

临床教学质量评价系统的研究进展

刘伟文, 黄才埔, 李永强*

南方医科大学第三附属医院全科医学科, 广东 广州

收稿日期: 2025年7月30日; 录用日期: 2025年8月29日; 发布日期: 2025年9月10日

摘要

背景: 教学医院承担医疗服务和人才培养双重职责, 临床教师专业发展是提升培训质量的核心。荷兰教学质量评估系统(System for Evaluation of Teaching Qualities, SETQ量表)为教师评价提供了重要的参考价值, 并在多个国家得到了广泛验证。方法: 本文对2009~2024年PubMed和中国知网收录的28篇SETQ相关文献, 重点分析其心理测量学特征及应用效果。结果: 在PubMed和中国知网上, 涉及“SETQ”的研究共30篇, 经过筛查后保留28篇。这些研究主要探讨了SETQ量表在不同应用场景下的心理测量特征, 包括信度分析、结构效度分析, 以及量表结果反馈的实施情况、人口特征与教育环境感知之间的相关性。结论: 基于SETQ量表的教师评价系统对临床医学教育具有积极意义, 有助于提高教师的教学质量, 促进临床医学教育的持续发展。

关键词

教学质量评估, 临床医学, 住院医师规范化培训, 心理测量特征, 人口特征

Research Progress of Clinical Teaching Quality Evaluation System

Weiwen Liu, Caipu Huang, Yongqiang Li*

Department of General Practice, The Third Hospital of Southern Medical University, Guangzhou Guangdong

Received: Jul. 30th, 2025; accepted: Aug. 29th, 2025; published: Sep. 10th, 2025

Abstract

Background: As healthcare institutions dedicated to both medical service and talent development, teaching hospitals rely on the professional growth of clinical faculty as the cornerstone of enhancing training quality. The Dutch-developed System for Evaluation of Teaching Qualities (SETQ) offers a robust benchmark for appraising medical educators and has been validated across multiple countries.

*通讯作者。

文章引用: 刘伟文, 黄才埔, 李永强. 临床教学质量评价系统的研究进展[J]. 教育进展, 2025, 15(9): 665-672.

DOI: 10.12677/ae.2025.1591722

Methods: This study examines 28 articles on the System for Evaluation of Teaching Qualities (SETQ), retrieved from both PubMed and the China National Knowledge Infrastructure (CNKI) between 2009 and 2024, focusing specifically on their psychometric characteristics and implemented effectiveness. **Results:** This study identified 30 articles related to the “SETQ” on PubMed and the CNKI. After screening, 28 were included for analysis. These studies primarily investigate the psychometric properties of the SETQ questionnaire across various contexts, including reliability analysis, construct validity analysis, as well as the implementation of feedback based on its results, and the correlation between demographic characteristics and perceptions of the educational environment. **Conclusion:** The SETQ-based teacher evaluation system holds positive significance for clinical medical education, as it contributes to enhancing teaching quality and promoting the sustainable development of clinical medical training.

Keywords

Teacher Quality Assessment, Clinical Medicine, Standardized Training for Residents, Psychometric Characteristics, Demographic Characteristics

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 前言

自 2013 年我国住院医师规范化培训全面启动以来,制度建设已基本完成,住培工作逐渐由形式规范向内涵质量提升转变。带教师资队伍的建设是提升住培质量的关键环节[1]。然而,目前国家层面对临床师资的准入、认定和评价等方面缺乏统一标准,导致临床师资的带教能力参差不齐,师资评价机制不完善,激励机制也不健全。这些问题将阻碍住培工作的进一步推进,直接影响住培的质量和可持续发展。临床师资的胜任力,即师资应当具备的能力,是评价临床师资能力的综合指标。本文通过综述国内外临床教学质量评价系统(SETQ 量表)的相关研究,旨在为建立一套适应中国国情的、客观科学的临床师资评价量表提供参考,以此为临床师资队伍建设提供新的思路。

