Annals of Medicine

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/iann20

Metaverse-based simulation: a scoping review of charting medical education over the last two decades in the lens of the 'marvelous medical education machine'

Vitaliy Popov, Natalie Mateju, Caris Jeske & Kadriye O. Lewis

To cite this article: Vitaliy Popov, Natalie Mateju, Caris Jeske & Kadriye O. Lewis (2024) Metaverse-based simulation: a scoping review of charting medical education over the last two decades in the lens of the 'marvelous medical education machine', Annals of Medicine, 56:1, 2424450, DOI: 10.1080/07853890.2024.2424450

To link to this article: https://doi.org/10.1080/07853890.2024.2424450

Published online: 13 Nov 2024.

ANNALS OF MEDICINE
2024, VOL. 56, NO. 1, 2424450
https://doi.org/10.1080/07853890.2024.2424450

REVIEW ARTICLE

OPEN ACCESS Check for updates

Metaverse-based simulation: a scoping review of charting medical education over the last two decades in the lens of the 'marvelous medical education machine'

Vitaliy Popov[a], Natalie Mateju[a], Caris Jeske[a] and Kadriye O. Lewis[b]

[a]Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, MI, USA; [b]Children's Mercy Kansas City, Department of Pediatrics, UMKC School of Medicine, Kansas City, MO, USA

ABSTRACT

Background: Over the past two decades, the use of Metaverse-enhanced simulations in medical education has witnessed significant advancement. These simulations offer immersive environments and technologies, such as augmented reality, virtual reality, and artificial intelligence that have the potential to revolutionize medical training by providing realistic, hands-on experiences in diagnosing and treating patients, practicing surgical procedures, and enhancing clinical decision-making skills. This scoping review aimed to examine the evolution of simulation technology and the emergence of metaverse applications in medical professionals' training, guided by Friedman's three dimensions in medical education: physical space, time, and content, along with an additional dimension of assessment.

Methods: In this scoping review, we examined the related literature in six major databases including PubMed, EMBASE, CINAHL, Scopus, Web of Science, and ERIC. A total of 173 publications were selected for the final review and analysis. We thematically analyzed these studies by combining Friedman's three-dimensional framework with assessment.

Results: Our scoping review showed that Metaverse technologies, such as virtual reality simulation and online learning modules have enabled medical education to extend beyond physical classrooms and clinical sites by facilitating remote training. In terms of the Time dimension, simulation technologies have made partial but meaningful progress in supplementing traditional time-dependent curricula, helping to shorten learning curves, and improve knowledge retention. As for the Content dimension, high-quality simulation and metaverse content require alignment with learning objectives, interactivity, and deliberate practice that should be developmentally integrated from basic to advanced skills. With respect to the Assessment dimension, learning analytics and automated metrics from metaverse-enabled simulation systems have enhanced competency evaluation and formative feedback mechanisms. However, their integration into high-stakes testing is limited, and qualitative feedback and human observation remain crucial.

Conclusion: Our study provides an updated perspective on the achievements and limitations of using simulation to transform medical education, offering insights that can inform development priorities and research directions for human-centered, ethical metaverse applications that enhance healthcare professional training.

KEY MESSAGES

- The evolution of simulation technology and the emerging metaverse applications have significantly extended medical education beyond physical boundaries and time constraints, enabling learners to access a wider range of learning experiences thereby preparing them for the rapidly changing healthcare environment.
- Learning analytics and automated metrics from metaverse-enabled simulation systems have improved competency evaluation and formative feedback mechanisms. However, integration into high-stakes testing is limited, and qualitative feedback and human observation are still crucial.
- The use of technology in medical education has advanced significantly, but problems still exist with access, content quality, and integration into high-stakes assessments. These issues call for more innovation and research to find the best ways to incorporate learning analytics, metaverse applications, and fair, human-centered training methods.

ARTICLE HISTORY
Received 12 October 2023
Revised 12 August 2024
Accepted 11 October 2024

KEYWORDS
Medical education; clinical simulation; metaverse; emerging technologies; eXtended reality; virtual reality; augmented reality; mixed reality

CONTACT Vitaliy Popov vipopov@umich.edu Department of Learning Health Sciences, University of Michigan Medical School, University of Michigan, School of Information, Victor Vaughan, 217, 1111 Catherine St, Ann Arbor, MI 48109, USA
Supplemental data for this article is available online at https://doi.org/10.1080/07853890.2024.2424450

Introduction

In medical education, artificial intelligence and eXtended Reality (XR) applications represent a cutting-edge intersection between technology and healthcare, offering transformative opportunities to enhance the training of medical professionals [1,2]. Through the emergence of multimedia technologies, World Wide Web accessibility, and widespread networked computing, this transformative shift has drastically changed the teaching and learning in the entire medical school educational system. Two decades ago, Friedman [3] proposed a visionary concept of the 'marvelous medical education machine' while acknowledging the potential of technology and criticizing the shortcomings of medical education. He envisioned a system where learners could access high-quality medical education content through computer-based learning, simulations, and other technological means at any time, from anywhere. This approach would allow learners to train at their own pace, unconstrained by the limitations of traditional classroom settings, physical location, or rigid curricula. Thus, Friedman's perspective on medical education in the United States has become 'stuck' in three dimensions: physical space, time, and content, which was rooted in his belief that the traditional model of medical education had not adapted effectively to the changing needs and opportunities of the modern world. Each dimension from Friedman's perspective is defined as below:

1. *Physical Space*: Friedman argued that traditional medical education, based on physical institutions like medical schools and classrooms, restricted access, especially for those in rural areas. He believed technology, like computer-based learning and telemedicine, could make medical education more accessible by eliminating the need for physical presence.
2. *Time*: The 'time' dimension in medical education pertains to rigid schedules and timelines in traditional programs, which may not suit diverse student needs. Friedman proposed flexible learning models, allowing students to learn at their own pace, aided by technology for adaptability in medical education.
3. *Content*: This dimension relates to the curriculum and content covered in traditional medical education. Friedman argued that the curriculum could be outdated, slow to adapt to new medical discoveries, and overly focused on theoretical knowledge rather than practical skills and patient-centered care. Friedman proposed a more dynamic and adaptable approach to delivering content in medical education. He believed that technology could help continuously update educational materials, ensuring students access the latest and most relevant medical information.

The physical space of medical schools has not kept up with the increasing demand for medical professionals. Many medical schools have faced challenges in expanding their physical facilities to accommodate the growing demand for medical education [4]. The time required to complete a medical degree has remained remarkably stable over the past several decades, and the number of required courses has increased [5,6].

With the advent of the metaverse technology, the marvelous medical education machine has now become a reality. Metaverse is a combination of 'meta' and 'universe' which describes a parallel or virtual environment linked to the physical world [7]. More specifically, the Metaverse is an umbrella term for a network of interconnected virtual spaces and immersive technologies, where users typically wear a head-mounted device to explore and interact with a blended physical and digital world. Application of Metaverse or Meta platforms in medical education offers new promising technological and pedagogical affordances for the development of healthcare professionals' competencies ranging from patient care, medical knowledge, procedural training to interprofessional teamwork, clinical reasoning, and critical thinking.

The machine envisioned by Friedman in 2000 would provide endless opportunities for repetitive practice of clinical skills, exposure to a wide range of patients and conditions, and experimentation through 'what-if' scenarios. Key capabilities of this envisioned machine became fully or partially realized through the range of metaverse applications in recent years, including clinical representation using virtual patients (interactive computer-based simulations of clinical encounters for diagnosis, treatment planning, and communications training [8]), procedural skills trainers with haptic feedback for practicing technical skills like laparoscopy or robot-assisted surgery [9,10], sensory rendering of symptoms through virtual reality [11], and multimodal learning analytics and automated scoring models to provide near real-time feedback on team and/or learner performance [12–14].

Analyzing the evolution within each dimension of Friedman's framework offers a structured approach for evaluating the advancement toward the envisioned metaverse in medical education while pinpointing existing deficiencies. A systematic understanding of how metaverse applications or marvelous medical education machines can leverage their unique affordances to best support procedural training, sharpen clinical decision-making skills, and fortify teamwork mechanisms, among various other aspects, which is currently lacking in medical education. Thus, we examined the evolution of simulation technology and the emergence of metaverse applications in medical professionals' training through the lens of Friedman's three constraints in medical education: physical space, time, and content, along with an additional dimension of assessment by utilizing a scoping review methodology.

Including Assessment as a key dimension in the framework enabled us to systematically review the diverse methods employed to evaluate learning experiences and outcomes in metaverse applications for medical education. Our approach taken to examine evaluation methods would fill a significant gap in the existing literature. The following four research questions guided this scoping review:

1. In what ways have metaverse-based simulation technologies changed the physical space aspect of medical education, impacting learning locations and resource accessibility?
2. How have metaverse-based simulation technologies changed, if at all, the way that medical education is delivered in terms of time, including when lessons are taught, how flexible the schedule is, and how feedback is provided in real-time?
3. How has the evolution of simulation technologies and metaverse applications affected the quality and availability of educational content in medical training?
4. What changes have occurred in the assessment methods in medical education due to metaverse-based simulation technologies (e.g. high-stakes assessments, fairness in assessment, and competency evaluation)?

Materials and methods

This scoping review used a systematic approach to identify, select, and synthesize relevant studies using Friedman's idea of the 'marvelous medical education machine' in the following four domains.

- *Time* refers to when educational interactions and events occur, often requiring learners and faculty to participate simultaneously. Friedman [3] argues that medical education is 'stuck' in time because events are bound to set schedules. Additionally, the concept of 'time' in the context of medical education and simulation technology is linked to the idea of the learning curve or proficiency gain over time. The use of simulation technology may have a significant impact on the learning curve by potentially accelerating the time it takes to acquire new skills.
- *Space* refers to the learning environment where educational experiences, interactions, and events occur physically or virtually, often in classrooms and healthcare delivery settings tied to an academic medical center. Despite the fact that delivery mechanisms have been less restricted to specific locations in recent years, medical education has been 'stuck' in space for decades [3].
- *Content* refers to the biomedical topics that are being taught. This covers a broad spectrum of areas, such as anatomy, physiology, diagnostics, surgical procedures, clinical reasoning, and patient care among others. In response to advancements in medical knowledge and changes in healthcare, some medical schools have expanded their curricula to include more courses or new areas of study [3].
- The '*Assessment*' dimension was included because, in the past two decades, assessment has become increasingly important and widespread in medical education in general and in clinical simulation, in particular, to evaluate trainees' clinical competency and readiness for practice [15,16]. As technology advances, multimodal data sources, such as video, audio, physiological data, and user interactions, can be collected and analyzed to examine meaningful associations, observe trends, and provide precision-guided feedback to each team or learner individually for deliberate practice and skills remediation [12,17,18].

We detailed the key steps involved in our review process, including a protocol, search strategy, eligibility criteria, critical appraisal for inclusion, data charting, and analysis used to achieve our study purpose as below.

Protocol

A protocol was developed based on the recommendations of the Prisma extension for scoping reviews reporting guidelines (PRISMA ScR) [19,20].

Search strategy

An experienced health sciences librarian generated search terms based on the Friedman articles published in Academic Medicine (1999) and Medical Teacher (2000). Reference tracking was also performed on both versions. The original search strategies were developed in PubMed and translated as appropriate to the other databases: (((Computer-Assisted Instruction[mh] OR computer-assisted instruction[tiab] OR digital simulation[tiab] OR educational innovation[tiab] OR Educational technology[mh] OR educational technology[tiab] OR new media technolog*[tiab] OR performance gains[tiab] OR Problem-based Learning[mh] OR problem-based learning[tiab] OR serious games[tiab] OR simulation-based education[tiab] OR virtual reality[mh] OR virtual reality[tiab] OR video games[mh] OR video games[tiab]) AND (Education, Medical/trends OR diffusion of innovation[mh]) AND (Educational Measurement[mh] OR educational measurement*[tiab] OR Clinical Competence[mh] OR clinical competence[-tiab] OR Clinical Decision-Making[mh] OR clinical decision making[tiab] OR learning retention[tiab])).

The detailed search strings and Boolean operators for each database are included in the Search Report (Supplementary Appendix A).

A systematic search of MEDLINE (PubMed), EMBASE (Embase.com), CINAHL (EBSCOhost), Scopus (Elsevier), Web of Science Core Collection (Clarivate Analytics), and ERIC (Proquest) to identify articles on the topic of innovation in medical education, especially the idea of it being stuck in time, space, and content, was conducted. To reduce language bias, abstracts for articles in languages other than English were evaluated during the screening process. Findings are reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [21], elaboration, and explanation [19].

All searches were completed by 18 December 2020. Citations were imported into EndNote X9.3.3 (Thomson Reuters, New York, NY, USA) for deduplication, then exported into Excel (Microsoft Office 2016) for analysis.

Screening and selection

Eligibility criteria

The initial phase of our literature screening process was centered on three key inclusion and exclusion criteria.

- *Language Inclusion:* Only studies conducted in the English-language studies were considered.
- *Time Frame:* Our scope included articles published within the past two decades, in alignment with the original publication of Friedman's seminal work in 2000.
- *Quality Assurance:* To ensure the utmost rigor, authenticity, and quality, we focused exclusively on peer-reviewed articles published in reputable scholarly journals.

Our subsequent screening phase entailed the selecting studies based on the content and focus areas.

Content and Focus Criteria:

- Studies focus on medical education.
- Studies involve the use of education technology, clinical simulation, or any form of extended reality.
- Studies explore e-learning applications.
- Studies provide insights related to at least one of Friedman's dimensions: Time, Space, Content, or Assessment.

Then, our focus was directed exclusively toward studies within the field of health professions education, differentiating from investigations centered on medical technology, for instance. Our final step involved the exclusion of publications based on the following criteria:

Exclusion Criteria:

- Studies focus on medical technology that is not related to education.
- Studies are conducted outside of the health professions education setting.
- Studies describe problem-based learning without any technology integration.
- Studies only discuss medical school curriculum changes without the inclusion of technology
- Studies involve learners who are not medical students, nurses or nursing students, residents, or other healthcare professionals.

Critical appraisal included publications

To mitigate selection bias, we employed several strategies. First, two reviewers (NM, CJ) independently screened titles and abstracts, with a third reviewer (VP) resolved any disagreements. The review team met regularly to discuss and resolve any ambiguities in study selection. Second, we used a pre-defined screening form based on our inclusion/exclusion criteria to ensure consistent application across all potential studies. To appraise the quality, study design, and assess the risk of bias of included studies, we utilized the Mixed Methods Appraisal Tool (MMAT) [22] (Table 1).

Table 1. Quality appraisal of the included studies using Mixed Methods Appraisal Tool (MMAT) [22].

