Al-Rafidain J Med Sci. 2025;9(2):1-15.
DOI: https://doi.org/10.54133/ajms.v9i2.2310

A J M S
Online ISSN (2789-3219)

Review Article

Mapping the Landscape: A Bibliometric and Narrative Review of Programmatic Assessment in Medical Education

N Juni Triastuti[1,2]*, Tri Nur Kristina[3], Titi Savitri Prihatiningsih[4]
[1]*Department of Medical Education, Faculty of Medicine, Universitas Muhammadiyah Surakarta, Indonesia;* [2]*Doctoral Program, Faculty of Medicine, Universitas Diponegoro, Semarang, Indonesia;* [3]*Department of Medical Education, Faculty of Medicine, Universitas Diponegoro, Semarang, Indonesia;* [4]*Department of Medical Education, Faculty of Medicine, Universitas Gadjah Mada, Yogyakarta, Indonesia*
Received: 10 July 2025; Revised: 25 August 2025; Accepted: 29 August 2025

Abstract

Background: Programmatic assessment is an evaluative framework designed to support learner development through continuous, multi-source, and integrated data collection, enabling more informed and robust summative decision-making. In medical education, it offers a holistic approach to assessing competencies. However, implementation practices and outcomes vary across institutions, highlighting the need for a comprehensive understanding of this approach. ***Objective***: This study aims to synthesize the current literature on programmatic assessment in medical education by conducting a combined review and bibliometric analysis. The focus is on its implementation, core components, challenges, and educational impacts. ***Methods***: This review was conducted in a Scopus, database for studies published between 2005 and 2025. Articles were included if they discussed programmatic assessment in undergraduate or postgraduate medical education. Thematic analysis was applied to identify recurring concepts. Additionally, from 1781 articles, a bibliometric analysis of 256 publications from Scopus (2005–2025) was performed to examine publication trends, key contributors, institutions, and funding sources. ***Results***: Of 1781 articles screened, eight studies met the inclusion criteria. Key components of programmatic assessment included the use of multiple formative tools, structured feedback, mentorship, and panel-based summative decisions. Implementation challenges involved faculty resistance, administrative burdens, and limited facilitator training. Reported benefits included enhanced student engagement and positive learning perceptions. Bibliometric analysis indicated increasing research interest, peaking in 2022. ***Conclusions***: Programmatic assessment holds considerable promises for enhancing assessment quality in medical education, though sustained institutional support and further research are essential.

Keywords: Bibliometric analysis, Competency-based assessment, Formative feedback, Medical education, Programmatic assessment.

رسم خريطة المشهد: مراجعة ببليومترية وسردية للتقييم البرنامجي في التعليم الطبي

الخلاصة

الخلفية: التقييم البرنامجي هو إطار تقييم مصمم لدعم تطوير المتعلم من خلال جمع البيانات المستمر والمتعدد المصادر والمتكامل، مما يتيح اتخاذ قرارات تلخيصية أكثر استنارة وقوة. في التعليم الطبي، يقدم نهجا شاملا لتقييم الكفاءات. ومع ذلك، تختلف ممارسات التنفيذ ونتائجه من مؤسسة إلى أخرى، مما يبرز الحاجة إلى فهم شامل لهذا النهج. **الهدف**: تهدف هذه الدراسة إلى تجميع الأدبيات الحالية حول التقييم البرنامجي في التعليم الطبي من خلال إجراء مراجعة مشتركة وتحليل ببليومتري. ينصب التركيز على تنفيذه ومكوناته الأساسية والتحديات والآثار التعليمية. **الطرائق**: أجريت هذه المراجعة في قاعدة بيانات Scopus للدراسات المنشورة بين عامي 2005 و 2025. تم تضمين المقالات إذا ناقشت التقييم البرنامجي في التعليم الطبي الجامعي أو الدراسات العليا. تم تطبيق التحليل المواضيعي لتحديد المفاهيم المتكررة. بالإضافة إلى ذلك، من بين 1781 مقالة، تم إجراء تحليل ببليومتري ل 256 منشورا من Scopus (2005-2025) لفحص اتجاهات النشر والمساهمين الرئيسيين والمؤسسات ومصادر التمويل. **النتائج**: من بين 1781 مقالة تم فحصها، استوفت ثماني دراسات معايير الاشتمال. تضمنت المكونات الرئيسية للتقييم البرنامجي استخدام أدوات تكوينية متعددة، والتغذية الراجعة المنظمة، والإرشاد، والقرارات الختامية القائمة على اللوحات. تضمنت تحديات التنفيذ مقاومة أعضاء هيئة التدريس والأعباء الإدارية وتدريب الميسرين المحدود. تضمنت الفوائد المبلغ عنها تعزيز مشاركة الطلاب وتصورات التعلم الإيجابية. أشار التحليل الببليومتري إلى زيادة الاهتمام البحثي، وبلغ ذروته في عام 2022. ***الاستنتاجات***: يحمل التقييم البرنامجي وعودا كبيرة لتعزيز جودة التقييم في التعليم الطبي، على الرغم من أن الدعم المؤسسي المستمر والمزيد من البحث ضروريان.

* ***Corresponding author***: N J. Triastuti, Department of Medical Education, Faculty of Medicine, Universitas Muhammadiyah Surakarta, Indonesia; Email: njt140@ums.ac.id
Article citation: Triastuti NJ, Kristina TN, Prihatiningsih TS. Mapping the Landscape: A Bibliometric and Narrative Review of Programmatic Assessment in Medical Education. *Al-Rafidain J Med Sci*. 2025;9(2):1-15. doi: https://doi.org/10.54133/ajms.v9i2.2310

INTRODUCTION

Over the past two decades, medical education has witnessed a significant paradigm shift in assessment practices, moving from isolated high-stakes examinations toward more continuous, holistic, and integrated approaches [1,2]. One such approach, known as *programmatic assessment*, has gained prominence for its ability to collect multiple low-stakes data points over time to inform high-stakes decisions and support learner development [3,4]. Programmatic assessment integrates feedback, mentorship, and decision-making in a cohesive framework that prioritizes both formative and summative purposes [5,6]. The increasing recognition

of its potential to enhance educational outcomes has led to a growing body of literature examining its theoretical underpinnings, implementation challenges, and impact on student learning [7,8]. Programmatic assessment has emerged as a comprehensive approach to evaluating learner competencies in medical education, integrating multiple assessment methods over time to provide a holistic view of student performance and development [9,10]. Unlike traditional assessments that often rely on isolated, high-stakes exams, programmatic assessment emphasizes continuous data collection, learner feedback, and reflective practice, thereby supporting formative learning and summative decision-making processes [11,12]. This approach aligns with the growing recognition of competency-based medical education, which prioritizes the demonstration of clinical skills, professional behavior, and cognitive abilities in real-world settings [13,14]. Several studies have demonstrated the feasibility and effectiveness of programmatic assessment across different stages of medical training, including undergraduate internships, residency programs, and specialized clinical training [10,12,13]. Furthermore, learner perspectives indicate that balancing structured assessment frameworks with learner autonomy and agency is crucial for fostering engagement and motivation [9,11]. Simulation-based assessments and tailored feedback mechanisms have been identified as valuable adjuncts that enhance skill acquisition and promote reflective learning [15,16]. Despite these advances, challenges persist in standardizing programmatic assessment practices and ensuring equitable evaluation across diverse learner profiles, particularly in addressing disparities related to engagement and performance [14]. Additionally, the development and validation of context-specific assessment instruments remain essential for optimizing feedback quality and educational outcomes [16]. Assessment in medical education has evolved significantly over recent years, transitioning from traditional, episodic examinations to more comprehensive, longitudinal approaches that emphasize continuous learner development and competency-based outcomes [17]. Programmatic assessment (PA) represents one such innovative framework designed to integrate diverse assessment methods systematically over time, providing rich data to inform both summative decisions and formative feedback [18]. This approach aligns closely with the principles of competency-based medical education (CBME), promoting a holistic understanding of learners' knowledge, skills, and professional behaviors in authentic clinical contexts [19]. The rising interest in programmatic assessment is reflected in an expanding body of literature that explores its theoretical underpinnings, implementation strategies, and educational impact [20]. Despite its demonstrated potential, the practical adoption of PA poses challenges, including resource demands, faculty readiness, and the need for robust data management systems [21,22]. Moreover, geographic disparities in research output suggest that low- and middle-income countries may face additional barriers to implementation, underscoring the importance of context-sensitive adaptations [19]. Recent advances in digital technology have further catalyzed the development of programmatic assessment by enabling efficient data aggregation, learning analytics, and real-time feedback mechanisms [23]. These technological innovations promise to enhance the scalability and effectiveness of PA but require rigorous evaluation to address concerns related to usability, data security, and educational validity [24]. Despite this increasing interest, the literature on programmatic assessment remains dispersed across multiple disciplines, with limited synthesis of prevailing trends and patterns [25,26]. Bibliometric analysis offers a powerful method for systematically mapping the structure and evolution of scientific knowledge within a field [27,28]. By analyzing citation patterns, co-authorship networks, and thematic clusters, bibliometrics can uncover influential publications, emerging topics, and collaboration dynamics [29,30]. In contrast, narrative reviews allow for a deeper, interpretive understanding of complex educational phenomena by contextualizing findings and integrating diverse perspectives [31,32]. Together, these two methodologies can provide a comprehensive overview of the development and trajectory of programmatic assessment in medical education. A number of conceptual frameworks have guided the evolution of programmatic assessment, including constructivist learning theory, competency-based education, and assessment for learning [33,34]. Central to this approach is the notion of assessment as a continuous learning process rather than a one-time event [35,36]. The introduction of *assessment for learning* and *feedback literacy* into medical education has shifted the focus toward how students engage with and respond to assessment information [37,38]. Consequently, programmatic assessment not only evaluates learners but also supports their development through meaningful feedback and mentoring [39,40]. Although numerous studies have examined specific components of programmatic assessment, there remains a need for a comprehensive synthesis of the literature to map current trends, identify gaps, and guide future research. Reviews combined with bibliometric analyses provide valuable tools for achieving these aims by systematically collating and quantitatively analyzing the research landscape [18,41]. Nevertheless, the implementation of programmatic assessment is not without challenges. Issues such as faculty resistance, data overload, consistency of judgments, and the logistics of managing longitudinal assessment data remain significant barriers [9,42]. Furthermore, the success of programmatic assessment depends heavily on institutional culture, leadership commitment, and faculty development [43,44]. Understanding how these challenges are addressed across various contexts requires a systematic and integrative examination of the literature [45,46]. Such analysis can guide educators and policymakers in refining assessment systems and maximizing the benefits for learners. This study therefore aims to conduct a bibliometric and

narrative review of the literature on programmatic assessment in medical education. The bibliometric component will map publication trends, influential authors, and thematic structures, while the narrative synthesis will explore conceptual themes, implementation practices, and contextual enablers and barriers. By combining these approaches, we aim to provide a comprehensive and evidence-informed overview that informs future research, policy, and practice in medical assessment [47,48]. Therefore, this study undertakes a review and bibliometric analysis to comprehensively examine the scope, nature, and evolution of programmatic assessment research in medical education. By doing so, it aims to inform educators, policymakers, and researchers about prevailing themes, regional contributions, and emerging directions, thereby facilitating the strategic advancement of programmatic assessment globally.

