e-ISSN : 2597-6052

MPPKI MEDIA PUBLIKASI PROMOSI KESEHATAN INDONESIA

Review Articles Open Access

Programmatic Assessment in Medical Education: A Bibliometric and Narrative Review of Its Potential Contributions to Health Literacy and Health Promotion Competencies

N Juni Triastuti[1,2*], Tri Nur Kristina[3], Titi Savitri Prihatiningsih[4]
[1]Doctoral Study Program of Medical and Health Science, Universitas Diponegoro, Semarang, Indonesia
[2]Department of Medical Education, Faculty of Medicine, Universitas Muhammadiyah Surakarta, 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

*Corresponding Author: E-mail: njt140@ums.ac.id

ARTICLE INFO

Manuscript Received: 09 Oct, 2025
Revised: 07 Dec, 2025
Accepted: 18 Dec, 2025
Date of Publication: 02 Feb, 2026
Volume: 9
Issue: 2
DOI: 10.56338/mppki.v9i2.8789

KEYWORDS

Programmatic Assessment;
Medical Education;
Bibliometric;
Narrative Review;
Health Literacy;
Health Promotion

ABSTRACT

Introduction: Programmatic assessment (PA) has emerged as a transformative framework in medical education, emphasizing longitudinal evaluation, meaningful feedback, and data-informed decision-making to enhance learner development. Its role extends beyond competency measurement, contributing to broader educational goals such as strengthening health literacy and health promotion competencies—core attributes for future healthcare professionals. This bibliometric and narrative review aimed to map global research trends, influential contributors, and thematic patterns in programmatic assessment while synthesizing evidence on its potential to advance health literacy and health promotion competencies in medical education.

Methods: A comprehensive search was conducted in Scopus database for articles published between 1977 and 2025. Inclusion criteria of eligible studies were those addressing programmatic assessment in medical or health professions education—including portfolios, decision processes, or programmatic feedback—and explicitly linking these approaches to health literacy or health promotion competencies. The dataset comprised 822 records and a final sample of 166 publications eligible for bibliometric analysis. Bibliometric indicators were analysed using VOSviewer, including publication trends, author networks, organizational and country contributions, and keyword co-occurrence. Narrative synthesis integrated empirical findings on PA implementation and its educational outcomes.

Results: A total of 1092 authors, 502 organizations, and 44 countries contributed to 1737 indexed keywords, of which 167 met inclusion thresholds. Thematic clustering identified four dominant domains: feedback and learning analytics, competency-based assessment, digital and AI-supported assessment, and professional identity formation. Recent literature emphasizes PA's capacity to cultivate reflective practice, communication skills, and health advocacy—key components of health literacy and promotion.

Conclusion: Programmatic assessment represents an integrative and evidence-driven approach that not only enhances learning and assessment quality but also fosters essential competencies in health literacy and promotion, aligning medical education with 21st-century public health priorities.

Publisher: Fakultas Kesehatan Masyarakat Universitas Muhammadiyah Palu

INTRODUCTION

Programmatic assessment (PA) has matured from a conceptual innovation into a widely discussed approach for designing assessment systems that simultaneously serve learning and high-stakes decision-making. Originally framed by van der Vleuten and colleagues, PA foregrounds continuous sampling of learner performance across multiple methods, deliberate aggregation of low-stakes data, and the use of narrative feedback and expert judgement to guide progression decisions (1-3). Recent consensus work and empirical syntheses have refined the core principles of PA and highlighted both its pedagogical promise and the practical complexity of implementation across contexts. These foundational accounts situate PA as an assessment philosophy that is particularly well suited to competency-based medical education (CBME) because it privileges developmental trajectories, meaningful feedback, and longitudinal judgement over single high-stakes examinations (1,3,4).

The last half-decade has seen an increase in applied research and design studies that explore how PA is operationalized in varied curricular, cultural, and resource contexts. Multi-institutional design investigations and program evaluations have documented the range of design choices—portfolios, progress committees, workplace-based assessment instruments, and digital platforms—through which PA principles are enacted, while also documenting recurring implementation barriers such as faculty workload, feedback quality, and curricular alignment (5-7). Contemporary scholarship therefore emphasizes that PA is not a single blueprint but a family of design decisions whose educational impact depends on fidelity to core principles, institutional capacity, and socio-cultural fit. This variation makes bibliometric and narrative mapping especially valuable: bibliometrics can surface trends, geographic patterns, and influential works, while narrative synthesis can unpack the mechanisms by which PA affects learning and competence development (1,5).

Health literacy—patients' and populations' capacities to access, understand, appraise, and use health information—has been linked to health behaviors and outcomes and is an emergent competency domain for clinical practice (8). Recent consensus and empirical studies have articulated concrete knowledge, skills, and attitudinal items for health literacy education, and curricula and workshops have demonstrated that targeted instruction improves trainee confidence and communication practices such as teach-back and jargon avoidance (8,9). Parallel efforts to define and assess health promotion competencies have underscored the need for integrative, longitudinal learning experiences that prepare graduates to promote population health through counselling, community engagement, and system-level thinking (10,11).

Given the longitudinal, developmental, and feedback-centered features of PA, it is plausible that PA offers fertile ground for reliably developing and assessing health literacy and health promotion competencies. PA's emphasis on multiple low-stakes assessments, rich narrative feedback, reflective practice, and aggregated judgement aligns with pedagogies for behavioral and communicative competencies that require iterative practice, coaching, and contextualized evaluation (5,12). Moreover, contemporary frameworks for novel competency domains—such as digital health literacy and population-level health promotion—explicitly call for integrative assessment strategies that span classroom, simulated, and workplace contexts; PA offers a conceptual architecture to do exactly that (6,13).

Despite conceptual fit, empirical evidence linking PA to improved health literacy and health promotion competencies remains scattered. This mixed and emerging literature motivates a combined bibliometric and narrative review: bibliometrics will quantify the growth, geographic distribution, and disciplinary intersections of PA research (e.g., connections to health literacy, communication, and health promotion), while a narrative synthesis will extract mechanisms, design choices, and evaluation outcomes that explain how (and when) PA supports these specific competencies. Such a review can inform educators and policymakers who seek to integrate PA into curricula intentionally targeted at producing graduates capable of promoting health and communicating effectively with diverse populations. The present study therefore maps the contemporary evidence base and synthesizes insights to guide future research, curriculum design, and assessment policy.

Concurrently, there has been growing international attention to health literacy and health promotion competencies as explicit learning outcomes for medical graduates. Emerging evidence shows that structured, feedback-rich assessment systems can enhance learners' capacity to interpret, communicate, and apply health information in clinical contexts (14). Competency-based assessment approaches—such as workplace-based assessments, portfolios, and longitudinal feedback—have also been associated with improved communication, shared decision-making, and preventive care behaviors, which are core domains of health literacy and health promotion (15).

Additionally, studies within medical and public health education indicate that assessments emphasizing reflective practice and integrative decision-making contribute to better learner performance in community-oriented and population-health competencies.

METHOD

This study employs a clear and systematic approach to ensure the reliability and validity of the findings. Below are the components of the methodology:

Research Type

This study employed a mixed-methods review combining a quantitative bibliometric analysis with a complementary narrative synthesis. The bibliometric component was designed to map publication trends, influential authors, journals, countries, and thematic clusters related to programmatic assessment (PA) and its intersections with health literacy and health promotion competencies. The narrative component interpreted content and methodological features of included primary studies to explain mechanisms, implementation choices, and reported educational outcomes. A convergent integrative logic guided synthesis so that bibliometric patterns informed purposive sampling for deeper narrative analysis (16,17).

