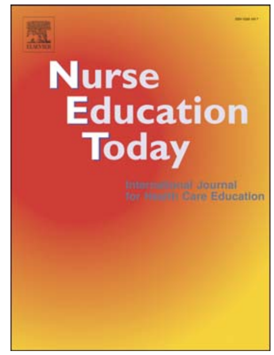
Accepted Manuscript



Simulation-based training for nurses; systematic review and meta-analysis

Pål A. Hegland, Hege Aarlie, Hilde Strømme, Gro Jamtvedt

PII: S0260-6917(17)30073-4
DOI: doi:10.1016/j.nedt.2017.04.004
Reference: YNEDT 3517

To appear in: *Nurse Education Today*

Received date: 14 June 2016
Revised date: 28 February 2017
Accepted date: 7 April 2017

Please cite this article as: Hegland, Pål A., Aarlie, Hege, Strømme, Hilde, Jamtvedt, Gro, Simulation-based training for nurses; systematic review and meta-analysis, *Nurse Education Today* (2017), doi:10.1016/j.nedt.2017.04.004

Simulation-based training for nurses; systematic review and meta-analysis

Hegland, Pål A.[1], Aarlie, Hege[2], Strømme, Hilde[3], Jamtvedt, Gro[4]

Western Norway University of Applied Sciences, Department of Health Studies, Norway – and –

Centre of Evidence-Based Practice, Western Norway University of Applied Sciences, Norway

1 Corresponding author – pal.a.hegland@hotmail.com

2 Second reviewer. Western Norway University of Applied Sciences, Bergen, Norway

3Librarian, Norwegian Institute of Public Health, Oslo, Norway

4 Oslo and Akershus University College of Applied Sciences, Norway

Word count: 3834 (excluded references in text, and bibliography)

Acknowledgement

There are no conflicts of interest in this systematic review except from a genuine interest in the topic. The experts in the field have only contributed with references, and there have been no economic funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

[1] Corresponding author – pal.a.hegland@hotmail.com
[2] Second reviewer. Bergen University College, Bergen, Norway
[3] Librarian, Norwegian Institute of Public Health, Oslo, Norway
[4] Oslo and Akershus University College of Applied Sciences, Norway

Background: Simulation-based training is a widespread strategy to improve health-care quality. However, its effect on registered nurses has previously not been established in systematic reviews. The aim of this systematic review is to evaluate effect of simulation-based training on nurses' skills and knowledge.

Methods: We searched CDSR, DARE, HTA, CENTRAL, CINAHL, MEDLINE, Embase, ERIC, and SveMed+ for randomised controlled trials (RCT) evaluating effect of simulation-based training among nurses. Searches were completed in December 2016. Two reviewers independently screened abstracts and full-text, extracted data, and assessed risk of bias. We compared simulation-based training to other learning strategies, high-fidelity simulation to other simulation strategies, and different organisation of simulation training. Data were analysed through meta-analysis and narrative syntheses. GRADE was used to assess the quality of evidence.

Results: Fifteen RCTs met the inclusion criteria. For the comparison of simulation-based training to other learning strategies on nurses' skills, six studies in the meta-analysis showed a significant, but small effect in favour of simulation (SMD -1.09, CI -1.72 to -0.47). There was large heterogeneity (I^2 85%). For the other comparisons, there was large between-study variation in results. The quality of evidence for all comparisons was graded as low.

Conclusion: The effect of simulation-based training varies substantially between studies. Our meta-analysis showed a significant effect of simulation training compared to other learning strategies, but the quality of evidence was low indicating uncertainty. Other comparisons showed inconsistency in results. Based on our findings simulation training appears to be an effective strategy to improve nurses' skills, but further good-quality RCTs with adequate sample sizes are needed.

Keywords: simulation, clinical competence, quality improvement, nurses, knowledge, skills, systematic review.

