首都医科大学附属北京安贞医院、国家卫生健康委药具管理中心. 心血管病药品临床综合评价技术指南（2022 年版 试行）[EB/OL]. [2023-07-10]. <http://cpia.net.cn/uploads/20220630/59831d29ddfe35745066ea54956a9b0f.pdf>
- [33] 国家药物和卫生技术综合评估中心（国家卫生健康委卫生发展研究中心）、国家儿童医学中心、国家卫生健康委药具管理中心. 儿童药品临床综合评价技术指南（2022 年版 试行）[EB/OL]. [2023-07-10]. <http://cpia.org.cn/uploads/20220630/eef00a18bd641c9b3bbf213193df2730.pdf>
- [34] 国家药物和卫生技术综合评估中心（国家卫生健康委卫生发展研究中心）、国家癌症中心、国家卫生健康委药具管理中心. 抗肿瘤药品临床综合评价技术指南（2022 年版 试行）[EB/OL]. [2023-07-10]. <https://www.hnysfw.com/data/article/1656481023744041713.pdf>
- [35] 国家药物和卫生技术综合评价中心. 药品临床综合评价质量控制指南(2024 年版 试行)[EB/OL]. [2023-07-10]. <http://nhei.cn/nhei/zxfb/202412/d33db3fcddf94daeabe49cfa95fa24f9.shtml>
- [36] Cuschieri S. The STROBE guidelines. Saudi J Anaesth[J]. 2019 Apr;13(Suppl 1):S31-S34.
- [37] Cohen L, Manion L, Morrison K. Research methods in education (8thed) [M]. London:Routledge, 2017:5-17.
- [38] 崔延强主译. 研究设计与协作指导:定性、定量与混合研究的路径[M]. 重庆:重庆大学出版社, 2007:3-19.
- [39] 万学红, 卢雪峰. 诊断学[M]. 北京:人民卫生出版社, 2018:1, 89, 237.
- [40] 尚红, 王兰兰. 实验诊断学[M]. 北京:人民卫生出版社. 2015:2-8.
- [41] 黄正新, 王召亿, 李鑫辉. 冠心病的中西医诊断思维差异认识的比较分析[J]. 陕西中医, 2017, 38(10):1374-1375.
- [42] 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(09):972-975.
- [43] 关敬高. 中西医诊断方法上的差异[J]. 医学与哲学, 1997, (07):56.
- [44] 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(03):223-227.

[45] Cheng A, Kessler D, Mackinnon R, et al. International Network for Simulation-based Pediatric Innovation, Research and Education (INSPIRE) Reporting Guidelines Investigators. Reporting guidelines for health care simulation research: Extensions to the CONSORT and STROBE statements. *BMJ Simul Technol Enhanc Learn*. 2016 Jul 24;2(3):51-60.

[46] 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(03):293-300.

[47] Gray CM, Grimson F, Layton D, et al. A Framework for Methodological Choice and Evidence Assessment for Studies Using External Comparators from Real-World Data. *Drug Saf*. 2020 Jul;43(7):623-633.

[48] 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(07):779-786.

[49] 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(01):38-43.

[50] Seeger JD, Davis KJ, Iannacone MR, et al. Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. *Pharmacoepidemiol Drug Saf*. 2020 Nov;29(11):1382-1392.

[51] Jahanshahi M, Gregg K, Davis G, et al. The Use of External Controls in FDA Regulatory Decision Making. *Ther Innov Regul Sci*. 2021 Sep;55(5):1019-1035.

[52] 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, (02):124-127.

[53] 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.

[54] 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(04):385-387.

[55] 李正翔. 医疗机构药品遴选指南[M]. 化学工业出版社:202209. 119.

[56] 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(07):769-776.

[57] 徐小微. 中国药品综合评价指南参考大纲(第二版)第一章药品安全性评价指南[J]. *药品评价*, 2015, 12(08):7.

