

健康结局指标效应量决策阈值制定方法指南



刘雅菲^{1,2#}, 朱鸿飞^{3,4,5#}, 李颖^{1,2}, 赖鸿皓^{1,2}, 周奇⁶, 罗旭飞⁶, 陈耀龙^{6,7,8,9,10}, 邓宏勇¹¹, 杜亮¹², 靳英辉¹³, 孔德昭¹⁴, 黎国威¹⁵, 李江¹⁶, 李玲¹², 廖星¹⁷, 刘杰¹⁸, 史楠楠^{17,19}, 商洪才²⁰, 田金徽^{6,7,8,9,10}, 王晓辉¹, 吴大嵘²¹, 杨丰文²², 杨克虎^{6,7,8,9,10}, 于长禾²⁰, 张家兴²³, 赵晨¹⁷, 葛龙^{1,2,7,8,9,10}

1. 兰州大学公共卫生学院社会医学与卫生事业管理研究所 (兰州 730000)
2. 兰州大学公共卫生学院循证社会科学研究中心 (兰州 730000)
3. 复旦大学公共卫生学院 (上海 200032)
4. 复旦大学公共卫生安全教育部重点实验室 (上海 200032)
5. 复旦大学国家卫生健康委员会卫生技术评估重点实验室 (上海 200032)
6. 兰州大学基础医学院循证医学中心 (兰州 730000)
7. 兰州大学健康数据科学研究院 (兰州 730000)
8. 甘肃省循证医学与临床转化重点实验室 (兰州 730000)
9. 世界卫生组织指南实施与知识转化合作中心 (兰州 730000)
10. GRADE 中国中心 (兰州 730000)
11. 上海中医药大学中医健康协同创新中心 (上海 201203)
12. 四川大学华西医院中国循证医学中心 (成都 610041)
13. 武汉大学中南医院循证与转化医学中心 (武汉 430071)
14. 辽宁中医药大学附属医院 (沈阳 110032)
15. 暨南大学附属广东省第二人民医院 (广州 510317)
16. 中国医学科学院肿瘤医院 (北京 100021)
17. 中国中医科学院临床与基础医学研究所 (北京 100700)
18. 中国中医科学院广安门医院 (北京 100700)
19. 中国中医药循证医学中心 (北京 100700)
20. 北京中医药大学东直门医院 (北京 100700)
21. 广州中医药大学第二附属医院 (广东省中医院) (广州 510120)
22. 天津中医药大学循证医学中心 (天津 301617)
23. 贵州省人民医院药剂科 (贵阳 550002)

【摘要】 临床循证决策过程中需要基于最佳证据, 综合考虑干预措施利弊大小、证据确信度、成本与可行性等因素。通过制定患者健康结局指标效应量决策阈值 (decision threshold of effect size, DTES), 能够帮助利益相关者理解干预措施利弊大小、评估证据确信度和解读研究结果。MERGE 工作组充分借鉴国际经验, 通过小组讨论、半结构化访谈、专家共识和试点应用, 开发了一套包括 8 个步骤的健康结局指标 DTES 制定方法指南, 包括: 必要性评估、组建工作组、遴选患者重要结局、基于最佳证据制作情景示例、设计专家函询问卷、分析专家函询结果、面对面专家共识以及推广应用与后效评价。DTES 指南的开发, 以期为利益相关者制定不同领域健康结局指标的 DTES 提供方法指导。

【关键词】 效应决策阈值; 核心结局指标; MERGE; 方法学

A guideline for determining the effect size decision threshold of health outcomes

LIU Yafei^{1,2}, ZHU Hongfei^{3,4,5}, LI Ying^{1,2}, LAI Honghao^{1,2}, ZHOU Qi⁶, LUO Xufei⁶, CHEN Yaolong^{6,7,8,9,10}, DENG Hongyong¹¹, DU Liang¹², JIN Yinghui¹³, KONG Dezhao¹⁴, LI Guowei¹⁵, LI Jiang¹⁶, LI Ling¹², LIAO Xing¹⁷, LIU Jie¹⁸, SHI Nannan^{17,19}, SHANG Hongcai²⁰, TIAN Jinhui^{6,7,8,9,10}, WANG Xiaohui¹

DOI: 10.7507/1672-2531.202308133

基金项目: 国家自然科学基金项目 (编号: 82204931); 中央级公益性科研院所基本科研业务费 (编号: ZZ15-WT-05); 中国中医科学院科技创新工程 (编号: CI2021A05502); 甘肃省科技计划资助 (编号: 22JR5RA733)

通信作者: 葛龙, Email: gelong2009@163.com

#共同第一作者



WU Darong²¹, YANG Fengwen²², YANG Kehu^{6, 7, 8, 9, 10}, YU Changhe²⁰, ZHANG Jiaxing²³, ZHAO Chen¹⁷,
GE Long^{1, 2, 7, 8, 9, 10}

