

DOI:10.13288/j.11-2166/r.2022.03.017

标准与规范

中医药真实世界研究技术规范 ——证据质量评价与报告



中华中医药学会《中医药真实世界研究技术规范》制订组

2020 年国家药品监督管理局发布的《真实世界证据支持药物研发与审评的指导原则（试行）》^[1]，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。同年，国家药品监督管理局公布《中药注册管理专门规定（征求意见稿）》^[2]，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。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。

本规范旨在规范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，因此在使用本规范时，还需紧密结合国内外最新研究成果。本规范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，虽然并非为指导设计真实世界研究而制定，但规范内容可以用来间接作为设计工作的借鉴。本规范由中华中医药学会发布，标准编号为 T/CACM1361—2021，发布时间为 2021 年 6 月 30 日，实施时间为 2021 年 6 月 30 日。

1 编制过程

1.1 成立起草组

本规范起草工作组由 24 位成员组成，其中博士学位 17 人、硕士学位 4 人、本科学历 3 人，包括

13 位临床研究方法学专家、4 位临床医生、2 位临床药理专家、1 位伦理审查专家、4 位统计学专家。

1.2 文献检索与梳理

1.2.1 检索策略 检索国内外文献：1) 数据库包括中国知网 (CNKI)、万方数据知识服务平台 (WF)、维普网 (VIP)、中国生物医学文献服务平台 (SinoMed)、PubMed、Cochrane Library、Embase，检索时间范围为各数据库建库至 2020 年 11 月 30 日；2) 国内外政府卫生机构网站近 5 年文件。根据各数据库情况，将“真实世界”相关检索词同时在“题名、摘要、关键词”中检索，将“证据”“方法学”“质量”“评价”相关检索词在“全文”中检索。

1.2.2 文献纳入标准 1) 涉及真实世界研究偏倚情况、真实世界证据质量评价、真实世界研究报告的文献；2) 文献类型包括政策法规、指南及专家共识、方法学研究、临床研究、综述、学位论文、专著；3) 文献中仅提及偏倚、证据质量、报告规范，但未对具体内容提供有用信息者仅作为文献计量分析。4) 如果文章观点明确属于二次引用，即追溯原文，此文章纳入文献计量分析；5) 同一内容重复发表的文献，仅保留首次发表的文献；6) 不限定地域和语种。

1.2.3 文献排除标准 1) 无法获取全文的文献，如会议摘要、保密文献；2) 文献内容雷同者，仅保留信息较多的文献，其他予以排除。

1.3 共识过程

根据文献研究结果形成初稿，经过 12 次内部讨论会、2 次专家共识会，以及对国内真实世界研究资深专家深度访谈等环节，完善规范构架、主要内容及中医特色内容等。通过函审形式发送给国内方法学领域 30 位专家审理，专家均为副高级及以上职称，得到回函 30 封，回收率 100%。专家分别

基金项目：国家自然科学基金 (82074282)；国家重点研发计划 (2019YFC1709300)；中药临床研究与评价重点实验室项目

✉ 通讯作者：yutong_fei@163.com

来自北京市、上海市、天津市、广东省、河北省、黑龙江省、湖北省、江苏省、江西省、辽宁省、山东省、四川省的 26 家单位。根据专家意见继续完善相关内容,并经过中华中医药学会的再次审查后同意发布。

2 真实世界研究设计分类与真实世界证据

2.1 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)》^[1]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。本规范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(包括前瞻性、回顾性、双向性队列研究)、注册研究、病例-对照研究、横断面研究、病例系列和病例报告。除此之外,仍有不能归为上述类型的,以数据为导向的医疗管理信息分析与大数据挖掘类研究。经典临床研究设计与真实世界研究设计具体信息见附件 1(请扫描本文二维码获取)。

2.2 真实世界证据

真实世界证据来源于真实世界研究,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。在真实世界研究设计所形成证据的基础上,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;此外,还包括基于医疗管理信息分析与大数据挖掘所形成的研究证据。

3 真实世界证据的质量评价

3.1 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)》^[1],评价真实世界证据应依从两个主要原则:一是真实世界证据是否可以支持需要回答的临床问题;二是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,本规范在不同类型真实世界证据质量的评价方法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。

3.2 真实世界研究常见偏倚

真实世界研究的常见偏倚包括三大类^[3-9]。1) 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、特发性偏倚、病程长度偏倚、竞争风险;2) 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;3) 混杂偏倚:适应证混杂、合并用药混杂、沾染与干扰。真实世界研究常见偏倚的类型见附件 2(请扫描本文二维码获取)。

