

Original Article

**Beyond the Station: A Critical Review of Technological Innovation, Validity, and the Evolving Role of the OSCE in Modern Medical Education**

*Hossin M. I<sup>1</sup>, Mahfuz A<sup>2</sup>, Saiham K.S<sup>3</sup>, Ali S M Y<sup>4</sup>, Islam M.S<sup>5</sup>, Nahar K<sup>6</sup>, Mahabub SM E.E<sup>7</sup>, Ahmad M.R.<sup>8</sup>*

**Abstract**

**Background:** The Objective Structured Clinical Examination (OSCE) is a widely used method for assessing clinical competence, but its validity and role in modern medical education are increasingly debated. The rapid integration of technology and the shift toward longitudinal assessment models necessitate a critical re-evaluation of the OSCE's function. **Methods:** This narrative review synthesises evidence from a comprehensive analysis of the literature on OSCEs, focusing on technological innovations, psychometric properties, and assessment frameworks. A thematic analysis was conducted on peer-reviewed articles, systematic reviews, and consensus statements retrieved from databases such as PubMed, Scopus, and Web of Science, with a particular emphasis on publications from the last five years to capture current trends. **Results:** The review identifies three major themes. First, technology, including virtual OSCEs (vOSCEs), virtual reality, and digital scoring systems, is transforming assessment delivery and feedback, though vOSCEs face limitations in evaluating physical examination skills. Second, while OSCEs offer standardisation, concerns persist about their authenticity and construct validity, prompting a need for a more nuanced, evidence-based approach to validity arguments. Third, the future of assessment lies in programmatic models, where the OSCE serves as one component, triangulated with workplace-based assessments like DOPS, to make holistic, high-stakes decisions about competence. **Conclusion:** The OSCE must evolve from a standalone, high-stakes examination to a dynamic, technology-enhanced element within a broader programmatic assessment framework. This integrated approach leverages the OSCE's strengths in standardisation while addressing its limitations through authentic, workplace-based observations, ensuring a more valid, reliable, and educationally sound evaluation of clinical competence.

**Keywords:** OSCE, medical education, clinical skills assessment, technology-enhanced assessment, virtual OSCE, validity, reliability, programmatic assessment, workplace-based assessment

1. Dr Md. Immam Hossin, PhD Fellow, Bangladesh Medical University, (OSD), DGHS, Mohakhali, Dhaka,
2. Dr Anika Mahfuz, Resident Doctor, Radiology and Imaging, Sir Salimullah Medical College, Dhaka,
3. Dr. Kazi Shantono Saiham, Consultant radiologist, Doctors Care General Hospital, Brahmanbaria
4. Dr S M Yunus Ali, Associate Professor, Department of Urology, Bangladesh Medical University, Dhaka,
5. Prof. Dr. Md. Saiful Islam, Professor, Department of Urology, Bangladesh Medical University, Dhaka,
6. Dr Kamrun Nahar, Professor of Forensic Medicine, International Medical College.
7. Dr.SM Ear-E-Mahabub, Associate Professor, Department of Cardiology, Bangladesh Medical University (BMU), Dhaka.
8. Dr Md. Rasel Ahmad, Assistant Professor, Medical Education, Department of Public Health and Informatics, Bangladesh Medical University, Dhaka.

**Address of correspondence:** Dr. Md. Rasel Ahmad, Assistant Professor, Medical Education, Department of Public Health and Informatics, Bangladesh Medical University, Dhaka, Bangladesh.  
Email: [raselbds42@bsmmu.edu.bd](mailto:raselbds42@bsmmu.edu.bd). Orchid ID-0009-0009-2936-0169

## Original Article

### Introduction

The assessment of clinical competence is a fundamental pillar of medical education, ensuring that healthcare professionals develop the knowledge, skills, and behaviours required for safe and effective patient care. Since its introduction in 1975, the Objective Structured Clinical Examination (OSCE) has become a globally recognised method for assessing competencies such as history taking, physical examination, communication, and professionalism through timed, standardised stations.<sup>1</sup> Its structured design supports objectivity and fairness compared with traditional clinical assessments and is therefore widely used for high-stakes decisions, including licensure and progression.<sup>4</sup>

