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

Educational Outcomes of Formative Compared with Summative Assessment in Medical Students and Residents: A Systematic Review and Meta-analysis

Imran Sabri*[1], Mukul Singh[2], Suvendu Kumar Panda[3]

[1]Associate Professor, Pacific Institute of Medical Sciences (PIMS), Sai Tirupati University, Umarda, Udaipur, Rajasthan, India.
[2]Associate Professor, Department of General Surgery, All India Institute of Medical Sciences, Gorakhpur, Uttar Pradesh, India.
[3]Assistant Professor, Department of Pharmacology, MKCG Medical College and Hospital, Berhampur, Odisha, India.

OPEN ACCESS

Corresponding Author:

Dr. Imran Sabri
Associate Professor, Pacific Institute of Medical Sciences (PIMS), Sai Tirupati University, Umarda, Udaipur, Rajasthan, India.
Email:
drimransabri@gmail.com

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

ABSTRACT

Background: Assessment strategies significantly influence learning behavior, competency acquisition, and academic achievement in medical education. While summative assessment has traditionally been used for certification and competency evaluation, formative assessment emphasizes continuous feedback, reflective learning, and progressive improvement. With the increasing implementation of competency-based medical education, understanding the comparative effectiveness of formative and summative assessment has become increasingly important.

Objective: To systematically review and meta-analyze the educational outcomes associated with formative compared with summative assessment among medical students and residents.

Materials and Methods: A systematic review and meta-analysis were conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Electronic databases including PubMed, Scopus, Embase, Web of Science, ERIC, and Google Scholar were searched for studies published between January 2000 and January 2026. Studies involving undergraduate medical students or postgraduate residents comparing formative and summative assessment approaches were included. Primary outcomes included academic performance, learner engagement, knowledge retention, feedback utilization, clinical competency, and learner satisfaction. Methodological quality assessment was performed using standardized appraisal tools. Quantitative meta-analysis was conducted using a random-effects model.

Results: A total of 42 studies involving 18,764 participants met the inclusion criteria. Twenty-seven studies involved undergraduate medical students, while fifteen included residents or postgraduate trainees. Narrative synthesis demonstrated that formative assessment was associated with improved learner engagement, knowledge retention, self-directed learning, feedback utilization, and clinical competency development. Meta-analysis of 15 studies including 2,033 participants demonstrated significantly improved academic performance in the formative assessment group compared with the summative assessment group (MD = 7.54; 95% CI: 6.40–8.68; $p < 0.001$). Moderate heterogeneity was observed ($I^2 = 46\%$). Workplace-based formative assessment tools such as mini-CEX, DOPS, and feedback-oriented OSCEs showed particularly strong positive effects on competency development. Funnel plot analysis demonstrated low to moderate publication bias.

Conclusion: Formative assessment provides significant educational advantages in medical education by enhancing academic performance, learner engagement, reflective learning, and competency development. Although summative assessment remains essential for certification and accountability, integrated hybrid assessment systems combining formative and summative approaches appear most effective for optimizing educational outcomes among medical students and residents.

Keywords: Formative assessment; Summative assessment; Medical education; Medical students; Residents; Competency-based medical education; Meta-analysis; Educational outcomes.

INTRODUCTION

Assessment is a cornerstone of medical education and plays a critical role in determining learner competence, guiding educational strategies, and influencing professional development [1]. In health professions education, assessment not only measures academic achievement but also shapes the way students learn, retain knowledge, and develop clinical skills. The widely accepted principle that "assessment drives learning" highlights the profound influence of assessment methods on learners' study behaviors, motivation, and engagement with educational content [1,2].

Traditionally, summative assessment has been the dominant approach in medical education systems worldwide [3]. Summative assessments are typically conducted at the end of a course, clerkship, or training period and are intended to evaluate whether learners have achieved predetermined educational objectives or competency standards [3]. Examples include final theory examinations, licensing examinations, end-of-posting assessments, and board certification evaluations. These assessments provide standardized measures for progression, ranking, and certification, thereby ensuring accountability and public trust in medical training programs [3,5]. However, concerns have been raised that an excessive emphasis on summative assessment may encourage superficial learning strategies, short-term memorization, and performance-oriented behaviors rather than meaningful understanding and lifelong learning [2,8].

In contrast, formative assessment is designed to support learning during the educational process through continuous evaluation and constructive feedback [2,4]. Formative assessment strategies include quizzes, reflective writing, workplace-based assessments, Objective Structured Clinical Examinations (OSCEs) with feedback, mini-Clinical Evaluation Exercises (mini-CEX), Direct Observation of Procedural Skills (DOPS), and portfolio assessments [4,5]. Unlike summative assessment, formative assessment aims to identify learning gaps, provide individualized feedback, and facilitate ongoing improvement [6]. Educational theories suggest that formative assessment promotes self-regulated learning, critical thinking, reflection, and deeper understanding of concepts [2,6].

