·专稿·

第四代医学教育考核评价:精准教育评价

田蕾　孙宝志
中国医科大学医学教育评价与改革研究院,沈阳 110122
通信作者:孙宝志,Email: bzsun@cmu.edu.cn

【摘要】 医学教育考核评价的发展经历了从追求尽可能公平地测量,到成为一个整体系统。现代信息技术的发展给医学教育考评的精准化和个性化提供了基础条件。精准化、智能化、过程化是第四代考评模式的改革需求。精准教育以及精准教育考评的优势在于以胜任力为导向、以学员为中心、以病人为中心,通过个性化、持续分析所设定的教育目标与临床结果之间的关系,向“教—学”双方提供更有意义的反馈,提高了教学的质量和效率、考评的效率和公平性。同时,精准教育评价所适用的范围得以向临床决策、临床思维、全球健康视野等方向实现拓展。

【关键词】 以病人为中心;　考核评价;　精准教育;　精准教育评价
【中图分类号】 R-05
DOI: 10.3760/cma.j.cn116021-20250401-02127

Fourth-generation medical education assessment and evaluation: precision education evaluation

Tian Lei, Sun Baozhi
Institute of Health Professions Education Assessment and Reform, China Medical University, Shenyang 110122, China
Corresponding author: Sun Baozhi, Email: bzsun@cmu.edu.cn

【Abstract】 Medical education assessment and evaluation have transitioned from equitable measurement to becoming a holistic system. The development of modern information technology has laid the foundation for the precision and personalization in medical education assessment and evaluation. Precision, intelligence and process-driven approaches are the reform requirements of the fourth-generation assessment and evaluation. The advantages of precision education and precision education assessment and evaluation are competency-oriented, learner-centered, and patient-centered, which can provide more meaningful feedback to both sides of the "teaching-learning" process through personalized and continuous analysis of the relationship between the set educational objectives and clinical outcomes, thereby improving the quality and efficiency of teaching and the efficiency and fairness of assessment and evaluation. Moreover, the scope of precision education evaluation can be expanded to include clinical decision-making, clinical thinking, and global health perspectives.

【Key words】 Patient-centered;　Assessment and evaluation;　Precision education;　Precision education evaluation
DOI: 10.3760/cma.j.cn116021-20250401-02127

医学教育考核评价为医学教育服务。医学教育的最终目的是培养以病人为中心,能够安全、高质量地提供医疗服务的医学人才,他们中的一些人在未来还可能成为医疗团队的领导者[1-2]。20世纪50年代以来,随着经济发展、社会文明和科学技术的进步,医学教育考核评价也经历了相应的发展阶

段，所使用的方法、技术、理念的每一次进步都是在前一次基础上进行改进，始终朝着保障与推动更高质量、更公平医学服务质量的目标前进。

纵观发展历程，按照考评的目的和方式，医学教育考核评价基本可以被划分为3个时代[3]：测量式考评（assessment as measurement）、判断式考评（assessment as judgement）、全系统式考评（assessment as a system），以及目前医学教育界正在结合现代信息技术发展优势，所提出的下一代医学教育考核评价[4]——基于大数据和人工智能技术，以病人为中心的精准教育考评。

1 医学教育考核评价历史简述

1.1 测量式考评

测量式考评是指用量化形式对学员的知识掌握情况进行描述。20世纪60年代，为了改进当时医学教育考评中存在的偏向主观、不够准确的问题，基于后实证主义[5]，以“可实现、结构化、标准化和客观”的测量式考评理念和方法迅速兴起。测量式考评借鉴了测试心理学（test psychology）的理论和技术，认为能力可以并且应该通过数量或分数来进行观察、控制和表示，并通过难度、区分度、效度和信度等指标来衡量考评质量。医学教育学者受此启发，将医学专业能力定义为一组性格特征（personality traits）的组合，主要包括“知识”“技能”“态度”和“解决问题的能力”，认为这些特征是普遍适用且相互独立的，每个特征都可以被独立测量。例如，理论知识的掌握是专业表现的核心组成部分，能够通过观察和测量，用经过验证的数字分数来表示。测量式考评有效地减少了考评的主观性，增加了考评结果的客观性和科学性[3]。然而直接将测试心理学的评价技术应用在医学教育考核评价中也存在一些问题。首先，测试心理学中衡量测试质量所使用的结构效度和信度并不能体现试题的内容质量，无法反映教育目标与教学内容的实际掌握情况。其次，受到测试心理学中性格特征分类的影响，强调客观区分“高分组—低分组”“通过—不通过”，却忽视了教育评价的过程性，弱化了持续跟踪学习进展的意义。

