Li *et al. BMC Emergency Medicine* (2025) 25:94
https://doi.org/10.1186/s12873-025-01256-2

BMC Emergency Medicine

RESEARCH Open Access

Effectiveness of extended reality technologies in cardiopulmonary resuscitation training: a bayesian network meta-analysis

Xiangmin Li[1,2], Xinbo Yin[1,2], Guoqing Huang[1,2] and Xiaokai Wang[1,2*]

Abstract

Background High-quality cardiopulmonary resuscitation (CPR) is critical to cardiac arrest patients. Extended Reality (XR) technologies, including Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR), provide immersive and interactive training, potentially enhancing CPR outcomes. This network meta-analysis compared the effectiveness of XR-based CPR training to traditional face-to-face methods.

Methods A Bayesian network meta-analysis was conducted following PRISMA guidelines. We systematically searched PubMed, Cochrane Library, Web of Science, EMBASE, and CNKI for randomized controlled trials (RCTs) comparing XR-based and traditional CPR training. Primary outcomes included chest compression depth and rate; secondary outcomes assessed full chest wall recoil. The CINeMA tool (GRADE framework) was used to assess evidence quality. Statistical analyses were performed using Stata 15 SE and ADDIS software with random-effects models.

Results 11 RCTs (1,190 participants) were included. MR showed the improvement in chest compression depth (SMD = 10.96; 95% CI, 0.95 to 20.82) compared to VR and traditional methods. For full chest wall recoil, AR outperformed VR (SMD = 48.57; 95% CI, 19.56 to 79.75) and traditional methods (SMD = 52.95; 95% CI, 25.94 to 80.48). However, no significant differences were observed for chest compression rate. SUCRA rankings placed MR as most effective for compression depth (87.4%) and AR for full chest wall recoil (99.1%). Evidence quality was moderate to high, with minor downgrades for imprecision. No publication bias was detected.

Conclusions XR technologies, particularly MR and AR, significantly improve chest compression depth and full chest wall recoil in comparing with face to face CRP training, offering a flexible and engaging approach to CPR training. Further studies are needed to evaluate long-term skill retention and real-world impact.

Clinical trial number Not applicable.

Keywords Cardiopulmonary resuscitation training, Extended reality, Augmented reality, Virtual reality, Mixed reality

*Correspondence:
Xiaokai Wang
xiaokaiwang@csu.edu.cn
[1]Emergency Department, Xiangya Hospital, Central South University, Changsha, Hunan, China
[2]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, Hunan, China

Background

The high quality of cardiopulmonary resuscitation (CPR) is essential for improving recovery outcomes in cardiac arrest patients, both in cases of in-hospital cardiac arrest (IHCA) and out-of-hospital cardiac arrest (OHCA) [1, 2]. Improving the CPR training methods is a critical component of Basic Life Support (BLS) and plays a significant role in increasing the return of spontaneous circulation (ROSC). Traditional BLS training methods, which often combine in-person instruction with video-based teaching have several limitations. These include constraints such as trainer availability, venue scheduling, and logistical challenges [3, 4]. Additionally, in large-scale training sessions, limited interaction between participants and trainers will hinder the effectiveness of skill acquisition and retention [5]. To address these challenges, the integration of advanced and immersive simulation technologies is enhancing CPR training within BLS education, offering more effective and engaging learning experiences.

Extended Reality (XR), such as Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR) are transformative technologies through immersive and interactive learning [6, 7]. These advanced tools provide a highly interactive and realistic training environment, enhance skill retention rates, and improve the overall effectiveness of CPR training programs. AR overlays digital content, like graphics or data, onto the physical world using devices like smart glasses, enriching real-world experiences without direct interaction with physical objects. VR immerses users in fully digital, three-dimensional environments via head-mounted displays (HMDs), offering realistic simulations or fictional worlds enhanced by spatial audio and haptic feedback [8]. MR combines the strengths of both, merging virtual and real environments seamlessly. Using advanced devices like holographic displays, MR enables users to interact naturally with digital content, such as manipulating holograms or integrating real-world actions into virtual workflows [9]. Leveraging the unique features of XR, this technology can address the limitations of traditional CPR training by enabling the training of more individuals with improved quality, even under resource constraints. XR also enhances accessibility and flexibility, allowing participants to engage in training more conveniently [10]. Moreover, its innovative and immersive approach adds a sense of novelty, making the learning process more enjoyable and satisfying for participants, and boosting engagement and motivation.

Currently, Sun et al. have compared the effectiveness of VR and AR to traditional CPR training [6]. However, their study lacks the granularity to distinguish between AR and VR and does not include MR methods. Additionally, no direct comparisons currently exist between traditional CPR training methods and the three XR-based approaches (AR, VR, and MR) in terms of training effectiveness. To address this gap, this study employs a network meta-analysis to evaluate the effectiveness of various CPR training methods. The findings aim to provide robust evidence supporting the integration of XR technologies into CPR training.

Methods

Study design

In this Bayesian network meta-analysis, we compared the effectiveness of XR-based CPR training with traditional face-to-face CPR training.

Data sources and search strategy

This network meta-analysis (NMA) protocol was registered in the prospective register of systematic reviews (CRD42024609853). The NMA was following the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for network meta-analyses [11]. A comprehensive search was performed in PubMed, the Cochrane Library, Web of Science, EMBASE, and CNKI to identify cluster randomized controlled trials comparing the effectiveness of XR (Extended Reality) and face-to-face CPR training. The search period spanned from January 1, 1970, to December 1, 2024. Additionally, manual searches of references from included studies and relevant systematic reviews were conducted to supplement and gather related literature. The following search strategy was employed: "(virtual reality OR augmented reality OR mixed reality) AND (Cardiopulmonary Resuscitation OR Heart Arrest OR Sudden Cardiac Arrest OR CPR OR basic life support OR chest compression)". No restrictions were placed on the language of publication.

Study selection and eligibility criteria

The inclusion criteria were cluster randomized controlled trials (RCTs) comparing the effectiveness of XR-based CPR training with traditional face-to-face CPR training. The outcomes assessed CPR performance, including chest compression depth, rate and full chest wall recoil. (Table 1). Exclusion criteria included studies focused on pediatric or neonatal CPR training, those that did not report CPR performance outcomes, non-RCT experiments, incomplete or duplicate publications of relevant data, non-human studies, and reviews or study protocols.