Introduction

Healthcare services offer complex, and advanced treatment for patients. Therefore highly competent and skilled healthcare providers are needed to secure patient safety (Grol, Berwick, & Wensing, 2008). Studies show that errors in healthcare are a risk for patient safety that in many cases can be prevented (Patel, Kannampallil, & Shortliffe, 2015). Patient safety, and quality improvement are therefore important issues in today's society (Institute for Healthcare Improvement, 2015). There are several tools for quality improvement such as evidence-based guidelines, or clinical audits (Ivers et al., 2012; NICE, 2014). Another strategy used to improve performance among healthcare workers, and students, is simulation-based training.

Simulation-based training is practising realistic scenarios using a specialized manikin, computer software, or humans playing the role as patient (Society for Simulation in Healthcare, 2014; The International Nursing Association for Clinical Simulation & Learning, 2013). The setting can be high-fidelity, where manikins and equipment are advanced, also called technology-enhanced simulation. It can also be low-fidelity where the equipment is less advanced (Healthy Simulation, 2014; Salas, Paige, & Rosen, 2013). The most specialized manikins today simulate the physiology of humans with pulse, blood pressure, and secretion of sweat and tears. The facilitator has the ability to regulate the parameters according to the actions initiated by the health workers, using specialized computer software (Healthy Simulation, 2014).

Previously published systematic reviews on simulation-based training for students in health-profession education, showed large effects on students' knowledge and skills, and moderate effects on patient-related outcomes (Cant & Cooper, 2010; Cook et al., 2011). Simulation-based training for critical care nurses in continuing education programmes, can improve adherence to recommendations about safe medication (Jansson, Kääriäinen, & Kyngäs, 2012). Further technology-enhanced simulation in emergency medicine seems to have effect on several outcomes (Ilgen, Sherbino, & Cook, 2013). In addition, a qualitative study among midwifery students showed that simulation created links between theory and practice, and provided a safe learning environment (Lendahls and Oscarsson, 2017)

One systematic review that summarized evidence for graduated nurses separately found only one cohort study. Since this learning strategy is widely used for nurses, it is relevant to evaluate its effect. The aim of this systematic review is therefore to summarize the effect of simulation-based training on nurses' knowledge and skills.

Methods

Inclusion Criteria

To be considered relevant for inclusion in the systematic review, the studies had to be randomised controlled trials (RCT) evaluating the effect of simulation-based training for graduated nurses, or graduated nurses in continuing education. Skills and/or knowledge had to be the primary outcomes in the trials. The studies were eligible for inclusion if they were written in English, German, Norwegian, Swedish, or Danish.

Comparisons

Relevant comparisons for the systematic review were simulation-based training to other learning strategies, different simulation strategies compared to each other, or different organisation of the simulation training.

Identification of studies

Systematic searches were performed in The Cochrane Database of Systematic Reviews (Wiley), Database of Abstracts of Reviews of Effects (CRD), Health Technology Assessment Database (CRD), Cochrane Central Register of Controlled Trials (Wiley), CINAHL (EBSCO), MEDLINE (OVID), Embase (OVID), ERIC (EBSCO), and SveMed+. Searches were performed 5 September 2014, with an update search 15 December 2016. Search strategies were reviewed by an experienced research librarian (Sampson et al., 2009). Search terms included ‘simulation’, ‘technology-enhanced simulation’, ‘computer-based simulation’, ‘nurs*’, ‘skills’, ‘knowledge’ among others (complete list in appendix E). No limitations pertaining to language, publication year, or study design were applied. An example of a complete search strategy is presented in Appendix A, and complete search strategies are available upon request. Hand searches were performed in the journals Clinical Simulation in Nursing, and Simulation in Healthcare for the years 2013 to December 2016 as they were indexed in MEDLINE until January 2013 at the time of the primary search. Previously identified systematic reviews and primary studies were screened for relevant references. Experts in the field were contacted for additional published or unpublished research. Finally Clinical Trials (ClinicalTrails.gov, 2014) were searched with the text word ‘simulation’, to identify unpublished or ongoing studies.