METHODS

This study employed a combined review and bibliometric analysis methodology to comprehensively map and quantify the research landscape on programmatic assessment in medical education. The review approach was chosen to systematically identify and characterize the breadth and depth of relevant literature, while bibliometric analysis facilitated the quantitative examination of publication trends, citation patterns, and research collaboration networks [49]. The bibliometric component followed established protocols for science mapping, using Scopus as the primary data source due to the comprehensive indexing of medical education journals [27,50]. A structured search strategy was conducted using the following search terms: “programmatic assessment” AND “medical education”. Filters were applied to include only peer-reviewed articles published between January 2005 and March 2025 in English, focusing on undergraduate and postgraduate medical education contexts [51,52]. Bibliometric data such as author keywords, citations, co-authorship networks, and journal sources were extracted and analysed using *VOS viewer* [28,30]. The narrative component was guided by principles of integrative review methodology to identify and synthesize recurring concepts, challenges, and innovations in implementing programmatic assessment (53, 54). Eligible articles for narrative synthesis were selected based on relevance to three core domains: 1) theoretical foundations, 2) practical implementation, and 3) institutional and learner-related factors. Full texts were independently reviewed by two authors, and data were extracted and coded using thematic synthesis procedures [55, 6]. Disagreements were resolved through discussion until consensus was achieved. The narrative synthesis focused on key constructs such as feedback practices, mentoring roles, data integration, and learner engagement, drawing from both qualitative and mixed-methods studies [57-59]. To ensure methodological transparency and reproducibility, this review followed the PRISMA-ScR guidelines [47,60]. Screening and eligibility decisions were documented using a PRISMA flow diagram. Quality assessment was not conducted for the bibliometric data due to the descriptive nature of the method [61]. For narrative synthesis, article inclusion was based on conceptual richness and relevance rather than methodological rigor, in line with scoping review principles [62,63]. The integration of bibliometric mapping with thematic interpretation provides a holistic understanding of the scholarly evolution and implementation dynamics of programmatic assessment in medical education [64,65].

Literature Search Strategy

A comprehensive search was conducted on Scopus that ensured coverage of both medical education and assessment research domains. The search strategy combined controlled vocabulary and free-text terms related to “programmatic assessment" and “medical education." Searches were limited to articles published between January 2005 and March 2025 to capture the most recent evidence and developments [18,19].

Inclusion and exclusion criteria

Studies were included if they 1) focused on programmatic assessment in undergraduate or postgraduate medical education, 2) reported empirical data, conceptual frameworks, or reviews relevant to PA implementation, challenges, or outcomes, and 3) were published in English. Excluded were articles addressing unrelated assessment types, non-medical education contexts, editorials without original data, and conference abstracts [17,24].

Data Extraction and Charting

Two independent reviewers extracted data using a standardized form capturing publication year, country of origin, study design, assessment modalities, outcomes measured, and key findings. Discrepancies were resolved through discussion or consultation with a third reviewer to ensure reliability [21,66].

Bibliometric Data Collection

Bibliometric data, including authorship, citation counts, journal impact factors, and institutional affiliations, were exported into VOS viewer for visualization and network analysis. Key metrics analyzed included annual publication trends, prolific authors and institutions, keyword co-occurrence, and international collaboration patterns [18,23]. These analyses provided insights into research productivity, influential contributors, and thematic emphasis in programmatic assessment literature [22,41].

Quality Appraisal

While the reviews traditionally do not emphasize quality appraisal, this study incorporated a critical evaluation of methodological rigor in included empirical studies using established tools such as the Mixed Methods Appraisal Tool (MMAT) to

contextualize findings and identify evidence gaps [17,19].

Ethical Considerations

As this study utilized publicly available literature, formal ethical approval was not required. The review adhered to PRISMA-ScR guidelines to ensure transparency and reproducibility [67]. This review was conducted based on the methodological framework developed by Arksey and O'Malley [68], further refined by Levac [69] and the Joanna Briggs Institute (JBI) [70]. Purpose and Research Question. This bibliometric study aims to systematically chart the existing body of literature on programmatic assessment within medical education, identify key thematic trends, and uncover areas that remain underexplored. The central research question guiding this investigation is, "How is programmatic assessment implemented in medical education, and what are its core components, models, defining characteristics, and methods of evaluation?"

Study Selection Process

The selection of studies followed a two-phase screening protocol. The initial phase involved evaluating titles and abstracts to determine relevance. In the second phase, full-text articles were retrieved and assessed in detail against the inclusion and exclusion criteria to ensure methodological rigor and topical alignment (Figure 1).

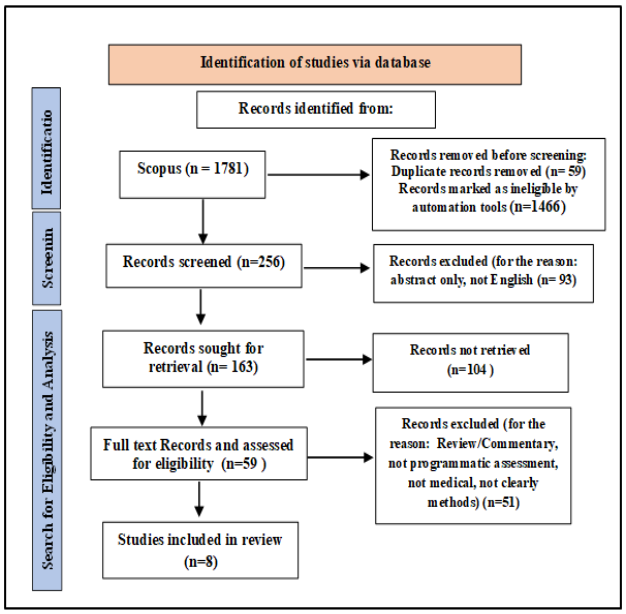

Figure 1: Study Selection process.

Bibliometric Analysis

A bibliometric analysis was conducted to quantitatively assess the research landscape, trends, and intellectual structure of literature related to programmatic assessment in medical education. This component complemented the review by providing an objective overview of publication productivity, citation patterns, collaboration networks, and thematic developments in the field. The search included publications from 2005 (the approximate emergence of the concept in medical education) to 2025 and was limited to English-language and peer-reviewed journal articles. Editorials, reviews, commentaries, letters, and non-peer-reviewed literature were excluded. The subjects were limited to medicine and health professions and article document types from the sources of journals and conference proceedings and in the English version. All search results were exported in CSV format for further processing. Data extraction included extracted data, study objectives, educational setting, components of programmatic assessment, assessment methods, roles of facilitators/mentors, and key findings related to effectiveness or implementation challenges. The data were analysed descriptively and presented in narrative form as well as in tables. The reporting of results followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Reviews) guidelines [67].

RESULTS

From 1781 articles, a total of 256 were screened, eight studies met the inclusion criteria and were analyzed in this review. These studies were published between 2005 and 2025 and originated from various medical education systems in the Netherlands, Canada, Australia, the United Kingdom, and several Asian countries. The majority of the research was conducted within the context of undergraduate medical education, with some studies derived from clinical or postgraduate education settings. The total participants in this review were 934 participants. The review identified a total of 256 articles and eight studies that met the inclusion criteria, spanning publications from 2005 to 2025. The majority of the literature was published in the last decade, reflecting a growing interest in programmatic assessment within medical education. The bibliometric analysis included 256 publications retrieved from the Scopus database. Annual publication volume increased steadily from 2005, with a marked acceleration after 2015. The highest number of publications was recorded in 2022, indicating sustained research interest. Table 1 shows a summary of the studies analyzed in the study. The chart shows publication trends from 2005 to 2025, with a clear growth pattern peaking in 2022 (55 publications) and 2021 (42 publications). There's been steady growth since 2014, with a slight decline in recent years (Figure 2). Analysis revealed 300 distinct authors, with the top five authors collectively contributing to 22% of the publications. Co-authorship networks indicated strong collaboration clusters primarily within North America and Europe, with emerging contributions from Asia and Oceania.

The data shows papers with an average of 5 authors with a median of 4. Collaboration has remained relatively stable over time, with one outlier paper having 37 authors. Most papers have 2-8 authors (Figure 3).

