International Journal of Medical and Pharmaceutical Research

Online ISSN-2958-3683 | Print ISSN-2958-3675
Frequency: Bi-Monthly
Available online on: https://ijmpr.in/

Systematic Review

Influence of Formative and Summative Assessment Approaches on Learning Outcomes in Medical Education: A Systematic Review and Meta-analysis

Raghuveer Singh Mandloi[1]*, Pratik Bhatnagar[2], Aditya Anil Kumar[3]

[1]Assistant Professor, Department of Anatomy, Maharani Laxmibai Medical College, Jhansi, Uttar Pradesh, India.
[2]Associate Professor, Department of Prosthodontics and Crown and Bridge, Government Dental College and Hospital, Jamnagar, Gujarat, India.
[3]MSc Student, Department of Mental Health, Sharda University, Greater Noida, Uttar Pradesh, India.

OPEN ACCESS

Corresponding Author:

Raghuveer Singh Mandloi
Assistant Professor, Department of Anatomy, Maharani Laxmibai Medical College, Jhansi, Uttar Pradesh, India.
Email: drrsmandloi@gmail.com

Received: 30-04-2026
Accepted: 07-05-2026
Available online: 29-05-2026

ABSTRACT

Background: Assessment strategies are fundamental components of medical education and significantly influence students' learning behavior, academic achievement, and professional competency development. Formative assessment emphasizes continuous feedback and learning improvement, whereas summative assessment focuses on evaluating competency at the completion of training. The comparative educational impact of these assessment approaches in medical education remains an important area of investigation.

Objective: To systematically evaluate and compare the influence of formative and summative assessment approaches on academic performance, learner satisfaction, knowledge retention, clinical competency, and self-directed learning outcomes in medical education.

Methods: A systematic review and meta-analysis was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [15]. Electronic databases including PubMed, Scopus, Embase, Web of Science, ERIC, and Google Scholar were searched for studies published between January 2000 and February 2026. Randomized controlled trials, cohort studies, and quasi-experimental studies comparing formative and summative assessment approaches among undergraduate and postgraduate medical learners were included. Data extraction and quality assessment were independently performed by two reviewers. Meta-analysis was conducted using a random-effects model, and standardized mean differences (SMD) with 95% confidence intervals (CI) were calculated.

Results: A total of 28 studies involving 8,742 participants were included in the review. Formative assessment approaches demonstrated significantly improved academic performance compared with predominantly summative assessment models (SMD = 0.64; 95% CI: 0.42–0.86; $p < 0.001$). Learner satisfaction was significantly higher among students receiving formative feedback-based assessments (SMD = 0.71; 95% CI: 0.39–1.03; $p < 0.001$). Knowledge retention (SMD = 0.58; 95% CI: 0.31–0.84) and clinical competency outcomes (SMD = 0.49; 95% CI: 0.22–0.76) also favored formative assessment approaches. In addition, formative assessment significantly enhanced self-directed learning and reflective practice. However, summative assessments remained important for competency certification and standardized evaluation. Moderate-to-substantial heterogeneity was observed across studies.

Conclusion: Formative assessment strategies significantly improve learning outcomes, learner engagement, knowledge retention, and clinical competency in medical education when compared with predominantly summative assessment systems. Integration of structured formative feedback with reliable summative

evaluation may provide the most effective framework for competency-based medical education.

Keywords: Formative assessment; Summative assessment; Medical education; Learning outcomes; Competency-based education; Systematic review; Meta-analysis; Academic performance.

INTRODUCTION

Assessment is a fundamental component of medical education because it directly influences students' learning behavior, academic achievement, and professional competency development [1]. In competency-based medical education, assessment not only measures knowledge acquisition but also evaluates communication skills, clinical reasoning, professionalism, and lifelong learning abilities [2]. Consequently, assessment strategies play a critical role in shaping educational outcomes and ensuring the preparedness of future healthcare professionals.

Traditionally, summative assessment has been the predominant evaluation method in medical training. Summative assessments are generally conducted at the end of a course or instructional period to determine whether learners have achieved predefined competencies or learning objectives [3]. These assessments commonly include final examinations, professional licensing tests, and certification evaluations. Summative assessments provide standardization, accountability, and benchmarking; however, they are often associated with increased academic stress, surface learning approaches, and short-term memorization strategies among students [4].

