Impact of simulation-based education on the performance assessment, knowledge retention and mentality of nursing students: A systematic reviews and meta-analysis

Yun Xu  
Nanjing University

Yue Xu  
Nanjing University of Chinese Medicine

Qing Wang  
Nanjing University of Chinese Medicine

Shizheng Du  
Nanjing University of Chinese Medicine

Xing Jiang  
Nanjing University of Chinese Medicine

Guihua Xu (✉ Guihua.xu@njucm.edu.cn)  
Nanjing University of Chinese Medicine

Research Article

Keywords: simulation-based education, performance assessment, systematic review, meta-analysis

DOI: https://doi.org/10.21203/rs.3.rs-464093/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background: Simulation-based education is a new type of teaching method that is suitable for clinic-related majors. This article aims to determine its impact on clinical skill assessment, performance maintenance, and mentality of nursing students.

Methods: After searching PubMed, Embase, Web of Science, The Cochrane Library, MEDLINE, and EBSCO database, we conducted a systematic collection of randomized controlled trials (RCTs) on the impact of simulation-based education on performance evaluation and mentality of nursing students. The retrieval time limit was from the establishment of the database to March 12, 2021. Meta-analysis was performed using RevMan5.4 software, and two researchers used ROB2.0 software to evaluate the risk of bias in the included literature. The quality evaluation of the outcome indicators was performed using GRADEpro. A standardized mean difference (SMD) with a 95% confidence interval (CI) was used in estimating the pooled effects of RCTs.

Results: A total of 21 RCTs were performed, including 1683 nursing students. The results of meta-analysis showed that simulation-based education significantly improved the assessment scores of nursing students (SMD = 1.46, 95% CI = 1.02 to 1.90, \( P < 0.00001 \)), self-confidence (SMD = 1.19, 95% CI = 0.48 to 1.90, \( P = 0.001 \)), satisfaction (SMD = 0.86, 95% CI = 0.13 to 1.60, \( P = 0.02 \)), and knowledge retention (SMD = 1.89, 95% CI = 0.76 to 2.87, \( P = 0.0008 \)), and the difference was statistically significant. The results of subgroup analysis showed that long-term intervention (MD = 1.44, 95% CI = 0.62, 2.26, \( P = 0.0006 \)) and short-term intervention (MD = 1.46, 95% CI = 0.95, 1.97, \( P < 0.00001 \)) improved the performance assessment of nursing students, and the difference was statistically significant.

Conclusions: Simulation-based education can significantly improve the assessment scores, self-confidence, and satisfaction of the nursing students, and the scores can be effectively retained for a period of time. Limited by the quality of the number of included studies, the above conclusions need to be verified by high-quality studies.

Background

Clinical skills are essential to nursing teaching, which is directly related to the skill levels of nursing students. High-quality and effective classroom education is a prerequisite for cultivating nursing students’ skills. Skill teaching programs for nursing students had been mostly based on the traditional teaching of “teacher demonstration–student imitation–teacher guidance,” which had the problems of single method, lack of context, lack of concreteness, and low professional experience. To meet the demand for high-quality nursing services, practical training and teaching models have become an educational consensus [1]. In recent years, clinical simulation education has developed rapidly and has been widely used in clinical teaching programs for nursing students [2, 3]. Simulation-based education refers to the simulation of real clinical work scenes through close integration with multimedia technology, computer programming technology, modern electronic communication technology, and other technical means. Providing medical students with a risk-free and standardized learning environment help them in improving their clinical abilities [4]. The advantage of this approach is that it shortens the clinical practice learning curve of medical students, provides them with a risk-free, standardized, and error-free learning environment and standardized, consistent, and rich learning and training materials, thereby solving the problem of insufficient teaching resources and problems in standardized assessment [5]. Moreover, the teaching effect and quality of teaching are improved to a certain extent. It has a significant effect on students’ assessment scores [6], knowledge retention [7], self-confidence [8], and satisfaction [9]. Therefore, this study refers to the AMSTAR 2.0 entry requirements [10] for systematic reviews/meta-analysis and reports in accordance with the requirements of PRISMA. The objectives were as follows: to analyze the impact of simulation-based education on the performance assessment and mentality of nursing students and provide evidence of high methodological quality and rigorous reporting quality of simulation-based education and its contribution to the improvement and maintenance of the performance and changes in the mentality of nursing students.

Methods

Study identification and selection

Randomized controlled trials (RCTs) on the impact of simulation-based education on the performance assessment and mentality of nursing students were obtained from PubMed, Embase, Web of Science, The Cochrane Library, MEDLINE, and EBSCO. The search time limit was from database construction until March 12, 2021. The clinical trial registration website (http://www.Clinical Trial.gov and http://www.chictr.org.cn) was also searched, and the references of the included literature were searched to supplement the acquisition of relevant literature. Subject words were combined with free words, and the search terms included terms for intervention (including simulation-based education, simulator, simulation, and manikin), topic (performance assessment, skills, abilities, and effect), learners (education nursing, nursing student, nursing scholar, internship, and residency), and research category (randomized controlled trial, RCT) regardless of whether the allocation concealment or blind method was used. The retrieved articles were analyzed using NoteExpress 3.2.0.7535. In the selection of documents, the title and abstract were read first, and after obviously irrelevant documents were excluded, the full texts were read. Two authors (YX and YX) screened
the reports independently to determine the eligibility of the reports according to the following inclusion criteria: the experimental group adopts simulated teaching or simulator teaching; the control group adopts traditional teaching methods; other intervention measures are required to be consistent between the experimental group and the control group; intervention objects are nursing students, and the students’ grade and gender, and education level are not limited; and outcomes assessment is focused on individual performance. The following exclusion criteria were used: the control group uses a simulator to teach, even though the simulation level of the simulator is lower than that of the experimental group; the intervention objects are nursing staff who receive training in the society after graduation or the intervention objects include other medical majors in addition to nursing students; non-Chinese English literature; the original research data are missing or the data cannot be extracted; and the full text cannot be obtained. Any disagreement was resolved through discussion or negotiation with a third party (QW).