Table 1: Summary of studies included in the review

Author (Year)	Country	Study design	Participants	Focus on programmatic assessment/objective	Research findings
Hauf *et al*. (2014) [13]	America	Descriptive analysis	28 internship students	Identifying the feasibility of programmatic assessment during the first year of internship in emergency medicine	Multi-source assessment methods employed during internships effectively identified first-year internship competencies, with attainment levels between 70% and 80%; however, aseptic technique was the area with the poorest performance.
Roberts *et al*. (2022) [11]	Australia	Qualitative	112 graduate entry students	To explore students' perspectives of programmatic assessment	Successful implementation of programmatic assessment necessitates balancing learner autonomy, assessment structures, and institutional culture, supported by mechanisms such as dedicated learning advisors and regular portfolio advisory group meetings to facilitate learner guidance and promote reflective practice.
Schut et al. (2018) [9]	Netherlands	Qualitative	26 students	To explore students' perspectives of assessment stakes using programmatic assessment	Programmatic assessment impacts both assessment outcomes and the effective use of learning data to improve student learning, with faculty–student relationships fostering student autonomy to ensure learners feel empowered and secure in directing their own educational progress through assessment.
Wilkinson *et al*. (2011) [10]	New Zealand	Descriptive	701 students	To describe and evaluate a programmatic system to handle student assessment results	The programmatic assessment detected more students in difficulty rather than the traditional one particularly problem with professionalism.
Barbagallo *et al*., (2024) [12]	Australia	Mixed methods	radiation oncology medical physics registrars (trainees)	To update and enhance the training program for radiation oncology medical physicists in Australasia by developing a structured, programmatic assessment model that ensures standardized, consistent, and ongoing evaluation of trainees' competencies, while also incorporating emerging technologies and promoting a dynamic curriculum framework	This study presents a comprehensive update of the radiation oncology medical physics training program in Australasia conducted by the ACPSEM, describing a four-phase approach encompassing program review, stakeholder engagement, curriculum development, and the design of a structured programmatic assessment model.
Rich *et al*, (2022) [14]	Canada	Qualitative	17 medical residency academic advisor	To explore how resident archetypes characterized by engagement and performance influence the functioning and processes of programs of assessment across multiple medical residency programs	The study revealed that resident engagement and performance substantially affect assessment processes, with less effort required for highly engaged, high-performing residents, whereas more time is devoted to interpreting limited or problematic data from disengaged, lower-performing residents. These variations can produce adaptive or maladaptive consequences for program effectiveness, underscoring the importance of implementing strategies to mitigate disparities and ensure equity in assessments.
Adam *et al*, 2022 [15]	South Africa	Qualitative	20 obstetrics and gynecology rotation of sixth year medical students	To investigate whether the use of medium-fidelity obstetric simulation scenarios fosters cooperative learning among undergraduate medical students,	The study demonstrated that simulation serves as a safe and effective complement to experiential learning by enabling students to identify knowledge gaps, enhance clinical skills, and promote collaboration and reflective practice. Participants reported that simulation improved their ability to integrate theoretical knowledge with practical application, boosted their confidence, and facilitated the development of essential graduate attributes for managing obstetric emergencies.
Ainin *et al*. (2023) [16]	Indonesia	Quantitative	30 pre-clinical undergraduate medical students in their final year	This study aimed to develop a valid and reliable instrument to evaluate the implementation of Programmatic Assessment (PA) from the perspective of undergraduate medical students in Indonesia. Grounded in psychometric principles, the instrument comprised 41 Likert-scale items and 3 yes/no questions and was validated by experts to support constructive feedback and foster dialogue between students and educators. The results indicate the necessity for further validation across diverse educational settings and highlight the critical role of continuous feedback and reflective practice in student-centered learning.	The results showed that students rated the overall questionnaire items positively, with an average score of 3.34 out of 4; however, they expressed dissatisfaction with the supportive activities component, underscoring the need to enhance continuous feedback and reflective opportunities within the assessment system.

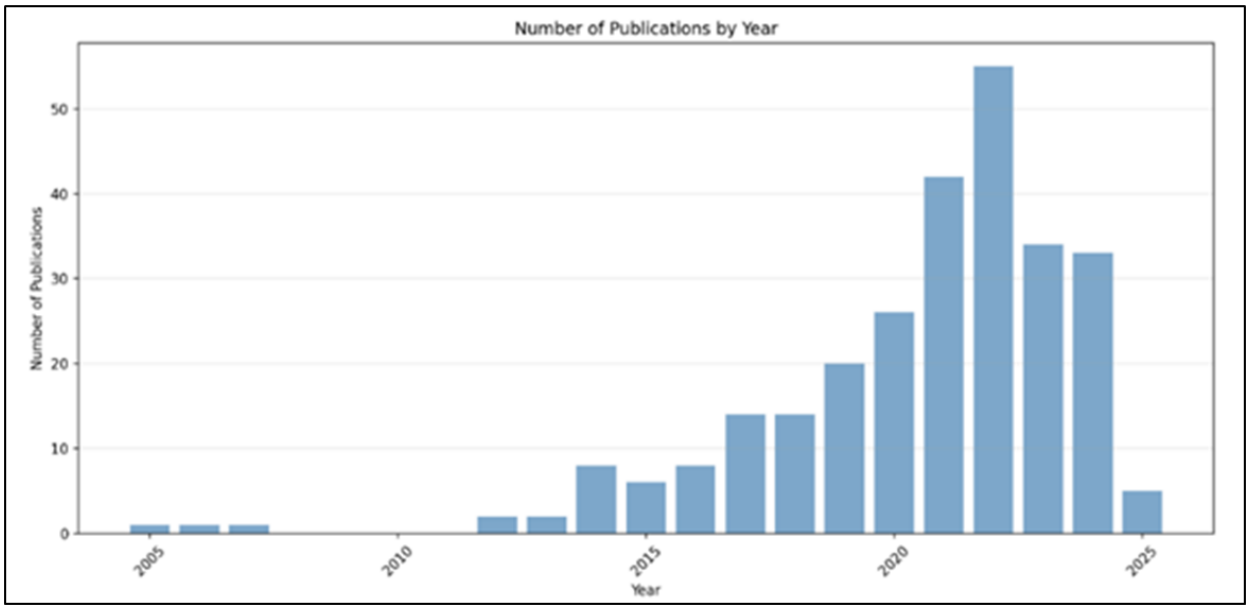

Figure 2: Number of publications by year.

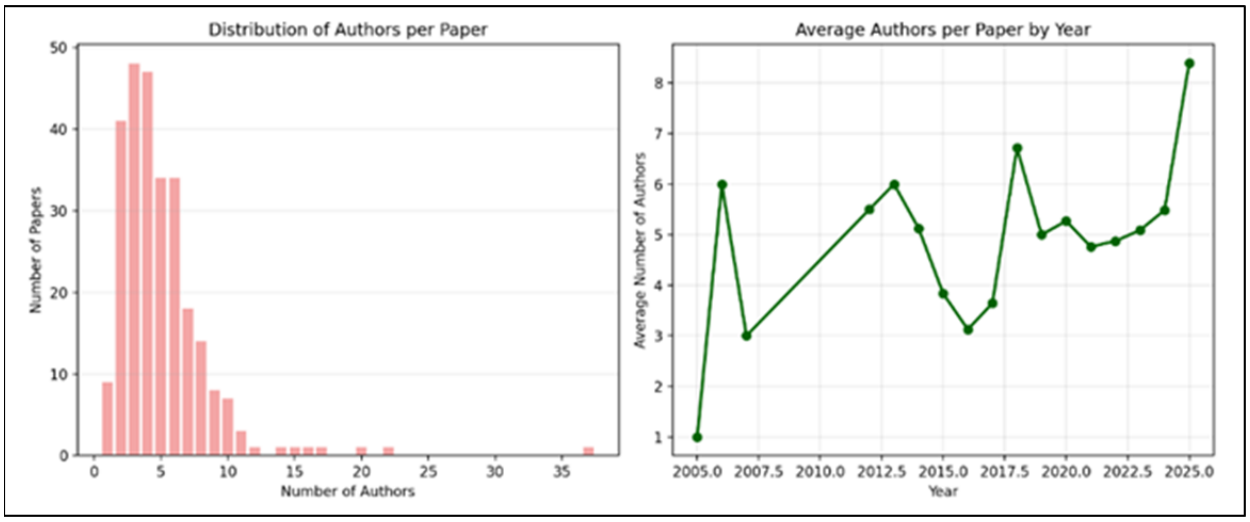

Figure 3: Distribution of authors per paper.

The top 10 most prolific authors were Cheung W.J.: 7 papers; Palermo C.: 6 papers; Santen S.A.: 5 papers; Ross S.: 4 papers; van der Vleuten C.: 4 papers; Chan T.M.: 4 papers; Vaughan B.: 4 papers; Moore K.: 4 papers; Marty A.P.: 3 papers; and Heeneman S.: 3 papers. The international collaboration rate was 81.6%. The analysis shows papers average 5 authors with high international collaboration (81.6%). Cheung W.J. leads with 7 publications. Collaboration has remained relatively stable over time. From 300 authors retrieved from Scopus databases, 12 meet the threshold of the minimum 3 documents per author (Figure 4). Minimum occurrence keywords: 3 keywords from 226 keywords, 37 meet the threshold. In the programmatic assessment item, there were 30 links with a total link strength of 71 and an occurrence of 21. Top 10 countries by publication count include United States: 253 publications; Canada: 166 publications; Australia: 134 publications; Netherlands: 113 publications; United Kingdom: 45 publications; India: 31 publications; Malaysia: 28 publications; Switzerland: 26 publications; Iran: 20 publications; South Africa: 20 publications. The average multi-institutional collaboration rate was 75.2%. The US leads with 253 publications, followed by Canada (166) and Australia (134). Multi-institutional collaboration averages 75.2% across years (Figure 5). Co-citation mapping highlighted seminal works by Van der Vleuten and colleagues as foundational. Keyword co-occurrence analysis clustered around themes such as "feedback," "competency-based education," "formative assessment," and "faculty development," indicating core research foci. From 226 keywords, 23 meet the threshold of the minimum 5 keywords. Mean citations per paper was 9.5; median citations per paper was 4.0; the most cited paper had 556 citations; and papers with 0 citations numbered 46. The top 5 journals were Medical Science Educator: 18 papers; AEM Education and Training: 13 papers; Education Sciences: 10 papers; BMJ Open: 8 papers; and American Journal of Pharmaceutical Education: 7 papers. Medical Science Educator leads with 18 papers. The most cited paper has 556 citations (rheumatology treatment guidelines). The average

citations per paper is 9.5, with 46 papers having zero citations (Figure 6).