1. Department of Social Science and Health Management, School of Public Health, Lanzhou University, Lanzhou 730000, P. R. China
2. Evidence-Based Social Science Research Center, School of Public Health, Lanzhou University, Lanzhou 730000, P. R. China
3. School of Public Health, Fudan University, Shanghai 200032, P. R. China
4. Key Laboratory of Public Health Safety, Fudan University, Shanghai 200032, P. R. China
5. National Health Commission Key Laboratory of Health Technology Assessment, Fudan University, Shanghai 200032, P. R. China
6. Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou 730000, P. R. China
7. Institute of Health Data Science, Lanzhou University, Lanzhou 730000, P. R. China
8. Key Laboratory of Evidence Based Medicine and Knowledge Translation of Gansu Province, Lanzhou 730000, P. R. China
9. WHO Collaborating Center for Guideline Implementation and Knowledge Translation, Lanzhou 730000, P. R. China
10. Chinese GRADE Center, Lanzhou 730000, P. R. China
11. TCM Health Collaborative Innovation Center, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, P. R. China
12. Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, Chengdu 610041, P. R. China
13. Center for Evidence-Based and Translational Medicine, Zhongnan Hospital of Wuhan University, Wuhan 430071, P. R. China
14. Affiliated Hospital of Liaoning University of Traditional Chinese Medicine, Shenyang 110032, P. R. China
15. Guangdong Second Provincial General Hospital Affiliated to Jinan University, Guangzhou 510317, P. R. China
16. Cancer Hospital Chinese Academy of Medical Sciences, Beijing 100021, P. R. China
17. Institute of Basic Research in Clinical, China Academy of Chinese Medical Sciences, Beijing 100700, P. R. China
18. Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing 100700, P. R. China
19. China Center for Evidence-Based Traditional Chinese Medicine, Beijing 100700, P. R. China
20. Dongzhimen Hospital, Beijing University of Traditional Chinese Medicine, Beijing 100700, P. R. China
21. The Second Affiliated Hospital of Guangzhou University of Traditional Chinese Medicine (Guangdong Hospital of Traditional Chinese Medicine), Guangzhou 510120, P. R. China
22. Evidence-Based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, P. R. China
23. Department of Pharmacy, Guizhou Provincial People's Hospital, Guiyang 550002, P. R. China

Corresponding author: GE Long, Email: gelong2009@163.com

【Abstract】 In the clinical evidence-based decision-making process, factors including benefits and harms, evidence certainty, cost and feasibility of interventions should be considered based on the best evidence. The development of decision threshold of effect size (DTES) for patient health outcomes can help stakeholders understand the benefits and harms of interventions, assess evidence certainty, and interpret research results. Based on international experience, the MERGE Working Group through group discussions, semi-structured interviews, expert consensus, and pilot application developed a set of 8-step guidelines for the development of DTES for health outcomes. These steps included necessity assessment, formation of working groups, selection of patient important outcomes, creation of scenarios based on the best evidence, design of expert consultation questionnaires, analysis of expert consultation results, face-to-face expert consensus, dissemination of the application and reevaluation. The DTES guideline development is intended to provide methodological guidance for stakeholders to develop DTES for health outcome indicators in different domains.

【Key words】 Decision threshold of effect size; Core outcomes; MERGE; Methodology

临床循证决策过程中应基于最佳证据, 综合考虑干预措施利弊大小、证据确信度、成本与可行性等因素后做出决策^[1,2], 推荐分级的评估、制定与评价 (grading of recommendations assessment, development and evaluation, GRADE) 工作组建议将干预措施利弊大小划分为无意义 (trivial or small, unimportant or no effect)、较小、中等或较大效应四类, 对应的 3 个阈值分别为较小效应阈值 (small effect threshold=T1)、中等效应阈值 (moderate effect threshold=T2) 和大效应阈值 (large effect threshold=T3)。通过制定患者健康结局指标效应量决策阈值 (decision threshold of effect size,

DTES), 能够帮助利益相关者理解干预措施利弊大小、评估证据确信度和解读研究结果。

目前, DTES 多被用于进行结果解读或证据质量评价相关的方法学指导, 或基于具体适用范围研究 DTES 的制定。然而, 这些研究主要关注 DTES 的应用, 并未详细描述 DTES 的制定过程和步骤结果, 因而国内外尚无能够指导制定 DTES 的方法标准^[3]。为此, MERGE 工作组——一个专注于循证医学方法学创新并在多学科交叉、融合和应用的团队, 在充分借鉴国际经验基础上, 组织多学科领域的专家开发了针对患者健康结局指标效应量决策阈值制定的方法学指南, 以期利益相关者制定不

同领域健康结局指标的 DTES 提供方法指导。

1 方法

本研究通过以下 3 个步骤构建方法指南：

① 通过系统评价与方法学核心小组讨论形成 DTES 制定方法指南初稿；② 以药物治疗新型冠状病毒感染 (coronavirus disease 2019, COVID-19) 动态快速循证实践要览^[4]的制订为例，对方法指南初稿进行试点应用，确认各步骤的规范性与可操作性；③ 基于半结构化深入访谈与德尔菲共识收集多学科领域专家意见，修改完善方法规程并拟定终稿。

1.1 形成方法指南初稿

计算机检索 CNKI、WanFang Data、CBM、VIP、PubMed、Web of Science 和 Embase 数据库，补充检索谷歌网站，中文检索词包括：阈值、指南、Meta 分析、系统评价、卫生技术评估等；英文检索词包括：threshold*、guideline*、meta analys*、systematic review、health technology assessment*等。2 名研究者独立筛选文献以获取国内外已发表的 DTES 制定相关研究，提取纳入研究中 DTES 制定的整体思路及相关步骤，进行描述性分析。

