

Original Paper

Effectiveness of Virtual Simulations Versus Mannequins and Real Persons in Medical and Nursing Education: Meta-Analysis and Trial Sequential Analysis of Randomized Controlled Trials

Nan Jiang^{1*}, MD; Yuelun Zhang^{2*}, PhD; Siyu Liang³, MD; Xiaohong Lyu⁴, MD; Shi Chen³, MD; Xiaoming Huang⁵, MD; Hui Pan³, MD

¹4 + 4 Medical Doctor Program, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

²Medical Research Center, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

³Department of Endocrinology, Key Laboratory of Endocrinology of National Health Commission, Translation Medicine Centre, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China

⁴Department of Breast Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

⁵Department of Family Medicine & Division of General Internal Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

*these authors contributed equally

Corresponding Author:

Shi Chen, MD

Department of Endocrinology

Key Laboratory of Endocrinology of National Health Commission

Translation Medicine Centre, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences

No.1 Shuaifuyuan, Wangfujing, Dongcheng District

Beijing, 100730

China

Phone: 86 13683136205

Email: cspunch@163.com

Abstract

Background: Virtual simulation (VS) is a developing education approach with the recreation of reality using digital technology. The teaching effectiveness of VSs compared to mannequins and real persons (RPs) has never been investigated in medical and nursing education.

Objective: This study aims to compare VSs and mannequins or RPs in improving the following clinical competencies: knowledge, procedural skills, clinical reasoning, and communication skills.

Methods: Following Cochrane methodology, a meta-analysis was conducted on the effectiveness of VSs in pre- and postregistration medical or nursing participants. The Cochrane Library, PubMed, Embase, and Educational Resource Information Centre databases were searched to identify English-written randomized controlled trials up to August 2024. Two authors independently selected studies, extracted data, and assessed the risk of bias. All pooled estimates were based on random-effects models and assessed by trial sequential analyses. Leave-one-out, subgroup, and univariate meta-regression analyses were performed to explore sources of heterogeneity.

Results: A total of 27 studies with 1480 participants were included. Overall, there were no significant differences between VSs and mannequins or RPs in improving knowledge (standard mean difference [SMD]=0.08; 95% CI -0.30 to 0.47; $I^2=67%$; $P=.002$), procedural skills (SMD=-0.12; 95% CI -0.47 to 0.23; $I^2=75%$; $P<.001$), clinical reasoning (SMD=0.29; 95% CI -0.26 to 0.85; $I^2=88%$; $P<.001$), and communication skills (SMD=-0.02; 95% CI: -0.62 to 0.58; $I^2=86%$; $P<.001$). Trial sequential analysis for clinical reasoning indicated an insufficient sample size for a definitive judgment. For procedural skills, subgroup analyses

showed that VSs were less effective among nursing participants (SMD=-0.55; 95% CI -1.07 to -0.03; $I^2=69%$; $P=.04$). Univariate meta-regression detected a positive effect of publication year ($\beta=.09$; $P=.02$) on communication skill scores.

Conclusions: Given favorable cost-utility plus high flexibility regarding time and space, VSs are viable alternatives to traditional face-to-face learning modalities. The comparative effectiveness of VSs deserves to be followed up with the emergence of new technology. In addition, further investigation of VSs with different design features will provide novel insights to drive education reform.

Trial Registration: PROSPERO CRD42023466622; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=466622

(*J Med Internet Res* 2024;26:e56195) doi: [10.2196/56195](https://doi.org/10.2196/56195)

KEYWORDS

artificial intelligence; clinical virtual simulation; medical education; meta-analysis; nursing education; virtual patient; virtual reality

Introduction

The ultimate goal of health profession education is to promote the transfer of theoretical knowledge into clinical practice. Studies in cognitive psychology suggest that recall and application of information are best in learning environments similar to workplace [1]. Clinical simulation, a technique that replicates real experiences with guided experiences, has thereby gained popularity in the last 2 decades and is currently the main tool for training health care professionals [2]. There are five main categories of simulation: (1) low-tech simulators (ie, models or mannequins); (2) standardized patients (SPs); (3) screen-based computer programs; (4) high-fidelity computer simulators integrated with visual, audio, and touch cues; and (5) computer-driven, full-length mannequins with realistic anatomy and physiology [3]. With the help of simulation, health care professionals refine the knowledge, skills, and attitudes needed to deliver quality patient care while patients are protected from unnecessary risk [3]. In medical and nursing education, simulation-based teaching has demonstrated superior effectiveness compared to didactic teaching [4,5]; moreover, it reduces anxiety and increases the confidence of students entering practice [6,7]. While simulation-based education may bring downstream benefits of reducing medical errors and improving patient care, the current level of evidence is still low due to a lack of high-quality studies with patient-centered outcomes [8-12].

Driven by COVID-19 social distancing and modern technological advancement, there has been a marked shift toward learning on digital platforms or software. Virtual simulation (VS) refers to the 2D or 3D recreation of reality by digital technology, where students are able to interact with digital scenes, instruments, and characters [13,14]. Compared to conventional simulations (eg, mannequins and SPs), VSs can be undertaken flexibly with no limit of time and space [15]; they can provide a more realistic experience with the aid of artificial intelligence (AI) and virtual reality (VR) when certain pathological findings cannot be readily expressed [16-18]. For educators, VSs can potentially save costs due to reduced

instructor time, manpower, and space resources [19]. One recent analysis found that VR simulations required 22% less time and 40% lower cost than traditional high-fidelity simulations to achieve the same learning outcomes [20]. Documentation and evaluation of student performance can also be automated with digital technologies [19].

Despite these promising advantages, the comparative effectiveness of training health professionals using mannequins or real persons (RPs; eg, SPs, role-play, actual patients) and VSs remains uncertain. Previous meta-analyses have shown that compared to traditional education (eg, lectures, reading exercises, in-class group discussions, mannequins, SPs), digital education is at least as effective in training medical students' communication skills [15]; virtual patients (VPs) can more effectively improve skills and can be at least as effective in enhancing knowledge [21]; VR can more effectively improve not only skills but also knowledge [22]. The comparison with an extensive category of "traditional education" is insufficient to determine the value and viability of implementing VS as substitutes for mannequins and RPs, so specific recommendations for educational reform cannot be made. Thus, this study aims to investigate the teaching efficacy of VSs versus mannequins and RPs in a fine-grain spectrum of clinical competencies, including knowledge, procedural skills, clinical reasoning, and communication skills.

Methods

Reporting Guidelines

This study follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) guidelines [23]. [Multimedia Appendix 1](#) provides a PRISMA checklist of this meta-analysis. The protocol was registered in PROSPERO (CRD42023466622).

Eligibility Criteria

This study included randomized controlled trials (RCTs) from inception to August 2024. Inclusion and exclusion criteria are listed in [Textbox 1](#).

Textbox 1. Inclusion and exclusion criteria.

<p>Inclusion criteria</p> <ul style="list-style-type: none"> • Year: From inception to August 2024 • Study design: Randomized controlled trials • Language: English • Pre- and post-registration medical/nursing students and staff • Intervention: Virtual simulations • Comparison: Mannequins or real persons • Outcome: Objective posttest ratings for knowledge, procedural skills, clinical reasoning, or communication skills • Data availability: Yes <p>Exclusion criteria</p> <ul style="list-style-type: none"> • Intervention: Blended simulations where virtual simulations and mannequins or real persons were both applied • Outcome: Self-evaluations

The four outcomes of interest were defined as follows: (1) knowledge: remembering and understanding basic concepts, measured by question-based theoretical tests [24]; (2) procedural skills: following a series of steps to accomplish certain tasks, demonstrated by simulation exams [24]; (3) clinical reasoning: evaluating and reacting to clinical situations, demonstrated by simulation examinations [24]; and (4) communication skills: interacting with patients or colleagues verbally and nonverbally, demonstrated by simulation examinations [24].

Search Strategy

A thorough search was carried out in the Cochrane Library, PubMed, Embase, and Educational Resource Information Centre databases by specifying the desired population, intervention, and comparison. Search terms were selected judiciously, with specific terms customized for each database (Multimedia Appendix 2). Additionally, we searched Google and reference lists of the selected studies to retrieve other relevant publications.

Search records were imported into the EndNote library (version X9; Clarivate). After eliminating duplicates, the remaining studies underwent eligibility screening by 2 independent reviewers (NJ and SYL) according to predefined inclusion and exclusion criteria (Textbox 1). The initial screening process involved the assessment of titles and abstracts for relevance. Subsequently, full-text screening was performed, and the rationale for exclusion was documented in the PRISMA flowchart. In cases of discrepancies between the 2 reviewers, a third reviewer (SC) was consulted to reach a consensus. Study authors were contacted for crucial missing information. Finally, all researchers agreed on the final literature to be included in the analysis.