Search strategy and information sources

We searched Scopus for records published between 1 January 1977 and the date of the search (to capture the development of PA since foundational work), with focused analyses of literature published in the last five years (2020–2025). Search strings combined controlled vocabulary and keywords for “programmatic assessment”, “programmatic evaluation”, “assessment program”, “health literacy”, and “health promotion”, and were translated from Scopus database. Bibliographic exports included full-record metadata and cited references (when available). Grey literature and conference proceedings from major health professions education meetings (e.g., Ottawa, AMEE) were sought to reduce publication bias. Search methods followed established bibliometric guidance to ensure reproducibility (search logs, database field mapping, and de-duplication procedures) (16,17).

A comprehensive literature search was conducted using the Scopus database to identify relevant publications on programmatic assessment within medical and health professions education. The search strategy employed a structured Boolean search string designed to capture both educational and health promotion dimensions of programmatic assessment. The search query was as follows:

(TITLE-ABS-KEY ("programmatic assessment" OR "programmatic evaluation" OR "programmatic approach" OR "assessment program") AND TITLE-ABS-KEY ("medical education" OR "health professions education" OR "clinical education" OR "undergraduate medical education" OR "postgraduate medical education") OR TITLE-ABS-KEY ("health literacy" OR "health promotion" OR "health behavior" OR "community health empowerment" OR "health communication" OR "lifestyle diseases" OR "health education")).

Eligibility criteria and study selection

Empirical and conceptual records were eligible if they explicitly addressed programmatic assessment in undergraduate or postgraduate medical or health professions education or examined assessment approaches that operationalized core PA principles (e.g., portfolios, progress/fail decision processes, programmatic feedback systems), and if they included explicit links to health literacy or health promotion outcomes or competencies. Exclusion criteria comprised studies limited to single high-stakes exams without programmatic intent, articles not in English, or publications lacking accessible metadata. Title/abstract screening and full-text review were performed in duplicate with discrepancies resolved by consensus; reasons for exclusion at full text were documented in a PRISMA-style flow diagram. Critical descriptive data (design, setting, PA components, competency targets, outcomes) were extracted into a structured spreadsheet for synthesis (1,18).

Sample Size

The final corpus consisted of 822 open-access publications and 166 that met all inclusion criteria. These studies were identified through systematic screening of titles/abstracts followed by full-text assessment, and the selection flow is documented in the PRISMA diagram to ensure methodological transparency.

Coding Rigor

To ensure analytic rigor, we developed a structured coding framework derived from the study objectives and preliminary scoping of the literature. The independent reviewers applied the initial codes to a pilot subset to refine category definitions, after which they coded the full dataset separately. Discrepancies were resolved through consensus discussions. All coding rules and definitions were documented in a finalized codebook.

Interpretive Reliability

Interpretive reliability was strengthened through methodological triangulation combining bibliometric indicators with narrative synthesis, assessment of inter-rater agreement on a randomized subset (Cohen's $\kappa \geq 0.70$ as the acceptable threshold). Representative excerpts from included studies were incorporated to enhance transparency and support thematic interpretations.

Bibliometric data processing and analysis

Cleaned bibliographic data were analyzed for descriptive publication trends (annual outputs, journal distribution, country and institutional contributions) and network metrics (co-authorship, co-citation, bibliographic coupling, and keyword co-occurrence). We used BibTeX/CSV exports for data cleaning (harmonizing author names and institutional affiliations) and analyzed networks with VOSviewer and Bibliometrix (R) to produce cluster maps, thematic evolution diagrams, and measures of centrality and citation impact. We adhered to recommended bibliometric practice for multi-database analyses, including transparent record merging, deduplication, normalization of citation counts, and sensitivity checks. Visualization parameters and cluster-resolution choices are reported to enable reproducibility (16,17).

Narrative synthesis, quality appraisal, and reporting

For the narrative review we purposively sampled studies revealed by bibliometric prominence and those explicitly addressing health literacy or health promotion outcomes. We used thematic synthesis to identify mechanisms by which PA designs purportedly support the development and assessment of health literacy and health promotion competencies (for example: iterative workplace-based feedback, reflective portfolio tasks, and progress committee deliberations). Methodological quality and risk of bias for empirical studies were appraised using appropriate checklists (e.g., MMAT for mixed-methods; CASP for qualitative studies; and ROBINS-I for non-randomized interventions) and used to contextualize confidence in reported outcomes. Findings are reported following PRISMA for the systematic elements, with separate sections describing bibliometric outcomes, thematic narrative synthesis, methodological limitations, and implications for curriculum and assessment policy (16,17).

This search yielded a total of 822 records. After the removal of one duplicate, 821 unique records remained. The dataset was then refined by limiting the results to the *Medicine* and *Health Professions* subject areas, resulting in 557 records. To ensure the inclusion of original and peer-reviewed evidence, the search was further restricted to *journal articles* and to those published in *English* (n=391). Finally, to enhance transparency and accessibility of data sources, only *open access* articles were included, yielding a final sample of 166 publications eligible for bibliometric analysis. These records formed the dataset used for subsequent mapping and narrative synthesis.

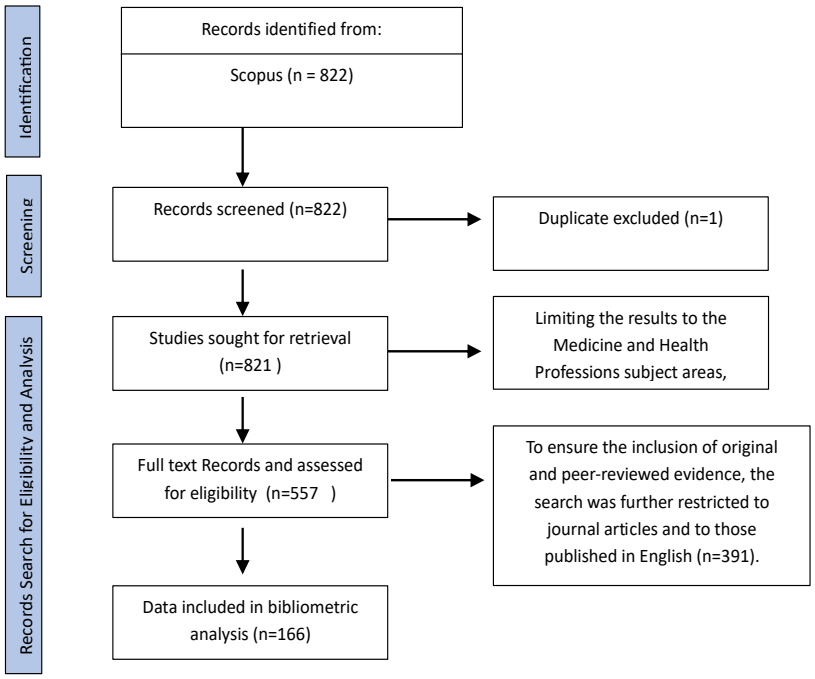
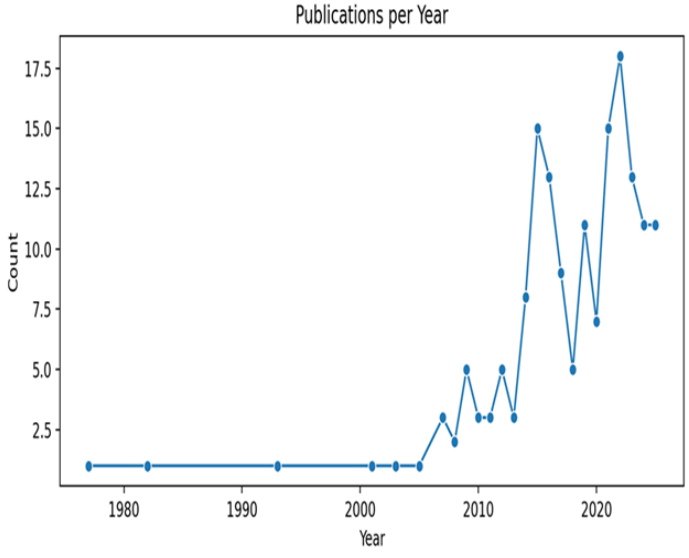


Figure 1. Study Selection process

RESULTS

Publication Trends

The field's foundation still centers on medical education, assessment, training, and programmatic approaches, while recent momentum is clustering around newer topics (e.g., AI-related terms and programmatic assessment). Average authors per paper was 6.1. This is the overall average team size per publication (Figure 2).