3.3 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。例如,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;其次,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。

3.3.1 评价研究对象的设定 1) 应考虑中医的病、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,以及用语是否规范。2) 应考虑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) 应考虑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(ICD)是否具有同等一致性等。

3.3.2 评价暴露/干预措施的设定 1) 应根据研究目的判断暴露或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、分类是否处于同一水平;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围。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；明确不同煎煮方式、剂型是否带来疗效差异。3) 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。4) 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。5) 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。6) 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。7) 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。

3.3.3 评价非暴露/对照措施的设定 如果需要设置对照，则应评价非暴露/对照措施的选择是否合理：1) 应判断对照措施/非暴露定义的清晰性，以及对其选择是否体现了研究目的。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对所研究疾病公认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，应考虑其理由是否充分、合理。2) 非暴露/对照措施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) 要判断对照组患者的主要预后特征或病情特征与暴露组/干预组的特征是否可比。

3.3.4 判定评价结局的设定是否合理 1) 结局指标的设置应与中医临床实践需要相一致，同时应考虑研究结局在国际及现代医学视角下的可理解与可应用的程度。建议主要结局采用国际公认的结局指标，次要结局可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指标，如未经合理测评的由多维度指标构成的有效率等。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该证候量表科学性不强或相应评分的变化是由其他证候变化所导致的。2) 结局指标的选择可参考相应核心结局指标集，如有效性试验核心结局指标测量（COMET），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，对所有试验所用药物的已知不良反

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安全性观测不足而导致的。要注意及时对不良事件是否要归属于不良反应做出判断。考察不良事件/反应的用语应符合国际相关规范，如国际医学用语辞典（MedDRA）^[10] 等。4) 评价结局指标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，如最小临床重要差异。

3.3.5 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用药时间与相配合的西药用药时间的关系。

3.3.6 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，或存在合并用药等情况时，应判断其是否采用了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的盲法。还需特别注意中医药个体化治疗结论的外推性问题等。

3.4 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。内部真实性评价主要通过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、结局和研究场所与拟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。真实世界证据质量评价工具和报告规范汇总表见附件3（请扫描本文二维码获取）。

经典研究设计类型直接形成的真实世界证据质量的评价方法：1) 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；2) 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；3) 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质量的评价方法：1) 优先选择真实世界研究特异性评价工具及条目进行评价；2) 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；3) 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；4) 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。

3.4.1 观察性研究的证据质量评价方法 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 (GRACE)^[11]、NOS 量表^[12] 中队列研究部分的条目、CASP 清单^[13] 中队列研究清单等。针对病例-对照研究证据的质量评价，应首先采用现有评价工具，如 NOS 量表^[12] 中病例-对照部分的条目，CASP 清单^[13] 中病例-对照清单等，还可以借鉴真实世界观察性研究评估工具 (ArRoWs)^[14]。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，如美国卫生保健质量和研究机构 (AHRQ)^[15] 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。针对病例系列报告证据的质量评价，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 (NICE)^[16]、澳大利亚乔安娜布里格斯研究所 (JBI)^[17] 和加拿大卫生经济研究所 (IHE)^[18] 研制的病例系列质量评价工具。针对病例报告证据的质量评价，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。

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。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-对照研究设计产生的证据，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。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，建议 ArRoWs 评价条目^[14] 和 NOS 量表^[12] 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。还应注意真实医疗环境下，对照组选择是否恰当，以及其他混杂因素的处理。

3.4.2 干预性研究证据的质量评价方法 基于随机对照设计的证据质量评价方法：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。实用性随机对照试验和技能型随机对照试验产生的证据均属于经典研究设计直接形成的真实世界证据类型，但技能型随机对照试验与本规范适用目的不符，故本段仅详细说明对实用性随机对照试验的证据质量评价方法。应选择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 (RoB)^[19]。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 2.0^[20] 及其他国际上较为公认的工具如改良 Jadad 量表^[21] 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，例如在该治疗领域中所选的干预措施、对照措施等是否为当前最具竞争力的最佳临床实践措施；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。

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：真实世界随机交叉试验证据、真实世界单病例随机对照试验证据、真实世界随机征求许可试验证据、真实世界基于患者意愿的随机对照试验证据，均属于经典随机对照研究设计类型改良后形成的真实世界证据。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^[19]、RoB 2.0^[20]、改良 Jadad^[21]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利因素还是不利因素。