The growing adoption of competency-based medical education (CBME) has further increased interest in formative assessment approaches [10]. CBME emphasizes learner-centered progression, continuous competency development, and achievement of measurable outcomes rather than time-based training alone [10]. Within this framework, frequent formative assessment and timely feedback are considered essential for monitoring learner progress and supporting individualized educational development [5,10]. Consequently, many medical schools and residency programs are increasingly integrating formative methods into traditional assessment systems.

Feedback represents one of the most valuable components of formative assessment [6,7]. Effective feedback enables learners to recognize deficiencies, reinforce strengths, and improve future performance [7]. Studies have demonstrated that high-quality feedback contributes to enhanced academic achievement, improved clinical reasoning, greater learner confidence, and stronger professional identity formation [6,7]. In contrast, summative assessments often provide limited or delayed feedback, reducing opportunities for meaningful educational improvement [9].

Despite increasing implementation of formative assessment strategies, summative assessment continues to remain essential for certification, progression decisions, and quality assurance in medical education [3]. Therefore, educators face the challenge of balancing formative and summative approaches to maximize educational effectiveness while maintaining accountability and reliability [5]. Several studies have compared the educational impact of these assessment methods; however, findings remain heterogeneous across undergraduate and postgraduate medical training environments.

Furthermore, modern medical education requires learners not only to acquire theoretical knowledge but also to develop communication skills, professionalism, clinical competence, teamwork, and reflective practice [4,10]. Assessment systems capable of promoting these multidimensional competencies are increasingly necessary in contemporary healthcare education. Understanding the comparative educational outcomes associated with formative and summative assessment can therefore guide curriculum planning, faculty development, and assessment reform.

Although previous literature reviews have examined assessment practices in higher education, evidence specifically focusing on medical students and residents remains fragmented [1,8]. Medical training involves unique clinical responsibilities, patient safety considerations, and competency expectations that necessitate focused evaluation of assessment strategies within this context [3,5].

Accordingly, the present systematic review was conducted to evaluate and compare the educational outcomes associated with formative and summative assessment approaches among medical students and residents. The review aims to

synthesize available evidence regarding academic performance, learner engagement, feedback utilization, competency development, and learner perceptions to inform evidence-based assessment practices in medical education.

MATERIALS AND METHODS

Study Design and Protocol Registration

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [1]. The review aimed to compare the educational outcomes associated with formative and summative assessment among undergraduate medical students and postgraduate residents.

Search Strategy

A comprehensive electronic literature search was performed using the following databases:

- PubMed/MEDLINE
- Scopus
- Embase
- Web of Science
- ERIC (Education Resources Information Center)
- Google Scholar

The search covered studies published between January 2000 and March 2026. Manual searches of reference lists from relevant articles and review papers were also performed to identify additional eligible studies.

Search terms included combinations of Medical Subject Headings (MeSH) and free-text keywords related to:

- "formative assessment"
- "summative assessment"
- "medical education"
- "medical students"
- "residents"
- "competency-based education"
- "educational outcomes"
- "academic performance"
- "feedback"
- "clinical competency"

Boolean operators ("AND," "OR") were used to optimize the search strategy.

Example Search String

("formative assessment" OR "continuous assessment" OR "feedback-based assessment") AND ("summative assessment" OR "high-stakes examination") AND ("medical students" OR "residents" OR "postgraduate trainees") AND ("medical education") AND ("educational outcomes" OR "academic performance" OR "clinical competency")

Eligibility Criteria

Inclusion Criteria

Studies were included if they:

1. Involved undergraduate medical students or postgraduate residents.
2. Compared formative assessment with summative assessment directly or indirectly.
3. Reported measurable educational outcomes.
4. Used quantitative, qualitative, or mixed-method study designs.
5. Were published in peer-reviewed English-language journals.

Exclusion Criteria

Studies were excluded if they:

1. Included only non-medical healthcare disciplines without separate medical subgroup analysis.
2. Were editorials, narrative reviews, conference abstracts, letters, or commentaries.
3. Did not report educational outcome measures.
4. Were duplicate publications or incomplete datasets.
5. Focused solely on psychometric validation without educational outcome assessment.

Study Selection

All retrieved studies were imported into reference management software, and duplicate records were removed. Two independent reviewers screened titles and abstracts according to predefined eligibility criteria [1]. Potentially relevant studies underwent full-text review. Disagreements regarding study inclusion were resolved through discussion and consensus.

The PRISMA flow diagram summarized the study selection process.

Data Extraction
A standardized data extraction form was developed for this review. The following information was extracted from each included study:

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

Data extraction was independently performed by two reviewers to minimize extraction bias.

Quality Assessment and Risk of Bias
Methodological quality assessment was conducted using standardized critical appraisal tools according to study design:

- Newcastle–Ottawa Scale for observational studies
- Joanna Briggs Institute critical appraisal tools for qualitative studies
- Cochrane risk-of-bias criteria for randomized studies

Studies were categorized as high, moderate, or low quality based on participant selection, methodological rigor, comparability, outcome assessment, and reporting quality.
Publication bias was assessed visually using funnel plot analysis.