Year	Number of publications
2025	11
2024	11
2023	13
2022	18
2021	15
2020	7
2019	11
2018	5
2017	9
2016	13
2015	15
2014	8
2013	3
2012	5
2011	3
2010	3
2009	5
2008	2
2007	3
2005	1
2003	3
2001	1
1993	1
1982	1
1977	1

Figure 2. Number of Publications Trends per Year

The temporal distribution of publications demonstrated a substantial growth in scholarly output over the past two decades, indicating an accelerating interest in programmatic assessment and its applications in medical and health professions education. Early research activity was sporadic, with isolated publications appearing between 1977 and the early 2000s, reflecting the exploratory phase of assessment scholarship. From 2007 onward, however, the number of publications began to increase steadily, signaling the emergence of programmatic assessment as a distinct area of inquiry within educational research.

A notable surge in publication frequency occurred from 2015 to 2020, during which consistent annual outputs were observed. This period marks the consolidation of programmatic assessment concepts into mainstream educational discourse, characterized by intensified discussion of feedback systems, longitudinal assessment design, and competency-based learning frameworks. The upward trend continued through 2021 and 2022, years that recorded high levels of publication activity, suggesting expanding international engagement and institutional implementation of programmatic approaches across medical schools and training programs.

The peak of publication activity was reached between 2023 and 2025, indicating a significant recent escalation in scholarly contributions. Specifically, 2025 showed the highest concentration of studies, underscoring the sustained academic attention and growing momentum in refining the theoretical and practical aspects of programmatic assessment. This increase may also reflect the broader integration of digital innovations, such as artificial intelligence and data-driven feedback analytics, into assessment systems, leading to new research trajectories and methodological advancements.

Geographical Distribution

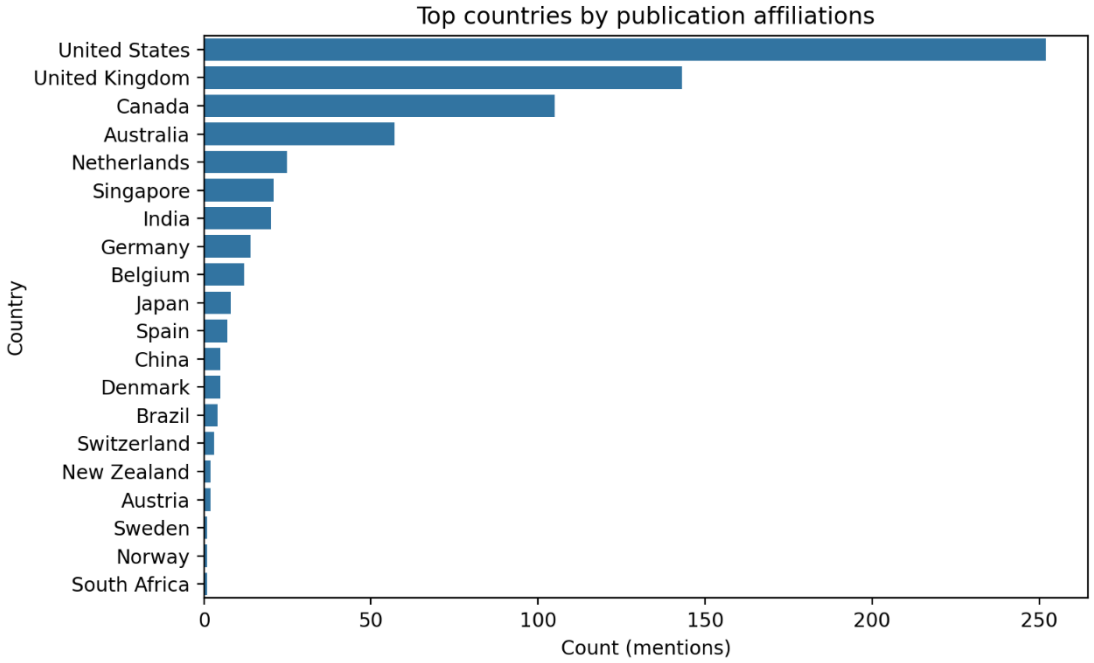


Figure 3. Top countries by publication affiliations

The country-level analysis of publication affiliations revealed that research on programmatic assessment in medical education is predominantly concentrated within five high-income nations: the United States, the United Kingdom, Canada, Australia, and the Netherlands. These countries collectively represent the major contributors to the global discourse on assessment reform and educational innovation in health professions training. The United States demonstrated the highest publication output, reflecting its strong academic infrastructure, diverse institutional collaborations, and sustained investment in competency-based medical education. The United Kingdom and Canada

followed closely, both serving as key centers of pedagogical research and policy development, particularly in integrating feedback mechanisms and longitudinal assessment frameworks into curricular design.

Australia and the Netherlands also showed significant scholarly engagement, contributing influential studies that emphasize reflective practice, workplace-based assessment, and professional identity formation. The prominence of these countries aligns with their established networks of medical education research and their early adoption of programmatic assessment principles within accreditation and quality assurance systems (Figure 3).