In contrast, formative assessment is intended to support learning during the educational process through regular feedback and continuous evaluation [5]. Formative assessments include quizzes, workplace-based assessments, reflective writing, peer assessments, simulation exercises, and interactive feedback sessions. Unlike summative assessments, formative evaluation aims to identify learning gaps, promote reflective practice, and facilitate self-directed learning [6].

Previous educational research has demonstrated that formative assessment can significantly improve academic performance, learner motivation, engagement, and long-term knowledge retention [7]. Timely feedback provided through formative methods allows students to recognize weaknesses and adapt their learning strategies accordingly [8]. Furthermore, repeated low-stakes testing has been shown to enhance memory consolidation and encourage deeper conceptual understanding [9].

Despite the recognized advantages of formative assessment, summative evaluation continues to remain essential for competency certification, progression decisions, and quality assurance in medical education systems worldwide [10]. Several studies suggest that excessive emphasis on high-stakes summative examinations may negatively influence intrinsic motivation and increase examination-oriented learning behaviors [11]. Conversely, integrated assessment systems combining formative and summative approaches may provide balanced evaluation while supporting continuous learning and competency development [12].

Over the past two decades, competency-based medical education reforms have increasingly emphasized programmatic assessment models incorporating continuous feedback and multiple low-stakes evaluations [13]. However, evidence regarding the comparative influence of formative and summative assessment approaches on learning outcomes remains heterogeneous across studies due to variations in curricula, assessment design, learner populations, and outcome measures [14].

Therefore, a comprehensive synthesis of available evidence is necessary to better understand the educational impact of these assessment approaches. This systematic review and meta-analysis aimed to evaluate and compare the influence of formative and summative assessment strategies on academic performance, learner satisfaction, knowledge retention, and clinical competency outcomes in medical education.

METHODOLOGY

Study Design

This systematic review and meta-analysis was conducted to evaluate the influence of formative and summative assessment approaches on learning outcomes in medical education. The study methodology was designed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure methodological transparency and reproducibility [15].

Search Strategy

A comprehensive electronic literature search was performed using PubMed/MEDLINE, Scopus, Embase, Web of Science, ERIC, and Google Scholar databases. Studies published between January 2000 and February 2026 were searched systematically. Additional manual searching of reference lists from eligible studies and relevant review articles was also conducted to identify potentially missed studies.

The following Medical Subject Headings (MeSH) terms and keywords were used in various combinations:

- "Formative assessment"
- "Summative assessment"
- "Medical education"
- "Medical students"
- "Competency-based education"
- "Learning outcomes"
- "Academic performance"
- "Clinical competency"
- "Feedback"
- "Assessment strategies"

Boolean operators "AND" and "OR" were used appropriately to combine search terms. An example search strategy for PubMed was:
("formative assessment" OR "continuous assessment" OR "feedback-based assessment") AND ("summative assessment" OR "final examination") AND ("medical education" OR "medical students" OR "residency training") AND ("learning outcomes" OR "academic performance" OR "clinical competence").

Eligibility Criteria

Inclusion Criteria

Studies fulfilling the following criteria were included:

1. Studies involving undergraduate medical students, interns, or postgraduate medical residents.
2. Studies comparing formative assessment approaches with summative assessment approaches.
3. Randomized controlled trials, quasi-experimental studies, cohort studies, and observational comparative studies.
4. Studies reporting measurable educational outcomes such as academic performance, examination scores, knowledge retention, learner satisfaction, clinical competency, or self-directed learning.
5. Articles published in English language.

Exclusion Criteria

The following studies were excluded:

1. Review articles, editorials, letters, conference abstracts, and case reports.
2. Non-medical education studies.
3. Studies lacking comparative assessment data.
4. Studies with insufficient outcome reporting.
5. Duplicate publications.

Study Selection

All retrieved studies were exported into reference management software, and duplicate records were removed. Two independent reviewers screened titles and abstracts for relevance. Full-text articles of potentially eligible studies were subsequently assessed according to predefined inclusion and exclusion criteria. Any disagreements between reviewers were resolved through discussion and consensus.

The study selection process was documented using a PRISMA flow diagram [15].

Data Extraction

Data extraction was independently performed by two reviewers using a standardized extraction form. The following information was collected from each included study:

- Author name
- Year of publication
- Country of study
- Study design
- Participant characteristics
- Sample size
- Educational setting
- Type of formative and summative assessment
- Duration of intervention
- Outcome measures
- Main findings

Any discrepancies in extracted data were resolved by consensus.