Data extraction

Two researchers (YX and YX) independently screened the literature and extracted and cross-checked the data. Disagreements were resolved through discussion or negotiation with a third party (QW). Information that is not yet determined but is important to this research was obtained by contacting the original research author via email or phone. The extracted data included first author, publication year, country, study design, comparison, sample size, sample source (age and gender), type of simulation, duration, outcome measures, and results. The key elements of bias risk evaluation, outcome indicators, and outcome measurement data of concern were also obtained. Outcome indicators comprised the main outcome indicators: test scores and knowledge retention (that is, the score obtained by measuring the assessment results again at a certain interval after the intervention) and secondary outcome indicators: self-confidence, self-efficacy, satisfaction, and anxiety.

Quality evaluation

Two investigators independently evaluated the risk of bias in the included studies and cross-checked the results. Disagreements were resolved through discussion or negotiation with a third party. The RCT bias risk assessment tool ROB2.0, which was recommended by Cochrane manual, was used in risk of bias assessment, which had six parts: randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, selection of the reported result, and overall bias.

Statistical analysis

RevMan5.4 software was used for statistical analysis. A standardized mean difference (SMD) with a 95% confidence interval (CI) was used in the estimation of the pooled effects of the RCTs. Heterogeneity among the results of the included studies was analyzed with \( \chi^2 \) test (\( \alpha = 0.1 \)), and the degree of heterogeneity was quantitatively evaluated by combining with \( I^2 \). When no statistical heterogeneity was observed among the results of each study, the fixed effects model was used for analysis. When statistical heterogeneity was observed between the results of each study, the source of the heterogeneity was analyzed, and after the influence of obvious clinical heterogeneity was excluded, the random-effects model was used for analysis. The level of meta-analysis was set at 0.05. Obvious clinical heterogeneity was treated through subgroup analysis.

Results

Study characteristics

A total of 2053 related articles were obtained during the initial inspection, and after a layer-by-layer screening, 21 RCTs were finally included [11-31], which included 1683 nursing students. The literature screening process and results are shown in Figure 1. The publication years of the included literature ranges from 2006 to 2020, and five of the studies are conducted in Turkey, five in Jordan, two in the UK, two in China, and two in Singapore. The United States, Brazil, France, Portugal, and Japan have only one article each. The research objects are mainly nursing students from different grades, mainly students in the third grade whose ages range from 10.29 years to 33.00 years. The sample sizes of the studies range from 31 to 146. The detailed information of the 21 studies is shown in Table 1.

Table 1. Characteristics of the included studies.
| Study          | Country | Design                        | Comparison                                                                 | Samples | Sample source          | Age (EG/CG)       |
|---------------|---------|-------------------------------|---------------------------------------------------------------------------|---------|------------------------|-------------------|
| Kahraman et al., 2019 | Turkey | quasiexperimental             | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 72      | third-year nursing students | 21.91±0.86       |
| Basak et al., 2019    | Turkey | RCT                           | simulation-based training group vs. standard training group               | 71      | second-year nursing students | 19.50±0.57       |
| Tamaki et al., 2019    | Japan  | RCT                           | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 38      | third-year nursing students | 21.30±0.62/21.00±0.00 |
| Pamela et al., 2013   | America| pretest-posttest two-group randomized experimental | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 79      | nursing students /      |                   |
| Alinier et al., 2006  | UK     | pretest-posttest two-group randomized experimental | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 99      | second-year nursing students | 29.3±7.5/33.0±8.4 |
| JU et al., 2019       | China  | RCT                           | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 132     | first-year nursing students | /                 |
| Stayt et al., 2015    | UK     | two centre phase II single RCT | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 98      | nursing students        | 26.3±7.0/29.5±8.4 |
| José et al., 2019     | Portugal| RCT                           | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 42      | nursing students        | 19.29±0.46/20.29±2.19 |
| AlAmrani et al., 2017 | Turkey | RCT                           | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 72      | first-year nursing students | 19.69±1.23/20.22±2.17 |
| Loai, 2020            | Jordan | RCT(pre-test-post-test)       | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 76      | nursing students        | 20.36±0.54/20.47±0.55 |
| Aloush, 2019          | Jordan | two-arm RCT with pre-post tests | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 131     | fourth-year nursing students | 21.0±2.1         |
| Loai, 2016            | Jordan | pretest-posttest RCT          | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 84      | nursing students        | 19.60±1.02/19.10±1.05 |
| Ahmad et al., 2015    | Jordan | pretest-posttest RCT          | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 91      | nursing students        | 20.40±0.98        |
| Ahmad et al., 2014    | Jordan | two-arm RCT                   | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 90      | nursing students        | 20.1±1.9/19.6±1.6 |
| Banu et al., 2019     | Turkey | RCT                           | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group | 59      | first-year nursing students | /                 |
| Authors         | Country | Study Design               | Intervention Details                                                                                                                                                                                                                                                                                                                                 | N   | Group Description | Score |
|-----------------|---------|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-------------------|-------|
| Ayse et al., 2019 | Turkey  | mixed-methods              | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group                                                                                                                                                                                                                                                                                     | 107 | third-year nursing students | 20.77±1.22 |
| Li et al., 2019  | China   | pretest-posttest two-group randomized experimental | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group                                                                                                                                                                                                                                                                                     | 100 | nursing students   | 20.60±0.78/20.2±1.00 |
| Sok et al., 2011 | Singapore | pretest-posttest RCT      | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group                                                                                                                                                                                                                                                                                     | 31  | third-year nursing students | 21.45±1.55 |
| Sok et al., 2012 | Singapore | pretest-posttest RCT      | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group                                                                                                                                                                                                                                                                                     | 31  | third-year nursing students | 21.47±1.43/21.44±1.41 |
| Antonia et al., 2020 | France  | prospective multicenter study | simulation-based training group vs. standard training group                                                                                                                                                                                                                                                                                                                                                       | 146 | second-year nursing students | 24.0±6.4/25.0±6.5 |
| Raphael et al., 2020 | Brazil  | pretest-posttest RCT      | Pre-intervention test vs. Post intervention test / simulation-based training group vs. standard training group                                                                                                                                                                                                                                                                                     | 34  | nursing students   | 22.3 |

Description of the intervention in the studies

All interventions in the study are simulation tutorials or trainings, and 13 articles [13, 14, 16-19, 21-26, 31] articles use scenario simulation teaching, which creates hospital-like treatment environments for students. One of the articles [28] uses simulator training, that is, high-level simulators, such as a robotic arm, is used in helping nurse students learn blood pressure measurement techniques. Seven of the articles [11, 12, 15, 20, 27, 29, 30] use a combination of simulators and simulation scenarios for education. The intervention times of most studies are short, mostly 1–2 hours, and only six studies have an intervention period of more than 1 week. The PICO information of 21 studies is shown in Table 2.