Data extraction and quality assessment

Two reviewers (X.Y. and X.W.) independently extracted relevant data parameters. In case of disagreement, the arbitration was conducted by a third reviewer. The following information was extracted: author, publication year, study location, sample size, background of the participants, XR used in the intervention group, and outcomes

Table 1 Descriptive details of studies that were included in the network meta-analysis

Author and year	Counrty/region	Type of participants	Sample size	XR used in the intervention group	Outcomes
Sungur 2023 [7]	Netherlands	Layperson	60	MR	CC rate and CC depth
Hubail 2022 [15]	United Kingdom	Layperson	26	VR	CC rate, CC depth and CC recoil
Chang 2023 [16]	Taiwan	Layperson	45	VR	CC recoil
Alcazar Artero 2024 [17]	Spain	Layperson	63	VR	CC rate and CC depth
Zhou 2022 [18]	China	Medical students	40	VR	CC rate and CC depth
Hou 2022 [19]	China	Layperson	163	AR	CC rate and CC depth
Hou 2022 [20]	China	Layperson	28	AR	CC rate and CC depth
Aranda-García 2023 [21]	Spain	Layperson	60	AR	CC rate, CC depth and CC recoil
Nas 2020 [22]	Netherlands	Layperson	252	VR	CC rate, CC depth and CC recoil
Castillo 2023 [23]	Spain	Student of Health Sciences	241	VR	CC depth and CC recoil
Leary 2020 [24]	United States	HCPs	100	AR	CC rate and CC depth

HCPs: Health care providers
CC: Chest Compression
AR: Augmented Reality
VR: Virtual Reality
MR: Mixed Reality

of interest. To ensure comprehensive and accurate data extraction, we developed a standardized extraction form based on the Cochrane Handbook guidelines. For studies reporting multiple time points, we extracted data from the most immediate post-intervention assessment. When studies reported both unadjusted and adjusted analyses, we prioritized the adjusted data. For outcomes reported as medians with ranges or interquartile ranges, we converted these to means and standard deviations using established statistical methods. All extracted data underwent double verification, with any discrepancies resolved through consensus meetings. The study quality was assessed by two authors (G.H. and X.W.) according to Cochrane Collaboration's tool. It includes six aspects: sequence generation, allocation consideration, blinding, incomplete outcome data, no selective outcome reporting, and other sources. RevMan software (v 5.3) was only used for the risk of bias summary. In case of disagreement, the arbitration will be conducted by the corresponding author.

Outcomes
The primary outcomes focused on CPR performance metrics, such as chest compression depth and rate. Secondary outcomes assessed the proportion of participants meeting CPR guideline quality criteria, including full chest recoil rate (also referred to as complete re-expansion or adequate recoil rate), Risk ratios (RRs) were calculated with 95% confidence intervals (CIs) for dichotomous data. Continuous data that had the same measure unit (i.e. depth and rate of chest compressions) were calculated as mean differences (MDs) with 95% CIs.

Methodological quality assessment
The Cochrane Risk of Bias Tool for RCTs (RoB 2.0) was used to evaluate the methodological quality of the included studies. Two reviewers independently assessed the studies, with any disagreements resolved through discussion with a corresponding author. The RoB 2.0 tool evaluates bias across several domains, including bias from the randomization process, deviations from intended interventions, missing outcome data, outcome measurement, and selection of reported results. Based on these assessments, studies were classified as having low risk, some concerns, or high risk of bias.

Statistical analysis
Stata 15 SE was used to generate network diagrams, while ADDIS software (version 1.16.8) was employed for network meta-analysis, with all analyses defaulting to a random-effects model. Node-split analysis was conducted to assess the consistency between direct and indirect comparisons. A P-value > 0.05 indicated consistency, and a consistency model was used; otherwise, an inconsistency model was applied [12]. If node-split analysis was not feasible, results from both models were reported. The potential scale reduction factor (PSRF) was used to assess model convergence, with a PSRF value of 1 indicating approximate convergence [13]. Effect sizes were expressed as network OR values with 95% credible intervals (CrIs), along with output rankings and the surface under the cumulative ranking curve (SUCRA).

In our Bayesian models, we employed minimally informative priors to minimize undue influence on posterior distributions while maintaining computational stability.

Specifically, we used normal distributions with mean 0 and variance 10,000 for effect parameters, allowing the observed data to primarily drive our results. For variance parameters, we employed half-normal priors with scale parameter 5, permitting reasonable between-study heterogeneity without allowing extreme values that could destabilize the model. Model convergence was carefully monitored throughout the analysis process using standard diagnostic approaches [13].

R software (version 3.6.1) was used for heterogeneity and sensitivity analyses. Following the guidelines in the Cochrane Handbook, heterogeneity was considered insignificant if Q-values were less than the degrees of freedom, P-values were greater than 0.10, and I^2 values ranged from 0 to 40%. Conversely, data were deemed heterogeneous if Q-values exceeded the DF, P-values were below 0.10, and I^2 values ranged from 75 to 100%. Sensitivity analysis was performed by switching from a random-effects model to a fixed-effects model; if the results remained consistent, sensitivity was considered low, indicating that the findings were stable and reliable [14]. The datasets, R scripts and detailed package version information are provided in the Supplementary Material.

Assessment of publication bias

To assess the presence of publication bias, Egger's regression test was employed. This test evaluates the asymmetry of the funnel plot by performing a linear regression of the effect sizes on their standard errors. A significant intercept in the regression model indicates potential publication bias. In addition, a funnel plot was created to visually inspect the symmetry of the included studies. The test was conducted using the metabias function from the R meta package, with a significance threshold set at $p < 0.05$.

Evidence quality in network meta-analysis

The Confidence in Network Meta-Analysis (CINeMA) tool was used to systematically evaluate the quality of evidence in the network meta-analysis. CINeMA is based on the GRADE framework and assesses evidence across six domains: within-study bias, reporting bias, indirectness, imprecision, heterogeneity, and incoherence. Specifically, the results of the network meta-analysis (including direct and indirect comparisons) were imported into the CINeMA tool. For each comparison, CINeMA provides an overall confidence rating based on the six domains, categorizing evidence quality into four levels: high, moderate, low, and very low. All ratings were independently performed by two reviewers, and disagreements were resolved through discussion.

Transitivity and network meta-analysis assumptions

Our network meta-analysis assumes transitivity across the compared interventions, meaning that participants receiving different interventions across studies would be eligible to receive any of the interventions being compared. To ensure this assumption holds, we carefully selected studies with similar participant characteristics (age ranges, professional backgrounds) and standardized outcome measurements. We evaluated potential effect modifiers across studies, including training duration, participant experience levels, and technological specifications of XR devices. While some variation exists, we determined that these differences were not substantial enough to violate the transitivity assumption.

For intervention comparability, we ensured all CPR training protocols adhered to standard guidelines (American Heart Association or European Resuscitation Council) regardless of delivery method, maintaining fundamental skill components across traditional and XR-based approaches. This approach helped maintain the validity of both direct and indirect comparisons within our network.

Results

Study selection and characteristics of the included studies

Our search strategy identified a total of 1525 citations. After removing duplicates and screening titles and abstracts, the full texts of 150 studies were reviewed for eligibility. Finally, 11 studies were included in meta-analyses [7, 15–24]. Figure 1 illustrates the study flowchart and selection process. In total, 1190 participants were analyzed with a network meta-analysis, including 10 two-arm studies and 1 three-arm study (Table 1). In the included literature, we studied three kinds of interventions (AR, VR, MR) and the fact-to-face Control group. The included participants had varying levels of medical knowledge. Most studies recruited individuals from the general public without formal medical education, while others included medical students with basic medical knowledge and students majoring in health sciences. According to the descriptions in the included studies, none had prior experience in CPR training.