Qualitative studies (15)	Can't tell	No	Yes
1.1. Is the qualitative approach appropriate to answer the research question?	0	0	15
1.2. Are the qualitative data collection methods adequate to address the research question?	0	0	15
1.3. Are the findings adequately derived from the data?	0	0	15
1.4. Is the interpretation of results sufficiently substantiated by data?	0	0	15
1.5. Is there coherence between qualitative data sources, collection, analysis, and interpretation?	0	0	15
Randomized controlled trials (26)			
2.1. Is randomization appropriately performed?	0	0	26
2.2. Are the groups comparable at baseline?	1	0	25
2.3. Are there complete outcome data?	0	0	26
2.4. Are outcome assessors blinded to the intervention provided?	10	0	16
2.5 Did the participants adhere to the assigned intervention?	0	0	26
Non-randomized studies (47)			
3.1. Are the participants representative of the target population?	0	0	47
3.2. Are measurements appropriate regarding both the outcome and intervention (or exposure)?	0	0	47
3.3. Are there complete outcome data?	0	0	47
3.4. Are the confounders accounted for in the design and analysis?	7	3	37
3.5. During the study period, is the intervention administered (or exposure occurred) as intended?	0	0	47
Quantitative descriptive studies (39)			
4.1. Is the sampling strategy relevant to address the research question?	1	0	38
4.2. Is the sample representative of the target population?	0	0	39
4.3. Are the measurements appropriate?	0	0	39
4.4. Is the risk of nonresponse bias low?	17	2	20
4.5. Is the statistical analysis appropriate to answer the research question?	1	0	38
Mixed method studies (15)			
5.1. Is there an adequate rationale for using a mixed methods design to address the research question?	0	0	15
5.2. Are the different components of the study effectively integrated to answer the research question?	0	0	15
5.3. Are the outputs of the integration of qualitative and quantitative components adequately interpreted?	0	0	15
5.4. Are divergences and inconsistencies between quantitative and qualitative results adequately addressed?	1	2	12
5.5. Do the different components of the study adhere to the quality criteria of each tradition of the methods involved?	0	0	15

This tool was chosen for its ability to assess various study designs, including qualitative, quantitative, and mixed methods research.

Out of the final pool of 173 selected publications for final analysis, most papers have been published since 2010 (N=131, 75%) (Supplementary Appendix B). They also have been published in a wide variety of scholarly journals (total number of unique journals is 138) and cover a wide range of medical specialties, including top fields: medical education and informatics (31 journals), surgery (20 journals), ophthalmology & ENT specializations (9 journals), obstetrics and gynecology (7 journals), pediatrics (6 journals), emergency medicine (6 journals), and others.

The studies were conducted in various countries around the world, including. North America (USA, Canada): 56% (99 studies), Europe (Germany, United Kingdom, Netherlands, Greece, Denmark, Sweden, Portugal, Croatia, Ireland, Italy): 32% (57 studies), Asia and Middle East/transcontinental country (Malaysia, Korea, Iran, Turkey): 7% (11 studies), Australia: 4% (5 study), New Zealand: 1% (1 study). Most studies were quantitative (N=112 studies, 65%), relying primarily on pre- and post-testing (knowledge tests, skills assessments, surveys, observational checklists (rating technical skills, behaviors, etc.), performance metrics (time, errors, efficiency, etc.), surveys and tests scores. Only a few studies were qualitative (N=15, 9%), using interviews, focus groups, observations, field notes, and reflective journals; or mixed methods (N=15, 9%), using a combination of performance data and focus groups, interviews, and/or observations. Thirty-one papers (N=31, 18%) did not fall under the MMAT classification criteria. These papers were not empirical studies, but rather case studies, case series, commentaries, or manuscripts describing specific simulation scenarios. Their primary focus was on illustrating the application of emergent technology or clinical simulation and its integration into medical curricula. Although not empirical in nature, these publications generated novel research questions, built theories, and provided valuable insights into one or more dimensions (space, time, content, assessment) in the context of medical education.

Data charting and analysis

Before data extraction, a draft extraction table was developed in Microsoft Excel to align with the scoping review research questions. Three authors extracted the final data (VP, NM, and CJ), before manually assembling and resolving discrepancies. A coding scheme was developed based on Friedman's [3] framework with the four dimensions (Figure 1) to thematically analyze included publications. All 173 papers included in the final review were analyzed and coded based on the coding scheme. This involved a rigorous process of reading and re-reading the selected papers to identify

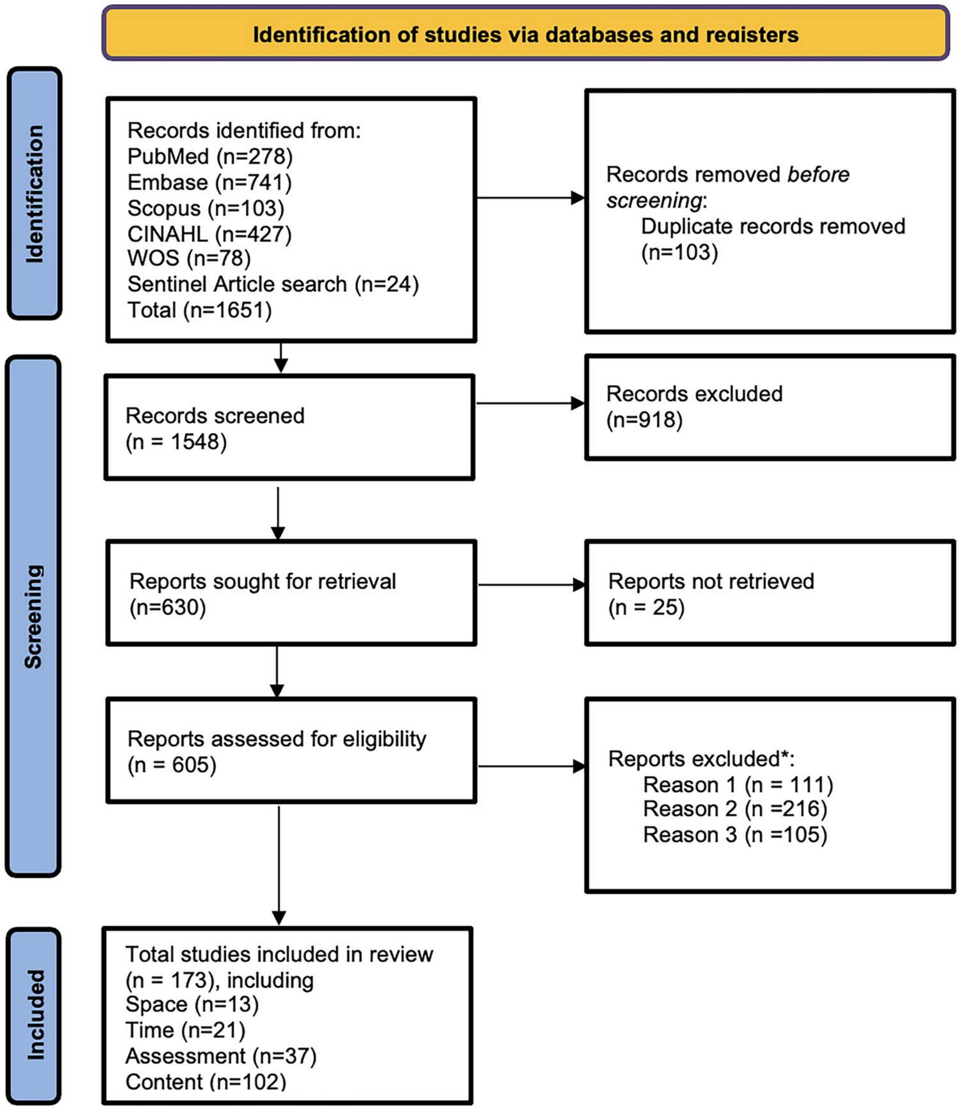

Figure 1. PRISMA Flow diagram of studies on innovation in medical education, especially the idea of it being stuck in time, space, and content. *Reason 1: Studies are outside of the health professions education setting. Reason 2: Studies are not related to the space, time, content, and assessment based on Friedman's operationalization. Reason 3: Participants were not medical students, nurses or nursing students, residents, other healthcare professionals.

significant patterns and trends within each dimension. This iterative process allowed for the extraction and coding of key ideas into themes. As themes emerged, they were systematically categorized to help map and interpret the study's complex data landscape.

Results

The scoping review revealed that advancements in simulation technology have greatly impacted medical education in all four dimensions of physical space, time, content, and assessment. In the original Friedman's model [3], the dimension of space was not intersecting with the other dimensions because it was 'stuck' (see Figure 2, left). The fixed locations (classrooms or hospitals) and predetermined times for learning activities meant that all trainees had to progress through the material at the same rate, regardless of their individual learning curves or prior knowledge. The content was primarily dictated by faculty or available patients in experiential learning setting, rather than being flexible to learner needs, meant that trainees could not slow down on challenging topics or speed up through familiar ones. In the updated 'unstuck' medical education model (see Figure 2, right), space is no longer isolated but converges with content and time, illustrating the shift towards increased independence from these dimensions in medical education. The spatial dimension has greatly expanded due to advances in technology, which has expanded

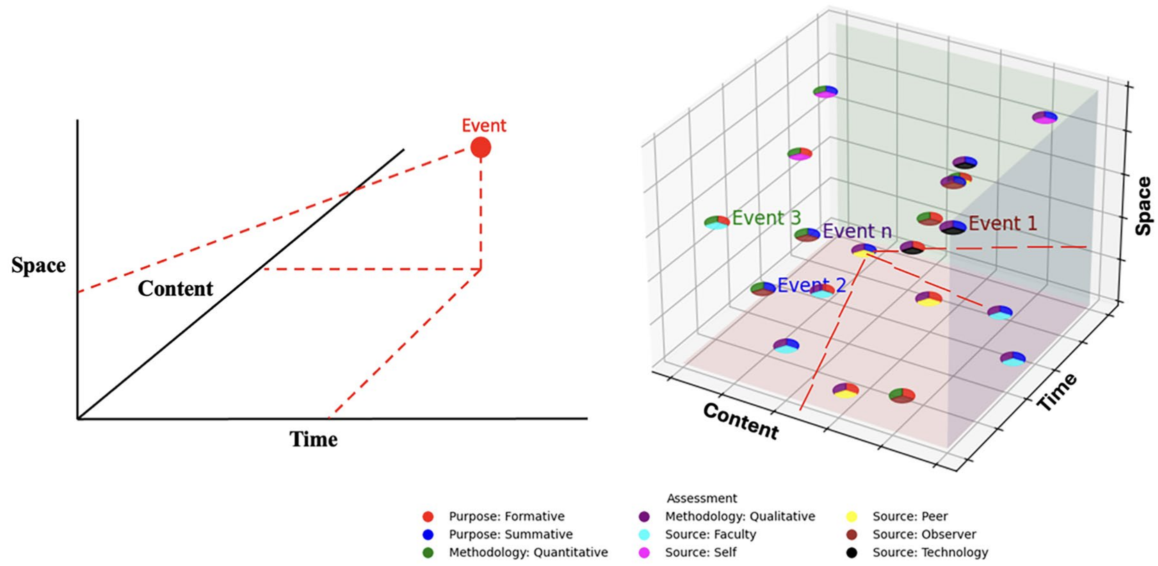

Figure 2. Original depiction by Friedman [3] (left) and two decades later (right) of medical education as a process with teaching and learning events that can occur anywhere, anytime, and cover any relevant topic as needed. The colored pie chart markers (image on the right) for each event represent different combinations of assessment modalities based on their purpose, methodology, and data source.

beyond classrooms and hospitals to include virtual simulations, mobile learning, and remote sites, which allow for flexibility in the location of learning. Temporally, on-demand access and self-paced progression become normalized in medical education nowadays. The proliferation of content dimension as evidenced in this review, illustrates a wide range of topics customizable to individual learner needs. Furthermore, assessment in the updated model (Figure 2, right) has become more multifaceted and multimodal. It can occur at any point in the learning process, using a variety of methods (quantitative, qualitative, or mixed) and sources (faculty, self, peer, observer, or technology), serving both formative and summative purposes. To exemplify this updated model, we can draw on an innovative eXtended Reality International Grand Rounds study, which utilized XR technology to present complex medical cases to trainees [23]. During bedside rounds in limited-space grand rounds, the XR application has enabled this experience for trainees to overcome the constraints of space and has connected tens of trainees to the same content at the same time to allow for a learning event to occur at scale.

The increasing trend in the total number of papers over the years also signifies the ongoing evolution and importance of simulation technology in transforming and unsticking the traditional constraints of medical education. However, there is a clear discrepancy in the concentration of research efforts across the key themes, with the content being most heavily researched, followed by assessment, time, and finally space (Figure 3). The imbalance across categories, with Content having the most papers, reflects the literature's emphasis on developing new educational content and curricula using simulation/XR technologies, while fewer studies focused specifically on overcoming time or physical space limitations. To be included under the Content category, papers had to focus on novel instructional approaches/techniques/curriculum or improvement of existing ones with the overarching goal of fostering innovative teaching and learning practices and improving trainees' technical and/or nontechnical skills.

Figure 3 illustrates the number of publications per year from 2000 to 2020, marked by an initial emphasis on e-learning and basic virtual simulations (2000–2005), progressing to a substantial increase in studies on virtual patients and procedural simulators (2006–2013), and a rapid expansion phase (2014–2020) characterized by significant growth in XR research, enhanced assessment methodologies using simulation data, and the incorporation of AI. Importantly, more recent empirical studies are shifting away from superiority, non-inferiority, or equivalence trials comparing XR to traditional teaching methods and are moving toward the research paradigm surrounding how to optimally integrate XR technologies for maximum impact in terms of learning gains and transfer of learning from simulated experiences into clinical practice.