Table 2. PICO of the included studies.
| Study                      | Samples(EG/CG) | Male/Female(EG) | Male/Female(CG) | Intervention                                | Duration | Control                        | Outcome measures |
|---------------------------|----------------|----------------|-----------------|--------------------------------------------|----------|-------------------------------|------------------|
| Kahraman et al., 2019     | 36/36          | 6/30           | 3/33            | Scenario simulation                        | 1w       | standard training              |                  |
| Basak et al., 2019        | 35/36          | 0/35           | 0/36            | Scenario simulation and simulator          | 1w       | standard training and curriculum lecture |                  |
| Tamaki et al., 2019       | 20/18          | 0/20           | 1/17            | Scenario simulation                        | 80min    | standard training              |                  |
| Pamela et al., 2013       | 40/39          | 0/40           | 0/39            | Scenario simulation                        | 90min    | standard training              |                  |
| Alinier et al., 2016      | 49/50          | 7/42           | 9/41            | Scenario simulation                        | 6h       | standard training              |                  |
| JU et al., 2019           | 66/66          | 12/54          | 14/52           | Scenario simulation                        | 8w       | standard training              |                  |
| Stayt et al., 2015        | 50/48          | 5/45           | 6/42            | Scenario simulation                        | 1h       | standard training              |                  |
| José et al., 2019         | 21/21          | 0/21           | 2/19            | Scenario simulation and simulator          | 2m       | standard training              |                  |
| AlAmrani et al., 2017     | 36/36          | 12/24          | 14/22           | Scenario simulation and simulator          | 2w       | standard training              |                  |
| Loai, 2020                | 38/38          | 16/22          | 14/24           | Scenario simulation                        | 3m       | standard training              |                  |
| Aloush, 2019              | 65/66          | 59/72          |                 | Scenario simulation                        | 6h       | standard training              |                  |
| Loai, 2016                | 42/42          | 21/21          | 20/22           | Scenario simulation                        | 3h20min  | standard training              |                  |
| Ahmad et al., 2015        | 47/44          | /              |                 | Scenario simulation and simulator          | 40min    | standard training              |                  |
| Ahmad et al., 2014        | 45/45          | 10/35          | 9/36            | Scenario simulation and simulator          | 4h       | standard training              |                  |
| Banu et al., 2019         | 32/27          | /              |                 | simulator                                  | 2d       | standard training              |                  |
| Ayse et al., 2019         | 53/54          | 7/46           | 7/47            | Scenario simulation and simulator          | 4h       | standard training              |                  |
| Li et al., 2019           | 50/50          | 23/27          | 21/29           | Scenario simulation                        | 2w       | standard training              |                  |
| Sok et al., 2011          | 15/16          | 9/22           |                 | Scenario simulation                        | 6h       | standard training              |                  |
| Sok et al., 2012          | 15/16          | 1/14           | 1/15            | Scenario simulation                        | 6h       | standard training              |                  |
| Antonia et al., 2020      | 73/73          | 14/59          | 8/65            | Scenario simulation and serious game       | 2h       | standard training              |                  |
| Raphael et al., 2020      | 17/17          | 7/27           |                 | Scenario simulation                        | 40h      | standard training              |                  |

**Abbreviations:** Test scores: Knowledge retention: Self-confidence: Self-efficacy: Satisfaction: Anxiety.

**Quality of the studies**
The proportion of items assessed for bias risk in the RCTs is shown in Figure 2, and the traffic light chart is shown in Figure 3. The RCT bias risk assessment tool ROB2.0 was used in risk assessment. The risk of bias is shown in Figures 2 and 3. Seven trials (33.33%) have a low risk of bias for the randomization process. All trials have low risk of bias for deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of reported result. Seven trials (33.33%) have low risk of bias for overall bias. All the studies included in this article describe the random process in detail, and the main reason for the high risk of bias in the randomization process is that most of these studies do not mention the process of allocation concealment. A small number of studies use single-blind or double-blind studies, and the research subjects and data analysts are blinded. However, the intervention measures of the research mainly focus on courses and training. The blinding method is unsuitable for most research subjects and interveners. Therefore, we believe that the lack of blinding will not affect the measurement of final outcome indicators.

**Meta-analysis for outcome measures**

In the included studies, a meta-analysis was conducted on the performance and mentality of nursing students after intervention. The data evaluated in each area is summarized below.

**Test scores**

Fifteen RCTs [12-14, 16-22, 25-27, 30, 31] consisting of 1114 nursing students were included. The analysis results of the random-effects model show that the test scores of the simulation-based education group is higher than those of the control group (SMD = 1.46, 95% CI = 1.02–1.90, \( P < 0.00001 \); Figure 4).

**Knowledge retention**

Seven RCTs [20, 22-25, 27, 30] consisting of 507 nursing students were included. The analysis results of the random-effects model show that the degree of knowledge retention of the simulation-based education group is higher than that of the control group (SMD = 1.89, 95% CI = 0.76–2.87, \( P = 0.0008 \); Figure 5).

**Self-confidence**

Nine RCTs [11, 13, 15, 23-26, 28, 31] consisting of 665 nursing students were included. The analysis results of the random-effects model show that the self-confidence of the simulation-based education group is higher than that of the control group (SMD = 1.19, 95% CI = 0.48–1.90, \( P = 0.001 \); Figure 6).

**Self-efficacy**

Three RCTs [21, 28, 30] consisting of 233 nursing students were included. The analysis results of the random-effects model show that the self-efficacy of the simulation-based education group is higher than that of the control group (SMD = 0.24, 95% CI = −0.27–0.74, \( P = 0.36 \); Figure 7).

**Satisfaction**

Two RCTs [29, 30] consisting of 188 nursing students were included. The analysis results of the random-effects model show that the satisfaction of the simulation-based education group is higher than that of the control group (SMD = 0.86, 95% CI = 0.13–1.60, \( P = 0.02 \); Figure 8).

