

ICS **.***.**
C**



团体标准

T/CACM ****—20**

中医药真实世界研究技术规范 证据质量评价与报告

Technical Specifications for Real-World Studies of Traditional Chinese Medicine
Quality appraisal and reporting specification

(文件类型：公示稿)

(完成时间：2021年6月)

20**-**-**发布

20**-**-**实施

中华中医药学会发布

目 次

前 言.....	II
引 言.....	III
1 范围.....	1
2 规范性引用文件.....	1
3 术语及定义.....	1
4 真实世界研究设计分类与真实世界证据.....	2
4.1 真实世界研究设计分类.....	2
4.2 真实世界证据.....	3
5 真实世界证据的质量评价.....	3
5.1 真实世界证据质量评价的总体原则与方法.....	3
5.2 真实世界研究常见偏倚.....	3
5.3 评价中医药真实世界证据质量的特殊考虑.....	3
5.4 不同类型真实世界证据质量的评价方法.....	5
6 真实世界证据的报告规范.....	8
6.1 观察性研究的证据报告规范.....	9
6.2 干预性研究的证据报告规范.....	9
6.3 证据综合的报告规范及要点.....	10
7 真实世界证据分级.....	10
附录 A（资料性）经典临床研究设计与真实世界研究设计.....	11
附录 B（资料性）真实世界研究常见偏倚的类型.....	13
B.1 选择偏倚.....	13
B.2 信息偏倚.....	15
B.3 混杂偏倚.....	17
附录 C（资料性）真实世界证据质量评价工具和报告规范汇总表.....	20
参考文献.....	24

前 言

《中医药真实世界研究技术规范》包括如下标准：

- T/CACM ***. 1 中医药真实世界研究技术规范 数据库构建和数据预处理；
- T/CACM ***. 2 中医药真实世界研究技术规范 统计分析计划制定；
- T/CACM ***. 3 中医药真实世界研究技术规范 证据质量评价与报告；
- T/CACM ***. 4 中医药真实世界研究技术规范 伦理审查；
- T/CACM ***. 5 中医药真实世界研究技术规范 基于证据的中药有效性及安全性评价；

本文件是该系列规范的第 3 个规范。

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》规定的规则起草。

本文件由北京中医药大学、中国中医科学院西苑医院、中药临床疗效和安全性评价国家工程实验室、中药临床研究与评价重点实验室提出。

本文件由中华中医药学会归口。

本文件起草单位：北京中医药大学循证医学中心、北京中医药大学国际循证中医药研究院、中国中医科学院西苑医院、中药临床疗效和安全性评价国家工程实验室、中药临床研究与评价重点实验室、广州中医药大学第二附属医院、北京大学公共卫生学院、广州中医药大学第一附属医院，中国人民大学。

本文件主要起草人：费宇彤、柴倩云、高蕊、陆芳、訾明杰、孙明月、杨忠奇、陈大方、刘建平。

本文件起草人：罗懿婧、冯玉婷、夏如玉、吴大嵘、刘少南、梁昌昊、李高彪、李迅、荣红国、金雪晶、曹卉娟、张颖、杨红、邢景丽、王萍。

引 言

2020年1月国家药品监督管理局发布《真实世界证据支持药物研发与审评的指导原则（试行）》，是我国首个关于真实世界证据支持药物研发与审评的指导文件。2020年5月，国家药监局公布《中药注册管理专门规定（征求意见稿）》，该规定强调了中药的研制应当符合中医药理论，注重体现整体观及中医药原创思维，注重临床实践基础；特别提出中药人用经验是指在长期临床实践中积累的用于满足临床需求，具有一定规律性、可重复性的关于中医临床诊疗认识的概括总结。申请注册的中药具有人用经验的，可根据人用经验对药物安全性、有效性的支持程度，合理减免相应的申报资料。以上规定凸显了在中药新药的研发中真实世界证据的重要性。真实世界证据的质量直接影响着证据的科学性、真实性、适用性等，对临床决策具有重要意义。然而目前真实世界研究质量评价方面尚缺乏统一标准。

本文件旨在建立基于真实世界证据的质量评价规范和报告规范，探索真实世界证据的分类、证据评价原则及要点等，为真实世界证据在中药新药研究中的应用定位和合理使用形成指导。由于真实世界研究的证据质量评价工具的研发尚未成熟，因此在使用本文件时，还需紧密结合国内外的最新研究成果。本文件也可以为科研目的的真实世界证据质量评价和报告提供参考。本文件并非为指导设计真实世界研究而制定，虽然规范内容可以用来间接作为设计工作的借鉴。

中医药真实世界研究技术规范 证据质量评价与报告

1 范围

本文件规定了基于循证医学方法的真实世界证据质量评价与报告规范。

本文件适用于中药新药研发工作中真实世界证据质量评价及规范报告的参考,亦可作为医疗机构、中药研发企业、中医药科研机构的人员对各类中医药临床研究证据的质量评价。

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。

T/CACM 022—2017 中医真实世界研究技术规范通则

国家药监局 2021 年第 27 号通告 用于产生真实世界证据的真实世界数据指导原则(试行)

国家药监局 2020 年第 1 号通告 真实世界证据支持药物研发和审评的指导原则(试行)

国家药监局 2020 年 中药注册管理专门规定(征求意见稿)

国家药监局,国家卫健委 2020 年 药物临床试验质量管理规范

国家食品药品监督管理总局 2015 年第 83 号 中药新药临床研究一般原则

吴阶平医学基金会中国胸部肿瘤研究协作组 2018 年 真实世界研究指南

世界医学大会 2013 年 赫尔辛基宣言

3 术语及定义

下列术语和定义适用于本文件。

3.1

真实世界数据 Real-World data, RWD

来源于日常所收集的各种与患者健康状况和/或诊疗及保健有关的数据。并非所有的真实世界数据经分析后就能产生真实世界证据,只有满足适用性的真实世界数据经恰当和充分地分析后才有可能形成真实世界证据。。

[来源:国家药监局 2021 年第 1 号通告用于产生真实世界证据的真实世界数据指导原则(试行)]

3.2

真实世界证据 Real-World evidence, RWE

通过对适用的真实世界数据进行恰当和充分的分析所获得的关于药物的使用情况和潜在获益-风险的临床证据。

[来源：国家药监局 2021 年第 1 号通告用于产生真实世界证据的真实世界数据指导原则（试行）]

3.3

真实世界研究 Real-World Research/Study, RWR/RWS

针对临床研究问题,在真实世界环境下收集与研究对象健康状况和/或诊疗及保健有关的数据（真实世界数据）或基于这些数据衍生的汇总数据，通过分析，获得药物的使用价值及潜在获益-风险的临床证据（真实世界证据）的研究过程。

[来源：国家药监局 2021 年第 1 号通告用于产生真实世界证据的真实世界数据指导原则（试行）]

3.4

人用经验 Human Experience

在中医药理论指导下，在长期临床实践中积累的用于满足临床需求，具有一定规律性、可重复性的关于中医临床诊疗认识的概括总结。

[来源：国家药监局 2020 年中药注册管理专门规定（征求意见稿）]

3.5

经典临床研究设计 Classical Clinical Study Design

目前临床流行病学中的常见临床研究设计，主要包括随机对照试验设计、非随机对照试验设计、队列研究设计、病例对照研究设计、横断面研究设计、病例系列研究设计及个案报告等。

[来源：刘建平.循证中医药临床研究方法.人民卫生出版社 2020.]

4 真实世界研究设计分类与真实世界证据

4.1 真实世界研究设计分类

《真实世界证据支持药物研发和审评的指导原则（试行）》（国家药监局 2020 年第 1 号通告）将真实世界研究的基本设计分为：实用临床试验、使用真实世界证据作为外部对照的单臂试验、观察性研究。本文件在此基础上扩展了随机对照设计、非随机对照设计、无对照单臂试验和注册研究设计。其中观察性研究包括利用真实世界数据和研究型数据开展的队列研究（前瞻性、回顾性、双向性）、注册研究、病例对照研究、横断面研究、病例系列和

病例报告。（附录 A）

4.2 真实世界证据

真实世界证据来源于真实世界研究,真实世界证据又可分为原始研究证据和二次研究证据,即原始研究证据的综合。在 4.1 真实世界研究形成的证据的基础上,还存在二次研究证据,如基于真实世界队列研究的系统综述与 meta 分析证据、基于实用性对照试验的系统综述和 meta 分析证据等。

5 真实世界证据的质量评价

5.1 真实世界证据质量评价的总体原则与方法

依据《真实世界证据支持药物研发和审评的指导原则(试行)》(国家药监局 2020 年第 1 号通告),评价真实世界证据应依从两个主要原则:真实世界证据是否可以支持需要回答的临床问题;已有的真实世界数据是否通过科学的研究设计、严谨的组织实施及合理的统计分析得到所需的真实世界证据。

对真实世界证据的质量评价,首先应明确研究类型及生成证据的类型;其次评估其研究设计类型可能存在的偏倚风险和证据综合时需要考虑的方法学要素,本文件 5.4 中所列各种评价工具分别体现了各研究类型中需着重考虑的偏倚风险;最后充分考虑证据的内部真实性和外部真实性,根据评价原则、设计类型和证据综合的评价方法进行综合评价。中医药真实世界证据还需要结合中医药自身的特点进行评价。