2. 引言

教学医院的使命不仅在于提供卓越的患者护理,更在于为未来的医疗人才提供系统而专业的培训。大量研究已证明,培训质量直接影响学员和医生在专业成长过程中的学习成果,同时也深刻影响患者的就医体验。如今,提升专科医疗培训质量已成为业界的共识,这与医疗培训的现代化进程密切相关,而临床教师的专业发展则是这一进程的关键环节。即便是经验丰富的教师,往往对自身作为教师的长处和短处、住院医师对其行为的看法、提升教学技能的潜力以及作为榜样的有效性缺乏深入的洞察力。因此,为临床教师提供关于培训质量的反馈,不仅是一种激励,也是一种支持。尽管评估方法多种多样,但最直接有效的方式之一是对住院医师进行问卷调查。作为培训的直接受益者,学员对教师的优缺点有着最为直观的感受,因此在评价教师时具有重要的发言权。为确保培训质量,荷兰阿姆斯特丹学术医学中心的麻醉学部门开发了一种独特的评估模型——教学质量评价系统(SETQ),旨在检测培训师的教学水平。该评估方法在国际上已相当成熟,但在中国的应用范围仍然相对有限。SETQ 量表模型作为一种评价工具,已经具备了一定的理论基础和丰富的应用经验。本综述将统计与描述 SETQ 在不同国家的研究内容及成果[2]-[8]。

3. 方法

在 PubMed 官网及中国知网上,以“SETQ”和“System for Evaluation of Teaching Qualities”为关键

词进行检索，涉及 2010 年至 2024 年的相关文献共 30 篇。在对所有文献进行阅读后，排除了 2 篇不相关的文章，具体标题和原因见表 1。最终纳入了 28 篇涉及使用 SETQ 对实习生、研究生或住院医师进行教育环境评估及相关研究的文献，其中近 5 年内发表的文献有 13 篇。在纳入的文献中，英文文献 25 篇，中文文献 5 篇，有 12 篇研究性论文的影响因子大于 3。随后，将这些文章中的研究方法、人口特征和研究结果进行了汇总和分析。流程图见图 1。

Table 1. Excluded literature (2 studies) with rationale for exclusion
表 1. 未纳入的 2 篇文献及排除原因

标题	作者	原因
Variability of residents' ratings of faculty's teaching performance measured by five-and seven-point response scales	Maarten P M Debets 1, Renée A Scheepers 2, Benjamin C M Boerebach 3, Onyebuchi Aarah 4 5 6 7, Kiki M J M H Lombarts 3	该研究仅比较了 5 分和 7 分制对量表结果的影响，两个量表测量的教学表现结果(几乎)相似。
Evaluating clinicians' teaching performance	Benjamin C.M. Boerebach	博士项目描述