Screening and selection of studies

Two review authors screened all titles and abstracts independently. The selection process was piloted by reading the first 50 titles and abstracts to calibrate understanding of inclusion and exclusion criteria (Higgins & Deeks, 2011, Ch. 7). We obtained full-text articles of all studies that did not clearly meet the exclusion criteria. The same two reviewers read all full-text articles for final inclusion. Disagreements in all stages were solved by discussion until consensus was reached (Higgins & Deeks, 2011, Ch. 7).

Data extraction

A pre-defined data-extraction form was developed. Data from the included studies were extracted by one person, and quality checked by a second person. We extracted the following data: author name, publication year, number of participants, interventions and comparisons, outcomes, country, and effect measures.

Assessing risk of bias and grading the evidence

The Cochrane Collaboration's Risk of Bias Tool (Higgins, Altman, & Sterne, 2011, Ch. 8.5), was used to evaluate risk of bias in included studies. We used the Guideline Development Tool (GRADE Working Group, 2012; Guyatt et al., 2011) to assess quality of the evidence for the following comparisons: Simulation-based training versus other learning strategies, high-fidelity simulation versus other simulation strategy, and different organisation of simulation training. The grading was made per comparison for each outcome, and was assessed as high, moderate, low, or very low quality (Guyatt et al., 2008).

Data synthesis

We planned to do a quantitative synthesis, by conducting meta-analyses when there was low clinical diversity in the studies (Deeks, Higgins, & Altman, 2011, Ch. 9). We also planned narrative syntheses, if meta-analyses were not possible to conduct. The statistical data were entered to Review Manager 5.3 (The Nordic Cochrane Centre, 2014).

For continuous measures we calculated Standardized Mean Difference (Inverse Variance, random effects model), and 95% confidence interval (CI), and for dichotomous measures we calculated Risk Ratio with 95% CI (Mantel-Haenszel, random effects model) (Deeks et al., 2011, Ch. 9).

In the meta-analysis, between-study consistency (heterogeneity) was calculated with I^2 statistics, which estimates the percentage of the variability not due to chance. An I^2 value > 50% indicates substantial heterogeneity. A p-value was also calculated, and $p < 0.05$ indicates significant between-study heterogeneity (Deeks et al., 2011, Ch. 9).

Results

Identification of studies and study selection

Fourteen-hundred and seventy-five potentially relevant studies were identified through the database searches, hand searching, and screening of reference lists. After screening by two reviewers independently as described in the methods section, fifty-eight articles were selected for full-text review. Fifteen studies met the inclusion criteria, and were included in this systematic review, see Figure A screening and selection process. Excluded studies are presented in Appendix B.