However, contemporary medical education is being reshaped by rapid technological innovation and an evolving philosophy of assessment. The expansion of digital learning-accelerated during the COVID-19 pandemic-has contributed to the development and uptake of virtual OSCEs (vOSCEs) and other technology-enhanced simulations.<sup>2</sup> At the same time, assessment practice is shifting away from isolated, one-off summative examinations toward longitudinal, programmatic assessment models that integrate multiple sources of evidence to support holistic judgments of competence.<sup>5</sup>

Despite its enduring popularity, the OSCE faces important challenges that create a critical gap in the literature. One major concern is the ongoing debate about validity and authenticity. Critics argue that simulated patients and scripted scenarios may inadequately represent the complexity and unpredictability of real clinical practice, raising doubts about whether OSCE performance extrapolates to

workplace performance.<sup>2</sup> These concerns are amplified in vOSCEs, where limited ability to assess hands-on physical examination may threaten construct validity.<sup>3</sup> Although technology offers opportunities to improve scoring efficiency, feedback quality, and immersion (eg, through virtual reality), its use is often fragmented and implemented without a coherent framework.<sup>10</sup> Most importantly, there remains limited synthesis that positions the OSCE within a modern programmatic assessment approach-moving beyond viewing the OSCE as a standalone “gatekeeper” and instead examining how OSCE findings can be combined with workplace-based assessments (WBAs) such as Direct Observation of Procedural Skills (DOPS), mini-CEX, multisource feedback, and portfolios to triangulate competence more robustly.<sup>5</sup>

This narrative review critically examines the current status and future development of OSCEs in medical education. Specifically, it aims to: (1) appraise how technological innovations (including vOSCEs, virtual reality, and digital scoring systems) influence delivery, validity, and educational value; (2) evaluate evidence supporting or challenging validity and reliability across traditional and virtual formats using contemporary validity thinking; and (3) propose an approach for integrating OSCEs into programmatic assessment, combining standardisation with authentic workplace-based observations to improve the defensibility of high-stakes decisions.<sup>5</sup>

### Methods

A narrative review methodology was employed to analyze the evolving role of the OSCE in modern medical education. This approach was selected because it enables interpretive synthesis across

## Original Article

diverse literature types, including empirical studies, theoretical perspectives, and review articles, thereby supporting a broad and critical appraisal of complex educational questions.<sup>4</sup>

A comprehensive literature search was conducted across PubMed, Scopus, and Web of Science using the terms “Objective Structured Clinical Examination,” “OSCE,” “technology in medical education,” “virtual OSCE,” “OSCE validity,” “OSCE reliability,” “programmatic assessment,” and “workplace-based assessment.” No date limits were applied to capture the historical development of OSCEs, while prioritizing literature from the past five years to reflect recent debates and technological advances. Titles and abstracts were screened for relevance. Eligible studies included peer-reviewed research articles, systematic/scoping reviews, and influential perspective pieces addressing technological

**Results:** The following are three tables summarising the key aspects of Technology Integration, Validity and Reliability, and

integration, psychometric properties (validity/reliability), and/or the role of OSCEs within broader assessment systems. Studies that were purely descriptive of local OSCE implementation without critical analysis, or focused narrowly on a single technical feature without educational implications, were excluded.

Full texts were reviewed in detail, and a thematic analysis was conducted. The themes were identified through an inductive approach, with two reviewers independently categorizing the data based on recurring patterns. A consensus process was followed to resolve any discrepancies between reviewers. The identified themes aligned with the review objectives: (i) technological integration, (ii) validity and reliability, and (iii) the OSCE’s role in programmatic assessment. Findings were synthesized into thematic categories and interpreted to develop a coherent narrative and forward-looking implications.<sup>2,5</sup>

Programmatic Assessment in the context of OSCEs.