Clinical simulation has been increasingly incorporated into training across diverse medical and surgical specialties over the past 20 years. Studies show

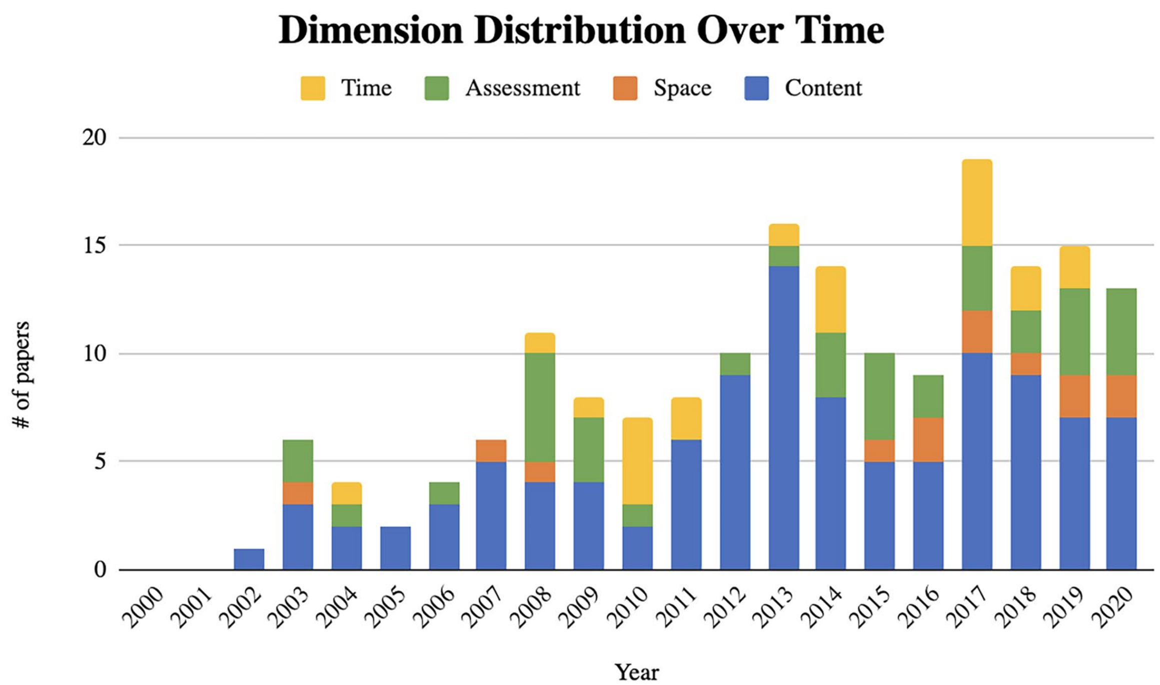

Figure 3. Number of publications in each dimension space, time, content, and assessment from 2000 to 2020.

adoption in areas like anesthesiology (e.g. simulators for endotracheal intubation, epidural placement, fiberoptic intubation [24,25]), radiology (e.g. VR simulation for ultrasound-guided procedures and CT interpretation [26,27]), obstetrics (e.g. simulators for obstetric emergencies like postpartum hemorrhage and shoulder dystocia [28]), pediatrics (e.g. virtual simulation for neonatal resuscitation and pediatric trauma management [29]). Traditionally, simulation-based training has involved using a mannikin as a patient in a simulated patient care setting. To improve accessibility, resource utilization, and the learner experience, there is a growing movement towards using virtual reality (VR)-based simulation training in healthcare [30,31].

Overall, there has been significant advancement in simulation technology for developing technical skills in specialties that require high levels of technical expertise, such as surgery (see Figure 4). There has also been increased use of simulation for training diagnostic reasoning skills and retaining knowledge across medical specialties, not just technical skills. For example, virtual patients for internal medicine education and cardiology cases [32,33], or obstetrics VR simulators to improve knowledge in managing complications [34,35]. The use of simulation technology has had a significant impact on the 'space' aspect of medical education. The number of papers focusing on the spatial aspect of simulation technology in medical education has grown gradually over the years, although this area remains less explored compared to other aspects. The total number of papers focusing on assessment in simulation technology in medical education is 37, indicating a steady interest in this area. Simulation technology has also affected the 'time' aspect of medical education, with 21 papers focusing on this topic, especially in recent years, with a peak in 2017.

Table 2 summarizes the results of the subgroup analysis based on specific technologies and target populations to provide deeper insights into which methods are most effective for different educational goals and learner groups. In 37% of the studies ($n = 65$), XR-type of technologies were the most prevalent, demonstrating particular effectiveness in surgical, anatomical learning, and procedural skills development. E-learning platforms were the second most common at 24% ($n = 42$), utilized across different learner groups with reported effectiveness in knowledge acquisition and facilitating self-paced learning. Simulation-based training at 20% ($n = 35$) showed strong performance improvements in team-based scenarios and crisis management. Serious games and gamification at 10% ($n = 18$) approaches proved engaging for medical students and showed effectiveness in teaching decision-making skills and procedural knowledge.

Reported sample sizes varied widely, ranging from 12 to 287 participants, with a mean of 72 and a median of 41. The most common interventions included VR simulations for immersive, risk-free practice, simulator-aided training courses using physical or computer-based models, web-based learning modules for flexible self-paced study, serious games to enhance decision-making and procedural skills, and e-learning platforms for remote

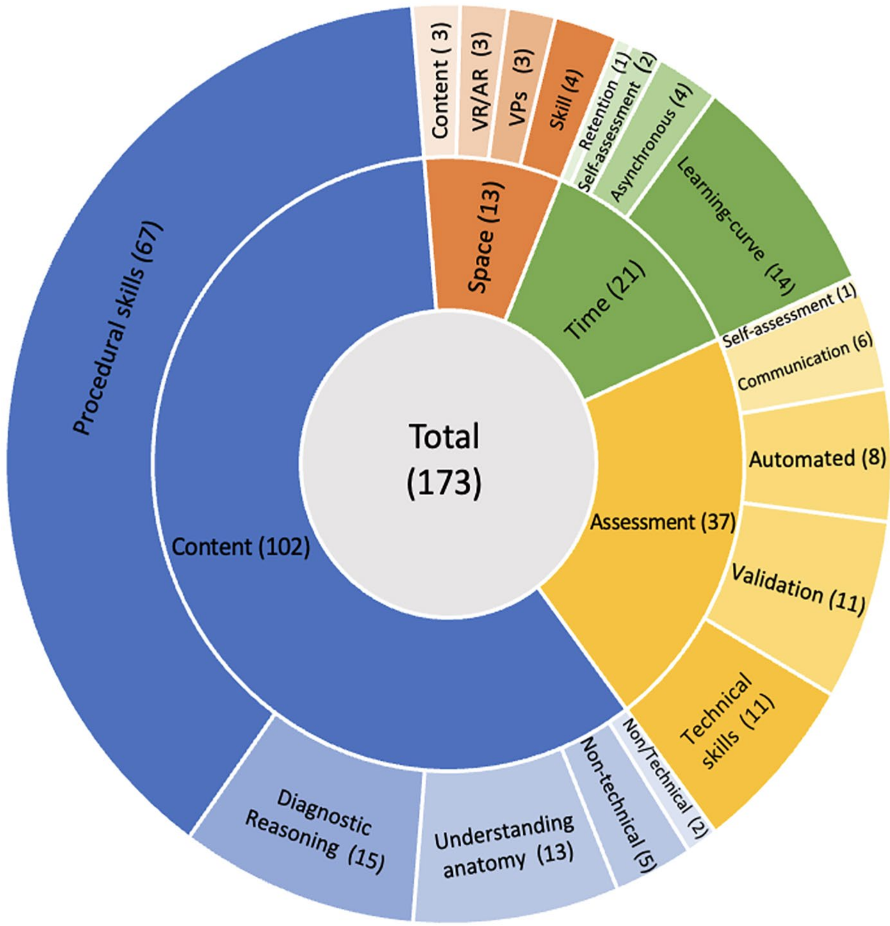

Figure 4. A comprehensive overview of 173 papers, organized by color and broken down by dimension with their corresponding themes. The inner circle of the chart depicts four main categories: Content in blue (102 papers), assessment in yellow (37 papers), time in green (21 papers), and space in orange (13 papers).

Table 2. Results of the subgroup analysis that maps educational technology, target learner group, and intended learning objective and effectiveness.

Technology *n = number of studies with specific technology	Target trainee populations	Educational goals	Effectiveness	Example studies
eXtended Reality (XR)-enriched simulations n = 65 (37%)	• Medical students • Surgical/emergency/internal medicine residents • Anesthesiology trainees • Radiology trainees • Pediatric trainees	• Anatomical understanding • Surgical skills • Procedural skills • Visuospatial reasoning	• Improved technical skills • Reduced learning curve • Enhanced visuospatial understanding and • Improved clinical performance	[36–40]
E-learning platforms n = 42 (24%)	• Medical students • Physicians (CME) • Nursing students • Pharmacy students • Physical and occupational therapists	• Knowledge acquisition • Self-paced learning • Remote access to content	• Improved knowledge retention • Increased accessibility • Flexible learning schedules	[41–43]
Simulation-based training n = 35 (20%)	• Emergency/internal/family medicine residents • Obstetrics and gynecology trainees • Critical care teams • Anesthesiology trainees • Surgical trainees • Pediatric trainees	• Team-based training • Crisis management skills • Procedural skills	• Improved team communication • Enhanced decision-making in critical situations • Increased procedural confidence	[44–46]
Serious games/ gamification n = 18 (10%)	• Medical students • Surgical trainees • Pharmacy students	• Decision-making skills • Procedural knowledge • Engagement in learning	• Increased motivation and engagement • Improved retention of procedural steps • Enhanced clinical decision-making	[47–49]
Other technologies (e.g. mobile apps, web-based tools) n = 13 (8%)	• Medical students • Nursing students • Allied health professionals	• Just-in-time learning • Self-assessment • Supplemental learning	• Improved access to resources • Enhanced self-directed learning • Complementary to traditional methods	[50,51]

access to educational content. Researchers measured effectiveness through various outcomes, with performance metrics (accuracy, time to completion, procedural efficiency) being the most common (68%), followed by learner perceptions (e.g. usability, realism) (55%), skill acquisition and retention (42%), knowledge gains (both self-reported and measured knowledge increases) (39%), and clinical applicability, such as impact on patient care and clinical practice (30%).

Studies that demonstrated *performance improvements* in terms of faster completion times and increased procedural accuracy were focused on both complex tasks like laparoscopic cholecystectomy and simpler, repetitive procedures, such as IV insertions [52,53]. The structured nature of training modules and real-time feedback mechanisms were key factors contributing to these improvements. *Usability and realism of simulation* tools were highlighted in ~55% of the studies. High-fidelity simulations were particularly effective in complex, team-based environments [54], while low-fidelity models proved beneficial for fundamental skill acquisition [55]. In terms of *knowledge gains,* these gains were most salient with learners from emergency medicine, critical care, and surgical disciplines [39,44,56]. Interactive and adaptive learning environments tend to enhance learners' diagnostic abilities and treatment management skills post-simulation intervention [57,58].

We structured the study results by revisiting key predictions made by Friedman [3] and analyzing to what extent they have come true, partially come true, or have not yet been realized based on the evolution of simulation and evidence from more recent research. Thus, we grouped the predictions into categories of physical space, time, content, and assessment. It is important to note the significant heterogeneity in terms of technologies, educational contexts, learner groups, and measured outcomes across the 173 included studies, which should be considered when interpreting the synthesized results presented in the following sections.

Physical space: expanding medical education beyond physical limitations through metaverse-based simulation technologies

The Friedman paper [3] predicted simulation technology would allow medical training to occur anytime and anywhere, but this prediction has proven only partially true. While web-based virtual simulations have certainly increased accessibility and flexibility, most simulation training still depends on dedicated centers with specialized equipment rather than being ubiquitously available anywhere. The studies included in the 'Space' dimension describe research on how metaverse-like technologies have allowed medical education to overcome dependence on physical locations. The selected papers related to the 'physical space' dimension fell into four categories: (1) virtual reality, (2) content-based e-learning, (3) skills-based e-learning, and (4) virtual patients.

Within the space domain, 8 out of the 13 studies were focused on skill and content-based e-learning. Web-based modules provide continuous skill practice and content learning unconstrained by place. For example, Guetterman et al. [59] showed virtual human programs could effectively teach communication skills remotely. By using a virtual patient, case-based online modules, or modern video conferencing tools, learners are freed from the restrictions of having to interact with standardized patients only in person.

Another major focus was virtual and augmented reality simulation. Based on the studies included in this scoping review, ~60% utilized some form of extended reality technologies like VR and AR, while 40% involved physical manikins, task trainers, or standardized patients. Historically, simulation training employed physical manikins, limiting access. Virtual reality simulation increases accessibility and scalability while reducing demands on resources and facilitators [60]. Analysis of the review studies shows a steady upward trend in papers on 'virtual reality' and 'medical education' since 2010, with growth acceleration around 2016–2017. Several review papers attribute the uptick in VR medical simulation after 2010/2011 to factors like improved graphics, haptics, AR integration, and motion tracking [61,62]. Earlier VR simulators had more rudimentary visual graphics and virtual anatomy. Newer systems like the VOXEL-MAN TempoSurg [63] and Visible Ear Simulator [64] feature enhanced 3D rendering and physics-based lighting for ultra-realistic illustrations of human anatomy. In surgical education, while early VR simulators had limited force feedback, new models incorporate advanced force-feedback haptics for highly realistic feeling of tissue properties and tool interactions during surgical tasks [65]. Overall, the studies demonstrate expanding capabilities to replicate clinical scenarios through digital environments, 3D visualizations, and immersive virtual worlds.

Unsticking medical education in time: improving learning curves, retention, and practice opportunities

Our analysis revealed significant developments in how medical education has been 'unstuck' from traditional time constraints over the past two decades. Friedman's paper [3] discusses the dimension of 'Time' in medical education and makes two predictions: the use of simulators will provide limitless practice opportunities and will eliminate the need for regimented, lockstep curricula. These predictions have become partially true. In short, modern simulators offer repeated practice on demand, but curricula are still fairly regimented, and full individualization has not occurred. According to Friedman, requiring faculty and learners to participate in class or training simultaneously in the same place can limit the time available to acquire new skills. Applying new technologies in medical education offers flexibility for developing competencies in areas like patient care, knowledge, procedures, teamwork, collaborative diagnostic reasoning, and critical thinking. The papers in this category show how increased flexibility in the time aspect of medical education has enhanced knowledge retention, quicker skill acquisition, and more efficient performance feedback loops. Four key themes emerged: (a) *accelerated learning curves*, (b) *improved retention*, (c) *increased opportunities for asynchronous learning*, and (d) *automated precision feedback for deliberate practice*.

Twelve studies across specialties like gastroenterology, obstetrics, orthopedics, and perfusion demonstrate that simulation-based training can significantly shorten the *length of learning curves* to attain competency for various technical skills compared to traditional clinical training alone. For instance, Loukas et al. [66] found a VR simulator significantly accelerated medical students' learning curve for intravenous cannulation. Yudkowsky et al. [67] showed that practice on an augmented reality/haptic simulator with a library of virtual brains improved neurosurgery residents' ability to successfully perform ventriculostomy cannulation on the first pass in both simulated and live procedures. In another study by Andersen et al. [68], structured and distributed virtual reality simulation practice before cadaveric dissection led to lower cognitive load and improved performance compared to standard practice during training for novice surgeons.

Research in the last two decades reveals that incorporating technology-enhanced simulation may result in superior *retention of clinical knowledge and skills* compared to traditional teaching methods alone. Studies across medical disciplines have shown that online learning, virtual simulation, and spaced education enable durable retention of knowledge, technical skills, and critical reasoning abilities for months to years after initial training. For example, in a randomized trial, Maagard et al. [39] showed VR simulator training enabled laparoscopic skills to be retained for up to 18 months, with only some decline after 6 months.

Four studies in this review demonstrated the value *of asynchronous online learning* compared to required synchronous lectures and training, which Friedman criticized for inflexibility. Early studies in the 2000s showed online discussion forums and computer-based modules could improve critical thinking and interaction beyond scheduled lectures [69–71]. Growth of learning management systems and e-learning platforms enabled increased adoption of asynchronous online learning in medicine [72]. Methods like online spaced education, personalized modules, and virtual patients enhance learning outcomes for knowledge, diagnostic skills, etc. [43,73,74]. For example, Cook et al. [73] found that spaced education improved long-term retention of clinical knowledge compared to traditional methods.