Data Extraction

Two reviewers (NJ and SYL) independently extracted data in a structured form in Microsoft Excel, including author, publication date, country, participants, sample size, type of intervention, type of comparison, and outcomes. If a study had multiple posttest measurements, only the first measurement was

recorded due to concern about the learning effect. If participants were rated by both SPs and independent evaluators, the ratings of the latter were preferred. If an outcome was assessed by multiple rating scales, the scale with the highest intraclass correlation coefficient was chosen. If an outcome was reported as multiple items instead of a total score, the primary item was selected, and if that was impossible, the mean score of all items was calculated. If a study had both mannequin and RP as control groups, each pairwise comparison was included separately while splitting the shared intervention group approximately evenly among the comparisons. The outcomes were recorded as mean values and SDs. For studies that reported outcomes as median values and IQRs or as mean values and CIs, these measures were converted to mean values and SDs using methods described in previous literature [25,26].

Risk of Bias Assessment

Two reviewers (NJ and SYL) independently assessed the risk of bias of the included studies using the Cochrane tool [25]. The following domains were considered: random sequence generation, allocation sequence concealment, blinding of participants or personnel, blinding of outcome assessment, complete outcome data, selective reporting, and other sources of bias. Discrepancies were resolved by discussion with a third author. Studies were not excluded from data extraction or analysis based on bias assessment scores.

Statistical Analysis

Since data were measured using different tools, the mean differences were recalculated into standardized mean differences (SMDs). The results are displayed in forest plots, with pooled SMDs computed using random-effects meta-analysis models. We interpreted the effect size as small (SMD=0.2), moderate (SMD=0.5), or large (SMD=0.8) [25]. Publication bias was assessed by the Egger regression asymmetry test.

Cochran Q (chi-square test) was used to evaluate statistical heterogeneity, with a statistical significance threshold at $P < .10$. I^2 statistic was adopted to quantify the degree of heterogeneity, where I^2 was categorized as unimportant (0%-40%), moderate

(30%-60%), substantial (50%-90%), and considerable (75%-100%) [25]. To investigate possible sources of heterogeneity, the leave-one-out method was first applied. Next, for each outcome, we conducted subgroup analyses by discipline (medicine; nursing), level (undergraduate; nonundergraduate [graduate students or clinical staff]), and comparison (mannequin; RP) to see if $I^2 < 50\%$ in both groups [27]. Univariate random-effects meta-regression analyses were then performed to investigate whether heterogeneity between trials could be attributed to year of publication, age of participants, discipline, level, and comparison. Multivariate meta-regression was not performed due to the limited number of studies and the risk of overfitting.

Trial sequential analysis (TSA) was used to evaluate the strength of evidence and adjusted for potential errors. The analysis set specific values for type 1 and 2 errors (5% and 20%) and used these values to calculate trial sequential monitoring boundaries, futility boundaries, and the required information size (RIS) [28]. The mean difference to generate RIS was set to detect a mean difference of 2.0 between intervention and comparison. The variance was estimated by pooling all included trials. Heterogeneity was adjusted based on the estimated ratio between the variance in the random-effects model and the variance in the fixed-effects model. Finally, a graphical evaluation was used to determine if the cumulative Z curve met defined thresholds.

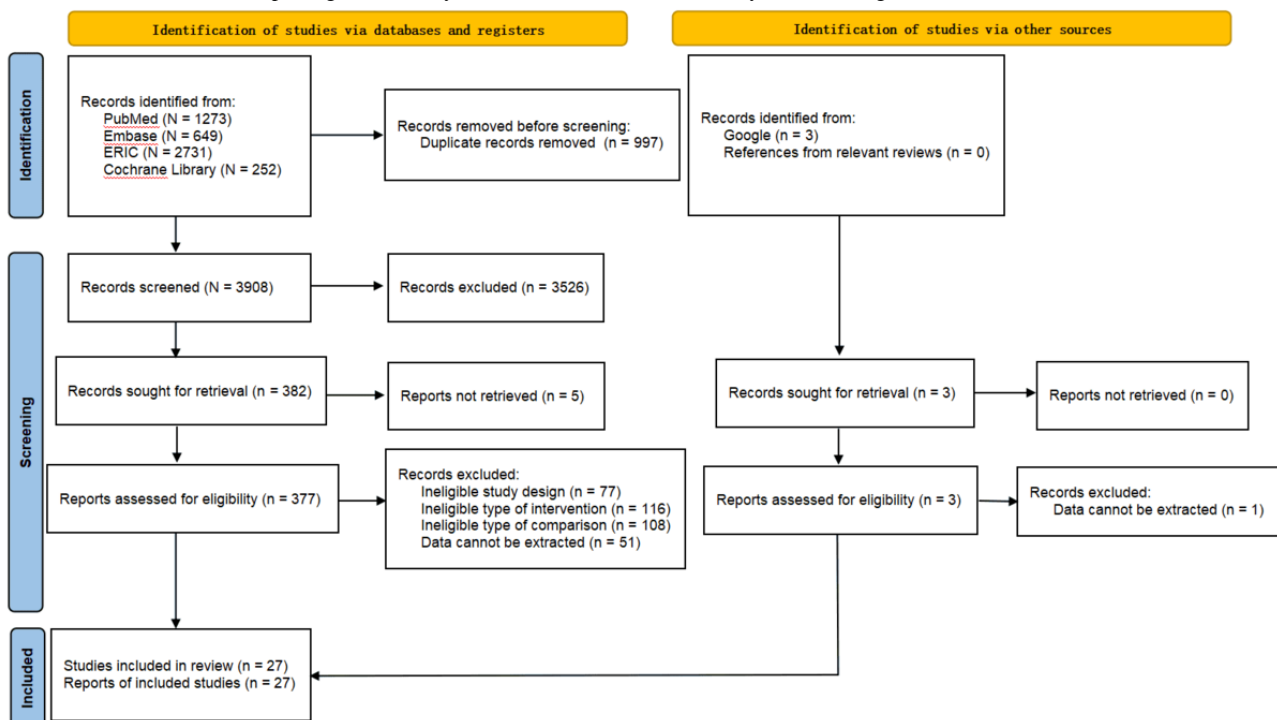
All data analyses were performed using package meta of R software (version 4.2.3; R Foundation for Statistical Computing) [29] and TSA software (version 0.9.5.10 beta; Copenhagen Trial Unit). Two-sided P values $< .05$ were considered to be statistically significant.

Results

Included Studies

After the selection process, a total of 27 studies with 1480 participants were included (Figure 1). In terms of discipline, participants in 14 studies were from the field of medicine [16,18,30-41]; participants in 11 studies were from the field of nursing [42-52]; and the remaining 2 studies enrolled mixed participants from medicine and nursing [53,54]. In terms of level, participants in 14 studies were undergraduate students [16,31,35,39,42-46,50-54]; participants in 8 studies were graduate students or clinical staff [18,32-34,36-38,40,41,49]; and the remaining 2 studies enrolled mixed undergraduate and nonundergraduate participants [30,48]. In terms of comparison, 13 studies compared VSs and mannequins [33-36,38,40,44-51]; 10 studies compared VSs and RPs [16,18,30,31,37,39,41-43,52-54]; and the remaining study included mannequins and RPs as comparison groups simultaneously [32]. Detailed information and extracted data for each study are summarized in Multimedia Appendix 3 [16,18,30-54].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) flow diagram.



Risk of Bias

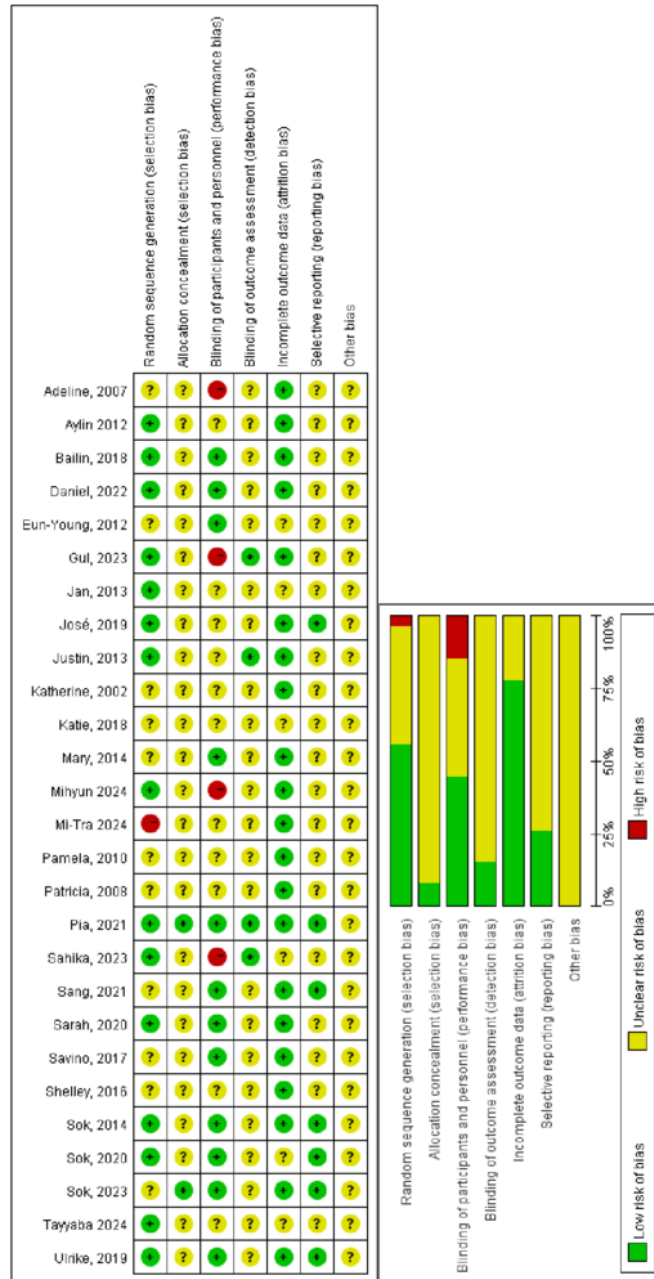
Following Cochrane guidance [25], the risk of bias was assessed for all the included studies. The results are summarized in Figure 2 [16,18,30-54]. Out of 27 studies, 15 studies described an adequate random sequence generation method; 11 studies did not provide a clear description; and the remaining 1 study