Insert Figure A here

Study characteristics

The fifteen included studies were published between 2005 and 2016, three of which before 2010. Nine studies were conducted in the USA (Arnold et al., 2013; Corbridge et al., 2010; Hebbar et al., 2015; Johnson et al., 2012; Keleekai et al., 2016; Maneval et al., 2012; Rutherford-Hemming et al., 2016; Schneider et al., 2006; Weiner et al., 2011), one in Australia (Cioffi, Purcal, & Arundell, 2005), two in Belgium (De Regge, Calle, De Paepe, & Monsieurs, 2008; Monsieurs et al., 2012), one in Finland (Jansson et al., 2016), one in Singapore (Liaw et al., 2015), and one in Norway (Simonsen et al., 2014). The studies had enrolled a total of 852 registered nurses. Twelve studies were conducted in hospitals (Arnold et al., 2013; De Regge et al., 2008; Hebbar et al., 2015; Jansson et al., 2016; Keleekai et al., 2016; Liaw et al., 2015; Maneval et al., 2012; Monsieurs et al., 2012; Rutherford-Hemming et al., 2016; Schneider et al., 2006; Simonsen et al., 2014; Weiner et al., 2011), whereas one study was in a military training programme (Johnson et al., 2012). Five studies had enrolled nurses in specialized wards (Arnold et al., 2013; Hebbar et al., 2015; Jansson et al., 2016; Rutherford-Hemming et al., 2016; Weiner et al., 2011), and three had enrolled anaesthesia-, midwifery, or advanced practice nursing students (Cioffi et al., 2005; Corbridge et al., 2010; Johnson et al., 2012). The final seven had enrolled regular nurses (De Regge et al., 2008; Keleekai et al., 2016; Liaw et al., 2015; Maneval et al., 2012; Monsieurs et al., 2012; Schneider et al., 2006; Simonsen et al., 2014). Nine trials compared simulation-based training to other learning strategies (Cioffi et al., 2005; Corbridge et al., 2010; Hebbar et al., 2015; Johnson et al., 2012; Keleekai et al., 2016; Liaw et al., 2015; Maneval et al., 2012; Rutherford-Hemming et al., 2016; Weiner et al., 2011), and two studies compared high-fidelity simulation to other simulation strategies like computer-based, or low-fidelity (Arnold et al., 2013; Johnson et al., 2012). Two studies compared interactive e-learning simulation to other learning strategies (Schneider et al., 2006; Simonsen et al., 2014). Finally, three studies compared alternative organisation of simulation-based training (De Regge et al., 2008; Jansson et al., 2016; Monsieurs et al., 2012). Manikins were used for training in eleven studies (Arnold et al., 2013; Corbridge et al., 2010; De Regge et al., 2008; Hebbar et al., 2015; Jansson et al., 2016; Johnson et al., 2012; Keleekai et al., 2016; Liaw et al., 2015; Maneval et al., 2012; Monsieurs et al., 2012; Weiner et al., 2011). Two studies used a fellow participant playing the role as patient (Cioffi et al., 2005; Rutherford-Hemming et al., 2016). Skills were evaluated in twelve trials (Arnold et al., 2013; Cioffi et al., 2005; De Regge et al., 2008; Hebbar et al., 2015; Jansson et al., 2016; Johnson et al., 2012; Keleekai et al., 2016; Liaw et al., 2015; Monsieurs et al., 2012; Rutherford-Hemming et al., 2016; Schneider et al., 2006; Weiner et al., 2011), whereas nine trials evaluated knowledge (Arnold et al., 2013; Cioffi et al., 2005; Corbridge et al., 2010; Jansson et al., 2016; Keleekai et al., 2016; Maneval et al., 2012; Rutherford-Hemming et al., 2016; Simonsen et al., 2014; Weiner et al., 2011). Characteristics of the included studies are presented in Table A.

Insert Table A here

Risk of Bias (RoB) in included studies

Risk of bias assessment is presented in Figure B. The risk of bias in the included studies was overall unclear to high due to issues with allocation concealment, blinding, and incomplete outcome data. Three studies (Arnold et al., 2013; Hebbar et al., 2015; Monsieurs et al., 2012) were assessed to have an overall high risk of bias, eight studies (Cioffi et al., 2005; Corbridge et al., 2010; De Regge et al., 2008; Jansson et al., 2016; Johnson et al., 2012; Keleekai et al., 2016; Rutherford-Hemming et al., 2016; Liaw et al., 2015) were assessed to have an unclear risk, and only four studies (Maneval et al., 2012; Simonsen et al., 2014; Schneider et al., 2006; Weiner et al., 2011) to have an overall low risk of bias.

Insert Figure B here

The quality of evidence for simulation-based training versus other learning strategies, high-fidelity simulation versus other simulation strategy, and different organisation of simulation-based training was assessed as low for both outcomes. The grading of documentation is presented in Appendices C.1–C.3 GRADE Evidence profiles or Summary of findings tables.

Simulation-based training versus other learning strategies

Six studies were eligible for meta-analysis on this comparison for nurses' skills (Cioffi et al., 2005; Hebbar et al., 2015; Johnson et al., 2012; Keleekai et al., 2016; Rutherford-Hemming et al., 2016; Weiner et al., 2011), see Figure C. Summary of findings are presented in table B. All six studies measured effect with predefined scoring-sheets, unique to the individual study.