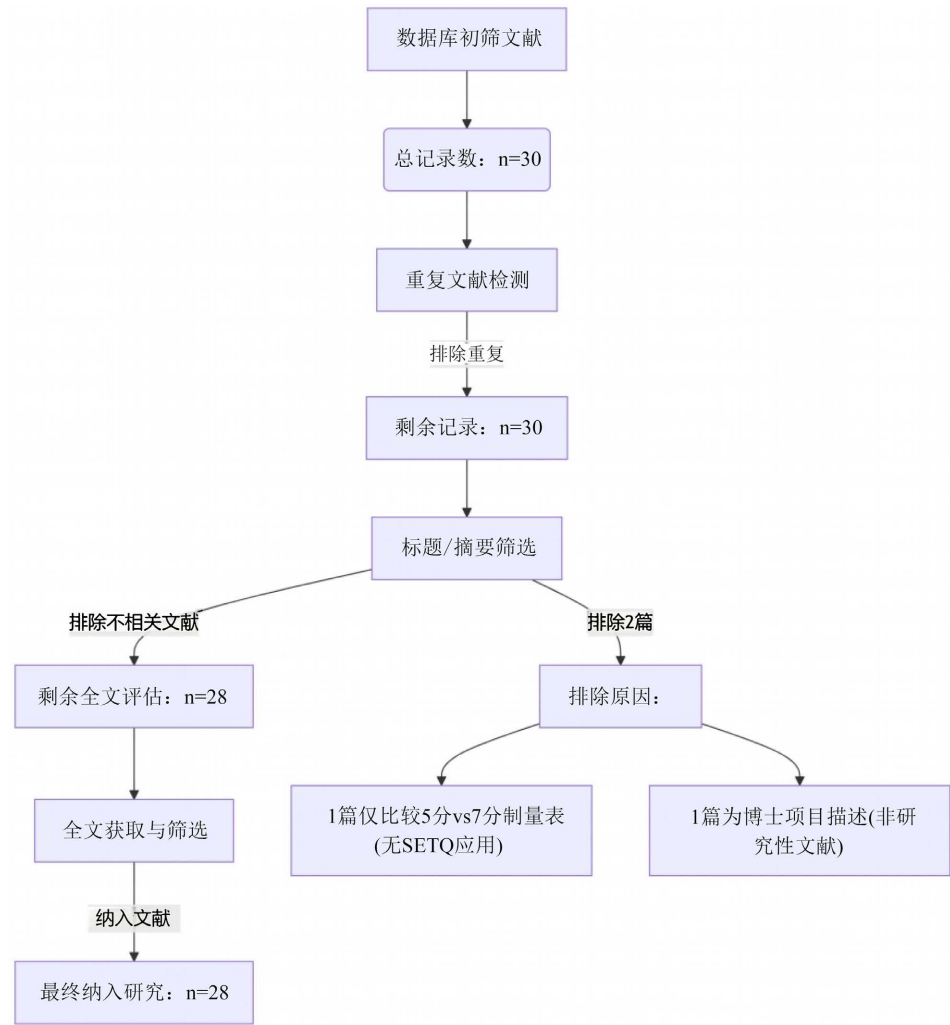


Figure 1. Flowchart
图 1. 流程图

4. 结果

共有 28 项研究涉及荷兰、巴林王国、美国、瑞士和伊朗等地区。SETQ 在定量研究中的应用主要集中在信度、效度测试以及与学习成果的相关性分析。信度测试通过 Cronbach's Alpha 系数评估量表的内部一致性,通常 Alpha 值大于 0.7 表明量表具有较好的可靠性。效度测试确保量表能够准确测量教学质量的各个维度,包括通过因子分析(如探索性因子分析 EFA 和验证性因子分析 CFA)验证教学质量的不同维度是否有效并且独立。此外,SETQ 还通过回归分析和路径分析探讨教学质量与学生学习成果(如学术成绩和学习满意度)之间的关系,发现教学质量的提高通常与学生的学习效果呈正相关。通过这些定量研究,SETQ 能够为教育管理者提供科学依据,帮助提高教学质量并优化教学实践。在定性研究方面,主要通过分析开放性反馈、访谈和焦点小组讨论等方式,深入探讨学生对教学质量的理解和反馈。定性研究的一个重要应用是内容分析,通过对量表中开放性问题的反馈进行分析,提取学生关于教学清晰度、课堂互动和教师可接触性等方面的评价。除了量表数据,定性研究还通过对学生和教师的访谈和焦点小组讨论,进一步理解教学质量的内涵,揭示量表评分背后的原因及潜在问题。此外,由于 SETQ 在不同国家和文化背景中应用,定性研究还特别关注文化适应性和情境理解,通过与不同文化背景的学生和教师的互动,研究者能够评估 SETQ 量表是否需要调整,以适应不同地区或学术文化的需求[1] [9]-[12]。

4.1. SETQ 量表的心理测量特征

心理测量特性是心理学领域的基本概念,涉及心理测量工具的可靠性和有效性,通常包括信度、效度、敏感性、偏移性和实用性。Kiki Lombarts 等人(2009)设计的初始 SETQ 量表包含 24 个项目,这表明每位参与者填写问卷所需的时间不会很长,并预示着该量表具有较高的实用性。

在 13 篇关于量表研制或改良及心理测量特征研究的文献中,均使用了 Cronbach's α 系数评估 SETQ 量表的信度。Kiki Lombarts 等人最初开发了住院医师版和教师自评版的量表,其中住院医师量表的内部一致性从“对住院医师的专业态度”的 0.89 到“目标交流”和“对住院医师的相关评价”的 0.94 不等。相比之下,教师自我评价量表的 Cronbach's 系数较低,从“学习氛围”的 0.57 到“目标沟通”的 0.86 不等[13]。随后,Kiki Lombarts 等人将量表的应用范围从最初的麻醉学扩展至妇产科和外科。妇产科住院医师量表的内部一致性为 0.86 到 0.89 [6],而外科住院医师量表的内部一致性则在 0.90 到 0.98 之间[8]。随着时间的发展,Kiki Lombarts 等人进一步加强了 SETQ 问卷的效度和信度验证,其学员个人问卷的内部一致性系数介于 0.85 至 0.91 之间,教师自我评价问卷的内部一致性系数介于 0.73 至 0.84 之间[14]。