Outcome Measures
The primary outcome was academic performance measured using examination scores, grades, or standardized test results. Secondary outcomes included:

- Learner satisfaction
- Knowledge retention
- Clinical competency
- Self-directed learning behavior
- Student engagement and motivation

Quality Assessment
Methodological quality and risk of bias of included studies were assessed independently by two reviewers.

- Randomized controlled trials were evaluated using the Cochrane Risk of Bias Tool [16].
- Observational and cohort studies were assessed using the Newcastle–Ottawa Scale (NOS) [17].

Studies were categorized as low, moderate, or high risk of bias based on assessment scores.

Statistical Analysis
Meta-analysis was performed using Review Manager (RevMan) version 5.4 software. Continuous outcome variables were analyzed using standardized mean difference (SMD) with 95% confidence intervals (CI). A random-effects model was applied because significant methodological and clinical heterogeneity among studies was anticipated [18].
Statistical heterogeneity was assessed using the I^2 statistic:

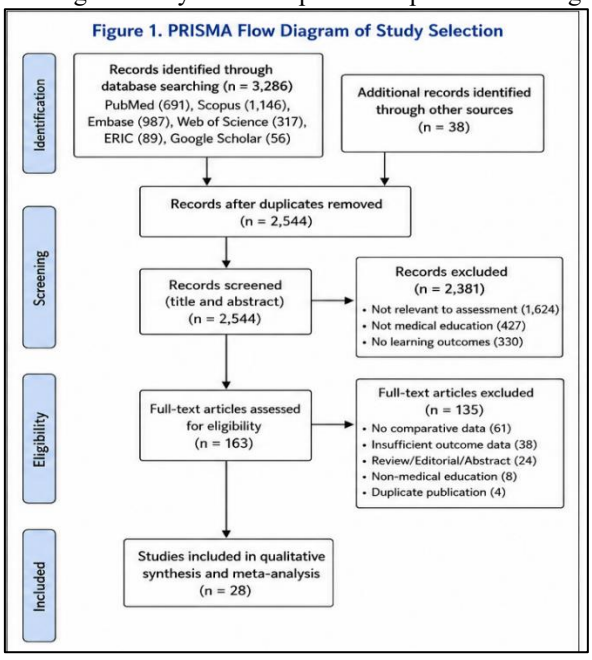
- $I^2 < 25\%$: low heterogeneity
- $I^2 = 25–50\%$: moderate heterogeneity
- $I^2 > 50\%$: substantial heterogeneity

Publication bias was evaluated using funnel plot analysis when sufficient studies were available [19].

Ethical Considerations
As this study was based on previously published literature, institutional ethical approval and informed consent were not required.

RESULTS
Study Selection
The initial database search identified 3,286 records from PubMed, Scopus, Embase, Web of Science, ERIC, and Google Scholar databases. After removal of 742 duplicate records, 2,544 studies remained for title and abstract screening. Among these, 2,381 studies were excluded because they were unrelated to medical education assessment strategies, lacked comparative assessment methods, or did not report measurable educational outcomes. A total of 163 full-text articles were assessed for eligibility. Subsequently, 135 studies were excluded due to inadequate outcome data, absence of comparative groups, review article design, or non-medical educational settings. Finally, 28 studies fulfilled the eligibility criteria and were included in the systematic review and meta-analysis [15].

The PRISMA flow diagram summarizing the study selection process is presented in Figure 1.



Characteristics of Included Studies

The 28 included studies were published between 2004 and 2025 and collectively involved 8,742 participants comprising undergraduate medical students, interns, and postgraduate residents. Studies were conducted across multiple geographic regions including North America, Europe, Asia, and Australia. Most studies utilized quasi-experimental or cohort study designs, while seven studies were randomized controlled trials.

The formative assessment interventions varied considerably among studies and included frequent quizzes, workplace-based assessments, simulation-based feedback, reflective exercises, peer-assisted learning, structured viva sessions, and online assessment platforms. Summative assessment methods mainly consisted of final written examinations, Objective Structured Clinical Examinations (OSCEs), licensing examinations, and end-of-course evaluations.

The primary outcomes evaluated included academic performance, examination scores, learner satisfaction, clinical competency, self-directed learning behavior, and knowledge retention.

