Revista Española de Educación Médica

UNIVERSIDAD DE MURCIA

Beyond technological realism: high-fidelity clinical simulation for competency-based training in medical students. A systematic review.

Beyond technological realism: high-fidelity clinical simulation for competence-based training in medical students. A systematic review.

Yoiler Batista Garcet [1,] Lissette Beatriz Villavicencio Cedeño [2,] Jacqueline Macías Mendoza [3,] Marieta del Jesús Azúa Menéndez

[1-3] ybatista@sangregorio.edu.ec, https://orcid.org/0000-0002-7851-576 ; https://orcid.org/0000-0002-8779-9091 ; San Gregorio University of Portoviejo, Department of Medicine, Portoviejo, Manabí, Ecuador.

[2] liso_villa@hotmail.com, https://orcid.org/0000-0001-8475-3947 ; National University of Tumbes, Peru; Cristo Rey Educational Unit, Portoviejo, Manabí, Ecuador.

[4] Southern State University of Manabí, Professor of Clinical Laboratory Science, Jipijapa, Manabí, Ecuador. https://orcid.org/0000-0002-5601-6621

* Correspondence: ybatista@sangregorio.edu.ec

Received: 4/8/26; Accepted: 1/9/26; Published: 3/9/26

Summary.

Introduction: High-fidelity clinical simulation is used to train complex skills before students interact with patients independently, although doubts remain about the added value that technological realism offers over more economical pedagogical designs. **Objective:** To determine the effectiveness of high-fidelity clinical simulation in developing skills in medical students and to identify the pedagogical conditions associated with improved performance. **Methods:** A systematic review was conducted according to PRISMA 2020 guidelines. Studies published in PubMed/MEDLINE, Scopus, Web of Science, ERIC, CINAHL, SciELO, and LILACS from January 2010 to August 2, 2026, were included. Quasi-experimental and experimental studies involving medical students were also included. The synthesis was performed narratively using SWiM. RoB 2 or ROBINS-I was used to assess the risk of bias. **Results:** Three hundred records were found, and 30 studies containing a total of 4,216 students were included. High-fidelity simulation showed consistent benefits in collaborative work, clinical reasoning, impairment identification, and procedural efficiency, especially when it involved intentional practice, observations, and debriefing. However, in direct comparisons with active or low-fidelity strategies, superiority was variable. Retention was assessed by ten studies, and clinical outcomes and transfer to real patients were excellent. Overconfidence and stress were identified as potentially harmful consequences. **Conclusions:** Regarding individual skills, high-fidelity simulation is most beneficial for integrated and dynamic skills. Effectiveness depends primarily on the consistency between objectives, debriefing, assessment, difficulty, and practice. It is recommended to select the minimum fidelity required for competence and to verify transfer and retention.

Keywords: clinical simulation; high fidelity; medical education; competency-based training; medical students; deliberate practice; debriefing; systematic review.

Abstract.

Introduction: High-fidelity clinical simulation is used to train complex competencies before students undertake autonomous patient care, yet uncertainty remains regarding the added value of

technological realism over less costly educational designs. **Objective:** To evaluate the effectiveness of high-fidelity clinical simulation for competence development in medical students and identify the educational conditions associated with better outcomes. **Methods:** A systematic review was conducted in accordance with PRISMA 2020. Studies published from January 2010 to 2 August 2026 were searched in PubMed/MEDLINE, Scopus, Web of Science, ERIC, CINAHL, SciELO and LILACS. Experimental and quasi-experimental studies involving undergraduate medical students were eligible. Narrative synthesis followed SWiM; risk of bias was assessed using RoB 2 or ROBINS-I. **Results:** Three hundred records were identified and 30 studies involving 4,216 students were included. High-fidelity simulation frequently improved procedural performance, clinical reasoning, recognition of deterioration and teamwork, particularly when deliberate practice, feedback and structured debriefing were present. However, direct comparisons with lower-fidelity or active educational alternatives showed inconsistent superiority. Ten studies assessed retention; transfer to real patients and clinical outcomes were rarely examined. Stress and overconfidence were potential adverse effects. **Conclusions:** High-fidelity simulation is most useful for integrated, dynamic competencies rather than isolated skills. Effectiveness depends primarily on alignment among objectives, difficulty, practice, debriefing and assessment. The minimum sufficient fidelity should be selected for each competence, with explicit evaluation of retention and transfer.

Keywords: clinical simulation; high fidelity; medical education; competence-based education; medical students; deliberate practice; debriefing; systematic review.

1. Introduction

For competency-based medical education to be effective, students must demonstrate observable performance that integrates technical skills, communication, professionalism, clinical reasoning, and patient safety (1, 2). This perspective modifies the logic of the curriculum, as learning is no longer measured solely by exposure time, but by the ability to act competently in clinical situations. Consequently, pedagogical strategies must provide opportunities to practice in safe conditions, observe performance, and receive feedback before assuming progressively autonomous responsibilities in patient care. Clinical simulation addresses this need by reproducing care situations in controlled environments, where it is possible to practice, make mistakes, and analyze their consequences without generating risk to the patient (3, 4). High-fidelity simulation typically incorporates immersive environments, computerized mannequins, and systems capable of modifying physiological parameters in response to the participants' decisions. Among medical students, experiences published in the Spanish Journal of Medical Education have shown a favorable perception of high-fidelity clinical simulation and its usefulness in strengthening clinical skills and confidence during learning (5). However, fidelity is a multidimensional concept that includes physical, functional, psychological, conceptual, and sociological components. Therefore, a technologically advanced simulator may be pedagogically inadequate if it does not reproduce the essential elements of the task or competency that is intended to be developed (6, 7).