5.2 真实世界研究常见偏倚

文献报道中真实世界研究的常见偏倚包括以下三类(附录 B)。

a)选择性偏倚:入院率偏倚、罹患率偏倚、存活者治疗性选择偏倚、检出征候偏倚、无应答偏倚、志愿者偏倚、易感性偏倚、时间效应偏倚、渠道偏倚、健康使用者偏倚、特异性偏倚、病程长度偏倚、竞争风险;

b)信息偏倚:错分偏倚、调查员偏倚、观察者偏倚、难以测量的时间偏倚、测量偏倚、非死亡时间偏倚、诱导偏倚、因果倒置、校正中介作用;

c)混杂偏倚:适应证混杂、合并用药混杂、沾染与干扰。

5.3 评价中医药真实世界证据质量的特殊考虑

首先应考虑中医药真实世界证据能否科学合理并且精准地回答所需解决的临床问题,体现中医药特色。研究证据中的关键特征(即患者人群、干预措施、对照措施及类型、结局指标)是否均紧密围绕研究目的设定;研究证据中的结果结论是否具有内部真实性和外部真实性。

a)评价人群的设定是否合理应考虑:

- 1) 中医的病、证、症、征及体质等中医诊断标准是否明确,诊断方式是否可靠。如不同临床医师的诊断水平是否一致,不同诊断设备的诊断结果是否统一,以及用语

的规范性；

2) 研究证据中的重要特征是否与研究目的对应的要求具有相似性，如病情轻重、病程时长、年龄分布、性别构成、合并症种类、辅助治疗种类及其他社会经济学特征；

3) 西医诊断与中医诊断是否需要同时出现，以及诊断名称与国际疾病分类编码（international classification of diseases, ICD）的一致性。

b)评价暴露/干预措施的设定是否合理：

1) 应根据研究目的判断暴露/干预措施定义的清晰性和设置的合理性，明确干预措施不应违背中医理论，例如中医干预措施的分类是否科学、分类是否同一水平；联合用药是否具有合理性、充分性；

2) 应明确所研究药物的组成及剂量，例如固定处方/协定方、经典方、院内制剂、名老中医经验方等，应明确其药物组成及加减药物的范畴；明确各味中药的具体剂量或增减剂量范畴；应明确所研究药物的剂型和煎煮方式。例如中药饮片应明确自煎或代煎，是否根据药性先煎后下；院内制剂应明确其制备工艺，包括但不限于药物是分开煎煮再混合发放，或直接混合后煎煮发放；中药颗粒剂亦应明确制备方法。不同煎煮方式、剂型是否可能带来疗效变异。非药物治疗，若涉及操作，需要考虑实施的质量和操作者的个人水平。

3) 评价中医药用药的连续性和疗程的长短，有交叉设计性质的研究证据，还需要评价干预措施洗脱期的设定是否合理；

4) 评价是否存在药物样本留存，以及能否提供各种药学参数和药物质量控制证据；

5) 还应考虑中医医嘱对患者日常生活的影响，例如饮食宜忌、调畅情志、调节作息等，尤其需考虑其对慢性疾病患者远期预后的可能影响；

6) 考量研究方案中的上述设定与可获得的真实世界数据及其分类之间的一致性。

c)如需要设置对照，评价对照措施/暴露的选择是否合理：

1) 应判断对照措施/暴露定义的清晰性和选择是否体现研究目的，例如，当以比较两种药物的有效性和安全性为主要目的时，条件允许的情况下是否选择了已知的、对所研究的疾病公认的最为有效和安全的阳性药；当与另一种中药或中医类干预措施/暴露对照时，需要评价对照的疗效和安全性的证据基础是否扎实；当选择空白对照或外部对照时，其理由是否充分合理；

2) 对照措施/暴露的选择须与干预措施/暴露的药物作用等方面保持较好的同质性，且应符合日常临床实践模式；如果属于非同类干预措施作为对照，需要说明理由；

3) 要判断对照组别患者的主要预后因素特征或病情特征与暴露组/干预组的特征是否可比。

d)评价结局的设定是否合理：

1) 结局指标的设定应与中医临床实践需要一致，同时顾及研究结局的国际及现代医学视角下的可理解并可利用的程度。建议主要结局采用国际公认的结局指标，次要结局加强对患者报告结局（PRO）的使用，如生命质量、患者主观评价及健康总体评价等。不建议使用没有公认标准的结局，如未经合理测评的由多维度指标构成的有效率结局等。关于证候结局的使用需要谨慎，因为证候评分的变化不一定直接代表疗效，

例如变化的分值也可能是由于证候量表有欠科学或变化为其他证候所致。

2) 关于结局指标选择的合理性,也可以参考国际核心结局指标集,如 COMET 核心结局研究,和已经发表的高质量指南、系统综述或随机对照试验等。

3) 中药临床研究必须重视对安全性的评价,包括对所有试验用到的药物的已知不良反应和未知的不良事件均需要进行严格周密的评判,并依据其严重程度,按照要求上报和处理。不合常理的低不良事件率往往会被怀疑是由于对安全性观测不足而导致的。要注意对不良事件是否要归属于不良反应做出及时的判断。考察不良事件/反应用语与国际相关规范,如国际医学用语辞典(Medical Dictionary for Regulatory Activities, MedDra)和世界卫生组织不良反应术语集(WHO Adverse Reaction Terminology, WHO-ART)等的一致性。

4) 评价是否包含卫生经济学结局、依从性信息,以及结局的效应量是否具有临床实际价值(最小临床重要差异, minimal clinical important difference, MCID)。

e)评价研究时长是否足够:应根据研究目的判断中医药暴露/干预时长及随访时长,例如,当探索有效性相关结局时,应根据短期疗效、长期疗效分别设定暴露/干预/随访时间;当探索暴露/干预对结局发生的影响时,应使随访时长足够达到结局发生。应判断待研究的中药开始和结束使用时间与相配合的西药的用药时间的关系。

f)评价中医药真实世界证据质量的其他特殊考虑:将真实世界研究用于中医药临床研究及中药研发时,还应考虑中医药在真实世界环境中的适宜性、便捷性、可及性等。当用于支持中药研发豁免 II 期临床试验时,应根据研究目的和中医理论,严格评判证据对患者人群的中医特质、中药组成及剂量范畴、中药剂型及煎煮方式、暴露/干预及随访时长、联合治疗的限定和体现是否合理。当适用人群复杂、暴露因素/干预方式过多,存在合并用药等情况时,应判断其是否采用必要措施处理相关混杂因素。还需考虑证据是否在恰当的时候采用了结局信息采集者、结局评价者和统计分析人员盲法。还需特别注意中医药个体化治疗结论的外推性问题等。具体参见相关规范如《中医药真实世界研究技术规范统计分析计划制定》。

5.4 不同类型真实世界证据质量的评价方法

评价真实世界证据的质量,可以分为内部真实性评价和外部真实性评价。本文件对二者都有涉及。内部真实性评价主要通过对产生各种证据的研究进行方法学偏倚风险评价来实现。外部真实性评价,即外推性的优劣,需要结合产生证据的人群、干预或暴露、对照、结局和研究场所与意欲外推的情境的符合程度来判断。

真实世界证据质量评价需区分针对经典研究设计类型直接形成的真实世界证据的评价和针对经典研究设计类型改良后形成的真实世界证据的评价两种类型。(附录 C)

——经典研究设计类型直接形成的真实世界证据的质量评价方法:

- a) 将国际现有研究证据的偏倚风险评价工具及条目作为必要条目进行评价;
- b) 将特殊研究类型在方法学偏倚方面的特殊考虑作为补充条目;
- c) 根据所评价的证据类型,进一步考量其应用于为中药研发提供证据时可能存在的偏倚问题及外推性问题。

——经典研究设计类型改良后形成的真实世界证据的质量评价方法:

- a) 优先选择真实世界研究特异性评价工具及条目进行评价；
- b) 若无针对该类型的真实世界评价工具及条目，则将国际现有研究类型的评价工具及条目作为必要条目，针对真实世界研究特征，进行必要的补充；
- c) 若两种评价工具及条目同时存在，建议结合使用，并凸显真实世界特异性条目；
- d) 根据所评价的证据类型，进一步考量其应用于为中药研发提供证据时可能存在的偏倚问题及外推性问题。