**Anxiety**
Two RCTs [12, 28] consisting of 131 nursing students were included. The analysis results of the random-effects model show that the anxiety of the simulation-based education group is lower than that of the control group (SMD = −0.36, 95% CI = −1.69–0.96, \( P = 0.59 \); Figure 9).

**Subgroup analysis**

The meta-analysis results show that \( I^2 > 50\% \) of many results indicates high heterogeneity between the studies. To determine the source of heterogeneity, we conducted a subgroup analysis based on the length of the intervention. The short-term intervention included six studies [12, 14, 21, 25, 26, 30], and the main intervention time was within a few hours or 2 days. The long-term intervention included 15 studies [11, 13, 15-20, 22-24, 27-29, 31], and the intervention time ranged from 1 week to 3 months. Given that few related studies involve knowledge retention, self-confidence, self-efficacy, satisfaction, and anxiety, we conducted a subgroup analysis on the index of assessment performance. The results are as follows.

**Test scores**

After the studies were grouped according to time, the heterogeneity of the meta-analysis results was significantly reduced. Thus, a random-effects model was used for analysis. After subgroup analysis, the total \( I^2 = 90 \), the heterogeneity is still high, indicating that the main source of heterogeneity has no relationship with intervention measures. In fact, 20 studies [11-27, 29-31] adopt scenario simulation teaching, and the detailed settings of the scenario simulation of each study may be the main source of heterogeneity. Six studies were included [12, 14, 21, 25, 26, 30] in the long-time group. The long-time intervention group is significantly better than the control group in terms of reducing test scores (MD = 1.44, 95% CI = 0.62, 2.26, \( P = 0.0006 \)). Nine studies [13, 16-20, 22, 27, 31] were included in the short-time group. Short-time intervention significantly reduces test scores (MD = 1.46, 95% CI = 0.95, 1.97, \( P < 0.00001 \)). The results are shown in Figure 10.

**GRADEpro evidence assessment**

In the evaluation of GRADEpro’s level of evidence, the meta-analysis results have three pieces of intermediate evidence: one piece of low-level evidence and two pieces of very low-level evidence. The overall level of evidence is low. The main reasons for degradation are inconsistency and imprecision. In terms of inconsistency, most indicators are statistically heterogeneous, that is, \( I^2 > 50\% \), and they are all downgraded to one level. Two main problems were determined in terms of imprecision. On the one hand, the sample sizes of some indicators do not meet the optimal sample size, that is, the total sample size of the total continuous variable is less than 400. On the other hand, the effect sizes of some indicators cross the invalid line, that is, \( P \geq 0.05 \). Through downgrade treatment, both types of indicators have been downgraded two times. No downgrade treatment was performed on the other three aspects of grade evaluation. In terms of limitations, some studies do not use blinding, and the allocation concealment report is insufficient, although they are not downgraded because blinding and allocation concealment have little effect on the experimental results. In terms of indirectness, although the intervention measures are not completely consistent (different types of simulation scenarios), no significant difference was found between PICO and the research question, and thus it is not downgraded. In terms of publication bias, the literature search in this study is comprehensive, the included studies do not involve commercial interests, and no clear evidence of a risk of bias is found. Hence, publication bias will not be downgraded. See Table 3 for details.

**Table 3. Evidence assessment of outcomes.**
### Outcome indicators (number of studies)

| Outcome indicators | Sample size | Limitation | Inconsistency | Indirectness | Imprecision | Publication bias | Effect size | Quality of evidence |
|--------------------|-------------|------------|---------------|---------------|-------------|-----------------|-------------|---------------------|
| Test scores | 560/554 | 0 | -1 | 0 | 0 | 0 | SMD=1.46,95%CI (1.02~1.90) | B:middle-level |
| Knowledge retention | 254/253 | 0 | -1 | 0 | 0 | 0 | SMD=1.81,95%CI (0.76~2.87) | B:middle-level |
| Self-confidence | 334/331 | 0 | -1 | 0 | 0 | 0 | SMD=1.19,95%CI (0.48~1.90) | B:middle-level |
| Self-efficacy | 119/114 | 0 | -1 | 0 | -2 | 0 | SMD=0.24,95%CI (-0.27~0.74) | B:middle-level |
| Satisfaction | 94/94 | 0 | -1 | 0 | -1 | 0 | SMD=0.86,95%CI (0.13~1.60) | B:middle-level |
| Anxiety | 68/63 | 0 | -1 | 0 | -2 | 0 | SMD=-0.36,95%CI (-1.69~0.96) | B:middle-level |

**Abbreviations:** Reason for deduction: Inconsistency: $I^2>50$%; Imprecision: continuous variable, the total sample size of the two groups of research subjects is less than 400 cases or the confidence interval crosses the invalid line.

### Discussion

Through GRADE evaluation, different types of evidence were obtained, which showed that simulation education has a certain contribution to the improvement of the assessment performance, self-confidence, and satisfaction of nursing students. Moderate-quality evidence shows that compared with traditional teaching methods, simulated teaching is more conducive to improving assessment results, self-confidence, and knowledge retention among nursing students. Low-quality evidence shows that students' satisfaction with the curriculum after simulated education is significantly higher than that of students who underwent traditional teaching programs. Evidence with extremely low quality shows that the impact of simulated education on the self-efficacy and anxiety of nursing students is not different from that of traditional education.

The acquisition and application of knowledge is the ultimate goal of teaching. An efficient and real teaching environment can promote students' clinical skills [32]. A number of studies have shown that [33, 34] rich simulation teaching can improve the assessment results of nursing students, enhance their clinical skill levels, and even replace clinical practice. The results of this research showed that compared with traditional teaching methods, simulated education can significantly improve the assessment results of nursing students. Students' scores were retested from 40 days to half a year after the intervention, and the degree of knowledge retention of nursing students in the simulation teaching group was significantly higher than that of the traditional teaching method group. This finding may be related to the use of simulated scenes and simulated instruments in simulation teaching, which enabled students to visualize abstract knowledge and convert them to concrete operations and left a deep impression on them. A study [35] confirmed that scene simulation can help students in memorizing information. Therefore, simulation teaching can be incorporated into teaching because it enables students to apply knowledge in advance, thereby enhancing the mastery rate of knowledge.