Cioffi et al. (2005) found a 7-point difference on an ungraded scale on skills for data collection segments. Hebbar et al. (2015) found a 2.6-point difference on a 7-point scale. Johnson et al. (2012) had a 47.77-point difference on a 107-point scale, Keleekai et al. (2016) found a 9.7-point difference, Rutherford-Hemming et al. (2016) found 18.6-point difference on a scale 0–100, and Weiner et al. (2011) found no difference between the groups, with only one point in difference. We pooled the results in a random effects meta-analysis, and found a standardized mean difference (SMD) -1.09 (CI -1.72 to -0.47). The pooled effect size for these studies shows a significant, but possibly small effect in favour of simulation. Further, the heterogeneity between the studies is high with I^2 85%, and $p < 0.00001$, indicating uncertainty with the results.

Insert Figure C here

Further, the meta-analysis on computer-based simulation to other learning strategies (Figure C) showed SMD -1.06 (CI -1.50 to -0.62), which is a significant effect in favour computer-based simulation. Johnson et al. (2012) found a 7.33 point difference, Liaw et al., (2015) found an 8.25-point difference on a 45-point scale, and Schneider et al. (2006) found a 6.45-point difference, all in favour of intervention group.

Our analysis shows a positive effect of simulation-based training compared to other learning strategies on nurses' skills, with a p-value < 0.0007 in the meta-analysis. On the other hand, the grading of the evidence showed low quality on this comparison, and our confidence in these finding is therefore limited.

Six studies evaluated effect of simulation on nurses' knowledge, all measured by tests, see Appendix D Forest Plot 1, and summary of findings table B. The population varied from newly educated nurses, to specialized nurses. Because of that, we assessed clinical diversity between the studies to be too significant to pool the total effect sizes. Cioffi et al. (2005)

found a 2-point difference on an ungraded scale in favour of simulation training. Corbridge et al. (2010) found a 0.1-point difference in a 5-point scale in favour of simulation. Keleekai et al. (2016) found a 3.5-point difference on a 22-point scale in favour of simulation. Rutherford-Hemming et al. (2016) found a 1.3-point difference in favour of simulation. Weiner et al. (2011) found a 1.8-point difference on a 55-point scale in favour of intervention, and finally Maneval et al. (2012) found a 0.39-point difference on a 33-point scale. Only Keleekai et al. (2016) found a significant difference in favour of intervention (SMD -1.68 (CI -2.28 to -1.08)). None of the remaining studies showed a significant difference between groups as they have broad confidence interval that cross over the line of no effect. The grading of evidence shows low quality, which makes our confidence in the results of this analysis low.

Insert Table B here

One study compared computer-based simulation to classroom teaching on outcome knowledge. Simonsen et al. (2014) found a non-significant difference in knowledge scores between groups. These findings are therefore not conclusive.

High-fidelity simulation versus other simulation strategies

Two studies compared high-fidelity simulation to other simulation strategies on nurses' skills. However Arnold et al. (2013) did not report results due to missing data from one group. The analysis is therefore based on one study (Johnson et al., 2012) comparing high-fidelity simulation to CD-ROM. Johnson et al. (2012) found a 40.44-point difference on a 107-point scale in favour the high-fidelity group. These results are statistically significant with $p < 0.0001$ and SMD -2.42 (CI -3.35 to -1.49). However, the evidence was graded low quality on this comparison, and we cannot draw any conclusions from this one study with relatively few participants. Summary of findings are presented in table C.