5.4.1 观察性研究的证据质量评价方法

5.4.1.1 经典研究设计类型直接形成的真实世界证据的评价方法

针对队列研究证据的质量评价，应首先采用国际现有评价工具，如高质量比较效果研究准则(good research for comparative effectiveness, GRACE)、NOS 量表 (Newcastle-Ottawa Scale, NOS) 中队列研究部分的条目、CASP 清单 (critical appraisal skill program, CASP) 中队列研究清单等。针对病例对照研究证据的质量评价，应首先采用国际现有评价工具，如 NOS 量表中病例对照部分的条目，CASP 清单中病例对照清单等，以及可以借鉴真实世界观察性研究评估工具 (Assessment of Real-World Observational Studies, ArRoWs)。针对横断面研究证据的质量评价，应首先采用国际现有评价工具，如 AHRQ (Agency for Healthcare Research and Quality, AHRQ) 横断面研究评价标准。注册研究所搭建的病例注册库或干预措施注册库，适用于支撑多种目的的观察性研究，如传统的队列研究、病例对照研究、病例-队列研究、横断面研究，也可以基于注册库设计随机对照试验和进行复杂的数据挖掘。当对注册研究证据进行质量评价时，应明确研究问题和设计类型。首先采用相应设计类型的评价工具及方法。针对病例系列研究证据的质量评价，应首先采用国际现有评价工具，如英国国立临床优化研究所 (National institute for clinical excellence, NICE)、澳大利亚 (The Joanna Briggs institute, JBI) 和加拿大卫生经济研究所 (Institute of Health Economics, IHE) 研制的病例系列质量评价工具。针对病例报告研究证据的质量评价，应首先采用国际现有评价工具，如澳大利亚 JBI 研制的病例报告质量评价工具。

5.4.1.2 经典研究设计类型改良后形成的真实世界证据的质量评价方法

对于该类型研究，应首先考虑真实世界特异性评价工具，并将其与经典工具结合使用。该领域尚处于发展早期，虽然国际上已经有一定研究进展，但并未形成非常成熟且公认的评价工具，如 ArRoWs 评价条目。建议 ArRoWs 和 NOS 量表联合使用。在基于已有数据库构建回顾性队列时，还需考虑真实世界中常见偏倚如入院率偏倚、竞争风险以及人群易感性损耗等偏倚对结果的影响。以及真实医疗环境下，对照组选择恰当与否，以及其他混杂因素。

5.4.2 干预性研究证据的质量评价方法

5.4.2.1 基于随机对照设计的证据质量评价方法

5.4.2.1.1 经典研究设计类型直接形成的真实世界证据的质量评价方法

首先应选择国际现有随机对照试验方法学质量评价工具及条目评价，如采用 Cochrane

协作网的偏倚风险评价工具（risk of bias tool, RoB）。值得注意的是，尽管实用性随机对照试验不要求受试者盲法和干预实施者盲法，但因此而导致的信息偏倚风险却仍旧真实存在。RoB 2 及其他国际上较为公认的工具如改良 Jadad 量表等也可考虑使用。此外，还需要考虑实用性随机对照试验在设计实施时的特殊性及潜在偏倚，例如治疗领域和干预措施、对照措施等是否为当前最具竞争力的最佳临床实践；是否具有足够的可以用于评价的病例数（特别是临床结局罕见的情况）；参与实用性随机对照试验的各试验中心甚至不同的数据库之间对结局的评价和报告方法是否一致；当盲法不可行时，应考虑非盲对结局变量（特别是患者报告的结局）可能产生的影响，可酌情着重使用受盲法与否影响较小的结局（如生存数据、显著性疾病进展等），以减少盲法不完全或非盲带来的可能偏倚。

5.4.2.1.2 经典研究设计类型改良后形成的真实世界证据的评价方法

经典随机对照研究设计类型改良后形成的真实世界证据，国际上尚未研制相应的质量评价方法，因此宜采用经典研究设计类型的质量评价工具，如 RoB、RoB 2、Jadad 等，并在此基础上根据改良后的真实世界证据的特殊性进行评价。真实世界随机交叉证据质量评价，还应考虑洗脱期的设置是否充分合理，是否存在顺序效应和剩余作用，以及脱落、失访病例对结果可靠性影响。真实世界单病例随机对照试验证据质量评价，还应考虑真实世界环境下，仅对 1 位患者进行非安慰剂、非盲法的试验，其结果局限性和可靠性的评价。真实世界随机征求许可试验证据质量评价还应考虑换组率及其带来的稀释效应。真实世界基于患者意愿的随机对照试验证据质量评价需考虑到非随机组涉及到的观察性检验相关（除暴露率和现患率外）的显著性检验，必要时进行危险度估计分析有助于确定患者意愿因素是有利因素还是不利因素。

5.4.2.2 基于非随机对照设计的证据质量评价方法

无对照单臂试验证据的质量评价，尚未发现国际上有相应的质量评价方法。可以考虑从如下方面进行评价：研究目的是否清晰，纳入排除标准是否明确合理，是否在多中心收集病例，如果为确证性研究是否合理计算了样本量并遵照执行，患者的纳入是否为连续病例，主要干预措施的规定是否清楚明确，联合干预措施是否规定清楚明确，研究结局及其测量方法是否在方案中明确规定并且严格执行了，干预前后是否均测量了结局指标，缺失数据是否在可接受范围内，是否应用了合理的统计学检验来评价相关结局指标并完善地进行了报告，是否充分报告了不良事件并进行了合理的因果推断，是否对影响疗效的其他因素进行了合理地考虑与分析，在可能的情况下是否对结局信息采集者、结局评价者施盲，研究结果是否支持研究结论，利益冲突是否被有效控制。

使用真实世界数据/非安慰剂研究数据/观察性研究数据作为对照的单臂试验证据质量评价也尚未发现国际规范，建议在前述评价要点的基础上，进一步考虑外对照数据来源数据集或研究其本身所处的医疗环境、医疗技术、诊断标准、结局的测量和分类、患者的基线水平、干预复杂程度、数据质量等方面是否可以保证单臂试验与其外对照相关特性的可比性，以及其对研究结果精确性、研究结论可靠性和外推性等方面带来的挑战。使用真实世界证据作为外对照的时候，上述局限可能由于数据的复杂性而变得更为需要慎重对待。为克服或减少这

些局限：

- a) 须确保所采集的数据符合真实世界数据的适用性要求；
- b) 采用平行外部对照设计要优于历史对照，平行外部对照可采用疾病登记模式，须保障数据记录尽可能完整、准确；
- c) 应采用恰当的统计分析方法，如合理利用倾向评分方法，虚拟匹配对照方法等；
- d) 要充分使用敏感性分析和偏倚的定量分析来评价已知或已测的混杂因素和未知或不可测量的混杂因素以及模型假设对分析结果的影响。

经典非随机对照研究设计类型改良后形成的真实世界证据，尚未发现国际上针对性的方法学质量评价工具，因此宜采用经典非随机对照研究设计类型的质量评价工具，如非随机对照试验方法学评价指标（methodological index for non-randomized studies, MINORS）、非随机对照试验偏倚风险评价工具（risk of bias in non-randomized studies of interventions, ROBINS-I）等。根据实际情况可以考虑多种评价方法结合使用。此外，还需此基础上根据改良后的真实世界证据的特殊性进行评价。

5.4.3 证据综合的质量评价方法

针对真实世界原始研究证据的综合，参考现有证据综合质量评价工具，如评价单项证据综合（系统综述）实施质量的 AMSTAR（Assessing the methodological quality of systematic reviews），用于干预性研究证据综合的 AMSTAR 2，以及评价 meta 分析结果证据等级的 GRADE(An Emerging Consensus on Rating Quality of Evidence and Strength of Recommendations, GRADE)体系。由于真实世界原始研究类型种类繁多，在进行证据综合时，建议将原始研究同一类别的进行综合，对于跨类别合并的证据，目前尚缺少评价工具。如果确需合并，建议拆分亚组及进行敏感性分析；对于定量综合的评价，上述原始研究偏倚风险的评价是其评价的重要方面，但除此以外，还尚需考虑合并的异质性、证据的直接性、合并结果的精确性、发表偏倚的存在、利益冲突的影响等。此领域也属于新兴领域，针对性工具还比较缺乏，现有工具可供参考，但需要在使用时结合实际情况酌情决定是否增加相关评价维度。

6 真实世界证据的报告规范

为更好地使用和评价真实世界证据，研究者应准确、充分地报告真实世界证据的设计、实施和分析，不同类型的真实世界证据可参照目前国际通用的相应的报告规范（附录 C）。在进行证据报告规范评价时，也可从以下 3 个方面进行：

- a) 将国际现有报告规范及条目作为必要条目进行评价；
- b) 当特殊研究类型尚无国际现有报告规范，可将其在实施方面的特殊考虑作为补充条目；
- c) 根据所评价的证据类型，进一步考量其应用于中药研发时需要报告的细节。

6.1 观察性研究的证据报告规范

前瞻性队列研究证据、回顾性队列研究证据、双向性队列研究证据、病例对照证据、巢式病例对照证据、横断面研究证据，此类属于观察性研究直接形成的真实世界证据，其报告内容首先应参照国际通用报告规范——《加强观察性流行病学研究报告的质量声明》[The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies]。

真实世界前瞻性队列研究设计、真实世界双向性队列研究设计、真实世界巢式病例对照研究设计，此类属于经典研究设计采用真实世界日常或医疗数据形成的真实世界证据，其报告内容首先应参照国际通用报告规范——《常规收集医疗卫生数据开展观察性研究的报告规范》[The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement, RECORD]，以及针对药物流行病学研究的版本(RECORD-PE)清单等，并结合各研究证据类型的特点进行评价。