1.2 方法指南初稿的试点应用

以制订“药物治疗新型冠状病毒感染快速动态循证要览”^[4]为背景，应用 DTES 指南初稿制定 COVID-19 患者重要结局指标的 DTES。通过预检索 COVID-19 结局指标 DTES 总结现有研究基础，进而判断针对国内感染者与可及治疗方案制定 DTES 的必要性。组建工作组并邀请共识专家，在检索核心结局指标集 (core outcome sets, COS) 以遴选重要结局指标、基于系统评价 (systematic review, SR) 构建情景示例、设计调查 DTES 的专家函询问卷后，通过一轮德尔菲专家函询和一轮面对面专家共识会议，针对目标人群分组确定了 COVID-19 重要结局指标的 T1、T2 与 T3，该 DTES 已在“药物治疗新型冠状病毒感染快速动态循证要览”^[4]中报告。

在应用方法规程初稿的同时，细化步骤的操作要点，及时发现并解决使用该方法时可能出现的问题。基于试点应用 DTES 工作组成员的讨论，采纳有关专业术语解释、专家函询前培训及情景示例证据遴选的主要修改意见。

1.3 形成方法指南终稿

采用目的抽样，遵循信息饱和原则^[5]，邀请资深卫生研究方法学专家，就如何优化方法规程进行半结构化深入访谈。根据扎根理论对访谈材料进

行编码分析，通过开放式编码、主轴编码与核心编码 3 个步骤概括整理专家的修改意见得到主范畴，并进一步完善各步骤形成修改稿。综合考虑专业领域、地域分布及性别等因素组建共识专家组，使用德尔菲法达成共识，使用专家积极系数与权威系数 (Cr) 反映专家可靠性和权威性，最终根据共识建议修改框架，经专家审议拟定终稿。

2 结果

2.1 构建方法指南结果

经检索后共获得 7 项研究，根据研究目的，3 项研究^[6-8]针对具体目标人群与结局指标展开 DTES 制定，4 项研究^[9-12]涉及 DTES 制定的思路与需要考虑的因素。以 SR 结果为基础，起草方法规程初稿。邀请 5 名工作年限在 5~15 年、博士学历且具有副高级及以上职称的方法学家组成核心小组，通过多轮讨论、修改形成初稿。

同时，基于“药物治疗新型冠状病毒感染快速动态循证要览”^[4]中试点应用工作组的讨论结果，采纳了 3 条有关专业术语解释、专家函询前培训及情景示例证据遴选的主要修改意见。邀请 6 名资深卫生研究方法学专家进行半结构化深入访谈，通过三级编码概括得到共识专家组组建、DTES 调查表设计、调查形式选择、资料整理分析、专家共识会议召开、方法指南整体建议 6 个主范畴，进一步完善形成修改稿。

综合考虑专业领域、地域分布及性别等因素，邀请来自北京、广东、甘肃、贵州、四川、湖北、辽宁 7 个省 (市) 数十家科研院所的 14 名专家组建共识专家组。专家的专业方向覆盖临床医学、临床药理学、公共卫生、卫生政策与卫生研究方法学。工作组向专家发放函询问卷，经过第一轮德尔菲调查后，专家积极系数 (问卷回收率) 为 100%，Cr 为 0.874，共识条目同意率在 78.57%~100% 之间，收到 30 条开放式专家意见。本轮所有共识条目同意率均大于 70%，无需进行第二轮德尔菲调查，继而修改步骤及操作要点，对操作要点进一步解释形成方法指南终稿并发给专家核查确认，经专家审议最终拟定 8 个步骤的方法指南：① 评估 DTES 制定的必要性；② 组建 DTES 工作组；③ 遴选重要结局指标；④ 基于最佳证据构建情景示例；⑤ 设计 DTES 专家函询问卷；⑥ 进行专家函询并分析结果；⑦ 面对面专家共识会议；⑧ 推广与后效评价。

2.2 健康结局指标 DTES 制定方法指南

健康结局指标 DTES 制定方法指南包含 8 个步

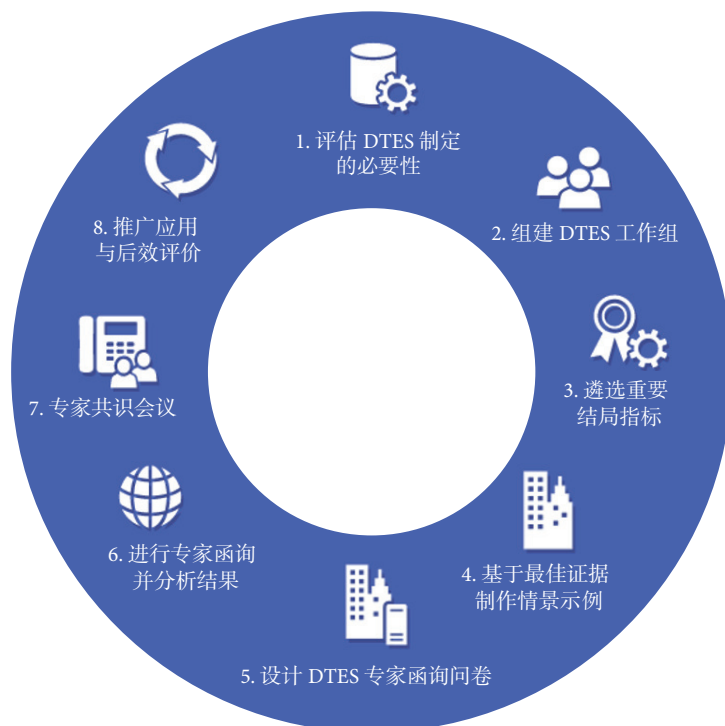


图1 DTES 制定步骤

骤(图1),各步骤操作要点及举例说明如下:

2.2.1 步骤1: 评估 DTES 制定的必要性 在后续步骤开始前,应先判断是否需要开展结局指标 DTES 的制定工作。首先系统检索目标主题是否已有研究报告相关结局指标 DTES,若不存在此类研究,可按本指南开展 DTES 的制定工作。

当存在已发布的 DTES 时,首先应评估当前发布的 DTES 是否符合目标主题要求,通常需判断其使用范围如人群分组、临床分型等是否与本研究目标一致。其次应评估 T1、T2 和 T3 是否均已制定,且已有 DTES 的结局指标是否已涵盖本研究关注的全部结局。需注意,T1 与最小临床意义差值(minimal clinically important difference, MCID)的含义一致^[10,13]。如果经评估认为已发布的 DTES 适用范围符合要求,T1、T2 和 T3 均已制定,且可使用已发布的 DTES。当已有的 DTES 不齐全、部分结局指标还需要确定 DTES 或适用范围的差异不可接受时,仍可遵循本指南开展制定工作。

2.2.2 步骤2: 组建 DTES 工作组 DTES 工作组应至少包含指导委员会、秘书组、证据评价组和共识专家组4个部分,成员的相关要求与主要职能见表1。成员可根据目标主题拟解决的问题(临床或公共卫生问题)进行调整,一般情况下应包括相关领域专家及卫生主管部门或卫生保健服务提供者等利益相关者,理想情况下还应包含患者、高危人群与公众^[7,14,15]。其中,共识专家组成员应以具备目

标主题相关研究经验,并对卫生研究方法学有一定了解的临床专家或公共卫生专家为主。此外,研究者还应邀请熟悉关注目标干预措施有效性相关证据的卫生研究方法学家参与共识。在条件允许的情况下建议访谈公众或患者代表等目标受众,邀请具备足以理解访谈内容的文化水平的利益相关者,调查其对结局指标重要性的评价及干预措施有效性的体验。由于公众或患者代表的观点基于个体而非群体视角,因此并不直接参与 DTES 的制定,可将公众或患者代表的观点作为专家讨论时的参考资料呈现。

在确定工作组成员后,所有成员需填写利益冲突声明表并交由秘书组备案,各类利益冲突的声明应在制定过程中基于实际情况动态更新。由指导委员会负责审核成员的利益冲突声明,存在直接或间接经济利益冲突的成员应予以排除,若存在其他类型利益冲突如专业或知识产权冲突,应予以公开。

2.2.3 步骤3: 遴选患者重要结局指标 研究者应基于专家、决策实施者与目标受众的意见^[16],采用文献研究、定性访谈、面对面专家共识会议、德尔菲法中的一种或多种方法遴选需要制定 DTES 的患者重要结局指标^[3,17]。

首先系统检索并筛选关注疾病的 COS^[18],遴选流程如图2所示。如果存在 COS,可以使用 COMET 工作组提出的 COS-STAD^[19] 进行方法学质量评价,以判断该 COS 是否可作为参考依据。COS-STAD

表 1 DTES 制定工作组的要求与职责

组别	专业/领域	主要职能
指导委员会	资深专业领域专家和方法学家	监督DTES制定过程, 提供咨询和指导, 管理利益冲突
秘书组	学会、协会或承担单位工作人员	调研结局指标, 组织专家参与DTES函询与共识会议, 记录制定过程, 撰写并发布、传播
证据评价组	具备循证医学知识及能力的专业人员	已发布DTES、核心结局指标集、情景示例证据的检索与评价, 制作调查表中各个结局指标的情景示例或目标干预措施的负担危害报告
共识专家组	临床或公共卫生专家、方法学家	遴选重要结局指标, 填写DTES专家函询问卷、投票和共识

