·教学改革·

基于“评学融合”的《医学免疫学》考试评价改革创新与实践

庞维　冯辉　曹妍　郑丽　孙逊
中国医科大学基础医学院免疫学教研室，沈阳 110122
通信作者：冯辉，Email：hfeng@cmu.edu.cn

【摘要】 医学教育评价是人才培养的导向核心，其改革成效直接关系到立德树人根本任务的落实以及医学教育的高质量发展。本研究基于“评学融合”理念，对评价目标、过程实施、命题策略及反馈机制进行了系统重构，构建了《医学免疫学》课程“三维度四阶段”评价体系。实践表明，该体系具有目标导向明确、动态追踪及时、评价精准全面的特点，并能显著提升学生的综合能力。本研究不仅为医学教育评价改革搭建了可操作的框架，更通过体系化的实践路径，为培养具有新时代临床胜任力的医学人才提供了范式参考。

【关键词】 医学教育；　考试评价；　评学融合；　医学免疫学；　形成性评价
【中图分类号】 R392；G642
基金项目：中华医学会教育分会和全国医学教育发展中心 2023 年医学教育研究课题（2023B333）；2025 年度中国医科大学本科教学改革研究项目（YDJG2025017）
DOI: 10.3760/cma.j.cn116021-20250715-02168

Reforming innovation and implementation of examination assessment in the course of Medical Immunology based on "assessment-learning integration"

Pang Wei, Feng Hui, Cao Yan, Zheng Li, Sun Xun
Department of Immunology, College of Basic Medical Sciences, China Medical University, Shenyang 110122, China
Corresponding author: Feng Hui, Email: hfeng@cmu.edu.cn

【Abstract】 Medical education assessment serves as a core guiding element for talent cultivation, and the effectiveness of its reform is directly related to the fulfillment of the fundamental task of fostering virtue and talent through education and the high-quality development of medical education. Based on the concept of "assessment-learning integration", this study systematically reconstructed the assessment objectives, implementation processes, proposition strategies, and feedback mechanisms to establish a "three-dimension four-phase" evaluation system for the Medical Immunology course. Empirical evidence has demonstrated that this system is characterized by clear goal orientation, timely dynamic monitoring, and precise and comprehensive evaluation, and it can significantly enhance students' comprehensive abilities. This work not only provides a practical framework for the reform of medical education assessment but also, through its systematic practical pathway, provides a paradigm for cultivating medical talents with clinical competence in the new era.

【Key words】 Medical education;　Examination assessment;　Assessment-learning integration;　Medical immunology;　Formative assessment
Fund program: Chinese Medical Association Education Branch and National Center for Health Professions Education Development 2023 Annual Medical Education Research Project (2023B333); 2025 China Medical University Undergraduate Education Reform Research Project (YDJG2025017)
DOI: 10.3760/cma.j.cn116021-20250715-02168

教育评价改革是医学教育创新的关键突破点。中共中央、国务院《深化新时代教育评价改革总体方案》明确指出要“改进结果评价，强化过程评价，探索增值评价，健全综合评价”[1]。作为质量保障核心环节，医学考试评价需兼顾知识测量与临床思维、职业素养等能力评估[2-3]。当前我国医学教育评价体系因发展时间短、经验不足，存在多方面问题，如评价标准界定不科学，过度侧重知识点记忆而忽视核心能力；内容覆盖范围局限于教材，缺少必要的前沿进展追踪及人文素养的考核；传统考前划重点-考后给分数模式导致反馈机制失效等。因此，亟须构建精准评价范式。

从全球来看，现代教育评价范式正经历从“为了学习的评价”（assessment for learning）向“作为学习的评价”（assessment as learning）不断探索和演进的过程[4]。这一“评学融合”理论强调评价应融入教学过程，其特征体现为：形成性与终结性评价结合、评价活动具有学习价值、构建教学—学习—评价动态闭环[5-6]。中国医科大学率先将“评学融合”这一理念付诸实践，构建了系统完善的评价改革框架：以“评学融合”理念为指导，以提升教师评价素养为基础，通过智能题库建设确保“测得准”，借助学习分析技术实现“用得好”，最终达成“学得透”的教育目标。该框架已深入开展了一系列教育教学评价改革[7]。