The accumulated evidence on simulation-based education shows favorable effects on knowledge, skills, and performance, especially when interventions incorporate repetition, feedback, deliberate practice, and structured debriefing (8-14). However, significant limitations persist in the literature, including variable high-fidelity definitions, small sample sizes, the use of instruments with insufficiently documented validity evidence, the predominance of assessments immediately following the intervention, and the frequent use of self-reported outcomes, such as satisfaction and confidence, instead of objective indicators of competence (15). In this regard, specific instruments have also been developed to assess the quality and satisfaction of medical students with clinical simulation, highlighting the importance of distinguishing these perceptual variables from the objective demonstration of competency performance (16). Furthermore, greater technological complexity can increase costs, infrastructure needs, maintenance requirements, and personnel without necessarily guaranteeing a proportional improvement in learning.

Recent systematic reviews have identified benefits of simulation on the performance of medical students and have specifically analyzed the use of high-fidelity simulators (17, 18). Likewise, a systematic review recently published in the Spanish Journal of Medical Education showed that the use of anatomical models and simulated scenarios can promote the development of clinical-surgical skills, theoretical learning, performance, and satisfaction among medical students (19). However, this background information does not allow us to conclusively establish whether greater technological fidelity offers consistent educational advantages compared to active strategies or lower-fidelity modalities, nor under what pedagogical conditions such advantages occur. Consequently, the need remains to analyze high fidelity from a competency-based perspective, rather than solely focusing on the sophistication of the device. The relevant question is not only whether simulation is more realistic, but also what competencies it helps develop, against what comparator, through what instructional design, and with what evidence of retention and transfer.

In this context, the present review expands upon previous syntheses by integrating analyses of fidelity levels, educational comparators, competencies and assessment instruments, instructional design characteristics, learning retention, transfer of learning, the emotional effects associated with the experience, and the risk of bias. In this way, the focus shifts from technological realism to the alignment between the fidelity used, the competency to be developed, and the pedagogical conditions that underpin learning.

This systematic review aimed to evaluate the effectiveness of high-fidelity clinical simulation in developing competencies in medical students and to identify the pedagogical conditions associated with better outcomes. Additionally, the levels and characteristics of fidelity used, the instruments and competencies assessed, learning retention, transfer of learning, and the risk of bias in the included studies were analyzed.

2. Methods

2.1. Reporting design and guidelines

A systematic review of research on educational interventions was conducted. The PRISMA 2020 statement (20) guided the identification, selection, synthesis, and communication of the results. Due to the diversity in populations, comparators, instruments, and follow-up periods, the synthesis principles without meta-analysis from the SWiM guide (21) were used. The protocol was not prospectively recorded; this was considered a limitation, and the methodological decisions used are explicitly described to promote reproducibility.

2.2. Review Question

The question was formulated using the PICO method: In medical students (P), does high-fidelity clinical simulation or a physiologically responsive simulation environment (I) improve, compared with traditional methods such as case discussion, video, teamwork, technical skills, communication, overall performance, or other active approaches (C), applied knowledge, clinical reasoning, or transfer and retention (O)? Anxiety, mental workload, stress, and confidence were considered as secondary outcomes when related to learning or implementation.

2.3. Sources of information and period

The search was conducted in the CINAHL, ERIC, LILACS, PubMed/MEDLINE, SciELO, Scopus, and Web of Science Core Collection databases. This was supplemented by reference tracing and manual searches for studies identified in relevant reviews. The search period spanned from January 1, 2010, to August 2, 2026. Publications in English, Spanish, and Portuguese were considered.

2.4. Search Strategy

Free-text and controlled vocabulary related to competencies, undergraduate medical education, and high-fidelity simulation were combined. PubMed was based on the following equation:

(("Simulation Training"[Mesh] OR "high-fidelity simulation"[tiab] OR "high fidelity simulation"[tiab] OR "maniquin*"[tiab] OR "human patient simulator"[tiab]) AND ("medical education, undergraduate"[Mesh] OR "medical students"[tiab]) AND (competenc*[tiab] OR skill*[tiab])) AND ("2010/01/01"[Date - Publication]: "2026/08/02"[Date - Publication]).

The syntax was adapted to the fields and thesauri of each database. The supplementary material includes complete strategies, as well as dates and returns by source.

2.5. Eligibility criteria

- Experimental, quasi-experimental or controlled investigations that carry out prior and subsequent measurements.
- Individuals enrolled in undergraduate medical programs; mixed samples were only accepted if the medical-related data were distinct.
- Interventions in which the essential element is a highly complex hybrid environment, a physiologically responsive mannequin, or high fidelity.
- Analysis of a minimum of one learning outcome, performance, retention, transfer or emotional variable related to the educational experience.
- Original articles that have been peer-reviewed and are fully accessible.

Studies focusing solely on standardized patients, studies involving only professionals or residents, purely descriptive works, protocols, conference abstracts, editorials, and reviews were excluded. Publications lacking sufficient data to detail the intervention or outcomes were also eliminated.

2.6. Selection of studies

The retrieved records were entered into a bibliographic management system, and duplicates were removed. Two reviewers independently conducted the study selection process in two phases. In the first phase, both independently assessed the titles and abstracts of the identified records; subsequently, potentially eligible articles were examined in full text. Discrepancies arising during both phases were resolved through discussion and consensus between the two reviewers. The participation of a third reviewer was not required. The reasons for excluding the studies reviewed in full text were recorded in a categorical manner. The process of identifying, screening, determining eligibility, and including studies was represented using a PRISMA 2020 flow diagram.

2.7. Data Extraction

Two reviewers independently extracted the data using a pre-standardized matrix. The extracted data were then compared between the two reviewers, and any discrepancies were resolved through discussion and consensus. The matrix included information on author, year, country, study design, sample, curriculum level, clinical area, definition of high fidelity, comparator, duration and number of sessions, as well as competency objectives, prebriefing, deliberate practice, feedback, debriefing, instruments used, and assessment timing.