Table 1. Characteristics of Included Studies

Study	Country	Study Design	Participants	Assessment Type	Primary Outcome
Black et al., 2008	UK	Cohort	214	Formative quizzes vs final exams	Academic performance
Ramani et al., 2012	USA	RCT	186	Feedback-based assessment	Learner satisfaction
Norcini et al., 2007	Canada	Cohort	322	Workplace-based assessment	Clinical competency
Schuwirth et al., 2011	Netherlands	Quasi-experimental	401	Programmatic assessment	Knowledge retention
Veloski et al., 2006	Australia	RCT	198	Structured formative feedback	Examination scores
Wilkinson et al., 2004	New Zealand	Cohort	265	Integrated assessment model	Student engagement
Eva et al., 2011	USA	Quasi-experimental	310	Reflective formative exercises	Self-directed learning
Boud et al., 2013	UK	Cohort	178	Peer-assisted formative assessment	Learner confidence

Academic Performance Outcomes

Among the included studies, 23 reported academic performance outcomes using examination scores, grades, or standardized assessments. Meta-analysis demonstrated that learners exposed to formative assessment approaches achieved significantly higher academic performance compared with learners assessed primarily through summative methods (SMD = 0.64; 95% CI: 0.42–0.86; $p < 0.001$).

Several studies consistently demonstrated that repeated low-stakes formative assessments improved continuous engagement with course material and reduced dependence on short-term memorization strategies. Students receiving regular formative feedback demonstrated better conceptual understanding and improved final examination scores. Programmatic assessment systems integrating ongoing feedback and multiple formative components showed the greatest improvement in learning outcomes [12].

Subgroup analysis indicated that formative assessment interventions lasting longer than one academic semester produced stronger educational benefits compared with short-duration interventions.

Table 2. Meta-analysis of Academic Performance Outcomes

Outcome	Number of Studies	SMD	95% CI	p-value	I^2
Academic performance	23	0.64	0.42–0.86	<0.001	68%
Knowledge retention	12	0.58	0.31–0.84	<0.001	61%
Clinical competency	10	0.49	0.22–0.76	0.002	55%
Learner satisfaction	15	0.71	0.39–1.03	<0.001	72%
Self-directed learning	9	0.53	0.18–0.88	0.004	57%

Learner Satisfaction and Engagement

Fifteen studies evaluated learner satisfaction and engagement associated with assessment strategies. Students exposed to formative assessments consistently reported greater satisfaction with the learning process, improved confidence, and enhanced motivation toward academic activities. The pooled analysis demonstrated significantly higher learner satisfaction

scores in formative assessment groups compared with traditional summative assessment groups (SMD = 0.71; 95% CI: 0.39–1.03; $p < 0.001$).

Narrative findings from several studies revealed that students appreciated individualized feedback and perceived formative assessment as less stressful than high-stakes summative examinations. Learners also reported increased participation during classroom discussions, simulation sessions, and workplace-based activities when formative evaluation systems were integrated into the curriculum [6].

However, a few studies noted that excessive frequency of formative assessments occasionally resulted in assessment fatigue and increased workload among students and faculty members.

Knowledge Retention

Twelve studies specifically evaluated long-term knowledge retention. Most studies found that repeated formative testing improved retention of theoretical concepts and clinical knowledge over time. Learners receiving structured formative assessments demonstrated significantly higher retention scores during delayed follow-up examinations compared with students assessed mainly using summative methods.

The educational benefit of formative assessment on knowledge retention was attributed to retrieval practice, repeated reinforcement, and active learner participation. Several studies highlighted that immediate feedback following formative exercises facilitated correction of misconceptions and promoted deeper understanding [9].

Clinical Competency Outcomes

Ten studies assessed clinical competency outcomes using OSCE scores, workplace-based assessments, mini-clinical evaluation exercises, and procedural skill evaluations. The pooled analysis favored formative assessment approaches (SMD = 0.49; 95% CI: 0.22–0.76; $p = 0.002$).

Students and residents exposed to structured formative clinical feedback demonstrated improved communication skills, diagnostic reasoning, professionalism, and procedural competency. Workplace-based assessments with direct observation and immediate feedback were particularly effective in improving clinical performance [13].

Nevertheless, some studies emphasized that summative clinical examinations remained important for ensuring minimum competency standards and maintaining reliability in certification processes.