本研究基于“评学融合”理念，以《医学免疫学》课程为载体，通过系统重构评价目标、创新评价方法、优化评价内容，着力拓展具有中国特色的医学教育评价改革新路径，为培养德才兼备的卓越医学人才提供有力支撑。

1 “评学融合”理念在《医学免疫学》考试评价体系中的创新实践

《医学免疫学》作为连接基础与临床的关键桥梁学科，具有理论复杂、临床关联强且学科发展迅速的特点，这对课程评价体系提出了较高要求。在过往实践中，课题组发现当前评价体系存在一些亟待优化的方面。其一，评价目标与培养目标之间的衔接不够紧密，在一定程度上难以充分保障培养目标的有效达成；其二，评价方式较为侧重于静态结果，对于学生学习过程的动态监测以及基于此的教学反馈机制尚不完善；其三，考试命题设计存在一定局限性，较多关注记忆性知识的考查，而对学生综合能力的评价力度相对不足。鉴于此，本课程依托“评学融合”框架，对考试评价体系进行了系统重构。

1.1 “三位一体”目标体系的构建

针对评价目标与培养目标脱节的核心痛点，课题组通过分析临床医学人才岗位胜任力模型和本校临床医学专业人才培养目标，从知识、能力和素质 3 个维度，建立了《医学免疫学》课程“三位一体”的教学目标体系（图 1）。该体系以培养目标为起点，明确界定与之严格匹配的教学目标（教什么）和考核目标（考什么），确保三者环环相扣、一一对应。此举旨在彻底改变讲什么考什么的旧模式，使考核评价真正成为衡量培养目标达成度的有效工具，有力支撑医学人才培养质量提升。

1.2 “四阶段多层次”评学融合评价体系构建

本研究构建了贯穿教学全周期的动态评价体系，即四阶段多层次的评学融合体系（图 2）。该体系具有三大特征：一是阶段性，设置预习—课堂—复习—总结四阶段评价；二是动态性，通过课前—课中—课后评价数据随时进行教学动态分析及反馈；三是智能化，依托教学平台实现学习过程全记录。在评价结构上，过程性评价（占比 40%）与终结性评价（占比 60%）相配比。一课三评价模式，即通过开展课前预习、课中互动答题、课后试题小测，平台实时收集学生学习数据并进行大数据分析，动态且全面地掌握学生学习全过程。期末考试后，依托教学平台的成绩分析模块，对试卷信息进行深度挖掘，并通过反馈机制优化教学内容与方法，形成评价—反馈—改进的闭环评价模式，持续提升教学质量。这一体系不仅实现了对学生学习过程的全程跟踪与精准评价，也为教学质量的持续改进提供了有力支撑。

1.3 终结性考核试题命制及分析

1.3.1 基于双向细目表制定命题方案

在医学教育评价改革中，试题命制质量直接关系到人才培养的成效评估。为此，课题组建立了以双向细目表为核心的标准化命题流程。以教学大纲为基准，通过双向细目表实现：①内容维度精准覆盖：重点考核内容（掌握类占比 60%~70%），次重点（熟悉类占比 20%~30%），扩展内容（了解类占比<10%），确保核心知识点覆盖率达 90% 以上。②认知维度梯度设计：采用改良版 Bloom 分类法，设置回忆型（30%~40%）、理解（25%~35%）、应用（20%~30%）和分析（10%~15%）四级认知层次。