[58] 梅丹, 张翠莲, 孙璐璐. 中国药品综合评价指南参考大纲(第二版)第四章药品质量评价指南[J]. *药品评价*, 2015, 12(08):12-14.

[59] 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(08):8-11.

[60] 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(08):11-12.

[61] 唐金陵. 证据金字塔与循证实践新思路[J]. *中华流行病学杂志*, 2011, 32(12):1193-1198.

[62] 徐志浩, 蒋理添. 基于药物经济学的卫生决策应用[J]. *现代养生(下半月版)*, 2021, 21(12):121-123.

[63] 石富豪, 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.

[64] 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.

- [65] 吴久鸿. 中国药品综合评价指南参考大纲(第二版)第六章临床药物经济性评估指南[J]. 药品评价, 2015, 12(08):16-19.
- [66] 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, (01):1-3+64.
- [67] 王文菁. 核心竞争力观念对当代企业管理理念的影响[J]. 商场现代化, 2021, (18):90-92.
- [68] 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(07):782-787.
- [69] 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(08):15-16.
- [70] 金有豫, 赵志刚. 中国药品综合评价指南参考大纲(第二版)第八章药品信息服务评价指南[J]. 药品评价, 2015, 12(08):21-23.
- [71] 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.
- [72] 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(02):60-64.
- [73] 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.
- [74] 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(2021年第12卷总第60卷)——上海市法学会知识产权法研究会文集. 华东政法大学国际法学院, 2021:7.
- [75] 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(04):438-444+452.
- [76] 赖俊恺, 王斌, 姚晨, 等. 从真实世界数据到临床研究数据的标准转化研究[J]. 中国食品药品监管, 2021, (11):39-46.
- [77] 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.
- [78] 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.
- [79] 陈瑜, 胡雁, 周英凤, 等. 推动证据向临床转化(十三)证据临床转化的研究设计[J]. 护士进修杂志, 2020, 35(19):1748-1750.
- [80] 曾繁典, 郑荣远, 詹思延, 等. 药物流行病学[M]第2版, 北京:中国医药科技出版社. 2016:80-87
- [81] 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(07):862-866.
- [82] 陈新林, 李先涛, 胡镜清, 等. 复杂干预的设计和评价方法[J]. 辽宁中医药大学学报, 2013, 15(04):81-83.
- [83] 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(04):240-247.
- [84] 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.
- [85] 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(01):23-26.
- [86] 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(04):496-500.
- [87] 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(03):229-234.
- [88] 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(05):381-385.
- [89] 于炳好. 医疗机构提高慢性病病例报告质量方法探讨[J]. 中国药物经济学, 2014, 9(03):319.
- [90] 王艺. 临床医学研究常用设计方案实施方法第 7 讲横断面研究[J]. 中国实用儿科杂志, 2008, (07):558-560.
- [91] 李业棉, 赵芃, 杨崙惠, 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.
- [92] 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(01):155-160.
- [93] 侯晓铭, 戴国华, 高武霖, 等. 中医临床疗效评价队列研究中统计学方法的选择[J]. 中西医结合心脑血管病杂志, 2020, 18(22):3813-3816.
- [94] 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(07):594-600.
- [95] 刘澳林. 应用专家共识方法制订中医药领域病例一对照研究技术规范的研究[D]. 广州中医药大学, 2020.
- [96] 高武霖. 中医临床疗效评价的队列研究方法及其关键技术[D]. 山东中医药大学, 2019.
- [97] 刘梦舟. 面向队列研究的数据校验方法设计与实现[D]. 浙江大学, 2018.
- [98] 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(04):494-496.
- [99] 王梅, 王建华, 张抗, 等. 中医药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.
- [100] 王梅. 中医药疗效评价队列研究方法学探索[D]. 辽宁中医药大学, 2016.
- [101] 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.
- [102] 刘玉秀. 临床随机对照研究设计[C]//江苏省医学会临床流行病学分会. 江苏省临床研究大会暨江苏省第四次临床流行病学学术会议资料汇编. 中国人民解放军南京总医院, 2017:11.
- [103] 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(02):350-352.
- [104] 万朝敏. 临床医学研究常用设计方案实施方法非随机对照研究[J]. 中国实用儿科杂志, 2008, (05):398-400.
- [105] Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0: The Cochrane Collaboration[EB/OL]. 2011, [2024-09-22]: Available from: www.cochrane-handbook.org.
- [106] 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(09):1025-1030.
- [107] 指南与共识[J]. 中华消化外科杂志, 2022, 21(1):C71-C76.
- [108] 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(01):26-34.
- [109] 北京协和医院罕见病多学科协作组, 中国罕见病联盟. 多准则决策分析应用于罕见病药品临床综合评价的专家共识(2022)[J]. 协和医学杂志, 2022, 13(02):235-254.

