

Review Article

Advancements in simulation-based assessments for health professionals: insights from a scoping review

Zahra Siavashpour^{1,2} , Zohreh Khoshgoftar¹ , Somaye Sohrabi¹ ,
Sarvin Bosak^{3*} 

¹ Department of Medical Education, School of Medical Education and Learning Technologies, Shahid Beheshti University of Medicine, Tehran, Iran

² Department of Radiotherapy Oncology, Shohada Tajrish Educational Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

³ Department of Nursing, School of Nursing and Midwifery, Dezful University of Medical Sciences, Dezful, Iran

Article info

Article history:

Received 10 Jan. 2025

Revised 26 Jan. 2025

Accepted 18 Nov. 2025

Published 19 Nov. 2025

*Corresponding author:

Sarvin Bosak, Department of Nursing, School of Nursing and Midwifery, Dezful University of Medical Sciences, Dezful, Iran.
Email: Sarvinbosak@gmail.com

How to cite this article:

Siavashpour Z, Khoshgoftar Z, Sohrabi S, Bosak S. Advancements in simulation-based assessments for health professionals: insights from a scoping review. *J Med Edu Dev*. 2025;18(4):140-152.
<http://dx.doi.org/10.61882/edcj.18.4.140>

Abstract

Background & Objective: Over the past 20 years, simulation has emerged as a key instrument in healthcare education, first for training and more recently for performance evaluation. However, the lack of clinical relevance and realism in traditional assessment methods has drawn criticism. Given their increasing use and the difficulty of assessing clinical competence, this scoping review was conducted to investigate the scope and features of Simulation-Based Assessments (SBAs) in health professions education.

Materials & Methods: This study adhered to the Joanna Briggs Institute (JBI) scoping review methodology. We searched Ovid MEDLINE, Scopus, Web of Science, CINAHL, APA PsycINFO, and Embase all the way through. We also manually searched two important journals, *Clinical Simulation in Nursing* and *Simulation in Healthcare*, for articles that were published between 2021 and 2024 and were related to the topic. We also looked through the reference lists of relevant reviews. After removing duplicates in EndNote X20, 49 studies that met the criteria were included and analyzed using descriptive and thematic content analysis in Microsoft Excel.

Results: The review pinpointed essential target demographics, applications, challenges, benefits, drawbacks, and requisite conditions pertaining to SBAs. Out of the 49 studies included, most were from the US and looked at specialists, emergency medicine providers, and nurses. Three main simulation modalities were identified: human participant (e.g., standardized patients), equipment-based simulators (both low and high-fidelity), and computer-based (virtual reality/screen-based). The challenges primarily pertained to the study's realism, validity, reliability, and feasibility.

Conclusion: More and more people in healthcare education are seeing simulation-based assessments as a useful way to test clinical competence and safety. Despite its advantages, SBA implementation faces significant challenges related to high cost and limitations in realism. Before adding SBAs to assessment frameworks, it is important to have clear rules and plans.

Keywords: health profession, education, simulation-based assessment, competency

Introduction

In its most general sense, assessment involves setting appropriate standards and applying expert judgment to evaluate quality [1]. During a student's educational journey, a variety of assessments are employed to evaluate progress, skill acquisition, and professional

preparedness. Nonetheless, apprehensions regarding the efficacy of conventional instruments, such as multiple-choice examinations, have led educators to transition towards more authentic and performance-oriented assessments [2]. Numerous reviews have scrutinized

assessment methodologies employed in both pre-registration and post-registration health professional education, frequently concentrating on the evaluation of learning outcomes or clinical competencies [3–5]. Evaluating clinical competency continues to be a complicated but necessary part of training and teaching health professionals.

There is now more awareness of how assessments affect learning behaviors, and more teachers are using performance-based methods [6]. Simulation-Based Assessment (SBA) has emerged as a promising method to enhance authentic evaluation in clinical education.

In the early 1900s, the military and aviation industries were the first to use simulation as a way to train and test people.

The first flight simulator was made in 1929 [7].

Since the 1950s, simulation technologies have undergone substantial transformations due to the emergence of computer-based systems, facilitating their swift integration into health professions educational settings. Simulation seeks to replicate actual patients, clinical situations, anatomical structures, and procedural tasks to enhance experiential learning [8].

The increase in simulation-based education can be linked to the development of diagnostic and therapeutic technologies, the scarcity of clinical educators, and the necessity to prepare for intricate or uncommon clinical situations.

Also, the increasing focus on patient safety has made simulation even more popular [9–11].

Because of this, health educators now widely agree that simulation is a useful method for teaching clinical and nonclinical skills.

Standardized Patients (SPs), anatomical models, part-task trainers, computerized high-fidelity human patient simulators, and virtual reality platforms are some of the simulation modalities used in health education [12,13].

Because they maximize learning opportunities and enhance patient safety by allowing learners to practice clinical skills prior to actual patient encounters, these simulation techniques are primarily utilized in preregistration training for health professionals [14, 15]. Through staged scenarios of increasing complexity that are customized to their proficiency levels, simulation offers a safe and controlled environment where learners can gradually develop clinical competencies. This is particularly valuable because working directly with patients for training can be expensive, time-consuming, difficult, dangerous, and morally dubious [11, 16].

Because SBA closely resembles actual clinical practice, it has become popular among health educators for assessing professional and student competencies. Previous literature reviews and systematic analyses have investigated the application of SBAs in health professions education. Although the advantages and uses of SBAs have been emphasized by systematic and narrative reviews, many of them do not fully examine or critically evaluate the various simulation modalities and how they affect learner outcomes [17]. Some reviews restrict the generalizability of their findings by concentrating only on particular occupations or environments [18].

Furthermore, it can be difficult to compile evidence and reach firm conclusions when terminology and assessment frameworks used in different studies differ. In order to properly map the available evidence and pinpoint knowledge gaps, researchers have emphasized the necessity of a more comprehensive and organized review process [19].

Comprehensive studies examining the breadth and depth of SBAs in health professions education are scarce, despite the widespread use of simulation in health education.

This kind of systematic review could provide insightful information.

A scoping review is a suitable methodology for methodically mapping the scope and depth of the current literature on SBAs in health professions education, given the diversity of simulation modalities and assessment practices. Scoping reviews, in contrast to traditional systematic reviews, are intended to clarify important concepts, identify research gaps, and guide future investigations [20].

They also permit a wider inclusion of study designs. In order to overcome the shortcomings of earlier reviews and provide educators and researchers with useful information, this study intends to conduct a scoping review that offers a comprehensive overview of how SBAs are used, assessed, and reported across various health professions.

Therefore, the purpose of this study is to investigate the use and reach of simulation-based assessments in health professions education.