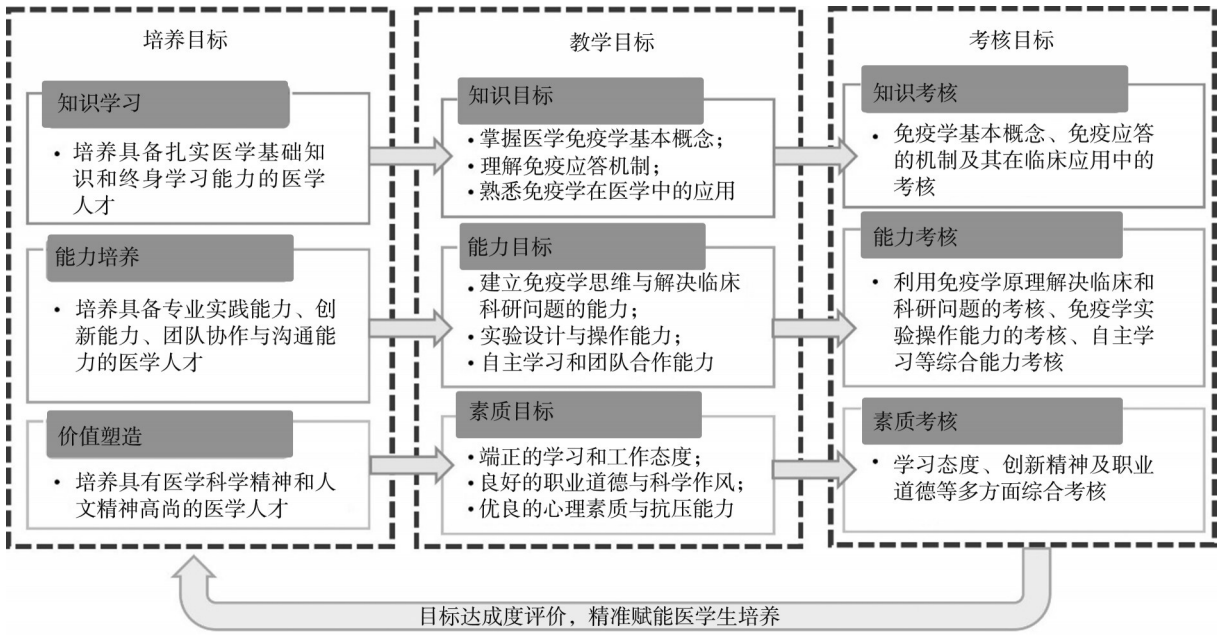


图 1　《医学免疫学》“三位一体”的目标体系

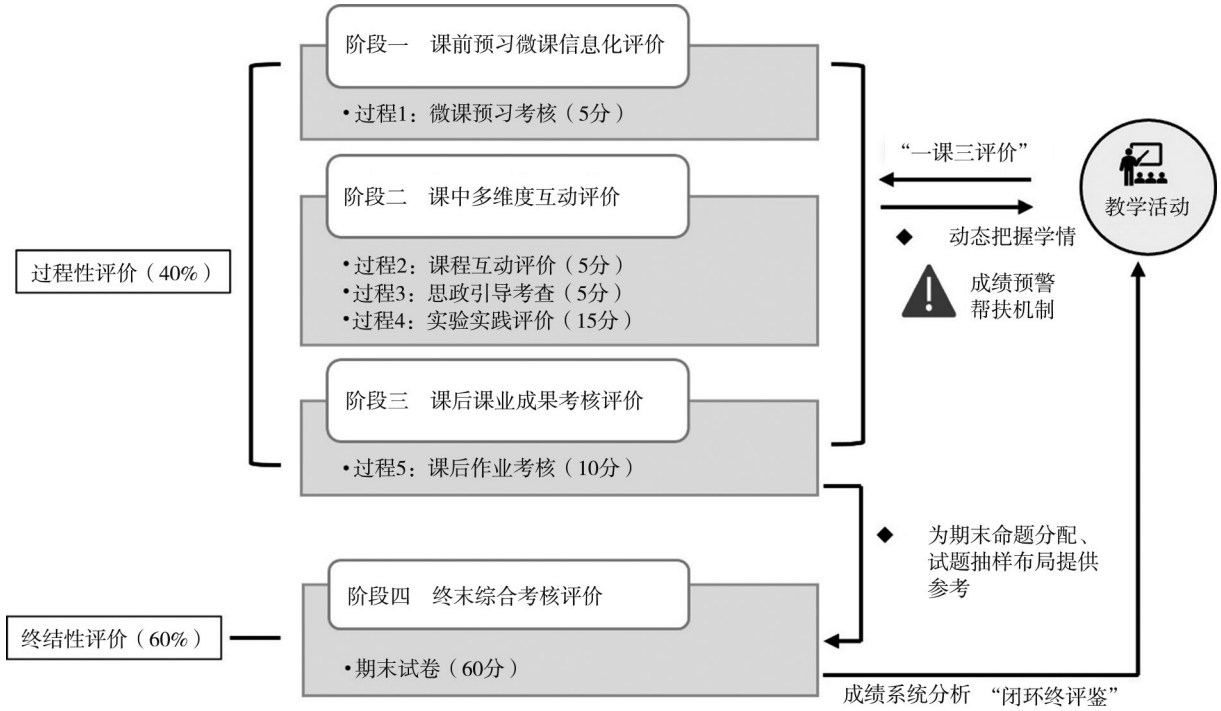


图 2　《医学免疫学》“四阶段多层次”的评学融合体系

1.3.2　多元化考核打造特色命题

在特色命题设计上，系统融入了多元化考核理念：强化基础医学阶段结合临床教学，可提升医学生学习兴趣与岗位胜任力[8]，通过融入临床案例可促进基础—临床融合；依托基础医学优势增设实验分析题，逐步培养学生科研思维；为培养医学生终身学习能力，突出自主学习能力考核，构建延展资源和过程跟踪闭环，借助在线平台提供拓展资源与自学题库，并将自学内容作为试题元素融入考核，可有效观测和评价学生的学习投入度与自主探究能力；同时，在题目中深度融入课程思政元素，在专业知识考查中自然渗透价值引领，实现专业素养培养与思想价值塑造的协同发展。

1.3.3　基于成绩分析系统的智能试卷分析

在考试结束后，课题组运用成绩分析系统对学生成绩及试题进行分析。依据难度（*P*）、区分度

（*D*）、信度（*R*）和效度（*r*）等质量评价指标进行试卷的综合评判（表 1）。同时也可以读取每道试题的难度（*P*）及区分度（*D*）等数值，通过智能化的成绩分析系统及时反映教学和命题评价中的问题。

表 1 试卷的“四度”数据

评价指标	试卷数据	标准[9-10]
难度（*P*）	0.73	*P*≥0.8 过易 0.6≤*P*<0.8 适中 0.4≤*P*<0.6 较难 *P*<0.4 过难
区分度（*D*）	0.34	*D*≥0.4 优秀 0.3≤*D*<0.4 良好 0.2≤*D*<0.3 一般 *D*<0.2 差
信度（*R*）	0.94	*R*>0.85、*r*>0.44 试卷质量合格 *R*>0.9、*r*>0.6 高质量试卷