病例系列证据目前尚无国际通用报告标准，建议对前述病例系列方法学质量评价工具中提及的要点进行详细报告，并参考 STROBE 标准中的部分内容。病例报告证据属于观察性研究直接形成的真实世界证据，其报告内容首先应参照国际通用报告规范——《基于共识的临床病例报告规范》(The CARE Guidelines: Consensus-based Clinical Case Reporting Guideline development)，当用于报告中药研发相关内容时，还需参照《中医病例报告建议条目》和《基于共识的中医病例报告规范》(Consensus-based recommendations for case report in Chinese medicine, CARC)，须重点报告古籍记载内容、文献检索的策略和文献分析的结果，用以证明本病例报告的价值；讨论与古籍或他人发表的相关文章进行比较，说明诊断及治疗上的异同。

6.2 干预性研究的证据报告规范

6.2.1 基于随机原理的试验性研究证据的报告规范

实用性随机对照试验证据的报告可参照目前国际通用的相关报告规范如《实用性随机对照试验报告规范》(Improving the reporting of pragmatic trials: an extension of the CONSORT statement)、《阶梯楔形群组随机对照试验报告规范：CONSORT 2010 声明扩展及解释》

(Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration)；真实世界随机交叉试验报告首先参照目前国际通用的相关报告规范《CONSORT 2010 声明：随机交叉试验扩展版》(CONSORT 2010 statement: extension to randomized crossover trials)；真实世界单病例随机对照试验报告首先参照目前国际通用的相关报告规范《CONSORT 声明单病例随机对照试验扩展版》[CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement]。真实世界随机征求许可试验、真实世界基于患者意愿的随机对照试验，尚无国际通用报告规范，但在报告该研究类型时须注意详细报告实际各组病例数及换组率的情况。当用于为中药研发提供证据时，可参照《草药干预措施随机对照试验报告：CONSORT 扩展声明》(Reporting Randomized Controlled Trials of Herbal Interventions: An Elaborated CONSORT Statement)、《中药复方临床随机对照试验

报告规范 2017》（CONSORT Extension for Chinese Herbal Medicine Formulas 2017: Recommendations, Explanation, and Elaboration, CONSORT-CHM Formulas）、《CONSORT 中医药单病例随机对照试验扩展版》（CONSORT extension for reporting N-of-1 trials for traditional Chinese medicine (CENT for TCM): Recommendations, explanation and elaboration）。当随机对照试验作为嵌入式研究、巢式研究，或使用队列及日常数据时（如注册研究、电子健康数据、管理数据库等），需参照目前国际通用的相关报告规范《使用队列或收集日常数据的随机对照试验的报告规范 CONSORT 扩展版》[CONSORT extension for the reporting of randomized controlled trials conducted using cohorts and routinely collected data (CONSORT-ROUTINE) checklist with explanation and elaboration]。

6.2.2 基于非随机对照设计的证据报告规范及要点

真实世界实用性同期非随机对照证据、真实世界历史对照证据、真实世界自身对照证据、真实世界交叉试验证据，均为非随机对照试验证据，应遵循《非随机对照试验报告规范》（Transparent Reporting of Evaluations with Nonrandomized Designs, TREND）；无对照单臂试验证据以及使用真实世界数据/非安慰剂研究数据/观察性研究数据作为对照的单臂试验证据，尚无国际通用报告规范，建议报告内容对前列单臂试验方法学质量要点中提及的内容进行详细报告。建议结合使用《草药干预措施随机对照试验报告：CONSORT 扩展声明》、CONSORT-CHM Formulas 中相关的要求。

6.3 证据综合的报告规范及要点

针对真实世界原始研究证据的综合，当进行系统综述和 Meta 分析时须参照《系统评价和 Meta 分析优先报告的条目》（Preferred Reporting Items for Systematic Reviews and Meta-Analyses, PRISMA 2020）；当用于报告中药研发相关内容时，还需参照其扩展版《中草药系统评价和 Meta 分析优先报告的条目》[Preferred Reporting Items for Systematic Reviews and Meta-Analyses: Extension for Chinese Herbal Medicines 2020 (PRISMA-CHM 2020)]。（附录 C）

7 真实世界证据分级

真实世界证据质量的评价结果依照上述相关方法学工具、条目及规范进行报告与评价，推荐方式遵循原有方法学评价结果。各类真实世界研究设计及分类尚未成熟，其设计类型级别排序尚未形成。真实世界研究的证据汇总结果的分级方法可以参考 GRADE 体系等，但要注意 GRADE 体系尚未针对真实世界证据起评点给出系统性建议。此外，建议结合使用《中医药真实世界研究技术规范 基于证据的中药有效性及安全性评价》等其他规范。

附录 A
(资料性)

经典临床研究设计与真实世界研究设计

分类	经典设计	限定/改良	真实世界研究设计
观察性研究设计			
	前瞻性队列研究设计（使用真实世界研究数据）	真实世界日常或医疗数据 /	真实世界前瞻性队列研究设计 前瞻性队列研究设计
	回顾性队列研究设计	/	回顾性队列研究设计
	双向性队列研究设计（使用前瞻性真实世界研究数据）	前瞻性真实世界日常或医疗数据 /	真实世界双向性队列研究设计 双向性队列研究设计
	注册研究	/	注册研究
	病例对照研究设计	/	病例对照设计
	巢式病例对照研究设计（使用真实世界研究数据）	真实世界日常或医疗数据 /	真实世界巢式病例对照研究设计 巢式病例对照研究设计
	横断面研究设计	/	横断面研究设计
	病例系列设计	/	病例系列设计
	病例报告设计	/	病例报告设计
干预性研究设计			
基于随机对照设计	实用性随机对照试验设计	/	实用性随机对照试验设计
	随机交叉临床试验设计	非安慰剂对照的研究数据	真实世界随机交叉临床试验设计
	单病例随机对照试验设计	非安慰剂对照的研究数据	真实世界单病例随机对照试验设计
	随机征求许可试验设计	非安慰剂对照的研究数据	真实世界随机征求许可试验设计
	基于患者意愿的随机对照试验设计	非安慰剂对照的研究数据	真实世界基于患者意愿的随机对照试验设计
	技能型随机对照试验设计	/	技能型随机对照试验设计

基于非随机对照设计	单臂试验设计	无对照/使用真实世界数据、非安慰剂研究数据或观察性研究数据作为对照	无对照单臂试验/使用真实世界数据/非安慰剂研究数据/观察性研究数据作为对照的单臂试验
	同期非随机对照设计	非安慰剂对照的研究数据	真实世界同期非随机对照设计
	历史对照设计	非安慰剂对照的研究数据	真实世界历史对照设计
	自身对照设计	非安慰剂对照的研究数据	真实世界自身对照设计
	交叉试验设计	非安慰剂对照的研究数据	真实世界交叉试验设计

附录 B

(资料性)

真实世界研究常见偏倚的类型

B.1 选择偏倚

选择偏倚是指由于选择研究对象的方法存在问题而使研究结果偏离真实情况而产生的偏倚。选择偏倚主要发生在研究设计阶段。真实世界研究常见的选择偏倚如下。

1) 入院率偏倚(admission rate bias) 又称伯克森偏倚(Berkson bias)和转介偏倚(referral bias)^[1], 指由于各种疾病的患者因所患疾病的严重程度、就医条件、对疾病的认识水平等因素而出现的不同的就医水平的现象, 使得以医院就诊患者为对象进行研究时产生的偏倚。如: 当研究某病 A 与因素 X 的关系时, 以 B 病患者为对照, 由于 A 病、B 病和暴露于因素 X 的入院率不同, 导致医院所得的样本不能反映人群中病例和对照人群的实际暴露情况, 而错误的估计暴露与疾病的关系。^[2]在病例对照研究中, 选择偏倚指利用住院病人或门诊就诊病人作为病例或对照的研究对象时, 由于病例和对照入院率不同, 使其住院率受到研究因素的影响, 从而导致研究结果产生偏倚。^[1]在使用数据库开展研究时, 需在设计阶段考虑所使用的数据库人群对源人群的代表性。^[3]

2) 罹患率偏倚(prevalence bias) 又称现患-新发病例偏倚(prevalence-incidence bias)或奈曼偏倚(Neyman's bias), 或患病率及发病率偏倚, 是指因现患病例与新病例的构成不同, 只研究典型病例而排除轻症或非典型病例以及现患病例暴露状态发生改变而导致的偏倚。如以医院为基础研究药物对冠心病心肌梗死发生的预后情况时, 由于急性心肌梗死发作后, 部分病例在送医院前死亡, 而常未被计算在该病的总发病患者数内; 而部分轻症病例, 发作后经一般医疗机构治疗得救, 或有些病例是无痛发作, 经检查才发现。这类病例都可能被排除在研究之外, 而影响对心肌梗死药物治疗预后研究的判定, 产生偏倚。^[2]病例对照研究中通常是纳入现患病例或存活病例, 而不包括死亡病例和那些病程很短已治愈的病例。因此, 所纳入的病例样本与新发病例在病情、病程、预后等方面不尽相同, 从而并不能代表所研究疾病的总体发病人群。^[1]可以采用新发病例或新用药者设计解决。^[3]