2.8. Assessment of the risk of bias

Randomized trials were assessed using RoB 2, taking into account allocation, intervention bias, missing data, measurement, and outcome selection (22). Non-randomized studies were assessed using ROBINS-I, taking into account confounding, selection, intervention classification, bias, missing data, measurement, and reporting (23). Judgments were reported by domain rather than a total score. Additional material includes the assessment matrix.

2.9. Synthesis Strategy

A narrative synthesis was conducted given the heterogeneity of methods and clinical settings. The studies were grouped into: a) technical skills and integrated performance; b) communication, leadership, and teamwork; c) comparison of fidelity levels or active tactics; d) retention and transfer; e) emotion-related variables; and f) knowledge and reasoning. Differences between groups and objective measures were highlighted. A meta-analysis was not performed because the instruments, scales, and comparators did not allow for the use of a common metric without losing sight of the educational context.

2.10. Ethical considerations

The review was conducted using published data and did not contain individually identifiable information; therefore, informed consent and ethics committee approval were not required. Transparency was maintained in the supplementary instruments used, in the tracking of sources, and in academic integrity.

3. Results

3.1. Selection of studies

Three hundred studies were identified: 292 in databases and eight through manual search and reference tracing. Seventy-four duplicate items and five records were removed before selection. One hundred and sixty-one of the 221 analyzed titles, along with their abstracts, were excluded. An attempt was made to retrieve 60 full texts; however, four could not be obtained. Of the 56 reports analyzed for eligibility, 26 were excluded because they did not meet the criteria for population (10 cases), intervention (9), outcome (8), design (3), or had insufficient data (2). The narrative synthesis ultimately included 30 studies (Figure 1).

3.2. General characteristics

The 30 studies (Table 1) were published between 2010 and 2026 and included 4,216 students; the median sample size was 100 (range: 20–615). Fourteen studies were from Europe, seven from North America, four from Latin America, four from Asia, and one from Africa. Sixteen used randomized or crossover designs, six used non-randomized comparator designs, and eight used single-group or before-and-after designs. Twenty-seven incorporated at least one objective measure of knowledge or performance; three focused primarily on anxiety, stress, or confidence. Ten studies conducted follow-up after the immediate assessment, but only a small number examined transfer to actual patients. The clinical areas were diverse: auscultation and cardiovascular or respiratory examination (n=6), emergency and critical care (n=9), pediatrics and neonatology (n=5), resuscitation and teamwork (n=3), diagnostic or airway procedures (n=3), geriatrics (n=1), neurology (n=1), and predominantly affective or workload studies (n=2). This heterogeneity justified the narrative synthesis.

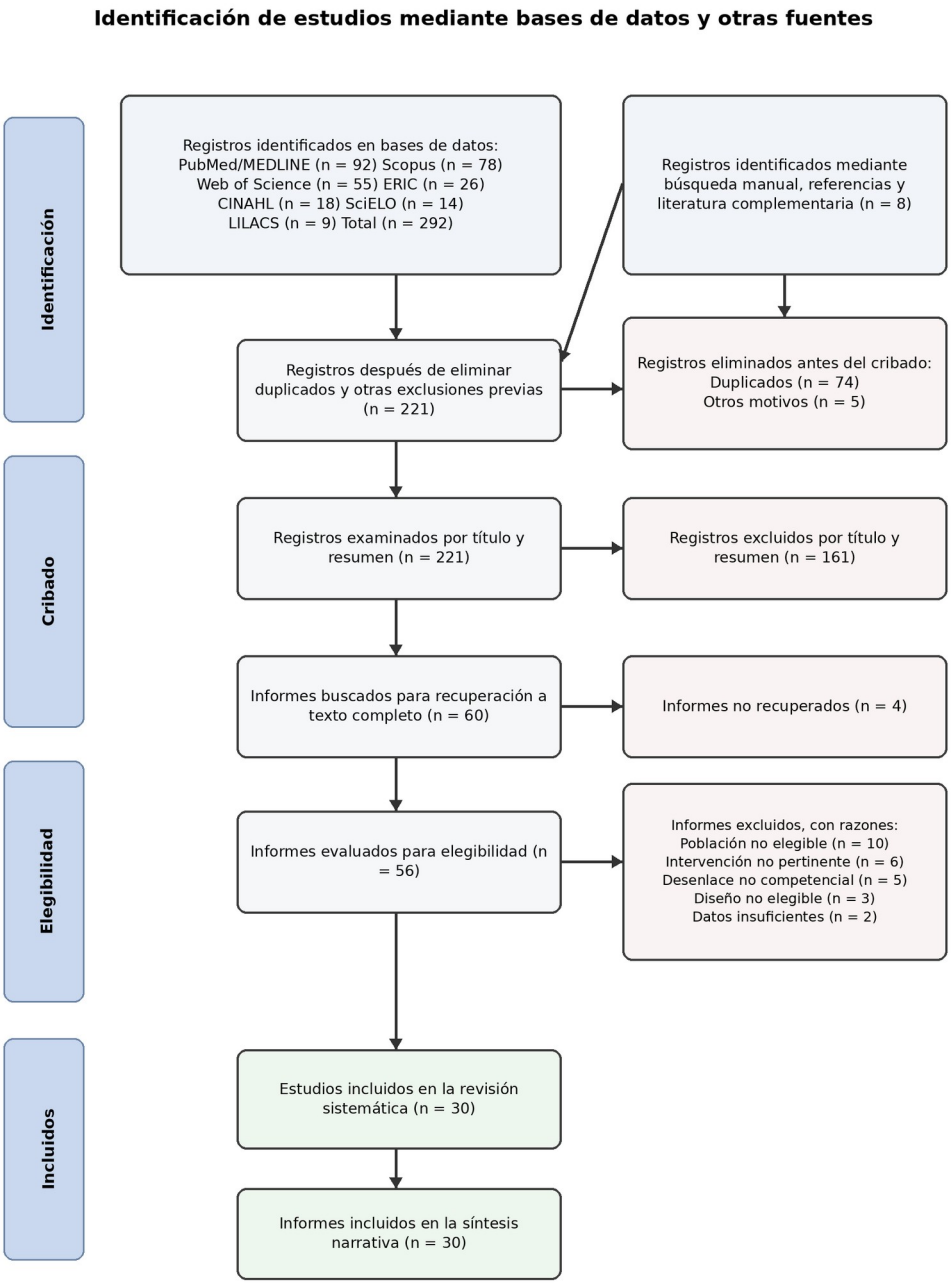


Figure 1. Flow diagram of the process following PRISMA 2020.