Risk of bias and quality of evidence assessment

The research included in this study used RevMan software (version 5.3) for the risk of bias summary. All included RCTs appeared to exhibit an acceptable level of methodological rigor based on standard criteria. Specifically, most trials provided clear details regarding random sequence generation and allocation concealment (e.g., sealed envelopes, random number tables, or computer-generated sequences). Funnel plots did not indicate any prominent publication bias across the included studies. Figure 2 summarizes the risk of bias assessments.

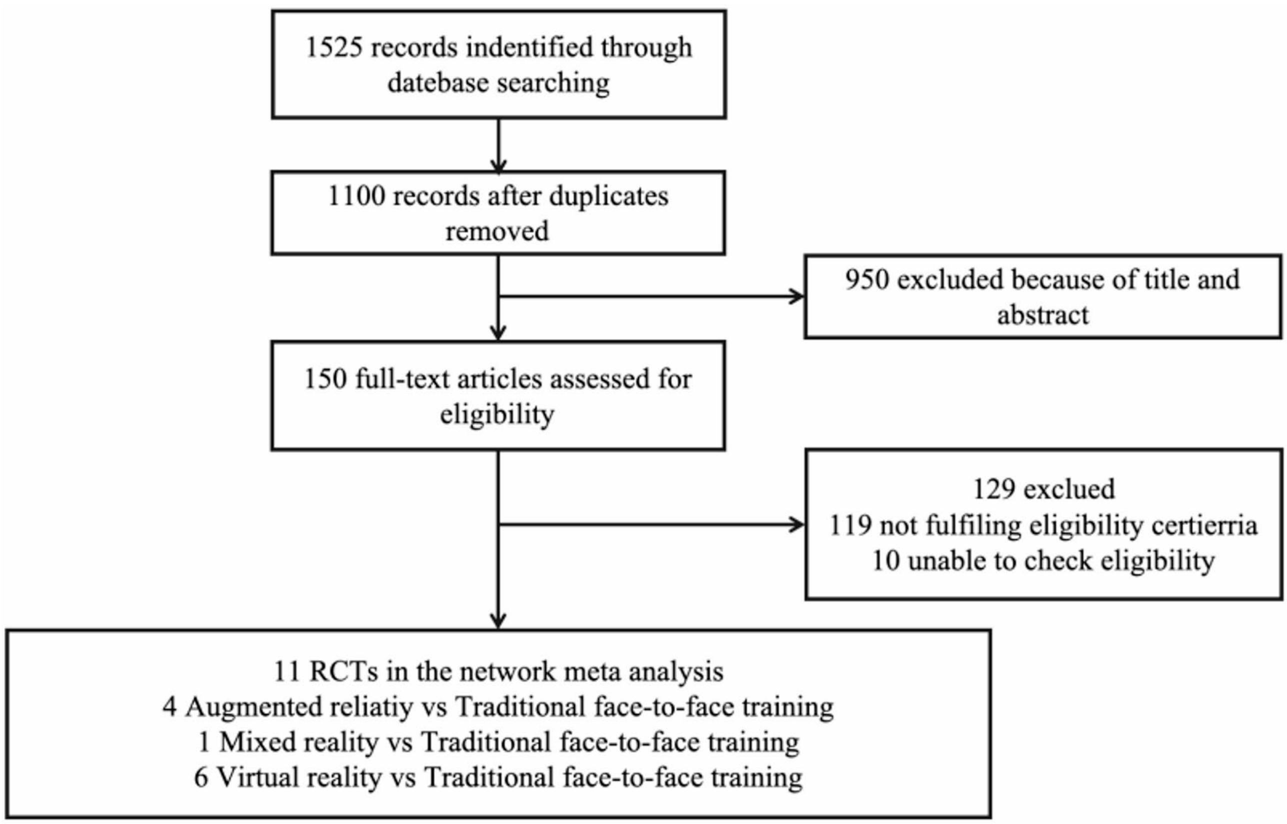

Fig. 1 Flow chart for study selection

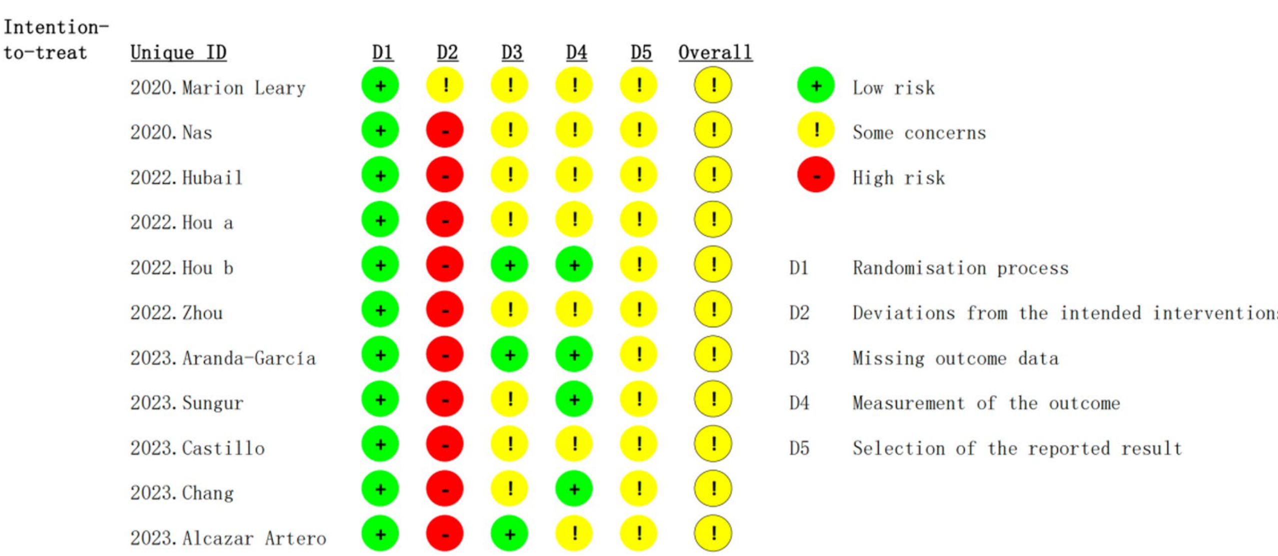

Fig. 2 Detailed analysis of the risk of bias for each included study using the cochrane collaboration's Tool

Depth of chest compression

Ten studies [7, 15, 17–24] (with 1208 participants) reported the depth of chest compression during CPR examination. Heterogeneity analysis indicated high heterogeneity (I^2 = 87.2%, $P<0.0001$), prompting the use of a random-effects model for analysis. As consistency analysis could not be conducted, both consistency and inconsistency models were applied. After 50,000 iterations, the Potential Scale Reduction Factor (PSRF) values approached 1, indicating good model convergence. A visual inspection of the results showed that the effect sizes and confidence intervals were consistent across the consistency and inconsistency models, suggesting good consistency. A total of 4 nodes were included in this NMA, with each node representing a different CPR training method (AR, VR, MR, and traditional training as CONTROL). The width of each line represents the number of direct comparisons between interventions; the analysis results are shown in Fig. 3.