One study had made this comparison on nurses' knowledge. Arnold et al. (2013) made three comparisons. For the comparison high-fidelity simulation versus low-fidelity simulation, they found a SMD -1.01 (CI -1.98 to -0.04). The results are statistically significant in favour of high-fidelity simulation. However, the results are uncertain due to a broad CI close up to the zero-effect line. The comparison high-fidelity simulation versus computer-based simulation shows a SMD 0.45 (-0.46 to 1.37). The results are in favour of the computer-based group, however not statistically significant. Finally low-fidelity simulation versus computer-based simulation found a SMD 1.38 (CI 0.32 to 2.43), and results statistically in favour of computer-based simulation. Se Appendix D Forest Plot 2.

The results indicate that computer-based simulation might be more effective for improving nurses' knowledge when compared to both high-, and low-fidelity simulation. Further, high-fidelity simulation might be superior to low-fidelity simulation. However, the quality of evidence is low, and our confidence in these results is therefore sparse, see table C.

Insert Table C here

Comparing different organisation of simulation

Three studies compared different organisation of simulation training on nurses' skills (De Regge et al., 2008; Jansson et al., 2016; Monsieurs et al., 2012), whereas one had made this comparison on knowledge (Jansson et al., 2016). Summary of findings for this comparison are presented in table D. De Regge et al. (2008) compared basic life-support training with one participant to one instructor in the intervention group, and six participants to one instructor in the control. For number of ventilations with correct volume, De Regge et al. (2008) found Median 4 with an inter-quartile range 0–7 in the intervention group, and Median 5 and an inter-quartile range 1–7 in the control group. For number of compressions with correct depth they found Median 59 with inter-quartile range 9–89 for the intervention group, and Median

33 and inter-quartile range 8–82 in the control group. The effect estimates in this study are highly imprecise with results in favour of the intervention in one skill measure, and in favour of control in the other skill measure.

Monsieurs et al. (2012) have compared a self-learning station versus traditional basic life-support training. They have presented results for five skills, and reported the results as dichotomous outcomes. We present these results in Appendix D Forest Plot 3. The Risk Ratio varies from 0.52 in one skills-measure to 1.14 in another. Also for this study, the results are highly inconsistent with results in favour of both intervention, and control group.

Jansson et al. (2016) compared feedback and debriefing, versus no feedback or debriefing in high-fidelity simulation. They found no differences between intervention and control group on neither skills nor knowledge, even though there is a small trend towards effect of feedback and debriefing. See Appendix D, Forest Plot 4.

Because of a high degree of inconsistency in the results, we cannot draw any overall clear conclusions based on the current studies. This is also due to low quality of the evidence.

Insert Table D here

Discussion

Based on our findings in this systematic review of fifteen included studies, simulation-based training seems to be effective in improving nurses' skills when compared to other learning strategies. However, there is uncertainty to these results due to weak documentation. For the other comparisons, we cannot draw any conclusions for registered nurses. The meta-analysis did show a significant but possibly small effect of simulation-based training compared to other learning strategies. This small effect could be due to only six eligible studies with small sample-sizes in the analysis. Further, the grading of evidence showed low quality, which means that we have limited confidence in these findings. In the sub-group analysis on computer-based simulation versus other learning strategies on outcome knowledge, we also found a significant effect in favour of simulation, and less statistical heterogeneity (I^2 23%). However, this analysis was also graded as low quality of evidence. For the other comparisons, several of the included studies present results in favour of simulation, but there is inconsistency in results between studies. Only three studies (Hebbar et al., 2015; Johnson et al., 2012; Rutherford-Hemming et al., 2016) showed a significant effect of high-fidelity simulation compared to other learning strategies on nurses' skills. Further, there seems to be a trend towards computer-based simulation as the most effective strategy on nurses' knowledge when compared to other simulation strategies, but the results are uncertain and not definitive. The quality of evidence is graded low on these comparisons as well. In our opinion, the lack of conclusive results in, and between the studies may be a result of small sample-sizes in most of the studies, as well as weakness in design for several of the trials.

In our first meta-analysis, there was large statistical heterogeneity with I^2 85%. We could not find any clinical diversity between the studies, nor any major differences in the risk of bias of the studies that could explain the differences in results. The total number of participants is also approximately the same. The study of Johnson et al. (2012) showed a larger effect on nurses' skills than the other studies. We chose to do two sub-group analyses due to differences in simulation strategy.