Automated scoring and mobile feedback technologies have been shown to help trainees save time, self-monitor progress, identify focus areas, and enable efficient deliberate practice to accelerate competency—advances that were not feasible through traditional training alone [75]. Tracking performance metrics longitudinally and with the use of artificial neural networks on procedural simulators allows monitoring of skill progression over time [76,77]. However, some research found limitations. Andersen et al. [64] showed VR assessment of mastoidectomy skills primarily measured efficiency rather than safe practices. Guided facilitation and supplemental feedback mechanisms were still needed.

These findings illustrate how technology has enabled more flexible, efficient, and personalized timing in medical education, addressing Friedman's concerns about rigid scheduling and pacing.

Content domain: skill and knowledge targets in simulation studies

Friedman [3] predicted that simulators would enable (a) on-demand practice of skills identified by faculty as priorities, (b) repetitive practice with variations on medical topics and cases, (c) recorded student performance data to provide feedback on areas needing work, and (d) self-paced learning and individual customization based on performance. This has been only partially realized, as the reviewed studies have shown successful application of simulation across many specialties and for various clinical skills (see Figure 5), but curriculum requirements still play a major role in guiding simulation content. The content dimension contains 102 papers focused on using simulation to improve medical students' skills and knowledge. These papers span multiple medical specialties, including

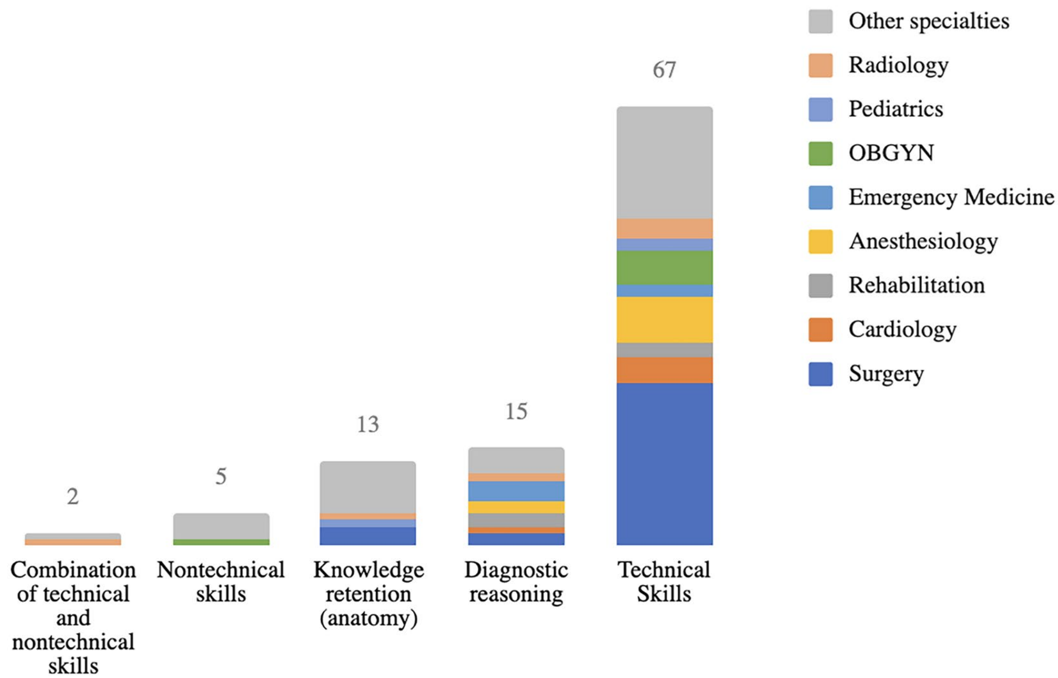

Figure 5. Distribution of specialties within content dimension and educational goal.

surgery, cardiology, rehabilitation, anesthesiology, emergency medicine, obstetrics/gynecology, pediatrics, radiology, and general clinical skills. The selected papers were further categorized into four main categories: non-technical skills, technical skills, diagnostic reasoning, and knowledge retention.

Sixty-seven of the 102 papers focused on the use of novel simulation-based training scenarios to improve trainees' technical skills for specific procedures; the majority of the papers were in surgical education. These studies demonstrated how simulation provides a low-risk training modality for trainees to gain proficiency in techniques before working with real patients. For example, Koch et al. [78] created a new virtual reality endoscopy simulator to significantly improve learners' performance of colonoscopies compared to baseline, with decreased procedure time and improved insertion depth. In a similar vein, Barsuk et al. [79,80] developed a mastery learning curriculum using high-fidelity simulation to improve central venous catheter insertion skills. Trainees went through an online module and then practiced on a mannequin until they demonstrated mastery based on a checklist. The simulator allowed endless opportunities to practice and receive detailed feedback on technique until reaching competency.

Fifteen studies focused on building learners' diagnostic reasoning and care management abilities through exposure to diverse clinical presentations and simulated clinical scenarios. For example, Giuliani et al. [81] used radiation oncology simulations of high-acuity, low-frequency cases to give participants practice determining appropriate responses. In this study, they selected 5 high-acuity, low-frequency clinical simulation scenarios in which participants had to determine how to manage the situation to get to the desired outcome. Participants highly valued these opportunities to strengthen clinical decision-making skills. Battaglia et al. [82] created an online simulation using a virtual diabetic patient which helped improve pharmacists' and pharmacy students' confidence in providing medication therapy management and their knowledge of how to deliver it. Learners could work at their own pace to master the delivery of medication therapy management.

Thirteen papers investigated using simulation to improve knowledge retention and anatomical understanding. For example, Vertemati et al. [83] developed an interactive virtual reality model for 3D visualization of patient-specific organs. This technology enhanced students' comprehension of organ structures. However, researchers cautioned about existing limitations in accurately modeling human anatomy and pathology processes digitally.

Other papers examined using novel simulation-based training scenarios to enhance non-technical skills like communication and teamwork. For instance, Chheang et al. [84] developed a multi-user virtual reality simulated environment to train interprofessional communication for anesthesiologists and surgeons during potential surgical complications in laparoscopic surgery. Bracq et al. [31] used a virtual reality simulation to train scrub nurses in error recognition and situation awareness by immersing them in a simulated operating room with embedded errors. The study results showed nurses who detected more errors had higher situation awareness, detected high-risk errors faster, and felt more immersed and satisfied with the experience.

Assessment: clinical simulation's bigger impact on formative vs. high-stakes assessment

Although the Assessment dimension was not explicitly discussed in the original paper [3], Friedman predicted that automated metrics and tracking from simulation technology would enable new competency-based assessment methods. In the last two decades, this prediction has proven partially true. Out of 173 studies reviewed, 37 pertained to assessment—the second largest category. Over the past two decades, technologies like virtual patients, learning analytics (refers to the 'measurement, collection, analysis and reporting of data about learners and their contexts…' [85], p. 1381), and immersive simulations have made formative competency evaluation more practical and informative, but have not yet transformed high-stakes assessments.

Seventeen studies out of 37 measured technical and non-technical skills using simulators and technology. These methods enhanced assessment quality, learning experience, and trainee proficiency [86,87]. Eleven studies provided validity evidence for specific simulation-based assessments, like tools to evaluate bronchoscopy skills (e.g. [88]). Nine studies examined automated, peer- and self-assessment, finding benefits like improved engagement and personalized feedback (e.g. [89,90]). For example, Guetterman et al. [59] showed that the virtual human application effectively assessed performance-based competence in breaking bad news to a cancer patient. Automated feedback after team-based simulations can be provided based on tracking non-technical skills through natural language processing and biosensor data [14,18,91].

However, the reviewed papers suggest that exclusively automated assessment lacks qualitative nuance, and validity evidence is still needed to integrate simulation into high-stakes testing [59,92,93]. Degree of assessment automation also depends on the simulation's main function (procedural skills *vs.* team-based training for interprofessional scenarios), fidelity (low or high) and number of learner at a time (single or multi-user scenarios). Advanced application of intelligent tutoring systems in procedural training, utilizing a state machine approach to provide real-time, adaptive feedback shows promising results [94]. While automation makes competency evaluation more practical, challenges remain in scaling such systems, data interpretability, trust and data privacy, infrastructure needs, and required data science expertise, especially for more complex, multi-user VR simulations.

Discussion

This scoping review analyzed the evolution of simulation technology, and emerging metaverse applications, over the past two decades to assess progress in transforming medical education around physical space, time, content, and assessment based on Friedman [3] paper. The results shed light on critical aspects of medical education, particularly in the context of simulation which has made partial but meaningful progress in expanding access beyond physical constraints, accelerating competency gain, and enabling more flexible educational experiences. However, some constraints around resources, content quality, curricular customization, and integration into high-stakes assessment persist.

Our findings both align with and extend those of a recent related review. While previous literature reviews and empirical studies on this topic have primarily focused on comparing XR to traditional teaching methods or educational technologies [1,95–97], our review recognizes that the field has evolved beyond this foundational question. Instead, our scoping review documents the progress over 20 years and pinpoints a critical shift in the research paradigm surrounding how to optimally integrate XR technologies across various dimensions of medical education for maximum impact. In addition, while few existing reviews focus on specific medical specialties, specific skills or learner groups (e.g. clinical anatomical education by McBain et al. [98]; intensive care unit staff training by Hill et al. [99]; surgical education [100]; cardiopulmonary resuscitation training by Trevi et al. [101], or nontechnical skills [102]), our study takes a broader approach and synthesizes findings across diverse areas of medicine and all levels of medical education in terms of impact on learning, usage, and effectiveness. In contrast to the review by Curran et al. [1] on artificial intelligence in medical education, which focused primarily on knowledge-based outcomes, our review highlights the potential of these technologies for both cognitive, behavioral, and psychomotor skill development (see Table 2). Additionally, our review of assessment methods in metaverse applications for medical education addresses a significant gap in the existing literature.

We mapped the current assessment landscape, highlighting the potential of multi-modal training ecosystems, and identified the shift toward personalized assessment, automation, and integration into high-stakes evaluations. The following discussion presents the implications and broader significance of the findings of our scoping review.

Extending medical education beyond physical boundaries

Our review's findings revealed that, in terms of the Space dimension, the selected articles fell into four groups: virtual reality, content-based e-learning, skills-based e-learning, and virtual patients. All these learning modalities allow learners to be free from the traditional constraints of physical classrooms and clinical practice settings, embracing innovative technologies and virtual platforms to provide a dynamic and flexible learning experience. While fewer studies directly addressed the space dimension, reviewed studies demonstrated how this multi-approach leverages e-learning [103], telemedicine [104,105], augmented reality [95,106], virtual simulations [36], and online resources to reach a wider audience at scale. At the same time, e-learning and metaverse-based simulations can support the development of the following six core competencies mandated by the Accreditation Council for Graduate Medical Education (ACGME) [72].

1. *Patient Care*: This competency focuses on the ability to deliver safe, evidence-based, and patient-centered care. E-learning and agent-based simulations have the potential to improve clinical skills, diagnostic reasoning, and knowledge of medical conditions. Safe practice of diagnosis, treatment planning, and care delivery is made possible by virtual patients that mimic real-world clinical interactions and react dynamically in response to learner activities [107].
2. *Medical Knowledge*: This competency entails the knowledge acquisition and clinical application of medical knowledge. E-learning tools, such as online courses and virtual libraries, can aid in the acquisition and retention of medical knowledge. Immersive 3D visualizations of anatomy and pathology can enhance understanding of structural relationships and disease processes [108].
3. *Practice-Based Learning and Improvement*: Medical professionals should be dedicated to lifelong learning and continuous improvement of their practice. E-learning platforms can support self-assessment, reflective practice, and quality improvement initiatives. Artificial Intelligence-driven dashboards that track skills progression across virtual and real clinical experiences to identify focus areas, knowledge gaps, and learning curves.
4. *Interpersonal and Communication Skills*: Effective communication is essential when interacting with patients, their families, and the medical staff. E-learning modules and multi-user virtual environments can help medical trainees develop their communication skills, including breaking bad news, counseling, and teamwork.
5. *Professionalism*: Professionalism includes ethical behavior, integrity, and accountability. Case-based XR-enhanced scenarios can include modules on medical ethics, cultural competency, and professionalism in healthcare.
6. *Systems-Based Practice*: A key component of this competency is understanding and navigating the healthcare system, advocating for patient safety, and collaborating effectively with others in the healthcare system. Virtual hospitals and clinics that simulate coordinated care across departments and professions to understand system interactions and quality improvement processes.

Overall, all modalities mentioned above can collectively contribute to a comprehensive and adaptable medical education system, enabling all levels of trainees to acquire both theoretical knowledge and practical skills while accommodating various learning preferences. Moreover, by removing the barriers of geography and limited physical resources, medical training on digital platforms allows learners to access a diverse array of clinical cases, collaborate with peers and experts globally, and tailor their learning journeys to individual needs. While extending medical education beyond physical boundaries through technology-enhanced learning modalities offers benefits, effectively embracing flexibility and time independence in medical education presents challenges.

Embracing flexibility and time-independence in medical education

Friedman [3] advocated for more flexible learning models, where learners could progress at their own pace, accessing educational materials without following a traditional curriculum with set timelines for completion. He believed that technology could facilitate this flexibility and adaptability in medical education. The rigid schedules and fixed timelines that have characterized medical education for generations can be limiting or not accommodating the diverse needs of today's learners [109], especially those who have family responsibilities or wish to pursue education later in life. By embracing more flexible and asynchronous learning approaches, medical institutions can empower students and practitioners to tailor their education to their unique needs and circumstances [103]. Technology-enhanced asynchronous learning modalities are well suited to help instructors meet several challenges of medical education, including (1) the need and desire to promote self-directed learning, (2) providing flexible learning opportunities, (3) offering continuous (24 h/day/7 days a week) availability for learners, and (4) engaging learners through collaborative learning communities to gain significant learning and augment continuous professional development.

While time-related studies were less numerous ($n = 21$), they revealed important trends. The results of our review indicated that there is a need for asynchronous learning. Embracing technology, such as online modules and virtual simulations, can further facilitate this shift towards a more accessible and adaptive medical education system. While predictions around ubiquitous access and individualized content have not fully materialized, gains have occurred in offering asynchronous learning opportunities that supplement traditional curricula. A key metaverse-enabled capability is asynchronous online learning, providing flexibility beyond physical and scheduling constraints. This is consistent with other reviews showing online learning as an effective supplement, rather than a replacement, for in-person instruction [110,111]. A blended approach balancing asynchronous modules with some synchronous activities and peer interactions is ideal [112]. However, truly ubiquitous access and individualized content envisioned for metaverse learning have yet to fully materialize. For example, current best practices for team-based, acute-care instruction involve manikin-based medical simulation are limited or unavailable in many community, rural, and under-resourced hospitals, leading to inequity in access to training and contributing to disparities in care [113]. With an average headset cost of $430 [114], remote training can be accessed by populations around the world on learners' own schedules for a fraction of the cost of conventional medical simulation training. It is crucial to acknowledge that the implementation of this technology may still raise significant hurdles that must be carefully considered especially in low and middle-income countries (LMICs) [115].