此外,SETQ 量表经过改良,形成了 SETQ Smart 量表,其住院医师版的内部一致性提升至 0.87 到 0.98,教师自评版的内部一致性则为 0.74 到 0.96 [15]。在逐一排除每个问题进行分析并删除分数低于 0.70 的问题后,研究发现即便删除某些问题,测量信度(Cronbach's alpha)也未显著改善,这表明所有问题之间存在较强的相关性。

在国际应用方面,Ahmed Al Ansari 等人(2017)将 SETQ 量表引入巴林王国,并根据实际情况进行了改良,研究其在临床医学专业本科教育中的心理测量学特性。其住院医师量表的内部一致性从 0.94 提升至 0.97,教师自评量表的内部一致性从 0.81 提升至 0.91 [16]。Fateme Tahmasebi Boldaji 等人则将 SETQ 量表翻译为波斯语,并在伊朗的临床医学教育阶段评估其适用性,内部一致性从 0.722 提升至 0.882 [17]。同年,Jan Breckwoldt 等人将 SETQ 量表引入瑞士,并结合 CanMEDS 2015 项目改良为 SwissSETQ 量表,研究其心理测量学特性,内部一致性范围从 0.718 提升至 0.974。由于个别项目的回答数据可能呈现偏态

分布, Cronbach's alpha 有时会低估内部一致性程度, 因此还使用了“欧米茄总值”和“最大下限”(GLB)作为衡量测验可靠性的替代指标[18]。

4.2. SETQ 量表结果对教师的影响

Kiki Lombarts 与 Tjitske S.R. van Engelen 等人(2016)对 SETQ 量表的结果进行了后续研究, 调查住院医师规范化培训项目的负责人。结果显示, 学员对导师教学表现的反馈(即量表结果)未被忽视, 绝大多数导师会在小组会议上跟进这些教学表现的反馈[19]。同年, Renéem. Van der Leeuw 等人在荷兰进行的一项前瞻性多中心队列研究探讨了住院医师填写的 SETQ 量表各个方面如何影响教师教学绩效的改善。研究表明, 建立临床教学绩效评价体系有助于提升住院医师规范化培训的临床教学绩效, 表现优秀的教师也能改善住院医师的认知能力[20]。

此外, Raymund L. Yong 等人(2022)将 SETQ 量表引入美国后, 对神经外科学员实施改良版教学质量评估系统(SETQ)进行评估, 每 6 个月进行一次, 持续了 4 年。结果显示, 在这 4 年期间, SETQ 的实施与教师教学表现的显著改善相关, 并指导了培训计划中的课程变动, 从而改善了教育结果[21]。

4.3. SETQ 量表与其他量表

4.3.1. 乌得勒支工作投入量表与杰弗逊共情量表

Lenny Lases 等人(2014)对荷兰 271 名外科和妇科住院医师进行问卷调查, 采用 SETQ 量表收集住院医师对教师教学表现的评价。同时, 研究者邀请住院医师使用简易乌得勒支工作投入量表(Utrecht Work Engagement Scale)和杰弗逊共情量表(Jefferson Scale of Physician Empathy)进行工作投入和共情的标准化测量。通过广义估计方程进行线性回归分析, 发现住院医师的敬业度和同理心与对师资的评价之间存在正相关关系[9]。

4.3.2. 住院医师学习氛围量表

Maas Jan Heineman 等人(2014)使用 SETQ 量表评估住院医师对教师的教学表现, 同时应用荷兰住院医师学习氛围量表(D-RECT)评估住院医师的学习氛围。在 11 个学习氛围分量表中, 有 3 个分量表与良好的教学表现显著相关, 分别是: “指导和评估”、“工作适应住院医师能力”和“正规教育”。因此, 改善住院医师规范化培训的学习氛围对教师个人教学绩效的评价具有积极影响[22]。