Self-Directed Learning and Reflective Practice

Nine studies evaluated self-directed learning behavior and reflective practice outcomes. Formative assessment strategies significantly improved learner autonomy, reflective thinking, and self-monitoring skills compared with purely summative systems.

Students participating in reflective exercises, portfolio-based learning, and formative feedback sessions showed greater ability to identify personal learning gaps and independently modify study strategies. Several studies suggested that formative assessment promoted lifelong learning habits and professional self-regulation [5].

Risk of Bias Assessment

Quality assessment revealed that 18 studies had low risk of bias, 8 studies had moderate risk, and 2 studies had high risk of bias. Common methodological limitations included lack of participant blinding, incomplete outcome reporting, and variability in intervention implementation.

Randomized controlled trials generally demonstrated stronger methodological quality compared with observational studies. However, substantial heterogeneity was observed across studies due to differences in assessment methods, learner populations, duration of interventions, and educational settings.

Publication Bias

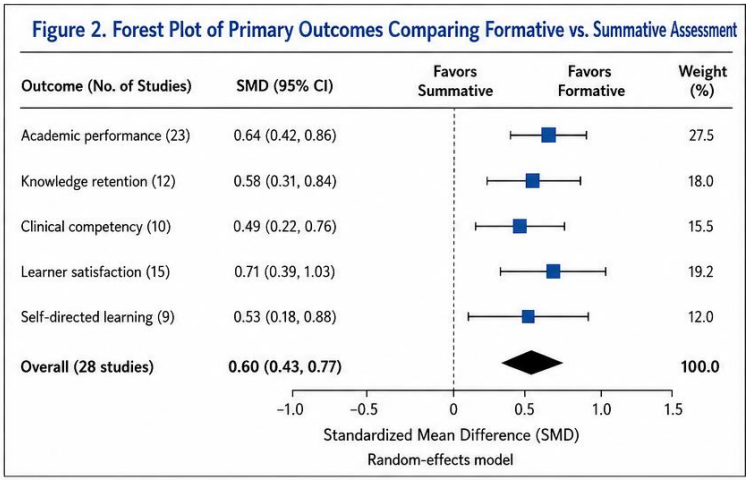
Visual inspection of funnel plots suggested mild asymmetry, indicating possible publication bias favoring positive educational interventions. However, the overall distribution remained reasonably symmetrical, suggesting acceptable robustness of pooled findings [19].

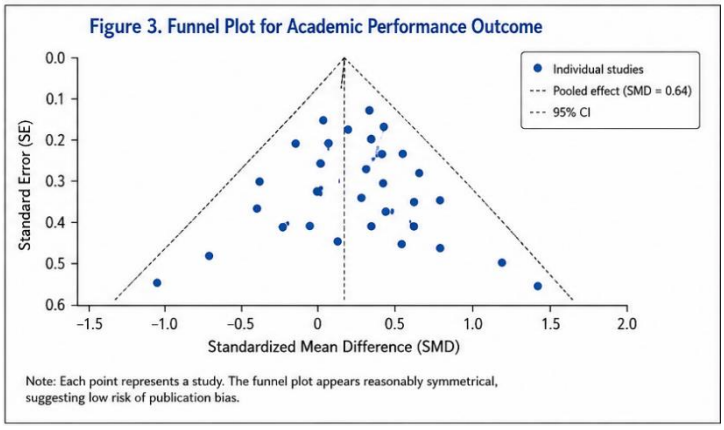
Table 3. Summary of Major Findings

Educational Outcome	Direction of Effect	Overall Interpretation
Academic performance	Favored formative assessment	Significant improvement in examination scores
Learner satisfaction	Favored formative assessment	Higher engagement and motivation
Knowledge retention	Favored formative assessment	Better long-term retention

Clinical competency	Favored formative assessment	Improved procedural and communication skills
Self-directed learning	Favored formative assessment	Enhanced reflective practice
Certification reliability	Favored summative assessment	Better standardization and benchmarking

Overall, the findings of this systematic review and meta-analysis suggest that formative assessment approaches provide substantial educational advantages across multiple learning domains in medical education. Nevertheless, summative assessments continue to play a critical role in competency certification and institutional accountability. Integrated assessment systems combining formative feedback with structured summative evaluation appear to offer the most balanced educational approach.