The results indicated that both MR and CONTROL outperformed VR (MR: SMD = 10.96, 95% CI: 0.95 to 20.82; Control: SMD = 4.38, 95% CI: 0.45 to 8.31), while no statistically significant differences were observed among the other comparisons. According to SUCRA rankings, MR (87.4%) was the highest, followed by AR (10.2%), Control (2.3%), and VR (0.2%).

To evaluate potential publication bias, Egger's regression test was performed. The test result showed no evidence of funnel plot asymmetry (t=-0.19, df = 9, $p = 0.8515$). The bias estimate was − 0.4300 with a standard error (SE) of 2.2321, which is close to zero, indicating symmetry. These findings suggest that there is no significant publication bias in the included studies. Additionally, the visual inspection of the funnel plot further supported the absence of substantial asymmetry.

The evidence quality of the network meta-analysis was evaluated using the CINeMA tool. For the comparison between Control and AR (4 studies), most dimensions showed no significant concerns, although some imprecision and incoherence were noted, leading to a high confidence rating. For the comparison between Control and VR (6 studies), certain within-study bias, major heterogeneity, and minor incoherence resulted in a moderate confidence rating. The comparison between Control and MR (1 study) was rated as moderate due to limited sample size and heterogeneity concerns. Comparisons involving AR vs. VR, AR vs. MR, and VR vs. MR lacked direct evidence; while indirectness was not a major concern, issues with heterogeneity and incoherence limited the reliability of the evidence, leading to moderate or high confidence ratings. Overall, the results indicate high evidence quality for some comparisons, but the reliability of certain comparisons was constrained by limited sample sizes and indirectness.

Rate of chest compression

Nine studies [7, 15, 17–22, 24] (with 967 participants) reported the rate of chest compression during CPR

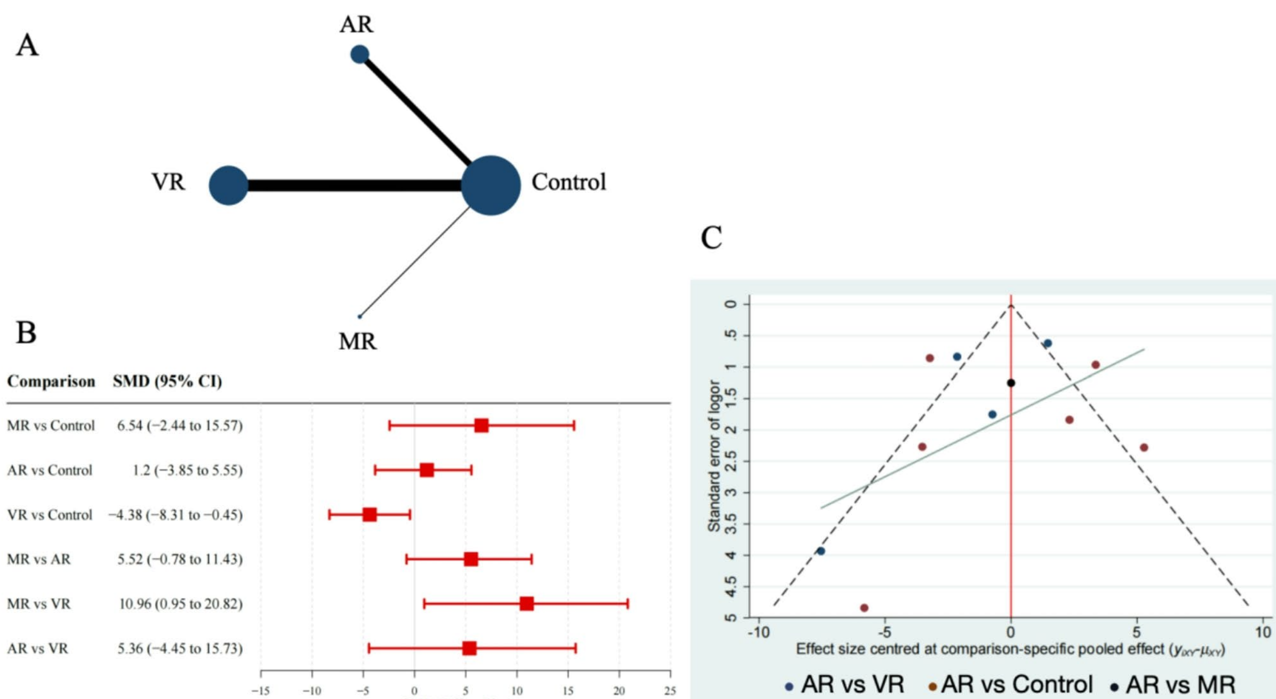

Fig. 3 Network meta-analysis for chest compression depth. (**A**) The network plot shows the depth of chest compression during CPR examination between nodes (blue circles). Each node represents a XR training methods or Control. The size of each node represents the included trainees, and the width of each line represents the number of direct comparisons between training methods. The connecting line noted the number of trial-level comparisons between the two nodes. (**B**) Forest plot showing the standardized mean differences (SMD) for comparisons of interventions based on the compression depth outcome. The plot displays the SMD effect sizes and their 95% confidence intervals, with comparisons ordered by the probability of being the most effective: MR (87.4%) > AR (10.2%) > Control (2.3%) > VR (0.2%). The red squares represent the point estimates, scaled according to their precision, while the horizontal lines indicate the 95% confidence intervals. The vertical red line at 0 represents no difference between the compared interventions. (**C**) The funnel plot, showing all studies spreading nearly evenly on both sides of the average, indicated no obvious publication bias

examination. Heterogeneity analysis indicated high heterogeneity (I^2 = 95.6%, $P<0.0001$), prompting the use of a random-effects model for analysis. The consistency and inconsistency model analyses indicated good convergence, with consistent effect sizes and confidence intervals, suggesting strong overall consistency. A total of 4 nodes were included in this NMA, with each node representing a different CPR training method (AR, VR, MR, and CONTROL). The width of each line represents the number of direct comparisons between interventions; the analysis results are shown in Fig. 3. The width of each line represents the number of direct comparisons between interventions; the analysis results are shown in Fig. 4.

No significant differences were found among any of the comparisons. Based on SUCRA rankings, MR (51.1%) was the highest, followed by AR (14.7%), Control (25.3%), and VR (8.9%).

To assess the potential for publication bias, Egger's regression test was conducted. The results showed no evidence of funnel plot asymmetry (t = 0.11, df = 8, $p = 0.9189$). The bias estimate was 0.3219 with a standard error (SE) of 3.0647, indicating that the intercept was close to zero. Additionally, visual inspection of the funnel plot revealed a symmetrical distribution of studies. These findings suggest that there is no significant publication bias, and the results of the meta-analysis are likely robust.

The evidence quality of the network meta-analysis was assessed using the CINeMA tool. For the comparison between Control and AR (4 studies), no significant concerns were noted regarding within-study bias, reporting bias, indirectness, or heterogeneity; however, some imprecision and incoherence resulted in a moderate confidence rating. The comparison between Control and VR (5 studies) also showed no major bias, and despite some imprecision and incoherence, the evidence quality was rated as high. For the comparison between Control and MR (1 study), limited sample size and concerns about imprecision and incoherence led to a moderate confidence rating. The comparisons involving AR vs. VR, AR vs. MR, and VR vs. MR lacked direct evidence but were rated as either high or moderate due to low risk of bias and good indirectness, despite some imprecision and incoherence.