allocated participants to intervention and comparison arms based on the order in which they volunteered. Allocation concealment was not explicitly mentioned in most studies except 2. Blinding participants was impossible to avoid in this type of research but should not raise a major concern. Four studies had high-risk performance bias since researchers were not blinded and might give biased ratings; 12 studies were low-risk due to blinding of

personnel; the others did not provide a clear description. The risk of detection bias was rated as unclear in most studies except 4, which clearly stated that statistical analysts were blinded. Most studies had low-risk attrition bias because no participants were lost, while participant dropout in 6 studies had an unclear

impact on outcome assessment. Due to the absence of protocols, most studies had unclear reporting bias; 7 studies with protocols provided were rated as low risk. Given the difficulty of evaluating other biases (eg, volunteer bias), the risk was judged to be unclear for all studies.

Figure 2. Risk of bias summary [16,18,30-54]. +: low risk of bias; -: high risk of bias; ?: unclear risk of bias.



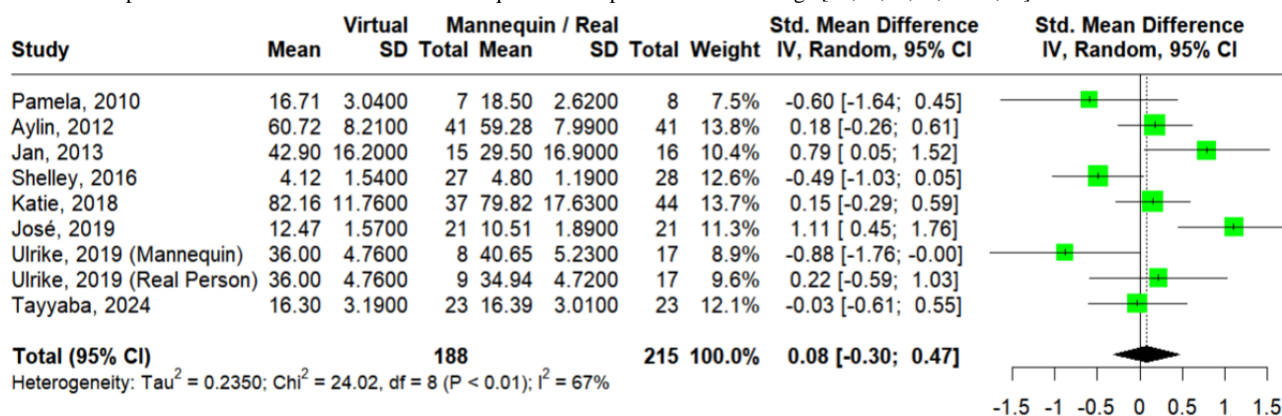
Effects of Interventions

Knowledge

A total of 8 studies assessed the outcome of knowledge [32,35,37,41,44-46,50]. The Egger test showed no statistically significant publication bias (P=.73). In Figure 3 [32,35,37,41,44-46,50], the pooled effect size did not reflect a

significant difference between VSs and mannequins or RPs on the knowledge outcome (SMD=0.08; 95% CI -0.30 to 0.47; I²=67%; P=.002). TSA showed that the “inner wedge” area had been reached, indicating strong evidence that further studies would hardly be able to change the statistically insignificant result (Figure S1 in Multimedia Appendix 4). This result was consistent according to the leave-one-out analysis (Figure S2 in Multimedia Appendix 4 [32,35,37,41,44-46,50]).

Figure 3. Forest plot of virtual simulations versus mannequins or real persons for knowledge [32,35,37,41,44-46,50].



For subgroup analysis by discipline (Figure S3a in Multimedia Appendix 4 [32,35,37,41,44-46,50]), 4 studies [32,35,37,41] and 4 studies [44-46,50] enrolled participants from medicine and nursing, respectively. Neither of these groups showed a significant difference between VSs and mannequins or RPs (SMD=-0.05; 95% CI -0.62 to 0.51; I²=59%; P=.04 for medicine and SMD=0.22; 95% CI -0.39 to 0.82; I²=78%; P<.001 for nursing), with no significant subgroup difference at P=.53.

For subgroup analysis by level (Figure S3b in Multimedia Appendix 4 [32,35,37,41,44-46,50]), 5 studies [35,44-46,50] and 3 studies [32,37,41] enrolled undergraduate and nonundergraduate participants, respectively. Neither of these groups showed a significant difference between VSs and mannequins or RPs (SMD=0.31; 95% CI -0.21 to 0.84; I²=75%; P<.001 for undergraduate and SMD=-0.25; 95% CI -0.71 to 0.22; I²=29%; P=.24 for nonundergraduate), with no significant subgroup difference at P=.12.

For subgroup analysis by comparison (Figure S3c in Multimedia Appendix 4 [32,35,37,41,44-46,50]), 2 studies [37,41] and 5 studies [35,44-46,50] included RP and mannequin arms, respectively; Weber et al [32] compared web-based simulators to both mannequins and RPs. Neither mannequins nor RPs was

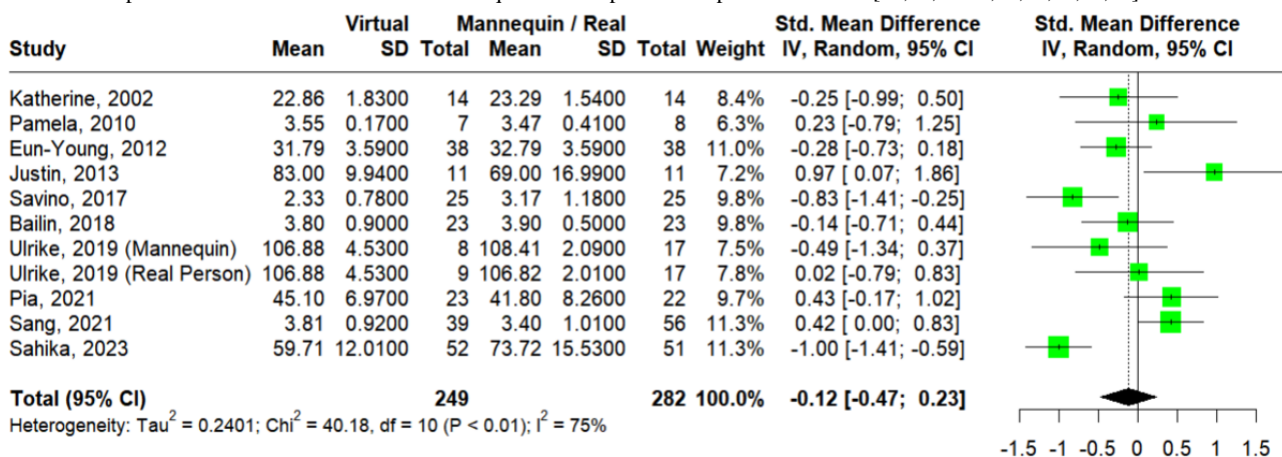
significantly different from VSs (SMD=0.15; 95% CI -0.40 to 0.71; I²=77%; P<.001 for mannequin and SMD=-0.06; 95% CI -0.48 to 0.37; I²=0%; P=.48 for RP), with no significant subgroup difference at P=.56.

Univariate meta-regression revealed that year of publication (β=.01; P=.92), age of participants (β=-.05; P=.29), nursing discipline (β=.27; P=.52), undergraduate level (β=.58; P=.14), and real person comparison (β=-.25; P=.58) had no effects on knowledge scores (Table S1 in Multimedia Appendix 4).

Procedural Skills

A total of 10 studies assessed the outcome of procedural skills [16,18,32-34,36,37,42,49,51]. The Egger test showed no statistically significant publication bias (P=.40). In Figure 4 [16,18,32-34,36,37,42,49,51], the pooled effect size did not reflect a significant difference between VSs and mannequins or RPs on the outcome of procedural skills (SMD=-0.12; 95% CI -0.47 to 0.23; I²=75%; P<.001). TSA showed that the RIS had been reached, implying a conclusive statistically insignificant result (Figure S4 in Multimedia Appendix 4). This result was consistent according to the leave-one-out analysis (Figure S5 in Multimedia Appendix 4 [16,18,32-34,36,37,42,49,51]).

Figure 4. Forest plot of virtual simulations versus mannequins or real persons for procedural skills [16,18,32-34,36,37,42,49,51].