Self-confidence is a key factor affecting students' assessment results and clinical skill application. Having a high degree of self-confidence can reduce students' sense of tension in the clinic and enable them to display knowledge that they have learned [36]. The results of this study showed that compared with traditional education methods, clinical simulation education can significantly enhance the self-confidence of nursing students. This may reduce the fear of the unknown environment when actually contacting the patient instead of having been in contact with the simulation scenarios. However, no difference in self-efficacy was observed between the simulated and traditional teaching groups. This result is different from the results of some studies [37], and this difference may be related to the small number of included studies, small sample size, and large heterogeneity. A study [38] showed that students are prone to nervousness, low confidence, anger, and other emotions during the assessment process. These emotions affected their performance. Simulation education can alleviate this psychology. Therefore, strengthening the simulation teaching of trial practice in advance is necessary because this approach allows students to play their actual roles in their respective clinics.
Student satisfaction with the course is an important indicator for evaluating the course and which can reflect the effectiveness and interest of a teaching method. Some studies [39, 40] showed that learning methods and learning environment directly affect learners’ learning interest, learning continuity, and satisfaction. The present study showed that students’ satisfaction with simulation education was higher than that of traditional medical education possibly because simulation teaching is more vivid, interesting, and authentic. Increase in learning satisfaction has a significant effect on willingness to continue learning [41]. A real teaching environment can strengthen people’s learning enthusiasm [35], thereby enhancing their sense of satisfaction. Scenes, teaching aids, and auxiliary teaching methods can increase students’ skills in recognizing possibilities.

Anxiety is a common emotion in the examination process of students. A certain degree of anxiety can promote students’ performance, but excessive anxiety has a considerable adverse impact on student performance [42]. In the present study, no difference among the effects of different education methods on nursing students’ anxiety was found, and this result was different from the results of previous studies [43]. The possible reason was the small number of included studies and the low level of evidence. Thus, sample size should be increased, and the experimental design should be optimized.

To determine the source of heterogeneity, we divided the included literature into two subgroups for in-depth analysis according to the intervention time of simulated education. The intervention time of the long-term intervention group was more than 1 week, whereas the intervention time of the short-term intervention group was within 1 day. The results of the subgroup analysis suggested that regardless of the length of the intervention, the impact of simulation education on the assessment results of nursing students is better than that of traditional teaching. After subgroup analysis, the heterogeneity of the meta-analysis was reduced. These results indicated that intervention time is an important confounding factor in the meta-analysis of the present study.

Strengths and limitations

GRADEpro and RoB2.0 bias risk assessment methods were used in evaluating the quality of the evidence, making the conclusions objective and scientific. For the first time, a meta-analysis was performed on the retention of nursing students’ knowledge through simulation education. Positive results were obtained, which provided support for the effectiveness of simulation research. However, some shortcomings in this research were observed. Knowledge acquisition and retention are closely related to intervention time and test time, but the measurement of knowledge retention in this study cannot exclude other interfering factors. For example, the measurement of the subjects’ performance after a period of time cannot exclude the period. The subjects themselves also consolidated and reviewed their knowledge. The meta-analysis results of this study were mostly heterogeneous possibly because of the different simulation scenarios. Given that most of the interventions had short durations, our subgroup analysis only focused on the index of assessment performance. Observing the heterogeneity of other indicators through subgroup analysis is impossible. The included studies did not focus on the indicators of nursing students’ mood. Satisfaction, self-efficacy, and anxiety are important factors that affect nursing students’ clinical skills [8], and data collection in this area should be strengthened in the follow-up. Owing to the particularity of the interventions in this study, the included studies seldom used blinding, and many studies did not mention the allocation concealment process. Although not downgraded in the present study, randomization and allocation should be further strengthened. Moreover, blinding should be performed to improve the quality of evidence.

Conclusion

Clinical simulation education not only significantly improves the performance of nursing students but also affects their self-confidence in operation and satisfaction with the curriculum. It has a significant effect on the long-term retention of nursing students’ performance. This study analyzed the impacts of simulated education programs with different intervention times on the assessment results of nursing students. The results of the subgroup analysis showed that regardless of the length of intervention time, the scores of nursing students after simulated education were better than those of students who underwent traditional education programs. However, postgraduate simulation education had little effect on the levels of anxiety and self-efficacy of nursing students.

Declarations

Authors’ contributions

YX and YX conceived and designed the study. YX wrote the paper. YX and YX contributed to the critical review and revised the manuscript. YX, YX, QW, SZD, XJ and GHX selected the available articles, extracted the data, and analyze the data. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.
Availability of data and materials

The datasets analysed during the current study are available in the PubMed, Embase, Web of Science, The Cochrane Library, MEDLINE, and EBSCO repository.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable.

Acknowledgements

The author also thanks the editors and reviewers for their detailed and constructive comments which greatly improved the manuscript.

References

1. E D, BJ S, MR G: Using virtual reality in medical education to teach empathy. Journal of the Medical Library Association : JMLA 2018, 106(4):498-500.
2. Fealy S JDHA: The integration of immersive virtual reality in tertiary nursing and midwifery education: a scoping review. Nurse Educ Today 2019, 79:14-19.
3. M B, E C, H P, R S, C H, B L, C T, J P, A T, L N: Simulation for emergency nurses (SIREN): A quasi-experimental study. NURS EDUC TODAY 2018, 68:100-104.
4. CS P, LR R, E Y, N H, JR B, KG W, JT S, RJ M: Acquisition of critical intraoperative event management skills in novice anesthesiology residents by using high-fidelity simulation-based training. ANESTHESIOLOGY 2010, 112(1):202-211.
5. Levine Al DSSA: New York:Springer; 2013.
6. R M, B B, M N: Simulation-based mastery improves nursing skills in BSc nursing students: a quasi-experimental study. BMC nursing 2021, 20(1):10.
7. A M, Y W, MC V, P F, DB J, SD S, GM F, LS F: Long-term knowledge retention following simulation-based training for electrosurgical safety: 1-year follow-up of a randomized controlled trial. Surgical endoscopy 2016, 30(3):1156-1163.
8. TD O, ALR V, PME O, IME O, AS M: Simulation-based Training for Pelvic and Breast Physical Examination: Effect on the Anxiety and Self-confidence of Medical Students. Revista brasileira de ginecologia e obstetricia : revista da Federacao Brasileira das Sociedades de Ginecologia e Obstetricia 2020, 42(11):739-745.
9. CC H, HS K, HC L, HF L, TP C, BO L: Effects of simulation-based learning on nursing students’ perceived competence, self-efficacy, and learning satisfaction: A repeat measurement method. NURS EDUC TODAY 2021, 97:104725.
10. JR C: Editorial Perspective: Time for Another Grading System-From PRISMA to AMSTAR 2. Global spine journal 2020, 10(5):674-675.
11. A A, S S: The impact of high-fidelity simulation on knowledge, critical thinking, and clinical decision-making for the management of pre-eclampsia. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics 2020, 150(3):354-360.
12. B VD, L ZA: The effects of training with simulation on knowledge, skill and anxiety levels of the nursing students in terms of cardiac auscultation: A randomized controlled study. NURS EDUC TODAY 2020, 84:104216.
13. T T, A I, Y Y, M F, M T, Y I, M K, Y Q, M T: The effectiveness of end-of-life care simulation in undergraduate nursing education: A randomized controlled trial. NURS EDUC TODAY 2019, 76:1-7.
14. A K, M G, Ş B, D Z, H U, S AS, HN ÇÖ, Z B: The effect of simulation-based education on childhood epileptic seizure management knowledge, skills, and attitudes of nursing students. Epilepsy & behavior : E&B 2019, 100:106497.
15. T B, A D, E I: The effect of simulation based education on patient teaching skills of nursing students: A randomized controlled study. *Journal of professional nursing : official journal of the American Association of Colleges of Nursing* 2019, 35(5):417-424.