与罹患率偏倚相类似的是, 长期用药者偏倚是指当患者在服用某种药品时, 有的患者会由于存在不适反应或者感觉药效不好而放弃继续服用该药品, 而长期用药者(prevalent user)往往是那些身体耐受性较好、治疗效果较好的患者。如果过多招募这些患者将导致低估药品的危害并高估药品的获益。^[1]

3) 存活者治疗性选择偏倚(survivor treatment selection bias) 对于病死率高的疾病, 生存时间长者才有机会接受后期治疗, 而对尚未接受后期治疗就死亡的患者, 则会被默认为未接受治疗。例如, 一个艾滋病患者队列, 在研究之初所有患者都没有接受过抗病毒治疗, 一年后某抗病毒新药上市, 存活者开始逐步接受该新药治疗, 而已死亡者则被看作未治疗者。如果比较治疗者和未治疗者的生存期, 则会发现治疗者的生存期明显长于未治疗者, 因为那些治疗者是可以存活到接受后期治疗阶段的患者, 其特征与那些无法生存到接受后期治疗阶段的“未治疗”患者自然有所不同。^[1]

4) 检出征候偏倚(detection signal bias, protopathic bias) 又称揭露伪装偏倚(unmasking

bias), 或暴露偏倚 (exposure bias)。^[2] 该偏倚是指某种因素虽然其本身并不是因其所研究疾病的病因, 但由于该因素能引起或促进与研究疾病症状和体征相似的症候出现, 从而有利于筛查出某些该疾病的早期患者, 使研究者误认为该因素即为病因^[1]。也可以理解为, 在选择病例时, 部分病例因为某种与所研究疾病无关的症状或体征而就医, 从而提高了所研究疾病的发现机会, 而产生偏倚。著名的例子是, 在研究雌激素与子宫内膜癌的关系中, 因为服用雌激素会致绝经期妇女子宫出血而增加子宫内膜癌的发现机会, 而错误的推断服用雌激素与子宫内膜癌发生有关^[2]。检出征候偏倚容易过高地估计暴露程度, 因而发生了系统误差, 最终可能得出该征候因素与该疾病有联系的错误结论。为了避免这类错误, 可以延长收集病例的时间, 使其超过由早期向中、晚期发生的时间, 则检出病例中暴露者的比例会趋于正常。^[3]

5) 无应答偏倚 (non-respondent bias) 指研究对象因各种原因对研究的内容产生不同的反应, 不予回答而产生的偏倚。无应答者的暴露或患病状况与应答者可能不同, 如果无应答者比例较高, 则使以有应答者为对象的研究结果可能存在严重偏倚。此外, 失访也可以认为是一种特殊的无应答, 因为研究对象未能按计划被随访, 造成了研究样本的选择偏倚。^[2]

6) 志愿者偏倚 (volunteer bias) 与一部分人无应答相反的情况是有一部分人特别乐意或自愿接受调查或测试。这些人往往是比较关心自身健康或自觉有某种疾病, 而想得到检查机会的人。他们的特征或经历不能代表目标人群。^[2]

在药物流行病学研究中, 自我选择偏倚 (self-selection bias) 或自我退出偏倚 (self-deselection bias) 与此较为类似。自我选择偏倚, 即当研究对象在选择是否参与某项研究或者是否接受某项治疗的时候, 往往会基于各种考虑以决定自己是否参与。很多情况下参与者与未参与者的个体特征会有所差异, 这些差异有可能直接或间接地与暴露因素或者研究结局之间存在联系, 从而引起偏倚。对于已参与队列研究的研究对象, 失访往往是特别需要考虑的问题。失访原因虽然多种多样, 但很多情况下是与暴露因素或者研究结局有关, 而不完全是单纯的随机性失访。由此可以引起研究结果产生偏倚, 称之为自我退出偏倚。^[1]

7) 易感性偏倚 (susceptibility bias) 指研究对象是否发生疾病不仅与暴露有关, 还与其自身对暴露的易感性有关。也就是由于各比较组研究对象的易感性不同而产生的偏倚。这类偏倚在传染性疾病的研究中最为常见。^[2] 同时, 在设计、实施药品服用者队列研究时, 对药品不能耐受者将随着时间的推移逐渐退出队列, 从而导致仍然留在队列中的患者对该药品会相对更为耐受。如果上述现象在研究设计和数据分析中不能得到正确的处理, 则会出现偏倚。^[1] 例如, 基于真实世界数据模拟经典 RCT 的关于绝经后激素治疗与冠心病风险的观察性研究中, 研究人员对接受激素治疗的患者进行随访, 并与非激素治疗者进行了比较。因为激素治疗组包含了许多已接受多年其他治疗的患者, 这排除了早期治疗结束后、激素治疗前发生的心血管事件, 使得这一部分的激素使用者对结局的易感性降低。这类偏倚也可能发生在采用阳性对照的研究中, 如在治疗组的患者中曾有一段时间的其他药物使用的经历, 而另一组患者大多为首次接受药物治疗新药者设计中, 在治疗开始前评估所有的协变量, 并确保这些协变量不是治疗和结局因果链上的中介变量。^[4]

8) 时间效应偏倚 (time effect bias) 对于肿瘤、冠心病等慢性疾病, 从开始暴露于内外危险因素到发病有一个漫长的过程。因此, 在研究中如果把暴露后即将发病的人、已发生

早期病变而未能检出的人作为非病例，就会产生此类偏倚。^[2]

9) 渠道偏倚 通常情况下，患者在就医的时候会自然地选择到某些特定的医院或门诊就医，或医生会有选择性地开药给某些特定的患者。例如，病情严重的患者倾向于到知名的大医院就诊；在某新药刚上市的时候，往往病情更严重或者对其他现有药品效果不好的患者会先使用该新药；或考虑安全性时，医生会倾向于给体质较好的患者推荐使用该药物。总之，由于种种就医、诊疗等原因导致不同医院患者或不同药品服用者的病情或个体特征不尽相同，由此引起的偏倚为渠道偏倚。^[1]

10) 健康使用者偏倚 (healthy user bias) 是指具有某些健康行为的患者也倾向于依从其他健康行为 (有效的药物治疗、饮食、体力活动等)。此类参与者比较关心自身健康，或自觉患有某些疾病，而想得到治疗或其他健康行为的引导。此类参与者的特征或经历通常不能代表目标人群。^[3]

11) 特发性偏倚 (protopathic bias) 是指由疾病或其他结局事件的基线特征而导致某种特别疗法或暴露开始、停止或改变。在真实世界病因研究中，可能会因纳入一些在疾病临床前期发生行为改变的对象而发生偏倚，当这种暴露状态的改变涉及到某种药物或治疗时，称为特发性偏倚。如研究对象因出现了与临床前期有关的症状或体征后自觉地减少或去除了某种暴露，当这些对象被纳入病例对照研究的对照组时，可能会影响危险因素与疾病的联系，而当这些对象因其早期症状或体征而被纳入病例组时，又有可能因其行为改变而只反映了其终止某种暴露后的效果而不是原有的长期暴露的效果。在研究设计阶段，应尽最大可能从整体上理解与疾病进展相关的病理生理学机制。^[3]

12) 病程长度偏倚 (length bias) 由于慢性疾病的病程长短不同，研究更容易纳入病程长的患者。例如在肿瘤药物研究中，肿瘤恶性程度低的患者病程可能更长、更易成为研究样本，而恶性程度较高的患者可能因病程较短而死亡，并未被及时纳入研究，从而低估了结局风险。因此，在研究方案中应对研究对象疾病分期、病程长短等作以具体规定。^[5]

13) 竞争风险 (competing risks) 对于长期随访的生存数据，一般生存分析只关注一个终点事件，而临床实践中研究对象的结局事件往往并不唯一，如果随访期内研究对象发生了其他结局而致使其不可能发生研究目标事件，这些目标结局以外的结局事件即称为竞争风险事件。例如，在某项肿瘤预后的登记注册研究中，结局事件是患者复发，如果随访期间患者因肿瘤死亡、因其他疾病死亡等，就不可能再发生肿瘤复发。传统的方法是将复发前死亡的个体按照删失数据处理，事实上高估了目标结局事件的发生率，导致估计偏差。这类问题应该选择竞争风险模型 (competing risk model)，考虑多种潜在结局，估计各原因危险率 (cause-specific hazard rate)。^[5]

B.2 信息偏倚

信息偏倚 (Information bias)，又称为观察偏倚或测量偏倚，是指研究过程中进行信息收集时产生的系统误差。测量方法的缺陷、诊断标准不明确或资料的缺失遗漏等都是信息偏倚的来源。

1) 错分偏倚 (Misclassification bias) ①药物错分：真实世界研究中药物暴露信息一般通过医院 EMR、医保数据、药物销售记录等电子数据库识别提取，诸多因素可能导致错分。