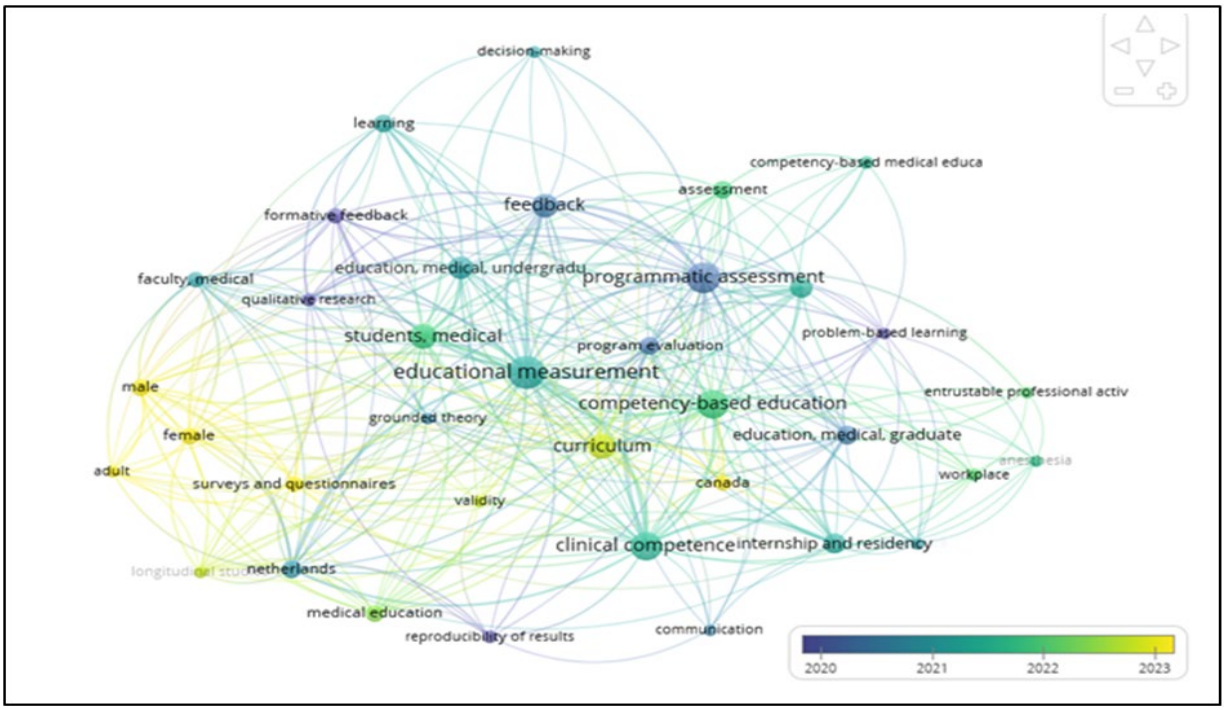

Figure 4: Network visualization of authors.

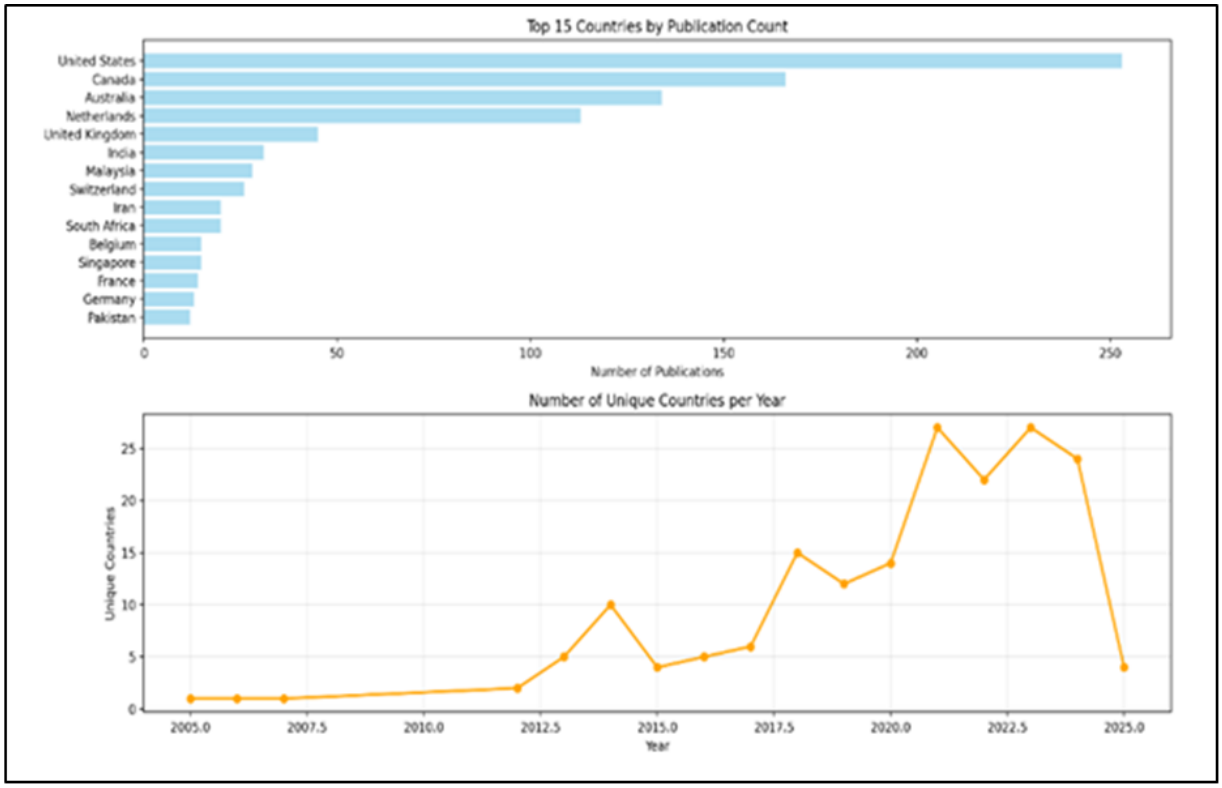

Figure 5: Geographical distribution of paper.

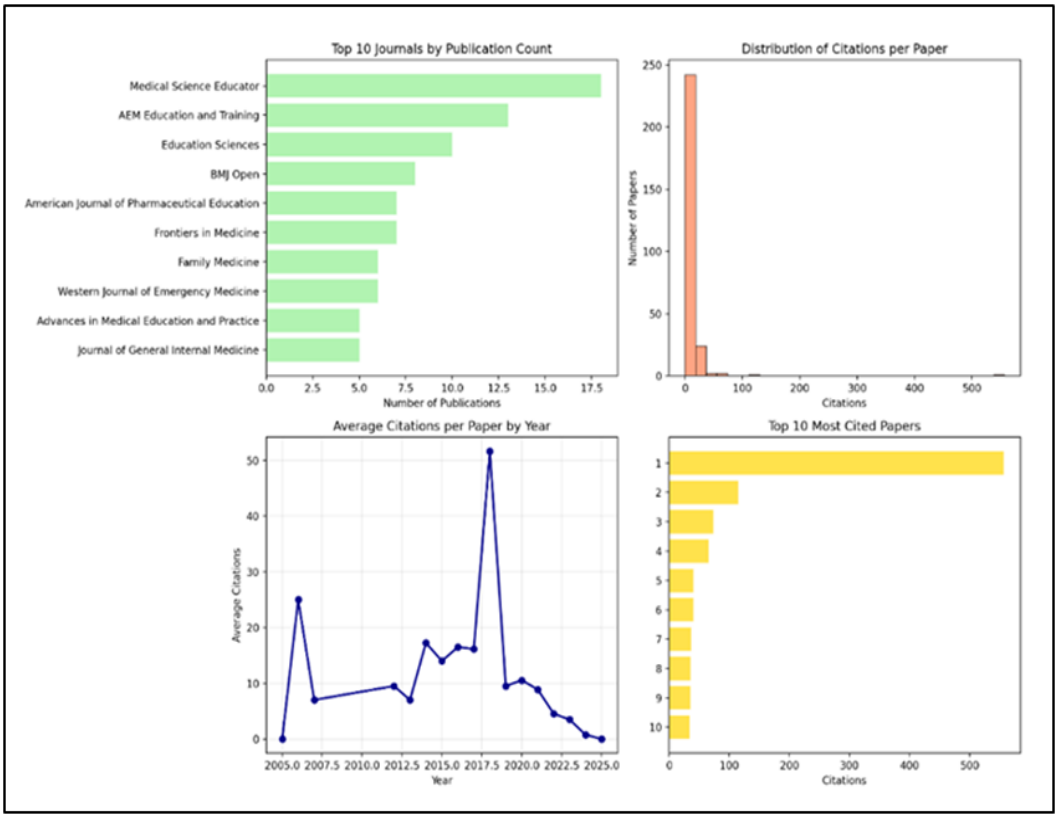

Figure 6: Distribution of citations.

Recent keyword trends suggested growing attention to digital assessment tools, learner autonomy, and program evaluation methodologies. The top 10 emerging themes in medical education were assessment (201 mentions), competency (75 mentions), evaluation (61 mentions), curriculum (58 mentions), communication (20 mentions), simulation (15 mentions), clinical skills (11 mentions), professionalism (11 mentions), interprofessional (10 mentions), and stress (10 mentions). The most frequent author keywords were assessment (50 papers), medical education (33 papers), programmatic assessment (20 papers), education (19 papers), and curriculum (12 papers). Assessment dominates as the primary theme (201 mentions), followed by competency-based education and curriculum evaluation. Technology adoption shows gradual growth, with simulation and digital learning gaining traction. Articles are predominantly research papers (78.4%) (Figure 7). The integration of and bibliometric analyses provides a comprehensive overview of the evolution of programmatic assessment research in medical education. The field has progressively matured from conceptual discussions to practical applications involving faculty roles and technological innovations. However, the bibliometric findings also underscore the need for more diverse geographical representation and robust empirical studies to validate theoretical models (Figure 8).

DISCUSSION

Programmatic assessment (PA) has emerged as a transformative approach in medical education, offering a more comprehensive and developmental method for evaluating learners. One of its most recognized strengths is the ability to gather multiple low-stakes assessments over time, leading to a more valid and reliable basis for high-stakes decisions [71,72]. This model emphasizes assessment for learning, encouraging continuous feedback, learner reflection, and professional identity formation [37,73,74]. Additionally, the integration of feedback and mentorship within PA enhances students' self-regulated learning and feedback literacy [57,75]. Numerous studies have confirmed that when well-implemented, PA supports longitudinal competency development and aligns effectively with competency-based medical education (CBME) frameworks [65,76-78]. Furthermore, PA systems promote shared decision-making and collaborative assessment cultures, fostering more educationally meaningful judgments [58,79]. The reviewed studies encompassed a variety of research designs, including descriptive, evaluative, qualitative, quantitative, and conceptual narrative reviews [2,80]. A key finding across the literature is that programmatic assessment is not a single assessment tool but a comprehensive, longitudinal framework made up of multiple components [80]. Consistently, the studies underscore the use of diverse assessment instruments—such as the Mini-Clinical Evaluation Exercise (Mini-CEX), Direct Observation of Procedural Skills (DOPS), electronic portfolios, and logbooks—to evaluate learners' progress [81,82].