16. SY L, JJ R, A S, KY P: Rescuing A Patient In Deteriorating Situations (RAPIDS): A simulation-based educational program on recognizing, responding and reporting of physiological signs of deterioration. *RESUSCITATION* 2011, 82(9):1224-1230.

17. LC S, C M, B R, S M, T S: Recognizing and managing a deteriorating patient: a randomized controlled trial investigating the effectiveness of clinical simulation in improving clinical performance in undergraduate nursing students. *J ADV NURS* 2015, 71(11):2563-2574.

18. PL L, S J: Nursing students’ clinical judgment regarding rapid response: the influence of a clinical simulation education intervention. *Nursing forum* 2013, 48(1):61-70.

19. SM A: Lecture-based education versus simulation in educating student nurses about central line-associated bloodstream infection prevention guidelines. *Journal of vascular nursing : official publication of the Society for Peripheral Vascular Nursing* 2019, 37(2):125-131.

20. AA A, MM A: High-fidelity simulation effects on CPR knowledge, skills, acquisition, and retention in nursing students. *WORLDV EVID-BASED NU* 2014, 11(6):394-400.

21. J L, E L, L G, R Z, R Z, Q C, Y L, H W, Q M, H W: Effects of Simulation-Based Deliberate Practice on Nursing Students’ Communication, Empathy, and Self-Efficacy. *The Journal of nursing education* 2019, 58(12):681-689.

22. RRO C, SM M, JCA M, VRD C, MS A: Effectiveness of simulation in teaching immunization in nursing: a randomized clinical trial. *REV LAT-AM ENFERM* 2020, 28:e3305.

23. G A, B H, R G, C H: Effectiveness of intermediate-fidelity simulation training technology in undergraduate nursing education. *J ADV NURS* 2006, 54(3):359-369.

24. LI T: Effect of Simulation on the Confidence of University Nursing Students in Applying Cardiopulmonary Assessment Skills: A Randomized Controlled Trial. *The journal of nursing research : JNR* 2017, 25(4):289-295.

25. LI T, A T: Effect of simulation on knowledge of advanced cardiac life support, knowledge retention, and confidence of nursing students in Jordan. *The Journal of nursing education* 2014, 53(1):38-44.

26. LI T: Effect of simulation modules on Jordanian nursing student knowledge and confidence in performing critical care skills: A randomized controlled trial. *International journal of Africa nursing sciences* 2020, 13:100242.

27. A T, LI T: Effect of Cardiac Arrhythmia Simulation on Nursing Students’ Knowledge Acquisition and Retention. *WESTERN J NURS RES* 2015, 37(9):1160-1174.

28. BT E, T G B, SG MA: Comparison of the effects of two different teaching methods in blood pressure measurement training: a randomized controlled study. *BLOOD PRESS MONIT* 2019, 24(6):294-298.

29. AB MA A, DB B: Comparative value of a simulation by gaming and a traditional teaching method to improve clinical reasoning skills necessary to detect patient deterioration: a randomized study in nursing students. *BMC MED EDUC* 2020, 20(1):53.

30. JM P, PP M, AR J, RP C: Clinical Virtual Simulation in Nursing Education: Randomized Controlled Trial. *J MED INTERNET RES* 2019, 21(3):e11529.

31. SY L, AS S, JY R, PK A: Assessment for simulation learning outcomes: a comparison of knowledge and self-reported confidence with observed clinical performance. *NURS EDUC TODAY* 2012, 32(6):e35-e39.

32. van Lankhuijzen L, A D, van Roosmalen J, G Z, AS A: A systematic review of the effectiveness of training in emergency obstetric care in low-resource environments. *BJOG : an international journal of obstetrics and gynaecology* 2010, 117(7):777-787.

33. Burns HK, Odonnell J, Artman J: High-fidelity Simulation in Teaching Problem Solving to 1st-Year Nursing Students: A Novel Use of the Nursing Process. *CLIN SIMUL NURS* 2010, 6(3):e87-e95.

34. KT D: The essentials of debriefing in simulation learning: a concept analysis. *Nursing education perspectives* 2009, 30(2):109-114.

35. Michael M, Katie W, Komal B: Working memory is limited: Improving knowledge transfer by optimising simulation through cognitive load theory. *BMJ Simulation and Technology Enhanced Learning* 2016, 2(4):131-138.

36. Hyeonsook J: A Study of Convergence on Experiences of Clinical Performance and Self-Confidence of Core Basic Nursing Skills, Clinical Competence in Nursing Students. *Journal of the Korea Convergence Society* 2018, 9(11).

37. JH D, DL M, KS K, MD M: Simulation-based education improves student self-efficacy in physiotherapy assessment and management of paediatric patients. *BMJ MED EDUC* 2019, 19(1):463.

38. CHEN X, Zhang DJ, Li SF, Li Y: The Influence of Middle School Students’ Psychological Quality on Academic Performance: A Model of Multiple Mediation Effects. *Journal of Southwest University (Natural Science Edition)* 2021, 43(02):14-19.