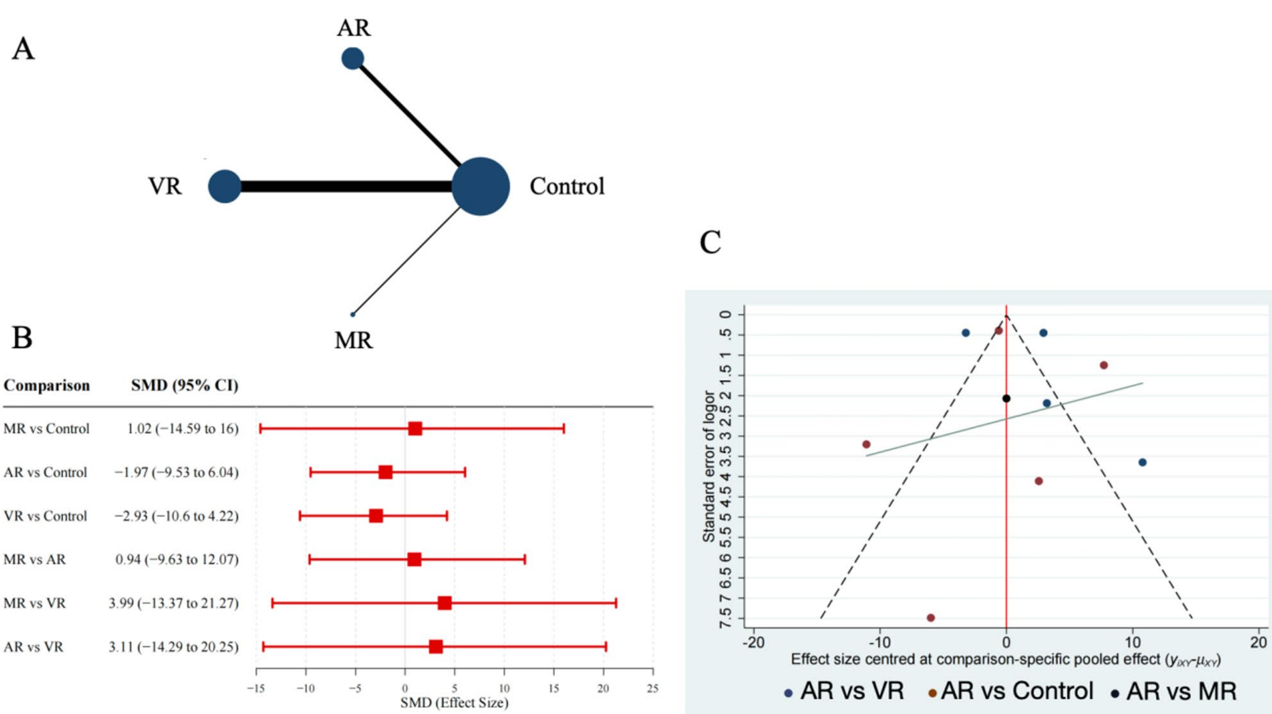

Fig. 4 Network meta-analysis for chest compression rate. (**A**) The network plot shows the chest compression rate during CPR examination between nodes (blue circles). Each node represents a XR training methods or Control. The size of each node represents the included trainees, and the width of each line represents the number of direct comparisons between training methods. The connecting line noted the number of trial-level comparisons between the two nodes. (**B**) Forest plot showing the standardized mean differences (SMD) for comparisons of interventions based on the compression rate outcome. The plot displays the SMD effect sizes and their 95% confidence intervals. Comparisons are ordered by the probability of being the most effective (SUCRA ranking): MR (51.1%) > AR (14.7%) > Control (25.3%) > VR (8.9%). Red squares represent point estimates, scaled according to their precision, while horizontal lines indicate the 95% confidence intervals. The vertical red line at 0 represents no difference between the compared interventions. (**C**) The funnel plot, showing all studies spreading nearly evenly on both sides of the average, indicated no obvious publication bias

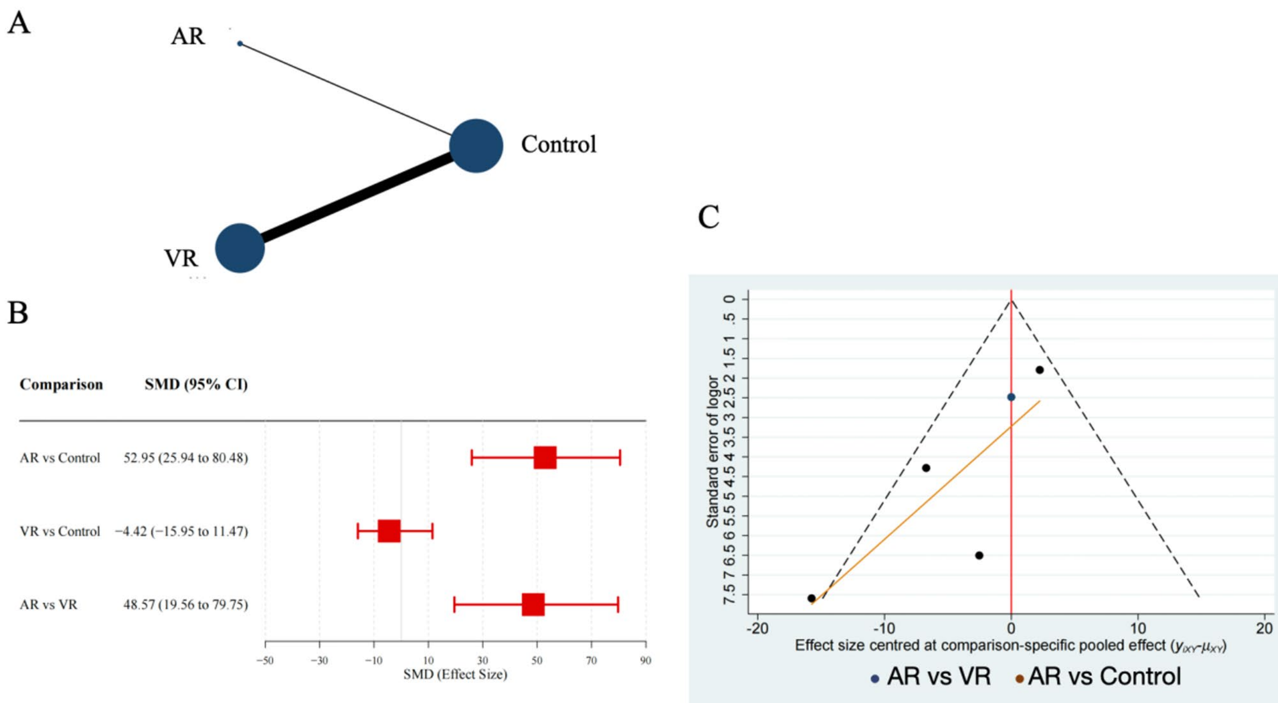

Fig. 5 Network meta-analysis for full chest wall recoil rate. (**A**) The network plot shows the rate of full chest wall recoil during CPR examination between nodes (blue circles). Each node represents a XR training methods or Control. The size of each node represents the included trainees, and the width of each line represents the number of direct comparisons between training methods. The connecting line noted the number of trial-level comparisons between the two nodes. (**B**) Forest plot showing the standardized mean differences (SMD) for comparisons of interventions based on the rate of full chest wall recoil outcome. The plot illustrates the SMD effect sizes and their 95% confidence intervals. Comparisons are ranked by probability of effectiveness (SUCRA ranking): AR (99.1%) > Control (0.07%) > VR (0.02%). The red squares represent the point estimates, scaled according to their precision, while the horizontal lines indicate the 95% confidence intervals. The vertical red line at 0 represents no difference between the compared interventions. (**C**) The funnel plot illustrating potential publication bias among the included studies