效度（*r*）	0.45	*R*<0.6、*r*<0.4 试卷质量较差

注：本研究各数值的计算方法为：*P*，通过率/得分率法；*D*，高低分组法；*R*，同质性信度；*r*，结构效度。

以 2022 级临床医学专业学生（105 人）《医学免疫学》考试成绩分析为例，试卷总体难度（*P*=0.73）处于中等水平，符合命题预期。区分度（*D*=0.34）达到良好标准，表明试题能有效区分不同能力水平的学生。信度反映试卷的可靠性和成绩稳定性，本次考试信度为 0.94，说明测试结果具有一致性。本次效度为 0.45，结合其他指标，整体试卷质量合格。从题型分析来看，不同题型表现存在差异。选择题难度适中（*P*=0.77），区分度良好（*D*=0.36）；以记忆为主的简答题难度最低（*P*=0.89），区分度较弱（*D*=0.21），而偏重应用能力的分析题难度较高（*P*=0.65），区分度优秀（*D*=0.48）。试题的难度和区分度数据不仅能反映命题的精准度，还可用于评估试题是否匹配学生能力层次，从而确保测试结果真实有效。

1.4 教学总结及反馈

及时的教学反思和反馈至关重要。第一，教学问题诊断与改进。组织授课教师开展多轮教学反思研讨会，综合课程考核数据对比、教学流程复盘和多学科交流等方式梳理问题，重点优化课程目标达成度、重构过程性评价体系、提升阶段性考核效度。第二，学生参与式质量提升计划。为实现当年问题当年改，教学团队实施假期专项提升计划，结合多元反馈与精准帮扶提升学习效果。组织 3 场线上活动构建质量改进闭环，通过学生座谈会、试题分析坊和经验分享会收集意见。针对后 30% 学生免疫学机制难学的反馈，建立伙伴帮扶计划，定制辅导、突破重难点。通过闭环管理保障教学问题解决，提升临床医学人才培养质量。