39. Yang N GPDS: Yang N, Ghislandi P, Dellantonio S. Online collaboration in a large university class supports quality teaching. *Educational Technology Research and Development* 2018, 66(3):671-691.
40. García-Cabrero B, Hoover ML, Lajoie SP, Andrade-Santoyo NL, Quevedo-Rodríguez LM, Wong J: Design of a learning-centered online environment: a cognitive apprenticeship approach. *Educational Technology Research & Development* 2018, 66(3):1-23.

41. Dai XL, Guo K, Liu L: An Empirical Study on the Influencing Factors of MOOC Learner Satisfaction—Based on the Structural Equation Analysis of the "Chinese University MOOC" Learner Questionnaire. *Modern distance education* 2017(02):17-23.

42. Teimouri Y, Goetze J, Plonsky L: SECOND LANGUAGE ANXIETY AND ACHIEVEMENT: A META-ANALYSIS. *STUD SECOND LANG ACQ* 2018, 41(2):1-25.

43. Y U, EY C, G S, E U, U K: Effect of simulation-based learning on first clinical day stress and anxiety levels of nursing students in Turkey. *JPMA. The Journal of the Pakistan Medical Association* 2020, 70(9):1505-1509.

**Figures**

Figure 1

PRISMA flow diagram.
**Figure 2**

Risk of bias graph.

| Study ID                | Randomized process | Deviations from intended interventions | Missing outcome data | Measurement of the outcome | Selection of the reported result | Overall  |
|-------------------------|--------------------|----------------------------------------|----------------------|---------------------------|---------------------------------|----------|
| Kohran et al., 2019     | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Roak et al., 2019       | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Tomski et al., 2019     | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Pumla et al., 2013      | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Alinier et al., 2019    | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Ju et al., 2019         | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Stayt et al., 2015      | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Jose et al., 2019       | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Almamuni et al., 2017   | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Lodi, 2020              | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Alouab, 2019            | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Lodi, 2016              | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Ahmad et al., 2015      | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Ahmad et al., 2014      | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Sma et al., 2019        | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Aya et al., 2019        | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Li et al., 2019         | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Sok et al., 2011        | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Sok et al., 2012        | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Antonio et al., 2020    | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |
| Raphael et al., 2020    | ⬤                   | ⬤                                      | ⬤                    | ⬤                         | ⬤                                | Low risk |

**As percentage (intention-to-treat)**

- **Overall Bias**
- **Selection of the reported result**
- **Measurement of the outcome**
- **Missing outcome data**
- **Deviations from intended interventions**
- **Randomization process**

- **Low risk**
- **Some concerns**
- **High risk**
Figure 3

Risk of bias summary.

| Study or Subgroup | Experimental Mean | SD | Total | Control Mean | SD | Total | Std. Mean Difference IV, Random, 95% CI |
|-------------------|------------------|----|-------|--------------|----|-------|---------------------------------------|
| Ahmad et al., 2014 | 6.89             | 1.12 | 45    | 5.29         | 1.05 | 45    | 1.46 [0.99, 1.93]                     |
| Ahmad et al., 2015 | 7.3             | 3.1  | 47    | 1.9          | 2.4  | 44    | 1.82 [1.32, 2.31]                     |
| AlAmrani et al., 2017 | 28.20          | 20.02 | 36    | 14           | 16   | 36    | 0.76 [0.26, 1.24]                     |
| Alouch, 2019       | 10.6            | 2.87 | 65    | 9.5          | 4.01 | 66    | 0.31 [-0.03, 0.68]                    |
| José et al., 2019  | 2.32            | 1.44 | 21    | 0.64         | 2.09 | 21    | 0.92 [0.28, 1.56]                     |
| JU et al., 2019    | 23.91           | 5.05 | 66    | 7.85         | 5.98 | 66    | 2.91 [2.42, 3.41]                     |
| Kahraman et al., 2019 | 1.91          | 1.63 | 35    | 1.5          | 1.89 | 35    | 0.23 [-0.24, 0.70]                    |
| Li et al., 2019    | 7.91            | 2.63 | 50    | 3.29         | 3.03 | 50    | 1.62 [1.16, 2.07]                     |
| Loai, 2020        | 18.33           | 3.42 | 38    | 12.69        | 2.69 | 38    | 2.21 [1.63, 2.78]                     |
| Pamela et al., 2013 | 29.84         | 14.89 | 40    | 17.75        | 18.29 | 39    | 1.31 [0.82, 1.80]                     |
| Raphael et al., 2020 | 2.69          | 1.96 | 17    | 1.6          | 3.78 | 17    | 0.32 [-0.35, 1.00]                    |
| Sok et al., 2011   | 9.76            | 2.97 | 15    | 1           | 2.32 | 16    | 5.1 [2.11, 4.32]                      |
| Sok et al., 2012   | 7.75            | 2.9   | 15    | 0.5          | 3.57 | 15    | 2.32 [1.36, 3.27]                     |
| Stayt et al., 2015 | 11.28           | 2.06 | 50    | 8.2          | 4.16 | 48    | 1.55 [1.09, 2.00]                     |
| Tamaki et al., 2019 | 2.3            | 1.3  | 20    | 0.17         | 1.36 | 18    | 1.57 [0.83, 2.31]                     |

Total (95% CI)  560 554  100.0%  1.46 [1.02, 1.90]

Heterogeneity: Tau² = 0.65; Chi² = 140.05, df = 14 (P < 0.00001); I² = 90%
Test for overall effect: Z = 6.57 (P < 0.00001)

Figure 4

Forest plot: effectiveness of simulation-based education interventions on test scores.