Fully recoil rate of chest compression

Five studies [15, 16, 21–23] (with 724 participants) reported the *fully recoil rate* of chest compression during CPR examination. Heterogeneity analysis indicated high heterogeneity (I^2 = 80.9%, $P = 0.0013$), prompting the use of a random-effects model for analysis. The consistency and inconsistency model analyses indicated good convergence, with consistent effect sizes and confidence intervals, suggesting strong overall consistency. A total of 3 nodes were included in this NMA, with each node representing a different CPR training method (AR, VR, and CONTROL). The width of each line represents the number of direct comparisons between interventions; the analysis results are shown in Fig. 5.

The results indicated that AR was superior to both VR (SMD: 48.57, 95% CI: 19.56–79.75) and CONTROL (SMD: 52.95, 95% CI: 25.94–80.48). Based on the SUCRA rankings, AR ranked highest at 99.1%, followed by CONTROL at 0.07%, and VR at 0.02%.

Egger's regression test showed no significant evidence of funnel plot asymmetry ($t = 0.32$, $df = 3$, $p = 0.765$). The bias estimate was 1.4532 with a standard error (SE) of 4.532, indicating that the intercept was close to zero. Visual inspection of the funnel plot also revealed no apparent asymmetry. These findings suggest no strong evidence of publication bias.

The quality of evidence in the network meta-analysis was assessed using the CINeMA tool. For the comparison between Control and AR (1 study), no significant concerns were observed regarding within-study bias, reporting bias, indirectness, imprecision, or heterogeneity, with only minor incoherence noted, resulting in a high confidence rating. The comparison between Control and VR (4 studies) showed good performance in terms of bias and indirectness but concerns about heterogeneity and incoherence led to a moderate confidence rating. For the comparison between AR and VR (no direct evidence available), good indirectness and low risk of bias were observed, with only minor incoherence, resulting in a high confidence rating.

Discussion

This study systematically evaluated the effectiveness of three extended reality technologies—augmented Reality, Virtual Reality, and Mixed Reality—in CPR training through a network meta-analysis, comparing them with traditional face-to-face training methods. Distinguishing from previous studies, this study focused on the comparative advantages and limitations of these three XR technologies against conventional training approaches, providing valuable insights for optimizing BLS education and training strategies. In the CPR standards recommended by the AHA, key quality control indicators include compression depth (5–6 cm), compression rate (100–120/min), and full chest wall recoil [25, 26]. These indicators are strongly associated with the return of spontaneous circulation (ROSC) and represent the core training components in traditional BLS training [27]. The results demonstrate that XR technologies offer certain advantages in improving CPR performance, particularly in critical metrics such as compression depth and full chest wall recoil.

Recent systematic reviews and meta-analyses on extended reality in medical education often either focus on single XR modalities, such as AR or VR, or group them together without distinguishing among different extended reality forms [6]. Moreover, very few studies have included MR in comparative evaluations. In contrast, our network meta-analysis provides a more comprehensive perspective by assessing AR, VR, and MR side by side, along with traditional face-to-face methods. This broader approach highlights differences in performance outcomes-most notably, the superior effect of MR on compression depth-which earlier reviews could not detect due to their narrower scope. Additionally, whereas some recent studies emphasize XR's effectiveness for either theoretical learning (e.g., anatomical concepts [28]) or single-skill acquisition (e.g., surgical suturing) [29], our work underscores its potential across multiple key indicators of CPR quality. Consequently, the nuanced insights provided here not only align with the general consensus that XR-based interventions can enhance medical training but also extend the literature by delineating the comparative advantages of each XR modality for improving CPR outcomes [30].

In this study, MR demonstrated a clear advantage in terms of chest compression depth, achieving the highest SUCRA ranking (87.1%), significantly outperforming AR, VR, and traditional training methods. This finding suggests that MR technology, by integrating virtual and real-world environments, provides a more realistic training experience and real-time feedback, enabling participants to better meet the depth requirements for chest compression [7, 31]. Additionally, MR allows learners to interact with both virtual content and their physical surroundings, making skill acquisition more intuitive and natural. This aligns with existing research on the advantages of MR in medical skills training. For example, Isidre et al. demonstrated that the use of MR in neurosurgical training improved medical students' satisfaction and engagement [32]. Similarly, Sivananthan et al. and Minty et al. explored the application of MR in clinical bedside teaching and Objective Structured Clinical Examinations (OSCEs), both of which reported good usability and positive outcomes [33, 34].

In the comparison of full chest wall recoil rates, AR demonstrated a significant advantage, with a SUCRA ranking nearly reaching 99.1%, far surpassing other training methods. Full chest wall recoil is a critical component of high-quality CPR, as incomplete chest wall recoil may affect resuscitation outcomes by increasing intrathoracic pressure and reducing venous return, coronary perfusion pressure, and myocardial blood flow [2]. By overlaying digital information onto real-world scenarios, AR provides learners with immediate feedback and operational guidance, significantly enhancing performance on this key indicator [35]. This finding supports the potential value of AR as a tool for reinforcing critical details in CPR training.

It is worth noting that VR performed worse than AR and MR in this study across various outcomes. This may be due to the lack of real physical feedback in VR, which makes it difficult for participants to accurately gauge compression force, thereby affecting the quality of their performance [36]. Additionally, the current limitations in the realism of VR simulations, especially in complex scenarios, may impact the trainees' attention and engagement [37]. Future advancements in VR technology should focus on enhancing tactile feedback and improving the integration with real-world scenarios to further enhance its training effectiveness.

The CINeMA analysis provided additional insights into the quality of evidence across the comparisons. Most comparisons were rated as having high or moderate confidence. Moderate ratings were largely attributed to imprecision and heterogeneity, such as the limited sample size of MR studies and variability in training scenarios across included trials. The lack of direct evidence for certain comparisons, such as AR vs. VR, further underscores the need for higher-quality RCTs to address these gaps.

Overall, this study confirms the effectiveness of XR technologies, particularly MR and AR, in CPR training and highlights their promising potential in CPR education. Compared to traditional face-to-face training, XR technologies not only offer more flexible and scalable training models but also improve learners' skill mastery and performance quality through immersive experiences and real-time feedback.