Co-citation and Keyword Analysis

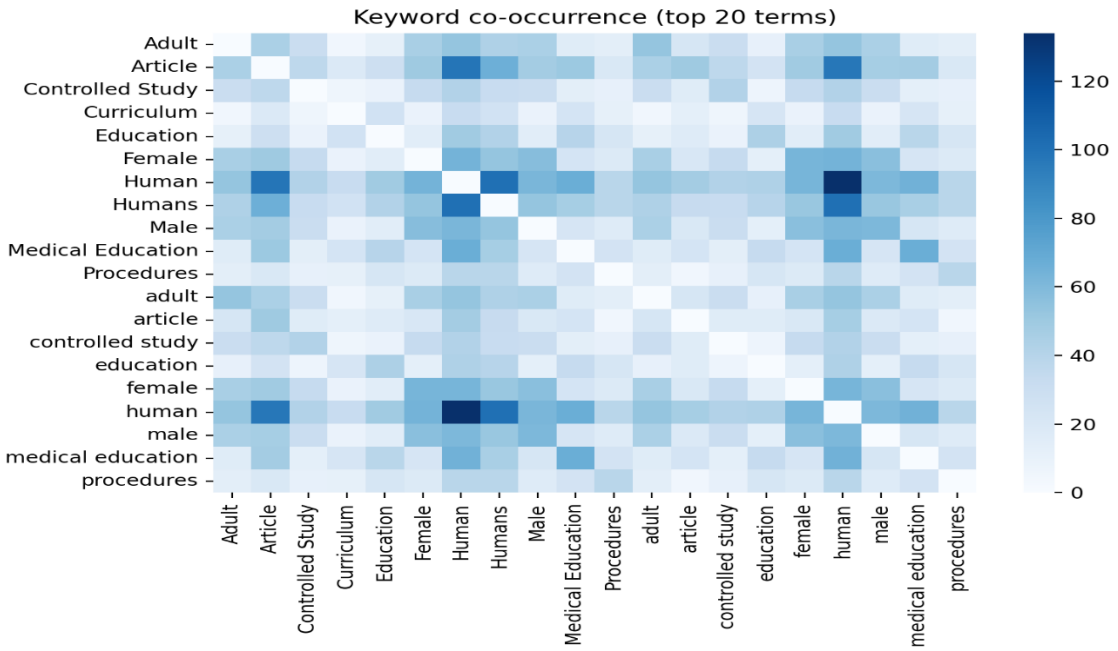


Figure 4. Top 20 keyword co-occurrence

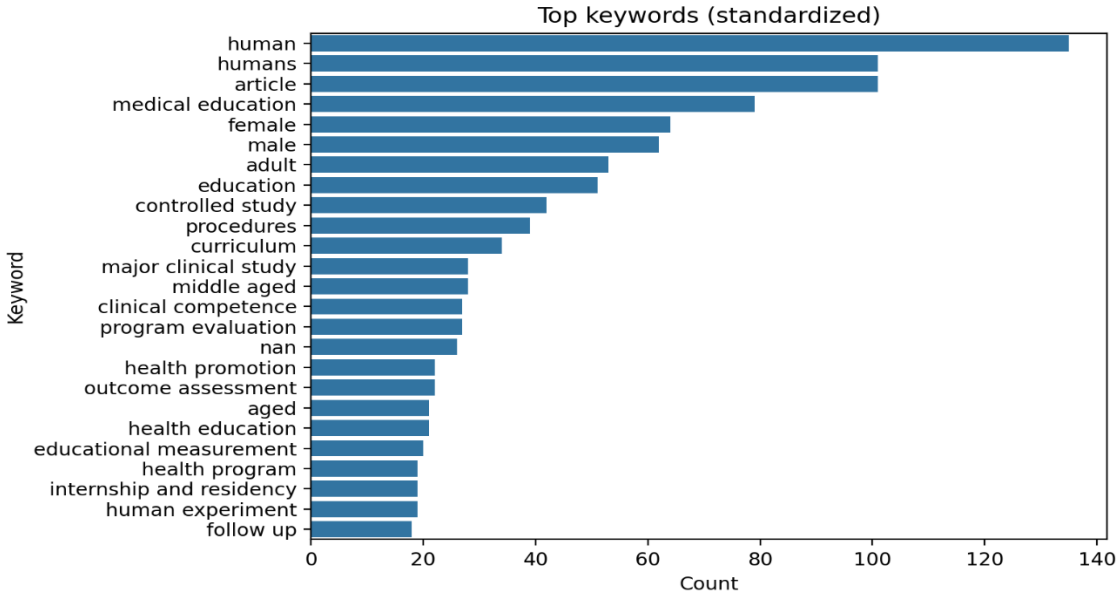


Figure 5. Emerging themes

The keyword frequency analysis from recent publications revealed distinct thematic patterns reflecting both the consolidation and evolution of research in programmatic assessment within medical education. The keyword "programmatic assessment" emerged as the most prevalent term (7 occurrences), affirming its centrality as the primary conceptual focus of the analyzed literature. Closely following were "training" (6 occurrences) and a cluster of contemporary and emerging topics—namely "artificial intelligence" (4), "ChatGPT" (4), "ophthalmology" (4), "feedback" (4), and "physician" (4)—indicating an increasing intersection between educational innovation, technology integration, and domain-specific applications (Figure 4,5).

The prominence of "artificial intelligence" and "ChatGPT" underscores a recent paradigm shift toward the exploration of AI-assisted assessment, data analytics, and feedback systems in medical education. These developments reflect the growing interest in leveraging digital tools to enhance formative assessment and personalized learning within programmatic frameworks. Simultaneously, the recurrence of discipline-oriented terms such as "ophthalmology" and "physician" suggests that programmatic assessment is being actively adapted and studied within specialized clinical training contexts.

Other frequently appearing keywords, including "undergraduate medical education," "competency-based education," "students," "student," "health survey," and "surveys and questionnaires" (each with three occurrences), point to sustained scholarly attention toward learner-centered outcomes, competency measurement, and empirical evaluation methods.

Journals and Citation Impact:

The journal-wise distribution of publications revealed a diverse range of sources contributing to the literature on programmatic assessment and related themes in medical education. A total of 78 publications were identified across 23 journals, with Health Technology Assessment emerging as the most productive outlet, publishing 13 articles and accumulating 648 citations, resulting in an average of approximately 49.85 citations per paper. This was followed by the Journal of Graduate Medical Education with seven publications and 115 citations (average 16.43 citations per paper) and AEM Education and Training with six publications and 44 citations (average 7.33 citations per paper). These findings indicate that while a small number of journals produce a higher volume of research outputs, citation impact varies considerably across publication venues. (Figure 6)

High-impact journals demonstrated markedly greater average citation counts despite fewer publications. Notably, Medical Education recorded the highest citation impact, with a single paper cited 873 times, followed by Ophthalmology Science (315 citations per paper) and the International Journal of Epidemiology (165 citations per paper). Similarly, British Journal of Ophthalmology (average 104 citations) and Advances in Health Sciences Education (average 92.5 citations) also showed substantial academic influence. This suggests that while such journals may not publish frequently on programmatic assessment, their contributions are highly influential and widely referenced within the scholarly community.

Mid-tier journals, such as BMC Public Health, American Journal of Pharmaceutical Education, BMJ Open, and International Journal of Medical Education, demonstrated consistent publication activity with moderate citation averages ranging between 12 and 22 citations per article. These outlets appear to play an essential role in disseminating applied and interdisciplinary research connecting educational practice, health promotion, and assessment frameworks. In contrast, journals such as BMC Medical Education (average 2.5 citations per paper) and JMIR Formative Research (average 4 citations per paper) showed lower citation averages, reflecting either their more recent publication timelines or more specialized readerships. (Figure 6)