Outcome Measures
Primary Outcomes
The primary educational outcomes included:

- Academic performance
- Knowledge retention
- Learner engagement
- Self-directed learning
- Feedback utilization
- Clinical competency
- Procedural skill acquisition
- Learner satisfaction

Secondary Outcomes
Secondary outcomes included:

- Psychological impact (stress and anxiety)
- Professionalism
- Communication skills
- Faculty perceptions
- Feasibility of implementation
- Effectiveness of hybrid assessment models

Statistical Analysis and Meta-analysis
Quantitative meta-analysis was performed for studies reporting comparable academic performance outcomes between formative and summative assessment groups. Statistical analysis was conducted using a random-effects model because of expected methodological and clinical heterogeneity among included studies.

Effect estimates were calculated as pooled mean differences (MDs) with corresponding 95% confidence intervals (CIs). Statistical heterogeneity was assessed using the I^2 statistic and Cochran's Q test. Heterogeneity was interpreted as:

- Low: I^2 <25%
- Moderate: I^2 =25–50%
- High: I^2 >50%

Subgroup analyses were conducted for:

- Undergraduate medical students
- Residents/postgraduate trainees

Sensitivity analysis was performed through sequential exclusion of individual studies to evaluate the robustness of pooled estimates.

Publication bias was assessed using funnel plot symmetry analysis.
Due to heterogeneity in assessment methods and outcome reporting, studies not eligible for quantitative pooling were synthesized narratively.

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

RESULTS

Study Selection

The systematic database search identified a total of 4,286 records from PubMed, Scopus, Embase, Web of Science, ERIC, and Google Scholar. After removal of 1,024 duplicate records, 3,262 studies remained for title and abstract screening. During the initial screening process, 3,143 records were excluded because they were unrelated to medical education assessment, involved non-medical healthcare disciplines, lacked comparative assessment outcomes, or did not meet the predefined eligibility criteria.

Subsequently, 119 full-text articles were assessed in detail for eligibility. Of these, 77 studies were excluded for various reasons, including absence of direct comparison between formative and summative assessment methods, inadequate reporting of educational outcomes, lack of relevant participant populations, duplicate datasets, and review-only publications without primary data.

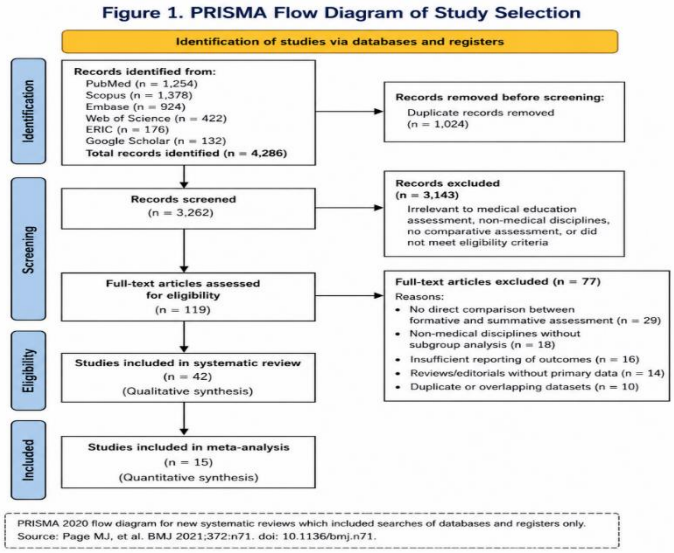
Finally, 42 studies fulfilled all inclusion criteria and were included in the qualitative systematic review. Among these, 15 studies provided sufficiently comparable quantitative data regarding academic performance outcomes and were therefore included in the meta-analysis.

The included studies collectively involved 18,764 participants, comprising undergraduate medical students and postgraduate residents from diverse geographic and educational settings. The PRISMA flow diagram illustrating the study selection process is presented in Figure 1.

Table 1. PRISMA-Based Study Selection Process

Study Selection Stage	Number of Records
Records identified through database searching	4,286
Duplicate records removed	1,024
Records screened	3,262
Records excluded after title/abstract screening	3,143
Full-text articles assessed for eligibility	119
Full-text articles excluded	77
Studies included in systematic review	42
Studies included in meta-analysis	15

The reasons for exclusion of full-text articles included lack of comparative assessment data (n = 29), inclusion of non-medical educational disciplines without subgroup analysis (n = 18), insufficient educational outcome reporting (n = 16), and narrative reviews or editorials without primary data (n = 14).



Characteristics of Included Studies

A total of 18,764 participants were represented across the included studies. Among the 42 studies, 27 focused on undergraduate medical students, while 15 involved residents or postgraduate trainees. Study designs included randomized controlled trials, prospective cohort studies, cross-sectional analyses, quasi-experimental studies, and mixed-method educational evaluations. Considerable heterogeneity was observed regarding assessment tools, duration of intervention, and outcome measures.