作为理论知识考试的主要形式，测量式考评仍然是医学教育考评的重要组成部分，是世界各国在医学院入学、毕业、专业执照考试和入职等关键节点所进行的终结性评价的主要方式。测量式考评在医学教育中的应用仍然有着重要的现实作用和意义。

1.2 判断式考评

20世纪90年代，医学教育界意识到仅用专业知识和临床推理并不能全面评价医学生能力，亟待建立涵盖专业精神、沟通能力及批判性思维等的综合评价体系。基于工作场所的考评（work-based assessment，WBA）和客观结构化临床考试开始在全球推广。

判断式考评基于对抽样、结构、信度和效度的深入把握，通过评判者的专业判断[6]，将多维观察转化为等级评定、分数、评语，为学员提供具体、个性化的反馈[3]。考评的质量直接受到各环节的评判质量影响，要求评判者兼具专业素养与标准化评估能力。研究发现，评判的差异主要源于评判者参照体系、观察视角及解释框架，这种差异既可能导致系统性偏倚，也可能反映不同评判者对于能力维度观察的互补意见[7]。为保障考评质量，需要通过系统培训、明确考评目的和细化考评标准等手段进行干预。近年来基于工作场所的考评引入了置信职业行为（entrustable professional activities，EPAs），通过聚焦具体医疗行为的动态发展轨迹，建立基于专业标准的评判标准，有效降低评判结果的差异[8]。相较于传统工作场景考评量表，EPAs展现出更强的公平性[9-10]。判断式考评与前一代考评相比有3个变化：能力不再被视为人格特质，回到了能力本身；考评回到了真实的工作环境，得以将更多维度（临床思维、沟通能力等）纳入考评内容；重新审视了人作为评判主体的作用[11]。

1.3 全系统式考评

2000年后，研究者运用系统理论审视教育、能力和考评的关系，认为“教育是动态解决问题的过程，需根据实际需求灵活调整策略[12]。”这促使医学教育考核评价从“以方法为导向”转向“以全系统为导向”。能力不再被拆解成独立的单元来进行评价，而是被视作一个整体。

在全系统式考评理念中，认为需要将不同类型

（如笔试、OSCE、WBA 评分等）和不同来源（如笔试分数、考官评价、同行评估、病人反馈）的考评方案（programmatic assessment）进行系统整合，实现考评效用的最大化。基于米勒金字塔[13]，考核体系覆盖知识、理解应用、操作表现和实际工作，结合笔试（知识）、判断式评估（应用）和大系统考评（实践），构建多层次综合体系。

至此，考评的功能已经超越了测量，从基于判断提升到全系统，根据多源信息的有效整合，作出最后的决策，从而确保每个具有决定性意义的决策背后都能提供清晰且透明的理由[14-16]，成为推动课程教学价值和学员学习质量的系统性整体工作。

2　全球医学教育第四代考核评价概念的提出

2.1　回应以胜任力为导向、以病人为中心的医学教育

三代考评都在回答如何考评、如何设计考评方法，对为什么考评却没有深度关注。尽管考评工作持续进行，但其结果能否转化为毕业生所获得的行业和社会认可与信赖仍存在不确定性。21 世纪初，第三代医学教育改革提出了以胜任力为导向的医学教育（competency-based medical education，CBME），以“病人及群体健康需求”为目标，强调医学教育成果与毕业生临床服务质量的直接关联[17]。

CBME 以病人为中心和以学员为中心[18-19]，强调基于行业标准定义教育目标，以此指导课程体系的设计与实施，并通过成果导向的考评机制，确保教育全过程始终指向培养具备高质量医疗服务能力的人才。这一改革推动医学教育从知识传授转向能力构建，回归以病人为中心的本质属性。目前，CBME 已从早期理论探索阶段，逐步发展为全球许多国家医学教育领域的实践框架。

然而，学员培养结果与培养目标之间仍然存在差距，低质量和不公平的医疗依然存在。培训资源不足可能是造成这个现象的原因之一，但是未能以学员为中心设计实行考评也是一个主要问题。根据教育目标，以能力发展为目标，创造有效的学习环境，迭代、创新考核评价系统，让学员对自己的教育过程进度有所掌控，逐步有序发展各项能力，并安全、公平、有效地运用在专业工作场所完成相关工作，是医学教育考评工作的改进方向。