For subgroup analysis by discipline (Figure S6a in Multimedia Appendix 4 [16,18,32-34,37,42,49,51]), 7 studies [16,18,32-34,36,37] and 3 studies [42,49,51] enrolled

participants from medicine and nursing respectively. Only the nursing group showed a significant difference between VSs and mannequins or RPs (SMD=0.06; 95% CI -0.33 to 0.46; I²=64%;

$P < .001$ for medicine and $SMD = -0.55$; 95% CI -1.07 to -0.03 , $I^2 = 69\%$; $P = .04$ for nursing), with no significant subgroup difference at $P = .07$.

For subgroup analysis by level (Figure S6b in [Multimedia Appendix 4](#) [16,18,32-34,36,37,42,49,51]), 3 studies [16,42,51] and 7 studies [18,32-34,36,37,49] enrolled undergraduate and nonundergraduate participants respectively. Neither of these groups showed a significant difference between VSs and mannequins or RPs ($SMD = -0.29$; 95% CI -1.10 to 0.52 ; $I^2 = 91\%$; $P < .001$ for undergraduate and $SMD = -0.04$; 95% CI -0.43 to 0.35 ; $I^2 = 56\%$; $P = .03$ for nonundergraduate), with no significant subgroup difference at $P = .59$.

For subgroup analysis by comparison (Figure S6c in [Multimedia Appendix 4](#) [16,18,32-34,36,37,42,49,51]), 5 studies [33,34,36,49,51] and 4 studies [16,18,37,42] included mannequin and RP arms, respectively; Weber et al [32] compared web-based simulators to both mannequins and RPs. Neither mannequins nor RPs was significantly different from VSs ($SMD = -0.22$; 95% CI -0.63 to 0.19 ; $I^2 = 56\%$; $P = .04$ for mannequins and $SMD = -0.00$; 95% CI -0.60 to 0.59 ; $I^2 = 86\%$;

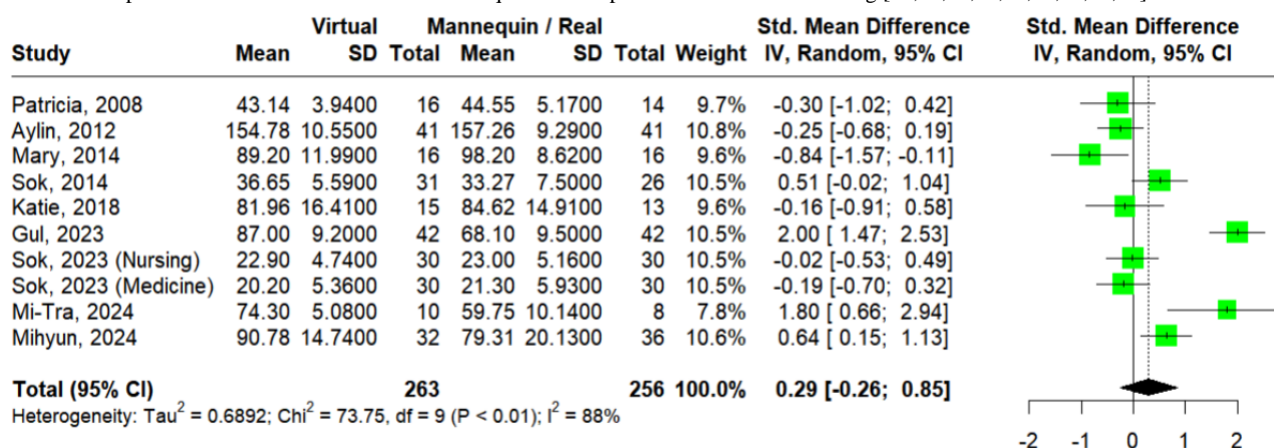
$P < .001$ for RPs), with no significant subgroup difference at $P = .57$.

Univariate meta-regression revealed that year of publication ($\beta = -.01$; $P = .71$), age of participants ($\beta = .05$; $P = .33$), nursing discipline ($\beta = -.60$; $P = .08$), undergraduate level ($\beta = -.25$; $P = .52$), and real person comparison ($\beta = .19$; $P = .61$) had no effects on procedural skill scores (Table S1 in [Multimedia Appendix 4](#)).

Clinical Reasoning

A total of 9 studies assessed the outcome of clinical reasoning [38,40,43,45,47,48,50,52,53]. The Egger test showed no statistically significant publication bias ($P = .88$). In [Figure 5](#) [38,40,43,45,47,48,50,52,53], the pooled effect size did not reflect a significant difference between VSs and mannequins or RPs on the outcome of procedural skills ($SMD = 0.29$; 95% CI -0.26 to 0.85 ; $I^2 = 88\%$; $P < .01$). TSA showed that the RIS had not been reached, implying further studies would be needed to verify the statistically insignificant result (Figure S7 in [Multimedia Appendix 4](#)). This result was consistent according to the leave-one-out analysis (Figure S8 in [Multimedia Appendix 4](#)).

Figure 5. Forest plot of virtual simulations versus mannequins or real persons for clinical reasoning [38,40,43,45,47,48,50,52,53].



For subgroup analysis by discipline (Figure S9a in [Multimedia Appendix 4](#) [38,40,43,45,47,48,50,52,53]), 2 studies [38,40] and 6 studies [43,45,47,48,50,52] enrolled participants from medicine and nursing respectively. Liaw et al [53] enrolled participants from both fields. Neither of these groups showed a significant difference between VSs and mannequins or RPs ($SMD = 0.35$; 95% CI -0.89 to 1.60 ; $I^2 = 81\%$; $P < .001$ for medicine and $SMD = 0.28$; 95% CI -0.39 to 0.95 ; $I^2 = 90\%$; $P < .001$ for nursing), with no significant subgroup difference at $P = .92$.

For subgroup analysis by level (Figure S9b in [Multimedia Appendix 4](#) [38,40,43,45,47,48,50,52,53]), 6 studies [43,45,47,50,52,53] and 2 studies [38,40] enrolled undergraduate and nonundergraduate participants, respectively; the study by Johnson et al [48] was excluded due to a mixed enrollment. Neither of these groups showed a significant difference between VSs and mannequins or RPs ($SMD = 0.36$; 95% CI -0.23 to 0.96 ; $I^2 = 89\%$; $P < .001$ for undergraduate and $SMD = 0.70$; 95% CI -1.35 to 2.75 ; $I^2 = 89\%$; $P < .001$ for nonundergraduate), with no significant subgroup difference at $P = .76$.

For subgroup analysis by comparison (Figure S9c in [Multimedia Appendix 4](#) [38,40,43,45,47,48,50,52,53]), 6 studies [38,40,45,47,48,50] and 3 studies [43,52,53] included mannequin and RP arms, respectively. Neither mannequins nor RPs was significantly different from VSs ($SMD = 0.06$; 95% CI -0.57 to 0.68 ; $I^2 = 75\%$; $P < .001$ for mannequins and $SMD = 0.60$; 95% CI -0.37 to 1.58 ; $I^2 = 93\%$; $P < .001$ for RPs), with no significant subgroup difference at $P = .35$.

Univariate meta-regression revealed that year of publication ($\beta = .08$, $P = .06$), age of participants ($\beta = -.02$; $P = .81$), nursing discipline ($\beta = -.05$; $P = .94$), undergraduate level ($\beta = -.27$; $P = .72$), and real person comparison ($\beta = .53$; $P = .35$) had no effects on clinical reasoning scores (Table S1 in [Multimedia Appendix 4](#)).

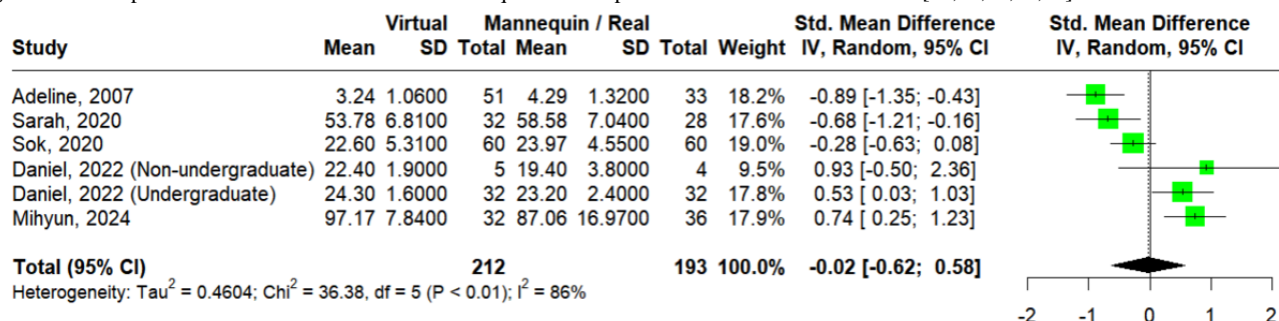
Communication Skills

A total of 5 studies assessed the outcome of communication skills [30,31,39,52,54]. The Egger test showed no statistically significant publication bias ($P = .55$). In [Figure 6](#) [30,31,39,52,54], the pooled effect size did not reflect a

significant difference between VSs and mannequins or RPs on the outcome of procedural skills (SMD=-0.02; 95% CI -0.62 to 0.58; $I^2=86%$; $P<.01$). TSA showed that the “inner wedge” area had been reached, indicating strong evidence that further

studies would hardly be able to change the statistically insignificant result (Figure S10 in [Multimedia Appendix 4](#)). This result was consistent according to the leave-one-out analysis (Figure S11 in [Multimedia Appendix 4](#) [30,31,39,52,54]).