基于非随机对照设计的证据质量评价方法:目前尚未发现国际上有针对无对照单臂试验证据的质量评价方法。可以考虑从如下方面进行评价:研究目的是否清晰,纳入与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,主要干预措施或联合干预措施的规定是否清楚明确,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,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的考虑与分析,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。

目前也尚未发现使用真实世界数据/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的国际规范。建议在前述评价要点的基础上,进一步考虑外部对照数据来源数据集或研究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部对照相关特性的可比性,以及其对研究结果的精确性、研究结论的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部对照的时候,上述情况所带来的挑战可能由于数据的复杂性而变得更加需要慎重对待。为克服或减少这些局限:1) 须

确保所采集的数据符合真实世界数据的适用性要求;2) 平行外部对照设计优于历史对照,因平行外部对照可采用疾病登记模式,更有利于保障数据记录尽可能完整、准确;3) 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;4) 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。

目前亦未发现国际上针对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(MINORS)^[22]、非随机对照试验偏倚风险评价工具(ROBINS-I)^[23]等。根据实际情况还可以考虑多种评价方法结合使用。此外,还需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。

3.4.3 医疗管理信息分析与大数据挖掘形成的证据的方法学评价工具及要点 医疗管理信息分析与大数据挖掘的研究方法尚未形成明确的体系,评价时应注意:1) 数据的完整性;2) 数据的准确性;3) 数据的丰富性;4) 数据的真实可靠程度;5) 数据的数量是否足够支持其所选用的数据挖掘方法;6) 数据内容、数量、质量及统计分析方法的使用是否足以获得研究结论;7) 数据是否对中医诊疗信息表达充分可靠,人群、干预或暴露、结局是否事先界定且分类清晰;8) 研究的局限性是否得到充分的讨论,并在结论中得到正确表达。当嵌套在上述研究设计分类中开展研究时,对其进行评价应首先明确其嵌套在何种研究设计类型形成的证据,再根据上述相应的证据评价方法进行质量评价。

3.4.4 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^[24],用于干预性研究证据综合的 AMSTAR 2^[25],以及评价证据等级的 GRADE 体系^[26]。由于真实世界原始研究类型种类繁多,在进行证据综合时,建议将同一类别的原始研究进行综合。而对于跨类别合并的证据,目前尚缺少评价工具,故如确需合并,建议拆分为亚组并进行敏感性分析。对于定量综合的评价,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评价维度。

4 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。在进行证据报告时,也可从以下3个方面进行:1)将国际现有报告规范及条目作为必要条目进行报告;2)尚无报告规范的特殊研究类型,可将其在实施方面的特殊考虑作为补充条目进行报告;3)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。详见附件3(请扫描本文二维码获取)。

4.1 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》(STROBE)^[27]。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,均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《常规收集医疗卫生数据开展观察性研究的报告规范》(RECORD)^[28],以及针对药物流行病学的版本(RECORD-PE)^[29-30]清单等,并结合各研究证据类型的特点进行报告。

针对病例系列证据,目前尚无国际通用的报告标准,建议参考前述病例系列方法学质量评价工具对其要点进行详细报告,并参考STROBE^[27]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《基于共识的临床病例报告规范》(CARE)^[31]。当用于报告中药研发相关研究时,还需参照《中医病例报告建议条目》^[32]和《基于共识的中医病例报告规范》(CARC)^[33],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,以证明相关病例报告的价值;应着重讨论与古籍或他人发表的相关论文的比较,说明其诊断及治疗上的异同。

4.2 干预性研究的证据报告规范

4.2.1 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《实用

性随机对照试验报告规范》^[34];真实世界随机交叉试验报告应首先参照《CONSORT 2010声明:随机交叉试验扩展版》^[35];真实世界单病例随机对照试验报告则首先参照《单病例随机对照试验CONSORT扩展版》(CENT)^[36]。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规范》^[37]、《中药复方临床随机对照试验报告规范》(CONSORT-CHM Formulas)^[38]、《中医药单病例随机对照试验CONSORT扩展版》(CENT for TCM)^[39]。当随机对照试验作为嵌入式研究、巢式研究,或使用了队列数据及日常数据进行研究时(如注册研究、电子健康数据、管理数据库等),需参照《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)^[40]。

4.2.2 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》(TREND)^[41]。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,建议参考前述单臂试验方法学质量要点中提及的内容进行详细报告。此外,还建议结合使用《草药干预措施随机对照试验报告规范》^[37]、《中药复方临床随机对照试验报告规范》^[38]中相关的要求。

4.3 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》(PRISMA 2020)^[42];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》(PRISMA-CHM 2020)^[43]。详见附件3(请扫描本文二维码获取)。

