◁教学与科研管理▷

研究者发起的临床研究分级分类评价指标体系的构建*

高巾媛，李梦诗，尤爱军，朱琪，薛芊△
海军军医大学第一附属医院/长海医院临床研究中心，上海 200433

【摘要】 **目的**　构建研究者发起的临床研究分级分类评价指标体系，划分项目等级以实现分级分类管理，为临床研究科学高效管理提供参考，促进研究成果产出与学科发展。**方法**　梳理相关文献、政策文件初步构建指标体系，经过两轮专家咨询筛选优化，采用SPSS 26.0与SAS 9.4进行统计分析。**结果**　两轮专家咨询有效回收率分别为100%和95%，专家权威系数为0.92和0.94，两轮专家意见协调性显著提高，结果差异有统计学意义。最终形成包含7个一级指标、19个二级指标、22个三级指标的评价体系。**结论**　该指标体系可实现临床研究分级分类管理，提升管理效能，推动管理向重质量、重结果转型，提高研究水平与成果产出，助力医院高质量发展。

【关键词】 临床研究；德尔菲法；分级分类评价指标体系
【文献标志码】 A　【文章编号】 1672-4232(2026)03-0088-05
【DOI编码】 10.3969/j.issn.1672-4232.2026.03.020

Construction of a Classification and Evaluation Index System for Clinical Research Initiated by Researchers

GAO Jin-yuan, LI Meng-shi, YOU Ai-jun, ZHU Qi, XUE Qian
The First Affiliated Hospital of Naval Medical University, Changhai Clinical Resarch Unit, Shanghai 200433, China

【**Abstract**】 **Objective:** To construct a clinical research grading and classification evaluation index system initiated by researchers and classify project levels to achieve grading and classification management so as to provide the references for scientific and efficient management of clinical research, and promote research output and disciplinary development. **Methods:** A preliminary indicator system was constructed by reviewing the relevant literature and policy documents. After two rounds of expert consultation and optimization, statistical analysis was conducted using SPSS 26.0 and SAS 9.4. **Results:** The effective recovery rates of the two rounds of expert consultation were 100% and 95%, respectively, with expert authority coefficients of 0.92 and 0.94. The coordination of expert opinions in the two rounds was significantly improved, and the difference in results was statistically significant. The final evaluation system consists of 7 first-level indicators, 19 second-level indicators, and 22 third-level indicators. **Conclusions:** This indicator system can achieve the graded and classified management of clinical research, improve management efficiency, promote the transformation of management to focus on quality and results, improve research level and output, and assist in the high-quality development of hospitals.

【**Key words**】 clinical research; Delphi method; grading and classification evaluation index system

研究者发起的临床研究（investigator-initiated trial，IIT），是由医疗卫生机构开展的，以个体或群体为研究对象，其核心目的并非用于药品、医疗器械等产品注册，而是围绕疾病的病因、诊断、治疗、康复及健康维护等方面开展的科学研究活动[1]。本研究所指的临床研究特指研究者发起的临床研究。近年来，随着国家及各医疗机构对临床研究重视程度的不断提升，IIT数量呈现日益增多的趋势。与临床试验（industry-sponsored trial，IST）及基础科学研究相比，IIT的研究类型复杂多样，但目前缺乏成熟的实施规范与统一的评价体系，且不同项目的实施风险存在差异，加之项目内部监管力度不足，导致IIT整体质量参差不齐[2]，给IIT项目评价工作带来较大困难。

建立符合IIT项目特点的评价指标体系，并通过评价结果实现项目分级分类管理，不仅有利于降低项目实施风险，还能有效提升项目整体质量。近几年，国家出台了一系列关于完善科技成果评价机制的文件，强调要优化科技成果评价体系，加强科研项目分类评价[3]。临床研究作为科学研究领域中极为关键的组成部分，其评价体系也面临着诸多优化与调整需求[4]。

目前，IIT项目评价过程中，大多沿用与普通科研项目类似的评价方式，主要通过同行评议开展。这种评价方式主观性较强，且评价过程中往往过度侧重发表论文数量、论文影响因子高低及所获各类奖项等量化成果，未充分考虑IIT项目类型复杂、研