Figure 6. Forest plot of virtual simulations versus mannequins or real persons for communication skills [30,31,39,52,54].



Since all 5 studies compared VSs to RPs, only subgroup analyses by discipline and level could be conducted. For subgroup analysis by discipline (Figure S12a in [Multimedia Appendix 4](#) [30,31,39,52,54]), 3 studies [30,31,39] and 1 study [52] enrolled participants from medicine and nursing, respectively; and the study by Liaw et al [54] was excluded due to a mixed enrollment. Only the nursing group showed a significant difference between VSs and mannequins or RPs (SMD=-0.14; 95% CI -0.97 to 0.69; $I^2=86%$; $P<.001$ for medicine and SMD=0.74; 95% CI 0.25-1.23 for nursing), with no significant subgroup difference at $P=.07$.

For subgroup analysis by level (Figure S12b in [Multimedia Appendix 4](#) [30,31,39,52,54]), 4 studies enrolled undergraduate participants [31,39,52,54] and the study by Sapkaroski et al [30] enrolled both undergraduate and nonundergraduate participants. Neither of these groups showed a significant difference between VSs and mannequins or RPs (SMD=-0.12; 95% CI -0.75 to 0.51; $I^2=88%$; $P<.001$ for undergraduate and SMD=0.93; 95% CI -0.50 to 2.36 for nonundergraduate), with no significant subgroup difference at $P=.19$.

Univariate meta-regression revealed that year of publication ($\beta=.09$; $P=.02$) had a positive effect on communication skill scores, whereas age of participants ($\beta=-.07$; $P=.84$), nursing discipline ($\beta=.88$; $P=.33$), undergraduate level ($\beta=-1.05$; $P=.32$), and real person comparison ($\beta=.19$; $P=.61$) had no effects (Table S1 in [Multimedia Appendix 4](#)). Since no studies compared VSs to mannequins, the effect of real person comparison could not be evaluated.

Discussion

Principal Findings

The aim of this meta-analysis was to evaluate the effectiveness of VSs in comparison with mannequins and RPs in medical and nursing education. We found that VSs and conventional simulations were not statistically different in teaching knowledge, procedural skills, clinical reasoning, and communication skills. Among nursing participants, VSs were more effective in improving procedural skills. For the training

of communication skills, VSs tended to be increasingly effective with time.

This meta-analysis followed the original protocol of PROSPERO (CRD42023466622). TSA and meta-regression were added to further evaluate the accumulative evidence and identify potential sources of heterogeneity. We believe that these analyses do not negatively affect the reliability of our results; instead, they offer novel insights and strengthen the existing content. Here, we state the aforementioned deviations to ensure our study’s transparency.

Comparison With Prior Work

There was no significant difference in knowledge gained between VSs and mannequins or RPs. This result aligns with a previous meta-analysis comparing 2D VSs with traditional education methods [21], while other analyses suggested that 3D VSs were more effective [22,55-57]. In our meta-analysis, almost all studies adopted 3D VSs, but the comparison was specifically narrowed down to conventional simulations. Our pooled effect size tended to favor VSs. Indeed, it can be beneficial to use VSs for knowledge delivery. First, digital platforms or software provide repeatable, information-rich feedback for learning [58]. Second, immersive technology (ie, devices that provide a sense of realism and immersion in the computer-generated world through sensory stimuli) such as VR promotes students’ interest with enhanced satisfaction, self-efficacy, and engagement [59]. Although VSs are viable alternatives to mannequins and RPs, undesirable usability may hinder the learning process. For instance, Cobbett and Snelgrove-Clarke [46] reported in their study that nearly half of the students expressed dislike for the VS due to technological issues such as the “online program was slow” and “didn’t know where to find things.” Consequently, the development of VSs should consider features that enhance ease of use and users’ level of satisfaction (eg, step-by-step instructions, user-friendly interfaces) so that students allocate their time appropriately on learning instead of wrestling with the platform or software.

There was no significant difference in improving procedural skills between VSs and mannequins or RPs, while our pooled effect size tended to favor mannequins or RPs. Subgroup analyses showed that VSs were less effective among nursing

participants. These results were inconsistent with previous meta-analyses comparing VSs with traditional education methods [21,22,57]. Such discrepancies may be due to different definitions of comparison. While VSs are advantageous for practicing and mastering skills relative to didactic teaching, this is not necessarily true if they are compared with mannequins and RPs. A gap exists between digital environments and real circumstances, and the abstraction of procedures may not be transferable to reality [60]. Notwithstanding, VSs are cost-effective alternatives that enable learners to repeat the training freely without worrying about time, space, and patient harm [61]. VSs may be designed to achieve better skill acquisition. First, the addition of haptic technology, for example, offers more realistic hands-on experiences [36]. Second, too much immersion may not be desirable as it can lead to cognitive overload and hamper procedure learning [42].

There was no significant difference in fostering clinical reasoning between VSs and mannequins or RPs. Again, this result does not conform to previous meta-analyses which compared VSs to no intervention or traditional education methods [21,62]. Our pooled effect size tended to favor VSs. It has been suggested that VSs, with a large number of customizable clinical scenarios, are highly suitable for improving clinical assessment and decision-making [63]. In particular, the following elements may be beneficial: a training duration of more than 30 minutes, varied clinical scenarios, nonimmersive 2D digital environments, and postscenario feedback [62]. Similar to procedural skills, nonimmersive VSs appear to be more effective than their immersive counterparts. According to the cognitive load theory, 3D digital environments can increase cognitive load as students get distracted by irrelevant stimuli [62]. More data are still required to reach a conclusive result regarding the comparative effectiveness of VSs in improving clinical reasoning.

There was no significant difference in enhancing communication skills between VSs and mannequins or RPs. This is in line with the findings of a previous meta-analysis comparing digital education to traditional learning [15]. Our pooled effect size tended to favor mannequins or RPs. It is noteworthy, however, that more recent studies tended to favor VSs. A possible explanation is that advances in technology can increasingly overcome the long-standing lack of realism and means for interaction in VSs [30]. Greater immersion can elicit a greater sense of “being there” for users [64]. This psychological experience is critical for communication training, to make learners feel as if they are having face-to-face conversations with digital avatars so that higher levels of empathy can be induced. Moreover, VSs have fewer limits on time or place,

eliminate the costs of SP training, and offer a variety of clinical scenarios available for student communication training. These properties render VSs with great potential as novel methods for teaching communication skills.

Strengths and Limitations

To the best of our knowledge, this is the first meta-analysis comparing the effectiveness of VSs and conventional simulations. Strengths of this study include pragmatic research questions, well-defined outcome measures, comprehensive searches of the most up-to-date RCTs (ie, the gold standard for evaluating the effectiveness of interventions), and critical analyses of the evidence. Our results indicate that VSs are viable alternatives to face-to-face learning modalities. For institutions, time and money saved can be invested in other research and teaching projects. For students, the removal of time, space, and location limitations leads to greater freedom of learning and wider popularization of education.

On the other hand, limitations must be acknowledged when interpreting the results of this study. First, researchers in several studies were not blinded and might introduce bias when assessing participants. Second, certain subgroup or meta-regression analyses were not feasible or included only one entry due to the limited number of studies. Further analyses may be possible when more RCTs are published. Third, the heterogeneity of the included studies could not be fully explained by leave-one-out, subgroup, or meta-regression analyses. This lack of explanation may be attributed to other unspecified sources of methodological and subject heterogeneity.

Conclusions

Overall, this meta-analysis did not find significant differences between VSs and mannequins or RPs in improving knowledge, procedural skills, clinical reasoning, and communication skills. For procedural skill acquisition, VSs were less effective than mannequins or RPs among nursing participants. For communication skill development, VSs exhibited increasing potency with more advanced technology over time. Given other prominent advantages of VSs (eg, cost-effectiveness, flexibility regarding time or space), it is worth considering these tools as alternatives to traditional education methods. As technology continues to evolve, the comparative effectiveness of VSs will need to be reevaluated and updated. In particular, it is anticipated to witness the integration of emerging large language models (eg, ChatGPT) in VSs and see if they can revolutionize educational practices. We also recommend future research to investigate the effectiveness of VSs across various design configurations.

Acknowledgments

This study was funded by the Peking Union Medical College Hospital Teaching Programme (X361011).

Authors' Contributions

HP, SC, and XH contributed equally to the work. NJ, YZ, SC, and XH contributed to the conception and design of the study. NJ performed search strategies. SL and XL selected articles, extracted data, and assessed the risk of bias. NJ and YZ analyzed data. NJ prepared the original draft. All authors reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[\[DOCX File , 29 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Search strategies.