One reason the results of Johnson et al. (2012) differs from the other studies, could be the motivation of the participants for this learning strategy. The nurses who are about to work in a war setting might to a greater extent acknowledge the need of training on realistic situations. The situations they are about to face is substantially different from what they have met in their previous practice, and of a kind they will not meet in a secure hospital setting in the USA. This could make them more motivated to train with simulation, to be prepared for the situations they are faced with in a war. However this is only our assumption, and there could be other reasons for the results.

Simulation-based training is one of the strategies used to improve quality in healthcare. However, simulation-based training may not be the best strategy for implementing a new guideline, or government directives. Then workshops or clinical audits are frequently used strategies to increase health-workers adherence to the guidelines. These strategies have shown small to moderate effect when used alone (Forsetlund et al., 2009; Ivers et al., 2012). However, evidence show that the strategies used in quality improvement should be tailored to the context where they are implemented (Baker et al., 2015). In our opinion quality improvement in healthcare need several approaches to be successful, depending on what area to be improved. Simulation-based training is in our view most feasible for practical issues, whereas implementation of, for example, guidelines should use other approaches. However, simulation-based training could be a part of a tailored implementation strategy, for example in combination with printed educational material (Giguère et al., 2012).

The inclusion criteria in this systematic review were relatively narrow, which may have contributed to our findings of non-conclusive results. If a broader range of study-designs and a wider population had been considered eligible, we would have included more studies to synthesize. On the other hand, we chose the narrow criteria to make a highly focused review.

This systematic review has a thorough literature search, reviewed by an experienced librarian. In phases of the process that are critical, we have been two review authors screening, and selecting independently. We are therefore confident that all relevant studies have been identified, and included in the review. We have not contacted any of the authors of the primary studies for additional information. Our analyses were based solely on information found in the articles. This might have influenced the risk of bias assessment and grading of evidence, since the authors could have contributed additional information. However, we made his choice, because all studies then were considered equally in terms of available information. Further, we only have two outcomes in this review. It could have been relevant to evaluate patient outcomes, and economic implications of simulation-based training as well.

The other systematic reviews we identified included the same comparisons and outcomes as our review. The sources searched in the other reviews were to a great extent the same as the ones we searched. The other reviews identified more references than we did from their searches, but this could be due to our relatively restricted population of registered nurses. Some of the reviews had searched for unpublished literature in databases designed for such studies. We only searched for protocols of unpublished studies in Clinical Trials. Further, we have assessed risk of bias using Cochrane's RoB-tool, whereas the other reviews have used different tools. However, we find that the tools are comparable. We have not found that the other reviews have graded the quality of evidence using GRADE.

Further, the results from the other systematic reviews we identified differ from ours, in the sense that they have found a large significant effect of simulation training (Cook et al., 2011; Ilgen et al., 2013), whereas we found a possibly small effect. This difference could be due to the fact that we only included RCTs. Finally, we have identified, and included RCTs not included in the other reviews.

Simulation-based training is rather resource demanding, and the equipment needed for high-fidelity simulation is expensive. Therefore, an evaluation of the use and effectiveness of this intervention is important. Based on the trends in this review, and the positive effect of simulation-based training for students shown in other systematic reviews, we assume that further RCTs of good quality might conclude that simulation-based training is an effective learning-strategy for nurses. Based on our experience with simulation this is the case.

However, our findings in this review are not conclusive, and we cannot claim that simulation-based training is effective on nurses' skills and knowledge.

Further studies should use RCT-designs, and they should aim to include larger sample sizes of registered nurses. The researchers should also strictly follow the rules of the RCT-design when conducting the trial, in particular when it comes to allocation concealment, and blinding of the personnel analysing the results. The purpose of these studies should be to establish whether simulation-based training for nurses is an effective learning strategy on nurses' skills and knowledge. When this is established, different simulation strategies should be compared to each other to establish which strategy or organisation is the most effective.