### Technology Integration

| Author (Year)                           | Study Design                           | Key Findings                                                                                                                                                                              | Strengths                                                                                                                   | Limitations                                                                    | Outcomes/Implications                                                                                                                    |
|-----------------------------------------|----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| Speedwell Software (2024) <sup>11</sup> | Review and case study (eSystem module) | Virtual OSCEs (vOSCEs) are a successful and flexible alternative to traditional OSCEs, with student outcomes comparable to in-person exams. VR and AR enable complex scenario simulation. | The eSystem module provides a seamless, integrated solution, expanding assessable competencies beyond physical limitations. | Initial costs and accessibility barriers for advanced technologies like VR/AR. | Technology enhances OSCEs by providing flexible assessment options, expanding learning opportunities, and improving feedback efficiency. |
| Blamoun et al. (2021) <sup>6</sup>      | Guide and perspective on vOSCEs        | vOSCEs require significant changes, including adapting educational objectives and                                                                                                         | Practical tips for vOSCE implementation, including virtual breakout rooms and pre-brief                                     | Limited ability to assess physical examination skills, relying                 | vOSCEs offer a solution for maintaining assessment continuity during crises and training future clinicians in telehealth.                |

Bangladesh Journal of Medical Education 2026; 17(2); Hossin et al., publisher and licensee Association for Medical Education. This is an Open Access article which permits unrestricted non-commercial use, provided the original work is properly cited.

## Original Article

| Author (Year)              | Study Design               | Key Findings                                                                                                                                                                     | Strengths                                                                                                                            | Limitations                                                                                                                     | Outcomes/Implications                                                                                                         |
|----------------------------|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Soong (2021) <sup>10</sup> | Perspective on AI in OSCEs | training learners and SPs in virtual environments.<br>AI can automate scoring, generate adaptive scenarios, and offer real-time feedback on procedural and communication skills. | sessions, which enhance educator engagement.<br>Reduces examiner bias, offers immediate feedback, and creates realistic simulations. | on verbalization.<br>Challenges in developing reliable AI algorithms for clinical skills, along with privacy and bias concerns. | AI represents a future direction for OSCEs, promising personalized, objective assessment but requiring ethical consideration. |

## Validity and Reliability

| Author (Year)                        | Study Design                   | Key Findings                                                                                                                            | Strengths                                                                                                                  | Limitations                                                                                                      | Outcomes/Implications                                                                                           |
|--------------------------------------|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|
| Rodriguez et al. (2021) <sup>9</sup> | Feasibility study              | OSCE was perceived as more objective and less stressful than traditional methods, with enhanced validity via the Kane framework.        | Use of formal validity framework strengthens assessment design, and a curriculum-based blueprint ensures content validity. | Conducted at a single institution with a small sample, limiting generalizability.                                | OSCE proved to be a valid and reliable summative assessment tool, positively influencing student behaviors.     |
| Chan et al. (2023) <sup>2</sup>      | Systematic review              | vOSCEs are viable for assessing communication skills but struggle with physical examination assessment, threatening construct validity. | Comprehensive synthesis of evidence on the feasibility of remote assessment for specific competencies.                     | Inability to assess physical skills undermines construct validity, and heterogeneity of studies limits findings. | vOSCEs maintain assessment continuity but have limitations in evaluating the full scope of clinical competence. |
| Hsieh et al. (2014) <sup>14</sup>    | Development of an OSCE station | OSCE successfully assessed higher-order cognitive skills like clinical reasoning and evidence application.                              | Demonstrates OSCE's flexibility in evaluating complex, non-technical competencies.                                         | Study focused on a single station, limiting the broader validity and reliability findings.                       | OSCE can assess critical skills beyond basic procedures, contributing to effective patient care.                |

## Programmatic Assessment

## Original Article

| Author (Year)                           | Study Design                                | Key Findings                                                                                                                                             | Strengths                                                                                                 | Limitations                                                                                | Outcomes/Implications                                                                                              |
|-----------------------------------------|---------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| Dijkstra et al. (2018) <sup>5</sup>     | Conceptual paper on programmatic assessment | OSCE should be integrated within a broader programmatic assessment framework, supplemented by workplace-based assessments (WBAs) like DOPS and mini-CEX. | The paper offers a robust theoretical framework for modern assessment, emphasizing multiple data sources. | Conceptual paper with no empirical data on implementation or outcomes.                     | The framework redefines the OSCE's role, enhancing its validity through integration with real-world observations.  |
| Chan et al. (2023) <sup>2</sup>         | Systematic review                           | Sole reliance on vOSCEs is insufficient; physical skills cannot be assessed, highlighting the need for complementary assessment methods.                 | Identifies gaps in vOSCEs and indirectly supports a programmatic approach using diverse assessment tools. | Review focuses on OSCE comparison, without explicit discussion of programmatic assessment. | Emphasizes using multiple assessment methods, as no single tool can capture the full range of clinical competence. |
| Speedwell Software (2024) <sup>11</sup> | Review and case study                       | The OSCE has evolved to meet the needs of medical education, with technology supporting continuous assessment.                                           | Highlights OSCE's adaptability, making it suitable for flexible, longitudinal assessment.                 | No explicit discussion on integrating OSCE data with other assessment methods.             | OSCE's evolution, driven by technology, ensures its relevance and integration into future assessment models.       |