*基金项目：上海申康医院发展中心研究型医师创新转化能力培训项目（SHDC2022CRS045）
△通信作者

究内容具有创新性、研究周期较长、质量差异显著等固有特点，难以满足IIT项目的个性化评价需求。

本研究从临床研究的核心内涵出发，结合IIT项目的自身特点，构建临床研究项目分级分类评价指标体系，对IIT项目进行全面综合评价；根据评价结果划分项目等级，对等级较低的项目给予有针对性的支持与指导，规范其后续研究开展，提升研究质量；对等级较高的项目，在后续高级别项目推荐、奖项申报、荣誉评定等工作中给予相应的倾斜和政策支持[3]。通过实施分级分类评价，进一步规范临床研究项目的实施流程，推动临床研究项目管理向重质量、重实效、重结果的方向转变，激发研究人员的创新活力，助力高质量临床研究开展及高水平研究成果产出，全面提升医院的临床研究能力，同时为临床研究管理部门开展管理工作、制定相关评价制度提供科学依据。

1　研究资料与方法

1.1　资料来源

本研究通过查阅临床研究相关法律法规及文献资料[4-6]，结合临床研究管理工作经验及已有的临床研究质量评价指标体系[7]，总结临床研究分级分类评价要点，提取临床研究评价指标，形成临床研究分级分类评价指标体系初步框架，包括7个一级指标、23个二级指标和14个三级指标（见表1）。

1.2　专家咨询法

本研究采用专家咨询法（Delphi法），研究人员通过问卷形式征询专家意见，随后根据回收的问卷内容对指标体系进行修改后进行下一轮咨询，直至专家意见趋于一致[8]。本研究采用两轮Delphi法，专家遴选以具有高级职称、从事临床相关工作10年以上且在临床研究领域具备丰富经验为核心标准，选取临床研究发展较迅速的上海地区，筛选出临床医学、临床研究管理、流行病学及统计学领域的20名自愿参与本研究的专家。研究通过微信平台发放专家咨询问卷，问卷主要包含专家咨询表及专家基本情况调查表。专家收到问卷后，对各指标重要程度进行赋值（0～5分），并对存在异议的指标提出具体修改意见或建议。第一轮专家咨询结束后，研究人员结合专家提出的修改意见，同时参考统计分析得出的指标综合评分、变异系数等数据，对相关指标进行删减优化，进一步完善指标体系后，进入第二轮专家咨询[9-10]。

1.3　统计学方法

专家咨询回收后，通过SPSS 26.0和SAS 9.4软件对专家积极系数、专家权威系数以及肯德尔协调系数进行计算，以$P<0.05$为有统计学意义；计算各指标均数、标准差与变异系数进行指标筛选与排序。指标权重通过加权平均法计算所得，将专家i对指标j的重要性评分记为S_{ij}（i=1，2，3…；j=1，2，3…）；将专家i的权威系数记为Cr_i，第j个指标权重记为w_j计算公式如下[8]：

$$w_j=\frac{\sum_{i=1}^{n}\left(S_{ij}\times Cr_i\right)}{\sum_{i=1}^{n}Cr_i}$$

将w_j进行变量标准化处理，与所属一级指标的权重相乘，计算组合权重[8]。

2　结果

2.1　专家基本情况

本研究共邀请20名咨询专家，其平均年龄为（47.05±5.88）岁；职称构成方面，高级职称19人、中级职称1人；学历构成方面，博士研究生17人、硕士研究生3人；专业领域涵盖临床医学、临床研究管理、统计学等多个方向。所有专家均从事临床相关工作10年以上，且均具备参与或主持高质量临床研究项目的经验。研究过程中，对中途退出的专家予以排除。

表1　研究者发起的临床研究分级分类评价指标体系初步框架

一级指标	二级指标	三级指标
重要性	科学意义、临床意义	
创新性	理论创新、方法论创新、临床实践创新 、技术创新、产品创新	
科学性	研究目标、研究方案	研究基础、研究设计、研究类型、质量控制、统计学方法、风险评估与处理、研究参与者权益与保护
规范性	药物/器械临床样本管理、临床研究注册和登记备案、人类遗传资源管理、合同协议管理、临床研究保险	
可行性	受资助情况、研究实施难度、研究团队及资质投入情况	主要研究者、研究参与者、研究辅助人员、研究团队、经费投入、人员投入、时间投入
风险程度	研究本身风险、研究团队风险、研究参与者风险、利益冲突说明	
成果产出	预期成果产出情况	