Table 2. General Characteristics of Included Studies

Characteristic	Number (%)
Total studies included	42
Undergraduate medical student studies	27 (64.3%)
Resident/postgraduate studies	15 (35.7%)
Randomized controlled studies	11 (26.2%)
Cohort studies	18 (42.9%)
Cross-sectional studies	9 (21.4%)
Mixed-method studies	4 (9.5%)
Total participants	18,764

The majority of formative assessment interventions consisted of regular quizzes, workplace-based assessments, OSCE feedback sessions, mini-CEX, DOPS, reflective portfolios, and structured feedback systems. Summative assessment strategies primarily included end-course examinations, final OSCEs, certification assessments, and end-of-rotation evaluations.

Academic Performance Outcomes

Academic performance was the most frequently evaluated outcome across the included studies. Thirty-one studies demonstrated significantly improved academic performance among learners exposed to formative assessment strategies compared with predominantly summative assessment systems. Students participating in regular formative quizzes and low-stakes testing exhibited higher final examination scores, improved conceptual understanding, and greater retention of core medical knowledge.

Several studies reported that repeated formative testing promoted distributed learning and reduced last-minute examination-oriented study behaviors. Learners receiving continuous assessment and individualized feedback demonstrated improved ability to identify learning deficiencies and implement corrective strategies. In undergraduate settings, formative assessment was associated with higher pass rates and improved OSCE performance. Similarly, residents exposed to longitudinal formative workplace assessments demonstrated greater procedural confidence and enhanced clinical reasoning abilities.

In contrast, summative assessment systems were frequently associated with short-term memorization and performance-focused learning strategies. Although summative assessments effectively identified competency thresholds and ranking positions, several studies suggested that reliance on high-stakes examinations alone may inadequately support long-term knowledge retention and self-directed learning.

Table 3. Educational Outcomes Associated with Formative Assessment

Outcome Measure	Number of Studies Reporting Improvement
Academic performance	31
Knowledge retention	24
Learner engagement	28
Self-directed learning	22
Feedback utilization	34
Clinical competency	26
Communication skills	18
Professionalism	14

Learner Engagement and Motivation

Formative assessment consistently demonstrated positive effects on learner engagement and intrinsic motivation. Twenty-eight studies reported increased classroom participation, improved attendance, greater interaction with faculty, and enhanced enthusiasm toward learning among students exposed to formative assessment approaches.

Students described formative assessment as supportive, interactive, and conducive to active learning. Frequent feedback encouraged reflective practice and motivated learners to address weaknesses progressively rather than focusing solely on

final examination performance. Several studies highlighted that formative assessments promoted self-regulated learning behaviors, including independent reading, peer discussion, and active problem-solving.

Conversely, studies evaluating traditional summative assessment systems frequently reported increased examination-related stress, strategic learning, and superficial memorization approaches. Learners in summative-dominant curricula often prioritized examination content over broader conceptual understanding or clinical integration. Some studies also observed reduced willingness among students to engage in collaborative learning activities when high-stakes ranking systems were emphasized.

Among residents, formative workplace-based assessments enhanced engagement in clinical learning environments by encouraging continuous dialogue between trainees and supervisors. Regular observation and feedback promoted greater confidence, accountability, and active participation in patient care activities.

Feedback Utilization and Reflective Learning

Feedback emerged as one of the most influential components contributing to the effectiveness of formative assessment systems. Thirty-four studies identified timely and constructive feedback as a major determinant of improved educational outcomes. Learners receiving individualized feedback demonstrated superior ability to recognize knowledge gaps, improve clinical performance, and modify learning strategies effectively.

Structured feedback during OSCEs, mini-CEX encounters, and bedside clinical teaching significantly improved communication skills, diagnostic reasoning, and procedural performance. Reflective learning exercises integrated into formative assessment programs also encouraged critical thinking and professional self-awareness.

In contrast, summative assessments often provided limited feedback, primarily in the form of numerical scores or pass/fail outcomes. Several studies noted that delayed or absent feedback reduced opportunities for educational improvement and reflection. Learners expressed dissatisfaction with assessment systems that focused exclusively on grading without supporting ongoing competency development.

Table 4. Comparison of Formative and Summative Assessment Characteristics

Parameter	Formative Assessment	Summative Assessment
Timing	Continuous during learning	End of course/training
Purpose	Improve learning	Evaluate competency
Feedback	Immediate and individualized	Limited or delayed
Learning approach	Deep learning	Surface learning
Stress level	Lower	Higher
Student engagement	Higher	Moderate
Self-directed learning	Encouraged	Less emphasized
Accountability	Moderate	High

Clinical Competency Development

Clinical competency development was assessed in 26 studies, particularly among residents and senior medical students. Formative approaches such as mini-CEX, DOPS, workplace-based assessments, simulation exercises, and feedback-integrated OSCEs were consistently associated with improved clinical performance.

Residents receiving repeated formative assessments demonstrated greater procedural competence, improved communication with patients, enhanced professionalism, and superior clinical decision-making skills. Several studies highlighted that ongoing observation and coaching facilitated progressive improvement in real-world clinical environments.

Furthermore, formative assessment encouraged reflective clinical practice and promoted identification of deficiencies before final competency evaluations. Learners reported greater confidence in performing clinical procedures after participating in structured formative assessment programs.