例如，信息可及性：当以医保数据识别提取暴露分类时，通过自费用药、第三方支付、资助用药等途径的药物。信息错误：以 OTC 销售识别暴露分类时，购买药物者并非真正用药患者等。②暴露错分：测量暴露（治疗干预）和结局事件过程中可能会出现暴露错误分类和结局事件错误分类，许多因素可能导致暴露或结局事件的错误分类，包括识别“暴露”的时间窗的长短、患者是否失访以及患者通过数据库覆盖范围外的其他途径获得治疗药物等。可以采用分层分析、多元回归分析和倾向评分分析进行分类加以控制。③结局错分：疾病诊断编码、药物编码、程序算法、数据提取系统、结局指标完整性等在识别结局指标时均可能存在错分。不同的 EMR 系统，对疾病诊断和诊断编码的完整性和准确性存在差异。不同疾病 ICD 编码的准确性也存在较大差异。^[4-5] 诊断怀疑偏倚、暴露怀疑偏倚、回忆偏倚、报告偏倚也属于错分偏倚。^[1]

诊断怀疑偏倚（diagnostic bias） 研究者事先已经知道了研究对象的某些情况，如服用某种药物或具有某种已知的暴露因素，因而会在研究过程中更加仔细地寻找某种结果，但对于不具有这些情况的研究对象则不会这样，从而产生偏倚。诊断亚临床病例或鉴别是否为药物副作用时常发生诊断怀疑偏倚。有时也称为期望偏倚。^[2]

暴露怀疑偏倚（exposure suspicion bias） 发生于研究者事先知道研究对象患有某种疾病，而在资料收集过程中会对患病者比未患病者更仔细地收集暴露因素，从而产生偏倚。当研究者对致病因素与某病的关联有主观见解时，最容易产生这类偏倚。^[2]

回忆偏倚（recall bias） 指各比较组中研究对象回忆以前发生的事或经历时，在准确性和完整性上存在着系统差异而导致的偏倚。^[2]产生回忆性偏倚的原因很多，如：调查的因素或事件发生的频率很低或因调查研究对象对此记忆模糊或遗忘，还可能病例组的患者因患病而对过去的暴露史反复思索，甚至家属也帮助提供线索，以致于夸大了暴露情况，而对照组的非患者对调查不够重视，未认真回忆暴露史。^[3]

报告偏倚（reporting bias） 当研究因素涉及生活方式或隐私如收入、婚育史、婚外性行为时，研究对象会因种种原因隐瞒或编造虚假信息，从而产生的偏倚。^[2]

2) 调查员偏倚（interviewer bias） 是指调查员倾向性的诱导患者的回答以支持其预先的假设。需要充分培训调查员，防止先入为主的观念。^[3]

3) 观察者偏倚（observer bias） 是指根据预先知道暴露的分组情况而对结果做出主观判断。针对需要主观判断的结局，尽量使调查员处于盲态，不了解患者的暴露分组情况。^[3]

4) 难以测量的时间偏倚（immeasurable time bias） 是指药物暴露的时间无法准确测量或被记录。这就需要尽可能收集全面的药物暴露信息。^[3]

5) 测量偏倚（measurement bias） 调查员对研究变量和数据进行测量时产生的偏倚，如仪器未校正、操作不规范、调查方法不统一等。可以通过设置严格的调查和操作流程，并进行调查员培训加以控制。^[3]

6) 非死亡时间偏倚（immortal time bias） 又称无风险时间偏倚（guarantee time bias）或时间依赖性偏倚（time-dependent bias）^[5]或恒定时间偏倚^[4]，或被翻译成永恒时间偏倚^[6]。所研究队列在特定期间内观察对象未见死亡或未出现结局事件，这段随访期间称为非死亡时间。当在进入队列和首次出现暴露之间的“非死亡时间”内，观察对象被错误地分类或简单地被排除，且未能在统计学分析中进行恰当处理，则会发生非死亡时间偏倚，因为在“非死

亡时间”内这些观察对象并非真正的“非死亡”。^[1]如果在后续分析中将该段时间错误地计算为暴露人时或完全剔除(即当队列定义时间、暴露发生时间、随访开始时间,这三个时间不相匹配时),就可能导非死亡时间偏倚。^[4]可以在开始随访前通过评估队列的纳入与排除标准和暴露测量标准,来避免这类偏倚的发生,或者采用一种合适的人时(person-time)方法来减少这类偏倚。^[3]

7) 诱导偏倚(inducement bias) 研究者的询问技术不当,或为了取得阳性结果,诱使研究对象作出某一倾向性的回答。这种诱导偏倚往往表现为对试验组做诱导而对对照组不做诱导或负向诱导。^[2]

8) 因果倒置(reverse causation) 是指当治疗和结局之间的联系是结局的状态影响治疗的选择,而不是治疗影响结局这一事实所导致的,就会发生因果倒置。可以通过设置药物暴露后的一段滞后时间来进行控制。例如,在研究恶性肿瘤与镇痛药物的关联时,肿瘤在诊断之前引起的癌性疼痛可能导致镇痛药物的使用,此时肿瘤已经发生,药物对肿瘤的发病没有作用。随着疾病发展,肿瘤的临床特征逐渐明显而被诊断,在RWD中就会呈现患者肿瘤发生(实际为诊断)之前有镇痛药物使用记录,误导研究者认为镇痛药物与肿瘤发病风险之间存在关联。在一些具有较长诱导期,例如上述恶性肿瘤的研究中,因果倒置的现象更为突出,这通常可以通过设置药物暴露后的一段滞后时间来进行控制^[4]

9) 校正中介作用(adjusting for causal intermediaries) 是指在医疗数据库中进行药物有效性比较的病例对照研究中,一种常见的错误类型是协变量和暴露测量的时间窗不一致,比如暴露测量的时间窗更长,使得两者信息不同步而导致结果的偏倚。研究者应避免调整治疗和结局因果关系链上的变量,因为这种调整可能导致对总体治疗效果和不对中介变量进行调整的直接效果产生有偏估计。此外,将治疗的时间纳入需要调整的基线变量中也会导致错误地调整中介作用。某种药物的使用时间通常受到初始治疗选择的影响,接着会影响结局发生的风险,使得这个变量成为了治疗和结局之间的中介变量(intermediary)。^[4]

B.3 混杂偏倚

混杂(confounding)是指由于一个或多个外来因素的存在,掩盖、夸大或缩小研究因素与结局之间的关联,从而部分或全部地歪曲了两者间的真实联系,造成测量结果与真实结果的误差,称之为混杂。引起混杂的因素称为混杂因素。通常情况下,一个外来因素须具备以下情况才可能引起混杂:(1)该因素与研究结局有关,是其危险因素或影响因素之一。

(2)该因素与研究因素有关,并且在所研究的各组之间分布不均。(3)该因素不是研究因素与研究结局因果链上的中间变量(intermediate variable)。^[2]

药物流行病学中还比较常见的是适应证混杂(confounding by indication),即具有某些适应证或医学问题的患者更容易服用某种药品,从而容易对该药品的评价产生混杂作用,如医生更倾向于开具处方给那些具有潜在疾病的患者,而未服用该药品的患者本身发生疾病的风险就比较低,从而形成了适应证混杂。合并用药混杂(confounding by comedication):是指临床上存在合并用药的情况,这些合并用药给所研究药品的评价带来混杂作用。^[2]

利用真实世界数据开展的临床研究中,与处方选择性有关的混杂分为来源于患者、医生和医疗保险体制的三类混杂^[1]。(1)来源于患者的混杂:病情较重、久治不愈或对现有药

品无法耐受的患者,往往倾向于选择某种新药或更好的药品。这将会导致患者组间可比性较差,从而无法对药物的有效性或安全性进行准确评估如:接受了药品广告宣传的患者可能会特别要求医生开具该药物,若如此也可能产生混杂。(2)来源于医生的混杂:医生日常开具处方的行为往往存在很大倾向性。对于病情较轻或者常规药品可以很好控制的患者,医生可能不会开具另一种较昂贵的二线、三线药品或者尝试新上市药品,而往往可能是病情较重、已有药品效果不好,患者出现药品不良反应无法耐受及其他特殊的情况下,医生才会决定使用新药。另外,不同医生处方用药的选择性也会存在一定差异,也可能间接对于该药品的评价产生混杂作用。(3)来自于医疗保险体制的混杂:由于药品是否被纳入医疗保险的覆盖范围存在差异,导致不同经济条件的患者群体在临床实际中对药物的选择性使用倾向和行为可能有所不同。

此外,文献中还有报告其他混杂的现象,如时间依赖性混杂,它指的是同时充当混杂因素和中间步骤的变量。例如,在研究使用阿司匹林对心脏病死亡的影响时,心肌梗死事件便是一个时间依赖性混杂。因为发生心肌梗死是后续服用阿司匹林的原因,也是后续发生心脏死亡的危险因素;同时,阿司匹林使用也可预防心肌梗死的发生。因此,心肌梗死事件同时扮演了混杂(之前的心肌梗死导致患者使用阿司匹林)和中间步骤(随访心肌梗死事件的发生受阿司匹林使用的影响)的角色。可通过严格限制纳入与排除标准,对资料进行分层分析加以控制。复杂的时间依赖性混杂可采用边际结构模型(marginal structural models)、嵌套结构模型(structural nested models)等复杂统计分析方法。然而,即使是详细、完整的数据源,采用多种设计、分析方法,依然可能无法完全识别和测量潜在的混杂因素。^[7]

沾染(contamination)与干扰(co-intervention) 沾染是指对照组的对象意外地接受了试验组的试验措施,而使得两组的最终结果差异缩小的情况。如观察阿司匹林预防血栓性疾病的队列研究中,曾有对照组患者因感冒多次服用阿司匹林,从而对最终结果产生影响。而干扰是指试验组试验对象有意或无意地额外接受了研究以外的、类似试验措施的药物,而使结果夸大且有利于试验组,称为干扰。^[2]

参考文献

- [1] 曾繁典,郑荣远,詹思延,杜晓曦.药物流行病学[M].2版.北京:中国医药科技出版社,2016.
- [2] 刘建平,主编.循证中医药临床研究方法[M].北京:人民卫生出版社,2009:61-68.
- [3] 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.
- [4] 石舒原,赵厚宇,周庆欣,等.真实世界证据与随机对照试验:RCT DUPLICATE项目方法学介绍[J].药物流行病学杂志,2020,29(03):198-205.
- [5] 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.
- [6] 廖星,章轶立,谢雁鸣.真实世界研究标准:RECORD清单和GRACE清单的解读[J].中国中药杂志,2015,40(24):4734-4738.