[\[DOCX File , 14 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Characteristics of selected 27 randomized controlled trials.

[\[DOCX File , 30 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Trial sequential, leave-one-out, subgroup and meta-regression analyses.

[\[DOCX File , 4611 KB-Multimedia Appendix 4\]](#)

References

1. Khan K, Pattison T, Sherwood M. Simulation in medical education. *Med Teach*. 2011;33(1):1-3. [doi: [10.3109/0142159X.2010.519412](https://doi.org/10.3109/0142159X.2010.519412)] [Medline: [21182376](https://pubmed.ncbi.nlm.nih.gov/21182376/)]
2. Herrera-Aliaga E, Estrada LD. Trends and innovations of simulation for twenty first century medical education. *Front Public Health*. 2022;10:619769. [FREE Full text] [doi: [10.3389/fpubh.2022.619769](https://doi.org/10.3389/fpubh.2022.619769)] [Medline: [35309206](https://pubmed.ncbi.nlm.nih.gov/35309206/)]
3. Ziv A, Wolpe PR, Small SD, Glick S. Simulation-based medical education: an ethical imperative. *Acad Med*. 2003;78(8):783-788. [doi: [10.1097/00001888-200308000-00006](https://doi.org/10.1097/00001888-200308000-00006)] [Medline: [12915366](https://pubmed.ncbi.nlm.nih.gov/12915366/)]
4. McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Acad Med*. 2011;86(6):706-711. [FREE Full text] [doi: [10.1097/ACM.0b013e318217e119](https://doi.org/10.1097/ACM.0b013e318217e119)] [Medline: [21512370](https://pubmed.ncbi.nlm.nih.gov/21512370/)]
5. Shin S, Park JH, Kim JH. Effectiveness of patient simulation in nursing education: meta-analysis. *Nurse Educ Today*. 2015;35(1):176-182. [doi: [10.1016/j.nedt.2014.09.009](https://doi.org/10.1016/j.nedt.2014.09.009)] [Medline: [25459172](https://pubmed.ncbi.nlm.nih.gov/25459172/)]
6. Oliveira Silva G, Oliveira FSE, Coelho ASG, Cavalcante AMRZ, Vieira FVM, Fonseca LMM, et al. Effect of simulation on stress, anxiety, and self-confidence in nursing students: systematic review with meta-analysis and meta-regression. *Int J Nurs Stud*. 2022;133:104282. [doi: [10.1016/j.ijnurstu.2022.104282](https://doi.org/10.1016/j.ijnurstu.2022.104282)] [Medline: [35679634](https://pubmed.ncbi.nlm.nih.gov/35679634/)]
7. Watmough S, Box H, Bennett N, Stewart A, Farrell M. Unexpected medical undergraduate simulation training (UMUST): Can unexpected medical simulation scenarios help prepare medical students for the transition to foundation year doctor? *BMC Med Educ*. 2016;16:110. [FREE Full text] [doi: [10.1186/s12909-016-0629-x](https://doi.org/10.1186/s12909-016-0629-x)] [Medline: [27079898](https://pubmed.ncbi.nlm.nih.gov/27079898/)]
8. Nielsen RP, Nikolajsen L, Paltved C, Aagaard R. Effect of simulation-based team training in airway management: a systematic review. *Anaesthesia*. 2021;76(10):1404-1415. [FREE Full text] [doi: [10.1111/anae.15375](https://doi.org/10.1111/anae.15375)] [Medline: [33497486](https://pubmed.ncbi.nlm.nih.gov/33497486/)]
9. Luzzi A, Hellwinkel J, O'Connor M, Crutchfield C, Lynch TS. The efficacy of arthroscopic simulation training on clinical ability: a systematic review. *Arthroscopy*. 2021;37(3):1000-1007.e1. [doi: [10.1016/j.arthro.2020.09.018](https://doi.org/10.1016/j.arthro.2020.09.018)] [Medline: [33220467](https://pubmed.ncbi.nlm.nih.gov/33220467/)]
10. James HK, Chapman AW, Pattison GTR, Griffin DR, Fisher JD. Systematic review of the current status of cadaveric simulation for surgical training. *Br J Surg*. 2019;106(13):1726-1734. [FREE Full text] [doi: [10.1002/bjs.11325](https://doi.org/10.1002/bjs.11325)] [Medline: [31573088](https://pubmed.ncbi.nlm.nih.gov/31573088/)]
11. Sarfati L, Ranchon F, Vantard N, Schwiertz V, Larbre V, Parat S, et al. Human-simulation-based learning to prevent medication error: a systematic review. *J Eval Clin Pract*. 2019;25(1):11-20. [doi: [10.1111/jep.12883](https://doi.org/10.1111/jep.12883)] [Medline: [29383867](https://pubmed.ncbi.nlm.nih.gov/29383867/)]
12. Zendejas B, Brydges R, Wang AT, Cook DA. Patient outcomes in simulation-based medical education: a systematic review. *J Gen Intern Med*. 2013;28(8):1078-1089. [FREE Full text] [doi: [10.1007/s11606-012-2264-5](https://doi.org/10.1007/s11606-012-2264-5)] [Medline: [23595919](https://pubmed.ncbi.nlm.nih.gov/23595919/)]
13. Turner S, Harder N, Vigier D, Cooper A, Pinel K, Mitchell K. Lessons from implementing virtual simulations: a multi-program evaluation. *Clin Simul Nurs*. 2023;74:57-64. [doi: [10.1016/j.ecns.2022.09.001](https://doi.org/10.1016/j.ecns.2022.09.001)]
14. Lioce L, Anderson M, Diaz D, Robertson J, Chang T, Downing D, et al. *Healthcare Simulation Dictionary*. 2nd Edition. Rockville, MD. Agency for Healthcare Research and Quality; 2020.
15. Kyaw BM, Posadzki P, Paddock S, Car J, Campbell J, Tudor Car L. Effectiveness of digital education on communication skills among medical students: systematic review and meta-analysis by the digital health education collaboration. *J Med Internet Res*. 2019;21(8):e12967. [FREE Full text] [doi: [10.2196/12967](https://doi.org/10.2196/12967)] [Medline: [31456579](https://pubmed.ncbi.nlm.nih.gov/31456579/)]