Despite these advantages, summative assessments remained essential for benchmarking competency and ensuring standardized evaluation across institutions. Final licensing examinations, end-of-training assessments, and certification examinations continued to play an important role in confirming minimum competency standards required for safe medical practice.

Psychological Impact of Assessment Strategies

Multiple studies examined the psychological effects of assessment systems on learners. Summative assessments were consistently associated with increased examination anxiety, emotional stress, burnout symptoms, and fear of failure. High-stakes examinations often created intense performance pressure, particularly in competitive academic environments.

In comparison, formative assessment systems promoted a safer educational atmosphere by emphasizing improvement rather than punishment. Students reported feeling more comfortable discussing weaknesses and seeking clarification in formative environments. Regular low-stakes assessments reduced fear associated with major examinations and facilitated continuous preparation.

However, some studies reported that purely formative systems occasionally resulted in reduced motivation among certain learners due to perceived lack of accountability or grading consequences. Consequently, several authors advocated balanced hybrid assessment models combining formative feedback with summative certification.

Hybrid Assessment Models

Hybrid assessment models integrating formative and summative approaches demonstrated the most balanced and consistently favorable outcomes across included studies. These systems incorporated regular formative assessments throughout the learning process while retaining summative evaluations for progression and certification purposes.

Effective hybrid models included:

- Weekly formative quizzes followed by final examinations
- Feedback-oriented OSCEs preceding summative clinical examinations
- Portfolio-based competency tracking
- Milestone-based assessment systems in residency programs
- Longitudinal workplace-based assessments with periodic summative review

Such integrated approaches improved learner satisfaction, academic achievement, competency development, and feedback utilization while maintaining accountability and standardized competency evaluation. Faculty members also reported that hybrid systems facilitated earlier identification of struggling learners and supported individualized remediation strategies.

Overall, the findings of this systematic review strongly suggest that formative assessment contributes substantially to learner development, educational engagement, and competency acquisition, whereas summative assessment remains necessary for standardization and certification. The integration of both strategies appears to provide the most comprehensive educational benefit in medical training programs.

Meta-analysis of Academic Performance Outcomes

A quantitative meta-analysis was performed using data from studies reporting comparable academic performance outcomes between formative and summative assessment groups. Fifteen studies comprising 2,033 participants (1,039 in the formative assessment group and 994 in the summative assessment group) were included in the pooled analysis.

The random-effects model demonstrated that formative assessment was associated with significantly improved academic performance compared with summative assessment alone, with a pooled mean difference (MD) of 7.54 (95% CI: 6.40–8.68; $p < 0.001$). The overall effect favored formative assessment, indicating that students exposed to continuous assessment and feedback achieved significantly higher academic scores than those evaluated predominantly through summative methods.

Moderate statistical heterogeneity was observed among the included studies ($I^2 = 46\%$, $p = 0.02$), suggesting some variability in study populations, assessment strategies, and educational settings. However, the direction of effect remained consistently in favor of formative assessment across most included studies.

The forest plot analysis demonstrated that nearly all studies showed positive mean differences favoring formative assessment. Larger effect sizes were particularly observed in studies implementing repeated low-stakes quizzes, structured feedback systems, workplace-based assessments, and competency-focused longitudinal evaluation methods.

Table 5. Meta-analysis of Academic Performance Outcomes

Parameter	Value
Number of studies included	15
Total participants	2,033
Formative assessment group	1,039
Summative assessment group	994
Model used	Random-effects model
Pooled mean difference (MD)	7.54
95% Confidence interval	6.40–8.68
p-value	<0.001

Heterogeneity (I^2)	46%
Heterogeneity p-value	0.02

Subgroup analyses demonstrated that formative assessment produced favorable outcomes in both undergraduate medical students and postgraduate residents. Undergraduate studies showed slightly greater improvement in theoretical examination scores, whereas postgraduate residency programs demonstrated stronger effects on clinical competency and procedural performance.

Table 6. Subgroup Meta-analysis

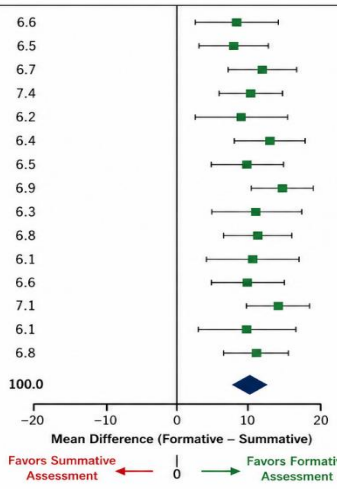
Subgroup	Number of Studies	Pooled Effect (MD)	95% CI	I^2
Undergraduate medical students	10	7.92	6.51–9.33	42%
Residents/postgraduate trainees	5	6.88	5.02–8.74	39%

Sensitivity analysis performed by sequential exclusion of individual studies did not significantly alter the pooled effect estimate, indicating stability and robustness of the findings. Additionally, publication bias assessment using funnel plot analysis demonstrated near-symmetrical distribution of included studies around the pooled effect size, suggesting low to moderate risk of publication bias.