2 改革实践效果分析及总结

2.1 考核目标与教学目标、培养目标的一致性实现

考试作为一种评价手段，其目的是服务培养目标。有什么样的目标，就应该有相应的评价来验证所培养的人才是否达到目标要求[11]。构建“三位一体”目标体系的核心价值在于彻底扭转评价与培养脱节的弊端，实现“以评促学、以评促教、评学融合”的闭环。其效果显著体现在：①确保评价效度：考核内容锚定目标（如岗位胜任力），避免考非所学和考非所用，真实反映培养目标达成度。②引导教与学：将考核目标作为教与学的指挥棒，聚焦核心能力培养，明确学生的学习方向。③驱动评价变革：推动考核从静态记忆转向动态、能力导向，强化能力的追踪。④提供改进依据：基于多维评价分析，实现教与学的精准调整与持续改进。这一体系从根本上保障了评价服务于人才培养的核心功能，显著提升了医学人才培养的精准性与质量，使考试评价真正成为衡量和保障临床医学人才免疫学素养与岗位胜任力的有效工具。

2.2 过程性评价的动态追踪与精准反馈

在过程性评价中，一课三评价智能监测与干预系统通过多维度分析，实现对学生学习状态的精准画像（图 3）。以本次授课中 105 名临床医学专业学生为例，过程性评价热图采用标准化评分（0~1 区间），可直观呈现学生课前预习完成度、课堂互动参与度和课后知识巩固度。针对学习困难学生（热图中偏红色），系统可以及时预警并启动帮扶机制，在教学中教师可针对不同层次学生提供个性化支持，助力学生学业发展。

同时，基于平台积累的千余条课后测试数据，构建了知识点掌握度雷达图，重点识别学生知识掌握薄弱点（如 MHC 限制性、免疫耐受机制和 T 细胞发育特点等）。这些数据不仅指导终结性考试实现基础题—综合题—创新题的优化配比，更推动形成动态教学调整机制。针对薄弱知识点，教师采用多种方式的教学手段，实现考核正确率提升。

因此，这一体系彻底解决了传统评价模式中缺乏动态评价的弊端。它不再是期末的一次性快照，而是提供了学习过程的连续影像，真正做到“评”与“学”的融合，使得教师能够随时了解学情、及时发现问题、精准实施干预、动态调整教学。学生也能

获得更及时、更具体的反馈，明确自身定位与努力方向。

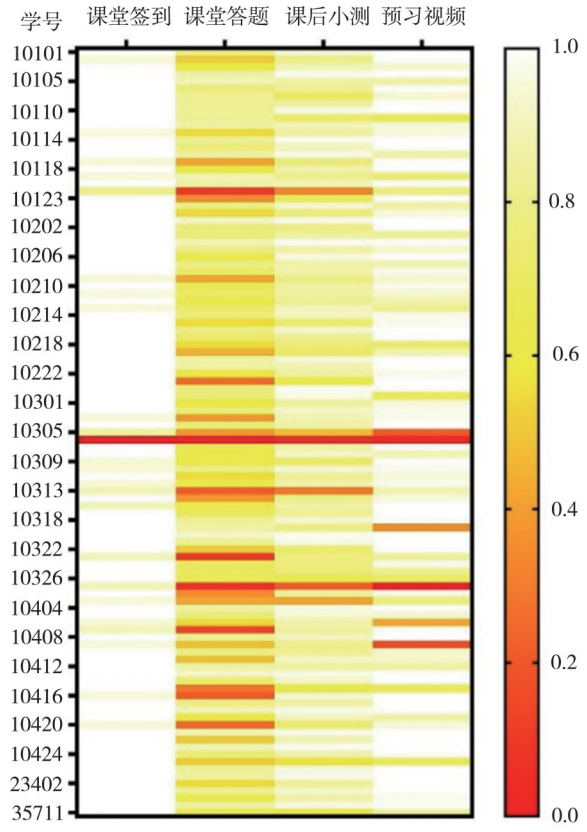


图3　过程性评价热图结果

2.3　命题改革聚焦系统全面评价

传统的终结性评价多为记忆性知识考核，难以全面评价学生的综合能力。为扭转这一问题，课题组在命题中强化解释型和问题解决型试题的出题质量，旨在实现对学生综合能力的精准评价。例如，将知识考查嵌入真实情境考核实践能力。基于高频知识点 $CD8^+$T 细胞的功能，在命题时配合 $CD3^+CD8^+$细胞的流式散点图，先让学生解读数据（识别细胞），再关联功能（图4），有效检验了学生知识迁移和解决实际问题的能力，考核层级从传统的记忆性考核上升为能力考核。该题成绩区分度也从0.33上升到0.58，将死记硬背与灵活运用的学生加以区分。命题的改革显著扭转了“重知识、轻能力”的考核倾向。通过提升能力题型权重、深化情境应用考查，并辅以严谨的试题分析，不仅有效评价了学生的高阶认知与实践能力，更精准诊断了其知识迁移等关键能力的薄弱环节，为后续针对性教学干预提供了实证依据，有力支撑了人才培养目标的达成。