Content quality in simulation and metaverse-based learning

For decades teaching and learning in medicine has centered on didactic lectures along with supplemental journal article readings. The universal usage of lectures is considered the most effective mode of information transfer in medical education, but this method of learning is associated with authoritarianism, poor lecturers, learners' passivity, and poor retention. Friedman's vision for medical education involved a more dynamic and adaptive approach to content delivery. He believed that technology could facilitate the continuous updating of educational materials, allowing students to access the most current and relevant information in the field of medicine. However, e-content development is a considerably new field that has emerged in response to the rapid advancements in technology and the increasing demand for digital learning resources [116]. This field encompasses the creation of educational content in various digital formats, such as e-books, online courses, interactive modules, movies, animations, simulations, interactive tests, interactive activities, and multimedia presentations. E-content development also requires pedagogical expertise with technical skills to design engaging and effective learning materials that cater to diverse learners' needs. The digital nature of e-content allows for flexibility in delivery, accessibility, and customization for multi-purpose. Thus, e-content quality plays an important role since it directly influences the success of the learning process and learner satisfaction. In both simulations and metaverse-based learning, content quality should align with educational objectives, promote active learning, and provide opportunities for learners to apply theoretical knowledge in practical, clinically relevant contexts. A study by Barsuk et al. [117] demonstrated that a simulation-based mastery learning curriculum for central venous catheter insertion, with online learning modules and deliberate practice on manikins, significantly reduced bloodstream infections compared to traditional ward-based training. This exemplifies the value of aligned objectives, active learning, and opportunities for deliberate practice.

Regular updates to content to reflect the latest advancements in medicine are also essential to maintain content quality [116]. Moreover, content should be designed to ensure accessibility and inclusivity to accommodate diverse learners, including those with disabilities, diverse demographic backgrounds (age, race, and gender), as well as individuals from diverse socioeconomic or cultural backgrounds, to ensure equitable access to medical education [109,118]. Metaverse environments for medical education should incorporate accessibility features, such as flexible user interfaces controllable through various modalities, closed captioning, adjustable text options, and wheelchair representations to accommodate disabilities. From the design perspective, the quality of simulation content depends on its realism and fidelity that use real-life clinical scenarios embedding robust feedback mechanisms. Simulations also vary in complexity, catering to learners at different stages of their education and training. Content should progress from basic skills to more advanced clinical scenarios, allowing learners to build upon their knowledge and skills incrementally. Metaverse-based learning also requires high-quality immersive and interactive environments that engage learners. These environments can simulate medical scenarios, anatomical structures, surgeries, or patient interactions within a virtual or augmented reality space.

Both simulation and metaverse-based learning require high-quality content to ensure that healthcare professionals receive the best possible training and education. As technology continues to evolve, e-content development will likely play an increasingly significant role in shaping the future of medical education, offering innovative ways to impart knowledge and facilitate lifelong learning.

Assessment gap

Assessment and evaluation are vehicles for educational improvement. For decades, educators have predominantly relied on traditional assessment approaches like tests and examinations to gauge learners, rank them, and deliver a final score or summary assessment. However, the advent of technological advancements has ushered in paradigm shifts in learners' expectations and teaching methods, rendering these traditional assessment approaches inadequate. To achieve valid and reliable gauges of learning process and gains within the technology-driven educational landscape, it is imperative to transition from traditional assessment to performance-based evaluation. Although a wide range of approaches are available to assess medical trainees' performance, such as direct observation, multisource feedback, milestones, and other appraisal forms, these tools may be poorly designed, too complicated, too long, or short and moreover, are not appropriate in every situation [119]. In fact, no single method can appropriately measure all aspects of learning in a digital environment, but learning analytics can be instrumental in evaluating learners' performance. This learner-produced intelligent data is a powerful analysis model to measure learners' success, discover information and social connections, and to predict learning outcomes, and/or to support the existing educational models.

The results regarding the impact of clinical simulations on formative and high-stakes assessment reveal an intriguing dimension of medical education. It becomes evident that clinical simulations offer substantial advantages in formative assessment, enabling students to receive feedback and improve their skills continuously. However, their role in high-stakes assessments, such as licensing exams, requires further exploration. The transition to Competency-Based Medical Education (CBME) aligns well with integrating XR technologies in high-stakes assessments. CBME system requires learners to demonstrate competence—the ability to independently perform tasks successfully and efficiently—before advancing to more challenging tasks or certification for independent practice [120]. To effectively implement CBME and leverage metaverse technologies for assessments, medical educators can adopt several practical solutions. For instance, in a virtual surgical simulation, the system could track metrics, such as hand movements, time taken for specific steps, and accuracy of incisions. This data can be immediately presented to the trainee post-assessment [121]. Another significant innovation emerged during the COVID-19 pandemic is the introduction of remote Objective Structured Clinical Examinations (OSCEs). These virtual examinations utilize video conferencing and interactive tools to simulate realistic patient encounters, allowing for assessment when in-person exams are not possible. Studies have shown that virtual OSCEs can be as effective as traditional formats in evaluating student competencies [122,123]. XR-based OSCEs provide standardized scenarios, enhancing objectivity in evaluations. Also, integration of XR assessment platforms with existing Learning Management Systems (LMS) and/or National Board of Medical Examiners (NBME) validated assessment instruments is important for streamlining operations and automatically updating student records. For example, NBME, a leading medical assessment organization, has recently acquired MedVR Education, XR platform for health care skill development.

Interactive e-learning modules with AR for anatomical dissections and quizzes may enhance consistency and objectivity in high-stakes assessments [124]. They provide standardized content, uniform learning experiences, precise scoring mechanisms, and adaptable difficulty levels while eliminating many of the variables that can introduce bias or inconsistency in traditional assessment methods. In addition, diagnostic reasoning simulations with automated agents can effectively evaluate clinical decision-making processes through complex case scenarios [107]. It is also beneficial to the medical education community to have increased attention to assessment and use of advanced learning analytics methods because this provides trainees with insight into their own learning and offers medical educators opportunities to make evidence-based interventions for the improvement of teaching and learning. According to the reviewed studies, types of learning analytics and data sources can include (1) performance metrics from simulation systems (e.g. time to complete procedure, errors made, efficiency of motions, etc.); (2) behavioral data tracking (e.g. communication patterns, leadership behaviors, coordination activities); (3) integrating and correlating simulation performance data with competency assessments from faculty observations in real clinical environments; (4) learner dashboards that visualize progress on skills acquisition longitudinally, revealing learning curves, strengths, and weaknesses. For competency assessment, metaverse capabilities allow continuous performance tracking and data-driven feedback. However, adoption remains limited for high-stakes examinations. Metaverse-enabled integrated systems that blend automated scoring with human observations could enable next-generation competency assessment. Optimal assessment blends human observations with sensor-based performance data [17,91]. Infrastructure costs and access barriers also remain.

Within technical skills training, a key emerging metaverse capability is the ecosystems that seamlessly integrate different fidelity levels to optimize technical skills gains. For instance, a technical skill like laparoscopic surgery in a metaverse training ecosystem could provide (a) 2D video box trainers that build initial familiarity (with anatomy, hand motions and instrument handling, e.g. Simball Box, LAP Mentor [125], (b) VR that develops core skills (e.g. spatial orientation, simulated laparoscopic operation, camera navigation [126]), and (c) mixed reality may overlay during actual laparoscopic procedures in the operating room to provide guidance and feedback [127,128]. This blend of modalities at different fidelity levels, enabled by metaverse connectivity, allows each to be leveraged at the appropriate stage. Data sharing across the ecosystem, including performance metrics, gaze, and eye-hand coordination data, further optimizes the training. This integrated metaverse with multi-modal curricula may allow trainees to progress across levels, ensuring each experience effectively builds toward mastery. In the reviewed studies published in recent years, we have seen these metaverse modalities begin to form metaverse ecosystems and training workflows in a wide range of medical specialties beyond surgical training.

Assessment and evaluation in clinical simulations and metaverse technologies hold the promise of authenticity and engagement. Learners are placed in realistic clinical scenarios, allowing for the observation of their clinical skills, communication, and decision-making abilities. While these methods offer numerous advantages in formative assessment and engagement, aligning them with the traditional curriculum poses challenges, particularly concerning high-stakes assessments. As long as these technologies are integrated thoughtfully, assessments are standardized, faculty training is provided, and evaluations of their effectiveness are continuously conducted, medical education can bridge the gap and harness the full potential of these innovative tools.

Research gaps and future directions

While the full vision of personalized and ubiquitous learning has not yet been achieved, gains have occurred in offering more flexible asynchronous opportunities that increase accessibility beyond physical constraints. Thoughtful integration of metaverse applications into training curricula is needed, evaluating their unique affordances while ensuring accessibility [129]. These findings underscore the need for continuous research and innovation in medical pedagogy, with the ultimate goal of producing highly competent and adaptable healthcare professionals. As part of our metaverse research agenda, we outline several key topics relating to the future of medical education and clinical simulation in particular:

- *Establishing best practices for integrating metaverse technologies into pedagogy.* Possible research directions may examine if virtual simulations and digital twin hospitals [130] in the metaverse enhance clinical skills and readiness for clinical workplace transitions compared to traditional clinical rotations alone. Develop evidence-based guidance on how much time learners, accounting for and accommodating their diverse needs, should spend in metaverse simulations *vs.* physical task trainers or simulators to reach competency benchmarks for specific procedures or training scenarios. There is also a need to resolve issues of virtual environment fidelity and accessibility for diverse learners. Future research should continue establishing evidence-based pedagogical frameworks that guide the design and implementation of immersive learning experiences in healthcare education (e.g. [131,132]).

- *Leveraging the affordances of computational artificial intelligence to optimize metaverse-enabled simulation experiences and assessment* The popularity and recent emergence of multimodal large language models (M-LLMs) marks a significant advancement in artificial intelligence capabilities for medical education [133]. These sophisticated AI systems, capable of processing and generating text, images, videos, and sound, offer unprecedented potential for enhancing learner experiences [134]. M-LLMs could accelerate production of the educational content, play a role of teaching assistant, and offer new ways to analyze learner data and provide personalized precision feedback. M-LLMs models could dynamically generate diverse, realistic clinical scenarios, including visual and auditory elements, as well as serve as personalized teaching assistants (e.g. explain complex concepts, practice breaking bad news, etc.). During simulations, M-LLMs could offer decision support by integrating information from various sources, such as simulated patient data, medical imaging, and lab results to assist in diagnosis, treatment planning, and team orchestration. Post-simulation, they could generate comprehensive debriefing materials, including personalized learning points based on the trainee's biosensor and behavior data. However, as medical training integrates these advanced AI systems, careful consideration must be given to ensuring the accuracy and reliability of AI-generated content, promoting critical thinking appropriate and preventing overreliance on AI, maintaining a balance between AI-driven and human-led instruction, addressing potential biases, misinformation in AI systems, protecting learner privacy and data security, and developing frameworks for explainable [135] and ethical use of AI in medical education.
- *Determining the utility and lasting impact of metaverse-enabled simulators in transferring skills learned on the simulated model to the clinical setting.* While many studies demonstrate immediate improvements in knowledge and procedural skills, future research should prioritize longitudinal research assessing learning gains and transfer of learning to clinical practice. Future studies should employ extended follow-up periods to evaluate the lasting impact of immersive learning experiences. One possible research direction could involve tracking the progress of trainees who utilize metaverse technologies for simulation training. Key components may include comparing patient outcomes before and after the implementation of metaverse technologies, assessing changes in the healthcare professionals' confidence and skill levels, and exploring how these technologies have influenced specific aspects of patient care, such as diagnosis accuracy, and treatment effectiveness. Another promising research topic could focus on the assessment of patient and provider satisfaction with metaverse-enabled telemedicine visits involving realistic avatars and virtual environments compared to video visits.
- *Broadening diversity and inclusion in application domains and learner groups.* Our review found a concentration of studies in surgery, urology, and anatomy. While these areas have seen significant advancements, other domains like public health training, rural and global health training, interdisciplinary teamwork, patient education remain underexplored [95]. Importantly, our review found limited research on how simulation and metaverse technologies impact learners from diverse backgrounds or with disabilities. Future research should prioritize inclusive design and evaluation of these technologies for all learners.
- *Prioritizing ethical considerations.* As metaverse technologies become more prevalent in medical education, ethical issues surrounding data privacy, fairness, trust in AI, bias in assessment, and equitable access need careful consideration [136]. Only 2% (3/173) of studies addressed these critical aspects, highlighting a significant gap in the current literature.

Practical applications of metaverse technologies for medical educators: a comparative analysis of affordances and limitations

Table 3 presents an overview of various metaverse technologies and their applications in medical education. Each technology offers unique affordances and faces distinct limitations in the context of medical training. Educators should consider these unique characteristics when selecting appropriate tools for their specific learning objectives and institutional resources. The choice of technology should align with the desired learning outcomes, available infrastructure, and target learner group to maximize educational impact.

Table 3. Comparative analysis of metaverse technologies: learning activities, affordances, and limitations.