4.3.3. 大五人格量表

Renée A Scheepers 等人(2014)对荷兰外科和非外科的主治医师使用翻译改良后的大五人格量表(The Validated 10-Item Big Five Inventory)进行人格特征评估, 并采用 SETQ 量表评价主治医师的整体教学表现。相关统计分析结果显示, 外向性较强的主治医师被一致评价为较好的导师, 而开放程度高的外科主治医师则被认为反馈不足, 认真负责的非手术主治医师在评估住院医师方面表现更为突出[23]。

4.3.4. Fleming VARK 学习风格量表(学习偏好)和 Staffordshire 教学风格评估量表(教学风格)

Karen J Dickinson 等人(2021)对美国外科住院医师和主治医师使用 SETQ 量表、学习风格量表和教学风格评估量表, 探讨主治医师的教学方式及总体学习偏好与住院医师基于其评价的表现之间的关系。研究表明, 教学风格可能影响外科教育者的有效性, 大多数外科教育者表现出多样化的学习偏好。未来的研究应以多机构分析为导向, 进一步探讨教学风格与教育效果之间的潜在关系[10]。

4.3.5. Maslach 职业倦怠量表、SEF:MED 情绪智力自我评估

James M Lewis 等人(2021)使用 Maslach 职业倦怠量表、情绪智力(EI)自我评估和 SETQ-Smart 量表评估外科教师的情绪智力、职业倦怠与教学能力。分析结果显示, 情绪智力与外科教师的教学能力呈正相

关, 而职业倦怠与住院医师对教师的教学评分之间的相关性较弱, 但趋势相似[11]。

4.3.6. 改善住院医师教育工具

Nancy J Zhou 等人(2021)围绕 JHFIRE 工具, 即住院医师教育改进联合会议(Joint Huddles for Improvement Resident Education)的介绍和使用, 采用 SETQ 量表对干预前后的外科教育质量进行评估。该工具在提升住院医师的外科教育方面显示出显著的主观效益。尽管有压倒性的积极反馈, 但在干预前和干预后的 SETQ 评分(4.25 ± 0.55 比 4.10 ± 0.88 , $p = 0.63$)未显示出总体差异[12]。

4.4. 国内师资评价

尽管国家层面对住院医师规范化培训师资提出了要求, 具体的实施和评估标准仍在探索和完善之中。各地培训基地医院根据自身实际情况进行尝试和创新, 力求建立最适合自己的住培师资评估体系[1]。

南京医科大学第一附属医院基于柯氏模型, 评估住院医师规范化培训中医学模拟师资的培训效果。这为提升培训基地的师资带教能力和培训内涵质量提供了有益的参考[24]。安徽省对连续四年开展的全科医学师资培训效果进行了研究, 比较各期学员的结业综合成绩, 以评价培训效果, 为进一步优化全科医学师资培训提供依据[25]。广西中医药大学则对壮医学教师进行了问卷调查, 调查实践教学师资队伍的现状及需求, 指出壮医学专业实践教学中的问题, 包括教学地位未受到足够重视、师资培养体系不健全、激励政策落实不到位以及教学管理制度缺乏执行力。这些问题对强化专业实践教学师资队伍的整体素养水平、提高壮医学专业实践教学效果具有重要意义[26]。

总体而言, 国外的师资评价模型及工具应用范围较广, 而国内的师资评价尚未达到成熟的模型阶段, 且多局限于个别住培医院内部。相关的师资评价往往仅限于简单的成绩分析或 360 度评价法[27] [28]。

5. 讨论

SETQ 量表最初由荷兰阿姆斯特丹大学学术医学中心开发, 旨在评估麻醉学专业教师的教学质量, 特别是加强临床教师在教学中的榜样作用。SETQ 通过个性化反馈系统对教学质量进行明确定义, 推动了教学质量的提升。自其推出以来, SETQ 在国际临床教学领域的应用取得了显著成果。该量表已被翻译成多种语言, 包括波斯语和英语, 并在荷兰、美国、伊朗等国家的不同医院和医学教学环境中得到广泛应用和验证。研究表明, SETQ 具有较高的信度和效度, 能够准确、客观地评估临床教师的教学表现, 这使其成为广泛适用的评估工具。