Materials & Methods

This scoping review was conducted using the methodological framework developed by Arksey and O'Malley, which comprises five core stages: (1)

identifying the research question; (2) identifying relevant studies; (3) study selection; (4) charting the data; and (5) collating, summarizing, and reporting the results [21]. The optional consultation stage was also incorporated. The review was conducted and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) checklist [22] and the Joanna Briggs Institute (JBI) Reviewers' Manual [23].

Stage 1: Identifying the research questions

Research questions were created to guide the search and identify relevant literature for this scoping review.

What are the target groups of SBA in health professions education?

What are the applications of SBA in health professions education?

What are the challenges of SBA in health professions education?

What are the advantages of SBA in health professions education?

What are the disadvantages of SBA in health professions education?

What are the necessary conditions for doing SBA in health professions education?

Stage 2: Identifying relevant studies

The following databases and search engines were used in the review:

Ovid MEDLINE, Embase, Scopus, Web of Science (WOS), Cumulative Index to Nursing and Allied Health Literature (CINAHL), and APA PsycINFO. The search strategy included studies published between 1994 and 2021 in the selected databases. In addition, two leading journals in simulation-based education—Clinical Simulation in Nursing and Simulation in Healthcare—were manually searched for studies published between

January 2021 and 2023 to ensure the inclusion of the most recent developments.

The key search terms included [physician* OR anesthesiologist* OR anesthetist* OR cardiologist* OR dermatologist* OR endocrinologist* OR gastroenterologist* OR hepatologist* OR "General Practitioner*" OR geriatrician* OR gerontologist* OR gynecologist* OR nephrologist* OR neurologist* OR obstetrician* OR oncologist* OR ophthalmologist* OR otolaryngologist* OR otologist* OR pathologist* OR pediatrician* OR neonatologist* OR pulmonologist* OR radiologist* OR rheumatologist* OR surgeon* OR urologist* OR nurse* OR paramedic* OR anatomist* OR audiologist* OR dentist* OR pharmacist*] AND [simulation OR simulator* OR "virtual reality" OR "augmented reality" OR "mixed reality" OR manikin* OR mannequin*] AND [assess].

All database searches were carried out by a knowledgeable librarian, who also managed the review's records and data. **Supplementary 1** outlines the search strategy. Our review specifically targeted practicing health professionals rather than students or trainees; therefore, we did not include terms like "resident," "intern," or "medical education." Our keyword selection was aligned with this focus to ensure relevance to the intended population.

Stage 3: Study selection

All studies were imported into EndNote X20, and duplicates were removed. Two independent reviewers screened titles, abstracts, and papers based on inclusion/exclusion criteria.

A third reviewer resolved discrepancies. Additional papers were found through reference list analysis. A PRISMA flow diagram [22] (**Figure 1**) illustrates the complete article selection procedure. **Table 1** lists inclusion and exclusion criteria.

Table 1. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
- From the beginning of 1994 until January of 2023	- Outside the mentioned period
- All SBAs in the health profession	- Review articles and books
- All research designs except reviews	- Non-health profession education
- Original research, editorials, conference papers, reports, and theses	- Lack of access to the full text of the article
- In all countries and ethnic groups	
- Studies from all countries were included, with no restrictions on the ethnic groups of the study populations	
- The full text of the article is available	

Abbreviations: SBAs, simulation-based assessments.

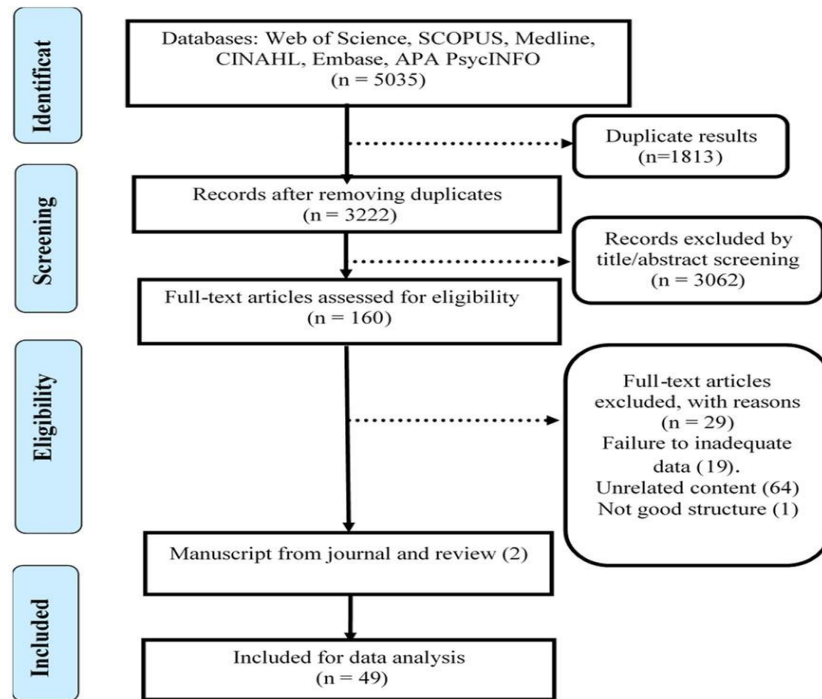


Figure 1. PRISMA Flow-diagram of the included studies in the review

Stage 4: Charting the data

Data were extracted and analyzed descriptively and thematically using Microsoft Excel. Data extraction was performed independently by two researchers using the JBI-recommended methodology [23] to minimize bias. Based on the study team's agreement, a data extraction tool [21] was made to capture the essential details about the source and conclusions relevant to the review questions. To calibrate the process and ensure consistency, two authors independently piloted the data extraction tool on the first five included studies. Their results were then compared, and any discrepancies were discussed to refine the tool and methodology before proceeding with the full data extraction. For each of the included papers, we collected the title, authors, year, type of publication, location of studies, study population, study design, publication type, simulation method, challenges, advantages, disadvantages, application, and useful tips for simulation (**Supplementary 2**). Finally, a total of 49 retrieved papers were determined to be relevant to the review questions.

Stage 5: Collating, summarizing, and reporting results

A descriptive technique was used to collate, summarize, and categorize the literature. The framework's fifth and

final step involved analyzing the results and reporting the findings, thereby addressing the scoping review's six research questions. In accordance with the JBI methodology for scoping reviews, no formal risk-of-bias assessment was conducted.

Results

There are two sections to the results presentation. First, a descriptive-narrative summary and overview of the included studies are presented (**Supplementary 3**). The six predetermined research questions are then used to organize and report the findings. The challenges identified were not included in the tabulated summary due to their interpretive and narrative nature; however, they are fully described in the relevant section of the text.