- [110] 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.
- [111] 夏如玉,金雪晶,柴倩云,等.运用多准则决策分析法的中药临床综合评价文献特征分析及设计要点[J].中医杂志,2022,63(02):119-124.
- [112] 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.
- [113] 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,2019,17(14):58-60.
- [114] 周罗晶,吴大嵘,欧爱华,等.卫生经济学评价方法在临床路径中的适用性、现状及应用思路[J].中国卫生经济,2010,29(01):52-54.
- [115] 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(03):353-357.
- [116] 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(01):31-35.
- [117] 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.
- [118] 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(04):307-311.
- [119] 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(05):644-648.
- [120] 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(06):716-718+723.
- [121] 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.
- [122] 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(04):727-731+733.
- [123] 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.
- [124] 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(05):751-756.
- [125] Roche JS, Clery MJ, Carter PM, et al. Tracking Assault-injured, Drug-using Youth in Longitudinal Research: Follow-up Methods. Acad Emerg Med. 2018 Nov;25(11):1204-1215.
- [126] 刘文贵.临床病例的远期随访方法[J].中国病案,2011,12(06):5-7.
- [127] 邓伟胜.医疗服务随访的方法与要求[J].医院管理论坛,2010,27(07):36-38.
- [128] 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,(06):618-620.
- [129] 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,(01):73-76.
- [130] 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,(01):83-85.
- [131] 杨军,张红叶,周北凡.心血管病前瞻性研究中对队列人群的随访方法[J].中国慢性病预防与控制,1996,(02):79-80+97.
- [132] 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(03):293-300.

[133] 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(06):745-749.

[134] 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(02):238-243.

[135] Prinsen CA, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” - a practical guideline. *Trials*. 2016 Sep 13;17(1):449.

[136] Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010 May;19(4):539-49.

[137] 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(03):239-243.

[138] 李峥, 刘宇. 混杂偏倚的控制——分层分析与多因素分析[J]. 中国护理管理, 2021, 21(02):279.

[139] 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.

[140] 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]. 山东大学, 2020.

[141] 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.

[142] 刘宝花, 王培玉. 危险因素研究中偏倚与混杂的控制[J]. 中华健康管理学杂志, 2011, 05(2):115-118.

[143] Berkson J. Limitations of the application of fourfold table analysis to hospital data. *Int J Epidemiol*. 2014 Apr;43(2):511-515.

[144] Andersen PK, Geskus RB, de Witte T, et al. Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidemiol*. 2012 Jun;41(3):861-70.

[145] 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.

[146] 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.

[147] 刘敏. 基于统计分析与张量分解的图像去噪算法研究[D]. 南京大学, 2021.

[148] 刘学英, 郭炯. 统计分析方法在市场调研前期的应用[J]. 产业创新研究, 2021, (13):50-51+102.

[149] Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*. 2006 Sep 4;185(5):263-267.

[150] Turner L, Shamseer L, Altman DG, et al. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev*. 2012 Nov 29;1:60.-67

[151] Tan ZW, Tan AC, Li T, et al. Has the reporting quality of published randomised controlled trial protocols improved since the SPIRIT statement? A methodological study. *BMJ Open*. 2020 Aug 26;10(8):e038283.