Limitations

However, this study has some limitations. First, there is heterogeneity across the included studies in terms of sample characteristics, training scenarios, and assessment methods, which may impact the robustness of the results. The high heterogeneity observed ($I^2 > 80\%$) reflects the diversity in XR implementation approaches and participant characteristics. Although we employed a Bayesian framework with minimally informative priors to allow the data to primarily drive our results, the substantial between-study variability necessitates caution in interpretation. Although we used a random-effects model and consistency analysis to minimize bias, further high-quality RCTs are needed to validate these findings. Second, this study primarily focused on short-term training effects and did not assess long-term skill retention or clinical outcomes, which are critical for evaluating the practical value of XR technologies in real-world emergency settings. Additionally, due to the differences in participant numbers and backgrounds across various studies, it was challenging to conduct meaningful subgroup analyses based on participants' training levels; while we recognize that training level could be an important effect modifier, the limited number of studies in each potential subgroup prevented us from conducting formal subgroup analyses to determine differential effects across training expertise levels. Moreover, only one study focused on MR, which may also limit the generalizability of its conclusions.

Future directions

The advancement of AI technology will open new avenues for applying XR to CPR training. Existing devices, such as Ray-Ban Meta and RayNeo glasses, capture images via an onboard camera and integrate them with AI large language models, enabling the system to analyze visual input and deliver feedback through speakers in the glasses' frames. This approach not only grants greater autonomy to CPR training but also holds potential for providing real-time feedback to rescuers during actual resuscitation, thereby improving CPR quality.

Conclusion

This study is the first to compare the effectiveness of three XR technologies with traditional CPR training methods. The results confirm the advantages of MR and AR in improving key CPR metrics like compression depth and full chest wall recoil rate. Further optimization of XR technology may improve its significant potential in enhancing CPR training outcomes and in practical applications within real-world emergency settings.

Abbreviations

CPR	Cardiopulmonary resuscitation
IHCA	In-hospital cardiac arrest
OHCA	Out-of-hospital cardiac arrest
ROSC	Return of spontaneous circulation
XR	Extended Reality
AR	Augmented Reality
VR	Virtual Reality
MR	Mixed Reality
HMDs	Head-mounted displays
NMA	Network meta-analysis
RCTs	Randomized controlled trials
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
CIs	Confidence intervals
RRs	Risk ratios
MDs	Mean differences
PSRF	Potential scale reduction factor
SUCRA	Surface under the cumulative ranking curve
CINeMA	Confidence in Network Meta-Analysis
OSCEs	Objective Structured Clinical Examinations

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12873-025-01256-2.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements
Not applicable.

Author contributions
XkW and XbY: conception and design of the paper and literature search. XbY and GqH: data collection and chatbot input. XkW and XmL contributed to the writing, revising, and submission of the final draft.

Funding
This work was supported by Peking Union Medical Foundation Ruiyi Emergency Medical Research Fund (Grant No.22322012016).

Data availability
The datasets generated and/or analyzed during this study are not publicly available due to data privacy imposed by the Xiangya Hospital, Central South University. However, the corresponding author may provide the data upon reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 24 January 2025 / Accepted: 29 May 2025
Published online: 07 June 2025

References

1. Varghese J, Thomas A, Prasad BS, Trikkur S, Nair S, Kumar G, Mohan N, Theresa MM. Improving cardiopulmonary resuscitation quality and resuscitation training by using Real-Time Audiovisual-Assisted devices. Cureus. 2024;16(9). https://doi.org/10.7759/cureus.68757.
2. Meaney PA, Bobrow BJ, Mancini ME, Christenson J, de Caen AR, Bhanji F, Abella BS, Kleinman ME, Edelson DP, Berg RA, Aufderheide TP, Menon V, Leary M. Cardiopulmonary resuscitation quality: improving cardiac resuscitation outcomes both inside and outside the hospital. Circulation Am Heart Association. 2013;128(4):417–35. https://doi.org/10.1161/CIR.0b013e31829d8654.
3. Einspruch EL, Lynch B, Aufderheide TP, Nichol G, Becker L. Retention of CPR skills learned in a traditional AHA heartsaver course versus 30-min video self-training: A controlled randomized study. Resusc Elsevier. 2007;74(3):476–86. https://doi.org/10.1016/j.resuscitation.2007.01.030.