Descriptive summary of the included studies

This review included 49 studies in total. The included studies were released from 2000 to 2023. Seven studies (14.3%) came from Canada, followed by Australia (5 studies; 10.2%), and the United Kingdom (5 studies; 10.2%, which includes one study from Scotland). Denmark and the Netherlands contributed two studies each (4.1%). Single studies (2%) originated from Switzerland and Finland. Medical specialists (22 studies;

44.9%) and nurses (9 studies; 18.4%) were the most frequently targeted groups. When categorized by clinical discipline, the largest proportions of studies targeted anesthesiology groups (12 studies; 24.5%), emergency medicine groups (12 studies; 24.5%), and surgical specialists (10 studies; 20.4%). Specialists in different medical specialties (22 studies; 9.44%) and nurses (9 studies; 18.4%) were the target groups for these studies.

The type of simulation methods used for SBAs in health professions education

We classified the simulation techniques into three groups:

Human participant: These comprise the use of cadavers [25], SPs [24, 25, 26], and simulated clinical environments [i.e., theater] [27] that use real or simulated humans as patients or actors.

Equipment-based simulators: anatomical models, APS [23], high fidelity manikins and synthetic models [25, 27], animal tissues and models [27, 28], bronchoscopy [29], colonoscopy [30], and transesophageal echocardiography simulators [31] are examples of simulators that use physical models or devices to replicate human or animal anatomy and physiology.

Computer-based simulators: These simulators generate virtual or enhanced environments using computer hardware or software.

Virtual OSCEs and virtual clinical stations [25], NeuroTouch, Cipro endoscopy, the Mimic™ Flex Virtual Reality (Mimic™ Flex VR) [28, 29], EyeSi [33], and virtual-reality (VR) EBUS simulator test [34] are a few examples.

Target groups examined in SBAs of HPE

The target groups assessed by SBA studies in HPE were categorized into three major categories:

Medical specialists: This category includes both specialist physicians (e.g., surgeons, cardiologists) [29–43] and primary care physicians/family physicians [35, 40, 44].

Emergency-based groups: These include Pre-Hospital Care Providers (PHPs) [45], emergency medicine personnel [25, 46], medical emergency technicians and paramedics [48], and intensive care providers [51].

Nursing and allied health professions: This group included nurses [24, 47–52], respiratory therapists [53], and pharmacy personnel [54, 55].

Applications of SBA methods in HPE

The applications of SBA methods in HPE can be grouped into four key domains:

Assessing integrated clinical competence: SBA methods are used to assess the competence, knowledge, performance, and performance sequence of health professionals in different domains and specialties [49, 56, 58–72]. SBA methods can also evaluate clinical judgment and diagnostic power [56, 60, 62, 65], critical thinking [60], and the technique of obtaining case histories by health professionals [73].

Evaluating skills and abilities: SBA methods are used to assess the technical and non-technical skills and abilities of health professionals, such as procedural skills, clinical skills, teamwork skills, leadership skills, interpersonal skills, patient counseling and guidance abilities, clinical reasoning and decision-making skills, and communication and remote patient management skills [42, 45, 56, 60, 61, 63, 65, 69, 72, 75, 76–81]. SBA methods can also assess deeply reflective and contemplative thinking skills [62], management of uncommon findings and rare cases [62, 85], examination skills for sensitive organs [62, 82], teamwork and interaction skills [62, 65, 72, 83, 84], leadership skills in emergencies [62], and risk management and preparedness in disasters and crises of health professionals [69].

Measuring formative and summative aspects: SBA supports high-stakes decisions such as board certification and licensing [21, 37, 39, 43, 86–88]. It also facilitates evaluation across Miller's pyramid levels [89].

Examining curriculum components: SBA is used to assess educational constructs not easily measured by traditional methods, such as ethics, professionalism, and humanism [37].

Challenges of SBA's methods in HPE

Challenges related to SBA were grouped into four categories:

Realism: Issues included limited psychological fidelity, difficulty generalizing to real-world settings, scenario complexity, and inability to simulate specific human characteristics [41, 58, 60].

Validity: Challenges involved defining and measuring competencies, limited assessment scope, a lack of validated tools, and questions about the high-stakes suitability [41, 56, 59, 60, 62].

Reliability: These included inconsistent scoring, scenario variability, sampling issues, and rater-related errors [60, 64–67].

Feasibility: Practical issues, such as high cost, lack of simulation culture, complexity in design, ethical concerns, and stakeholder coordination, were noted [41, 56, 61, 63, 65].

Advantages of SBA's methods in HPE

Advantages were grouped into four domains:

Quality-related advantages: These are advantages that relate to the accuracy, reliability, and validity of the SBA methods, such as high sensitivity, high specificity [20], high validity [26, 27, 29, 30, 42, 47, 53–55], objectivity of assessment [27, 56], structured assessment high reliability [63], and correlation between simulator and in real performance environment [42, 57].

Learning-related advantages: These are advantages that relate to the satisfaction, motivation, and improvement of the learners and the test takers, such as the satisfaction of learners and test takers [39, 53, 54, 58], reduction of learner and test-taker anxiety [25, 27, 28, 47, 59], complementarity of learning and testing [25, 27], provision of feedback [28, 32, 60, 61], increased interest in simulation [31], and ethical superiority [35, 60].

Realism-related advantages: These are advantages that relate to the similarity, authenticity, and applicability of the SBA methods to the real environment and conditions, such as similarity to the real environment [28, 40, 62], reduction of patient harm [28, 29, 31, 32, 35, 60, 61, 63], usefulness emergency conditions [23], case diversity [59], and replicability of simulation [58].

Equity-related advantages: These are advantages that relate to the fairness, diversity, and accessibility of the SBA methods for different groups and contexts, such as the reduction of disparities in assessment [58].

Disadvantages of SBA's methods in HPE

Our findings indicate that SBAs in HPE face three types of disadvantages, categorized by their type and severity. They are:

Major criticisms: These represent the most significant barriers to the widespread adoption and credibility of SBAs. They include fundamental psychometric concerns, such as limited evidence for validity [64] and fidelity issues (e.g., lack of realism) [63], as well as major practical constraints like high cost [28, 44, 59, 63, 64].

Moderate criticisms: These issues pose notable challenges to implementation and learner experience. They include inducing anxiety in some users [25], logistical concerns such as equipment disposal [28], and limitations in assessing complex or integrated skill sets [65].

Minor criticisms: These are the least important or relevant problems that may arise from SBAs. They include sampling errors and other factors that may vary depending on the context, design, or method of SBAs [25].

Necessary conditions for using SBAs in HPE

The necessary conditions for effective SBA were structured across four key dimensions: (1) Inputs, (2) Processes, (3) Outputs, and (4) Outcomes. These are summarized in **Table 2**, and visually represented in **Figure 2**.