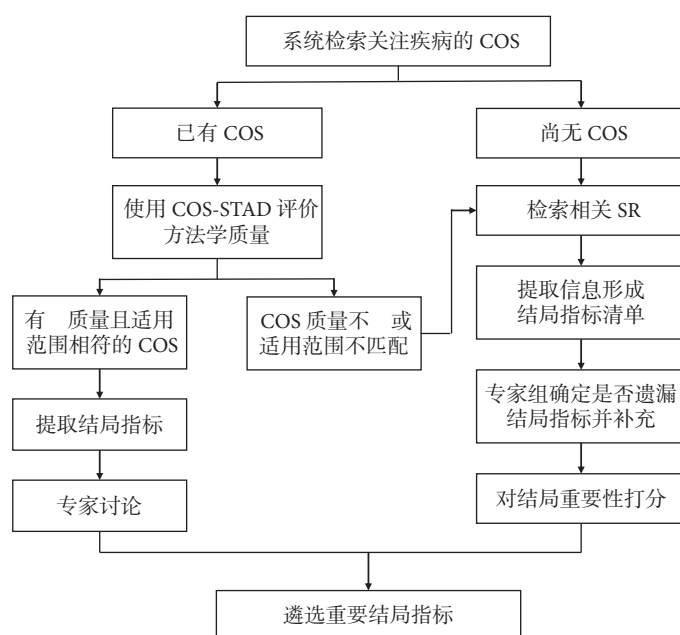


图 2 重要结局指标遴选流程

包含 11 个条目, 各条目最终评价结果为“是”(1 分)或“不是”(0 分)。针对高质量且符合研究范围的 COS, 提取 COS 的结局指标, 通过专家组对重要性打分排序并形成共识以确定重要结局指标。若无高质量 COS, 或 COS 适用范围不足以涵盖 DTES 应用范围(例如 COS 仅关注目标疾病的某一分型, 或 COS 仅关注中老年人而 DTES 目标人群还包括青少年与成人等), 则参考 WHO 指南制定手册与 GRADE 工作组建议的结局指标遴选与重要性排序方法^[16,17,20-22], 基于人群、干预措施与对照措施系统检索并筛选已发表的 SR, 提取多个 SR 所关注的结局指标以形成结局指标清单。由专家组确定是否有遗漏的相关结局并进行补充, 可要求专家组成员使用 9 分 Likert 评分系统对结局指标的重要性进行打分, 7~9 分代表“至关重要”, 4~6 分代表“重要”, 1~3 分代表“相对次要”。基于评分情况对结局指标进行重要性排序, 遴选“至关重要”与“重要”的结局指标。

2.2.4 步骤 4: 基于最佳证据制作情景示例 为便于专家理解调查目的, 直观呈现拟关注结局接受干

预措施后与基线风险的差异, 在 DTES 调查表中应针对各个结局构建相应的情景示例。基于确定的研究对象、干预措施、对照措施与结局指标制定检索策略与证据的纳入与排除标准。尽管基于随机对照试验的 SR 或 Meta 分析(systematic review/meta-analysis, SR/MA)被视为最佳证据主要来源^[23], 但最佳证据需同时考虑 SR/MA 质量和时效性两个因素, 若存在已发表的 SR/MA, 可使用 AMSTAR 2 工具(a measurement tool to assess systematic reviews)^[24,25]对其进行方法学质量评价。AMSTAR 2 工具包含 16 个条目, 关注 SR/MA 从选题设计、注册、数据提取、偏倚风险评估、统计分析、结果的解释与讨论到利益冲突报告的全过程, 7 个条目被视为关键条目。各条目的评价结果为“是”(回答正确且依据充分)、“部分是”(回答正确但依据不足)和“否”(无相关评价内容或评价不当), SR/MA 的整体方法学质量可被评为“高”(无或仅 1 个非关键条目存在缺陷)、“中等”(≥1 个非关键条目存在缺陷)、“低”(1 个关键条目存在缺陷)以及“极低”(≥1 个关键条目存

表 2 分类变量的情景示例

结局指标：机械通气率 人群或患者：COVID-19感染者 对照措施：常规支持治疗					
样本量 (研究数)	相对效应	绝对效应			GRADE证据质量
		基线风险	干预组风险	差值	
3 981 (7 RCTs)	RR=0.81	116/1 000	94/1 000	-22/1 000	低

表 3 连续型变量的情景示例

结局指标：机械通气时长 人群或患者：COVID-19感染者 对照措施：常规支持治疗				
样本量 (研究数)	基线值	干预组效应	差值	GRADE证据质量
201 (3 RCTs)	14.7天	13.5天	MD=-1.2天	低

在缺陷)。

参考 Cochrane 系统评价手册^[26]，当存在近两年内发布的高质量 SR/MA 时，可将该 SR/MA 作为情景示例的证据基础；若 SR/MA 为高质量但发布时间较早，应考虑更新该 SR/MA；如果不存在相关 SR/MA 或存在低质量 SR/MA，则检索原始研究证据制作新的 SR/MA。当存在多篇符合要求的 SR/MA 时，由指导委员会综合考虑各研究中临床分型、人群及地区等亚组分析的详细程度确定最终的证据基础。

为避免干预措施的具体名称对专家判断造成影响，情景示例中不报告具体的干预措施。由于结局指标 DTES 为效应点估计值的阈值，为避免呈现信息造成判断干扰，不显示该估计值的置信区间。情景示例的呈现方式参考 SoF 表^[27]，应报告结局指标、人群或患者、对照措施、证据的样本量和研究数量、GRADE 证据质量。除上述信息外，对于分类变量，呈现相对效应、常规对照的基线风险、干预组风险、干预与基线风险差；对于连续型变量，呈现常规对照的基线值、干预组效应、干预与基线值的差值。参考 BMJ 发表的题为“Drug treatments for COVID-19: living systematic review and network meta-analysis”的研究^[28]，表 2 与表 3 为示例。

2.2.5 步骤 5：设计 DTES 专家函询问卷 设计 DTES 专家函询问卷，通过德尔非法调查 DTES。问卷由前言、填写说明和 DTES 调查表及专家基本信息 4 个部分组成。在前言部分详细说明研究目的、意义和调查内容。填写说明应对 DTES 调查表中出现的专业术语均给予详尽的解释。