Table 1. Aggregate characteristics of the included studies.

Feature	Category	Result
Period	Years of publication	2010-2026
Participants	Total / Median / Range	4.216 / 100 / 20-615
Region	Europe; North America; Latin America; Asia; Africa	14; 7; 4; 4; 1

Design	Randomized or crossover; non-randomized controlled; one group/before-after	16; 6; 8
Measurement	At least one objective outcome; primarily affective/physiological	27; 3
Follow-up	Evaluation following the immediate post-test	10 studies
Synthesis	Meta-analysis	Not performed due to heterogeneity

3.3. Clinical knowledge and reasoning

Most studies showed pre-post improvement after simulation, but the advantage over active comparators was variable. McCoy et al. observed 93% correct answers with simulation versus 71% with instruction (24), while Cortegiani et al. reported a post-test score of 74.1 versus 65.5 in advanced life support (25). DeWaay et al. found better performance on a clinical examination of acute coronary syndrome (26), and Mutter et al. documented better reasoning using the Script Concordance Test when case analysis was supplemented with a high-fidelity manikin (27). The superiority was not universal. Solymos et al. found a greater immediate gain, but no difference at two weeks (28). Although satisfaction was higher with simulation (29), Couto et al. did not detect variations in knowledge acquisition or retention compared to case discussions. According to Moliterno et al., both methods increased confidence and knowledge, but the high-fidelity method was superior in terms of reasoning, communication, attitude, and leadership (30). These findings indicate that the advantage is most evident when the outcome involves merging dynamic information and taking action, rather than simply recalling content.

3.4. Technical skills and integrated performance

Skills requiring coordinated execution, sign recognition, and prioritization showed the most consistent effects. Auscultation, both in simulators and real patients, improved thanks to mastery learning (31). Kern et al. achieved superior performance in all five cardiovascular competencies of the OSCE (32), and Perlini et al. reported that recognition of heart sounds increased from 11% to 72%, and was sustained in a smaller sample (33). In neurological emergencies, repeated sessions produced competency progression, with faster acquisition of technical skills (34). In resuscitation and procedures, Berger et al. found a higher proportion of adequate CPR immediately after the intervention, although the advantage decreased at six months (35). Zhao et al. demonstrated better acquisition of echocardiographic views at one week and one month (36). In AIR-MASTER-SIM, training with videolaryngoscopy within a high-fidelity environment resulted in greater success, fewer attempts, and fewer complications than direct laryngoscopy (37); this study provides evidence of how task design and shared viewing can modify learning within the same simulation modality.

3.5. Non-technical skills

Communication, leadership, and teamwork benefited when they were explicit objectives and assessed using observational instruments. Gonçalves et al. found improvements in technical performance and TeamSTEPPS during pediatric resuscitation (38). Moliterno et al. reported large effect sizes for reasoning, communication, attitude, and leadership (Hedges' g 1.15–2.20) despite not finding superiority in knowledge (30). Nicolaides et al. showed that both fidelities improved teamwork; high fidelity was perceived as more realistic and favored some leadership components, although the overall objective score was not superior (39).

3.6. Is high fidelity superior?

Five studies directly compared high fidelity with a lower-fidelity or less technologically complex alternative. DeStephano et al. found no superiority of the high-fidelity obstetric simulator over a low-technology model (40). Massoth et al. also observed no advantage and detected greater overconfidence in the high-fidelity group (41). Nicolaides et al. reported objective findings that were equivalent to or advantageous with low fidelity, although the high-fidelity experience proved to be more realistic (39). Valente et al. (42) found that, while both modalities improved learning, retention at three months and confidence were superior with high fidelity. Up to the seventh or eighth week, Pal et al. found no distinctions in the OSCE (43). In summary, technological superiority was not a consistent quality. Benefits emerged when fidelity made it possible to represent fundamental functional relationships, such as physiological changes, clinical progress, the effects of decisions, or team coordination; and when the comparator did not provide equivalent opportunities for active practice. The difference between fidelity levels decreased when both groups were provided with participation, feedback, and an aligned task.

3.7. Withholding and transfer

Ten studies included follow-up beyond the immediate post-test. Retention benefits were found at one month in geriatrics and echocardiography (36, 44), at three months in pediatric respiratory emergencies (42), at four or six months with both methods in the study by Couto et al. (29), and up to two years in the analysis on acute asthma, with a lower decline in knowledge in high-fidelity assessments (45). Perlini et al. documented maintenance of auscultation skills in a subsample assessed years later (33). However, durability was not uniform. The immediate advantage in critical care disappeared at two weeks (28), and the difference in CPR quality diminished at six months (35). Transfer to real patients was exceptional: Butter et al. included bedside assessment and showed improved performance (31), but most studies assessed participants within the same setting or with instruments very close to the intervention. No study robustly demonstrated effects on clinical patient outcomes.