Table 2. Categories of the necessary conditions for using SBA in HPE

Aspect	Factor	Summary of findings
Design	Test objectives [23, 37]	Test objectives should be clearly defined and reviewed by a panel of peers or external experts. Evidence can be used to identify objectives, competencies, and skills.
	Performance standards [41]	Performance standards should be specified and agreed upon by assessors. Assessment checklists should be used to ensure consistency and accuracy.
	Validity [65]	Validity of assessment tools should be measured using frameworks such as Messick's framework of validity theory.
	Reliability [23, 24, 41, 65]	Reliability of assessment can be enhanced by increasing the number of cases, using blind assessors, and standardizing and training assessors.
	Assessment of examinees and standardized patients [23, 45]	Examinees and standardized patients should be trained before the assessment. Prior practice in a simulated testing environment can increase comfort and decrease anxiety for examinees.
	Criterion fidelity (real-world similarity) testing [23, 56]	Fidelity is a complex phenomenon that includes physical and psychological aspects. Reliable equipment enhances fidelity testing.
Implementation	Video recording of examinations [25]	Video recording of examinations using high-powered fiber-optic cameras and microphones is essential for performance assessment.
	Expert opinions on assessment [41]	Experts and knowledgeable sources should be involved in administering assessments. Physicians and faculty members are recommended for validating assessment tools and identifying test cases.
	Feedback [23, 24, 41]	Feedback is necessary for evaluating the assessment method and providing guidance for improvement.
	Curriculum changes when using SBA [57]	An appropriate curriculum for SBA should be designed, considering the field and level of test takers.
	Timing of the assessment [52]	Ample time should be allowed for test takers to complete assessments.
Input	Cost of assessment [52, 59]	SBA is an expensive method that requires careful financial planning and institutional support. Cost-effectiveness of this method should be considered.
	Different simulation methods and techniques [24]	Different simulation methods have specific conditions and advantages. Artificial models and animal tissues can be used for invasive procedures, while electronic simulation can evaluate knowledge and competency.
	The Application of SBA Technique [24]	Simulations used for competence assessment should strive for a high degree of fidelity, and their results should be interpreted carefully. Performance criteria can be identified using evidence.

Abbreviations: SBA, simulation-based assessments; HPE, Health Profession Education.

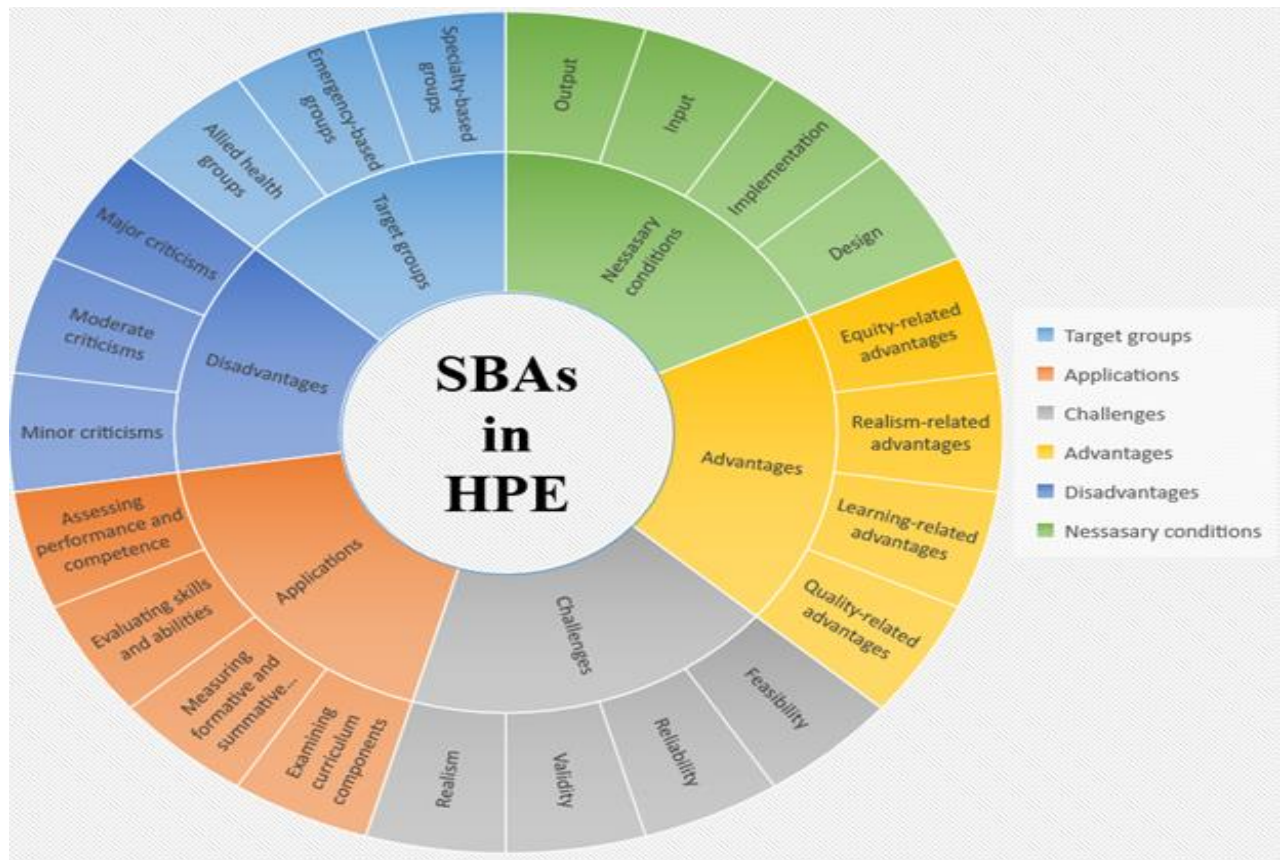


Figure2. Four structural dimensions identified as necessary conditions for SBAs in HPE

Discussion

By examining 49 pertinent studies, this review adds to the body of knowledge already available on simulation-based assessment in health professions education. It provides information on a number of aspects of the subject, such as target audiences, applications, simulation techniques, benefits, drawbacks, and the prerequisites for successful execution.

Our findings demonstrate that a range of simulation modalities are employed for assessment purposes in HPE. Standardized patient simulations and new computer-based simulation techniques are frequently used to assess competencies across a range of skills. A key advantage of these techniques, as identified in the literature, is the capacity for immediate feedback, alongside strong reported evidence of validity and a high degree of patient safety. Cost is still a common restriction, though. SBAs offer different degrees of fidelity and feedback and are versatile enough to accommodate a wide range of simulation modalities and technologies. These results are in line with Ryall et al.'s research, which highlighted how useful SBA is when paired with additional tools and a variety of simulation

scenarios [2]. SBAs, which cover a range of target groups across various disciplines and specialties, are extensively used and researched in health professions education. Three primary target groups were identified: physician groups (including specialists and family physicians) [29–44], emergency and critical care providers [25, 45, 46, 48, 51], and nursing & allied health professionals [24, 47, 49, 50, 52–55]. Similar classifications have been reported in a number of studies [2, 65], underscoring the flexibility and applicability of SBAs to the varied learning goals and needs of health professionals at different phases of their careers. Our research indicates that SBA methods are used in health professions education for a number of purposes, such as evaluating skills and abilities [42, 45, 56, 60, 61, 63, 65, 69, 72, 75–81], assessing performance and competence [49, 56, 58–72], measuring formative and summative outcomes [37, 39, 43, 86–88], and analyzing curriculum components [37]. These approaches are adaptable and can be used to address a variety of learning objectives and results. Nonetheless, there are still issues with SBA practice and literature. A lack of clarity and consistency in SBA

definitions and classifications persists, and many studies highlight the

ongoing challenge of establishing strong validity and reliability evidence for these assessment techniques [65].