2.2　精准教育的概念回应 CBME 以学员为中心的要求

精准教育的目标是“在正确的时间向正确的对象提供正确的教育干预[20]”，改善学员在医疗环境中面对日益增长的工作量、日益复杂的工作要求与知识要求，以及有限的工作和学习时间之间的矛盾。精准教育可以被定义为：“通过整合学员纵向数据（如学习经历、临床实践记录、电子健康档案等）和先进分析技术（如人工智能与决策支持工具），动态设计精准教育干预措施，以持续、及时、周期性的方式，满足个体学员的需求与目标，最终提升教育质量、临床或系统结果[21]”。精准教育的核心优势主要体现在高效、个性化和资源优化上。通过实时数据监测与分析提供即时反馈，解决传统医学教育中对数据收集和分析能力有限、反馈滞后的问题；根据学员纵向学习的各阶段发展来动态调整学习进度与内容进行个性化路径设计；利用自动化工具减轻教师工作负担，同时通过学员参与式设计增强学习自主性，实现资源优化。精准教育的理念里，固定学习时限被打破、强化用数据去驱动决策，融合临床实践与教育数据去提升干预措施的科学性，从而促进技术赋能教育公平[22]。

3　第四代医学教育考核评价

3.1　第四代医学教育考核评价的设想

第四代医学教育考核评价设想是服务于精准教育，以 CBME 双中心（以病人为中心、以学员为中心）为内核的“精准教育评价（precision education assessment）”[4]。

精准教育评价是通过数据和技术手段，对学员的表现进行过程性的、全面且个性化的考评。其目标是提供更有意义的反馈，帮助学员识别自己的优势和不足，从而优化学习过程。

精准教育评价的特点体现在通过提供贴合个人需求的反馈，依据每位学员的独特表现定制专属建议，让学习更具针对性；通过实现教育机构间学员数据的互通共享，保障学员在不同教育阶段的学习能够连贯有序，避免出现断层；凭借客观全面的数据模型，有效消除主观偏见，确保所有学员都能

得到公正的评估与反馈，让教育评价更加科学、公平。

一方面，技术革新为精准教育评价提供了条件支撑：人工智能、大数据、大语言模型、虚拟现实等新信息技术的应用，使持续收集和分析工作环境中的行为表现数据成为可能。如果能证实这些分析的结果可信，就可以将数据进行分析，进而转化为具体改进建议，这就为改善医学教育考评中长期存在的效率和公平性问题提供了新的思路。有了信息技术的赋能，考评的内容范围将可能拓展至学员的决策能力、思维、全球健康视野等维度，也将能够满足医学教育精准化、智能化的改革需求，形成高质量、公平的新型考评。另一方面，借助信息技术，实现全面、个性化、多维度、系统化的考评结果，生成公平、立体、有意义的反馈信息，帮助学员识别自身优势与不足，优化学习策略[4,21,23]。

3.2 利用医疗数据，寻找评价教育结果的证据和价值

医学教育评价领域正在探索信息技术支持下的数据为核心的考评模式。通过挖掘信息系统在学员实施医疗行为过程中记录的数据，对教育成果与患者预后等临床结果间的关联进行验证，开展数据驱动的持续的教育考评工作[24]。用这种方式减少对人的观察的依赖，借助数据所体现的客观性与全面性，可以有效规避评价者选择性注意和评价者差异导致的偏差，突破传统过程性评价只能通过各个时间段的考核所形成的技术瓶颈。

通过分析电子病例（electronic health records，EHR）所储存的患者医疗过程和预后的原始电子数据，可以持续地跟踪学员在临床轮转中的学习进展，评估学员在临床实习阶段和住院医师阶段的知识掌握运用情况、临床技能水平，以及对患者的影响，识别他们接触病例的类型，调整教育计划。已有研究通过查看学员在EHR中输入的信息进行评分，并填入仪表板（dashboard），向学员进行反馈，帮助学员调整学习进度，弥补学习过程中的不足。还有研究通过分析学员使用电子病例信息时点击的次数和点击的模式，尝试评价学员对工作流程的熟悉程度和完成工作的效率[25]。