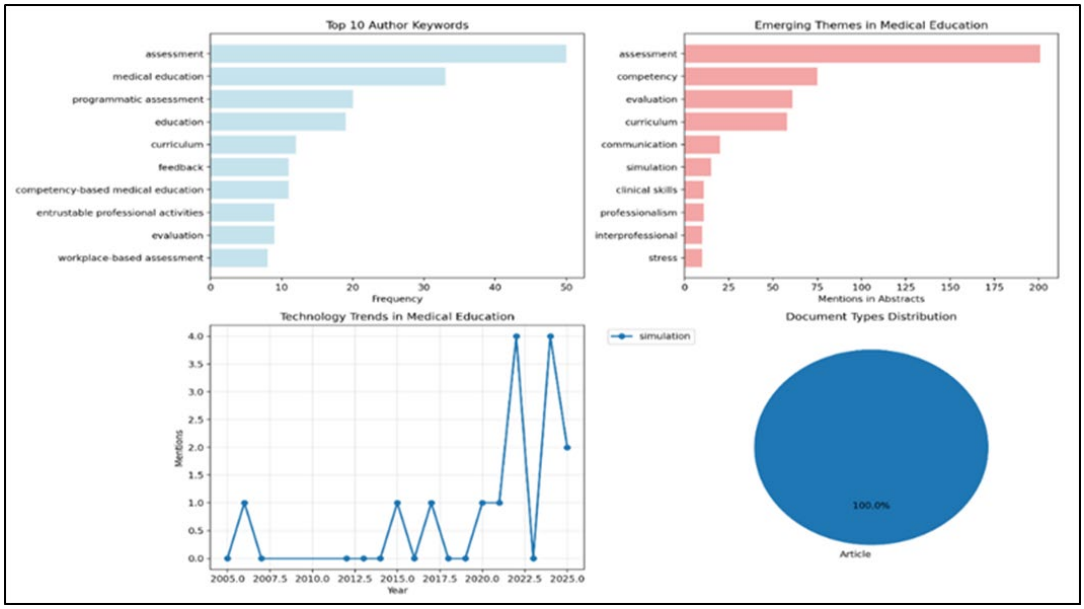

Figure 7: Distribution of emerging themes.

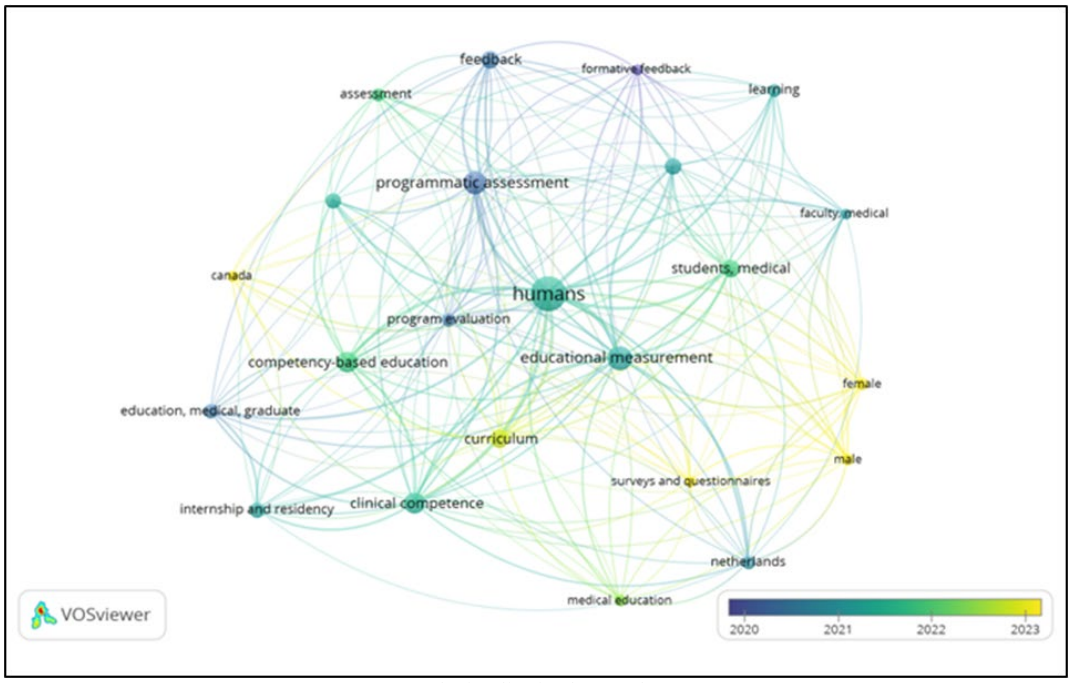

Figure 8: Network visualization keywords co-occurrence.

Assessments are administered repeatedly over time to capture the development of student competencies, moving away from reliance on isolated summative evaluations. Instead, the framework emphasizes the aggregation of formative data collected longitudinally, which forms the foundation for detailed and targeted feedback [2,83]. Feedback, predominantly provided by faculty or mentors, is central to fostering student self-reflection and guiding performance enhancement, with research highlighting that the quality of feedback is more critical than the quantity [80]. Moreover, most studies stress the importance of mentorship or academic coaching in helping learners interpret their assessment results, identify strengths and areas for improvement, and formulate individualized learning plans [80]. High-stakes decisions within this framework are made not on singular data points but rather through a holistic review of multiple sources of assessment data by an expert panel, thereby improving the reliability and fairness of the evaluation process [80,84,85]. Among the eight studies assessed for quality, five were deemed methodologically sound and sufficiently relevant, thus qualifying for inclusion in the review. Among these, Schut *et al*. [9]. The quality assessment of the eight studies on programmatic assessment indicated that five met the established inclusion criteria, demonstrating adequate methodological rigor and relevance for inclusion in the analysis. Roberts *et al*. [11] and Hauf *et al*. [13] also scored highly, with totals of 9 and 8, respectively, further confirming their reliability and contribution to the evidence base. Studies published by Wilkinson *et al*. [10] and Ainin *et al*. [16] were deemed suitable for inclusion, with each attaining a score of 7, signifying

adequate quality notwithstanding some minor methodological shortcomings. Conversely, studies by Adam *et al*. [15], Rich *et al*. [14], and Barbagallo *et al*. [12] with scores of 4, 2, and 0, respectively; these studies fell below the eligibility threshold, reflecting significant methodological limitations that precluded their inclusion. The studies emphasized that programmatic assessment is primarily intended to encourage lifelong learning, enhance the quality of feedback, mitigate the pressure of high-stakes examinations, and ensure a robust evidentiary basis for summative decisions [80,82,86]. Findings indicate that programmatic assessment contributes to more favorable student perceptions and fosters increased learner engagement, but also presents challenges related to data handling and administrative demands [87]. Several challenges were identified in the implementation of programmatic assessment. Many institutions lack the infrastructure required to integrate and manage multiple sources of assessment data effectively. Delivering high-quality feedback demands both time and specific competencies, which are not always readily available. Furthermore, insufficient training for faculty and students regarding the principles and practical application of programmatic assessment hampers effective adoption. Resistance to change is also common, as both educators and learners may be more accustomed to traditional assessment approaches [5,87]. A review of eight studies from diverse international contexts demonstrates the growing adoption of programmatic assessment (PA) in medical education, driven by its potential to promote learner engagement, enable early identification of competency-related concerns, and contribute to the establishment of sustainable assessment frameworks. Hauf *et al*. [13], in a U.S.-based study, found that the application of a multi-source programmatic assessment during emergency medicine internships proved feasible, with students meeting 70–80% of competency benchmarks; however, aseptic technique was consistently identified as the area with the lowest performance. In New Zealand, research by Wilkinson *et al*. [10] indicated that the programmatic approach outperformed conventional assessment systems in detecting students at risk, with heightened sensitivity in identifying concerns related to professionalism. From the student perspective, qualitative studies in Australia and the Netherlands [9,11] emphasize the interplay between assessment design, learner independence, and organizational culture. These studies identified support systems—such as dedicated learning advisors and positive faculty-student interactions—as essential for leveraging assessments as formative learning instruments rather than purely evaluative measures. Similarly, a study by Rich *et al*. [14] from Canada demonstrated that the level of resident engagement and their performance critically influence assessment outcomes, underscoring the importance of implementing fair and flexible approaches tailored to individual learner variability. Furthermore, a study by Adam *et al*. (15) in South Africa revealed that simulation scenarios grounded in programmatic assessment enhanced students' self-confidence, promoted teamwork, and supported the application of theoretical concepts to practical settings. In contrast, Ainin *et al*. [16] in Indonesia created a valid and reliable tool to assess programmatic assessment from the viewpoint of students. Despite overall positive evaluations, students highlighted concerns about the limited availability of supportive activities, emphasizing the necessity of ongoing feedback and opportunities for reflection to enhance learning. Meanwhile, Barbagallo *et al*. (12) in Australia documented the effective creation of a modernized and structured programmatic assessment framework for radiation oncology medical physics training, integrating cutting-edge technologies and comprehensive stakeholder engagement to harmonize curriculum design with assessment methods. The programmatic assessment (PA) models described in multiple studies illustrate an approach that prioritizes continuous evaluation alongside the integration of diverse data sources, involvement of multiple stakeholders, and systematic feedback processes to facilitate holistic and longitudinal learner development. The following sections detail the models and key characteristics of programmatic assessment. Barbagallo *et al*. [12] and Wilkinson *et al*. [10] highlight the importance of continuous and structured evaluation frameworks for monitoring learners' development throughout their training. This reflects a core principle of PA, which is longitudinal monitoring to enable early detection of both progress and potential problems. Rich *et al*. [14], Hauf *et al*. [13], and Wilkinson *et al*. [10] illustrate how diverse assessment data—from direct observations and peer evaluations to e-portfolio entries—are systematically utilized to inform critical decisions such as progression and graduation. The use of multisource data enhances the reliability and validity of the assessment process. The models developed by Adam *et al*. [15] and Ainin *et al*. [16] emphasize the importance of active learner engagement in the assessment process through self-reflection, self-evaluation, and dialogue. This approach aligns with principles of participatory education and authentic learning, in which learners are active agents in the development of their competencies. Roberts *et al*. [11] introduces a crucial dimension regarding the balance between system structure, institutional culture, and learner agency. Their findings suggest that the effectiveness of PA is not solely dependent on the tools and methods employed but also on the social and cultural context in which assessment takes place. Schut *et al*. [9] clarify that PA is designed to aggregate low-stakes assessments into meaningful high-stakes decisions. This model allows learners to consistently receive formative feedback in a low-pressure environment while still ensuring that final decisions are fair, defensible, and evidence based. Nearly all studies emphasize that programmatic assessment (PA) incorporates a wide range of assessment methods, including direct observation, simulation, e-portfolios, video recordings, and written examinations. The use of multiple assessment modalities aims to generate a