Additionally, SBA techniques aren't always properly incorporated into or matched with clinical settings and curricula [2]. According to Boulet et al. [59], it's critical to specify the assessment's goal and the particular competencies that will be assessed. These should be in line with the test taker's level and the assessment type.

In professional health education, it is crucial to evaluate a wide range of competencies, such as management, diagnosis, treatment, and teamwork. Therefore, for SBA to be effective, it must be designed to integrate the assessment of multiple integrated competencies, including knowledge, clinical reasoning, judgment, and communication skills [23, 27–31, 36, 39, 42, 43, 45, 52, 56–58, 60, 64, 65]. The necessity of striking a balance between the evaluation of competencies and specialized skills and the inherent limitations of simulators and assessment formats is highlighted by Alsulimani's study on high-fidelity SBA in emergency medicine. According to the study's findings, SBA shouldn't be the only evaluation technique used [65]. In a similar vein, other researchers also advise combining SBA with additional assessment methods, especially in formative and summative settings [39, 43, 86–88].

Creating tests that are appropriate for staff and specialist levels is another prerequisite for successful simulation-based assessment [25, 37, 61, 65]. According to Ennen et al. [44], teamwork scenarios should be more complex than those created for individual assessments, and they support the creation of simulation-based curricula that take into account the learner's field and skill level. In addition to suggesting that learners be assessed in a clinical setting using SBA in conjunction with other assessment techniques, Rizzolo et al. [52] suggest updating the curriculum to guarantee contextual and conceptual alignment with simulation activities.

Particularly in the areas of realism, validity, reliability, and feasibility, this review highlights significant obstacles and constraints in the application of SBAs in health professions education. These difficulties draw attention to the intricacy and unpredictability of SBAs as an evaluation technique, highlighting the necessity for more study and improvement. They also offer chances and practical suggestions for furthering SBA practice and research in the future.

Future efforts should focus on three key areas: (1) adopting more sophisticated simulators to enhance

fidelity [65], (2) implementing rigorous design and evaluation frameworks to strengthen validity and reliability [66], and (3) utilizing resources more effectively to improve cost-efficiency [67].

The benefits of SBAs are numerous, including those related to quality [23, 26, 27, 29, 30, 38, 42, 47, 52–55, 63], learning [21, 23, 24, 35, 43, 49, 50, 54], realism [24, 25, 27, 28, 31, 56, 57, 59], and equity [58]. These advantages show how useful SBAs are for assessing medical professionals. The effectiveness of SBAs, however, is moderated by several key variables. These include factors such as scenario quality, staff training, the quality of evaluation tools, and the alignment of the simulation with curriculum and practice [65]. Other influencing factors include the expectations of instructors and learners, the degree of realism achieved, the effectiveness of communication, and consideration of relevant ethical and legal issues [66]. Additionally, Ryall and colleagues [2] and Kononowicz and colleagues [67] discovered that SBAs and virtual patient simulations can enhance health professionals' learning outcomes.

As identified in our results, SBA methods face a range of disadvantages, from major limitations to more contextual challenges. These drawbacks present significant challenges to the optimal implementation and scalability of SBA. They do, however, also present opportunities for further study and application. Using inexpensive simulators and internet resources is one way to cut expenses [65]. To reduce anxiety and increase confidence, it is also advised to use formative and supportive SBA techniques, such as feedback, peer evaluation, and self-assessment [66]. A third recommendation is to use advanced and realistic simulators, such as virtual reality, augmented reality, and artificial intelligence, to improve the realism and validity of SBA methods [67]. In their discussion of the benefits and drawbacks of simulation education, Maloney and Haines emphasized the necessity of conducting economic analyses [68]. Simulation-based evaluation requires a lot of resources and works best with personalized formative feedback, according to Sinz et al. [69]. The SBA assessment method is costly and demands investor attention, according to Boulet et al. [63], and Rizzolo et al. also stress the importance of taking the method's cost-effectiveness into account [52].

In HPE, four primary facets of SBA were found to establish its prerequisites. In HPE, the planning and preparation of the SBA methods is referred to as the design aspect of SBA. The implementation facet involves the execution of the assessment, including

administration, logistics, and data collection (e.g., video recording, timing). The terms input and output, respectively, refer to various simulation approaches and strategies as well as the use of the SBA technique. These elements are necessary to guarantee the SBA methods' quality and legitimacy as well as to match them with the desired learning objectives and proficiencies of medical professionals. A critical part of the design is establishing well-defined test objectives, which should be evidence-based and peer-evaluated to ensure they reflect necessary competencies and skills. This result is consistent with the body of research suggesting that test objectives should be based on the curriculum's educational goals and objectives and should cover the important and relevant knowledge, abilities, and attitudes of the health profession [65].

To ensure consistency and accuracy, evaluators must establish and agree upon performance standards and use checklists. Henrico & Makkink found that using checklists improved the validity and reliability of a simulation-based test for anesthesia residents and helped students get feedback and corrective actions [66].

Key strategies to improve reliability include increasing the number of cases, using blind assessors, allowing adequate time for assessment, and standardizing and training assessors [52]. Additionally, according to Boulet et al., proper time management can improve assessment reliability [63]. Synthesize these into a coherent argument.

"To ensure and enhance validity, the literature emphasizes the development of high-quality scenarios through expert consensus and peer review [25, 44, 59, 60]. The involvement of well-trained standardized patients (SPs) and a focus on strong content validity are also critical factors [30, 34, 62, 64–67]. Peer review and the establishment of standards are necessary to preserve validity.

Scenario expert opinions are required [25, 44]. The role of simulator fidelity in validity is complex. While some experts advocate for high-fidelity simulators as a prerequisite for valid assessment [53, 68], other evidence suggests that low-fidelity simulations can effectively assess certain competencies, particularly for experts or when specific skills are targeted [29, 32]. This indicates that the required level of fidelity is context-dependent and should be matched to the specific learning or assessment objective.

The field of simulation-based assessment is changing as a result of the quick development of technologies like machine learning, augmented reality, and artificial

intelligence. Beyond conventional simulation modalities, these tools allow for data-rich, customized, and adaptive assessments. Promising approaches for enhancing formative and summative evaluations in real-time clinical settings include AI-driven feedback systems, AR-enhanced procedure trainers, and predictive analytics dashboards.