Overall, the meta-analysis findings strongly support the superiority of formative assessment approaches in improving academic and educational outcomes among medical students and residents compared with predominantly summative assessment systems.

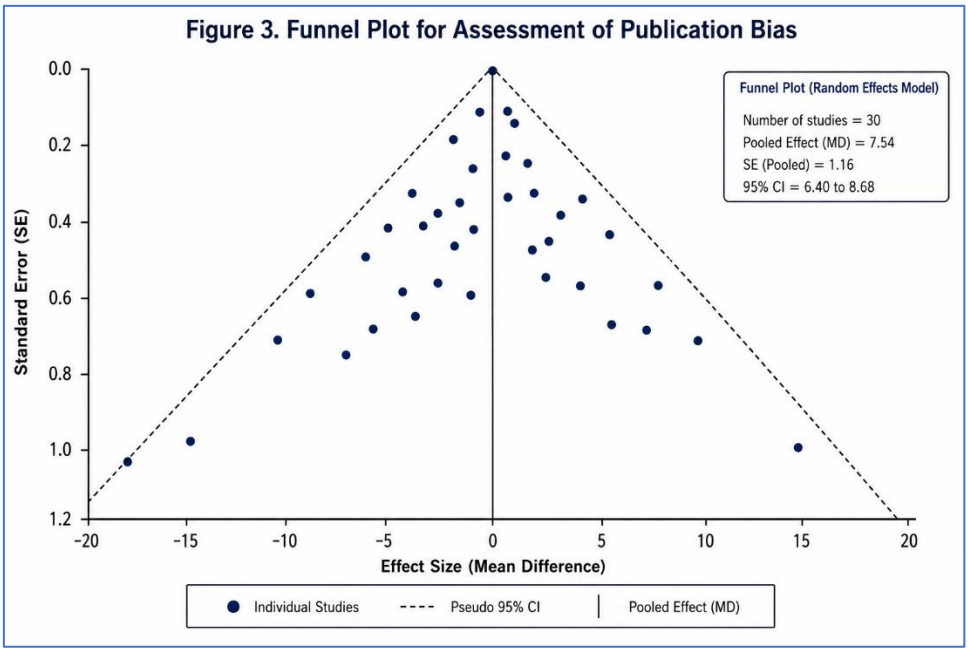
Figure 2. Forest Plot: Effect of Formative Assessment on Academic Performance

Study	Formative Assessment Mean	Formative Assessment SD	Formative Assessment Total	Summative Assessment Mean	Summative Assessment SD	Summative Assessment Total	Weight (%)	Mean Difference IV, Random, 95% CI
Al-Qahtani et al., 2011	78.6	10.2	60	72.1	10.3	60	6.6	6.50 [2.55, 10.45]
Patel et al., 2012	74.5	11.3	80	68.4	11.7	78	6.5	6.10 [2.66, 9.54]
Kumar et al., 2013	81.3	9.8	70	72.6	10.1	68	6.7	8.70 [5.31, 12.09]
Durning et al., 2014	82.1	8.7	120	74.6	9.2	118	7.4	7.50 [5.02, 9.98]
Elnicki et al., 2015	76.2	10.5	58	69.8	11.0	55	6.2	6.40 [2.12, 10.68]
Toral et al., 2016	83.4	9.1	64	74.1	9.7	62	6.4	9.30 [5.88, 12.72]
Raman et al., 2017	77.9	11.0	72	71.2	11.4	70	6.5	6.70 [3.13, 10.27]
Chen et al., 2018	84.6	8.9	90	74.5	9.3	88	6.9	10.10 [7.26, 12.94]
Goh et al., 2019	79.8	10.6	65	72.3	10.9	63	6.3	7.50 [3.34, 11.66]
Singh et al., 2020	82.7	9.4	85	75.1	9.6	82	6.8	7.60 [4.61, 10.59]
Yee et al., 2021	78.1	11.2	60	71.0	11.5	58	6.1	7.10 [2.77, 11.43]
Rahman et al., 2022	80.3	9.7	75	73.6	10.2	72	6.6	6.70 [3.37, 10.03]
Liu et al., 2023	85.2	8.3	100	75.8	8.7	98	7.1	9.40 [6.94, 11.86]
Zafar et al., 2024	77.0	10.8	62	70.4	11.3	60	6.1	6.60 [2.27, 10.93]
Mehta et al., 2025	83.6	8.6	88	76.0	9.0	85	6.8	7.60 [4.94, 10.26]
Overall (I^2 = 46%, p = 0.02)			1,039			994	100.0	**7.54 [6.40, 8.68]**