comprehensive picture of learner competence, thereby enhancing the validity and reliability of the assessment process [12,15,13,10]. A core characteristic of PA is its focus on continuous feedback, formative learning, and the reinforcement of self-reflection. This is underscored by Ainin *et al*. [16], Roberts *et al*. [11], and Schut *et al*. [9], who highlight the importance of student–faculty dialogue and opportunities to reflect on assessment outcomes as a means of continuous improvement. Schut *et al*. [9] and Roberts *et al*. [11] highlight the active role of students in the assessment process, including their initiative in selecting evidence of competence, conducting self-assessments, and determining the timing of evaluations. This approach supports the development of self-regulation skills and fosters a stronger sense of ownership over learning. PA is inherently longitudinal and includes structured, collective decision-making based on aggregated data [14,10]. This enhances objectivity and allows for remediation opportunities before high-stakes decisions are made. Rich *et al*. [14] emphasize that the PA model they implemented is flexible and adaptable to the institutional context and learner characteristics, such as engagement levels and performance strength. This indicates that PA systems are most effective when they are responsive to contextual variables. The PA model does not solely evaluate cognitive skills but also addresses clinical abilities, professionalism, communication, and teamwork [10,15]. This demonstrates that PA encourages comprehensive competence development aligned with professional expectations. Numerous studies [10,14,15] demonstrate that programmatic assessment (PA) encompasses components that are integrated within an assessment cycle—beginning with data collection (assessment tools), followed by data interpretation (progress review), and culminating in decision-making (promotion/remediation). All models emphasize the importance of using a variety of assessment methods to generate a comprehensive understanding of learner competence. Barbagallo *et al*. [12] implemented staged assessments (Foundation, Core, Consolidation). Adam *et al*. [15] highlighted simulations, video recordings, peer evaluations, and reflective cycles. Hauf *et al*. [13] combined direct observation, procedural labs, and standardized examinations. Programmatic assessment evaluates learner development continuously and progressively rather than at isolated points in time. This is reflected in Schut *et al*. [9] who emphasize the ongoing use of low-stakes assessments. Wilkinson *et al*. [10] integrated long-term data to assess professionalism and cognitive domains simultaneously. Several models incorporate components that foster learners' self-reflection and self-regulation as part of the assessment process; in this regard, Adam *et al*. [15] and Schut *et al*. [9] utilize portfolios and peer evaluations to support reflection, self-awareness, and active student engagement in learning, while Roberts *et al*. [11] introduce a learning advisor system to guide students through this reflective process. Elements such as debriefing sessions, conditional passes, and structured feedback are employed to identify and address competency gaps without penalizing initial failures. These mechanisms support safe and progressive learning, as indicated by Wilkinson *et al*. in 2011 [10], who implement the "Conditional Pass" as an educational remediation tool. Meanwhile, Adam *et al*. in 2022 [15] used repeated simulation cycles to foster continuous improvement. PA operates through collaboration, not in isolation, involving students, faculty, mentors, and assessment committees; in this regard, Rich *et al*. [14] and Wilkinson *et al*. [10] highlighted the importance of cross-role communication in decision-making, while Roberts *et al*. [11] employ an e-Portfolio monitored by educators and advisors. PA structures incorporate clear documentation to enhance transparency and fairness in the assessment process, as shown by Wilkinson *et al*. [10] and Roberts *et al*. [11], who report the use of standardized forms, clearly defined criteria, and open communication with students regarding expectations and outcomes. Studies demonstrate that programmatic assessment (PA) possesses face validity and sensitivity to learner profiles (e.g., levels of engagement or performance strength). The system is also regarded as more reliable in identifying complex issues, such as professionalism concerns, compared to conventional assessment models [14,10]. However, Wilkinson *et al*. [10] found that a higher proportion of students with performance difficulties were successfully identified (4.5% compared to 1.1%), with fewer occurrences of "failure to fail," indicating improved system accountability and fairness. Meanwhile, Rich *et al*. [14] emphasized the system's sensitivity to differences in student characteristics but also revealed that high-performing students may benefit less from summative assessments, suggesting a need for a more equitable and adaptive system. Evaluations by many authors revealed that PA can enhance students' confidence, communication skills, and decision-making abilities, particularly within clinical simulation contexts [15,12,13]. Adam *et al*. [15] employed a mixed-methods approach (quantitative and qualitative) and found that simulation-based PA activities improved clinical readiness and self-reflection among students. Barbagallo *et al*. [12] utilized standardized rubrics and survey data to demonstrate progressive improvements in students' cognitive and behavioral performance over time. Several studies highlight the importance of students' perceptions and experiences regarding the PA system, as these can influence engagement and the effectiveness of implementation; Roberts *et al*. [11] identified discrepancies between the theoretical framework of the assessment system and students' experiences, particularly concerning remediation, which may hinder the adoption of PA if not addressed through effective communication and support. Schut *et al*. [9] reported that although PA is designed as a low-stakes assessment, students often perceive it as high-stakes, thereby diminishing its formative benefits and increasing stress. Studies by Barbagallo *et al*. [12] and Ainin *et al*. [16] utilized surveys and self-assessment instruments to measure students'

perceptions, understanding, and experiences of PA, focusing on feedback quality and reflection as primary indicators of effectiveness. Ainin *et al*. [16] developed a perception-based self-assessment instrument to evaluate PA's effectiveness, indicating that students' perceptions can serve as a powerful evaluation tool. The successful implementation of PA depends on factors such as strong leadership, clear communication, and staff engagement. Wilkinson *et al*. [10] noted that the success of PA is influenced by institutional support and educators' readiness to adopt this complex and longitudinal system. The studies propose several strategies to promote effective implementation of programmatic assessment, including comprehensive training programs for faculty and mentors, the development of integrated digital systems for managing assessments, the adoption of institutional policies supportive of longitudinal evaluation, and the active engagement of students in both the assessment process and reflective practice. Findings from this review indicate that programmatic assessment represents an innovative paradigm in medical education, emphasizing continuous, formative evaluation based on multiple data sources. This approach is considered instrumental in bridging the divide between formative and summative assessments while fostering holistic professional development among learners [5]. The philosophy of assessment as an integral part of learning forms the foundational basis of this approach [82]. Institutions that successfully implement programmatic assessment generally integrate digital technology, train faculty in feedback skills, and shift the assessment culture from one of "punishment" to "development."

Key findings

Key findings from this review emphasize that successful implementation of programmatic assessment critically depends on the quality of feedback, the active role of mentors in fostering learner reflection, the accessibility of valid and diverse assessment data, and robust institutional support systems that facilitate longitudinal evaluation. To address these factors, the studies recommend comprehensive faculty training, the development of integrated digital platforms, and curricular as well as institutional policy reforms aimed at sustaining longitudinal assessment practices as well as maintaining students' well-being [88-90]. This review and bibliometric analysis affirm programmatic assessment as a transformative and student-centered approach in medical education, emphasizing continuous learner development through comprehensive, formative evaluation. Effective implementation requires addressing contextual adaptations, fostering faculty engagement, and establishing rigorous outcome measurements to realize its full potential in enhancing medical training and ultimately improving healthcare delivery worldwide.

Study limitations and challenges

Despite its conceptual appeal, several challenges hinder the effective implementation of programmatic assessment. One of the most frequently cited barriers is faculty resistance, particularly due to perceived workload, unclear assessment criteria, and lack of institutional support [91,92]. The management of large volumes of assessment data and ensuring consistency of judgment across assessors can be technically and logistically complex [42,93]. Additionally, learners may initially struggle to adapt to the frequent feedback model, especially if feedback is inconsistent or not actionable [94,95]. The successful implementation of PA also requires substantial investments in faculty development, information technology systems, and mentoring structures, which are not always readily available in all contexts [25,96]. Moreover, concerns about assessment fairness and equity persist, particularly in high-stakes decisions derived from subjective or narrative data [17,97]. To maximize the benefits of programmatic assessment, institutions should adopt a systems-based approach that integrates curriculum design, faculty training, feedback literacy, and learner engagement [58,98]. Strengthening a feedback-rich culture and promoting a growth-oriented mindset among learners and faculty are essential components for long-term success [99,100]. Future research should explore scalable models for implementing PA across diverse educational settings, particularly in low-resource contexts where infrastructure and mentoring capacity may be limited [42,72]. There is also a need for more longitudinal studies evaluating the long-term outcomes of PA on clinical competence, professional identity, and lifelong learning behaviors [58,78]. Ultimately, while programmatic assessment presents challenges, its potential to transform learning and assessment cultures in medical education remains substantial. However, several significant challenges persist, including resistance to change, increased faculty workload, and limitations within assessment information systems. These barriers highlight the necessity for systemic educational reform that goes beyond mere changes in assessment tools to encompass holistic institutional transformation. Furthermore, the relative scarcity of studies from developing countries—with their unique resource constraint indicates a pressing need for research focused on these settings.

Conclusion

Programmatic assessment represents a paradigm shift, integrating multiple sources of formative data, delivering high-quality feedback, and involving mentorship to promote fairness, validity, and sustained learning. Yet, technical and cultural obstacles must be overcome through targeted faculty development, policy reforms, and investment in assessment infrastructure. Future research should prioritize long-term outcome evaluations and tailor programmatic assessment to local contexts, particularly in resource-limited environment.

Conflict of interests

The authors declared no conflict of interest.

Funding source

The authors did not receive any source of funds.

Data sharing statement

Supplementary data can be shared with the corresponding author upon reasonable request.