2.4　评价改革驱动学生能力提升

“评学融合”评价体系不仅能精准评估学生的综合能力，还能驱动学生能力的提升。对比未改革前2021级同层次临床医学学生（n=123）期末考试选择题（侧重点和难易程度相似）成绩，“评学融合”组学生回忆型与解释型试题的答题正确率略有提升，但无统计学差异（P>0.05）；而问题解决型试题的正确率明显高于对照组（P<0.05，图5A）。另外，在主观病例分析题中，“评学融合”组的得分率（65.31%）高于对照组（55.42%）。这表明评价改革可有效提升学生的高阶认知能力（如问题解决和临床决策）。

实施课前课后动态评价机制，可激发学生学习动力，增强自主学习意识。通过对课外拓展性视频（共7个，不计入过程性评价分数）的观看完成情况分析发现，“评学融合”组视频完成率（29.66% ± 2.71%）明显高于对照组（17.54% ± 4.58%，P<0.001，图5B），凸显过程性评价通过驱动学习行为变革，促进学生自主知识建构能力发展，最终实现终身学习能力培养的核心目标。

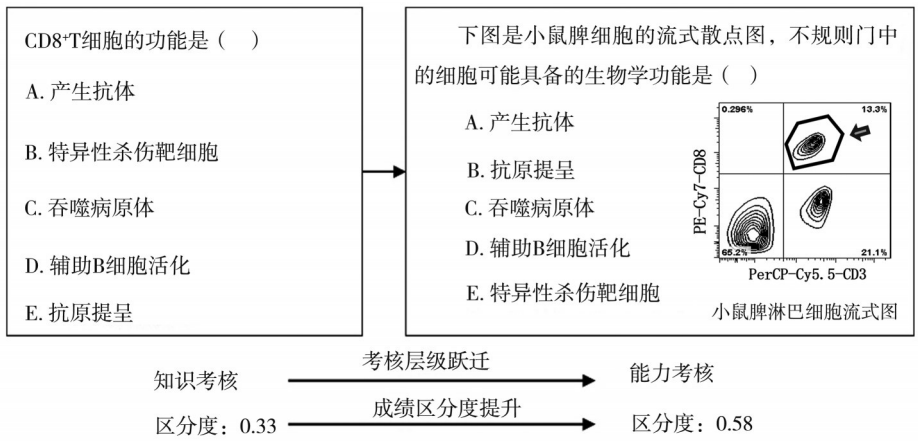
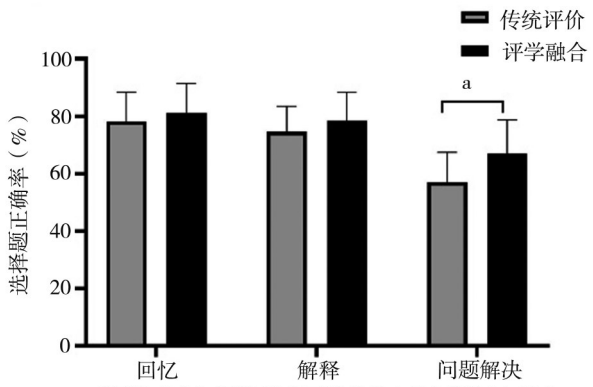
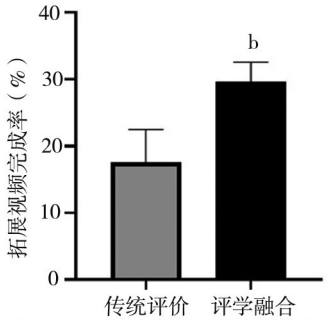