[7] 吴晶,朱玄,刘程宇.基于医疗大数据的比较效果研究设计要点述评与应用[J].中国药物经济学,2019,14(05):26-30.

CACM标准公示稿

附录 C
(资料性)
真实世界证据质量评价工具和报告规范汇总表

使用说明：下表提供了真实世界证据质量的方法学评价工具和报告规范，但并未在表中列出评价中医药相关证据质量时需要特殊考虑的内容，请结合本文件正文“5.3 评价中医药真实世界证据质量的特殊考虑”及规范中其他相关论述来使用。例如将真实世界证据应用于为中药研发提供证据时需考虑：研究的患者特征应与目标人群的相似；干预措施不应违背中医理论，且应符合日常临床实践模式；对照措施应符合日常临床实践模式，对照措施的选择须与干预措施的药理作用、剂型、给药途径等因素方面保持较好的同质性；结局指标的设置应与中医临床实践需要一致，同时顾及研究结局的国际及现代医学视角下的可理解并可利用的程度等。本汇总表仅基于国际现有工具和规范进行汇总，随着真实世界研究规范的不不断研制，还需使用者根据最规范和研究成果合理使用和评价。多来源、多结局的 RWS 质量评价仍需考虑其内容结构。本汇总表基于国际现有工具和规范构建，随着真实世界研究规范的不不断研制，还需使用者根据最新规范和研究成果合理使用和评价。

真实世界证据	评价工具	报告规范
观察性研究的证据质量评价及报告规范		
真实世界前瞻性队列研究证据	国际上并未形成非常成熟且公认的评价工具。可以参考 ArRoWs 评价条目	采用国际现有报告规范，如常规收集医疗卫生数据开展观察性研究 [The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement, RECORD]，以及针对药物流行病学研究的版本 (RECORD-PE) 清单等
真实世界双向性队列研究证据		
真实世界巢式病例对照研究证据		
前瞻性队列研究证据	采用国际现有评价工具，如 GRACE、NOS 量表中队列研究部分的条目，CASP 队列研究清单，以及可以借鉴 ArRoWs 评价条目中的通用条目及队列研究条目	采用国际现有报告规范，如加强观察性流行病学研究报告的质量声明 [The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies]
回顾性队列研究证据		
双向性队列研究证据		
病例对照研究证据	采用国际现有评价工具，如 NOS 量表中病例对照部分的条目，CASP 病例对照清单等，以	

	及可以借鉴 ArRoWs 评价条目中的通用条目及病例对照条目	
巢式病例对照研究证据	采用国际现有评价工具，如 NOS 量表、CASP 队列及病例对照条目	
横断面研究证据	采用国际现有评价工具，如 AHRQ 横断面研究评价标准。	
注册研究证据	采用相应设计类型的评价工具及方法	
病例系列证据	采用国际现有评价工具，如 NICE、JBI 和 IHE 研制的病例系列质量评价工具	尚无国际通用报告规范，建议报告内容对前列病例系列方法学质量评价工具中提及的要点进行详细报告，并参考 STROBE 标准中的部分内容。
病例报告证据	采用国际现有评价工具，如 JBI 研制的病例报告质量评价工具	采用国际现有报告规范，如基于共识的临床病例报告指南的制定（The CARE Guidelines: Consensus-based Clinical Case Reporting Guideline development）《中医病例报告建议条目》《基于共识的中医病例报告规范》（Consensus-based recommendations for case report in Chinese medicine）
干预性研究的质量评价及报告规范		
基于随机对照设计的证据质量评价		
实用性随机对照试验证据	采用国际现有评价工具，如 Cochrane RoB、RoB2，及其他国际较为公认的工具如改良 Jadad 量表	采用国际现有报告规范，如实用性随机对照试验报告规范（Improving the reporting of pragmatic trials: an extension of the CONSORT statement），使用队列或收集日常数据的随机对照试验的报告规范 CONSORT 扩展版 [CONSORT extension for the reporting of randomized controlled trials conducted using cohorts and routinely collected data (CONSORT-ROUTINE) checklist with explanation and elaboration]，及其他如阶梯楔形群组随机对照试验报告规范：CONSORT 2010 声明扩展及解释（Reporting of stepped wedge cluster

		randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration) 及其他扩展版等。结合草药干预措施随机对照试验报告: CONSORT 扩展声明(Reporting Randomized Controlled Trials of Herbal Interventions: An Elaborated CONSORT Statement)、中药复方临床随机对照试验报告规范 2017 (CONSORT Extension for Chinese Herbal Medicine Formulas 2017: Recommendations, Explanation, and Elaboration, CONSORT-CHM Formulas) 中与中药相关的要求。
真实世界随机交叉临床试验证据	采用国际现有评价工具, 如 Cochrane RoB、RoB2, 及其他国际较为公认的工具如改良 Jadad 量表, 并在此基础上根据改良后的真实世界证据的特殊性进行评价	采用国际现有报告规范, 如 CONSORT 声明随机交叉试验扩展版 (CONSORT 2010 statement: extension to randomized crossover trials) 结合草药干预措施随机对照试验报告和 CONSORT-CHM Formulas 中与中药相关的要求
真实世界单病例随机对照试验证据		采用国际现有报告规范, 如 CONSORT 声明单病例随机对照试验扩展版 (CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement]) CONSORT 中医药单病例随机对照试验扩展版 (CONSORT extension for reporting N-of-1 trials for traditional Chinese medicine (CENT for TCM): Recommendations, explanation and elaboration)
真实世界随机征求许可试验证据		尚无国际通用报告规范, 但在报告该研究类型时须注意详细报告实际各组病例数及换组率的情况, 结合使用草药干预措施随机对照试验报告和 CONSORT-CHM Formulas 中与中药相关的要求。
真实世界基于患者意愿的随机对照试验证据		
基于非随机对照设计的证据质量评价		
使用真实世界数据/非安慰剂研究数据/观察性研究数据作为对照的单臂试验证据	尚未找到国际规范。建议参考本文件正文中所列要点进行评价。	尚无国际通用报告规范, 建议报告内容对前列单臂试验方法学质量要点中提及的内容进行详细报告, 并结合使用草药干预措施随机对照试验报告和 CONSORT-CHM Formulas 中相关的要求。
无对照单臂试验证据		

实用性同期非随机对照证据	经典非随机对照研究设计类型的质量评价工具，如 MINORS、ROBINS-I，并在此基础上根据改良后的真实世界证据的特殊性进行评价	非随机对照试验报告规范 (Transparent reporting of evaluations with nonrandomized designs, TREND) 结合使用 草药干预措施随机对照试验报告和 CONSORT-CHM Formulas 中针对草药增设的条目报告
真实世界历史对照证据		
真实世界自身对照证据		
真实世界交叉试验证据		
证据综合的方法学评价工具及要点		
基于前述各种研究的系统综述与 meta 分析证据	参考现有证据综合质量评价工具，如评价单项证据综合 (系统综述) 实施质量的 AMSTAR2 及评价 meta 分析结果证据等级的 GRADE 体系	系统评价和 Meta 分析优先报告的条目 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, PRISMA 2020)；当用于报告中中药研发相关内容时，还需参照其扩展版中草药系统评价和 Meta 分析优先报告的条目 [Preferred Reporting Items for Systematic Reviews and Meta-Analyses: Extension for Chinese Herbal Medicines 2020 (PRISMA-CHM 2020)]