虽然 SETQ 最初是为麻醉学项目定制的, 但其广泛的适用性使其经过轻微修改后, 也能用于评估其他学科的教学人员, 如外科医生和非外科医生[21]。其广泛适用性可能源于各国医学培养体系的相似性。无论是在荷兰、美国, 还是在伊朗, 实习医生或住院医师都在临床教师的指导下参与实际的诊疗工作, 逐步积累经验和技能。因此, 尽管这些学员在不同的医疗体系中具有不同的身份称谓, SETQ 的基础设计使其能够在多个国家和背景下保持高水平的适用性和测量一致性。

尽管 SETQ 已展现出广泛的适用性, 研究人员在跨国或跨机构应用 SETQ 量表时仍需谨慎。由于各个住院医师项目和医学机构的文化和结构差异, SETQ 可能无法直接适用于所有专业或机构。虽然已有研究证明了 SETQ 的心理测量学质量, 但效度验证应被视为一个持续的过程。随着时间的推移, 测量工具需要根据科学、实用和心理测量学的反馈进行更新, 以确保在不同环境和人群中继续保持其相关性。未来的验证工作还应包括定期访谈相关人员、更新 SETQ 条目, 以及将量表结果与预期的教学表现进行关联性分析。

然而, SETQ 量表在实际应用中也存在一些局限性。尽管调查是匿名进行的, 但在一些规模较小的部门, 完全匿名的报告可能难以实现。研究发现, 匿名评分的结果往往低于公开评分, 且自愿参与可能导

致选择性偏倚。由于学员可能倾向于评价他们喜欢或高度认可的导师,量表结果可能倾向于较高的评价。此外,问卷调查还面临低回复率、学员遗忘完成问卷或由非指定人员填写等问题。

此外,该量表的心理测量学性能还需考虑潜在的混杂因素,如学员的年级、教育背景和对学习医学的满意度。未来的研究需要通过更全面的研究设计和对混杂因素的考虑,进一步评估 SETQ 量表的可靠性与有效性,以确保其在不同医学教育环境中的持续适用性。

6. 总结

通过对 SETQ 量表的评估,能够有效识别临床教师在教学过程中存在的问题与不足,从而为教师提供针对性的改进建议,促进临床教学质量的提升。此外,SETQ 量表的应用还助力于完善临床教师的评价体系,为医学教育的发展提供强有力的支持。

目前,国内外在师资评价研究方面的主要区别在于,国外相关的师资评价量表研究相对全面且系统。而相比之下,国内在此领域的研究仍显薄弱,缺乏成熟的师资评价体系。因此,未来的研究应更为注重师资评价系统的实际应用与推广,以培养合格的住院医师,并提升临床教学工作的规范化水平。

现有的师资评价系统尚未明确教学能力提升的分阶段目标与实施路径,亟需进一步细化具体的实施步骤,以增强系统的实用性和可操作性。尽管 SETQ 量表在临床教学中的应用已取得一定成果,但仍存在诸多研究方向有待进一步探讨。未来的研究可以从以下几个方面展开:完善 SETQ 量表的编制:持续提高量表的信度和效度,以确保其在不同教育环境中的适用性;探究与其他评价工具的关联性:通过建立更加科学的评价体系,提升师资评价的综合性和准确性;扩大 SETQ 量表的应用场景:在更多临床教学环境中应用 SETQ,以验证其普适性和有效性;深入分析评估结果的反馈机制:通过对 SETQ 评估结果的有效反馈,提高临床教师的教学质量及整体医学教育水平。通过对上述研究方向的深入探索,可以进一步推动 SETQ 量表的发展与应用,为提升临床教学质量以及促进医学教育改革做出更大贡献。