## Discussion

This review highlights that the OSCE is at a crucial point in its evolution. Once revolutionary for standardizing clinical skills assessment, OSCEs are now being challenged by technological advances and a broader shift toward holistic, longitudinal assessment models. The adoption of virtual OSCEs (vOSCEs) demonstrates the "double-edged" impact of technology. While vOSCEs have shown feasibility and acceptability, with evidence suggesting comparable performance outcomes to in-person formats, limitations in assessing physical examination skills remain a major concern. These limitations may threaten construct validity<sup>2,3</sup>, as physical examination is integral to clinical

competence and authenticity. This limitation is not solely logistical but also conceptual—because the ability to physically examine a patient is a core aspect of medical practice, and its absence in virtual formats reduces the authenticity of the assessment. Despite decades of OSCE use, these concerns persist due to structural barriers, such as technological limitations, and philosophical barriers related to the difficulty of replicating hands-on, interactive clinical experiences virtually. Consequently, technology should be positioned as a complement to, rather than a wholesale substitute for, hands-on assessment. Emerging technologies—such as virtual reality simulations—offer the potential to assess procedural and complex

## Original Article

scenarios in controlled settings, while digital scoring platforms may improve standardization, efficiency, and feedback workflows<sup>7,10</sup>.

The ongoing debate surrounding the OSCE's validity and reliability should be viewed as an opportunity to apply more contemporary validity thinking. Programmatic assessment scholarship highlights that validity is not a static property of a single instrument, but rather a defensible argument built from multiple sources of evidence and interpreted within context<sup>5</sup>. From this perspective, OSCEs contribute valuable standardized evidence, particularly for sampling defined competencies under comparable conditions. Reliability can be deemed acceptable when the OSCE design is robust—for example, with adequate station sampling, trained examiners, and standardized marking. Recent evidence suggests that virtual OSCE formats can also demonstrate fair internal consistency and acceptability when well-designed<sup>8</sup>. However, OSCE results represent only one "slice" of performance and should not be over-interpreted as a direct proxy for sustained workplace competence<sup>5</sup>. Structural and procedural barriers, such as insufficient training or improper adaptation to real-world environments, may explain why these issues persist despite decades of OSCE use.

A critical implication of this review is the need to integrate OSCEs within a programmatic assessment framework. In such models, competence decisions are strengthened through triangulation across multiple low- and high-stakes data points accumulated longitudinally<sup>6</sup>. OSCE results gain interpretive power when combined with workplace-based assessments (e.g., DOPS, mini-CEX, multisource feedback), portfolios, and knowledge assessments<sup>7</sup>. For example, performance in a

musculoskeletal examination OSCE station becomes more meaningful when interpreted alongside directly observed performance of similar tasks in real clinical settings. This approach supports proportional decision-making, where higher-stakes judgments require richer, diverse, and more authentic evidence<sup>8</sup>. By integrating OSCEs within this framework, they can evolve from isolated "gatekeeper" exams into structured components of a comprehensive assessment system that better reflects real-world practice and supports continuous learning and development<sup>2,3</sup>.

## Future Research

Future research should explore the longitudinal impact of AI-enhanced OSCEs on clinical competence development, particularly focusing on the potential for personalized, adaptive learning. Additionally, the effectiveness of integrating multiple assessment modalities, including vOSCEs and workplace-based assessments, needs to be tested in diverse healthcare settings to determine its impact on clinical performance over time. Research should also address the persistent challenges in assessing physical examination skills in virtual environments, and how these can be overcome with emerging technologies like haptic feedback and augmented reality.