针对每个重要结局指标，DTES 调查表根据目标人群分组或疾病临床分型构建一个或多个情景示例，每个情景示例下包含两个问题(框 1)。若存

在可用不同量表测量效应的结局指标，应针对 COS 指定的测量工具或 SR/MA 报告较为广泛使用的量表构建情景示例；若暂无相关 COS 或无法确定最为常用的量表，建议考虑每个量表均构建相应情景示例。为便于专家理解问题，贴合决策实际，明确要求专家基于群体效应视角对于干预组绝对风险或效应做出判断。

2.2.6 步骤 6：进行专家函询并分析结果 考虑到 DTES 专家函询问卷的专业性较强，为使专家能够充分理解调查内容，可在问卷发放前召开会议或制作教学视频对专家进行培训。

预先针对专家意见的集中程度与协调程度设定达成共识的标准，未达成共识的 DTES 应在反馈分析结果后再次共识。核对收集到的资料，进行双人双录。需注意问题 2 所得各组数据为干预组效应阈值，而研究目的为制定重要结局指标干预与基线差值 DTES。因此，在计算 DTES 前应将收集数据转化为干预与基线差值的阈值，即将基线风险或基线值减去专家所填写较小、中等或较大改变的数值。对转换后各组数值进行正态性检验，服从正态分布采用均数和标准差描述，否则采用中位数和四分位间距描述，所得均数或中位数即为各结局指标的 DTES。

2.2.7 步骤 7：面对面专家共识会议 当专家函询的分析结果存在未达成共识的结局指标时，向参与函询专家发出会议邀请，在会议中介绍会议议题并分别呈现达成共识及未达成共识的调查结果。专家在会议中针对未达成共识的 DTES 展开讨论，通过投票等方式确定该结局指标最终的 DTES。无论是否召开面对面专家共识，所有结局指标 DTES 都应由专家组最终确认通过。

2.2.8 步骤 8：推广应用与后效评价 DTES 可通

框 1 DTES 调查表示例

样本量 (研究数)	相对效应	绝对效应			GRADE 证据质量
		基线风险 ^b	干预组风险 ^c	差值 ^d	
6 700 (10 RCTs)	RR=0.48	85/1 000	41/1 000	-44/1 000	中等

RCTs: 随机对照试验; RR: 相对危险度。

请您根据实际情况填写相应答案。

问题 1: 对于亚组人群 P, 发生结局 O 的基线风险为 85/1 000, 干预后风险降至 41/1 000, 您认为这属于多大程度的改变? (C)

A: 无意义^e B: 较小的改变 C: 中等的改变 D: 较大的改变

问题 2: 对于亚组人群 P, 发生结局 O 的基线风险为 85/1 000, 当干预组风险为多少时, 分别对应较小、中等或较大的改变?^f

较小的改变

中等的改变

较大的改变

问题 1 为单选, 问题目的在于获取专家对具体情景示例中的干预组效应量大小的判断, 4 个选项分别为无意义、较小、中等与较大的改变。以表 2 和表 3 的情景为例, 专家应为干预组的绝对风险 (接受干预措施的人群或患者每 1 000 人中有 94 人有发生该结局的风险) 或干预组效应 (13.5 天) 选择效应量大小分类。问题 2 回答设置为滑块滑动条, 3 个选项分别调查专家认为当干预组绝对风险或干预组效应属于较小、中等或较大改变时的具体数值。坐标轴最小值设置为 0, 最大值不低于基线风险或基线值。a: 本次调查中所有数据来源于《药物治疗新型冠状病毒感染快速动态循证要览》证据基础纳入的研究; b: 基线事件发生风险指在没有积极治疗情况下每 1 000 人中有 85 人有发生结局 O 的风险。本次调查结局指标的基线风险数据主要来自世界卫生组织发布的 Therapeutics and COVID-19: living guideline; c: 干预事件发生风险指接受干预的目标人群 P, 每 1 000 人中有 41 人存在发生结局 O 的风险; d: 发生结局 O 的人群基线风险为 85/1 000 人, 接受干预措施之后, 风险降低至 41/1 000 人, 两者之间的差值为 -44/1 000 人, 即比较对照组, 干预措施可使每 1 000 人中多减少 44 人发生结局 O; e: 本案例中的“无意义”并不等同于无统计学意义, 而是指该干预措施对患者的临床结局并无实际能让患者感受到获益的改善。当点估计值大于最小临床意义差值时, 代表干预措施具有临床意义; f: 在判断阈值时, 应综合考虑结局指标对于人群或患者健康的重要性大小以及人群的风险等级。例如: 对于死亡率结局, 干预后 1 000 人中减少 10 人死亡的发生, 可视为有意义的。但对于腹泻等症状发生率结局, 干预后 1 000 人中减少 10 人腹泻的发生, 可能是无意义的。其次, 对于低风险人群 (轻型或中型感染者), 假设对照组的 7 天核酸转阴率为 163/1 000 人, 增加 50 人核酸转阴的发生, 或许可视为小到中等的效应。但对于高风险人群 (重型或危重型感染者), 假设对照组 7 天核酸转阴率为 200/1 000 人, 增加 50 人核酸转阴的发生可能为中等至较大的效应。