Significant gaps still exist in the application of SBA across a range of disciplines, despite its widespread recognition as a useful strategy in health professions education. Notably, there aren't many studies that concentrate on medical physicists, radiologists, radiation therapists, and operating room assistants. A significant research gap was highlighted by the SBA's failure to identify any studies specifically for healthcare professionals in Iran.

Additionally, this scoping review ran into a number of issues. Furthermore, research published in languages other than English and Persian, inaccessible, or incomplete was not included.

In addition, research published in languages other than English and Persian, inaccessible, or incomplete was not included.

Future studies ought to focus on topics like the validity and dependability of SBA tools, the function of remediation and feedback, and the wider effects and ramifications of simulation-based assessments. According to models like Miller's Pyramid, competency-based education frameworks place a strong emphasis on evaluating clinical performance at increasingly higher levels of expertise. These priorities are in line with these frameworks.

Innovation in assessment frameworks and simulation technologies will be crucial as the field develops. Future research should use strict methodological designs and concentrate on creating, assessing, and sharing efficient SBA techniques suited to various health professions in order to make significant strides.

Conclusion

In health professions education, SBAs have the potential to revolutionize the way clinical competency, technical proficiency, and patient safety are evaluated. Their special ability to offer safe, controlled, and authentic settings for evaluating complex competencies—like crisis management and diagnostic reasoning—that frequently elude conventional evaluation techniques is highlighted by this scoping review.

Significant obstacles still stand in the way of their widespread adoption, despite this promise.

High implementation costs, inconsistent realism among simulation modalities, and unanswered questions about their validity and dependability in high-stakes situations are a few of these. A strategic approach is necessary to fully utilize SBAs. To guarantee uniformity and equity, educators and organizations must create standardized standards for scenario design, scoring criteria, and rater training. Accessibility can also be increased by giving priority to affordable solutions like virtual reality platforms or blended simulations.

Importantly, SBAs ought to be incorporated into comprehensive assessment frameworks as supplementary instruments rather than as substitutes. Future studies should concentrate on verifying SBA metrics in a variety of medical specialties, investigating how AI can improve simulations, and assessing how these evaluations affect real clinical outcomes over the long run.

Ethical considerations

The Shahid Beheshti University of Medical Sciences [SBMU] review board approved the study protocols with the Ethics Code of IR.SBMU.SME.REC.1401.079.

Artificial intelligence utilization for article writing

During the preparation of this manuscript, ChatGPT (version 3.5) was utilized to enhance writing quality and correct grammatical errors.

Acknowledgment

We would like to thank our colleagues for their efforts in working on this project.

Conflict of interest statement

None

Author contributions

ZKh, SB, and ZS contributed to the study concept and design. The search protocol was designed by ZKh, SB, and ZS. ZS performed the literature searches and data acquisition.

Abstract and full-text screening were conducted by ZKh, SB, ZS, and SS. SS and ZS were responsible for the revision of the results and writing the manuscript. All authors reviewed and approved the final manuscript.

Funding

None

Data availability statement

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. O'Donovan BM, den Outer B, Price M, Lloyd A. What makes good feedback good? *Stud High Educ.* 2021;46(2):318–29. <https://doi.org/10.1080/03075079.2019.1630812>
2. Ryall T, Judd BK, Gordon CJ. Simulation-based assessments in health professional education: a systematic review. *J Multidiscip Healthc.* 2016;9:69–82. <https://doi.org/10.2147/JMDH.S92695>
3. Cant RP, Cooper SJ. Simulation-based learning in nurse education: systematic review. *J Adv Nurs.* 2010;66(1):3–15. <https://doi.org/10.1111/j.1365-2648.2009.05240.x>
4. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA.* 2002;287(2):226–35. <https://doi.org/10.1001/jama.287.2.226>
5. Swanson DB, Norman GR, Linn RL. Performance-based assessment: lessons from the health professions. *Educ Res.* 1995;24(5):5–11. <https://doi.org/10.3102/0013189X024005005>
6. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach.* 2007;29(9–10):855–71. <https://doi.org/10.1080/01421590701775453>
7. Jones F, Passos-Neto CE, Braghiroli OFM. Simulation in medical education: brief history and methodology. *Princ Pract Clin Res.* 2015;1(2). <https://doi.org/10.21801/ppcrj.2015.12.8>
8. Qayumi K, Pachev G, Zheng B, Ziv A, Koval V, Badieli S, et al. Status of simulation in health care education: an international survey. *Adv Med Educ Pract.* 2014;5:457–67. <https://doi.org/10.2147/AMEP.S65451>
9. Gaba DM. The future vision of simulation in health care. *BMJ Qual Saf.* 2004;13 Suppl 1:i2–i10. <https://doi.org/10.1136/qshc.2004.009878>
10. Koukourikos K, Tsaloglidou A, Kourkouta L, Papathanasiou IV, Iliadis C, Fratzana A, et al. Simulation in clinical nursing education. *Acta Inform Med.* 2021;29(1):15–20. <https://doi.org/10.5455/aim.2021.29.15-20>
11. Ziv A, Wolpe PR, Small SD, Glick S. Simulation-based medical education: an ethical imperative. *Simul Healthc.* 2006;1(4):252–6. <https://doi.org/10.1097/01.SIH.0000242724.08501.63>
12. Ker J, Mole L, Bradley P. Early introduction to interprofessional learning: a simulated ward environment. *Med Educ.* 2003;37(3):248–55.