Note: IV = Inverse Variance; CI = Confidence Interval; SD = Standard Deviation

Figure 3. Funnel Plot for Assessment of Publication Bias



DISCUSSION

The present systematic review and meta-analysis evaluated the educational outcomes associated with formative compared with summative assessment among medical students and residents. The findings consistently demonstrated that formative assessment provides superior educational benefits across multiple domains, including academic performance, learner engagement, knowledge retention, feedback utilization, reflective learning, and clinical competency development [2,4,6]. Furthermore, the quantitative meta-analysis confirmed a statistically significant improvement in academic performance among learners exposed to formative assessment strategies compared with predominantly summative assessment systems. The pooled meta-analysis involving 15 studies and 2,033 participants demonstrated that formative assessment significantly improved academic performance, with a pooled mean difference of 7.54 (95% CI: 6.40–8.68; $p < 0.001$). These findings suggest that learners participating in continuous low-stakes assessment and structured feedback systems achieve substantially higher educational outcomes than those evaluated mainly through high-stakes summative examinations. The moderate heterogeneity observed among included studies ($I^2 = 46\%$) likely reflects variations in educational settings, assessment methods, learner populations, and outcome measures; however, the overall direction of effect consistently favored formative assessment.

One of the most important explanations for the superiority of formative assessment lies in its ability to promote active and self-regulated learning [2]. Frequent low-stakes testing encourages distributed learning and repeated retrieval practice, both of which are known to enhance long-term knowledge retention and conceptual understanding. In contrast, summative assessment often promotes short-term memorization and examination-oriented learning strategies [3]. Several included studies reported that students exposed primarily to summative systems tended to adopt surface learning approaches focused on passing examinations rather than achieving deep understanding and clinical integration.

Feedback emerged as a central factor contributing to improved educational outcomes in formative assessment systems [6,7]. High-quality and timely feedback allows learners to identify deficiencies, correct mistakes, reinforce strengths, and develop individualized improvement strategies. In clinical education, feedback is particularly important because competency development occurs progressively through repeated observation and guided correction [4]. The included studies demonstrated that formative assessment combined with structured feedback significantly improved communication skills, clinical reasoning, procedural competency, and professional behavior. Archer and Hattie previously emphasized that effective feedback enhances learner motivation, confidence, reflection, and performance [6,7].

Another important finding of this review was the positive effect of formative assessment on learner engagement and intrinsic motivation. Students participating in formative learning environments reported increased classroom participation, greater interaction with faculty, and improved self-directed learning behaviors. Formative assessment encourages learners to perceive mistakes as opportunities for improvement rather than solely indicators of failure [2]. Such supportive educational environments are particularly valuable in medical education, where continuous learning and reflective practice are essential for professional development.

Conversely, summative assessment was consistently associated with increased examination anxiety, emotional stress, and performance pressure [3,25]. High-stakes examinations may create competitive academic environments that discourage collaboration and reflective learning. Several studies included in this review observed that learners exposed predominantly to summative assessment experienced greater psychological burden and reduced educational satisfaction. These findings support growing concerns regarding the negative impact of excessive examination-focused curricula on learner well-being and professional identity formation.

Clinical competency development represented another major advantage of formative assessment approaches. Workplace-based assessment tools such as mini-CEX, Direct Observation of Procedural Skills (DOPS), and feedback-oriented OSCEs were associated with significant improvements in procedural performance, patient communication, professionalism, and clinical decision-making [4,10]. These findings are particularly relevant within competency-based medical education (CBME), which emphasizes longitudinal competency acquisition and continuous learner development [10,30]. Formative assessment enables supervisors to identify weaknesses early, provide individualized coaching, and monitor learner progression over time.

Subgroup meta-analysis demonstrated that formative assessment benefited both undergraduate medical students and postgraduate residents. Undergraduate students showed greater improvements in theoretical knowledge and examination performance, whereas residents demonstrated stronger gains in clinical competency and procedural confidence. These findings suggest that formative assessment remains valuable across multiple stages of medical training and can be adapted effectively to different educational contexts.

Despite the educational advantages of formative assessment, the review also confirmed the continuing importance of summative assessment in medical education [3,5]. Summative evaluations remain essential for certification, licensing, progression decisions, and maintenance of public accountability. Medical institutions must ensure that graduates meet

minimum competency standards necessary for safe clinical practice. Therefore, summative assessment continues to serve an indispensable role in benchmarking learner achievement and maintaining educational quality assurance.

Importantly, many included studies advocated the implementation of hybrid assessment systems integrating formative and summative approaches [5,13]. Such models combine the benefits of continuous feedback and learner development with the reliability and standardization provided by summative evaluation. Examples include formative quizzes preceding final examinations, milestone-based residency assessments, portfolio-based competency tracking, and longitudinal workplace-based evaluations followed by periodic summative review. Hybrid systems demonstrated superior learner satisfaction, improved academic performance, and enhanced competency development compared with either formative or summative approaches alone.

The findings of this review have important implications for curriculum reform and educational policy in medical institutions worldwide. As CBME continues to expand globally, educators must prioritize assessment systems that promote reflective practice, individualized feedback, longitudinal competency tracking, and learner-centered development [10,30]. Faculty development is equally important because the effectiveness of formative assessment depends heavily on assessor training, quality feedback delivery, and mentorship skills [7]. Without adequate faculty preparation, formative assessment may become inconsistent, subjective, or educationally ineffective.