基金项目

南方医科大学第三附属医院 2024 年度教育教学课题(编号: JXZ202401)。

参考文献

- [1] 张黎峰, 张婕, 毛雯靖, 等. 国内外住院医师规范化培训师资胜任力研究进展[J]. 中国社会医学杂志, 2023, 40(3): 277-280.
- [2] Lombarts, K.M., Heineman, M.J. and Arah, O.A. (2010) Good Clinical Teachers Likely to Be Specialist Role Models: Results from a Multicenter Cross-Sectional Survey. *PLOS ONE*, 5, e15202. <https://doi.org/10.1371/journal.pone.0015202>
- [3] Lombarts, M.J., Arah, O.A., Busch, O.R., et al. (2010) Using the SETQ System to Evaluate and Improve Teaching Qualities of Clinical Teachers. *Nederlands Tijdschrift Voor Geneeskunde*, 154, A1222.
- [4] Schuwirth, L.W. (2010) Evaluation of Students and Teachers. *Nederlands Tijdschrift Voor Geneeskunde*, 154, A1677.
- [5] Arah, O.A., Hoekstra, J.B.L., Bos, A.P. and Lombarts, K.M. (2011) New Tools for Systematic Evaluation of Teaching Qualities of Medical Faculty: Results of an Ongoing Multi-Center Survey. *PLOS ONE*, 6, e25983. <https://doi.org/10.1371/journal.pone.0025983>
- [6] van der Leeuw, R., Lombarts, K., Heineman, M.J. and Arah, O. (2011) Systematic Evaluation of the Teaching Qualities of Obstetrics and Gynecology Faculty: Reliability and Validity of the SETQ Tools. *PLOS ONE*, 6, e19142. <https://doi.org/10.1371/journal.pone.0019142>
- [7] Arah, O.A., Heineman, M.J. and Lombarts, K.M. (2012) Factors Influencing Residents' Evaluations of Clinical Faculty Member Teaching Qualities and Role Model Status. *Medical Education*, 46, 381-389. <https://doi.org/10.1111/j.1365-2923.2011.04176.x>
- [8] Boerebach, B.C.M., Arah, O.A., Busch, O.R.C. and Lombarts, K.M. (2012) Reliable and Valid Tools for Measuring Surgeons' Teaching Performance: Residents' vs. Self Evaluation. *Journal of Surgical Education*, 69, 511-520.