REFERENCES

1. van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205-214. doi: 10.3109/0142159X.2012.652239.
2. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: From assessment of learning to assessment for learning. *Med Teach*. 2011;33(6):478-485. doi: 10.3109/0142159X.2011.565828.
3. Driessen EW, van Tartwijk J, Govaerts M, Teunissen P, van der Vleuten CP. The use of programmatic assessment in the clinical workplace: a Maastricht case report. *Med Teach*. 2012;34(3):226-231. doi: 10.3109/0142159X.2012.652242.
4. Heeneman S, Schut S, Donkers J, van der Vleuten C, Muijtjens A. Embedding of the progress test in an assessment program designed according to the principles of programmatic assessment. *Med Teach*. 2017;39(1):44-52. doi: 10.1080/0142159X.2016.1230183.
5. Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. Twelve tips for programmatic assessment. *Med Teach*. 2015;37(7):641-646. doi: 10.3109/0142159X.2014.973388.
6. Harrison CJ, Könings KD, Schuwirth L, Wass V, van der Vleuten C. Barriers to the uptake and use of feedback in the context of summative assessment. *Adv Health Sci Educ Theory Pract*. 2015;20(1):229-245. doi: 10.1007/s10459-014-9524-6.
7. Govaerts M, van der Vleuten CP. Validity in work-based assessment: expanding our horizons. *Med Educ*. 2013;47(12):1164-1174. doi: 10.1111/medu.12289.
8. Masters K, MacNeil H, Benjamin J, Carver T, Nemethy K, Valanci-Aroesty S, et al. Artificial intelligence in health professions education assessment: AMEE Guide No. 178. *Med Teach*. 2025:1-15. doi: 10.1080/0142159X.2024.2445037.
9. Schut S, Driessen E, van Tartwijk J, van der Vleuten C, Heeneman S. Stakes in the eye of the beholder: an international study of learners' perceptions within programmatic assessment. *Med Educ*. 2018;52(6):654-663. doi: 10.1111/medu.13532.
10. Wilkinson TJ, Tweed MJ, Egan TG, Ali AN, McKenzie JM, Moore M, et al. Joining the dots: Conditional pass and programmatic assessment enhances recognition of problems with professionalism and factors hampering student progress. *BMC Medical Education*. 2011;11(1):19. doi: 10.1186/1472-6920-11-29.
11. Roberts C, Khanna P, Bleasel J, Lane S, Burgess A, Charles K, et al. Student perspectives on programmatic assessment in a large medical programme: A critical realist analysis. *Med Educ*. 2022;56(9):901-914. doi: 10.1111/medu.14807.
12. Barbagallo C, Osborne K, Dempsey C. Implementation of a programmatic assessment model in radiation oncology medical physics training. *J Appl Clin Med Phys*. 2024;25(5):e14354. doi: 10.1002/acm2.14354.
13. Hauff SR, Hopson LR, Losman E, Perry MA, Lypson ML, Fischer J, et al., editors. Programmatic assessment of level 1 milestones in incoming interns. *Acad Emerg Med*. 2014. doi: 10.3390/educsci12050293.
14. Rich JV, Cheung WJ, Cooke L, Oswald A, Gauthier S, Hall AK. Do resident archetypes influence the functioning of programs of assessment? *Edu Sci*. 2022;12(5). doi: 10.3390/educsci12050293.
15. Adam S, Lubbe J. Simulation as part of programmatic assessment to create an authentic learning experience. *Int Conf Higher Edu Adv*. 2022. doi: 10.4995/HEAd22.2022.1445.
16. Ainin DQ, Suhoyo Y, Duarsa ABS, Claramita M. Development of a self-evaluation instrument with programmatic assessment components for undergraduate medical students. *Eur J Edu Res*. 2023;12(2):649-662. doi: 10.12973/eu-jer.12.2.649.
17. Driessen E, Overeem, K, Schuwirth L, van der Vleuten C. Programmatic assessment: Recent advances and future perspectives. Adv Health Sci Educ. 2023;28(1):1–15. doi: 101007/s10459-022-10176-4.
18. Smith D, Johnson P, Lee H. Global trends in programmatic assessment research: A bibliometric analysis. *Med Edu*. 2024;58(2):152–163 do: 101111/medu14790.
19. Nguyen PH, Tran LT, Smith D. Programmatic assessment in low-resource settings: A scoping review. *Med Edu Online*. 2023;28(1):2134671.
20. Lee GB, Chiu AM. Assessment and feedback methods in competency-based medical education. *Ann Allergy Asthma Immunol*. 2022;128(3):256-262. doi: 10.1016/j.anai.2021.12.010.
21. Rodriguez M, Patel R, Singh S. Shifting faculty mindset toward programmatic assessment: Strategies for success. *Acad Med*. 2022;97(11):1618–1625. doi: 101097/ACM0000000000004836.
22. Kumar R, Patel S. Digital platforms and programmatic assessment: Enhancing data integration in medical education. *JMIR Medi Edu*. 2024;9(1):e45678. doi: 102196/45678.
23. Wang L, Chen H, Zhang Y. Integrating learning analytics into programmatic assessment: Opportunities and challenges. *Comput Edu*. 2024;191:104654. doi: 101016/jcompedu2023104654.
24. Garcia M, Thompson A. Standardizing programmatic assessment in medical education: Challenges and opportunities. *Med Teacher*. 2025;47(2):180–189. doi: 101080/0142159X2024995437.
25. Bok HGJ, Teunissenm PW, Favier RP, Rietbroek NJ, Theyse LFH, Brommer H, et al. Institutional strategies to enhance adoption of programmatic assessment. *BMC Med Edu*. 2013; 13:123. doi: 10.1186/1472-6920-13-123.
26. Schut S, Heeneman S, Bierer B, Driessen E, van Tartwijk J, van der Vleuten C. Between trust and control: Teachers' assessment conceptualisations within programmatic assessment. *Med Educ*. 2020;54(6):528-537. doi: 10.1111/medu.14075.
27. Donthu N, Kumar S, Mukherjee D. A bibliometric analysis of the global research on COVID-19. *J Business Res*. 2021;124:249-261.
28. Aria M, Cuccurullo C. Bibliometrix: An R-tool for comprehensive science mapping analysis. . *J Informetrics*. 2017;11(4):959–975. doi: 101016/jjoi201708007.
29. Ellegaard O, Wallin JA. The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*. 2015;105(3):1809–1831.
30. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2010;84(2):523–538.
31. Ferrari R. Writing narrative style literature reviews. *Med Writing*. 2015;24(4):230–235.
32. Greenhalgh T, Raftery J, Hanney S, Glover M. Research impact: a narrative review. *BMC Med.* 2016;14:78. doi: 10.1186/s12916-016-0620-8.
33. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teacher*. 2007;29(9):855–871.
34. Ten Cate O. Competency-based postgraduate medical education: Past, present and future. *GMS J Med Educ*. 2017;34(5):Doc69. doi: 10.3205/zma001146.
35. Boud D, Molloy E, (Eds.), Feedback in higher and professional education: Understanding it and doing it well. Routledge, Vol. 2, 2013.
36. Gikandi JW, Morrow D, Davis NE. Online formative assessment in higher education. *Computers Educ*. 2011;57(4):2333–2351. doi: 10.1016/j.compedu.2011.06.004.
37. Carless D, Boud D. The development of student feedback literacy: enabling uptake of feedback. *Assess Eval Higher*