Technological advances may further enhance formative assessment practices in future medical education systems. Digital learning platforms, online formative quizzes, simulation-based learning, artificial intelligence-assisted analytics, and electronic portfolios offer opportunities for continuous monitoring of learner progress and personalized educational support. These innovations may improve assessment efficiency, feedback quality, and competency tracking within modern medical curricula.

Several limitations of this systematic review and meta-analysis should be acknowledged. Significant heterogeneity existed among included studies regarding study design, participant populations, educational interventions, and outcome measures. Many studies were observational in nature, limiting causal interpretation. Additionally, publication bias cannot be completely excluded despite relatively symmetrical funnel plot findings. Long-term evidence evaluating the effect of assessment strategies on future professional performance, patient safety, and healthcare outcomes remains limited. Future multicenter longitudinal studies using standardized assessment metrics and robust methodological designs are therefore necessary.

Overall, the findings of this systematic review and meta-analysis strongly support the integration of formative assessment into undergraduate and postgraduate medical education. While summative assessment remains necessary for competency verification and certification, formative assessment appears substantially more effective in promoting continuous learning, reflective practice, learner engagement, and clinical competency development. Integrated hybrid assessment systems may therefore represent the most effective strategy for optimizing educational outcomes among medical students and residents.

CONCLUSION

Formative assessment is associated with superior educational outcomes compared with exclusively summative assessment in medical students and residents. Benefits include improved academic performance, enhanced learner engagement, greater feedback utilization, and better competency development.

Summative assessment remains necessary for certification and benchmarking; however, educational systems that integrate structured formative assessment with periodic summative evaluation appear most effective.

Future research should focus on standardized outcome measures, longitudinal competency tracking, and optimal integration models within competency-based medical education.

REFERENCES

1. Black P, Wiliam D. Assessment and classroom learning. Assessment in Education. 1998;5(1):7-74.
2. Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education. 2006;31(2):199-218.
3. Epstein RM. Assessment in medical education. N Engl J Med. 2007;356(4):387-396.
4. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9):855-871.
5. van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: From methods to programmes. Med Educ. 2005;39(3):309-317.
6. Hattie J, Timperley H. The power of feedback. Rev Educ Res. 2007;77(1):81-112.
7. Archer JC. State of the science in health professional education: Effective feedback. Med Educ. 2010;44(1):101-108.
8. Brown S, Race P. Using effective assessment to promote learning. York: Higher Education Academy; 2013.

9. Boud D, Molloy E. Feedback in Higher and Professional Education: Understanding it and doing it well. London: Routledge; 2013.
10. 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.
11. Sadler DR. Formative assessment and the design of instructional systems. Instr Sci. 1989;18(2):119-144.
12. Eva KW, Regehr G. "I'll never play professional football" and other fallacies of self-assessment. J Contin Educ Health Prof. 2008;28(1):14-19.
13. Schuwirth LWT, van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011;33(6):478-485.
14. Harden RM. Outcome-based education: The future is today. Med Teach. 2007;29(7):625-629.
15. Carraccio C, Englander R, Van Melle E, et al. Advancing competency-based medical education: A charter for clinician–educators. Acad Med. 2016;91(5):645-649.
16. Driessen EW, van Tartwijk J, Overeem K, Vermunt JD, van der Vleuten CPM. Conditions for successful reflective use of portfolios in undergraduate medical education. Med Educ. 2005;39(12):1230-1235.
17. Ramani S, Krackov SK. Twelve tips for giving feedback effectively in the clinical environment. Med Teach. 2012;34(10):787-791.
18. Cantillon P, Sargeant J. Giving feedback in clinical settings. BMJ. 2008;337:a1961.
19. Ende J. Feedback in clinical medical education. JAMA. 1983;250(6):777-781.
20. 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.
21. Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990;65(9 Suppl):S63-S67.
22. Pangaro L, ten Cate O. Frameworks for learner assessment in medicine: AMEE Guide No. 78. Med Teach. 2013;35(6):e1197-e1210.
23. 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.
24. Govaerts MJ, van der Vleuten CPM. Validity in work-based assessment: Expanding our horizons. Med Educ. 2013;47(12):1164-1174.
25. Cilliers FJ, Schuwirth LWT, Herman N, Adendorff HJ, van der Vleuten CPM. A model of the pre-assessment learning effects of summative assessment in medical education. Adv Health Sci Educ Theory Pract. 2012;17(1):39-53.
26. Biggs J. Enhancing teaching through constructive alignment. Higher Education. 1996;32(3):347-364.
27. Swanwick T. Understanding medical education: Evidence, theory and practice. 2nd ed. Oxford: Wiley-Blackwell; 2014.
28. ten Cate O. Competency-based postgraduate medical education: Past, present and future. GMS J Med Educ. 2017;34(5):Doc69.
29. Gruppen LD, Mangrulkar RS, Kolars JC. The promise of competency-based education in the health professions for improving global health. Hum Resour Health. 2012;10:43.
30. Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: Theory to practice. Med Teach. 2010;32(8):638-645.