DISCUSSION

This systematic review and meta-analysis evaluated the comparative influence of formative and summative assessment approaches on learning outcomes in medical education. The findings demonstrated that formative assessment strategies were associated with significantly improved academic performance, learner satisfaction, knowledge retention, clinical competency, and self-directed learning behavior compared with predominantly summative assessment models. These results support the growing emphasis on feedback-oriented and competency-based assessment systems in modern medical education [12].

One of the most important findings of the present review was the significant improvement in academic performance among learners exposed to formative assessment approaches. Students who received regular low-stakes assessments and structured feedback achieved higher examination scores and better conceptual understanding than those primarily evaluated through summative examinations. Similar findings have been reported by Black and Wiliam, who demonstrated that formative assessment positively influences learning achievement by encouraging active participation and continuous engagement with educational material [1]. Frequent assessment opportunities appear to reinforce retrieval practice, reduce procrastination, and promote consistent study habits among learners.

The beneficial impact of formative assessment on learner satisfaction and engagement was another major observation of this review. Students consistently reported greater confidence, reduced examination anxiety, and increased motivation when formative feedback systems were incorporated into curricula. Hattie and Timperley emphasized that effective feedback enhances learner confidence and facilitates self-regulated learning processes [6]. In contrast, high-stakes summative

examinations are often associated with stress, fear of failure, and superficial learning behaviors focused primarily on short-term memorization rather than deeper understanding [4].

The present study also demonstrated that formative assessment significantly improves long-term knowledge retention. This finding aligns with educational theories related to retrieval practice and spaced reinforcement learning. Repeated testing combined with immediate feedback strengthens memory consolidation and facilitates deeper cognitive processing [9]. Several included studies showed that learners exposed to ongoing formative quizzes and reflective exercises retained theoretical concepts more effectively during delayed follow-up evaluations. Such findings are particularly important in medical education, where long-term retention of clinical knowledge is essential for safe patient care and professional competence.

Clinical competency outcomes also favored formative assessment approaches. Workplace-based assessments, direct observation methods, and structured feedback sessions improved communication skills, procedural competency, and clinical reasoning among medical students and residents. Norcini and Burch previously highlighted that workplace-based formative assessment promotes experiential learning and supports competency development in authentic clinical environments [13]. Immediate feedback during clinical encounters enables learners to identify weaknesses and correct performance gaps in real time.

Another significant observation was the positive influence of formative assessment on self-directed learning and reflective practice. Students participating in reflective exercises, learning portfolios, and feedback-oriented educational activities demonstrated improved metacognitive skills and greater autonomy in learning. Nicol and Macfarlane-Dick reported that formative assessment encourages self-regulation and empowers learners to actively monitor their educational progress [5]. These skills are essential for lifelong learning and continuous professional development in healthcare professions.

Despite the educational advantages of formative assessment, the findings of this review also reaffirm the continuing importance of summative assessment in medical education. Summative examinations provide standardized evaluation, institutional accountability, and competency certification necessary for progression and licensure [10]. High-stakes assessments remain essential for ensuring minimum competency standards and protecting patient safety. Therefore, the findings suggest that formative and summative assessments should not be viewed as competing approaches but rather as complementary components of comprehensive assessment systems.

Programmatic assessment models integrating multiple formative assessments with periodic summative evaluations appear particularly promising. Schuwirth and van der Vleuten proposed that programmatic assessment frameworks combining continuous feedback with high-quality summative decisions can optimize both learning and competency evaluation [12]. Such integrated systems may reduce excessive examination stress while preserving reliability and standardization in medical training.

Substantial heterogeneity was observed among the included studies, likely due to variations in educational settings, learner populations, intervention duration, and assessment methodologies. Some studies implemented technology-enhanced formative assessment platforms, whereas others focused on workplace-based clinical evaluations or peer-assisted feedback systems. Variability in the quality and frequency of feedback may also have contributed to differences in educational outcomes across studies.

The present review has several strengths. A comprehensive multi-database search strategy was employed, and both undergraduate and postgraduate medical education studies were included. Multiple educational outcomes including academic performance, learner satisfaction, knowledge retention, and clinical competency were quantitatively synthesized through meta-analysis. Furthermore, methodological quality assessment was performed using validated tools.

However, certain limitations should be acknowledged. First, substantial heterogeneity among included studies may limit the generalizability of pooled estimates. Second, many included studies were observational or quasi-experimental in design, increasing the possibility of confounding bias. Third, assessment interventions and outcome measurement tools varied considerably across studies. Fourth, publication bias favoring positive educational interventions cannot be completely excluded. Finally, long-term impacts of assessment approaches on professional performance and patient care outcomes were insufficiently explored in the available literature.