过同行评审期刊、媒体等途径传播, 可同时发布教学视频或培训手册等 DTES 的应用辅助材料。其主要应用方式包括辅助指南、SR 或卫生技术评估等研究背景下的数据分析结果解读, 评价效应估计值的不精确性, 以及作为推荐意见更新的先验阈值。

监测评估应用情况, 并建立合理的反馈机制。根据应用反馈与相关研究证据的动态监测情况 (例如定期检查 DTES 研究发表后的被引情况) 判断是否更新 DTES, 主要包括完善已制定的 DTES 和新增某个重要结局指标 DTES。

3 讨论

本研究在方法指南构建初期系统检索 DTES 制定相关研究, 在对此进行归纳总结的基础上撰写初稿, 同时充分参考权威机构制订且经过多个研究验证的相关方法工具与流程, 经方法学家多轮讨论一致通过初稿并进行试点应用; 为优化完善方法学框架, 根据试点应用多学科 DTES 工作组反馈、资深卫生研究方法学家访谈结果及一轮德尔菲函询

收集的多学科专家组正式意见最终拟定方法指南终稿, 构建了首个系统可操作的核心结局指标 DTES 制定方法指南。

该 DTES 制定方法指南的优点在于: ① 各步骤及操作要点清晰明确, 可操作性强, 解释举例部分列出具体方法流程的参考来源, 便于使用者理解; ② 明确的 DTES 可评估干预措施利弊、评价证据质量, 从而决定推荐意见方向与强度、判断推荐意见是否更新。基于国际 GRADE 标准^[29], 决定推荐强度的关键因素包括干预措施利弊平衡、证据质量、价值观和偏好以及成本大小。当确定结局指标的 DTES 后, 可明确干预措施的效应量大小与利弊方向, 评价证据质量的不精确性, 若干预利弊间差距越大、证据质量越高, 越适合作出强推荐, 反之适合作出弱推荐。同样, 当新证据出现时, 可通过 DTES 评估是否更新推荐意见, 确保指南的时效性和准确性。

鉴于时间与成本的限制, 本指南尚存在以下不足: ① 试点应用的范围不够广泛, 未来还应在

SR、卫生技术评估和指南等研究背景下应用,进一步验证方法指南;②DTES制定领域研究的数量较少,潜在的定量方法均未得到研究的应用与验证^[30,31],如考虑单一获益结局与单一危害结局的方法包括计算增量净健康效益(incremental net health benefit, INHB)、净临床效益(net clinical benefit, NCB)和需治疗人数(number needed to treat, NNT)等,同时考虑多个获益或危害结局的方法主要有多准则决策分析(multicriteria decision analysis, MCDA),限制了DTES制定方法的多样性;③临床决策注重患者的偏好与价值观,主要通过调查或系统检索目标受众的偏好与价值观研究收集观点,但将调查或SR结果融入决策形成过程的具体方法步骤少有报告^[3,32],考虑到DTES专家函询问卷的专业性较强,邀请目标受众直接参与DTES制定难度较大,方法指南提出通过访谈或问卷调查的方式收集受众对干预措施及结局指标效应量大小的观点,未来研究还需对科学系统方法展开进一步探索。

随着指南、SR和卫生技术评估等研究对制定并应用DTES以辅助解释数据分析结果、评价证据质量、制订推荐意见或形成结论的需求增加,今后有必要采取具体措施以推动框架的推广、应用与完善。其次,有必要针对DTES的有效性评价展开研究,以降低决策带来的风险,提高决策的可信度与稳定性。

4 小结

健康结局指标DTES制定方法指南可为临床决策的利益相关者衡量干预措施效应量大小提供可靠方法,促进研究者合理解读多种干预措施产生的健康影响、判断影响大小并评估证据精确性,对决策过程的科学化、规范化与透明化有着积极作用。

专家组成员(按姓氏首字母排序)

陈耀龙(兰州大学基础医学院循证医学中心)、邓宏勇(上海中医药大学中医健康协同创新中心)、杜亮(四川大学华西医院中国循证医学中心)、靳英辉(武汉大学中南医院循证与转化医学中心)、孔德昭(辽宁中医药大学附属医院)、黎国威(暨南大学附属广东省第二人民医院)、李江(中国医学科学院肿瘤医院)、李玲(四川大学华西医院中国循证医学中心)、廖星(中国中医科学院临床与基础医学研究所)、刘杰(中国中医科学院广安门医院)、商洪才(北京中医药大学东直门医院)、史楠楠(中国中医科学院中医临床基础医学研究所)、田金徽(兰州大学基础医学院循证医学中心)、王晓辉(兰州大学公共卫生学院)、吴大嵘(广

州中医药大学第二附属医院)、杨丰文(天津中医药大学循证医学中心)、杨克虎(兰州大学基础医学院循证医学中心)、于长禾(北京中医药大学东直门医院)、张家兴(贵州省人民医院药剂科)、赵晨(中国中医科学院中医临床基础医学研究所)