可穿戴设备和运动跟踪系统记录所收集到的工作流程与工作表现等相关的音频、视频、行为等数据结合人工智能和大型语言模型等新的分析方法，可能提供前技术时代所无法收集和处理的信息，如果分析的信度和效度合理，那么将能实现对每次临床工作流程、手术表现、沟通情况和效率进行持续、详细的评估[26-28]。

美国辛辛那提儿童医学中心、辛辛那提医学院的研究者在儿科和内科专业开发了住院医师敏感质量测量（resident-sensitive quality measures，RSQMs）和“实时归因及自动化医疗评估”（trainee attributable and automatable care evaluations in real-time，TRACERs）[29]，用临床绩效指标对个体受训者进行评估。RSQMs和TRACERs所需调用的数据基于EHR自动生成。

RSQMs通过小组共识方法制定的质量指标，专为评估住院医师个体表现设计，质量指标基于两大核心标准：聚焦特定疾病或治疗场景中直接影响诊疗质量的关键环节（如医嘱、病史采集、体格检查等行为）；能够排除团队协作或系统性因素，明确归因于住院医师个人职责范围内的操作。RSQMs的优势是通过实时自动化采集与疾病及医疗行为紧密相关的数据（如医嘱适当性、文档完整性等），将临床工作质量与教育考评直接关联，精准识别住院医师在诊疗流程中的表现差异，并提供针对性反馈以改进技能[30-31]，但是它的缺点是难以实现跨项目扩展[32]。

TRACERs是通过识别受训者提供医疗的临床场景，从EHR中提取信息以及计算不同医疗措施的完成率来开发的。TRACERs概念框架包括：对治疗和学员“有意义（meaningful）”、对个体学员“可归因（attributable）”、最少的人工输入“自动化（automatable）”、跨电子病例记录和培训环境“可扩展（scalable）”，以及适合形成性教育反馈循环“实时性（real-time）”[32]。与RSQMs相比，TRACERs实现了跨机构、跨系统的标准化数据提取；覆盖医学生至住院医师的全培训周期；利用实时数据生成动态反馈；通过精准识别学员在特定临床场景下的独立行为（如医嘱录入、病历文档）。

3.3 以学员为中心，设计精准教育评价的方案

选择恰当的评价依据。患者预后与治疗之间

的关系是复杂的，患者的预后不仅仅由治疗计划决定，还与可用的医院服务、社会经济环境及其他卫生保健专业人员的介入密切相关。在选择患者结果作为评价依据时，要确保它们能够真实反映学员的能力。

减少算法对评价数据的影响。评价过程中使用的分析模型、数据类型，都可能会影响输出结果。在创建使用这些数据的分析算法时，需要尽可能减少算法对评价数据的影响，确保公平[33]。

增加评价设计的透明度，让学员参与考评体系的建构。学员需要经常接受考评，却很少能够参与到考评系统的设计过程[34]。作为参与者，学员在提供意见和改进考评工具设计方面有自己的独特视角，可以帮助发现评价目的不明确的地方，成为创新思维和教育变革的重要推动力。将学员纳入考评系统设计，不仅能增强他们的认同感和投入意愿，还能建立信任，提升评价的效力[35]。

3.4　精准教育评价的挑战

对医疗数据的管理和使用规范将是精准教育评价的挑战。精准教育评价依赖大量的数据，数据的收集、存储和分析需要复杂的技术支持。分散在庞大的医学教育生态数据系统之间是否兼容，其中所存储的数据能否整合使用，是决定精准教育评价是否能够推广和应用的重要因素。此外，医学教育工作者接受由数据所驱动的教学与考评需要时间与文化上的适应。从数据角度本身，必须制定用于精准教育评价过程中的数据使用原则和标准，明确数据共享、复制、扩展等规范，确保数据支持跨机构、跨系统兼容使用。同时，确保数据的准确性和隐私保护不能忽视。

4　结语

医学教育的考评时代由考评的方式定义，在追求客观、真实、多维、综合的过程中稳步发展。从辩证的角度看，基于固定时间和条件的考评预测往往不准确，甚至无效。在设计考评方案时，需要优先关注反馈和学员发展，减少使用过时或缺乏证据支持的考评方法，保持考评透明度，建立个人与考评系统间的信任。