| Study or Subgroup | Experimental Mean | SD | Total | Control Mean | SD | Total | Std. Mean Difference IV, Random, 95% CI |
|-------------------|------------------|----|-------|--------------|----|-------|---------------------------------------|
| Ahmad et al., 2014 | 12.27           | 1.14 | 45    | 10.07        | 1.44 | 45    | 1.68 [1.20, 2.16]                     |
| Ahmad et al., 2015 | 12.2            | 3.81 | 47    | 7.2          | 2.79 | 44    | 1.48 [1.01, 1.94]                     |
| AlAmrani et al., 2016 | 61.72         | 7.53 | 49    | 56           | 9.46 | 50    | 0.68 [0.26, 1.07]                     |
| José et al., 2019  | 10.55           | 1.81 | 21    | 11.93        | 1.84 | 21    | -0.74 [-1.37, -0.11]                  |
| Li et al., 2019    | 12.7            | 2.09 | 40    | 7.3          | 3.09 | 42    | 1.55 [1.06, 2.05]                     |
| Loai, 2016        | 25.1            | 1.88 | 35    | 9.3          | 1.22 | 34    | 9.83 [8.07, 11.58]                    |
| Raphael et al., 2020 | 6.55          | 1.71 | 17    | 6.01         | 1.14 | 17    | 0.36 [-0.32, 1.04]                    |

Total (95% CI)  254 253  100.0%  1.81 [1.07, 2.87]

Heterogeneity: Tau² = 1.88; Chi² = 148.55, df = 6 (P < 0.00001); I² = 96%
Test for overall effect: Z = 3.36 (P = 0.0008)

Figure 5

Forest plot: effectiveness of simulation-based education interventions on knowledge retention.

| Study or Subgroup | Experimental Mean | SD | Total | Control Mean | SD | Total | Std. Mean Difference IV, Random, 95% CI |
|-------------------|------------------|----|-------|--------------|----|-------|---------------------------------------|
| AlAmrani et al., 2016 | 42.72           | 6.62 | 53    | 40.16        | 10.33 | 54    | 0.29 [0.09, 0.68]                     |
| Ayse et al., 2019  | 127.03          | 16.43 | 32    | 126.52       | 13.79 | 27    | 0.03 [-0.48, 0.58]                    |
| Basak et al., 2019 | 8.48            | 0.88 | 35    | 7.07         | 1.33 | 36    | 1.23 [0.72, 1.74]                     |
| Li et al., 2019    | 74.38           | 11.55 | 50    | 32.85        | 18.16 | 50    | 2.71 [2.16, 3.26]                     |
| Loai, 2016        | 8.01            | 1.86 | 42    | 4.9          | 1.96 | 42    | 1.13 [1.01, 2.11]                     |
| Loai, 2020        | 35.76           | 5.45 | 38    | 23           | 3.33 | 38    | 2.80 [2.16, 3.44]                     |
| Sok et al., 2012   | 24.53           | 6.56 | 15    | 20.63        | 6.05 | 16    | 0.60 [0.12, 1.33]                     |
| Tamaki et al., 2019 | 2.53            | 0.88 | 20    | 1.45         | 0.81 | 18    | 1.63 [0.89, 2.38]                     |

Total (95% CI)  334 331  100.0%  1.19 [0.48, 1.90]

Heterogeneity: Tau² = 1.11; Chi² = 135.09, df = 8 (P < 0.00001); I² = 94%
Test for overall effect: Z = 3.27 (P = 0.001)

Figure 6
Forest plot: effectiveness of simulation-based education interventions on self-confidence.

| Study or Subgroup | Experimental Mean | Experimental SD | Experimental Total | Control Mean | Control SD | Control Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|-------------------|------------------|-----------------|-------------------|--------------|------------|--------------|--------|----------------------------------------|
| Banu et al., 2019 | 29.38            | 4.96            | 32                | 29.81        | 3.54       | 27           | 32.2%  | -0.10 [-0.61, 0.42]                   |
| José et al., 2019 | 30.38            | 4.57            | 21                | 30.14        | 4.29       | 21           | 28.4%  | 0.05 [-0.55, 0.66]                    |
| JU et al., 2019   | 41.23            | 5.01            | 66                | 37.91        | 5.32       | 66           | 39.3%  | 0.64 [0.29, 0.99]                     |
| **Total (95% CI)**| **119**          | **114**         | **100.0%**        |              |            |              |        | **0.24 [-0.27, 0.74]**                |

Heterogeneity: $\tau^2 = 0.13; \, \chi^2 = 6.46, \, df = 2 \, (P = 0.04); \, I^2 = 69\%$

Test for overall effect: $Z = 0.92 \, (P = 0.36)$

**Figure 7**

Forest plot: effectiveness of simulation-based education interventions on self-efficacy.

| Study or Subgroup | Experimental Mean | Experimental SD | Experimental Total | Control Mean | Control SD | Control Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|-------------------|------------------|-----------------|-------------------|--------------|------------|--------------|--------|----------------------------------------|
| Antonia et al., 2020 | 8.5             | 1.6             | 73                | 7.6          | 1.7        | 73           | 57.7%  | 0.54 [0.21, 0.87]                      |
| José et al., 2019 | 9.04            | 0.55            | 21                | 7.47         | 1.58       | 21           | 42.3%  | 1.30 [0.63, 1.97]                      |
| **Total (95% CI)** | **94**          | **94**          | **100.0%**        |              |            |              |        | **0.86 [0.13, 1.60]**                 |

Heterogeneity: $\tau^2 = 0.22; \, \chi^2 = 3.96, \, df = 1 \, (P = 0.05); \, I^2 = 75\%$

Test for overall effect: $Z = 2.30 \, (P = 0.02)$

**Figure 8**

Forest plot: effectiveness of simulation-based education interventions on satisfaction.

| Study or Subgroup | Experimental Mean | Experimental SD | Experimental Total | Control Mean | Control SD | Control Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|-------------------|------------------|-----------------|-------------------|--------------|------------|--------------|--------|----------------------------------------|
| ALAmrani et al., 2017 | -8.5            | 11.36           | 36                | 3            | 10.54      | 36           | 50.2%  | -1.04 [-1.53, -0.54]                  |
| Banu et al., 2019 | 32.5            | 8.89            | 32                | 29.78        | 8.04       | 27           | 49.8%  | 0.32 [0.20, 0.83]                     |
| **Total (95% CI)** | **68**          | **63**          | **100.0%**        |              |            |              |        | **-0.36 [-1.69, 0.96]**               |

Heterogeneity: $\tau^2 = 0.85; \, \chi^2 = 13.81, \, df = 1 \, (P = 0.0002); \, I^2 = 93\%$

Test for overall effect: $Z = 0.54 \, (P = 0.59)$

**Figure 9**

Forest plot: effectiveness of simulation-based education interventions on anxiety.
Figure 10

Forest plot: effectiveness of different intervention time of simulation-based education interventions on test scores.