3.8. Confidence, anxiety, stress and calibration

Confidence frequently increased, but did not always correspond to competence. Vattanavanit et al. observed simultaneous increases in knowledge and confidence (46), while Massoth et al. identified overconfidence without objective superiority (41). Yu et al. reported decreased anxiety and increased confidence after two sessions (47). In contrast, Barbadoro et al. found elevated anxiety and cortisol levels in simulation, with a higher perceived burden in high-fidelity simulations (48), and Martín-Sánchez et al. showed that polytrauma and anaphylaxis scenarios produced particularly high levels of anxiety (49). These findings suggest a nonlinear relationship between arousal and learning. Moderate stress can increase engagement and psychological realism; excessive stress can impair working memory, decision-making, and psychological safety. Therefore, prebriefing, graded difficulty, and debriefing are regulatory mechanisms and not merely additional components.

3.9. Risk of bias

Randomized controlled trials revealed concerns primarily due to a lack of data on sequence concealment, the inability to blind participants and instructors, incomplete blinding of evaluators, and the absence of publicly available analytical protocols. In non-randomized studies, moderate risk was prevalent due to confounding, convenience sampling, historical comparators, or unequal exposure. Three studies focusing on subjective outcomes or before-and-after designs showed severe risk. Although the overall trend remained consistent for complex skills that were practiced regularly, confidence in the precise magnitude of the effect decreased due to instrument variety, a lack of validity evidence, and poor follow-up.

4. Discussion

This systematic review included 30 studies with 4,216 medical students. The main finding is that high-fidelity clinical simulation can improve performance, particularly when competencies require integrating dynamic information, prioritizing actions, making decisions, executing procedures, and coordinating teamwork. However, its educational value does not appear to depend solely on the degree of technological realism. The benefits were less consistent for isolated skills or declarative knowledge, which can also be developed through less complex and less costly active learning strategies.

These findings clarify the contribution of this review compared to previous syntheses on simulation in medical students (17, 18). While these reviews have established general benefits of simulation and high-fidelity simulators, this study specifically examines the conditions under which greater fidelity provides differential educational value. To this end, it integrates comparisons between fidelity levels and other active learning strategies, distinguishes between objective competency outcomes and self-reported variables, and jointly analyzes retention, transfer, emotional effects, and risk of bias. Furthermore, a review recently published in the Spanish Journal of Medical Education has shown benefits of anatomical models and simulated scenarios for the development of clinical-surgical competencies (19). This review expands on this perspective by specifically analyzing whether technological fidelity, in itself, is a determinant of better outcomes and what pedagogical components influence its effectiveness. This contribution shifts the analysis from the question of whether high fidelity "works" to a more relevant educational question: for which competency, compared to which alternative, and under what instructional conditions is its use justifiable?

The interpretation of the results is consistent with a multidimensional conception of fidelity, according to which the simulation design should reproduce the functional elements relevant to the educational objective, without indiscriminately increasing physical similarity (6, 7). A high-fidelity environment may be particularly justified when competence depends on uncertainty, interaction between participants, the temporal evolution of a scenario, or the physiological response to clinical decisions. Conversely, for certain skills, a lower-fidelity simulator may facilitate a greater number of repetitions, reduce extrinsic cognitive load, and offer a more favorable relationship between resources and learning. Therefore, the selection of the resource should adhere to the principle of minimum sufficient fidelity, understood as the level necessary to activate the cognitive, technical, and behavioral processes required for competence.

The observed effectiveness appears to be largely associated with the quality of the instructional design. Programs that incorporated explicit objectives, performance standards, repeated practice, immediate feedback, and opportunities for correction showed favorable results, as was the case in neurological emergency training (34) and in auscultation mastery training (31). These findings are consistent with previous evidence on deliberate practice and instructional design in simulation (8, 9, 11). Consequently, a highly immersive experience can be memorable, but without repetition, feedback, and structured debriefing, it does not necessarily translate into lasting learning.

Assessment constitutes another important source of heterogeneity. Satisfaction, perceived realism, and self-confidence provide relevant information about the educational experience, but do not in themselves constitute evidence of competence. Published studies have shown a favorable perception among medical students regarding high-fidelity simulation experiences (5) and have developed specific instruments to assess the quality of and satisfaction with clinical simulation (16). These contributions are relevant for evaluating the acceptability of the strategy; however, the results of this review reinforce the need to differentiate perceptual measures from objective performance. This distinction is especially important given the overconfidence identified by Massoth et al. (41), where a greater sense of preparedness was not accompanied by objective

superiority. Consequently, assessment should prioritize checklists, rubrics, OSCEs, and reasoning tests with evidence of validity, complemented—but not replaced—by measures of satisfaction and confidence. Furthermore, the reliability of the instruments, the blinding of the evaluators, effect sizes, and their confidence intervals should be reported.

Retention and transfer remain two of the major gaps in evidence. Some studies showed that learning was maintained for months or even years, while in others the initial advantage diminished rapidly. Furthermore, most studies assessed participants under conditions similar to those used during training, which may overestimate transfer. Truly competency-based training requires demonstrating that students can apply what they have learned to new situations that are progressively closer to real-world clinical practice. Therefore, future research should incorporate longitudinal follow-ups, previously untrained scenarios, assessment in supervised clinical settings, and, when methodologically feasible, indicators related to the quality and safety of care.

The emotional effects also require careful interpretation. High fidelity can increase engagement and psychological realism, but it can also elevate anxiety, mental workload, and the physiological stress response (48, 49). The relationship between emotional arousal and learning is probably not linear: certain levels of stress can promote engagement and processing of the clinical situation, while excessive stress can compromise working memory, decision-making, and psychological safety. In this context, prebriefing, familiarization with the environment, setting clear expectations, the fictional contract, progressive difficulty, and psychologically safe debriefing should be considered core components of instructional design. Stress, therefore, should not be used as an indicator of authenticity on its own.