参考文献

- 1 Alonso-Coello P, Schünemann HJ, Moher J, *et al.* GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ*, 2016, 353: i2016.
- 2 国家卫生健康委员会. 药品临床综合评价管理指南. 2021.
- 3 Kantorová L, Friessová T, Slezáková S, *et al.* Addressing challenges when applying GRADE to public health guidelines: a scoping review protocol and pilot analysis. *Int J Environ Res Public Health*, 2022, 19(2): 992.
- 4 孙铭谣, 田晨, 李颖, 等. 药物治疗新型冠状病毒感染快速动态循证要览. *中国循证医学杂志*, 2023, 23(7): 745-754.
- 5 孙晓娥. 扎根理论在深度访谈研究中的实例探析. *西安交通大学学报(社会科学版)*, 2011, 31: 87-92.
- 6 Neumann I, Quiñelen E, Nahuelhual P, *et al.* Using Explicit Thresholds were valuable for judging benefits and harms in partially contextualized GRADE Guidelines. *J Clin Epidemiol*, 2022, 147: 69-75.
- 7 Helsing LM, Zeng L, Siemieniuk RA, *et al.* Establishing thresholds for important benefits considering the harms of screening interventions. *BMJ Open*, 2020, 10(12): e037854.
- 8 Morgano GP, Mbuagbaw L, Santesso N, *et al.* Defining decision thresholds for judgments on health benefits and harms using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Evidence to Decision (EtD) frameworks: a protocol for a randomised methodological study (GRADE-THRESHOLD). *BMJ Open*, 2022, 12(3): e053246.
- 9 Hultcrantz M, Rind D, Akl EA, *et al.* The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol*, 2017, 87: 4-13.
- 10 Zeng L, Brignardello-Petersen R, Guyatt G. When applying GRADE, how do we decide the target of certainty of evidence rating. *Evid Based Ment Health*, 2021, 24(3): 121-123.
- 11 Zeng L, Brignardello-Petersen R, Hultcrantz M, *et al.* GRADE guidelines 32: GRADE offers guidance on choosing targets of GRADE certainty of evidence ratings. *J Clin Epidemiol*, 2021, 137: 163-175.
- 12 Schünemann HJ, Neumann I, Hultcrantz M, *et al.* GRADE guidance 35: update on rating imprecision for assessing contextualized certainty of evidence and making decisions. *J Clin Epidemiol*, 2022, 150: 225-242.
- 13 Guyatt GH, Osoba D, Wu AW, *et al.* Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*, 2002, 77(4): 371-383.
- 14 Qaseem A, Forland F, Macbeth F, *et al.* Guidelines International Network: toward international standards for clinical practice guidelines. *Ann Intern Med*, 2012, 156(7): 525-531.
- 15 ECBC. Evidence-based methodologies for public health. 2011.
- 16 WHO. WHO handbook for guideline development. 2012.
- 17 Schünemann HJ, Wiercioch W, Etzandia I, *et al.* Guidelines 2.0:



- systematic development of a comprehensive checklist for a successful guideline enterprise. *CMAJ*, 2014, 186(3): E123-E142.
- 18 Williamson PR, Altman DG, Bagley H, *et al*. The COMET handbook: version 1.0. *Trials*, 2017, 18(Suppl 3): 280.
- 19 Kirkham JJ, Davis K, Altman DG, *et al*. Core outcome set-standards for development: the COS-STAD recommendations. *PLoS Med*, 2017, 14(11): e1002447.
- 20 WHO. WHO handbook of guideline development (2nd edition). 2014.
- 21 丁泓帆, 杨楠, 邓围, 等. WHO 指南制定的基本原则和方法. *中国循证医学杂志*, 2016, 16(4): 471-477.
- 22 Schünemann H, Brożek J, Guyatt G, *et al*. GRADE handbook for selecting and rating the importance of outcomes (Updated October 2013). 2013.
- 23 Schünemann H, Brożek J, Guyatt G, *et al*. GRADE handbook for summarizing the evidence (Updated October 2013). 2013.
- 24 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, 358: j4008.
- 25 陶欢, 杨乐天, 平安, 等. 随机或非随机防治性研究系统评价的质量评价工具 AMSTAR 2 解读. *中国循证医学杂志*, 2018, 18(1): 101-108.
- 26 Higgins JPT, Thomas J, Chandler J, *et al*. Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022). 2022.
- 27 Guyatt G, Oxman AD, Akl EA, *et al*. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*, 2011, 64(4): 383-394.
- 28 Siemieniuk RA, Bartoszko JJ, Zeraatkar D, *et al*. Drug treatments for COVID-19: living systematic review and network meta-analysis. *BMJ*, 2020, 370: m2980.
- 29 Guyatt GH, Oxman AD, Kunz R, *et al*. Going from evidence to recommendations. *BMJ*, 2008, 336(7652): 1049-1051.
- 30 Puhan MA, Singh S, Weiss CO, *et al*. A framework for organizing and selecting quantitative approaches for benefit-harm assessment. *BMC Med Res Methodol*, 2012, 12: 173.
- 31 Yu T, Fain K, Boyd CM, *et al*. Benefits and harms of roflumilast in moderate to severe COPD. *Thorax*, 2014, 69(7): 616-622.
- 32 Santesso N, Schünemann H, Blumenthal P, *et al*. World Health Organization guidelines: use of cryotherapy for cervical intraepithelial neoplasia. *Int J Gynaecol Obstet*, 2012, 118(2): 97-102.

收稿日期: 2023-08-24 修回日期: 2023-12-21
本文编辑: 张洋