- <https://doi.org/10.1046/j.1365-2923.2003.01439.x>
13. Rosen KR. The history of medical simulation. *J Crit Care*. 2008;23(2):157–66.
<https://doi.org/10.1016/j.jerc.2007.12.004>
 14. Bland AJ, Topping A, Wood B. A concept analysis of simulation as a learning strategy in the education of undergraduate nursing students. *Nurse Educ Today*. 2011;31(7):664–70.
<https://doi.org/10.1016/j.nedt.2010.10.013>
 15. Seropian MA, Brown K, Gavilanes JS, Driggers B. An approach to simulation program development. *J Nurs Educ*. 2004;43(4):170–4.
<https://doi.org/10.3928/01484834-20040401-02>
 16. Tun JK, Alinier G, Tang J, Kneebone RL. Redefining simulation fidelity for healthcare education. *Simul Gaming*. 2015;46(2):159–74.
<https://doi.org/10.1177/1046878115576103>
 17. Foronda C, Gattamorta K, Snowden K, Bauman EB. Use of virtual clinical simulation to improve communication skills of baccalaureate nursing students: a pilot study. *Nurse Educ Today*. 2014;34(6):e53–e7.
<https://doi.org/10.1016/j.nedt.2013.10.007>
 18. Rizzolo MA, Kardong-Edgren S, Oermann MH, Jeffries PR. The National League for Nursing Project to Explore the Use of Simulation for High-Stakes Assessment: process, outcomes, and recommendations. *Nurs Educ Perspect*. 2015;36(5):299–303. <https://doi.org/10.5480/15-1639>
 19. Zhang L, Patterson F, Penman A, Forbes R. Understanding simulation-based learning for health professions students from culturally and linguistically diverse backgrounds: a scoping review. *Adv Health Sci Educ*. 2024.
<https://doi.org/10.1007/s10459-024-10384-6>
 20. Khalil H, Jia R, Moraes EB, Munn Z, Alexander L, Peters MDJ, et al. Scoping reviews and their role in identifying research priorities. *J Clin Epidemiol*. 2025;181:111712.
<https://doi.org/10.1016/j.jclinepi.2025.111712>
 21. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci*. 2010;5:69. <https://doi.org/10.1186/1748-5908-5-69>
 22. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–9.
<https://doi.org/10.1136/bmj.b2535>
 23. Peters MDJ, Godfrey CM, McInerney P, Soares CB, Khalil H, Parker D. The Joanna Briggs Institute reviewers' manual 2015: methodology for JBI scoping reviews. 2015.
<https://doi.org/10.46658/JBIRM-20-01>
 24. Heiner JD, Proffitt AM, McArthur TJ. The ability of emergency nurses to detect simulated long bone fractures with portable ultrasound. *Int Emerg Nurs*. 2011;19(3):120–4.
<https://doi.org/10.1016/j.ienj.2010.08.004>
 25. Srinivasan M, Hwang JC, West D, Yellowlees PM. Assessment of clinical skills using simulator technologies. *Acad Psychiatry*. 2006;30(6):505–15.
<https://doi.org/10.1176/appi.ap.30.6.505>
 26. Weidner AC, Gimpel JR, Boulet JR, Solomon M. Using standardized patients to assess the communication skills of graduating physicians for the comprehensive osteopathic medical licensing examination (COMLEX) level 2-performance evaluation (level 2-PE). *Teach Learn Med*. 2010;22(1):8–15.
<https://doi.org/10.1080/10401330903445604>
 27. Dunkley M. Competence assessment using simulation. *Minim Invasive Ther Allied Technol*. 2000;9(5):341–5.
<https://doi.org/10.3109/13645700009061456>
 28. Shah J, Darzi A, editors. Simulation and skills assessment. In: Proceedings of the International Workshop on Medical Imaging and Augmented Reality; 2001 Jun 10–12; Shatin, Hong Kong, China. IEEE; 2001. p. 8–12.
<https://doi.org/10.1109/MIAR.2001.930256>
 29. Steinfert DP, Yong YH, Byrne T, Gorelik A, Colt H, Irving LB. Assessment of bronchoscopic dexterity and procedural competency in a low-fidelity simulation model. *J Bronchology Interv Pulmonol*. 2018;25(3):198–203.
<https://doi.org/10.1097/LBR.0000000000000481>
 30. Black S, Nestel D, Kneebone R, Wolfe J. Assessment of surgical competence at carotid endarterectomy under local anaesthesia in a simulated operating theatre. *Br J Surg*. 2010;97(4):511–6. <https://doi.org/10.1002/bjs.6938>
 31. Crosby E. The role of simulator-based assessments in physician competency evaluations. *Can J Anaesth*. 2010;57(7):627–35.
<https://doi.org/10.1007/s12630-010-9323-3>
 32. Glavin RJ, Gaba DM. Challenges and opportunities in simulation and assessment. *Simul Healthc*. 2008;3(2):69–71.
<https://doi.org/10.1097/SIH.0b013e31817bb8f6>
 33. Grantcharov TP, Carstensen L, Schulze S. Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. *JSLS*. 2005;9(2):130–3.
<https://pubmed.ncbi.nlm.nih.gov/15984697/>
 34. Konge L, Annema J, Clementsen P, Minddal V, Vilmann P, Ringsted C. Using virtual-reality simulation to assess performance in endobronchial ultrasound. *Respiration*. 2013;86(1):59–65.
<https://doi.org/10.1159/000350428>
 35. Levine AI, Schwartz AD, Bryson EO, DeMaria Jr S. Role of simulation in US physician licensure and