From an institutional perspective, the results also do not justify acquiring highly complex technology as a separate curriculum objective. Investment in simulation requires teacher training, maintenance, technical support, dedicated time for faculty, available spaces, assessment systems, and curriculum governance that ensures its longitudinal use. The opportunity cost is particularly relevant: a highly complex simulator used only occasionally may generate fewer learning opportunities than a simpler technology systematically integrated into the training program. Therefore, institutional decisions should simultaneously consider pedagogical relevance, frequency of use, intended competencies, sustainability, and available resources. Table 2 summarizes these elements within a curriculum decision-making framework derived from the findings of this review.

Table 2. Evidence-derived curriculum decision framework.

Curriculum question	Recommended criteria	Quality indicator	Risk avoided
What must the student demonstrate?	Define an observable competency and its expected level.	Explicit objective, standard, and rubric.	Confusing exposure with competence.
What kind of loyalty is needed?	Select the minimum fidelity that reproduces the functional keys of the task.	Justification based on competence.	Technocentrism and unnecessary cost.
How will learning take place?	Include prebriefing, active practice, repetition, feedback, and debriefing.	Reproducible description of each phase.	Isolated experience without consolidation.
How will it be evaluated?	Prioritize objective performance and calibrate it with self-perception.	Valid instruments, trained evaluators, and accurate results.	Overconfidence and weak conclusions.
Is it transferable?	Evaluate retention and performance in new or clinical situations.	Longitudinal follow-up and transfer.	Overestimating the immediate post-test.
Is it sustainable?	Include staff, maintenance, teaching time and opportunity cost.	Longitudinal use and resource analysis.	Underutilized equipment.

Limitations

This review has several limitations. The protocol was not prospectively registered, and the heterogeneity of the interventions, comparators, instruments, and assessment times prevented a meta-analysis without compromising the educational interpretability of the results. Furthermore, the definitions of high fidelity varied across studies, some included small samples or historical comparators, and some outcomes were based on self-reported measures. Language restrictions may also have excluded potentially relevant evidence. Although the search extended to August 2026 and included recent studies, the rapid evolution of simulation technologies will necessitate subsequent updates. Finally, the quality and completeness of the primary studies' reporting limited the analysis of certain pedagogical components and the accuracy of the risk of bias assessments.

Among the strengths of this review are its exclusive focus on medical students, the differentiation between objective performance and perceived outcomes, the direct comparison between different levels of fidelity and active methodologies, and the combined inclusion of retention, transfer, emotional effects, and risk of bias. Unlike an interpretation focused solely on the overall effectiveness of simulation, this review allows for the identification of educational circumstances in which high fidelity provides added value and those in which a less complex alternative may suffice. This perspective transcends the "works/doesn't work" binary and positions fidelity as an instructional design variable subordinate to competence, context, and the purpose of the assessment.

5. Conclusions

- High-fidelity clinical simulation helps hone complex skills that require information integration, decision-making, procedural execution, and group coordination. It should not be assumed, however, that it is inherently superior to low-fidelity simulation or other active learning methodologies, as its educational value is based primarily on instructional design and the relationship between the resources used and the competency task. Timely feedback, structured debriefing, objective performance evaluation, and deliberate practice appear to be the main factors contributing to positive outcomes.

- Self-confidence, therefore, must be studied in conjunction with observable performance, as excessive immersion can generate stress or overconfidence in addition to educational benefits. Despite its potential, there is still little evidence on its sustained effectiveness, the transfer of learning to clinical practice, its impact on patient outcomes, and its cost-effectiveness. Therefore, training programs should implement the principle of minimum sufficient fidelity and develop longitudinal assessment procedures that allow for quantifying the true progress of competencies.

Supplementary material: Final matrix of included studies.

Authorship contributions: Conceptualization: LBVC and YBG; methodology: LBVC, J,MM, and YBG; data search and curation: LBVC and YBG; analysis and interpretation: all authors; drafting: LBVC, MAM, and YBG; critical review and editing: all authors; supervision: YBG. All authors approved the final version and assume responsibility for its content.

Funding: The research did not receive specific external funding.

Conflicts of interest: The authors declare no conflicts of interest related to this work.

Data availability: The search strategies, extraction matrix, and bias risk assessment are provided as supplementary material. No individual participant data were generated.

Declaration on the use of artificial intelligence: ChatGPT (OpenAI; version available August 2026) was used as an auxiliary tool for the structural organization, linguistic review, and formal standardization of the manuscript. The selection of sources, bibliographic verification, methodological assessment, scientific interpretation, and final approval are the responsibility of the authors, who critically reviewed the content and assume full responsibility for its accuracy and integrity.

6. References

1. Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: theory to practice. *MedTeach.* **2010,** 32(8), **638.** https://doi.org/10.3109/0142159X.2010.501190 .
2. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA.* **2002,** 287(2), **226.** https://doi.org/10.1001/jama.287.2.226 .
3. Gaba DM. The future vision of simulation in health care. *Qual Saf Health Care.* **2004,** 13(Suppl 1), i2-i10. https://doi.org/10.1136/qshc.2004.009878 .
4. Issenberg SB, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach.* **2005,** 27(1), 10. https://doi.org/10.1080/01421590500046924 .
5. Ramos G, Ardila Botero D. Perception of clinical simulation as a didactic tool in teaching postpartum hemorrhage in the Degree in Medicine. *Rev Esp Edu Med.* **2022,** 3(1), 31. https://doi.org/10.6018/edumed.501861 .
6. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. *Acad Med.* **2014,** 89(3), **387.** https://doi.org/10.1097/ACM.0000000000000130
7. Norman G, Dore K, Grierson L. The minimal relationship between simulation fidelity and transfer of learning. *Med Educ.* **2012,** 46(7), 636. https://doi.org/10.1111/j.1365-2923.2012.04243.x .
8. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. *Med Educ.* **2010,** 44(1), 50-63. https://doi.org/10.1111/j.1365-2923.2009.03547.x .
9. McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? *Acad Med.* **2011,** 86(6), 706. https://doi.org/10.1097/ACM.0b013e318217e119 .
10. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA.* **2011,** 306(9), 978-88. https://doi.org/10.1001/jama.2011.1234 .