## Conclusion:

The implications of this review for medical education are transformative. The OSCE must evolve from a standalone, high-stakes examination into a dynamic component of a longitudinal, programmatic assessment framework. This shift requires educators to strategically integrate technology, using virtual and digital tools to enhance feedback and simulate complex scenarios, while acknowledging their limitations in assessing physical skills. The core implication is the necessity of triangulation:

OSCE data should be systematically combined with authentic, workplace-based assessments like DOPS to create a holistic, trustworthy evidence base for competence. This approach ensures that high-stakes decisions are proportional and valid, moving beyond artificial simulations to reflect real-world performance. Ultimately, this evolution will foster a more robust, fair, and educationally sound system that better prepares resilient, competent clinicians for the complexities of modern healthcare.

## References

1. Harden RM, Gleeson FA, Stevenson M. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ*. 1975;9(1):83–97. doi:10.1111/j.1365-2923.1975.tb00909.x
2. Chan SCC, Wass V, Shulruf B, et al. Implementation of virtual OSCE in health professions education: a systematic review. *Med Educ*. 2023;57(7):783–95. doi:10.1111/medu.14878
3. Chan SCC, Wass V, Shulruf B, et al. Evaluating construct validity of virtual OSCEs in exceptional times: a multi-centre study. *BMC Med Educ*. 2025;25:147. doi:10.1186/s12909-025-03512-x
4. Dewan P, Kolhe K, Jain A. Objective structured clinical examination for teaching and assessment in medical education: issues of validity, reliability, and standard setting. *J Med Educ*. 2024;28(2):110–20. doi:10.1007/s01034-024-01690-0
5. Dijkstra J, van der Vleuten CPM, Schuwirth LWT. Assuring the quality of programmatic assessment: moving from reliability to trustworthiness. *Med Teach*. 2018;40(11):1107–13. doi:10.1080/0142159X.2018.1508251
6. Blamoun J, Goyal K, et al. A guide for medical students and residents preparing for formative and summative virtual OSCE. *Adv Med Educ Pract*. 2021;12:491–8. doi:10.2147/AMEP.S308577
7. Mühling T, Hube E, Wolff K, et al. Comparing virtual reality-based and traditional physical OSCEs for clinical competence assessment: a randomized controlled trial. *J Med Internet Res*. 2025;27(1):e55066. doi:10.2196/55066
8. O'Malley A, McCarthy M, Hannigan A, et al. A comparison of objective structured clinical examinations during and after COVID-19: student performance, reliability, and acceptability. *BMC Med Educ*. 2025;25:201. doi:10.1186/s12909-025-03659-9
9. Rodriguez-Gazquez MA, Padilha KG, Musso MF, et al. Development and implementation of an objective structured clinical examination in a resource-limited institution. *Int J Med Educ*. 2021;12:220–6. doi:10.5116/ijme.60fd.2f68
10. Soong C. Artificial intelligence in medical OSCEs: reflections and future directions. *Adv Med Educ Pract*. 2021;12:233–7. doi:10.2147/AMEP.S314231
11. Speedwell Software. Innovations in OSCE: technology and simulation in medical education [Internet]. Speedwell Software Blog; 2024 [cited 2025 Dec 27]. Available from: <https://www.speedwellsoftware.com/innovations-in-osce-technology-and-simulation-in-medical-education/>
12. Ahmad MR, Ahmed AHU, Hossin MI, Ahmed AKMK, Hasan MK, Mahmud A. Direct observation of procedural skills (DOPS) in enhancing clinical competencies of postgraduate medical and dental education: a scoping review. *IOSR J Dent Med Sci*. 2024;23(9):59–70. doi:10.9790/0853-2309085970
13. Ahmad MR, Ear-E-Mahabub SM, Kajol HR, Afrin A, Ahmed AHHU, Hossin MI. Enhancing clinical competency in oral implantology: the role of direct observation of procedural skills (DOPS) in postgraduate training. *Front Health Inform*. 2024;13(8):5954–5960. doi:10.63682/fhi1891
14. Hsieh MC, Chen CY, Lee SY, Chen YC, Lin CH. Development and evaluation of an objective structured clinical examination station to assess evidence-based medicine competencies. *Med Educ*. 2014;48(6):617–626. doi:10.1111/medu.12442