图4　命题改革促进能力考核层级跃迁



A.“评学融合”评价模式可提高学生的问题解决能力



B.“评学融合”评价模式可提升学生学习的主动性

注：a，$P<0.05$；b，$P<0.001$。

图5　评价改革驱动学生能力提升

3　未来展望

通过构建可操作、可量化、可持续的改进方案，《医学免疫学》评价体系已初步完成。目前正在开展以及未来重点计划开展的工作内容包括：①建立多维评价体系，实现精准评估。针对传统考试“评不准”“评不好”的弊端，重点完善过程性评价机制。一是引入学生自评与互评，将课堂表现、实验操作和小组协作纳入评价维度。二是联合临床专家开发典型免疫相关病例题库，推动考核内容从知识记忆向临床思维扩展。三是改革实验考核方式，以实验设计方案撰写替代传统实验报告，鼓励学生绘制思维导图等。②实施分层考核，促进个性化发展。根据学生学业基础差异，设计阶梯式评价方案。对前20%的拔尖学生，通过参与科研课题、撰写综述或参与学科竞赛等方式强化创新能力培养；对中等层次学生，设计文献抄读、案例分析等拓展性活动，激发学习潜能；对基础薄弱学生，采取知识点分级测试、错题针对性训练等措施夯实基础，建立动态达标机制，逐步提升其学习信心。③建立智能命题系统，提升考核科学性。基于历年考试数据，运用大数据技术对试题难度、区分度、信度及知识点覆盖情况进行系统分析，构建动态更新的免疫学试题库。开发AI智能组卷平台，自动生成符合双向细目表的试卷，根据教学目标智能调节试题认知层次，针对不同学生自动匹配差异化考核方案。确保试卷与课程目标契合度达到90%以上。期待通过系统性改革，使课程考核真正成为提升医学人才培养质量的有效抓手，为造就具有创新能力和临床胜任力的卓越医学人才提供支撑。

利益冲突　所有作者声明不存在利益冲突

作者贡献声明　庞维：论文构思、撰写和修订；冯辉：总体指导、论文修订；曹妍、郑丽：数据收集、整理；孙逊：论文审阅、修订

参考文献

[1] 中共中央，国务院. 深化新时代教育评价改革总体方案[EB/OL].（2020-10-13）[2025-07-01]. https://www.gov.cn/zhengce/2020-10/13/content_5551032.htm.

[2] 褚宏启. 以考试评价改革推进拔尖创新人才培养[J]. 中国考试，2024（10）：9-11. DOI: 10.19360/j.cnki.11-3303/g4.2024.10.001.

[3] Christoph Sprung，陈颖. 论高等教育中以学生发展为中心的考试[J]. 应用型高等教育研究，2022，7（4）：68-74. DOI: 10.3969/j.issn.1672-920X.2022.04.012.

[4] Dann R. Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice [J]. Assess Educ Princ Policy Pract, 2014, 21(2): 149-166. DOI: 10.1080/0969594X.2014.898128.

[5] 赖晓华. 高校教学模式与考试评价体系改革研究[J]. 课程教育研究，2017（12）：2-3.

[6] 申继亮，王辉，陈红艳. 深化考试评价改革 服务教育强国建设：第一届教育考试与评价改革论坛会议综述[J]. 中国考试，2025（1）：104-110. DOI: 10.19360/j.cnki.11-3303/g4.2025.01.011.

[7] 曲波，朱亚鑫，屈若祎. 医学教育“评学融合”新理念：从理论到实践[J]. 中国高等医学教育，2024（6）：1-4. DOI: 10.3969/j.issn.1002-1701.2024.06.001.

[8] 孙莉，郎明非，曾常茜，等. 从临床执业医师资格考试成绩分析加强基础与临床结合的教学改革效果[J]. 中华医学教育探索杂志，2023，22（2）：203-207. DOI: 10.3760/cma.j.cn116021-20201022-01173.

[9] 张春辉，席卫文，陈戈煜，等. 高校考试教育测量学部分参数的实证研究[J]. 中国高等医学教育，2024（10）：1-2. DOI: 10.3969/j.issn.1002-1701.2024.10.001.

[10] 杨志明. 考试组卷的若干测量学要求[J]. 教育测量与评价（理论版），2016（1）：62-64. DOI: 10.16518/j.cnki.emae.2016.01.012.

[11] 巩建闽，萧蓓蕾. 如何让考试发挥出人才培养的正能量[J]. 中国高等教育，2013（11）：33-37.

（收稿日期：2025-07-15）
（本文编辑：兰明娟）