2.2　专家积极系数

两轮专家咨询过程中，问卷有效回收率分别达到 100% 和 95%，表明专家参与本研究的积极程度较高，研究响应性良好。

2.3　专家权威系数

专家权威系数（Cr）体现专家判断的科学性与可靠性，其计算方法为专家对咨询内容的熟悉程度（Ca）与判断依据（Cs）的算术平均值。通常认为，专家权威系数≥0.70 即可满足研究要求[10]。本研究两轮咨询的专家权威系数分别为 0.92 与 0.94，均>0.90，说明专家权威程度较高，其评价意见具有较强的科学性和可信度。

2.4　专家意见协调系数

专家协调系数反映专家对各指标评价意见的分歧程度，取值范围为 0~1，数值越大表明专家意见的一致性越高[7]。本研究两轮专家咨询中，第二轮专家意见协调系数整体高于第一轮，且两轮指标咨询结果差异均有统计学意义（P<0.05），表明咨询结果具备一定程度的可靠性，咨询结果可取。

2.5　指标筛选

结合两轮专家反馈意见及统计分析结果，对指标体系进行修改完善。

一级指标：20 位专家意见基本一致，无需调整。

二级与三级指标：（1）一级指标“创新性”下，新增“创新性类别”“创新性等级”2 个二级指标；其中“创新性类别”包含理论创新、方法论创新、临床实践创新与技术创新 4 个三级指标，“创新性等级”包含一般创新、重要创新与颠覆性创新 3 个三级指标，同时删除三级指标“产品创新”。（2）一级指标“科学性”下，将原二级指标“研究基础”“研究设计”“质量控制”调整为三级指标，并归属于新增二级指标“研究方案”下；将原二级指标“风险评估与处理”与“研究参与者权益与保护”合并为“受试者权益与保护”，并将其归属于一级指标“规范性”下，同时删除原二级指标“研究类型”及“统计学方法”。（3）一级指标“可行性”下，删除“受资助情况”“研究参与者”“研究辅助人员”3 个二级指标，新增二级指标“既往研究实施情况”，并将原二级指标“研究实施难度”简化为“实施难度”。（4）一级指标“风险程度”下，删除“研究参与者风险”与“利益冲突说明”2 个二级指标。（5）一级指标“成果产出”下，将原二级指标“预期成果产出”修改为“产出类别”，并新增“有望改写指南/共识”“有望发表高影响力论著”“临床新技术/新策略”3 个三级指标，最终通过指标结构调整，实现指标层级的细化与优化，提高了评价体系的可操作性。

2.6　指标确立

综合上述指标筛选结果，最终确定临床研究分级分类评价指标体系（见表 2）。该体系包含 7 个一级指标、19 个二级指标与 22 个三级指标。一级指标相对权重从大到小依次为：科学性（19.47%）、创新性（17.79%）、重要性（17.26%）、可行性（13.45%）、规范性（13.42%）、成果产出（10.13%）、风险程度（8.47%）。其中，三级指标中组合权重较高的依次为研究设计（3.51%）、质量控制（3.24%）、研究基础（3.02%）及颠覆性创新（3.01%），这一结果提示，科学性与创新性在临床研究分级分类评价中占据核心重要地位，是评价临床研究质量的关键维度。