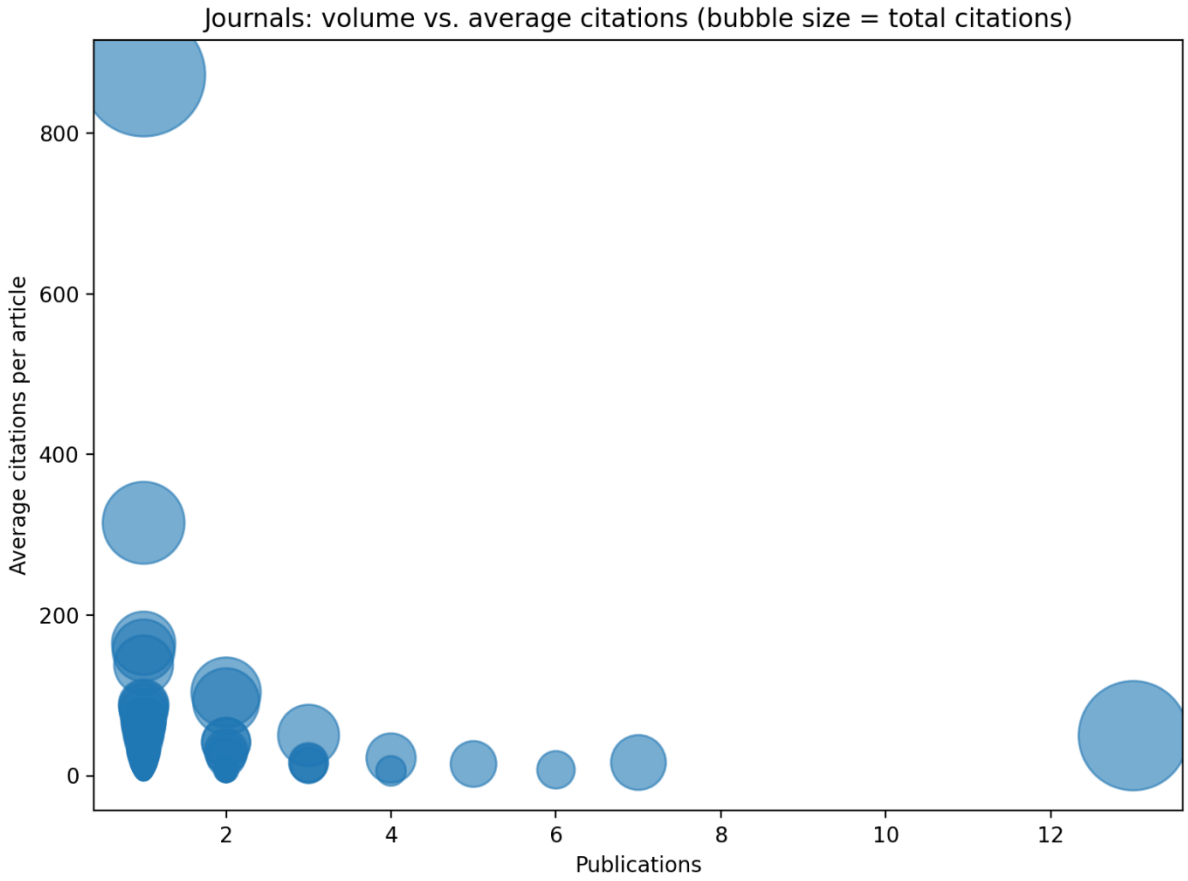


Figure 6. Journal volume and average citations

The analysis of recent publication trends by journal revealed a diverse but concentrated distribution of research output across prominent outlets in medical and health professions education. *AEM Education and Training* emerged as the most prolific journal in recent years, contributing six publications with a total of 44 citations, followed by *Education Sciences* with four publications and 24 citations. Other journals with multiple contributions included *Health Technology Assessment* (three publications, six citations), *BMC Public Health* (two publications, eight citations), *JMIR Formative Research* (two publications, eight citations), *Gerontology and Geriatrics Education* (two publications, seven citations), and *BMC Medical Education* (two publications, five citations). This distribution suggests that current research activity is concentrated within a select group of education-oriented and interdisciplinary health journals, reflecting both sustained interest and ongoing expansion of the field.

High-impact contributions, however, were not necessarily associated with publication volume. *Ophthalmology Science* recorded the highest citation count (315 citations from a single publication), followed by *British Journal of Ophthalmology* (49 citations) and *American Journal of Pharmaceutical Education* (21 citations). These findings indicate that while certain specialized or clinical journals publish fewer articles related to programmatic assessment, their individual contributions exert significant influence within their respective domains. Similarly, *Archives of Disease in Childhood* and *Public Health Nursing* each reported 17 citations from single papers, emphasizing the growing application of educational assessment frameworks in public health and pediatric contexts. A cluster of journals with moderate citation counts—such as *Journal of General Internal Medicine* (15 citations), *Journal of Postgraduate Medicine* (14), and *e-Clinical Medicine* (12)—demonstrates the broadening relevance of programmatic assessment research beyond traditional educational outlets into general medical and clinical research platforms. Meanwhile, emerging publication venues such as *JMIR Research Protocols*, *Medical Science Educator*,

and *Asian Bioethics Review* contributed smaller citation numbers but signify the diversification of scholarly discourse, particularly toward digital education, ethics, and methodological transparency.

Authorship and Collaboration

The analysis of author distribution per publication year revealed a steady and progressive increase in collaborative research within the field over time. In the early decades, publications were relatively sparse, with fewer contributing authors—typically between two and four per paper—from 1977 to the early 2000s. This pattern suggests a nascent stage of academic interest and limited collaborative engagement in the domain. Beginning in the mid-2000s, however, there was a noticeable expansion in both the number of publications and the average number of authors per year. For instance, from 2007 to 2015, several studies involved between six and ten authors per paper, reflecting a transition toward larger, multidisciplinary research teams and broader institutional cooperation.

From 2016 onwards, the trend toward multi-authorship became even more pronounced, indicating a maturation of the field and the establishment of cross-institutional collaborations. The years 2016 to 2020 consistently demonstrated author counts ranging from five to ten per publication, with multiple instances of maximum collaboration (ten authors). This escalation likely reflects the increasing complexity of research questions addressed, the adoption of international partnerships, and the growing emphasis on comprehensive, multi-perspective studies in health professions education.

The post-2020 period, particularly from 2021 to 2025, maintained this high level of scholarly collaboration, with the majority of publications including between six and ten co-authors. The year 2025 showed particularly strong activity, with several publications listing eight to ten authors, suggesting that programmatic assessment research has become both globally integrated and methodologically collaborative. This aligns with broader academic trends emphasizing interdisciplinary approaches, shared data interpretation, and collective authorship in response to complex educational challenges. (Figure 7)

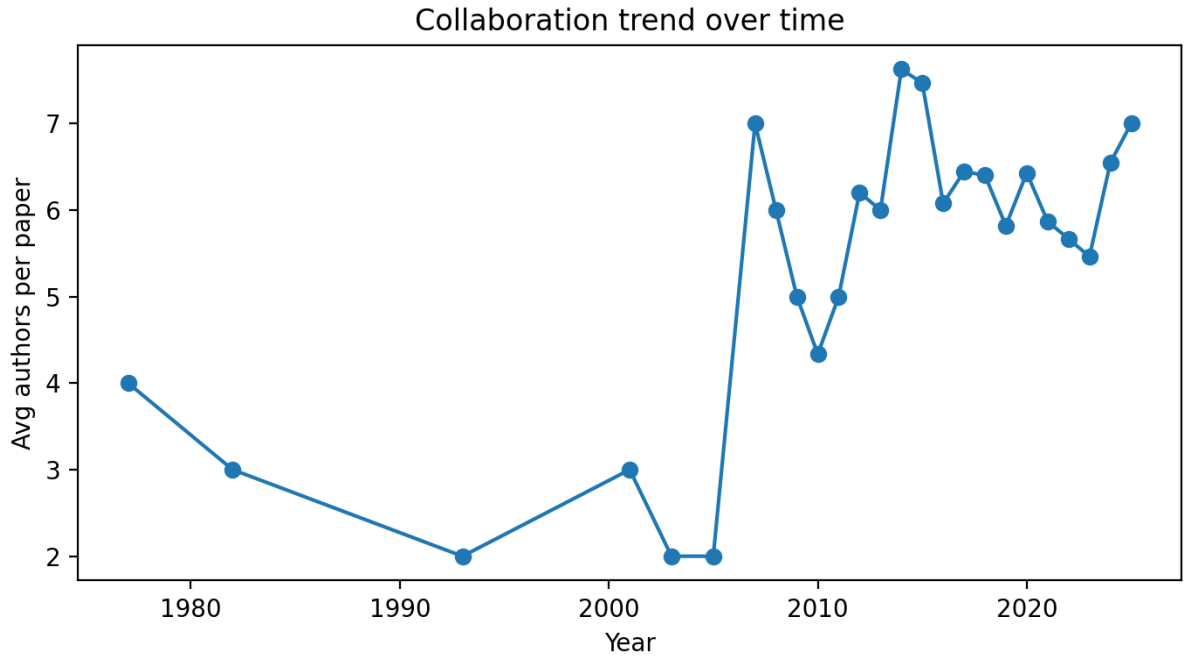


Figure 7. Collaboration trend over time

Synthesis of Bibliometric Findings

The synthesis of bibliometric findings revealed a concentrated yet globally distributed pattern of scholarly contribution in the field of programmatic assessment. Among the 1,092 identified authors, only seven met the inclusion threshold based on publication frequency and citation impact, indicating that a small core group of

researchers has driven much of the field's intellectual development. This concentration suggests the presence of key opinion leaders and collaborative research networks that have significantly influenced theoretical and methodological advancements in programmatic assessment scholarship.

Similarly, of the 502 contributing organizations, only five met the defined thresholds, highlighting the dominance of a limited number of highly productive academic institutions. These organizations likely serve as major research hubs, fostering interdisciplinary collaboration, methodological innovation, and the dissemination of programmatic assessment practices in medical education. The limited number of institutions exceeding the threshold also implies that while interest in the topic is widespread, sustained and high-impact research activity remains concentrated within a select few academic centers.