Technology type	Learning activities	Unique pedagogical and technological affordances	Limitations	Example studies
Virtual reality (VR)	• *Immersive Anatomy Exploration:* Engage with 3D anatomical models in an immersive environment for a deeper understanding of complex structures. • *Multi-user VR simulations:* Enable collaborative learning by providing an immersive experience and presenting standardized stimuli and infinite attempts. • *Surgical Practice Simulations:* Conduct surgeries in a risk-free virtual setting to improve skills and decision-making, with real-time feedback. • *Virtual Patient Interaction:* Interact with virtual patients to practice diagnostic reasoning, communication, and empathy.	• **High Immersion:** Creates a sense of presence and realism in a risk-free environment. Leveraging smart haptic gloves for VR surgery simulation • **Effective Skill Transfer:** May facilitate the transfer of learning to clinical practice. • **Increased Engagement:** Interactive and engaging, improving learning retention.	• **High Cost:** Advanced VR systems and maintenance are expensive. • **Limited Accessibility:** Requires specific hardware like VR headsets. • **Physical Discomfort:** Extended use may cause motion sickness or discomfort.	[36–40, 137,138]
Augmented reality (AR)	• *Enhanced Anatomical Study:* Overlay virtual images onto physical models or cadavers, providing extra insights during dissections. • *Augmented Medical Imaging:* Display 3D scans directly onto a patient's body for better visualization and surgical planning. • *Telestration for Remote Proctoring and Telementoring:* Enables remote surgical mentoring through real-time visual annotations, augmented reality overlays, and live streaming of expert hands, enhancing communication and guidance during procedures. • *Procedural Training Aids:* Provide step-by-step instructions overlaid onto patients or equipment, guiding medical procedures (e.g. medication administration or ACLS algorithm)	• **Enhanced Reality:** Adds digital layers to real-world views & objects, enhancing understanding. • **Easily Accessible:** Usable with common devices like smartphones and tablets, or HoloLens for more computationally demanding tasks. • **Real-Time Guidance:** Offers live guidance for procedures, improving accuracy and navigational tasks.	• **Dependence on Cameras and Sensors:** Effectiveness relies on camera resolution/ performance and hands/object detection, eye calibration, environment lighting (HoloLens). • **Potential for Extra Cognitive Load:** Overlaid information can be overwhelming/distracting and cause additional cognitive load.	[40, 139–142]
Mixed reality (MR)	• *Integrated Simulations:* Merge physical simulators with virtual elements for complex and realistic training scenarios. • *Coordinated Team Training:* Simulate scenarios requiring coordination among medical teams, enhancing teamwork and communication skills. • *Interactive Procedural Training:* Interact with both real and virtual objects to practice complex medical procedures (e.g. peripheral intravenous catheter placement).	• **Interactive Realism:** High interaction between real and digital objects provides comprehensive training. • **Team Coordination:** Enables realistic team training scenarios. • **Versatile Applications:** Combines the best of VR and AR for diverse applications.	• **High Cost:** Advanced MR systems and headsets are expensive. • **Complexity:** Requires sophisticated integration of real and virtual elements. • **Limited Field of View:** Users may experience a restricted field of view.	[46,143,144]
Simulation-based training (task-trainers, high-fidelity manikins)	• *Clinical Skills Enhancement(e.g.,:* Use lifelike simulations to practice clinical skills in a controlled environment. • *Critical Care Scenarios:* Simulate emergency scenarios to improve technical and nontechnical skills. • *Procedural Proficiency using Task-Trainers:* Practice procedures repeatedly to achieve proficiency (e.g. laparoscopic manual tasks)	• **Realistic Training:** Provides high-fidelity simulations that mimic real-life scenarios. • **Repeatability:** Allows for repeated practice to build proficiency. • **Immediate Feedback:** Offers instant feedback to correct mistakes and improve techniques.	• **Resource Intensive:** High setup and maintenance costs. • **Time-Consuming:** Requires dedicated time slots and simulation center capacity. • **Limited Scalability**: typically relies on expert human facilitation, which limits implementation at scale.	[41,42,44,45]
E-learning platforms	• *Self-Paced Learning:* Access course materials and lectures at anywhere, any time and pace. • *Continuing Medical Education:* Online modules for continual skill and knowledge updates. • *Cognitive Skill Development:* Interactive assessments to test and reinforce understanding.	• **Flexible Access:** Learn any time anywhere. • **Self-Directed:** Allows learners to regulate their own learning pace. • **Updated Content:** Easily update content for latest practices and guidelines.	• **Reduced Interaction:** Limited face-to-face interaction with educators and peers. • **Engagement Issues:** May be less engaging without in-person motivation. • **Limited to mainly cognitive task practice:** Primarily suitable for self-practice on cognitive tasks	[42,43,145]
Serious games	• *Gamified Learning:* Use game-based methods to teach clinical decision-making, diagnostic reasoning, and procedural knowledge. • *Interactive Scenarios:* Engage with interactive scenarios to solve medical cases. • *Motivational Learning:* Increased motivation through competitive elements and rewards.	• **Active Learning:** Engages learners actively, improving retention. • **Leveraging Intelligent Tutoring Strategies:** combining finite state machines to model different phases in the diagnostic process or offering personalized feedback	• **Quality Variability:** Quality and educational value can vary widely. • **Accessibility:** May require specific gaming hardware or software as well as programming abilities.	[47–49]

While the potential for innovative learning experiences is immense, it is crucial to acknowledge that the implementation of this technology may raise significant hurdles universal in nature and that must be carefully considered specific to low and middle-income countries (LMICs) [115]. On a global scale, technical challenges, ethical considerations, potential barriers for learners with disabilities are universal and equally pressing, with concerns about privacy, data security, and equitable access at the forefront. In LMICs, these universal challenges have an additional layer of complexity. Specifically, financial barriers driven by the high costs of equipment, maintenance, and infrastructure may be prohibitive for many institutions [115]. Connectivity issues, including unreliable internet and frequent outages, my disrupt learning flow. New technologies may also prove challenging in regions where digital literacy and familiarity with advanced systems are limited. Furthermore, cultural adaptation is essential to ensure that the metaverse aligns with local customs and healthcare needs [146]. To address these issues of equity and accessibility across diverse learner populations, in our prior work we suggested a range of potential solutions and considerations (see [115] for details).

To integrate metaverse applications powered by AI effectively into medical curricula, a comprehensive approach involving careful planning, stakeholder collaboration, and ongoing evaluation is essential. The detailed guidance with successful implementation examples is provided in our earlier work (see [147] for review). In brief, implementation should follow a structured process. This begins with comprehensive *readiness and needs assessments*, involving evaluation of the current technological infrastructure, identifying gaps in the curriculum that these technologies could address, and assessing the readiness of faculty and students to adopt these new tools. Second, *professional development for faculty and trainees* that includes proper orientation to navigate XR-enriched environments and practice movements (teleportation) and interactions with XR objects and associate their function with concepts from the real-world. Before full-scale adoption, pilot programs are crucial to identify and address potential issues related to technology or pedagogy. Third, *continuous evaluation and quality improvement* are vital to the success of these technological integrations. This involves collecting data on student engagement and performance as well as surveying both students and faculty about their experiences with XR-enhanced learning materials.

Further research should continue elucidating best practices for integrating metaverse applications into medical education, ensuring training innovations are equitable, human-centered, and drive mastery.

Limitations of the study

This scoping review has some limitations. Our review focused on publications between 2000 and 18 December 2020, a period we believed captured the rise and influence of metaverse technologies in medical education. However, this timeframe may not encompass all relevant developments. First, some innovative extended reality-based educational methods were documented before 2000. For instance, Hoffman and Vu [148] paper examined several virtual reality applications as teaching tools of the twenty-first century. Second, there are hundreds of new papers published since 2021. Consequently, the conclusions drawn from this study are specifically applicable to the chosen timeframe. To gain a more comprehensive understanding, future research should explore publications from other time periods, potentially uncovering additional insights on the most recent innovations that were not captured in this review.

Despite comprehensive searches across major databases, some relevant studies may have been missed. The quality of the studies included in this scoping review may vary, which could impact the overall reliability and validity of the synthesized findings. Some studies may have inherent methodological limitations that could affect the conclusions drawn. The process of thematically coding articles into dimensions involved subjective judgments by researchers and alternative classifications could be justified. However, the rigorous process of constant comparison and consensus building amongst coders adds validity. There is also potential for expanding the search strategy with an increased set of keywords and synonyms and additional databases, such as IEEE Xplore to include a broader range of immersive technologies. We also note that a scoping review primarily provides a broad map of the literature landscape on a topic rather than an exhaustive systematic analysis. Our aim was to analyze high-level trends over the past two decades, which this methodology achieved. Additionally, while we analyzed publication trends over time, more sophisticated bibliometric analyses could reveal deeper insights into research patterns. Our findings are best interpreted as identifying dominant themes and trajectories to date, which can inform future research directions.

While Friedman's framework provided a valuable structure for our analysis, it is important to acknowledge its limitations as an analytical lens. The predetermined categories of space, time, content, and assessment may have limited our ability to identify emerging themes or innovations that fall outside these dimensions. Additionally, the framework's focus on 'unsticking' medical education from traditional constraints may not fully capture the nuances of how technology integration occurs in practice. Future reviews could benefit from employing multiple theoretical frameworks, such as (a) focusing on adoption, implementation, and trends of educational technologies, and/or (b) examining learning processes and outcomes. For adoption, implementation, and trends, future studies could utilize the Diffusion of Innovations theory [149] to analyze technology spread, Normalization Process Theory [150], or the Gartner Hype Cycle [151] to examine institutional integration over time. To investigate learning processes and outcomes, researchers could employ the Cognitive Affective Model of Immersive Learning (CAMIL) [152] for immersive technologies or the Technology Acceptance Model [153] to understand educators' integration of technology with pedagogy and content.

Conclusion

Medical education is undergoing a profound transformation driven by advancements in technology. This transformation is not fundamentally about adopting new technologies, but rather about improving educational design, standardization, scale, and assessment. Traditional methods of teaching and learning are giving way to innovative approaches that harness the power of clinical simulations and metaverse technologies to enhance learning. The last two decades have witnessed promising innovative teaching and learning approaches. However, this shift poses significant questions about 'unsticking' medical education across key dimensions, including the alignment of assessment techniques with the traditional curriculum. Our study findings can inform simulation development and research priorities going forward. Realizing the full vision of the 'Marvelous Medical Education Machine' proposed by Friedman [3] will require metaverse ecosystems integrating virtual patients, procedural simulators, collaborative environments, multimodal learning analytics, and competency dashboards into an accessible, personalized, flexible, and validated training ecosystem.

Acknowledgements

We extend our appreciation to the Health Sciences Librarian at the University of Michigan Medical School for formulating search terms and facilitating the literature review.

Author contributions

All authors conceptualized, performed a literature search, drafted, and revised the manuscript. VP contributed to the study conception, design, and analysis of the selected paper. The first draft of the manuscript was primarily compiled by the first author, who synthesized the accumulated literature and mapped out the foundational arguments. NM and CJ performed all screening phases, analysis, and reporting findings. KOL critically revised the manuscript for intellectual content, contributing significant expertise, and valuable insights into various aspects of medical education and technology-enhanced learning. KOL's input greatly enriched the content of this scoping review. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Science Foundation under Grant number 2202451.

ORCID

Vitaliy Popov http://orcid.org/0000-0003-2348-5285
Kadriye O. Lewis http://orcid.org/0000-0001-5947-409X

Data availability statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

[1] Curran VR, Xu X, Aydin MY, et al. Use of extended reality in medical education: an integrative review. Med Sci Educ. 2023;33(1):275–286. doi: 10.1007/s40670-022-01698-4.

[2] James CA, Wheelock KM, Woolliscroft JO. Machine learning: the next paradigm shift in medical education. Acad Med. 2021;96(7):954–957. doi: 10.1097/ACM.0000000000003943.

[3] Friedman CP. The marvelous medical education machine or how medical education can be 'unstuck' in time. Med Teach. 2000;22(5):496–502. doi: 10.1080/01421590050110786.

[4] Adler B, Biggs WS, Bazemore AW. State patterns in medical school expansion, 2000–2010: variation, discord, and policy priorities. Acad Med. 2013;88(12):1849–1854.

[5] Densen P. Challenges and opportunities facing medical education. *Trans Am Clin Climatol Assoc*. 2011;122:48.

[6] Schwartz, C. C., Ajjarapu, A. S., Stamy, C. D., & Schwinn, D. A. Comprehensive history of 3-year and accelerated

US medical school programs: a century in review. Med Educ Online. 2018;23(1):1530557.
[7] Ritterbusch GD, Teichmann MR. Defining the metaverse: a systematic literature review. IEEE Access. 2023;11:12368–12377. doi: 10.1109/ACCESS.2023.3241809.
[8] Quail NPA, Boyle JG. Virtual patients in health professions education. Adv Exp Med Biol. 2019;1171:25–35. doi: 10.1007/978-3-030-24281-7_3.
[9] Gani A, Pickering O, Ellis C, et al. Impact of haptic feedback on surgical training outcomes: a randomised controlled trial of haptic versus non-haptic immersive virtual reality training. Ann Med Surg. 2022;83:104734. doi: 10.1016/j.amsu.2022.104734.
[10] Suresh D, Aydin A, James S, et al. The role of augmented reality in surgical training: a systematic review. Surg Innov. 2023;30(3):366–382. doi: 10.1177/15533506221140506.
[11] Isenstein EL, Waz T, LoPrete A, et al. Rapid assessment of hand reaching using virtual reality and application in cerebellar stroke. PLOS One. 2022;17(9):e0275220. doi: 10.1371/journal.pone.0275220.
[12] Echeverria V, Martinez-Maldonado R, Yan L, et al. HuCETA: a framework for human-centered embodied teamwork analytics. IEEE Pervasive Comput. 2023;22(1):39–49. doi: 10.1109/MPRV.2022.3217454.
[13] Goldbraikh A, D'Angelo A-L, Pugh CM, et al. Video-based fully automatic assessment of open surgery suturing skills. Int J Comput Assist Radiol Surg. 2022;17(3):437–448. doi: 10.1007/s11548-022-02559-6.
[14] Yan L, Echeverria V, Jin Y, et al. Evidence-based multimodal learning analytics for feedback and reflection in collaborative learning. Br J Educ Tech. 2024;55(5):1900–1925. doi: 10.1111/bjet.13498.
[15] Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. Adv Simul. 2016;1(1):31. doi: 10.1186/s41077-016-0033-y.
[16] Radkowitsch A, Fischer MR, Schmidmaier R, et al. Learning to diagnose collaboratively: validating a simulation for medical students. GMS J Med Educ. 2020;37(5):Doc51. doi: 10.3205/zma001344.
[17] Rochlen LR, Putnam EM, Tait AR, et al. Sequential behavioral analysis: a novel approach to help understand clinical decision-making patterns in extended reality simulated scenarios. Simul Healthc. 2023;18(5):321–325. doi: 10.1097/SIH.0000000000000686.
[18] Rosen MA, Dietz AS, Yang T, et al. An integrative framework for sensor-based measurement of teamwork in healthcare. J Am Med Inform Assoc. 2015;22(1):11–18. doi: 10.1136/amiajnl-2013-002606.
[19] Page MJ, McKenzie JE, Bossuyt PM, et al. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. J Clin Epidemiol. 2021;134:103–112. doi: 10.1016/j.jclinepi.2021.02.003.
[20] Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth. 2020;18(10):2119–2126. doi: 10.11124/JBIES-20-00167.
[21] Knobloch K, Yoon U, Vogt PM. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and publication bias. J Craniomaxillofac Surg. 2011;39(2):91–92. doi: 10.1016/j.jcms.2010.11.001.
[22] Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI. 2018;34(4):285–291. doi: 10.3233/EFI-180221.
[23] Salavitabar A, Popov V, Nelson J, et al. Extended reality international grand rounds: an innovative approach to medical education in the pandemic era. Acad Med. 2022;97(7):1017–1020. doi: 10.1097/acm.0000000000004636.
[24] Boulet JR, Murray DJ, Warner DS. Simulation-based assessment in anesthesiology. Anesthesiology. 2010;112(4):1041–1052. doi: 10.1097/ALN.0b013e3181cea265.
[25] Kennedy CC, Cannon EK, Warner DO, et al. Advanced airway management simulation training in medical education. Crit Care Med. 2014;42(1):169–178. doi: 10.1097/CCM.0b013e31829a721f.
[26] Davrieux CF, Giménez ME, González CA, et al. Mixed reality navigation system for ultrasound-guided percutaneous punctures: a pre-clinical evaluation. Surg Endosc. 2020;34(1):226–230. doi: 10.1007/s00464-019-06755-5.
[27] Ni D, Chan WY, Qin J, et al. A virtual reality simulator for ultrasound-guided biopsy training. IEEE Comput Graph Appl. 2011;31(2), 36–48. doi: 10.1109/MCG.2009.151.
[28] Marshall NE, Vanderhoeven J, Eden KB, et al. Impact of simulation and team training on postpartum hemorrhage management in non-academic centers. J Matern Fetal Neonatal Med. 2015;28(5):495–499. doi: 10.3109/14767058.2014.923393.
[29] Abulfaraj MM, Jeffers JM, Tackett S, et al. Virtual reality vs. high-fidelity mannequin-based simulation: a pilot randomized trial evaluating learner performance. Cureus. 2021;13(8):e17091. doi: 10.7759/cureus.17091.
[30] Bienstock J, Heuer A. A review on the evolution of simulation-based training to help build a safer future. Medicine. 2022;101(25):e29503. doi: 10.1097/MD.0000000000029503.
[31] Bracq M-S, Michinov E, Le Duff M, et al. Training situational awareness for scrub nurses: error recognition in a virtual operating room. Nurse Educ Pract. 2021;53:103056. doi: 10.1016/j.nepr.2021.103056.
[32] Botezatu M, Hult H, Tessma MK, et al. Virtual patient simulation for learning and assessment: superior results in comparison with regular course exams. Med Teach. 2010;32(10):845–850. doi: 10.3109/0142159100 3695287.
[33] Sobocan M, Turk N, Dinevski D, et al. Problem-based learning in internal medicine: virtual patients or paper-based problems? Intern Med J. 2017;47(1):99–103. doi: 10.1111/imj.13304.
[34] Letterie GS. How virtual reality may enhance training in obstetrics and gynecology. Am J Obstet Gynecol. 2002;187(3 Suppl):S37–S40. doi: 10.1067/mob.2002.127361.
[35] Madsen ME, Konge L, Nørgaard LN, et al. Assessment of performance measures and learning curves for use of a virtual-reality ultrasound simulator in transvaginal ultrasound examination. Ultrasound Obstet Gynecol. 2014;44(6):693–699. doi: 10.1002/uog.13400.
[36] Darras KE, Forster BB, Spouge R, et al. Virtual dissection with clinical radiology cases provides educational value to first year medical students. Acad Radiol. 2020;27(11):1633–1640. doi: 10.1016/j.acra.2019.09.031.