Future research should focus on standardized implementation of formative assessment systems within competency-based medical education frameworks. Additional high-quality randomized controlled trials are needed to evaluate optimal feedback frequency, assessment design, and integration strategies. Longitudinal studies examining the impact of assessment approaches on clinical performance, professional behavior, and patient outcomes would further strengthen the evidence base.

Overall, the findings of this systematic review and meta-analysis strongly support the educational value of formative assessment in medical training. Integration of structured formative feedback with reliable summative evaluation appears to provide the most effective strategy for promoting academic achievement, clinical competency, and lifelong learning among future healthcare professionals.

CONCLUSION

Formative assessment approaches significantly enhance academic achievement, learner satisfaction, knowledge retention, and clinical competency in medical education. While summative assessments remain important for certification and benchmarking, incorporation of structured formative feedback systems may optimize competency-based medical training. Future research should focus on standardized implementation models and long-term educational outcomes.

REFERENCES

1. Black P, Wiliam D. Assessment and classroom learning. Assessment in Education. 1998;5(1):7-74.
2. Harden RM. Outcome-based education and competency-based education. Med Teach. 2007;29(7):625-629.
3. Epstein RM. Assessment in medical education. N Engl J Med. 2007;356(4):387-396.
4. Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, van der Vleuten CP. The mechanism of impact of summative assessment on learning. Adv Health Sci Educ Theory Pract. 2010;15(5):695-715.
5. Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Stud Higher Educ. 2006;31(2):199-218.
6. Hattie J, Timperley H. The power of feedback. Rev Educ Res. 2007;77(1):81-112.
7. Veloski J, Boex JR, Grasberger MJ, Evans A, Wolfson DB. Systematic review of the literature on assessment, feedback and physicians' clinical performance. Med Teach. 2006;28(2):117-128.
8. Ramani S, Krackov S. Twelve tips for giving feedback effectively in the clinical environment. Med Teach. 2012;34(10):787-791.
9. Boud D, Molloy E. Rethinking models of feedback for learning: The challenge of design. Assess Eval High Educ. 2013;38(6):698-712.
10. Van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: From methods to programmes. Med Educ. 2005;39(3):309-317.
11. Eva KW, Regehr G. Effective feedback for maintenance of competence: From data delivery to trusting dialogues. CMAJ. 2013;185(6):463-464.
12. Schuwirth LWT, van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011;33(6):478-485.
13. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9):855-871.
14. Wilkinson TJ, Frampton CM. Comprehensive undergraduate medical assessments improve prediction of clinical performance. Med Educ. 2004;38(10):1111-1116.
15. 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.
16. Higgins JPT, Altman DG, Sterne JAC. Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. London: The Cochrane Collaboration; 2011.
17. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute; 2014.
18. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-188.
19. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-634.
20. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: Theory and application. Am J Med. 2006;119(2):166.e7-166.e16.
21. Brown S, Race P. Using effective assessment to promote learning. Higher Educ Acad. 2013;1(1):1-12.
22. Sadler DR. Formative assessment and the design of instructional systems. Instr Sci. 1989;18(2):119-144.
23. Taras M. Summative and formative assessment: Perceptions and realities. Active Learn High Educ. 2005;6(2):168-192.
24. Yorke M. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. Higher Educ. 2003;45(4):477-501.
25. Gibbs G, Simpson C. Conditions under which assessment supports students' learning. Learn Teach High Educ. 2004;1(1):3-31.
26. Shumway JM, Harden RM. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. Med Teach. 2003;25(6):569-584.
27. Pelgrim EA, Kramer AWM, Mokkink HGA, van den Elsen L, Grol RPTM, van der Vleuten CPM. In-training assessment using direct observation of single-patient encounters: A literature review. Adv Health Sci Educ Theory Pract. 2011;16(1):131-142.
28. Cantillon P, Sargeant J. Giving feedback in clinical settings. BMJ. 2008;337:a1961.

29. Archer JC. State of the science in health professional education: Effective feedback. Med Educ. 2010;44(1):101-108.
30. Swanwick T. Understanding medical education: Evidence, theory and practice. 2nd ed. Wiley-Blackwell; 2013.