[152] 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. *PLoS One*. 2013 Dec 26;8(12):e83138.
- [153] 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(01):1-4.
- [154] Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013: new guidance for content of clinical trial protocols. *Lancet*. 2013 Jan 12;381(9861):91-92.
- [155] Vandembroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int J Surg*. 2014 Dec;12(12):1500-1524.
- [156] Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015 Oct 6;12(10):e1001885.
- [157] 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(03):190-198+212.
- [158] 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(04):263-272.
- [159] Gagnier JJ, Kienle G, Altman DG, et al. The CARE Guidelines: Consensus-based Clinical Case Reporting Guideline Development. *Glob Adv Health Med*. 2013 Sep;2(5):38-43.
- [160] 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, (05):407-410.
- [161] Fu SF, Cheng CW, Zhang L, et al. Consensus-based recommendations for case report in Chinese medicine (CARC). *Chin J Integr Med*. 2016 Jan;22(1):73-79.
- [162] Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*. 2008 Nov 11;337:a2390.
- [163] Dwan K, Li T, Altman DG, et al. CONSORT 2010 statement: extension to randomised crossover trials. *BMJ*. 2019 Jul 31;366:14378.
- [164] Vohra S, Shamseer L, Sampson M, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *BMJ*. 2015 May 14;350:h1738.
- [165] Gagnier JJ, Boon H, Rochon P, et al. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med*. 2006 Mar 7;144(5):364-367.
- [166] Cheng CW, Wu TX, Shang HC, et al. CONSORT Extension for Chinese Herbal Medicine Formulas 2017: Recommendations, Explanation, and Elaboration. *Ann Intern Med*. 2017 Jun 27;167(2):112-121.
- [167] Li J, Hu JY, Zhai JB, et al. CONSORT extension for reporting N-of-1 trials for traditional Chinese medicine (CENT for TCM): Recommendations, explanation and elaboration. *Complement Ther Med*. 2019 Oct;46:180-188.
- [168] Kwakkenbos L, Imran M, McCall SJ, 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. *BMJ*. 2021 Apr

29;373:n857.

- [169] Vlahov D. Transparent Reporting of Evaluations with Nonrandomized Designs (TREND). *J Urban Health*. 2004 Jun;81(2):163-4.
- [170] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021 Mar 29;372:n71.
- [171] Zhang X, Tan R, Lam WC, et al. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Extension for Chinese Herbal Medicines 2020 (PRISMA-CHM 2020). *Am J Chin Med*. 2020;48(6):1279-1313.
- [172] 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. *Value Health*. 2022 Jan;25(1):3-9.
- [173] Chen Y, Yang K, Marušić A, et al. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. *Ann Intern Med*. 2017 Jan 17;166(2):128-132.
- [174] Ma LL, Wang YY, Yang ZH, et al. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? *Mil Med Res*. 2020 Feb 29;7(1):7.
- [175] Muckart DJ. Evidence-based medicine - are we boiling the frog? *S Afr Med J*. 2013 May 28;103(7):447-448.
- [176] Mulrow CD. The medical review article: state of the science. *Ann Intern Med*. 1987 Mar;106(3):485-488.
- [177] Higgins JP, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011 Oct 18;343:d5928.
- [178] Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019 Aug 28;366:14898.
- [179] Ibbotson T, Grimshaw J, Grant A. Evaluation of a programme of workshops for promoting the teaching of critical appraisal skills. *Med Educ*. 1998 Sep;32(5):486-91.
- [180] 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 Adv Nurs*. 2020 Apr;76(4):1082-1089.
- [181] Taylor R, Reeves B, Ewings P, et al. A systematic review of the effectiveness of critical appraisal skills training for clinicians. *Med Educ*. 2000 Feb;34(2):120-125.
- [182] Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996 Feb;17(1):1-12.
- [183] Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995 Feb 1;273(5):408-412.
- [184] Hartling L, Ospina M, Liang Y, et al. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *BMJ*. 2009 Oct 19;339:b4012.
- [185] Slim K, Nini E, Forestier D, et al. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ J Surg*. 2003 Sep;73(9):712-716.