注 1: RoB, risk of bias, Cochrane 协作网的偏倚风险评价工具; MINORS, methodological index for non-randomized studies, 非随机对照试验方法学评价指标; ROBINS-I, risk of bias in non-randomized studies of interventions, 非随机对照试验偏倚风险评价工具; GRACE, good research for comparative effectiveness, 高质量比较效果研究准则; NOS, Newcastle-Ottawa Scale, NOS 量表; CASP, critical appraisal skill program, 关键评估技术计划; ArRoWs, Assessment of Real-World Observational Studies, 真实世界观察性研究评估工具; AHRQ, Agency for Healthcare Research and Quality, AHRQ 横断面研究评价标准; NICE, National institute for clinical excellence, 英国国立临床优化研究所; JBI, Joanna Briggs institute, 澳大利亚乔安娜·布里格斯学院; IHE, Institute of Health Economics, 加拿大卫生经济研究所; CARC, Consensus-based recommendations for case report in Chinese medicine, 中医病例报告; AMSTAR, Assessing the Methodological Quality of Systematic Reviews, 评估系统评价的方法学质量; GRADE, the grading of recommendations assessment, development and evaluation, 推荐、评估、发展和评价分级

参考文献

- [1] 国家药品监督管理局.《真实世界证据支持药物研发与审评的指导原则(试行)》[S/OL].(2020-07-17)[2020-12-12] <http://www.nmpa.gov.cn/WS04/CL2138/373175.Html>
- [2] 国家药品监督管理局.《中药注册管理专门规定(征求意见稿)》[S/OL].(2020-04-30)[2020-12-12].<https://www.nmpa.gov.cn/xxgk/zhqyj/zhqyjyp/20200430154501315.html>
- [3] 吴阶平医学基金会,中国胸部肿瘤研究协作组.真实世界研究指南 2018[M/OL].(2018-08-03)[2021-05-07].<https://max.book118.com/html/2018/1108/5133311122001330.shtm>
- [4] 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.
- [5] 温泽淮,李玲,刘艳梅,等.实效性随机对照试验的技术规范[J].中国循证医学杂志,2019,19(7):794-802.
- [6] 中国临床医学真实世界研究施行规范专家委员会.中国临床医学真实世界研究施行规范[J].中华实验和临床感染病杂志(电子版),2017,11(6):521-525.
- [7] 吴宇侠,康德英,洪旗,等.临床随机对照试验的外部真实性及其评价的探讨[J].中华流行病学杂志,2011(5):514-518.
- [8] Fan AY. The methodology flaws in Hinman's acupuncture clinical trial, Part II: Zelen design and effectiveness dilutions. *J Integr Med.* 2015 May;13(3):136-9. doi: 10.1016/S2095-4964(15)60172-8. PMID: 26006026.
- [9] 温泽淮,李玲,刘艳梅,等.实效性随机对照试验的技术规范[J].中国循证医学杂志,2019,19(7):794-802.
- [10] 李欣霖,张雅静,刘建平,等.考虑患者偏好的部分随机临床试验在中医非药物治疗疗效评价中的设计与应用[J].中医杂志,2016,57(20):1750-1753+1767.
- [11] 于明坤,于河,王迪,等.目标值法在名老中医经验传承临床研究中的方法学价值和设计要点[J].中医杂志,2020,61(1):36-41.
- [12] 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,2016:80-87.
- [13] 刘建平.循证中医药临床研究方法[M].北京:人民卫生出版社,2009:61-68.
- [14] 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.
- [15] 石舒原,赵厚宇,周庆欣,等.真实世界证据与随机对照试验:RCT DUPLICATE项目方法学介绍[J].药物流行病学杂志,2020,29(3):198-205.
- [16] 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.
- [17] 廖星,章轶立,谢雁鸣.真实世界研究标准:RECORD清单和GRACE清单的解读[J].中国中药杂志,2015,40(24):4734-4738.

[18] 吴晶,朱玄,刘程宇.基于医疗大数据的比较效果研究设计要点述评与应用[J].中国药物经济学,2019,14(5):26-30.

[19] NANCY AD, SEBASTIAN S, BARBARA JM, et al. GRACE principles: recognizing high-quality observational studies of comparative effectiveness[J]. Am J Manag Care, 2010, 16(6): 467

[20] WELLS G, SHEA B, O'CONNELL D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomized studies in meta-analyses[Z/OL]. (2011-01-01)[2020-12-12]. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

[21] Critical Appraisal Skills Programme. CASP (Case Control Study) Checklist[Z/OL].(2019-01-01)[2020-12-12].<https://casp-uk.net/casp-tools-checklists/>

[22] 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[J]. Ann Epidemiol, 2021,55:57-63. doi: 10.1016/j.annepidem.2020.09.014.

[23] ROSTOM A, DUBE C, CRANNEY A, et al. Appendix D. Quality Assessment Forms: Cross-Sectional/Prevalence Study Quality[M/OL]//Celiac Disease:Evidence Reports/Technology Assessments, No. 104. Rockville: Agency for Healthcare Research and Quality (US).(2004-09-01)[2020-12-12].<http://www.ncbi.nlm.nih.gov/books/NBK35156/>

[24] National Institute for Clinical Excellence. Quality assessment for case series [EB/OL]. (2015-01-01)[2020-12-12].<http://www.nice.org.uk/guidance/cg3/resources/appendix-4-quality-of-case-series-form2>.

[25] MOOLA S, MUNN Z, TUFANARU C, et al. Chapter 7: Systematic reviews of etiology and risk[M/OL]//AROMATARIS E, MUNN Z (Editors). JBI Manual for Evidence Synthesis. JBI, 2020. <https://synthesismanual.jbi.global>.

[26] Institute of Health Economics. IHE Quality Appraisal Checklist for Case Series Studies[Z/OL]. (2016-01-01)[2021-05-11].<https://www.ihe.ca/advanced-search/ihe-quality-appraisal-checklist-for-case-series-studies>.

[27] HIGGINS JPT, GREEN S. Cochrane handbook for systematic reviews of interventions version 5.1.0 [M/OL].(2011-03-01)[2020-12-12].<http://www.cochrane-handbook.org>

[28] MATTHEW P, CHRIS W, DARIO S. RoB 2.0: A revised tool to assess risk of bias in randomized trials[Z/OL].(2016-11-01)[2021-05-11].
<https://training.cochrane.org/resource/rob-20-webinar>

[29] JADAD AR, MOORE RA, 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.

[30] 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.

[31] STERNE JA, HERNÁN MA, REEVES BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions[J]. *BMJ*,2016,355:i4919. doi: 10.1136/bmj.i4919.

[32] SHEA BJ, GRIMSHAW JM, WELLS GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews[J]. *BMC Med Res Methodol*, 2007,7:10. doi: 10.1186/1471-2288-7-10.

[33] LECLERCQ V, HILIGSMANN M, PARISI G, et al. Best-worst scaling identified adequate statistical methods and literature search as the most important items of AMSTAR2 (A measurement tool to assess systematic reviews) [J]. *J Clin Epidemiol*, 2020,128:74-82. doi: 10.1016/j.jclinepi.2020.08.011.

[34] GUYATT GH, OXMAN AD, KUNZ R, et al. What is "quality of evidence" and why is it important to clinicians? [J]. *BMJ*, 2008,336(7651):995-8. doi: 10.1136/bmj.39490.551019.

[35] VANDENBROUCKE JP, VON ELM E, ALTMAN DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration[J]. *Int J Surg*, 2014,12(12):1500-1524. doi:10.1016/j.ijsu.2014.07.014

[36] BENCHIMOL EI, 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.

[37] 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.

[38] 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.

[39] GAGNIER JJ, KIENLE G, ALTMAN DG, et al. The CARE Guidelines: Consensus - Based Clinical Case Reporting Guideline Development[J]. *Headache*,2013,53(10):1541-1547.

[40] 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.

[41] FU SF, CHENG CW, ZHANG L, et al. Consensus-based recommendations for case report in Chinese medicine (CARC)[J]. *Chin J Integr Med*, 2016, 22 (1): 73-79

[42] ZWARENSTEIN M, TREWEEK S, GAGNIER JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement[J]. *BMJ*, 2008, 337:a2390. doi: 10.1136/bmj.a2390.

[43] DWAN K, LI T, ALTMAN DG, et al. CONSORT 2010 statement: extension to randomised crossover trials[J]. *BMJ*, 2019, 366: 14378. doi: 10.1136/bmj.14378.

[44] Sunita, Vohra, Larissa, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement[J]. *BMJ (Clinical research ed.)*, 2015 355: i5381. doi: 10.1136/bmj.i5381.

[45] GAGNIER JJ, 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.

[46] CHENG CW, WU TX, SHANG HC, et al. CONSORT Extension for Chinese Herbal Medicine Formulas 2017: Recommendations, Explanation, and Elaboration[J]. *Ann Intern Med*,2017, 167(2):112-121.

[47] 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[J]. *Complement Ther Med*, 2019,46:180-188.https:// doi.org/10.1016/ j.ctim.2019.08.014.

[48] 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[J]. *BMJ*, 2021,373: n857. doi: 10.1136/bmj.n857

[49] VLAHOV D. Transparent Reporting of Evaluations with Nonrandomized Designs (TREND)[J]. *J Urban Health*, 2004,81(2):163-164.

[50] PAGE MJ, MCKENZIE JE, BOSSUYT PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews[J]. *BMJ*, 2021,372:n71. doi: 10.1136/bmj.n71.

[51] 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).[J] *Am J Chin Med*,2020,48(6):1279-1313.