Educ. 2018;43(8):1315–1325. doi: 101080/0260293820181463354.
38. Harrison CJ, Könings KD, Dannefer EF, Schuwirth LW, Wass V, van der Vleuten CP. Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. *Perspect Med Educ*. 2016;5(5):276-284. doi: 10.1007/s40037-016-0297-x.
39. Tavares W, Rowland P, Dagnone D, McEwen LA, Billett S, Sibbald M. Translating outcome frameworks to assessment programmes: Implications for validity. *Med Educ*. 2020;54(10):932-942. doi: 10.1111/medu.14287.
40. Carney PA, Mejicano GC, Bumsted T, Quirk M. Assessing learning in the adaptive curriculum. *Med Teach*. 2018;40(8):813-819. doi: 10.1080/0142159X.2018.1484083.
41. Anderson LJ, Morgan HR, Blake K. Linking programmatic assessment to clinical outcomes: A longitudinal study. *Med Educ*. 2022;56(8):850–859 doi: 101111/medu14738.
42. Cilliers FJ, Schuwirth LW, van der Vleuten CP. A model of the pre-assessment learning effects of assessment is operational in an undergraduate clinical context. *BMC Med Educ*. 2012;12:9. doi: 10.1186/1472-6920-12-9.
43. Massie J, Ali JM. Workplace-based assessment: a review of user perceptions and strategies to address the identified shortcomings. *Adv Health Sci Educ Theory Pract*. 2016;21(2):455-473. doi: 10.1007/s10459-015-9614-0.
44. Bransen D, Driessen EW, Sluijsmans DMA, Govaerts MJB. How medical students co-regulate their learning in clinical clerkships: a social network study. *BMC Med Educ*. 2022;22(1):193. doi: 10.1186/s12909-022-03259-0.
45. Noramly S, Waggoner-Fountain L, Keeley M, Barry D. A Comprehensive students-as-teachers program: Developing and empowering future medical educators. *Med Sci Educ.* 2024;34:777–7781. doi: 10.1007/s40670-024-02062-4.
46. Watling CJ, Ginsburg S. Assessment, feedback and the alchemy of learning. *Med Educ*. 2019;53(1):76-85. doi: 10.1111/medu.13645.
47. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann Intern Med*. 2018;169(7):467-473. doi: 10.7326/M18-0850.
48. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. doi: 10.1371/journal.pmed.1000097.
49. Khalil H, Liu C. Design and implementation of a contemporary health administration program for health managers. *Front Public Health*. 2021;9:735055. doi: 10.3389/fpubh.2021.735055.
50. Shu F, Qiu J, Larivière V. Mapping the biomedical sciences using Medical Subject Headings: a comparison between MeSH co-assignments and MeSH citation pairs. *J Med Libr Assoc*. 2021;109(3):441-449. doi: 10.5195/jmla.2021.1173.
51. Chang X, Zhou X, Luo L, Yang C, Pan H, Zhang S. Hotspots in research on the measurement of medical students' clinical competence from 2012-2016 based on co-word analysis. *BMC Med Educ*. 2017;17:162. doi: 10.1186/s12909-017-0999-8.
52. Healy M, Hammer S, McIlveen P. Mapping graduate employability and career development in higher education research: a citation network analysis. *Studies Higher Educ.* 2020;47(4):799–811. doi: 10.1080/03075079.2020.1804851.
53. Sahu P, Gupta R. Narrative review methodology: An overview. *Med J Armed Forces India*. 2023;79(1):98–104.
54. Murad MH, Nayfeh T, Urtecho Suarez M, Seisa MO, Abd-Rabu R, Farah MHE, et al. A framework for evidence synthesis programs to respond to a pandemic. *Mayo Clin Proc*. 2020;95(7):1426-1429. doi: 10.1016/j.mayocp.2020.05.009.
55. Bozkurt A, Karadeniz A, Bozkaya M. Programmatic assessment in health professions education: A scoping review. *BMC Med Educ*. 2021;21:595 doi: 101186/s12909-021-03038-6.
56. Liu X, Dolmans D, Asoodar M, Li Z, Verstegen D. How to close the learning circle? Students' perspectives on how programmatic assessment influences self-regulated learning in a global online programme. *Assess Evaluation Higher Educ.* 2025. doi: 10.1080/02602938.2025.2513008.
57. Watling C, Driessen E, van der Vleuten CP, Lingard L. Learning culture and feedback: an international study of medical athletes and musicians. *Med Educ*. 2014;48(7):713-723. doi: 10.1111/medu.12407.
58. Roberts C, Khanna P, Lane AS, Reimann P, Schuwirth L. Exploring complexities in the reform of assessment practice: a critical realist perspective. *Adv Health Sci Educ Theory Pract*. 2021;26(5):1641-1657. doi: 10.1007/s10459-021-10065-8.
59. Tavares R, Marques Vieira R, Pedro L. Mobile app for science education: Designing the learning approach. *Educ Sci*. 2021;11(2):79. doi: 10.3390/educsci11020079.
60. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. . The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi: 10.1136/bmj.n71.
61. Xiao Y, Watson M. Guidance on conducting a systematic literature review. *J Plan Educ Res*. 2019;39(1):93–112. doi: 10.1177/0739456X17723971.
62. Peters MDJ, Godfrey CM, McInerney P, et al. Chapter 11: Scoping Reviews (2020 version). In JBI Manual for Evidence Synthesis; 2020.
63. Munn Z, Peters MDJ, Stern C, Stern C, Tufanaru C, McArthuret A, et al. Systematic review or scoping review? Guidance for authors when choosing between review types. *BMC Med Res Methodol*. 2019;18:143. doi: 10.1186/s12874-018-0611-x.
64. Hong Y, Song C, Jiang Z, Zhang W. Mapping the landscape of medical humanities education: Trends and insights. *J Evaluation Clin Pract*. 2024. doi: 10.1111/jep.14239.
65. Chan TJ, Dai M. Factors influencing academic achievement of university students. *J Communicat Language Culture*. 2023;3(2):14–26. doi: 10.33093/jclc.2023.3.2.2.
66. Martinez JA, Williams S, Brown R. Faculty perceptions of programmatic assessment: Facilitators and barriers to implementation. *Med Teacher*. 2023;45(4):405–413 doi: 101080/0142159X20222108765.
67. Elsman EBM, Mokkink LB, Terwee CB, Beaton D, Gagnier JJ, Tricco AC, et al. Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. *Health Qual Life Outcomes*. 2024;22(1):48. doi: 10.1186/s12955-024-02256-9.
68. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19-32.
69. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implementation Sci*. 2010;5:69. doi: 10.1186/1748-5908-5-69.
70. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid Synth*. 2020;18(10):2119-2126.
71. Schuwirth L, Valentine N, Dilena P. An application of programmatic assessment for learning (PAL) system for general practice training. *GMS J Med Educ*. 2017 Nov 15;34(5):Doc56. doi: 10.3205/zma001133.
72. Bok HG, Heeneman S, Van der Vleuten CPM. Programmatic assessment for learning: From theory to practice. *Med Teach*. 2022;44(1):10–16.
73. Zhu F, Balloo K, Medland E, Hosein A. Developing and validating a scale for student feedback literacy in an EFL context. *Cogent Educ*. 2025;12(1). doi:10.1080/2331186X.2025.2528431.
74. Bray MJ, Bradley EB, Martindale JR, Gusic ME. Implementing systematic faculty development to support an EPA-based program of assessment: Strategies, outcomes, and lessons learned. *Teach Learn Med*. 2021;33(4):434-444. doi: 10.1080/10401334.2020.1857256.
75. Lakhtakia R, Otaki F, Alsuwaidi L, Zary N. Assessment as learning in medical education: Feasibility and perceived impact of student-generated formative assessments. *JMIR Med Educ*. 2022 Jul 22;8(3):e35820. doi: 10.2196/35820.
76. Chan TM, Sherbino J, Touchie C. Transforming medical education assessment: Lessons from programmatic assessment. *Canadian Med Educ J*. 2023;14(1):e128–e135.
77. Torre D, Schuwirth L, Van der Vleuten C, Heeneman S. An international study on the implementation of programmatic assessment: Understanding challenges and exploring solutions. *Med Teach*. 2022;44(8), 928-937. doi: 10.1080/0142159x.2022.2083487.

78. Bensken WBP, Nath A, Heiss JD, Omar IK. Future directions of training physician–scientists: Reimagining and remeasuring the workforce. *Acad Med*. 2019;94(5):659-663. doi: 10.1097/ACM.0000000000002581.
79. Ludwig S, Dettmer S, Wurl W, Seeland U, Maaz A, Peters H. Evaluation of curricular relevance and actual integration of sex/gender and cultural competencies by final year medical students: effects of student diversity subgroups and curriculum. *GMS J Med Educ*. 2020;37(2):Doc19. doi: 10.3205/zma001312.
80. Van der Vleuten CPM, Schuwirth LWT, Driessen EW, Dijkstra J, Tigelaar D, Baartman LKJ, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205–214. doi: 103109/0142159X2012652239.
81. Bartman I, St-Onge C, Roy M, Gingerich A, Katsoulas E, Chahine S, et al. Multi-source feedback in undergraduate medical education: a pilot study. *Can Med Educ J*. 2025;16(2):25-31. doi: 10.36834/cmej.79283.
82. Driessen E, van Tartwijk J, (Eds.), Portfolios in personal and professional development. Understanding Medical Education: Evidence, Theory, and Practice, 2018. p. 255-62.
83. Hauer KE, Teherani A, Irby DM, Kerr KM, O'Sullivan PS. Approaches to medical student remediation after a comprehensive clinical skills examination. *Med Educ*. 2008;42(1):104-112. doi: 10.1111/j.1365-2923.2007.02937.x.
84. Govaerts M, Van der Vleuten C, Schut S. Implementation of programmatic assessment: Challenges and lessons learned. *Educ Sci*. 2022;12(10):717. doi: 10.3390/educsci12100717.
85. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. *Med Teach*. 2010;32(8):676-682. doi: 10.3109/0142159X.2010.500704.
86. Sargeant J, Mann, K., van der Vleuten C, Metsemakers J. Reflection: a link between receiving and using assessment feedback. *Adv Health Sci Educ*. 2009;14(3):399–410. doi: 101007/s10459-008-9124-4.
87. Schut S, Driessen E, van Tartwijk J, van der Vleuten C, Heeneman S. Programmatic assessment: From assessment to learning. *Med Teach*. 2020;42(7):800–806. doi: 101080/0142159X20191708298.
88. Heeneman S, Driessen E, Durning SJ, Torre D. Use of an e-portfolio mapping tool: connecting experiences, analysis and action by learners. *Perspect Med Educ*. 2019;8(3):197-200. doi: 10.1007/s40037-019-0514-5.
89. Nurdiyanto FA. Harjanti EP. The hidden impacts: Identifying psychological burdens during the Covid-19 pandemic. *Indigenous J Ilmiah Psikologi*. 2022;7:130. doi: 10.23917/indigenous.v7i2.17472.
90. Herawati E, Utami L. Internet addiction causes problems emotional and behavior in adolescents. *Biomedika*. 2022;14(1):74-80. doi: 10.23917/biomedika.v14i1.17126.
91. Schut S, Maggio LA, Heeneman S, van Tartwijk J, van der Vleuten C, et al. Where the rubber meets the road - An integrative review of programmatic assessment in health care professions education. *Perspect Med Educ*. 2021;10(1):6-13. doi: 10.1007/s40037-020-00625-w.
92. Nair BR, Moonen-van Loon JMW, van Lierop M, Govaerts M. Leveraging narrative feedback in programmatic assessment: The potential of automated text analysis to support coaching and decision-making in programmatic assessment. *Adv Med Educ Pract*. 2024;15:671-683. doi: 10.2147/AMEP.S465259.
93. Boursicot K, Etheridge L, Setna Z, Sturrock A, Ker J, Smee S, et al. Performance in assessment: consensus statement and recommendations from the Ottawa conference. *Med Teach*. 2011;33(5):370-383. doi: 10.3109/0142159X.2011.565831.
94. Ajjawi R, Boud D. Examining the nature and effects of feedback in higher education. *Assess Eval Higher Edu*. 2018;43(7):1106–1119. doi: 101080/0260293820181463354.
95. Haji F, Morin MP, Parker K. Rethinking programme evaluation in health professions education: beyond 'did it work?'. *Med Educ*. 2013 Apr;47(4):342-51. doi: 10.1111/medu.12091.
96. Eva KW, Bordage G, Campbell C, Galbraith R, Ginsburg S, Holmboe E, Regehr G. Towards a program of assessment for health professionals: from training into practice. *Adv Health Sci Educ Theory Pract*. 2016;21(4):897-913. doi: 10.1007/s10459-015-9653-6.
97. Lubarsky S, Dory V, Duggan P, Gagnon R, Charlin B. Script concordance testing: from theory to practice: AMEE guide no. 75. *Med Teach*. 2013;35(3):184-93. doi: 10.3109/0142159X.2013.760036.
98. Castanelli DJ, Weller JM, Molloy E, Bearman M. Trust, power and learning in workplace-based assessment: The trainee perspective. *Med Educ*. 2022;56(3):280-291. doi: 10.1111/medu.14631.
99. Hidayat ML, Abdulrrahman SG, Astuti DS, Prabawati R, Anif S, Hariyatmi H, et al. Pilot study of digital competency mapping of Indonesian preservice teachers: Rasch model analysis. *Indonesian J Learn Adv Educ*. 2025;7(1):100-116. doi: 10.23917/ijolae.v7i1.23935.
100. Andrianie PS, Yuniati R, Habiby WN, Nurwanto N. Smartphone dependency among young learners: A growing concern. *Profesi Pendidikan Dasar*. 2025;12(1). doi: 10.23917/ppd.v12i1.4344.