Geographically, of the 44 countries represented in the dataset, seven surpassed the inclusion criteria, underscoring the international scope of engagement with programmatic assessment research. This reflects an increasing globalization of educational assessment discourse, with notable contributions emerging from regions actively integrating competency-based education and formative assessment into medical curricula. Such cross-national participation signifies both the adaptability of the programmatic assessment framework and the growing recognition of its relevance in diverse educational contexts.

In terms of thematic content, out of 1,737 extracted keywords, 167 met the occurrence threshold, suggesting a rich yet focused body of research. The prominent clustering of keywords around concepts such as "feedback," "competency-based education," "learning outcomes," and "assessment design" indicates a well-established research agenda, while emerging terms—such as "artificial intelligence" and "digital assessment"—signal the evolution of the field toward technological integration and data-informed educational practices.

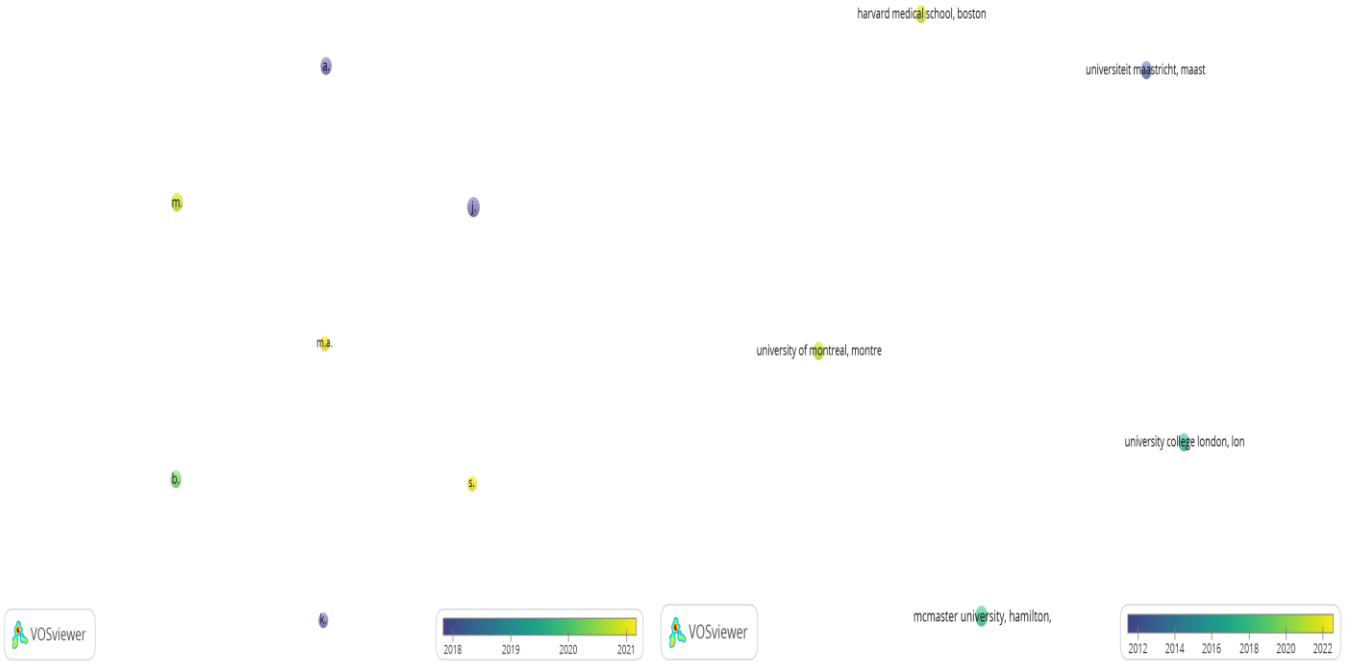


Figure 8. Network visualization

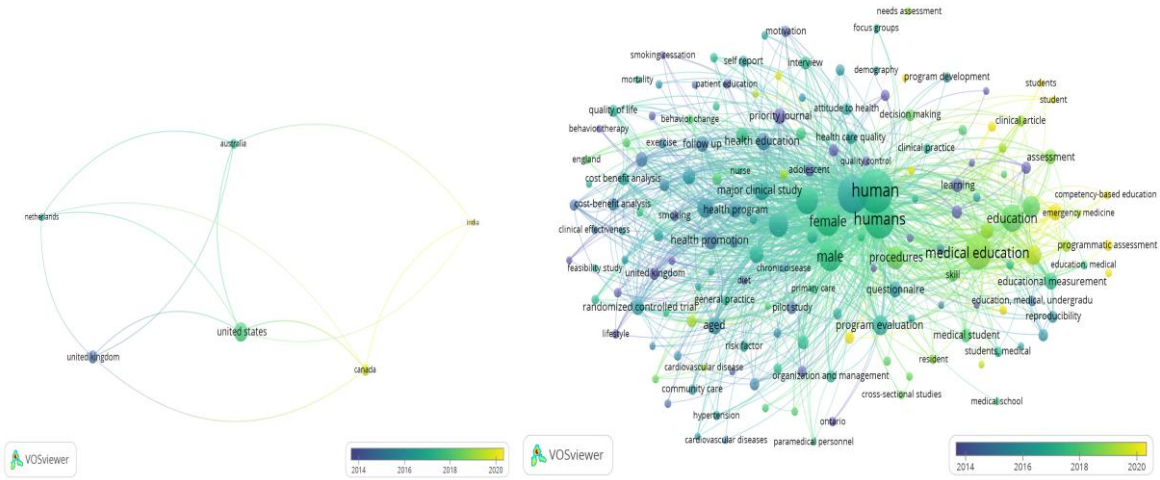


Figure 9. Keywords co-occurrence

The keyword co-occurrence analysis revealed several dominant thematic trends in recent publications related to programmatic assessment and medical education. The leading conceptual focus was "programmatic assessment" (7 occurrences), affirming its centrality as the primary research construct across the analyzed literature. Closely following were "training" (6 occurrences) and a cluster of emerging topics related to artificial intelligence (4), "ChatGPT" (4), "ophthalmology" (4), and "feedback" (4), demonstrating the growing intersection between educational technology, domain-specific applications, and formative assessment methodologies. (Figure 8, 9)

The prominence of keywords such as "artificial intelligence" and "ChatGPT" suggests a contemporary shift toward exploring the integration of AI-driven tools in assessment and feedback systems, reflecting the increasing adoption of generative technologies in health professions education. Meanwhile, the recurring appearance of "ophthalmology" indicates that programmatic assessment research is expanding into specialized clinical domains, highlighting discipline-specific adaptations of educational frameworks. Similarly, the presence of "feedback," "physician," and "training" underscores continued scholarly interest in competency development, longitudinal learning processes, and professional performance enhancement within structured assessment programs.

Additional keywords such as "undergraduate medical education," "competency-based education," "students," and "surveys and questionnaires" appeared three times each, illustrating a sustained emphasis on learner-centered approaches and empirical evaluation of educational outcomes. The frequent pairing of these terms implies that current studies often examine how programmatic assessment contributes to competency-based medical training and student learning experiences, particularly through the use of structured survey methodologies and reflective feedback mechanisms.

DISCUSSION

Bibliometric growth and publication trends

Our bibliometric analysis identified a clear acceleration in publications addressing programmatic assessment (PA) across health professions education over the last five years, with notable surges around consensus and implementation reports following the Ottawa 2020 meeting. Annual output increased substantially after 2020, and several special issues and methodological commentaries further concentrated literature on PA implementation and theory. These temporal patterns are consistent with prior syntheses and the Ottawa consensus outputs that framed PA principles and implementation priorities.(3, 4)

Core journals, geographic distribution, and influential works

Publications were concentrated in a small set of education and medical-teaching journals (e.g., *Medical Teacher*, *Perspectives on Medical Education*, *Education Sciences*), and originated predominantly from Europe, North America, and Australasia, with emerging contributions from East Asia and other regions. Influential conceptual and

review papers (including the Ottawa consensus statements and integrative reviews) and methodological guides (bibliometric and implementation guidance) were highly cited and functioned as intellectual hubs in co-citation networks (1,3,17).