5 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。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^[44],根据降级因素“偏倚风险”“不一致性”“间接性”“不

精确性”“发表偏倚”，及升级因素“效应量大”“剂量反应”“可能的剩余混杂因素”进行分级调整，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。

附：规范起草单位及起草组成员

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

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

✉通讯作者：yutong_fei@163.com

#与第一作者同等贡献

参与起草人：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

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)[EB/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)[EB/OL]. (2020-04-30)[2020-12-12]. <https://www.nmpa.gov.cn/xxgk/zhqyj/zhqyjyp/20200430154501315.html>.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] Medical Dictionary for Regulatory Activities (MedDRA). MedDRA term selection: points to consider release 4. 21 [EB/OL]. (2021-09-01)[2021-12-30]. <https://www.meddra.org/how-to-use/support-documentation/english/welcome>.
- [11] 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): 467-471.
- [12] 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 [EB/OL]. (2011-01-01)[2020-12-12]. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
- [13] Critical Appraisal Skills Programme. CASP (Case Control Study) checklist [EB/OL]. (2019-01-01)[2020-12-12]. <https://casp-uk.net/casp-tools-checklists/>.
- [14] COLES B, TYRER F, HUSSEIN H, et al. Development, content validation, and reliability of the Assessment of Real-World Observational Studies (ArRoWS) critical appraisal tool[J]. Ann Epidemiol, 2021, 55: 57-63. doi: 10.1016/j.annepidem. 2020. 09.014.
- [15] ROSTOM A, DUBÉ C, CRANNEY A, et al. Appendix D. Quality assessment forms: cross-sectional/prevalence study quality [M/OL]//Rockville: Agency for Healthcare Research and Quality (US), 2004. (2004-09-01)[2020-12-12]. <http://www.ncbi.nlm.nih.gov/books/NBK35156/>.
- [16] 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>.
- [17] The Joanna Briggs Institute. Checklist for Case Series [EB/OL]. (2021-04-01)[2021-06-12]. <https://jbi.global/critical-appraisal-tools>.
- [18] Institute of Health Economics. IHE quality appraisal checklist for case series studies [EB/OL]. (2016-01-01)[2021-05-11]. <https://www.ihe.ca/advanced-search/ihe-quality-appraisal-checklist-for-case-series-studies>.
- [19] HIGGINS JPT, GREEN S. Cochrane handbook for systematic reviews of interventions version 5. 1. 0 [EB/OL]. (2011-03-01)[2020-12-12]. <http://www.cochrane-handbook.org>.
- [20] MATTHEW P, CHRIS W, DARIO S. RoB 2. 0: a revised tool to assess risk of bias in randomized trials [EB/OL]. (2016-11-01)[2021-05-11]. <https://training.cochrane.org/resource/rob-20-webinar>.
- [21] 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.
- [22] 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.
- [23] STERNE JA, HERNÁN MA, REEVES BC, et al. ROB-INS-I: a tool for assessing risk of bias in non-randomised studies of interventions [J]. *BMJ*, 2016, 355: i4919. doi: 10.1136/bmj.i4919.
- [24] 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.
- [25] 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.
- [26] 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):995-998.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] GAGNIER JJ, KIENTLE G, ALTMAN DG, et al. The CARE guidelines: consensus-based clinical case reporting guideline development [J]. *Headache*, 2013, 53(10): 1541-1547.
- [32] 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.
- [33] 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.
- [34] 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.
- [35] DWAN K, LI T, ALTMAN DG, et al. CONSORT 2010 statement: extension to randomised crossover trials [J]. *BMJ*, 2019, 366: l4378. doi: 10.1136/bmj.l4378.
- [36] SUNITA V, LARISSA S, MARGARET S, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement [J]. *BMJ*, 2015, 350: h1738. doi: 10.1136/bmj.h1738.
- [37] 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.
- [38] 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.
- [39] 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. doi: 10.1016/j.ctim. 2019. 08. 014.
- [40] 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.
- [41] VLAHOV D. Transparent reporting of evaluations with nonrandomized designs (TREND) [J]. *J Urban Health*, 2004, 81(2):163-164.
- [42] 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.
- [43] 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.
- [44] BALSHEM H, HELFAND M, SCHÜNEMANN HJ, et al. GRADE guidelines: 3. Rating the quality of evidence [J]. *J Clin Epidemiol*, 2011, 64(4):401-406.

(收稿日期: 2021-06-11; 修回日期: 2021-10-25)

[编辑: 焦 爽]