3　讨论

3.1　评价指标体系的科学性与导向性

本研究在前期开展充分文献调研的基础上启动，运用 Delphi 法构建临床研究分级分类评价指标体系。在方法学的科学性、有效性和合理性方面，Delphi 法已经成熟并获得了学术界的广泛认可，同时在临床研究和相关领域的实际应用中也被广泛采用[11]。本研究以两轮专家咨询所获取的数据为基础，运用 Delphi 法进行数据处理，并结合专家意见开展指标筛选，最终确定了科学性、创新性、重要性、可行性、规范性、成果产出及风险程度 7 个一级指标。其中，科学性、创新性和重要性的权重显著高于可行性、规范性、成果产出及风险程度，提示研究设计的科学性、研究内容的创新性以及研究主题的重要性，是开展高质量临床研究的重要前提，也是临床研究评价工作中需重点考量的核心因素。而可行性、规范性及风险程度是高质量临床研究顺利开展的重要保障，也是临床研究实现高质量成果产出的必要条件。7 个一级指标下共包含 19 个二级指标与 22 个三级指标，较为全面地涵盖了临床研究全过程的评价要点，可为临床研究管理部门开展项目评价、制定相关评价制度提供科学依据，进而提升临床研究管理效能。

3.2　评价指标体系的可行性与实效性

本研究立足临床研究实际，基于临床研究全流程，构建临床研究分级分类评价指标体系并开展初步实践，具体举措如下：一是医院设立临床研究管理办公室，积极为致力于临床研究的医务人员开展临床研究相关培训，并提供“临床科研门诊”服务，

表2　研究者发起的临床研究分级分类评价指标体系及指标权重(%)

一级指标	二级指标	三级指标	组合权重	秩次
重要性(17.26)	科学意义(50.00)		8.67	3
	临床意义(50.00)		8.67	2
创新性(17.79)	创新性类别(52.00)	理论创新(25.00)	2.28	16
		方法论创新(24.00)	2.12	22
		临床实践创新(26.00)	2.22	19
		技术创新(25.00)	2.17	21
	创新性等级(48.00)	一般创新(31.00)	2.71	13
		重要创新(34.00)	2.95	12
		颠覆性创新(35.00)	3.01	11
科学性(19.47)	研究目标(50.00)		9.79	1
	研究方案(50.00)	研究基础(31.00)	3.02	10
		研究设计(36.00)	3.51	6
		质量控制(33.00)	3.24	8
规范性(13.42)	受试者权益与保护(18.00)		2.49	14
	药物/器械/临床样本管理(16.00)		2.18	20
	临床研究注册和登记备案(17.00)		2.32	15
	人类遗传资源管理(17.00)		2.27	17
	合同/协议管理(17.00)		2.24	18
	临床研究保险(15.00)		1.98	23
可行性(13.45)	既往研究实施情况(23.00)		3.04	9
	实施难度(25.00)		3.35	7
	研究团队及资质(27.00)	主要研究者(51.00)	1.85	24
		研究团队(49.00)	1.72	25
	投入情况(26.00)	经费投入(34.00)	1.20	33
		人员投入(0.33)	1.17	34
		时间投入(0.33)	1.15	35
风险程度(8.47)	研究本身风险(0.52)		4.42	4
	研究团队风险(0.48)		4.09	5
成果产出(10.13)	产出类别(1.00)	有望改写指南/共识(0.16)	1.59	26
		有望发表高影响力论著(0.15)	1.52	27
		临床新技术/新策略(0.15)	1.50	28
		有助于新器械/药物的研发(0.15)	1.49	29
		为解决临床问题提供新的思路(0.15)	1.48	30
		申请专利/软著(0.12)	1.25	32
		人才培养(0.13)	1.34	31

重点解决项目准备阶段的科学性论证与创新性提升问题。二是医院开展临床研究科学性审查工作，对临床研究的科学性、创新性、重要性、可行性及风险程度进行全面系统评价。三是医院搭建三级质控体系，根据项目风险程度对临床研究项目实施全流程质控，重点核查临床研究实施的规范性及成果产出的真实性、完整性。最后，通过各阶段综合评价结果对临床研究项目实施分级分类管理并有所侧重：等级较高的项目可入驻示范性研究型病房，实现高质量临床研究项目“一站式”高效实施；对等级较低的项目给予相应的支持与指导，规范其临床研究开展流程，提升研究质量。通过上述分级分类管理举措，医院的临床研究能力和水平得到显著提升，我院脑血管病中心牵头开展的两项大型RCT研究——PROTECT-MT研究、MAGIC-MT研究，于2024年同步在《柳叶刀》与《新英格兰医学杂志》两大国际顶级学术期刊上重磅发布[12-13]，充分彰显了医院临床研究分级分类管理的显著成效。