4. Hsieh M-J, Bhanji F, Chiang W-C, Yang C-W, Chien K-L, Ma MH-M. Comparing the effect of self-instruction with that of traditional instruction in basic life support courses—A systematic review. Resuscitation. 2016;108:8–19. https://doi.org/10.1016/j.resuscitation.2016.08.021.
5. Nabecker S, Huwendiek S, Theiler L, Huber M, Petrowski K, Greif R. The effective group size for teaching cardiopulmonary resuscitation skills– A randomized controlled simulation trial. Resusc Elsevier. 2021;165:77–82. https://doi.org/10.1016/j.resuscitation.2021.05.034.
6. Sun R, Wang Y, Wu Q, Wang S, Liu X, Wang P, He Y, Zheng H. Effectiveness of virtual and augmented reality for cardiopulmonary resuscitation training: a systematic review and meta-analysis. BMC Med Educ BioMed Cent. 2024;24(1):1–12. PMID:38970090.
7. Sungur H, van Berlo ZMC, Lüwa LM. Enhancing cardiopulmonary resuscitation training with mixed reality: improving cardiopulmonary resuscitation performance and enjoyment. Cyberpsychology Behav Soc Netw. 2024;27(6):379–86. PMID:38330413.
8. Kuyt K, Park SH, Chang TP, Jung T, MacKinnon R. The use of virtual reality and augmented reality to enhance cardio-pulmonary resuscitation: a scoping review. Adv Simul Adv Simul. 2021;6(1):4–11. https://doi.org/10.1186/s41077-021-00158-0.
9. Popov V, Mateju N, Jeske C, Lewis KO. Metaverse-based simulation: a scoping review of charting medical education over the last two decades in the lens of the marvelous medical education machine. Ann Med Engl. 2024;56(1):2424450. PMID:39535116.
10. Cheng A, Fijacko N, Lockey A, Greif R, Abelairas-Gomez C, Gosak L, Lin Y. Use of augmented and virtual reality in resuscitation training: A systematic review. Resusc Plus Authors. 2024;18(March):100643. https://doi.org/10.1016/j.resplu.2024.100643.
11. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JPA, Straus S, Thorlund K, Jansen JP, Mulrow C, Catalá-López F, Gøtzsche PC, Dickersin K, Boutron I, Altman DG, Moher D. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med United States. 2015;162(11):777–84. PMID:26030634.
12. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med Engl. 2010;29(7–8):932–44. PMID:20213715.
13. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. J Comput Graph Stat ASA Website. 1998;7(4):434–55. https://doi.org/10.1080/10618600.1998.10474787.
14. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, Thomas J. Updated guidance for trusted systematic reviews: a new edition of the cochrane handbook for systematic reviews of interventions. Cochrane Database Syst Rev. 2019;10(10):ED000142. https://doi.org/10.1002/14651858.ED000142. PMID: 31643080; PMCID: PMC10284251.
15. Hubail D, Mondal A, Al Jabir A, Patel B. Comparison of a virtual reality compression-only cardiopulmonary resuscitation (CPR) course to the traditional course with content validation of the VR course– A randomized control pilot study. Ann Med Surg Elsevier Ltd. 2022;73(December 2021):103241. https://doi.org/10.1016/j.amsu.2022.103241.
16. Chang YT, Wu KC, Yang HW, Lin CY, Huang TF, Yu YC, Hu YJ. Effects of different cardiopulmonary resuscitation education interventions among university students: A randomized controlled trial. PLoS ONE. 2023;18(3 March):1–12. PMID:36917584.
17. Alcázar Artero PM, Greif R, Cerón Madrigal JJ, Escribano D, Pérez Rubio MT, Alcázar Artero ME, López Guardiola P, Mendoza López M, Melendreras Ruiz R. Pardo Ríos M. Teaching cardiopulmonary resuscitation using virtual reality: A randomized study. Australas Emerg Care. 2024;27(1):57–62. https://doi.org/10.1016/j.auec.2023.08.002.
18. Zhou Y, Zhang W, Wang Y, Zhu X, Sheng H, Mao ECE. A study of comparison between virtual reality-based helicopter air medical training methods with traditional training methods. Chin J Emerg Resusc Disaster Med. 2022;17(7):963–7.
19. Hou L, Dong X, Li K, Yang C, Yu Y, Jin X, Shang S. Effectiveness of a novel augmented reality cardiopulmonary resuscitation self-training environment for laypeople in china: a randomized controlled trial. Interdiscip Nurs Res. 2022;1(1):43–50. https://doi.org/10.1097/nr9.0000000000000010.
20. Hou L, Dong X, Li K, Yang C, Yu Y, Jin X, Shang S. Comparison of augmented Reality-assisted and Instructor-assisted cardiopulmonary resuscitation: A simulated randomized controlled pilot trial. Clin Simul Nurs. 2022;68:9–18. https://doi.org/10.1016/j.ecns.2022.04.004.
21. Aranda-García S, Otero-Agra M, Berlanga-Macías C, Rodríguez-Núñez A, Barcala-Furelos R, Domingo J, Seijas-Vijande A, Fernández-Méndez F. New communication tool for basic life support training: smart glasses. A quasi-experimental study. Med Intensiva (English Ed. 2024;48(2):77–84. https://doi.org/10.1016/j.medine.2023.10.011.
22. Nas J, Thannhauser J, Vart P, Van Geuns RJ, Muijsers HEC, Mol JQ, Aarts GWA, Konijnenberg LSF, Gommans DHF, Ahoud-Schoenmakers SGAM, Vos JL, Van Royen N, Bonnes JL, Brouwer MA. Effect of Face-to-Face vs virtual reality training on cardiopulmonary resuscitation quality: A randomized clinical trial. JAMA Cardiol. 2020;5(3):328–35. PMID:31734702.
23. Castillo J, Rodríguez-Higueras E, Belmonte R, Rodríguez C, López A, Gallart A. Efficacy of virtual reality simulation in teaching basic life support and its retention at 6 months. Int J Environ Res Public Health 2023;20(5). PMID:36901106.
24. Leary M, McGovern SK, Balian S, Abella BS, Blewer AL. A pilot study of CPR quality comparing an augmented reality application vs. a standard Audio-Visual feedback manikin. Front Digit Heal. 2020;2(February):1–8. https://doi.org/10.3389/fdgth.2020.00001.
25. Olasveengen TM, Mancini ME, Perkins GD, Avis S, Brooks S, Castrén M, Chung SP, Considine J, Couper K, Escalante R, Hatanaka T, Hung KKC, Kudenchuk P, Lim SH, Nishiyama C, Ristagno G, Semeraro F, Smith CM, Smyth MA, Vaillancourt C, Nolan JP, Hazinski MF, Morley PT. Adult basic life support: 2020 international consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations. Circulation United States. 2020;142(16suppl1):S41–91. PMID:33084391.
26. Crowley C, Salciccioli J, Wang W, Tamura T, Kim EY, Moskowitz A. The association between mechanical CPR and outcomes from in-hospital cardiac arrest: an observational cohort study. Resusc Irel. 2024;198:110142. PMID:38342294.
27. Murphy TW, Windermere S, Morris T, Slish J, Holtzman L, Becker TK. Risk and ROSC - legal implications of bystander CPR. Resuscitation. 2020;151:99–102. https://doi.org/10.1016/j.resuscitation.2020.03.017. Epub 2020 Apr 4. PMID: 32259608.
28. Bogomolova K, van der Ham IJM, Dankbaar MEW, van den Broek WW, Hovius SER, van der Hage JA, Hierck BP. The effect of stereoscopic augmented reality visualization on learning anatomy and the modifying effect of Visual-Spatial abilities: A Double-Center randomized controlled trial. Anat Sci Educ United States. 2020;13(5):558–67. PMID:31887792.
29. Toni E, Toni E, Fereidooni M, Ayatollahi H. Acceptance and use of extended reality in surgical training: an umbrella review. Syst Rev Engl. 2024;13(1):299. PMID:39633499.
30. 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):937–58. https://doi.org/10.1007/s10648-020-09586-2.
31. Choi MJ, Kim KJ. Effects of team-based mixed reality simulation program in emergency situations. PLoS ONE. 2024;19(2 February):1–13. PMID:38422080.
32. Silvero Isidre A, Friederichs H, Müther M, Gallus M, Stummer W, Holling M. Mixed reality as a teaching tool for medical students in neurosurgery. Med (B Aires). 2023;59(10). https://doi.org/10.3390/medicina59101720.
33. Sivananthan A, Gueroult A, Zijlstra G, Martin G, Baheerathan A, Pratt P, Darzi A, Patel N, Kinross J. Using mixed reality headsets to deliver remote bedside teaching during the COVID-19 pandemic: feasibility trial of hololens 2. JMIR Form Res. 2022;6(5):e35674. https://doi.org/10.2196/35674.
34. Minty I, Lawson J, Guha P, Luo X, Malik R, Cerneviciute R, Kinross J, Martin G. The use of mixed reality technology for the objective assessment of clinical skills: a validation study. BMC Med Educ BioMed Cent. 2022;22(1):1–8. PMID:35999532.
35. Desselle MR, Brown RA, James AR, Midwinter MJ, Powell SK, Woodruff MA. Augmented and virtual reality in surgery. Comput Sci Eng. 2020;22(3):18–26. https://doi.org/10.1109/MCSE.2020.2972822.
36. Ruthenbeck GS, Reynolds KJ. Virtual reality for medical training: the state-of-the-art. J Simul Taylor Francis. 2015;9(1):16–26. https://doi.org/10.1057/jos.2014.14.
37. Baniasadi T, Ayyoubzadeh SM, Mohammadzadeh N. Challenges and practical considerations in applying virtual reality in medical education and treatment. Oman Med J Oman. 2020;35(3):e125. PMID:32489677.

Publisher's note