16. Han SG, Kim YD, Kong TY, Cho J. Virtual reality–based neurological examination teaching tool (VRNET) versus standardized patient in teaching neurological examinations for the medical students: a randomized, single-blind study. *BMC Med Educ.* 2021;21(1):493. [FREE Full text] [doi: [10.1186/s12909-021-02920-4](https://doi.org/10.1186/s12909-021-02920-4)] [Medline: [34526004](https://pubmed.ncbi.nlm.nih.gov/34526004/)]
17. Dante A, Marcotullio A, Masotta V, Caponnetto V, La CC, Bertocchi L, et al. From high-fidelity patient simulators to robotics and artificial intelligence: a discussion paper on new challenges to enhance learning in nursing education. 2020. Presented at: 10th International Conference on Methodologies and Intelligent Systems for Technology Enhanced Learning; June 17-19, 2020; L'Aquila, Italy. [doi: [10.1007/978-3-030-52287-2_11](https://doi.org/10.1007/978-3-030-52287-2_11)]
18. Pietersen PI, Jørgensen R, Graumann O, Konge L, Skaarup SH, Lawaetz Schultz HH, et al. Training thoracic ultrasound skills: a randomized controlled trial of simulation-based training versus training on healthy volunteers. *Respiration.* 2021;100(1):34-43. [doi: [10.1159/000509298](https://doi.org/10.1159/000509298)] [Medline: [33454705](https://pubmed.ncbi.nlm.nih.gov/33454705/)]
19. Choules AP. The use of elearning in medical education: a review of the current situation. *Postgrad Med J.* 2007;83(978):212-216. [FREE Full text] [doi: [10.1136/pgmj.2006.054189](https://doi.org/10.1136/pgmj.2006.054189)] [Medline: [17403945](https://pubmed.ncbi.nlm.nih.gov/17403945/)]
20. Bumbach MD, Culross BA, Datta SK. Assessing the financial sustainability of high-fidelity and virtual reality simulation for nursing education: a retrospective case analysis. *Comput Inform Nurs.* 2022;40(9):615-623. [doi: [10.1097/CIN.0000000000000916](https://doi.org/10.1097/CIN.0000000000000916)] [Medline: [35524349](https://pubmed.ncbi.nlm.nih.gov/35524349/)]
21. Kononowicz AA, Woodham LA, Edelbring S, Stathakarou N, Davies D, Saxena N, et al. Virtual patient simulations in health professions education: systematic review and meta-analysis by the digital health education collaboration. *J Med Internet Res.* 2019;21(7):e14676. [FREE Full text] [doi: [10.2196/14676](https://doi.org/10.2196/14676)] [Medline: [31267981](https://pubmed.ncbi.nlm.nih.gov/31267981/)]
22. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. *J Med Internet Res.* 2019;21(1):e12959. [FREE Full text] [doi: [10.2196/12959](https://doi.org/10.2196/12959)] [Medline: [30668519](https://pubmed.ncbi.nlm.nih.gov/30668519/)]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009;339:b2535. [FREE Full text] [doi: [10.1136/bmj.b2535](https://doi.org/10.1136/bmj.b2535)] [Medline: [19622551](https://pubmed.ncbi.nlm.nih.gov/19622551/)]
24. Sim JH, Abdul Aziz YF, Mansor A, Vijayanathan A, Foong CC, Vadivelu J. Students' performance in the different clinical skills assessed in OSCE: What does it reveal? *Med Educ Online.* 2015;20:26185. [FREE Full text] [doi: [10.3402/meo.v20.26185](https://doi.org/10.3402/meo.v20.26185)] [Medline: [25697602](https://pubmed.ncbi.nlm.nih.gov/25697602/)]
25. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.4. Cochrane. 2023. URL: <https://training.cochrane.org/handbook> [accessed 2023-08-22]
26. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol.* 2014;14:135. [FREE Full text] [doi: [10.1186/1471-2288-14-135](https://doi.org/10.1186/1471-2288-14-135)] [Medline: [25524443](https://pubmed.ncbi.nlm.nih.gov/25524443/)]
27. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21(11):1539-1558. [doi: [10.1002/sim.1186](https://doi.org/10.1002/sim.1186)] [Medline: [12111919](https://pubmed.ncbi.nlm.nih.gov/12111919/)]
28. Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis. *BMC Med Res Methodol.* 2017;17(1):39. [FREE Full text] [doi: [10.1186/s12874-017-0315-7](https://doi.org/10.1186/s12874-017-0315-7)] [Medline: [28264661](https://pubmed.ncbi.nlm.nih.gov/28264661/)]
29. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health.* 2019;22(4):153-160. [FREE Full text] [doi: [10.1136/ebmental-2019-300117](https://doi.org/10.1136/ebmental-2019-300117)] [Medline: [31563865](https://pubmed.ncbi.nlm.nih.gov/31563865/)]
30. Sapkaroski D, Mundy M, Dimmock MR. Immersive virtual reality simulated learning environment versus role-play for empathic clinical communication training. *J Med Radiat Sci.* 2022;69(1):56-65. [FREE Full text] [doi: [10.1002/jmrs.555](https://doi.org/10.1002/jmrs.555)] [Medline: [34706398](https://pubmed.ncbi.nlm.nih.gov/34706398/)]
31. O'Rourke SR, Branford KR, Brooks TL, Ives LT, Nagendran A, Compton SN. The emotional and behavioral impact of delivering bad news to virtual versus real standardized patients: a pilot study. *Teach Learn Med.* 2020;32(2):139-149. [doi: [10.1080/10401334.2019.1652180](https://doi.org/10.1080/10401334.2019.1652180)] [Medline: [31437006](https://pubmed.ncbi.nlm.nih.gov/31437006/)]
32. Weber U, Zapletal B, Base E, Hambrusch M, Ristl R, Mora B. Resident performance in basic perioperative transesophageal echocardiography: comparing 3 teaching methods in a randomized controlled trial. *Medicine.* 2019;98(36):e17072. [FREE Full text] [doi: [10.1097/MD.00000000000017072](https://doi.org/10.1097/MD.00000000000017072)] [Medline: [31490407](https://pubmed.ncbi.nlm.nih.gov/31490407/)]
33. Jiang B, Ju H, Zhao Y, Yao L, Feng Y. Comparison of the efficacy and efficiency of the use of virtual reality simulation with high-fidelity mannequins for simulation-based training of fiberoptic bronchoscope manipulation. *Simul Healthc.* 2018;13(2):83-87. [doi: [10.1097/SIH.0000000000000299](https://doi.org/10.1097/SIH.0000000000000299)] [Medline: [29621098](https://pubmed.ncbi.nlm.nih.gov/29621098/)]
34. Spadaro S, Karbing DS, Fogagnolo A, Ragazzi R, Mojoli F, Astolfi L, et al. Simulation training for residents focused on mechanical ventilation: a randomized trial using mannequin-based versus computer-based simulation. *Simul Healthc.* 2017;12(6):349-355. [FREE Full text] [doi: [10.1097/SIH.0000000000000249](https://doi.org/10.1097/SIH.0000000000000249)] [Medline: [28825930](https://pubmed.ncbi.nlm.nih.gov/28825930/)]
35. Ahlqvist JB, Nilsson TA, Hedman LR, Desser TS, Dev P, Johansson M, et al. A randomized controlled trial on 2 simulation-based training methods in radiology: effects on radiologic technology student skill in assessing image quality. *Simul Healthc.* 2013;8(6):382-387. [doi: [10.1097/SIH.0b013e3182a60a48](https://doi.org/10.1097/SIH.0b013e3182a60a48)] [Medline: [24096919](https://pubmed.ncbi.nlm.nih.gov/24096919/)]
36. LeBlanc J, Hutchison C, Hu Y, Donnon T. A comparison of orthopaedic resident performance on surgical fixation of an ulnar fracture using virtual reality and synthetic models. *J Bone Joint Surg Am.* 2013;95(9):e60, S1-e60, S5. [doi: [10.2106/JBJS.K.01284](https://doi.org/10.2106/JBJS.K.01284)] [Medline: [23636199](https://pubmed.ncbi.nlm.nih.gov/23636199/)]