11. Cook DA, Hamstra SJ, Brydges R, et al. Comparative effectiveness of instructional design features in simulation-based education. *MedTeach.* **2013,** 35(1), e867. https://doi.org/10.3109/0142159X.2012.714886 .
12. Motola I, Devine LA, Chung HS, Sullivan JE, Issenberg SB. Simulation in healthcare education: a best evidence practical guide. AMEE Guide No. 82. *Med Teach.* **2013,** 35(10), e1511. https://doi.org/10.3109/0142159X.2013.818632 .
13. Cheng A, Eppich W, Grant V, Sherbino J, Zendejas B, Cook DA. Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. *Med Educ.* **2014,** 48(7), 657. https://doi.org/10.1111/medu.12432 .
14. Rudolph JW, Simon R, Raemer DB, Eppich WJ. Debriefing as formative assessment: closing performance gaps in medical education. *Acad Emerg Med.* **2008,** 15(11), 1010. https://doi.org/10.1111/j.1553-2712.2008.00248.x .
15. 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 .
16. Padilla MJ, González J, Sarmiento F, Tripoloni D, Cohen Arazi L. Clinical simulation: validation of a quality and satisfaction survey in a group of medical students. *Rev Esp Edu Med.* **2024,** 5(1). https://doi.org/10.6018/edumed.591511.
17. McInerney N, Nally D, Khan MF, Heneghan H, Cahill RA. Performance effects of simulation training for medical students: a systematic review. *GMS J Med Educ.* **2022,** 39(5). https://doi.org/10.3205/zma001572 .
18. Ramah U, Keenoo BS. High-fidelity simulators in undergraduate medical education: a systematic review. *Cureus.* **2025,** 17(10), e95019. https://doi.org/10.7759/cureus.95019 .
19. Vade Martínez C, Painemal Rivera B, Serey Torres D, et al. Anatomical models and simulated scenarios for learning surgical skills in undergraduate and postgraduate medical students. A systematic review. *Rev Esp Edu Med.* **2025,** 6(1). https://doi.org/10.6018/edumed.638611 .
20. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* **2021,** 372. https://doi.org/10.1136/bmj.n71 .
21. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ.* **2020,** 368. https://doi.org/10.1136/bmj.l6890 .
22. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomized trials. Version 2. London: Cochrane; **2019** . https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2 .
23. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* **2016,** 355. https://doi.org/10.1136/bmj.i4919 .
24. McCoy CE, Menchine M, Anderson C, Kollen R, Langdorf MI, Lotfipour S. Prospective randomized crossover study of simulation vs. Didactics for teaching medical students the assessment and management of critically ill patients. *J Emerg Med.* **2011,** 40(4). https://doi.org/10.1016/j.jemermed.2010.02.026 .
25. Cortegiani A, Russotto V, Montalto F, et al. Effect of high-fidelity simulation on medical students knowledge about advanced life support: a randomized study. *PLoS One.* **2015,** 10(5). https://doi.org/10.1371/journal.pone.0125685 .
26. DeWaay DJ, McEvoy MD, Kern DH, Alexander LA, Nietert PJ. Simulation curriculum can improve medical student assessment and management of acute coronary syndrome during a clinical practice exam. *Am J Med Sci.* **2014,** 347(6). https://doi.org/10.1097/MAJ.0b013e3182a562d7 .
27. Mutter MK, Martindale JR, Shah N, Gusic ME, Wolf SJ. Case-based teaching: does the addition of high-fidelity simulation make a difference in medical students clinical reasoning skills? *Med Sci Educ.* **2020,** 30(1), 307. https://doi.org/10.1007/s40670-019-00904-0 .