- certification. *Mt Sinai J Med*. 2012;79(1):140–53. <https://doi.org/10.1002/msj.21291>
36. Lien CA, Warner MA, Rathmell JP. Simulation for assessment of the practice of board-certified anesthesiologists. *Anesthesiology*. 2017;127(3):410–2. <https://doi.org/10.1097/ALN.0000000000001745>
37. Macario A. Can physician performance be assessed via simulation? *Anesthesiology*. 2014;120(1):18–21. <https://doi.org/10.1097/ALN.0000000000000056>
38. Noveanu J, Amsler F, Ummenhofer W, von Wyl T, Zuercher M. Assessment of simulated emergency scenarios: are trained observers necessary? *Prehosp Emerg Care*. 2017;21(4):511–24. <https://doi.org/10.1080/10903127.2017.1302528>
39. Qian J, Rama A, Wang E, Wang T, Hess O, Khoury M, et al. Assessing pediatric life support skills using augmented reality medical simulation with eye tracking: a pilot study. *J Educ Perioper Med*. 2022;24(3):E692. https://doi.org/10.46374/volxxiv_issue3_qian
40. Ross BK, Metzner J. Simulation for maintenance of certification. *Surg Clin North Am*. 2015;95(4):893–905. <https://doi.org/10.1016/j.suc.2015.04.010>
41. Rovamo L, Mattila MM, Andersson S, Rosenberg P. Assessment of newborn resuscitation skills of physicians with a simulator manikin. *Arch Dis Child Fetal Neonatal Ed*. 2011;96(5):F383–9. <https://doi.org/10.1136/adc.2010.194043>
42. Sanford DI, Ma R, Ghoreifi A, Haque TF, Nguyen JH, Hung AJ. Association of suturing technical skill assessment scores between virtual reality simulation and live surgery. *J Endourol*. 2022;36(10):1388–94. <https://doi.org/10.1089/end.2022.0158>
43. Zupanc CM, Wallis GM, Hill A, Burgess-Limerick R, Riek S, Plooy AM, et al. Assessing colonoscopic inspection skill using a virtual withdrawal simulation: a preliminary validation of performance metrics. *BMC Med Educ*. 2017;17(1):118. <https://doi.org/10.1186/s12909-017-0948-6>
44. Ennen CS, Satin AJ. Training and assessment in obstetrics: the role of simulation. *Best Pract Res Clin Obstet Gynaecol*. 2010;24(6):747–58. <https://doi.org/10.1016/j.bpobgyn.2010.03.003>
45. Skube ME, Witthuhn S, Mulier K, Boucher B, Luszczek E, Beilman GJ. Assessment of pre-hospital hemorrhage and airway care using a simulation model. *J Trauma Acute Care Surg*. 2018;85(1):27–32. <https://doi.org/10.1097/TA.0000000000001800>
46. O'Reilly DJ. Computer simulation in mass emergency and disaster response: an evaluation of its effectiveness as a tool for demonstrating strategic competency in emergency department medical responders [dissertation]. Detroit (MI): Wayne State University; 2011.
47. Crannell C. Chemotherapy administration: using simulation case-based scenarios to assess chemotherapy competency. *Oncol Nurs Forum*. 2012;39(1):19–22. <https://doi.org/10.1188/12.ONF.19-22>
48. Heyes ME, Schnitzen L, Starr DG, Vacchiano C, Muckler VC, Thiemann L, et al. Recertification and reentry to practice for nurse anesthetists: determining core competencies and evaluating performance via high-fidelity simulation technology. *J Nurs Regul*. 2018;8(4):43–55. [https://doi.org/10.1016/S2155-8256\(17\)30181-3](https://doi.org/10.1016/S2155-8256(17)30181-3)
49. Hinton JE, Mays MZ, Hagler D, Randolph P, Brooks R, DeFalco N, et al. Measuring post-licensure competence with simulation: the nursing performance profile. *J Nurs Regul*. 2012;3(2):45–53. [https://doi.org/10.1016/S2155-8256\(15\)30219-2](https://doi.org/10.1016/S2155-8256(15)30219-2)
50. Hock S, Sergel M, Christian E, Carizey R, Perumalsamy P, Patwari R, et al. Development and assessment of a simulation model of focused and ongoing professional practice evaluations. *Ann Emerg Med*. 2016;68(4):S114–5. <https://doi.org/10.1016/j.annemergmed.2016.08.308>
51. Kardong-Edgren S, Hanberg AD, Keenan C, Ackerman AA, Chambers K. A discussion of high-stakes testing: an extension of a 2009 INACSL conference roundtable. *Clin Simul Nurs*. 2011;7(1):e19–24. <https://doi.org/10.1016/j.ecns.2010.02.002>
52. Rizzolo MA, Kardong-Edgren S, Oermann MH, Jeffries PR. The National League for Nursing project to explore the use of simulation for high-stakes assessment: process, outcomes, and recommendations. *Nurs Educ Perspect*. 2015;36(5):299–303. <https://doi.org/10.5480/15-1639>
53. LeFlore JL, Anderson M. Effectiveness of 2 methods to teach and evaluate new content to neonatal transport personnel using high-fidelity simulation. *J Perinat Neonatal Nurs*. 2008;22(4):319–28. <https://doi.org/10.1097/01.JPN.0000341364.60693.70>
54. Daupin J, Atkinson S, Bédard P, Pelchat V, Lebel D, Bussi eres JF. Medication errors room: a simulation to assess the medical, nursing and pharmacy staffs' ability to identify errors related to the medication-use system. *J Eval Clin Pract*. 2016;22(6):911–20. <https://doi.org/10.1111/jep.12558>
55. MacFarlane B, Bergin J, Peterson GM. Assessment and management of serotonin syndrome in a simulated patient study of Australian community pharmacies. *Pharm Pract (Granada)*.

- 2016;14(2):703.
<https://doi.org/10.18549/PharmPract.2016.02.703>
56. Lee KHK, Grantham H, Boyd R. Comparison of high and low-fidelity mannequins for clinical performance assessment. *Emerg Med Australas*. 2008;20(6):508–14. <https://doi.org/10.1111/j.1742-6723.2008.01137.x>
57. Thomsen ASS, Smith P, Subhi Y, Cour ML, Tang L, Saleh GM, et al. High correlation between performance on a virtual-reality simulator and real-life cataract surgery. *Acta Ophthalmol*. 2017;95(3):307–11.
<https://doi.org/10.1111/aos.13275>
58. Alotaibi FE, AlZhrani GA, Mullah MAS, Sabbagh AJ, Azarnoush H, Winkler-Schwartz A, et al. Assessing bimanual performance in brain tumor resection with NeuroTouch, a virtual reality simulator. *Oper Neurosurg (Hagerstown)*. 2015;11(1):89–98.
<https://doi.org/10.1227/NEU.0000000000000631>
59. Brazil V. Performance assessment and simulation fidelity for dummies. *Emerg Med Australas*. 2008;20(5):456–7.
<https://doi.org/10.1111/j.1742-6723.2008.01129.x>
60. Carroll JD, Messenger JC. Medical simulation: the new tool for training and skill assessment. *Perspect Biol Med*. 2008;51(1):47–60.
<https://doi.org/10.1353/pbm.2008.0003>
61. Hall AK, Chaplin T, McColl T, Petrosioniak A, Caners K, Rocca N, et al. Harnessing the power of simulation for assessment: consensus recommendations for the use of simulation-based assessment in emergency medicine. *CJEM*. 2020;22(2):194–203.
62. Decker S, Utterback VA, Thomas MB, Mitchell M, Sportsman S. Assessing continued competency through simulation: a call for stringent action. *Nurs Educ Perspect*. 2011;32(2):120–5.
<https://doi.org/10.5480/1536-5026-32.2.120>
<https://doi.org/10.1017/cem.2019.488>
63. Boulet JR. Summative assessment in medicine: the promise of simulation for high-stakes evaluation. *Acad Emerg Med*. 2008;15(11):1017–24.
<https://doi.org/10.1111/j.1553-2712.2008.00228.x>
64. Schuwirth LW, Van der Vleuten CP. The use of clinical simulations in assessment. *Med Educ*. 2003;37 Suppl 1:65–71.
<https://doi.org/10.1046/j.1365-2923.37.s1.8.x>
65. Bick JS, DeMaria Jr S, Kennedy JD, Schwartz AD, Weiner MM, Levine AI, et al. Comparison of expert and novice performance of a simulated transesophageal echocardiography examination. *Simul Healthc*. 2013;8(5):329–34.
<https://doi.org/10.1097/SIH.0b013e31829068df>
66. Holzmacher J, Sarani B, Puche A, Granite G, Shalin V, Pugh K, et al. Can hyper-realistic physical models of peripheral vessel exposure and fasciotomy replace cadavers for performance assessment? *J Trauma Acute Care Surg*. 2017;83(1):130–5.
<https://doi.org/10.1097/TA.0000000000001419>
67. Podolsky DJ, Fisher DM, Riff KWW, Zuker RM, Drake JM, Forrest CR. Assessing performance in simulated cleft palate repair using a novel video recording setup. *Cleft Palate Craniofac J*. 2020;57(6):687–93.
<https://doi.org/10.1177/1055665620913178>
68. Nabzdyk CS, Bittner EA. One (not so small) step for simulation-based competency assessment in critical care. *Crit Care Med*. 2018;46(6):1026–7.
<https://doi.org/10.1097/CCM.0000000000003101>