37. Andreatta PB, Maslowski E, Petty S, Shim W, Marsh M, Hall T, et al. Virtual reality triage training provides a viable solution for disaster-preparedness. *Acad Emerg Med*. 2010;17(8):870-876. [FREE Full text] [doi: [10.1111/j.1553-2712.2010.00728.x](https://doi.org/10.1111/j.1553-2712.2010.00728.x)] [Medline: [20670325](https://pubmed.ncbi.nlm.nih.gov/20670325/)]
38. Youngblood P, Harter PM, Srivastava S, Moffett S, Heinrichs WL, Dev P. Design, development, and evaluation of an online virtual emergency department for training trauma teams. *Simul Healthc*. 2008;3(3):146-153. [doi: [10.1097/SIH.0b013e31817bedf7](https://doi.org/10.1097/SIH.0b013e31817bedf7)] [Medline: [19088658](https://pubmed.ncbi.nlm.nih.gov/19088658/)]
39. Deladisma AM, Cohen M, Stevens A, Wagner P, Lok B, Bernard T, et al. Do medical students respond empathetically to a virtual patient? *Am J Surg*. 2007;193(6):756-760. [doi: [10.1016/j.amjsurg.2007.01.021](https://doi.org/10.1016/j.amjsurg.2007.01.021)] [Medline: [17512291](https://pubmed.ncbi.nlm.nih.gov/17512291/)]
40. Tran MT, Ahmad M, Patel K, Argyriou O, Davies A, Shalhoub J. Comparing virtual reality and simulation to teach the assessment and management of acute surgical scenarios: a pilot study. *Health Sci Rep*. 2024;7(7):e2245. [FREE Full text] [doi: [10.1002/hsr2.2245](https://doi.org/10.1002/hsr2.2245)] [Medline: [38983682](https://pubmed.ncbi.nlm.nih.gov/38983682/)]
41. Malik TG, Mahboob U, Khan RA, Alam R. Virtual patients versus standardized patients for improving clinical reasoning skills in ophthalmology residents. A randomized controlled trial. *BMC Med Educ*. 2024;24(1):429. [FREE Full text] [doi: [10.1186/s12909-024-05241-4](https://doi.org/10.1186/s12909-024-05241-4)] [Medline: [38649884](https://pubmed.ncbi.nlm.nih.gov/38649884/)]
42. Simsek-Cetinkaya S, Cakir SK. Evaluation of the effectiveness of artificial intelligence assisted interactive screen-based simulation in breast self-examination: an innovative approach in nursing students. *Nurse Educ Today*. 2023;127:105857. [doi: [10.1016/j.nedt.2023.105857](https://doi.org/10.1016/j.nedt.2023.105857)] [Medline: [37253303](https://pubmed.ncbi.nlm.nih.gov/37253303/)]
43. Sahin Karaduman G, Basak T. Is virtual patient simulation superior to human patient simulation: a randomized controlled study. *Comput Inform Nurs*. 2023;41(6):467-476. [doi: [10.1097/CIN.0000000000000957](https://doi.org/10.1097/CIN.0000000000000957)] [Medline: [36633879](https://pubmed.ncbi.nlm.nih.gov/36633879/)]
44. Padilha JM, Machado PP, Ribeiro A, Ramos J, Costa P. Clinical virtual simulation in nursing education: randomized controlled trial. *J Med Internet Res*. 2019;21(3):e11529. [FREE Full text] [doi: [10.2196/11529](https://doi.org/10.2196/11529)] [Medline: [30882355](https://pubmed.ncbi.nlm.nih.gov/30882355/)]
45. Haerling KA. Cost-utility analysis of virtual and mannequin-based simulation. *Simul Healthc*. 2018;13(1):33-40. [doi: [10.1097/SIH.0000000000000280](https://doi.org/10.1097/SIH.0000000000000280)] [Medline: [29373382](https://pubmed.ncbi.nlm.nih.gov/29373382/)]
46. Cobbett S, Snelgrove-Clarke E. Virtual versus face-to-face clinical simulation in relation to student knowledge, anxiety, and self-confidence in maternal-newborn nursing: a randomized controlled trial. *Nurse Educ Today*. 2016;45:179-184. [doi: [10.1016/j.nedt.2016.08.004](https://doi.org/10.1016/j.nedt.2016.08.004)] [Medline: [27537670](https://pubmed.ncbi.nlm.nih.gov/27537670/)]
47. Liaw SY, Chan SWC, Chen FG, Hooi SC, Siau C. Comparison of virtual patient simulation with mannequin-based simulation for improving clinical performances in assessing and managing clinical deterioration: randomized controlled trial. *J Med Internet Res*. 2014;16(9):e214. [FREE Full text] [doi: [10.2196/jmir.3322](https://doi.org/10.2196/jmir.3322)] [Medline: [25230684](https://pubmed.ncbi.nlm.nih.gov/25230684/)]
48. Johnson MP, Hickey KT, Scopa-Goldman J, Andrews T, Boerem P, Covec M, et al. Manikin versus web-based simulation for advanced practice nursing students. *Clin Simul Nurs*. 2014;10(6):e317-e323. [doi: [10.1016/j.ecns.2014.02.004](https://doi.org/10.1016/j.ecns.2014.02.004)]
49. Chang KKP, Chung JWY, Wong TKS. Learning intravenous cannulation: a comparison of the conventional method and the cathSim intravenous training system. *J Clin Nurs*. 2002;11(1):73-78. [doi: [10.1046/j.1365-2702.2002.00561.x](https://doi.org/10.1046/j.1365-2702.2002.00561.x)] [Medline: [11845758](https://pubmed.ncbi.nlm.nih.gov/11845758/)]
50. Durmaz A, Dicle A, Cakan E, Cakir S. Effect of screen-based computer simulation on knowledge and skill in nursing students' learning of preoperative and postoperative care management: a randomized controlled study. *Comput Inform Nurs*. 2012;30(4):196-203. [doi: [10.1097/NCN.0b013e3182419134](https://doi.org/10.1097/NCN.0b013e3182419134)] [Medline: [22228217](https://pubmed.ncbi.nlm.nih.gov/22228217/)]
51. Jung EY, Park DK, Lee YH, Jo HS, Lim YS, Park RW. Evaluation of practical exercises using an intravenous simulator incorporating virtual reality and haptics device technologies. *Nurse Educ Today*. 2012;32(4):458-463. [doi: [10.1016/j.nedt.2011.05.012](https://doi.org/10.1016/j.nedt.2011.05.012)] [Medline: [21664014](https://pubmed.ncbi.nlm.nih.gov/21664014/)]
52. Lee M, Kim SK, Go Y, Jeong H, Lee Y. Positioning virtual reality as means of clinical experience in mental health nursing education: a quasi-experimental study. *Appl Nurs Res*. 2024;77:151800. [doi: [10.1016/j.apnr.2024.151800](https://doi.org/10.1016/j.apnr.2024.151800)] [Medline: [38796255](https://pubmed.ncbi.nlm.nih.gov/38796255/)]
53. Liaw SY, Sutini, Chua WL, Tan JZ, Levett-Jones T, Ashokka B, et al. Desktop virtual reality versus face-to-face simulation for team-training on stress levels and performance in clinical deterioration: a randomised controlled trial. *J Gen Intern Med*. 2023;38(1):67-73. [FREE Full text] [doi: [10.1007/s11606-022-07557-7](https://doi.org/10.1007/s11606-022-07557-7)] [Medline: [35501626](https://pubmed.ncbi.nlm.nih.gov/35501626/)]
54. Liaw SY, Ooi SW, Rusli KDB, Lau TC, Tam WWS, Chua WL. Nurse-physician communication team training in virtual reality versus live simulations: randomized controlled trial on team communication and teamwork attitudes. *J Med Internet Res*. 2020;22(4):e17279. [FREE Full text] [doi: [10.2196/17279](https://doi.org/10.2196/17279)] [Medline: [32267235](https://pubmed.ncbi.nlm.nih.gov/32267235/)]
55. Chen FQ, Leng YF, Ge JF, Wang DW, Li C, Chen B, et al. Effectiveness of virtual reality in nursing education: meta-analysis. *J Med Internet Res*. 2020;22(9):e18290. [FREE Full text] [doi: [10.2196/18290](https://doi.org/10.2196/18290)] [Medline: [32930664](https://pubmed.ncbi.nlm.nih.gov/32930664/)]
56. Woon APN, Mok WQ, Chieng YJS, Zhang HM, Ramos P, Mustadi HB, et al. Effectiveness of virtual reality training in improving knowledge among nursing students: a systematic review, meta-analysis and meta-regression. *Nurse Educ Today*. 2021;98:104655. [doi: [10.1016/j.nedt.2020.104655](https://doi.org/10.1016/j.nedt.2020.104655)] [Medline: [33303246](https://pubmed.ncbi.nlm.nih.gov/33303246/)]
57. Liu K, Zhang W, Li W, Wang T, Zheng Y. Effectiveness of virtual reality in nursing education: a systematic review and meta-analysis. *BMC Med Educ*. 2023;23(1):710. [FREE Full text] [doi: [10.1186/s12909-023-04662-x](https://doi.org/10.1186/s12909-023-04662-x)] [Medline: [37770884](https://pubmed.ncbi.nlm.nih.gov/37770884/)]
58. Kiegaldie D, Shaw L. Virtual reality simulation for nursing education: effectiveness and feasibility. *BMC Nurs*. 2023;22(1):488. [FREE Full text] [doi: [10.1186/s12912-023-01639-5](https://doi.org/10.1186/s12912-023-01639-5)] [Medline: [38114940](https://pubmed.ncbi.nlm.nih.gov/38114940/)]

59. Ryan GV, Callaghan S, Rafferty A, Higgins MF, Mangina E, McAuliffe F. Learning outcomes of immersive technologies in health care student education: systematic review of the literature. *J Med Internet Res*. 2022;24(2):e30082. [FREE Full text] [doi: [10.2196/30082](https://doi.org/10.2196/30082)] [Medline: [35103607](https://pubmed.ncbi.nlm.nih.gov/35103607/)]
60. Gu Y, Zou Z, Chen X. The effects of vSIM for Nursing™ as a teaching strategy on fundamentals of nursing education in undergraduates. *Clin Simul Nurs*. 2017;13(4):194-197. [doi: [10.1016/j.ecns.2017.01.005](https://doi.org/10.1016/j.ecns.2017.01.005)]
61. Plotzky C, Loessl B, Kuhnert B, Friedrich N, Kugler C, König P, et al. My hands are running away—learning a complex nursing skill via virtual reality simulation: a randomised mixed methods study. *BMC Nurs*. 2023;22(1):222. [FREE Full text] [doi: [10.1186/s12912-023-01384-9](https://doi.org/10.1186/s12912-023-01384-9)] [Medline: [37370124](https://pubmed.ncbi.nlm.nih.gov/37370124/)]
62. Sim JJM, Rusli KDB, Seah B, Levett-Jones T, Lau Y, Liaw SY. Virtual simulation to enhance clinical reasoning in nursing: a systematic review and meta-analysis. *Clin Simul Nurs*. 2022;69:26-39. [FREE Full text] [doi: [10.1016/j.ecns.2022.05.006](https://doi.org/10.1016/j.ecns.2022.05.006)] [Medline: [35754937](https://pubmed.ncbi.nlm.nih.gov/35754937/)]
63. LaManna JB, Guido-Sanz F, Anderson M, Chase SK, Weiss JA, Blackwell CW. Teaching diagnostic reasoning to advanced practice nurses: positives and negatives. *Clin Simul Nurs*. 2019;26:24-31. [doi: [10.1016/j.ecns.2018.10.006](https://doi.org/10.1016/j.ecns.2018.10.006)]
64. Cummings JJ, Bailenson JN. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychol*. 2015;19(2):272-309. [doi: [10.1080/15213269.2015.1015740](https://doi.org/10.1080/15213269.2015.1015740)]

Abbreviations

AI: artificial intelligence

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses

RCT: randomized controlled trial

RIS: required information size

RP: real person

SMD: standardized mean difference

SP: standardized patient

TSA: trial sequential analysis

VP: virtual patient

VR: virtual reality

VS: virtual simulation

Edited by A Mavragani; submitted 09.01.24; peer-reviewed by J Zhang, YG Abdildin; comments to author 03.05.24; revised version received 23.06.24; accepted 16.09.24; published 05.12.24

Please cite as:

Jiang N, Zhang Y, Liang S, Lyu X, Chen S, Huang X, Pan H

Effectiveness of Virtual Simulations Versus Mannequins and Real Persons in Medical and Nursing Education: Meta-Analysis and Trial Sequential Analysis of Randomized Controlled Trials

J Med Internet Res 2024;26:e56195

URL: <https://www.jmir.org/2024/1/e56195>

doi: [10.2196/56195](https://doi.org/10.2196/56195)

PMID:

©Nan Jiang, Yuelun Zhang, Siyu Liang, Xiaohong Lyu, Shi Chen, Xiaoming Huang, Hui Pan. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 05.12.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.