- <https://doi.org/10.1016/j.jsurg.2012.04.003>
- [9] Lases, S.S.L., Arah, O.A., Pierik, E.G.J.M.R., Heineman, E. and Lombarts, M.J. (2014) Residents' Engagement and Empathy Associated with Their Perception of Faculty's Teaching Performance. *World Journal of Surgery*, **38**, 2753-2760. <https://doi.org/10.1007/s00268-014-2687-8>
 - [10] Dickinson, K.J., Bass, B.L., Graviss, E.A., Nguyen, D.T. and Pei, K.Y. (2021) How Learning Preferences and Teaching Styles Influence Effectiveness of Surgical Educators. *The American Journal of Surgery*, **221**, 256-260. <https://doi.org/10.1016/j.amjsurg.2020.08.028>
 - [11] Lewis, J.M., Yared, K., Heidel, R.E., Kirkpatrick, B., Freeman, M.B., Daley, B.J., et al. (2021) Emotional Intelligence and Burnout Related to Resident-Assessed Faculty Teaching Scores. *Journal of Surgical Education*, **78**, e100-e111. <https://doi.org/10.1016/j.jsurg.2021.09.023>
 - [12] Zhou, N.J., Kamil, R.J., Hillel, A.T., Tan, M., Walsh, J., Russell, J.O., et al. (2021) The Role of Preoperative Briefing and Postoperative Debriefing in Surgical Education. *Journal of Surgical Education*, **78**, 1182-1188. <https://doi.org/10.1016/j.jsurg.2020.11.001>
 - [13] Lombarts, K.M., Bux, M.J.L. and Arah, O.A. (2009) Development of a System for the Evaluation of the Teaching Qualities of Anesthesiology Faculty. *Anesthesiology*, **111**, 709-716. <https://doi.org/10.1097/aln.0b013e3181b76516>
 - [14] Boerebach, B.C.M., Lombarts, K.M. and Arah, O.A. (2016) Confirmatory Factor Analysis of the System for Evaluation of Teaching Qualities (SETQ) in Graduate Medical Training. *Evaluation & the Health Professions*, **39**, 21-32. <https://doi.org/10.1177/0163278714552520>
 - [15] Lombarts, K.M., Ferguson, A., Hollmann, M.W., Mallng, B., Arah, O.A., Berger, M.M., et al. (2016) Redesign of the System for Evaluation of Teaching Qualities in Anesthesiology Residency Training (SETQ smart). *Anesthesiology*, **125**, 1056-1065. <https://doi.org/10.1097/aln.0000000000001341>
 - [16] Al Ansari, A., Strachan, K., Hashim, S. and Ootom, S. (2017) Analysis of Psychometric Properties of the Modified SETQ Tool in Undergraduate Medical Education. *BMC Medical Education*, **17**, Article No. 56. <https://doi.org/10.1186/s12909-017-0893-4>
 - [17] Boldaji, F.T., Amini, M. and Parvizi, M.M. (2022) Psychometric Properties of the Persian Version of System for Evaluation of Teaching Qualities by Students: A Tool for Assessing Clinical Tutors from Students' Viewpoint. *Journal of Education and Health Promotion*, **11**, Article No. 92. https://doi.org/10.4103/jehp.jehp_1622_20
 - [18] Breckwoldt, J., Marty, A.P., Stricker, D., Stolz, R., Thomasin, R., Seeholzer, N., et al. (2022) Bottom-Up Feedback to Improve Clinical Teaching: Validation of the Swiss System for Evaluation of Teaching Qualities (SwissSETQ). *Swiss Medical Weekly*, **152**, w30137. <https://doi.org/10.4414/smww.2022.w30137>
 - [19] van Engelen, T.S., van der Leeuw, R.M. and Lombarts, M.J. (2016) Supervisors Being Assessed. *Nederlands Tijdschrift Voor Geneeskunde*, **160**, D419.
 - [20] Van Der Leeuw, R.M., Boerebach, B.C.M., Lombarts, K.M., Heineman, M.J. and Arah, O.A. (2016) Clinical Teaching Performance Improvement of Faculty in Residency Training: A Prospective Cohort Study. *Medical Teacher*, **38**, 464-470. <https://doi.org/10.3109/0142159x.2015.1060302>
 - [21] Yong, R.L., Cheung, W., Shrivastava, R.K. and Bederson, J.B. (2022) Teaching Quality in Neurosurgery: Quantitating Outcomes over Time. *Journal of Neurosurgery*, **136**, 1147-1156. <https://doi.org/10.3171/2021.2.jns203900>
 - [22] Lombarts, K.M., Heineman, M.J., Scherpier, A.J.J.A. and Arah, O.A. (2014) Effect of the Learning Climate of Residency Programs on Faculty's Teaching Performance as Evaluated by Residents. *PLOS ONE*, **9**, e86512. <https://doi.org/10.1371/journal.pone.0086512>
 - [23] Scheepers, R.A., Lombarts, K.M., van Aken, M.A.G., Heineman, M.J. and Arah, O.A. (2014) Personality Traits Affect Teaching Performance of Attending Physicians: Results of a Multi-Center Observational Study. *PLOS ONE*, **9**, e98107. <https://doi.org/10.1371/journal.pone.0098107>
 - [24] 冯玉波, 朱滨海, 占伊扬, 等. 基于柯式模型的住院医师规范化培训医学模拟师资培训效果评价[J]. 中国毕业后医学教育, 2023, 7(6): 438-441+446.
 - [25] 邢桂枚, 严汪龙, 阮三平, 等. 安徽省全科医学师资培训的效果评价[J]. 安徽医学专报, 2023, 22(1): 5-7.
 - [26] 张云, 刘儒鹏, 韦英才, 等. 壮医学专业实践教学师资队伍现状分析与对策研究[J]. 中医药管理杂志, 2022, 30(20): 101-103.
 - [27] 赵稳稳, 王荣英, 张金佳, 等. 全科医学师资质量评价体系建设研究进展[J]. 中国全科医学, 2024, 27(7): 784-788+800.
 - [28] Bezerra, N.C., Lombarts, K.M., Naiz, S.C.d.R., Castanhel, F.D. and Grosseman, S. (2024) System for Evaluation of Teaching Qualities (SETQ) Smart Para O Uso No Brasil: Versão Para Residente. *Revista Brasileira de Educação Médica*, **48**, e076. <https://doi.org/10.1590/1981-5271v48.3-2023-0217>