[37] Kang SG, Yang KS, Ko YH, et al. A study on the learning curve of the robotic virtual reality simulator. J Laparoendosc Adv Surg Tech A. 2012;22(5):438–442. doi: 10.1089/lap.2011.0452.
[38] Kulcsár Z, O'Mahony E, Lövquist E, et al. Preliminary evaluation of a virtual reality-based simulator for learning spinal anesthesia. J Clin Anesth. 2013;25(2):98–105. doi: 10.1016/j.jclinane.2012.06.015.
[39] Maagaard M, Sorensen JL, Oestergaard J, et al. Retention of laparoscopic procedural skills acquired on a virtual-reality surgical trainer. Surg Endosc. 2011;25(3):722–727. doi: 10.1007/s00464-010-1233-5.
[40] McGrath J, Kman N, Danforth D, et al. Virtual alternative to the oral examination for emergency medicine residents. West J Emerg Med. 2015;16(2):336–343. doi: 10.5811/westjem.2015.1.24344.
[41] Ferri P, Rovesti S, Magnani D, et al. The efficacy of interprofessional simulation in improving collaborative attitude between nursing students and residents in medicine. A study protocol for a randomised controlled trial. Acta Biomed. 2018;89(7-S):32–40. doi: 10.23750/abm.v89i7-S.7875.
[42] Gartmeier M, Bauer J, Fischer MR, et al. Fostering professional communication skills of future physicians and teachers: effects of e-learning with video cases and role-play. Instr Sci. 2015;43(4):443–462. doi: 10.1007/s11251-014-9341-6.
[43] Kerfoot BP, Fu Y, Baker H, et al. Online spaced education generates transfer and improves long-term retention of diagnostic skills: a randomized controlled trial. J Am Coll Surg. 2010;211(3):331–337.e1. doi: 10.1016/j.jamcollsurg.2010.04.023.
[44] Burnett G, Goldberg A, DeMaria SJr., et al. Knowledge retention after simulated crisis: importance of independent practice and simulated mortality. Br J Anaesth. 2019;123(1):81–87. doi: 10.1016/j.bja.2019.02.030.
[45] Eich C, Timmermann A, Russo SG, et al. Simulator-based training in paediatric anaesthesia and emergency medicine – thrills, skills and attitudes. Br J Anaesth. 2007;98(4):417–419. doi: 10.1093/bja/aem051.
[46] Satin AJ. Simulation in obstetrics. Obstet Gynecol. 2018;132(1):199–209. doi: 10.1097/aog.0000000000002682.
[47] Friedrich M, Bergdolt C, Haubruck P, et al. App-based serious gaming for training of chest tube insertion: study protocol for a randomized controlled trial. Trials. 2017;18(1):56. doi: 10.1186/s13063-017-1799-5.
[48] Kanthan R, Senger J-L. The impact of specially designed digital games-based learning in undergraduate pathology and medical education. Arch Pathol Lab Med. 2011;135(1):135–142. doi: 10.5858/2009-0698-OAR1.1.
[49] LeRoy Heinrichs W, Youngblood P, Harter PM, et al. Simulation for team training and assessment: case studies of online training with virtual worlds. World J Surg. 2008;32(2):161–170. doi: 10.1007/s00268-007-9354-2.
[50] Dearnley C, Haigh J, Fairhall J. Using mobile technologies for assessment and learning in practice settings: a case study. Nurse Educ Pract. 2008;8(3):197–204. doi: 10.1016/j.nepr.2007.07.003.
[51] Golenhofen N, Heindl F, Grab-Kroll C, et al. The use of a mobile learning tool by medical students in undergraduate anatomy and its effects on assessment outcomes. Anat Sci Educ. 2020;13(1):8–18. doi: 10.1002/ase.1878.
[52] Luciano CJ, Banerjee PP, Sorenson JM, et al. Percutaneous spinal fixation simulation with virtual reality and haptics. Neurosurgery. 2013;72 Suppl 1(Supplement 1):89–96. doi: 10.1227/NEU.0b013e3182750a8d.
[53] Vitish-Sharma P, Knowles J, Patel B. Acquisition of fundamental laparoscopic skills: is a box really as good as a virtual reality trainer? Int J Surg. 2011;9(8):659–661. doi: 10.1016/j.ijsu.2011.08.009.
[54] Happel CS, Lease MA, Nishisaki A, et al. Evaluating simulation education via electronic surveys immediately following live critical events: a pilot study. Hosp Pediatr. 2015;5(2):96–100. doi: 10.1542/hpeds.2014-0091.
[55] Reznek MA, Rawn CL, Krummel TM. Evaluation of the educational effectiveness of a virtual reality intravenous insertion simulator. Acad Emerg Med. 2002;9(11):1319–1325. doi: 10.1197/aemj.9.11.1319.
[56] Snyder CW, Vandromme MJ, Tyra SL, et al. Retention of colonoscopy skills after virtual reality simulator training by independent and proctored methods. Am Surg. 2010;76(7):743–746. doi: 10.1177/000313481007600732.
[57] Lu C, Ghoman SK, Cutumisu M, et al. Unsupervised machine learning algorithms examine healthcare providers. Perceptions and longitudinal performance in a digital neonatal resuscitation simulator. Front Pediatr. 2020;8:544. doi: 10.3389/fped.2020.00544.
[58] Solyar A, Cuellar H, Sadoughi B, et al. Endoscopic sinus surgery simulator as a teaching tool for anatomy education. Am J Surg. 2008;196(1):120–124. doi: 10.1016/j.amjsurg.2007.06.026.
[59] Guetterman TC, Kron FW, Campbell TC, et al. Initial construct validity evidence of a virtual human application for competency assessment in breaking bad news to a cancer patient. Adv Med Educ Pract. 2017;8:505–512. doi: 10.2147/AMEP.S138380.
[60] Pottle J. Virtual reality and the transformation of medical education. Future Healthc J. 2019;6(3):181–185. doi: 10.7861/fhj.2019-0036.
[61] Alaker M, Wynn GR, Arulampalam T. Virtual reality training in laparoscopic surgery: a systematic review & meta-analysis. Int J Surg. 2016;29:85–94. doi: 10.1016/j.ijsu.2016.03.034.
[62] Barsom EZ, Graafland M, Schijven MP. Systematic review on the effectiveness of augmented reality applications in medical training. Surg Endosc. 2016;30(10):4174–4183. doi: 10.1007/s00464-016-4800-6.
[63] Nash R, Sykes R, Majithia A, et al. Objective assessment of learning curves for the Voxel-Man TempoSurg temporal bone surgery computer simulator. J Laryngol Otol. 2012;126(7):663–669. doi: 10.1017/S0022215112000734.
[64] Andersen SAW, Mikkelsen PT, Sørensen MS. Expert sampling of VR simulator metrics for automated assessment of mastoidectomy performance. Laryngoscope. 2019;129(9):2170–2177. doi: 10.1002/lary.27798.

[65] Overtoom EM, Horeman T, Jansen F-W, et al. Haptic feedback, force feedback, and force-sensing in simulation training for laparoscopy: a systematic overview. J Surg Educ. 2019;76(1):242–261. doi: 10.1016/j.jsurg.2018.06.008.
[66] Loukas C, Nikiteas N, Kanakis M, et al. Evaluating the effectiveness of virtual reality simulation training in intravenous cannulation. Simul Healthc. 2011;6(4):213–217. doi: 10.1097/SIH.0b013e31821d08a9.
[67] Yudkowsky R, Luciano C, Banerjee P, et al. Practice on an augmented reality/haptic simulator and library of virtual brains improves residents' ability to perform a ventriculostomy. Simul Healthc. 2013;8(1):25–31. doi: 10.1097/SIH.0b013e3182662c69.
[68] Andersen SAW, Konge L, Sørensen MS. The effect of distributed virtual reality simulation training on cognitive load during subsequent dissection training. Med Teach. 2018;40(7):684–689. doi: 10.1080/0142159x.2018.1465182.
[69] Raupach T, Muenscher C, Anders S, et al. Web-based collaborative training of clinical reasoning: a randomized trial. Med Teach. 2009;31(9):e431–e437. doi: 10.1080/01421590903095502.
[70] Taradi SK, Taradi M. Expanding the traditional physiology class with asynchronous online discussions and collaborative projects. Adv Physiol Educ. 2004;28(1–4):73–78. doi: 10.1152/advan.00017.2003.
[71] Taradi SK, Taradi M, Radic K, et al. Blending problem-based learning with web technology positively impacts student learning outcomes in acid-base physiology. Adv Physiol Educ. 2005;29(1):35–39. doi: 10.1152/advan.00026.2004.
[72] Ruiz JG, Mintzer MJ, Leipzig RM. The impact of E-learning in medical education. Acad Med. 2006;81(3):207–212. doi: 10.1097/00001888-200603000-00002.
[73] Cook DA, Andriole DA, Durning SJ, et al. Longitudinal research databases in medical education: facilitating the study of educational outcomes over time and across institutions. Acad Med. 2010;85(8):1340–1346. doi: 10.1097/ACM.0b013e3181e5c050.
[74] Vallée A, Blacher J, Cariou A, et al. Blended learning compared to traditional learning in medical education: systematic review and meta-analysis. J Med Internet Res. 2020;22(8):e16504. doi: 10.2196/16504.
[75] Eaton M, Scully R, Schuller M, et al. Value and barriers to use of the SIMPL tool for resident feedback. J Surg Educ. 2019;76(3):620–627. doi: 10.1016/j.jsurg.2019.01.012.
[76] Mirchi N, Bissonnette V, Ledwos N, et al. Artificial neural networks to assess virtual reality anterior cervical discectomy performance. Oper Neurosurg. 2020;19(1):65–75. doi: 10.1093/ons/opz359.
[77] Smith CC, Huang GC, Newman LR, et al. Simulation training and its effect on long-term resident performance in central venous catheterization. Simul Healthc. 2010;5(3):146–151. doi: 10.1097/SIH.0b013e3181dd9672.
[78] Koch AD, Ekkelenkamp VE, Haringsma J, et al. Simulated colonoscopy training leads to improved performance during patient-based assessment. Gastrointest Endosc. 2015;81(3):630–636. doi: 10.1016/j.gie.2014.09.014.
[79] Barsuk JH, Ahya SN, Cohen ER, et al. Mastery learning of temporary hemodialysis catheter insertion by nephrology fellows using simulation technology and deliberate practice. Am J Kidney Dis. 2009;54(1):70–76. doi: 10.1053/j.ajkd.2008.12.041.
[80] Barsuk JH, McGaghie WC, Cohen ER, et al. Simulation-based mastery learning reduces complications during central venous catheter insertion in a medical intensive care unit. Crit Care Med. 2009;37(10):2697–2701. doi: 10.1097/CCM.0b013e3181a57bc1.
[81] Giuliani M, Gillan C, Wong O, et al. Evaluation of high-fidelity simulation training in radiation oncology using an outcomes logic model. Radiat Oncol. 2014;9(1):189. doi: 10.1186/1748-717X-9-189.
[82] Battaglia JN, Kieser MA, Bruskiewitz RH, et al. An online virtual-patient program to teach pharmacists and pharmacy students how to provide diabetes-specific medication therapy management. Am J Pharm Educ. 2012;76(7):131. doi: 10.5688/ajpe767131.
[83] Vertemati M, Cassin S, Rizzetto F, et al. A virtual reality environment to visualize three-dimensional patient-specific models by a mobile head-mounted display. Surg Innov. 2019;26(3):359–370. doi: 10.1177/1553350618822860.
[84] Chheang V, Fischer V, Buggenhagen H, et al. Toward interprofessional team training for surgeons and anesthesiologists using virtual reality. Int J Comput Assist Radiol Surg. 2020;15(12):2109–2118. doi: 10.1007/s11548-020-02276-y.
[85] Siemens G. Learning analytics. Am Behav Sci. 2013;57(10):1380–1400. doi: 10.1177/0002764213498851.
[86] Goh P-S, Sandars J. A vision of the use of technology in medical education after the COVID-19 pandemic. MedEdPublish. 2020;9(1):49. doi: 10.15694/mep.2020.000049.1.
[87] Stephenson CR, Bonnes SL, Sawatsky AP, et al. The relationship between learner engagement and teaching effectiveness: a novel assessment of student engagement in continuing medical education. BMC Med Educ. 2020;20(1):403. doi: 10.1186/s12909-020-02331-x.
[88] Davoudi M, Osann K, Colt HG. Validation of two instruments to assess technical bronchoscopic skill using virtual reality simulation. Respiration. 2008;76(1):92–101. doi: 10.1159/000126493.
[89] Gierl MJ, Latifi S, Lai H, et al. Automated essay scoring and the future of educational assessment in medical education. Med Educ. 2014;48(10):950–962. doi: 10.1111/medu.12517.
[90] Hulsman RL, van der Vloodt J. Self-evaluation and peer-feedback of medical students' communication skills using a web-based video annotation system. Exploring content and specificity. Patient Educ Couns. 2015;98(3):356–363. doi: 10.1016/j.pec.2014.11.007.
[91] Kolbe M, Boos M. Laborious but elaborate: the benefits of really studying team dynamics. Front Psychol. 2019;10:1478. doi: 10.3389/fpsyg.2019.01478.
[92] Chan T, Sebok-Syer S, Thoma B, et al. Learning analytics in medical education assessment: the past, the present, and the future. AEM Educ Train. 2018;2(2):178–187. doi: 10.1002/aet2.10087.
[93] Guetterman TC, Sakakibara R, Baireddy S, et al. Medical students' experiences and outcomes using a virtual hu-