先进的技术和数据资料分析方法并不能完全解决教育考评中的所有问题，也并不可能完全代替人的判断。综合评估和以理论为依据的群体决策过程依然重要[35]。伴随技术革新与社会观念的快速迭代，传统的工作经验和思维模式正面临前所未有的挑战。医学教育考评体系必将经历变革、发展与优化的过程。未来科技在医疗领域的深度融合，也许会推动医学模式发生根本性变革，医学教育考评的核心内容与方式也会随之调整以适应新的需求和任务。

利益冲突　所有作者声明不存在利益冲突

作者贡献声明　田蕾：撰写论文；孙宝志：论文构思、总体把关及审定

参考文献

[1] Hodge S. The origins of competency-based training [J]. Austral J Adult Learn, 2007(2): 179-209.

[2] Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: theory to practice [J]. Med Teach, 2010, 32(8): 638-645. DOI: 10.3109/0142159x.2010.501190.

[3] Schuwirth LWT, van der Vleuten CPM. A history of assessment in medical education [J]. Adv Health Sci Educ Theory Pract, 2020, 25(5): 1045-1056. DOI: 10.1007/s10459-020-10003-0.

[4] Drake CB, Heery LM, Burk-Rafel J, et al. A theoretical foundation to inform the implementation of precision education and assessment [J]. Acad Med, 2024, 99(4S Suppl 1): S30-S34. DOI: 10.1097/ACM.0000000000005604.

[5] Schumacher DJ, Kinnear B, Burk-Rafel J, et al. The next era of assessment: can ensuring high-quality, equitable patient care be the defining characteristic? [J]. Acad Med, 2024, 99(4S Suppl 1): S7-S13. DOI: 10.1097/ACM.0000000000005603.

[6] Epstein RM, Hundert EM. Defining and assessing professional competence [J]. JAMA, 2002, 287(2): 226-235. DOI: 10.1001/jama.287.2.226.

[7] Gingerich A, Ramlo SE, van der Vleuten CPM, et al. Inter-rater variability as mutual disagreement: identifying raters' divergent points of view [J]. Adv Health Sci Educ Theory Pract, 2017, 22(4): 819-838. DOI: 10.1007/s10459-016-9711-8.

[8] Valentine N, Wignes J, Benson J, et al. Entrustable professional activities for workplace assessment of general practice trainees [J]. Med J Aust, 2019, 210(8): 354-359. DOI: 10.5694/mja2.50130.

[9] Crossley J, Johnson G, Booth J, et al. Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales [J]. Med Educ, 2011, 45(6): 560-569. DOI: 10.1111/j.1365-2923.2010.03913.x.

[10] Crossley J, Jolly B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people [J]. Med Educ, 2012, 46(1): 28-37. DOI: 10.1111/j.

1365-2923.2011.04166.x.

[11] Norcini JJ, Blank LL, Arnold GK, et al. The mini-CEX (clinical evaluation exercise): a preliminary investigation [J]. Ann Intern Med, 1995, 123(10): 795-799. DOI: 10.7326/0003-4819-123-10-199511150-00008.

[12] Ulrich W. The quest for competence in systemic research and practice [J]. Syst Res Behav Sci, 2001, 18(1): 3-28. DOI: 10.1002/sres.366.

[13] Miller GE. The assessment of clinical skills/competence/performance [J]. Acad Med, 1990, 65(9 Suppl): S63-S67. DOI: 10.1097/00001888-199009000-00045.

[14] van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes [J]. Med Educ, 2005, 39(3): 309-317. DOI: 10.1111/j.1365-2929.2005.02094.x.

[15] van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose [J]. Med Teach, 2012, 34(3): 205-214. DOI: 10.3109/0142159X.2012.652239.

[16] Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. Twelve tips for programmatic assessment [J]. Med Teach, 2015, 37(7): 641-646. DOI: 10.3109/0142159X.2014.973388.

[17] Schumacher DJ, Santen SA, Pugh CM, et al. Foreword: the next era of assessment and precision education [J]. Acad Med, 2024, 99(4S Suppl 1): S1-S6. DOI: 10.1097/ACM.0000000000005609.

[18] Frenk J, Chen L, Bhutta ZA, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world [J]. Rev Peru Med Exp Salud Publica, 2011, 28(2): 337-341. DOI: 10.1590/s1726-46342011000200023.