Thematic clusters and keyword co-occurrence

Network and keyword analyses revealed distinct thematic clusters: (1) PA theory and principles (feedback, portfolios, progress committees); (2) implementation and faculty development (barriers, workload, feedback literacy); and (3) interfaces with competency domains such as communication, professionalism, and emergingly, health literacy and health promotion. While the PA–health-literacy/health-promotion linkage appears as an identifiable but smaller cluster, its presence across recent empirical and consensus documents suggests growing interest in using PA architectures to assess complex, communicative and population-health competencies (1,17).

Narrative synthesis: mechanisms, exemplars, and evidence strength

The narrative synthesis identified three recurrent mechanisms through which PA could support health literacy and health-promotion competencies: provision of repeated workplace-based practice with formative feedback (enabling deliberate practice), aggregation of multi-source narrative data to inform developmental decisions (allowing targeted remediation), and reflective portfolio tasks that cultivate metacognitive awareness for patient-centered communication. Empirical exemplars that explicitly link PA-like systems to health literacy or health-promotion outcomes are still limited but include recent curriculum-development and cohort-evaluation projects that map health-literacy competencies or health-promotion "passport" activities into longitudinal assessment frameworks. Overall, evidence tends toward feasibility, acceptability, and mechanism description rather than robust, multi-site outcome trials (19,20).

Study types, quality appraisal, and gaps

The corpus is dominated by descriptive case studies, single-institution programmed evaluations, qualitative investigations, and expert consensus statements; fewer controlled or comparative quantitative studies exist. Quality appraisals reflected common methodological limitations (small sample sizes, short follow-up, reliance on self-report and programmed metrics). Critical gaps include (a) validated outcome measures linking PA exposure to demonstrable improvements in learners' health-literacy practices or patient outcomes, (b) multi-site comparative studies that test different PA design choices for these competencies, and (c) implementation research that identifies context–mechanism–outcome configurations for successful scale-up. These gaps point to priorities for future research and for pragmatic evaluation of PA as a strategy to strengthen health literacy and health promotion competencies (5,20,21).

Interpretation of Key Findings

The combined bibliometric and narrative synthesis presented here indicates that programmatic assessment (PA) has consolidated as a prominent paradigm in health professions education, with accelerating publication activity and growing attention to implementation practice over the last five years. Our bibliometric mapping aligns with prior integrative and consensus work that characterized PA by its longitudinal data-collection logic, emphasis on low-stakes feedback, and aggregated high-stakes decision-making through expert judgement (1,3,22). These patterns help explain why PA is increasingly referenced as a suitable architecture for assessing complex, integrative competencies such as health literacy and health promotion—domains that require iterative practice, contextual judgement, and multi-source evidence rather than single-point testing (12,23).

Mechanistically, our narrative synthesis suggests three principal pathways by which PA may strengthen health literacy and health promotion competencies. First, PA's routine low-stakes encounters (workplace-based assessments, reflective portfolios, and formative entrustment decisions) create opportunities for deliberate practice and coached feedback on communication and health-promotion behaviors—activities central to health literacy responsiveness. Second, the aggregation of diverse assessment data and narrative feedback enables progress committees and mentors to identify longitudinal trajectories and to target remediation for recurring deficits in health-promotion skills. Third, PA's emphasis on reflective practice and learning portfolios fosters meta-cognitive growth in learners' ability to appraise and adapt their communicative strategies for diverse patient populations. These mechanisms resonate with

empirical and conceptual studies that identify feedback quality, mentoring, and curricular alignment as critical mediators of PA's educational impact (1,6,19).

Despite conceptual fit and plausible mechanisms, the empirical evidence directly linking PA to measurable improvements in health literacy outcomes remains emergent and heterogeneous. The literature we reviewed often reports favorable feasibility and acceptability indicators (e.g., richer feedback, better alignment with competency-based curricula) but stops short of demonstrating consistent, generalizable gains in patient-facing health literacy outcomes or population-level health promotion metrics. Many studies are single-institution case reports, programmed evaluations, or qualitative inquiries that describe processes and perceptions rather than robust controlled outcome measures. This evidence gap underscores the need for mixed-methods designs that combine PA implementation evaluation with objective assessments of learners' communication behaviors, patient understanding (e.g., teach-back performance), and downstream health-promotion activities (8,24,25).

Implementation realities create both opportunities and constraints for adopting PA specifically to advance health literacy and health promotion competencies. Important enablers include committed faculty development in high-quality feedback and mentoring, digital infrastructure to capture multi-source evidence, and institutional governance (progress committees) that can interpret aggregated data longitudinally. Conversely, reported barriers—including faculty workload, variable feedback literacy, inconsistent curricular integration, and cultural misalignment—may limit the potential benefits of PA if they are not explicitly considered in its design and scale-up strategies. Our findings therefore recommend that curriculum leaders pair PA adoption with targeted investments in faculty capacity-building (feedback and coaching), clear assessment blueprints linking tasks to health-literacy competencies, and iterative evaluation plans that monitor both process fidelity and learner outcomes (12,19).

Limitations and Cautions

This review has several limitations that should be acknowledged. First, although a comprehensive search strategy was employed across multiple databases, the inclusion was limited to articles published in English and within the last five years, which may have excluded relevant non-English or earlier foundational studies. Second, bibliometric analyses inherently depend on database indexing and citation metrics, which may underrepresent emerging scholarship from low- and middle-income countries or regional journals not indexed in Scopus or PubMed. Third, the narrative synthesis relied on the quality and heterogeneity of included studies, most of which were descriptive or qualitative in nature; hence, causal inferences between programmatic assessment (PA) implementation and measurable improvements in health literacy or health promotion competencies remain tentative. Moreover, grey literature, conference proceedings, and institutional reports were not systematically analyzed, potentially omitting valuable insights on practical implementation. Finally, the mapping between PA frameworks and health-literacy constructs remains conceptually exploratory, warranting further empirical validation and theoretical refinement through longitudinal, multi-institutional studies. Another limitation of this study is its reliance solely on open-access publications, which may introduce selection bias and reduce the comprehensiveness of the mapped literature.

Recommendations for Future Research

Finally, we articulate priorities for future research and policy. Empirical work should move beyond descriptive studies toward multi-site, comparative designs that evaluate PA models with pre-specified, validated measures of health literacy competence and health-promotion practice. Implementation research using realist or programmed-theory approaches can unpack context–mechanism–outcome configurations to show where PA succeeds or fails for these domains. Policymakers and accreditation bodies should consider endorsing assessment frameworks that explicitly include health literacy and health promotion as mapped outcomes within PA systems, while funders should support trials and longitudinal cohort studies that assess downstream patient and community effects. Taken together, the literature suggests that PA is offer a useful approach for supporting the development of health literacy and health promotion capabilities of future clinicians; deliberate design, implementation supports, and rigorous evaluation are needed to realize that promise as well as maintaining student's well-being (1,8,22,24-30).

CONCLUSION

This bibliometric and narrative review highlights that programmatic assessment has matured into a robust paradigm for integrating multiple assessment data points to foster reflective, feedback-driven, and competency-based learning in medical education. The findings reveal growing scholarly attention toward leveraging PA to strengthen complex competencies such as health literacy and health promotion, aligning with global educational goals for producing socially accountable and health-literate physicians. Despite the promise demonstrated in conceptual and pilot studies, empirical evidence linking PA implementation to demonstrable gains in learners' health-literacy practices or community-health outcomes remains limited. Future research should focus on developing validated assessment tools, outcome metrics, and multi-site interventions that capture the authentic impact of PA on learners' ability to communicate, advocate, and promote health effectively. Ultimately, embedding health-literacy and health-promotion competencies within a coherent programmatic-assessment framework may represent a strategic pathway toward more holistic, equitable, and socially responsive medical education systems.