3.3　研究不足与展望

近年来，临床研究发展势头迅猛，本研究基于临床研究项目评价的实际需求，构建了临床研究分级分类评价指标体系，但受研究时间及研究设计的限制，本研究仍存在一些不足。在专家邀请环节，尽管邀请了20名咨询专家，其数量及专业领域均符合Delphi法的实施要求，但20位专家均来自上海地区，导致专家的区域代表性相对不足。在指标构建方面，本研究的评价指标主要通过文献检索及参考临床研究领域专家的相关经验确定，但受检索方法不够全面、不同领域专家对临床研究评价的侧重点存在差异等因素制约，可能影响临床研究评价要点的全面覆盖。目前，该评价指标体系仍处于初始应用阶段，尚未经过大样本验证与优化，导致其认可度有待提升，进而给推广应用工作带来较大困难。针对上述不足，在后续研究中，将重点推进以下工作：首先，扩大专家邀请范围，吸纳不同地区、不同层级医疗机构的专家参与，提升专家的区域代表性

与专业覆盖面；其次，进一步验证与优化该指标体系，争取在多家医疗机构推广应用，不断完善指标体系的科学性、可行性与普适性。

参　考　文　献

[1] 国家卫生健康委，国家中医药局，国家疾控局. 关于印发医疗卫生机构开展研究者发起的临床研究管理办法的通知[EB/OL]. (2024-09-18). [2025-07-03]. https://www.nhc.gov.cn/qjjys/c100016/202409/3a3ad0a7b656420d9580b65f2321a623.shtml.

[2] 朱琪，李梦诗，金晶，等. 新时期研究者发起的临床研究管理实践[J]. 中华医学科研管理杂志，2023，36(4)：308-311.

[3] 陈卓，岳文涛，阴赪宏. 建立基础、临床科研项目结题评价体系及其应用[J]. 中华医学科研管理杂志，2022，35(4)：286-292.

[4] 李海燕，王雪莹. 浅析临床研究评价发展趋势[J]. 中华医学科研管理杂志，2020，33(6)：461-464.

[5] 康玫，李宪辰，赫慧琛，等. 研究者发起的临床研究过程质量管理探讨[J]. 中国新药杂志，2020，29(12)：1387-1390.

[6] 曹烨，王欣，曹玉，等. 我国研究者发起的临床研究管理现况调查与分析[J]. 中国新药与临床杂志，2018，37(7)：395-400.

[7] 吕文文，张维拓，谢丽，等. 研究者发起的临床研究项目实施过程质量评估指标构建探讨[J]. 中国新药与临床杂志，2019，38(2)：85-89.

[8] 创维. 研究者发起的临床研究风险评估量表的研制及实证研究[D]. 广州：南方医科大学，2021.

[9] 陈武朝，黄二丹，文宇，等. 公立医院分院区同质化评价指标体系构建与应用[J]. 中华医院管理杂志，2023，39(1)：1-5.

[10] 秦佳鑫，马晓静，胡翔. 公立医院绩效考核政策认同评价指标体系构建研究[J]. 中国医院管理，2021，41(8)：12-16.

[11] 钱梦星，徐正梅，孙亚林，等. 研究型医科大学应用技术类科技成果评价指标体系构建[J]. 解放军医院管理杂志，2015，22(6)：584-586.

[12] Liu J, Zhou Y, Zhang L, et al. PROTECT-MT Investigators. Balloon guide catheters for endovascular thrombectomy in patients with acute ischaemic stroke due to large-vessel occlusion in China (PROTECT-MT): a multicentre, open-label, blinded-endpoint, randomised controlled trial [J]. Lancet, 2024, 404(10468): 2165-2174.

[13] Liu J, Ni W, Zuo Q, et al. MAGIC-MT Investigators. Middle Meningeal Artery Embolization for Nonacute Subdural Hematoma[J]. N Engl J Med, 2024, 391(20): 1901-1912.

收稿日期：2025-05-22
修回日期：2025-07-17　　（编辑　张瀚予）