28. Solymos O, O'Kelly P, Walshe CM. Pilot study comparing simulation-based and didactic lecture-based critical care teaching for final-year medical students. *BMC Anesthesiol.* **2015,** 15. https://doi.org/10.1186/s12871-015-0109-6 .
29. Couto TB, Farhat SCL, Geis GL, Olsen O, Schvartsman C. High-fidelity simulation versus case-based discussion for teaching medical students in Brazil about pediatric emergencies. *Clinics (Sao Paulo).* **2015,** 70(6), 393. https://doi.org/10.6061/clinics/2015(06)02 .
30. Moliterno NV, Paravidino VB, Robaina JR, et al. High-fidelity simulation versus case-based discussion for training undergraduate medical students in pediatric emergencies: a quasi-experimental study. *J Pediatr (Rio J).* **2024,** 100(4), 422. https://doi.org/10.1016/j.jped.2024.03.007 .
31. Butter J, McGaghie WC, Cohen ER, Kaye M, Wayne DB. Simulation-based mastery learning improves cardiac auscultation skills in medical students. *J Gen Intern Med.* **2010,** 25(8), 780. https://doi.org/10.1007/s11606-010-1309-x .
32. Kern DH, Mainous AG, Carey M, Beddingfield A. Simulation-based teaching to improve cardiovascular exam skills performance among third-year medical students. *Teach Learn Med.* **2011,** 23(1), 15. https://doi.org/10.1080/10401334.2011.536753 .
33. Perlini S, Salinaro F, Santalucia P, Musca F. Simulation-guided cardiac auscultation improves medical students clinical skills: the Pavia pilot experience. *Intern Emerg Med.* **2014,** 9(2), 165. https://doi.org/10.1007/s11739-012-0811-z .
34. Sánchez-Ledesma MJ, Juanes JA, Gómez-Sánchez E, et al. Acquisition of competencies by medical students in neurological emergency simulation environments using high fidelity patient simulators. *J Med Syst.* **2016,** 40(6), 139. https://doi.org/10.1007/s10916-016-0496-3 .
35. Berger C, Brinkrolf P, Ertmer C, et al. Combination of problem-based learning with high-fidelity simulation in CPR training improves short and long-term CPR skills: a randomized single blinded trial. *BMC Med Educ.* **2019,** 19, 180. https://doi.org/10.1186/s12909-019-1626-7 .
36. Zhao Y, Zhang X, Li X, et al. Simulation-based training following a theoretical lecture enhances the performance of medical students in obtaining 20 standard cross-sectional transesophageal echocardiographic views: a prospective, randomized, controlled trial. *BMC Med Educ.* **2021,** 21, 336. https://doi.org/10.1186/s12909-021-02753-1 .
37. Puerma-Jiménez J, Pérez-Villares JM, Domínguez-Guzmán P, et al. High-fidelity simulation in airway management training: results from the AIR-MASTER-SIM randomized study. *Front Med (Lausanne).* **2026,** 13, 1829702. https://doi.org/10.3389/fmed.2026.1829702.
38. Gonçalves BAR, Melo MCB, Oliveira ACL, et al. Teamwork in pediatric resuscitation: training medical students on high-fidelity simulation. *Adv Med Educ Pract.* **2022,** 13, 697. https://doi.org/10.2147/AMEP.S365976 .
39. Nicolaides M, Theodorou E, Emin EI, et al. Team performance training for medical students: low vs high fidelity simulation. *Ann Med Surg (London).* **2020,** 55, 308. https://doi.org/10.1016/j.amsu.2020.05.042 .
40. DeStephano CC, Chou B, Patel S, Slattery R, Hueppchen N. A randomized controlled trial of birth simulation for medical students. *Am J Obstet Gynecol.* **2015,** 213(1), 91.e1. https://doi.org/10.1016/j.ajog.2015.03.024 .
41. Massoth C, Röder H, Ohlenburg H, et al. High-fidelity is not superior to low-fidelity simulation but leads to overconfidence in medical students. *BMC Med Educ.* **2019,** 19, 29. https://doi.org/10.1186/s12909-019-1464-7 .
42. Valente BCHG, Melo MCB, Liu PMF, et al. High and low-fidelity simulation for respiratory diseases pediatric training: a prospective and randomized study. *J Pediatr (Rio J).* **2023,** 99(5), 521. https://doi.org/10.1016/j.jped.2023.04.007 .
43. Pal B, Chong SV, Thein AW, Tay AGM, Soe HHK, Pal S. A randomized controlled trial study on the effectiveness of high-fidelity simulation in enhancing skills among undergraduate medical students. *Med J Malaysia.* **2024,** 79(4), 421. https://pubmed.ncbi.nlm.nih.gov/39086339/ .

44. Fisher JM, Walker RW. A new age approach to an age old problem: using simulation to teach geriatric medicine to medical students. *Age Aging.* **2014,** 43(3), 424. https://doi.org/10.1093/ageing/aft200 .
45. Grissa MH, Dhaoui R, Bel Haj Ali K, et al. Comparison of simulation and video-based training for acute asthma. *BMC Med Educ.* **2023,** 23. https://doi.org/10.1186/s12909-023-04836-7 .
46. Vattanavanit V, Kawla-Ied J, Bhurayanontachai R. High-fidelity medical simulation training improves medical students knowledge and confidence levels in septic shock resuscitation. *Open Access Emerg Med.* **2017,** 9, 1. https://doi.org/10.2147/OAEM.S122525 .
47. Yu JH, Chang HJ, Kim SS, et al. Effects of high-fidelity simulation education on medical students anxiety and confidence. *PLoS One.* **2021,** 16(5). https://doi.org/10.1371/journal.pone.0251078 .
48. Barbadoro P, Brunzini A, Dolcini J, et al. Stress responses in high-fidelity simulation and standard simulation training among medical students. *BMC Med Educ.* **2023,** 23, 116. https://doi.org/10.1186/s12909-023-04101-x .
49. Martín-Sánchez R, Castro Villamor MA, Rabanales-Soto J, et al. Threshold-anxiety in medical students performing a prehospital high-fidelity clinical simulation: randomized clinical trial. *Am J Emerg Med.* **2025,** 94, 103. https://doi.org/10.1016/j.ajem.2025.04.039 .
50. Swamy M, Bloomfield TC, Thomas RH, Singh H, Searle RF. Role of SimMan in teaching clinical skills to preclinical medical students. *BMC Med Educ.* **2013,** 13, 20. https://doi.org/10.1186/1472-6920-13-20 .
51. Swamy M, Sawdon M, Chaytor A, Cox D, Barbaro-Brown J, McLachlan J. A study to investigate the effectiveness of SimMan as an adjunct in teaching preclinical skills to medical students. *BMC Med Educ.* **2014,** 14, 231. https://doi.org/10.1186/1472-6920-14-231 .
52. Alluri RK, Tsing P, Lee E, Napolitano J. A randomized controlled trial of high-fidelity simulation versus lecture-based education in preclinical medical students. *MedTeach.* **2016,** 38(4), 404-9. https://doi.org/10.3109/0142159X.2015.1031734 .
53. Bernardi S, Giudici F, Leone MF, et al. A prospective study on the efficacy of patient simulation in heart and lung auscultation. *BMC Med Educ.* **2019,** 19, 275. https://doi.org/10.1186/s12909-019-1708-6 .



© 2026 University of Murcia. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 Spain License (CC BY-NC-ND) (http://creativecommons.org/licenses/by/4.0/)