AUTHOR'S CONTRIBUTION STATEMENT

All authors contributed equally to the conception, development, and preparation of this manuscript.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest related to this work.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

Generative AI tools were utilized solely to enhance the clarity and linguistic accuracy of the manuscript.

SOURCE OF FUNDING STATEMENTS

This research received no external financial support and was fully funded by the authors.

ACKNOWLEDGMENTS

The authors express their gratitude to Universitas Muhammadiyah Surakarta for providing support for this research.

BIBLIOGRAPHY

1. Schut, S., Driessen, E., van Tartwijk, J., van der Vleuten, C., & Heeneman, S. (2018). Stakes in the eye of the beholder: an international study of learners perceptions within programmatic assessment. Medical Education, 52(6), 654-663. doi:10.1111/medu.13532
2. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011;33(6):478-85. doi: 10.3109/0142159X.2011.565828. PMID: 21609177.
3. Heeneman S, de Jong LH, Dawson LJ, Wilkinson TJ, Ryan A, Tait GR, et al. Ottawa 2020 consensus statement for programmatic assessment–1. Agreement on the principles. Medical Teacher. 2021;43(10):1139-48.
4. Torre D, Rice NE, Ryan A, Bok H, Dawson LJ, Bierer B, et al. Ottawa 2020 consensus statements for programmatic assessment–2. Implementation and practice. Medical Teacher. 2021;43(10):1149-60.
5. Baartman L, van Schilt-Mol T, van der Vleuten C. Programmatic assessment design choices in nine programs in higher education. Frontiers in Education. 2022;7.
6. Govaerts M, Van der Vleuten C, Schut S. Implementation of Programmatic Assessment: Challenges and Lessons Learned. Education Sciences. 2022;12(10).
7. Roberts C, et al. Student perspectives on programmatic assessment in a large medical programme: A critical realist analysis. . Medical Education (2022). ;https://doi.org/10.1111/medu.14807.

8. Tsai HY, Lee, S.–Y. D., Coleman, C., Sørensen, K., et al. Health literacy competency requirements for health professionals: A Delphi consensus study in Taiwan. BMC Medical Education. (2024);24, Article 209. https://doi.org/10.1186/s12909-024-05198-4.
9. Stone M, Bazaldua, O., & Morrow, J. Developing health literacy communication practices for medical students. MedEdPORTAL,. (2021). ;Article 11091. https://doi.org/10.15766/mep_2374-8265.11091.
10. Magzoub ME. Going beyond competencies: Building blocks for a patient-centred medical curriculum. Medical Teacher, (2024). ;46(5). https://doi.org/10.1080/0142159X.2024.2412786. .
11. Blackford K, et al. A preliminary analysis of 'Passport to practice': approaches to assessing health promotion competencies in undergraduate curricula. Health Promotion International,. (2024). ; 2024. https://doi.org/10.1177/17579759241230065.
12. Torre D, Daniel M, Ratcliffe T, Durning SJ, Holmboe E, Schuwirth L. Programmatic Assessment of Clinical Reasoning: New Opportunities to Meet an Ongoing Challenge. Teach Learn Med. 2025 Jun-Jul;37(3):403-411. doi: 10.1080/10401334.2024.2333921..
13. Car J, Ong, Q. C., Erlikh Fox, T., et al. The Digital Health Competencies in Medical Education Framework: An international consensus statement based on a Delphi study. JAMA Network Open,. (2025). ; 8(1), e2453131. https://doi.org/10.1001/jamanetworkopen.2024.53131.
14. Dearnley C, Haigh, J., & Fairhall, J. Using mobile technologies for assessment and learning in practice settings: a case study. . Nurse Education in Practice, (2008). ; 8(3), 197-204.
15. Pelaccia T VR. Motivation in medical education. . Med Teach. 2017 39(2):136-140. doi: 10.1080/0142159X.2016.1248924. Epub 2016 Nov 21. PMID: 27866457.
16. Ahmed S, Waqar U , Lohana K.K , Khan D.A. A bibliometric analysis of the global research on COVID-19. National Journal of Health Sciences, (2020), 5, 119-126. DOI: 10.21089/njhs.53.0119.
17. Donthu N, Kumar, S., Mukherjee, D., et al. . How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285–296. (2021). ;https://doi.org/10.1016/j.jbusres.2021.04.070.
18. Schut S, Maggio LA, Heeneman S, van Tartwijk J, van der Vleuten C, Driessen E. 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.
19. Uijtdehaage S, Schuwirth LWT. Assuring the quality of programmatic assessment: Moving beyond psychometrics. Perspect Med Educ. 2018 Dec;7(6):350-351. doi: 10.1007/s40037-018-0485-y.
20. Govaerts MJB, van der Vleuten CPM, Holmboe ES. Managing tensions in assessment: moving beyond either-or thinking. Med Educ. 2019 Jan;53(1):64-75. doi: 10.1111/medu.13656.
21. Baartman, L. K. J., & Quinlan, K. M. (2024). Assessment and feedback in higher education reimagined: using programmatic assessment to transform higher education. Perspectives: Policy and Practice in Higher Education, 28(2), 5767. https://doi.org/10.1080/13603108.2023.2283118.
22. Schuwirth L.W.T., van der Vleuten C., Current Assessment in Medical Education: Programmatic Assessment. Journal of Applied Testing Technology, Vol 20(S2), 2-10, 2019.
23. Pedlow M. The key role of feedback in institutional readiness for programmatic assessment. . Assessment & Evaluation in Higher Education/Medical Education (2025). https://doi.org/10.1080/14703297.2025.2564775.
24. Greenfield J, & Qua, K. "It changed how I think"—Impact of programmatic assessment upon practicing physicians: A qualitative study. . Perspectives on Medical Education,. (2023). ;12(4), 301–310. https://pmc.ncbi.nlm.nih.gov/articles/PMC10403454/ PMC.
25. Schuwirth L., van der Vleuten C. Durning S.J., What programmatic assessment in medical education can learn from healthcare. Perspect Med Educ (2017) 6:211215. DOI 10.1007/s40037-017-0345-1.
26. Tandilangi TP, Rekawati, E., & Eriska, W. Improving Functional Status of Older Adults Through Health Literacy Media: A Systematic Review. Jurnal Berita Ilmu Keperawatan,. (2024) 17(2), 218–226. https://doi.org/10.23917/bik.v17i2.5063.
27. Ozturk FO A-AS. Health Literacy and Health Promotion Behaviors of Adolescents in Turkey. Journal of pediatric nursing. 2020 54:e31-e35. doi: 10.1016/j.pedn.2020.04.019. Epub 2020 May 20. PMID: 32446664.

28. Kusumaningrum T. Ky, Indriawan T., Widya M., Saputri, Pebrianti S., Liswanti A.,. Formation of Peer Educators in Information Dissemination Efforts to Prevent HIV Risky Behavior in Students,. Warta LPM. 2021, ;24:677. https://journals.ums.ac.id/index.php/warta/article/view/15035.
29. Herawati E, & Utami, L. Internet Addiction Causes Problems Emotional And Behavior In Adolescents. Biomedika 14(1), 74-80. (2022). ;DOI: 10.23917/biomedika.v14i1.17126.
30. Nurdiyanto F.A. HEP. The hidden impacts: Identifying psychological burdens during the Covid-19 pandemic,. Indigenous: Jurnal Ilmiah Psikologi, 7:130. 2022, ;DOI: https://doi.org/10.23917/indigenous.v7i2.17472.