- [186] Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016 Oct 12;355:i4919.
- [187] 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]. Accessed 16 Jan 2020, [2023-07-13]. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
- [188] Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. 2010 Sep;25(9):603-5.
- [189] Downes MJ, Brennan ML, Williams HC, et al. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open*. 2016 Dec 8;6(12):e011458.
- [190] Gagnier JJ, Kienle G, Altman DG, et al. The CARE guidelines: consensus-based clinical case report guideline development. *J Clin Epidemiol*. 2014 Jan;67(1):46-51.
- [191] Drummond MF, Richardson WS, O'Brien BJ, et al. 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? Evidence-Based Medicine Working Group. *JAMA*. 1997 May 21;277(19):1552-1557.
- [192] Vardell E, Malloy M. Joanna Briggs Institute: an evidence-based practice database. *Med Ref Serv Q*. 2013;32(4):434-42. doi: 10.1080/02763869.2013.837734. Erratum in: *Med Ref Serv Q*. 2014;33(1):122-123.
- [193] Hannes K, Lockwood C. Pragmatism as the philosophical foundation for the Joanna Briggs meta-aggregative approach to qualitative evidence synthesis. *J Adv Nurs*. 2011 Jul;67(7):1632-1642.
- [194] Sacks HS, Berrier J, Reitman D, et al. Meta-analyses of randomized controlled trials. *N Engl J Med*. 1987 Feb 19;316(8):450-455.
- [195] 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 Feb 15;7:10.
- [196] Shea BJ, Reeves BC, 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. *BMJ*. 2017 Sep 21;358:j4008.
- [197] Whiting P, Savović J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016 Jan;69:225-34.
- [198] Vlayen J, Aertgeerts B, Hannes K, et al. A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *Int J Qual Health Care*. 2005 Jun;17(3):235-242.
- [199] Cattani A, Teixeira PP, Eckert IDC, 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). *Br J Nutr*. 2023 Jan 14;129(1):66-76.
- [200] Hailey D. Toward transparency in health technology assessment: a checklist for HTA reports. *Int J Technol Assess Health Care*. 2003 Winter;19(1):1-7.
- [201] Critical Appraisal Skills Programme. CASP (CaseControlStudy) checklist

- [EB/OL]. 2019-01-01, [2023-07-14]. <https://casp-uk.net/casp-tools-checklists/>.
- [202] 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. *Ann Epidemiol*. 2021 Mar;55:57-63. e15.
- [203] Zhang J, Wang Y, Weng H, et al. Management of non-muscle-invasive bladder cancer: quality of clinical practice guidelines and variations in recommendations. *BMC Cancer*. 2019 Nov 6;19(1):1054.
- [204] Hooijmans CR, Rovers MM, de Vries RB, et al. SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol*. 2014 Mar 26;14:43.
- [205] Li BH, Yu ZJ, Wang CY, et al. A Preliminary, Multicenter, Prospective and Real World Study on the Hemostasis, Coagulation, and Safety of Hemocoagulase Bothrops Atriox in Patients Undergoing Transurethral Bipolar Plasmakinetic Prostatectomy. *Front Pharmacol*. 2019 Nov 27;10:1426.
- [206] Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med*. 2019 Jan 1;170(1):W1-W33.
- [207] Gold C, Erkkilä J, Crawford MJ. Shifting effects in randomised controlled trials of complex interventions: a new kind of performance bias? *Acta Psychiatr Scand*. 2012 Nov;126(5):307-314.
- [208] 石秀园. 我国药品临床综合评价认知现状及决策转化影响因素研究[D]. 昆明医科大学, 2021.
- [209] 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.
-