[19] Van Melle E, Frank JR, Holmboe ES, et al. A core components framework for evaluating implementation of competency-based medical education programs [J]. Acad Med, 2019, 94(7): 1002-1009. DOI: 10.1097/ACM.0000000000002743.

[20] Cook CR, Kilgus SP, Burns MK. Advancing the science and practice of precision education to enhance student outcomes [J]. J Sch Psychol, 2018(66): 4-10. DOI: 10.1016/j.jsp.2017.11.004.

[21] Desai SV, Burk-Rafel J, Lomis KD, et al. Precision education: the future of lifelong learning in medicine [J]. Acad Med, 2024, 99(4S Suppl 1): S14-S20. DOI: 10.1097/ACM.0000000000005601.

[22] Triola MM, Burk-Rafel J. Precision medical education [J]. Acad Med, 2023, 98(7): 775-781. DOI: 10.1097/acm.0000000000005227.

[23] Chan T, Sebok-Syer S, Thoma B, et al. Learning analytics in medical education assessment: the past, the present, and the future [J]. AEM Educ Train, 2018, 2(2): 178-187. DOI: 10.1002/aet2.10087.

[24] Smirnova A, Chahine S, Milani C, et al. Using resident-sensitive quality measures derived from electronic health record data to assess residents' performance in pediatric emergency medicine [J]. Acad Med, 2023, 98(3): 367-375. DOI: 10.1097/ACM.0000000000005084.

[25] Warm EJ, Kinnear B, Knopp MI, et al. Ambulatory long block: a model of precision education and assessment for internal medicine residents [J]. Acad Med, 2024, 99(4S Suppl 1): S57-S63. DOI: 10.1097/ACM.0000000000005610.[PubMed]

[26] Khosravi S, Bailey SG, Parvizi H, et al. Wearable sensors for learning enhancement in higher education [J]. Sensors (Basel), 2022, 22(19): 7633. DOI: 10.3390/s22197633.

[27] Tolsgaard MG, Pusic MV, Sebok-Syer SS, et al. The fundamentals of Artificial Intelligence in medical education research: AMEE Guide No. 156 [J]. Med Teach, 2023, 45(6): 565-573. DOI: 10.1080/0142159X.2023.2180340.

[28] Guerin S, Huaulmé A, Lavoue V, et al. Review of automated performance metrics to assess surgical technical skills in robot-assisted laparoscopy [J]. Surg Endosc, 2022, 36(2): 853-870. DOI: 10.1007/s00464-021-08792-5.

[29] Schumacher DJ, Gielissen K, Kinnear B. Competency-based medical education: connecting training outcomes to patient care [J]. Curr Probl Pediatr Adolesc Health Care, 2024, 54(10): 101675. DOI: 10.1016/j.cppeds.2024.101675.

[30] Schumacher DJ, Martini A, Holmboe E, et al. Developing resident-sensitive quality measures: engaging stakeholders to inform next steps [J]. Acad Pediatr, 2019, 19(2): 177-185. DOI: 10.1016/j.acap.2018.09.013.

[31] Schumacher DJ, Holmboe ES, van der Vleuten C, et al. Developing resident-sensitive quality measures: a model from pediatric emergency medicine [J]. Acad Med, 2018, 93(7): 1071-1078. DOI: 10.1097/ACM.0000000000002093.

[32] Burk-Rafel J, Sebok-Syer SS, Santen SA, et al. TRainee attributable & automatable care evaluations in real-time (TRACERs): a scalable approach for linking education to patient care [J]. Perspect Med Educ, 2023, 12(1): 149-159. DOI: 10.5334/pme.1013.

[33] Kinnear B, Caretta-Weyer H, Lam ACL, et al. Finding medicine's moneyball: how lessons from major league baseball can advance assessment in precision education [J]. Acad Med, 2024, 99(4S Suppl 1): S35-S41. DOI: 10.1097/ACM.0000000000005600.

[34] Burk-Rafel J, Jones RL, Farlow JL. Engaging learners to advance medical education [J]. Acad Med, 2017, 92(4): 437-440. DOI: 10.1097/ACM.0000000000001602.

[35] Marcotte K, Negrete Manriquez JA, Hunt M, et al. Trainees' perspectives on the next era of assessment and precision education [J]. Acad Med, 2024, 99(4S Suppl 1): S25-S29. DOI: 10.1097/ACM.0000000000005602.

（收稿日期：2025-04-01）
（本文编辑：唐宗顺）