man simulation to improve communication skills: mixed methods study. J Med Internet Res. 2019;21(11):e15459. doi: 10.2196/15459.
[94] Bissonnette V, Mirchi N, Ledwos N, et al. Artificial intelligence distinguishes surgical training levels in a virtual reality spinal task. J Bone Joint Surg Am. 2019;101(23):e127. doi: 10.2106/JBJS.18.01197.
[95] Tang KS, Cheng DL, Mi E, et al. Augmented reality in medical education: a systematic review. Can Med Educ J. 2020;11(1):e81–e96. doi: 10.36834/cmej.61705.
[96] Tene T, Vique López DF, Valverde Aguirre PE, et al. Virtual reality and augmented reality in medical education: an umbrella review. Front Digit Health. 2024;6:1365345. doi: 10.3389/fdgth.2024.1365345.
[97] Tursø-Finnich T, Jensen RO, Jensen LX, et al. Virtual reality head-mounted displays in medical education. Simul Healthc. 2023;18(1):42–50. doi: 10.1097/SIH.0000000000000636.
[98] McBain KA, Habib R, Laggis G, et al. Scoping review: the use of augmented reality in clinical anatomical education and its assessment tools. Anat Sci Educ. 2022;15(4):765–796. doi: 10.1002/ase.2155.
[99] Hill J, Hamer O, Breed H, et al. The range of uses of virtual reality for intensive care unit staff training: a narrative synthesis scoping review. Comput Assist Learn. 2023;39(3):869–882. doi: 10.1111/jcal.12787.
[100] Yi WS, Rouhi AD, Duffy CC, et al. A systematic review of immersive virtual reality for nontechnical skills training in surgery. J Surg Educ. 2024;81(1):25–36. doi: 10.1016/j.jsurg.2023.11.012.
[101] Trevi R, Chiappinotto S, Palese A, et al. Virtual reality for cardiopulmonary resuscitation healthcare professionals training: a systematic review. J Med Syst. 2024;48(1):50. doi: 10.1007/s10916-024-02063-1.
[102] Bracq M-S, Michinov E, Jannin P. Virtual reality simulation in nontechnical skills training for healthcare professionals: a systematic review. Simul Healthc. 2019;14(3):188–194. doi: 10.1097/SIH.0000000000000347.
[103] Lewis KO, Cidon MJ, Seto TL, et al. Leveraging e-learning in medical education. Curr Probl Pediatr Adolesc Health Care. 2014;44(6):150–163. doi: 10.1016/j.cppeds.2014.01.004.
[104] Adamkiewicz D, Atri L, Berman L, et al. Implementation of a telemedicine student clinical experience. Telemed J E Health. 2023;29(3):432–441. doi: 10.1089/tmj.2022.0127.
[105] Pérez Alonso N, Pardo Rios M, Juguera Rodriguez L, et al. Randomised clinical simulation designed to evaluate the effect of telemedicine using Google Glass on cardiopulmonary resuscitation (CPR). Emerg Med J. 2017;34(11):734–738. doi: 10.1136/emermed-2016-205998.
[106] Dhar P, Rocks T, Samarasinghe RM, et al. Augmented reality in medical education: students' experiences and learning outcomes. Med Educ Online. 2021;26(1):1953953. doi: 10.1080/10872981.2021.1953953.
[107] Richters C, Stadler M, Radkowitsch A, et al. Who is on the right track? Behavior-based prediction of diagnostic success in a collaborative diagnostic reasoning simulation. Large-Scale Assess Educ. 2023;11(1):3. doi: 10.1186/s40536-023-00151-1.
[108] Brown KE, Heise N, Eitel CM, et al. A large-scale, multiplayer virtual reality deployment: a novel approach to distance education in human anatomy. Med Sci Educ. 2023;33(2):409–421. doi: 10.1007/s40670-023-01751-w.
[109] Meeks LM, Herzer K, Jain NR. Removing barriers and facilitating access: increasing the number of physicians with disabilities. Acad Med. 2018;93(4):540–543. doi: 10.1097/acm.0000000000002112.
[110] Chernikova O, Heitzmann N, Stadler M, et al. Simulation-based learning in higher education: a meta-analysis. Rev Educ Res. 2020;90(4):499–541. doi: 10.3102/0034654320933544.
[111] Liu Q, Peng W, Zhang F, et al. The effectiveness of blended learning in health professions: systematic review and meta-analysis. J Med Internet Res. 2016;18(1):e2. doi: 10.2196/jmir.4807.
[112] Wolbrink TA, Burns JP. Internet-based learning and applications for critical care medicine. J Intensive Care Med. 2012;27(5):322–332. doi: 10.1177/0885066611429539.
[113] Harrington RA, Califf RM, Balamurugan A, et al. Call to action: rural health: a presidential advisory from the American Heart Association and American stroke association. Circulation. 2020;141(10):e615–e644. doi: 10.1161/CIR.0000000000000753.
[114] Statista. (June 28, 2024). Virtual reality (VR) headset average price in the United States from 2019 to 2029 (in U.S. dollars) [Graph]. In Statista. Retrieved November 04, 2024, from https://www.statista.com/forecasts/1338404/vr-headset-average-price-united-states
[115] Zaidi SSB, Adnan U, Lewis KO, et al. Metaverse-powered basic sciences medical education: bridging the gaps for lower middle-income countries. Ann Med. 2024;56(1):2356637. doi: 10.1080/07853890.2024.2356637.
[116] Bankar MN, Bankar NJ, Singh BR, et al. The role of e-content development in medical teaching: how far have we come? Cureus. 2023;15(8):e43208. doi: 10.7759/cureus.43208.
[117] Barsuk JH, Cohen ER, Potts S, et al. Dissemination of a simulation-based mastery learning intervention reduces central line-associated bloodstream infections. BMJ Qual Saf. 2014;23(9):749–756. doi: 10.1136/bmjqs-2013-002665.
[118] Timbi-Sisalima C, Sánchez-Gordón M, Hilera-Gonzalez JR, et al. Quality assurance in e-learning: a proposal from accessibility to sustainability. Sustainability. 2022;14(5):3052. doi: 10.3390/su14053052.
[119] Lewis KO, Hathaway SB, Bratcher D, et al. Current milestones assessment practices, needs, and challenges of program directors: a collective case study in a pediatric hospital setting. Cureus. 2021;13(4):e14585. doi: 10.7759/cureus.14585.
[120] Harriman D, Singla R, Nguan C. The resident report card: a tool for operative feedback and evaluation of technical skills. J Surg Res. 2019;239:261–268. doi: 10.1016/j.jss.2019.02.006.
[121] Bogar PZ, Virag M, Bene M, et al. Validation of a novel, low-fidelity virtual reality simulator and an artificial intelligence assessment approach for peg transfer laparoscopic training. Sci Rep. 2024;14(1):16702. doi: 10.1038/s41598-024-67435-6.
[122] Shaban S, Tariq I, Elzubeir M, et al. Conducting online OSCEs aided by a novel time management web-based system. BMC Med Educ. 2021;21(1):508. doi: 10.1186/s12909-021-02945-9.
[123] Shorbagi S, Sulaiman N, Hasswan A, et al. Assessing the utility and efficacy of e-OSCE among undergraduate medical students during the COVID-19 pandemic. BMC

Med Educ. 2022;22(1):156. doi: 10.1186/s12909-022-03218-9.
[124] Bölek KA, De Jong G, Henssen D. The effectiveness of the use of augmented reality in anatomy education: a systematic review and meta-analysis. Sci Rep. 2021;11(1):15292. doi: 10.1038/s41598-021-94721-4.
[125] Våpenstad C, Fagertun Hofstad E, Eivind Bernstein T, et al. Optimal timing of assessment tasks depending on experience level of surgical trainees. Minim Invasive Ther Allied Technol. 2020;29(3):161–169. doi: 10.1080/13645706.2019.1612441.
[126] Ashraf A, Collins D, Whelan M, et al. Three-dimensional (3D) simulation versus two-dimensional (2D) enhances surgical skills acquisition in standardised laparoscopic tasks: a before and after study. Int J Surg. 2015;14:12–16. doi: 10.1016/j.ijsu.2014.12.020.
[127] Fida B, Cutolo F, di Franco G, et al. Augmented reality in open surgery. Updates Surg. 2018;70(3):389–400. doi: 10.1007/s13304-018-0567-8.
[128] Volonté F, Pugin F, Bucher P, et al. Augmented reality and image overlay navigation with OsiriX in laparoscopic and robotic surgery: not only a matter of fashion. J Hepatobiliary Pancreat Sci. 2011;18(4):506–509. doi: 10.1007/s00534-011-0385-6.
[129] Popov V, Ruis AR, Cooke JM. Taking stock and looking ahead: evolution of accreditation feedback for simulation centers over 8 years using epistemic network analysis. Simul Healthc. 2023;18(1):1–7. doi: 10.1097/SIH.0000000000000638.
[130] Hassani H, Huang X, MacFeely S. Impactful digital twin in the healthcare revolution. BDCC. 2022;6(3):83. doi: 10.3390/bdcc6030083.
[131] Birkheim SL, Calogiuri G, Martinsen R. Advancing immersive virtual reality-based simulation practices: developing an evidence-based and theory-driven pedagogical framework for VR-based simulations of non-technical skills among healthcare professionals. Interact Learn Environ. 2023:3579–3591. doi: 10.1080/10494820.2023.2186896.
[132] Stanney KM, Skinner A, Hughes C. Exercisable learning-theory and evidence-based andragogy for training effectiveness using XR (ELEVATE-XR): elevating the ROI of immersive technologies. Int J Hum Comput Interact. 2023;39(11):2177–2198. doi: 10.1080/10447318.2023.2188529.
[133] Clusmann J, Kolbinger FR, Muti HS, et al. The future landscape of large language models in medicine. Commun Med. 2023;3(1):141. doi: 10.1038/s43856-023-00370-1.
[134] Abd-Alrazaq A, AlSaad R, Alhuwail D, et al. Large language models in medical education: opportunities, challenges, and future directions. JMIR Med Educ. 2023;9:e48291. doi: 10.2196/48291.
[135] Fiok K, Farahani FV, Karwowski W, et al. Explainable artificial intelligence for education and training. J Defense Model Simul. 2022;19(2):133–144. doi: 10.1177/15485129211028651.
[136] Arora VM, Carter K, Babcock C. Bias in assessment needs urgent attention—no rest for the "wicked". JAMA Netw Open. 2022;5(11):e2243143. doi: 10.1001/jamanetworkopen.2022.43143.
[137] Boutin J, Kamoonpuri J, Faieghi R, et al. Smart haptic gloves for virtual reality surgery simulation: a pilot study on external ventricular drain training. Front Robot AI. 2023;10:1273631. doi: 10.3389/frobt.2023.1273631.
[138] Lerner D, Mohr S, Schild J, et al. An immersive multi-user virtual reality for emergency simulation training: usability study. JMIR Serious Games. 2020;8(3):e18822. doi: 10.2196/18822.
[139] Aebersold M, Dunbar D-M. Virtual and augmented realities in nursing education: state of the science. Annu Rev Nurs Res. 2020;39(1):225–242. doi: 10.1891/0739-6686.39.225.
[140] Chytas D, Malahias M-A, Nikolaou VS. Augmented reality in orthopedics: current state and future directions. Front Surg. 2019;6:38. doi: 10.3389/fsurg.2019.00038.
[141] Fisher J, Gordon A, Pattinson J, et al. S-11: innovating for a reason: using new educational technologies to improve learning in geriatric medicine. Eur Geriatr Med. 2015;6:S166–S168. doi: 10.1016/s1878-7649(15)30568-4.
[142] Moro C, Štromberga Z, Raikos A, et al. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. Anat Sci Educ. 2017;10(6):549–559. doi: 10.1002/ase.1696.
[143] Morgan PJ, Pittini R, Regehr G, et al. Evaluating teamwork in a simulated obstetric environment. Anesthesiology. 2007;106(5):907–915. doi: 10.1097/01.anes.0000265149.94190.04.
[144] Rochlen LR, Putnam E, Levine R, et al. Mixed reality simulation for peripheral intravenous catheter placement training. BMC Med Educ. 2022;22(1):876. doi: 10.1186/s12909-022-03946-y.
[145] Probst H, Eddy D, Doughty J, et al. Integrating E-learning into postgraduate radiotherapy and oncology education: a case study. E-Learn Digit Media. 2009;6(4):363–371. doi: 10.2304/elea.2009.6.4.363.
[146] Lau KHV, Greer DM. Using technology adoption theories to maximize the uptake of e-learning in medical education. Med Sci Educ. 2022;32(2):545–552. doi: 10.1007/s40670-022-01528-7.
[147] Lewis KO, Popov V, Fatima SS. From static web to metaverse: reinventing medical education in the post-pandemic era. Ann Med. 2024;56(1):2305694. doi: 10.1080/07853890.2024.2305694.
[148] Hoffman H, Vu D. Virtual reality. Acad Med. 1997;72(12):1076–1081. doi: 10.1097/00001888-199712000-00018.
[149] Rogers EM. Diffusion of innovations: modifications of a model for telecommunications. In: Die Diffusion von Innovationen in der Telekommunikation. Berlin; Heidelberg: Springer Berlin Heidelberg; 1995. p. 25–38. doi: 10.1007/978-3-642-79868-9_2.
[150] Murray E, Treweek S, Pope C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med. 2010;8(1):63. doi: 10.1186/1741-7015-8-63.
[151] O'Leary DE. Gartner's hype cycle and information system research issues. Int J Account Inform Syst. 2008;9(4):240–252. doi: 10.1016/j.accinf.2008.09.001.
[152] Makransky G, Petersen GB. The cognitive affective model of immersive learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. Educ Psychol Rev. 2021;33(3):937–958. doi: 10.1007/s10648-020-09586-2.
[153] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. 1989;13(3):319. doi: 10.2307/249008.