Conclusion

Based on the findings in this systematic review no clear conclusions can be drawn. The methodical quality of the included studies varies too much, and the inconsistencies in results are significant. For the comparison simulation-based training versus other learning strategies, effect measure is statistically significant, but the results are uncertain due to heterogeneity. For the other comparisons, the results are not conclusive. The quality of the evidence was also graded low for all comparisons. In our opinion, the effect of simulation-based training for nurses seems to be positive on skills, whereas for knowledge effect has not been established. There is a need for new, good quality RCTs of reasonable size to establish a confidence in the effect-measures. However, we believe this systematic review is an important contribution to the knowledge in this topic.

Key points:

- Quality improvement in healthcare is important to increase patient safety
- Simulation-based training is one of several quality-improvement strategies used in healthcare
- Effect of simulation-based training on nurses' skills and knowledge is still uncertain due to lack of robust evidence

References

Arnold, J. J., Johnson, L. M., Tucker, S. J., Chesak, S. S., & Dierkhising, R. A. (2013). Comparison of Three Simulation-Based Teaching Methodologies for Emergency Response. *Clinical Simulation In Nursing, 9*(3), e85–e93. doi:10.1016/j.ecns.2011.09.004

Baker, R., Camosso-Stefinovic, J., Gillies, C., Shaw, E. J., Cheater, F., Flottorp, S., van Lieshout J, J. C. (2015). Tailored interventions to address determinants of practice. *Cochrane Database of Systematic Reviews 2012*(4).

Cant, R. P., & Cooper, S. J. (2010). Simulation-based learning in nurse education: systematic review. *Journal of Advanced Nursing, 66*(1), 3–15.

Cioffi, J., Purcal, N., & Arundell, F. (2005). A pilot study to investigate the effect of a simulation strategy on the clinical decision making of midwifery students. *Journal of Nursing Education, 44*(3), 131–134.

ClinicalTrails.gov. (2014). Retrieved from https://clinicaltrials.gov/

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., Hamstra, S. J. (2011). Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. *JAMA, 306*(9), 978–988.

Corbridge Susan, J., Robinson, F. P., Tiffen, J., & Corbridge Thomas, C. (2010). Online Learning versus Simulation for Teaching Principles of Mechanical Ventilation to Nurse Practitioner Students International Journal of Nursing Education Scholarship (Vol. 7).

De Regge, M., Calle, P. A., De Paepe, P., & Monsieurs, K. G. (2008). Basic life support refresher training of nurses: individual training and group training are equally effective. *Resuscitation, 79*(2), 283–287. doi:http://dx.doi.org/10.1016/j.resuscitation.2008.06.015

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (2011). Analysing data and undertaking meta-analyses. In J. P. T. Higgins & S. Green (Eds.), *Cochrane Handbook for systematic reviews of interventions version 5.1.0*: The Cochrane Collaboration 2011.

Forsetlund, L., Bjørndal, A., Rashidian, A., Jamtvedt, G., O'Brien, M. A., Wolf, F. M., Oxman, A. D. (2009). Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* (2).

Giguère, A., Légaré, F., Grimshaw, J., Turcotte, S., Fiander, M., Grudniewicz, A., Gagnon, M.-P. (2012). Printed educational materials: effects on professional practice and healthcare outcomes. *Cochrane Database of Systematic Reviews* (10).

GRADE Working Group. (2012). Guideline Development Tool (GDT). Retrieved from http://www.guidelinedevelopment.org/

Grol, R., Berwick, D. M., & Wensing, M. (2008). On the trail of quality and safety in health care. *BMJ, 336*(7635), 74–76. doi:10.1136/bmj.39413.486944.AD

Guyatt, G. H., Oxman, A. D., Akl, E., Kunz, R., Vist, G., & Brozek, J. (2011). GRADE guidelines 1. Introduction – GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology, 64*, 383–394.

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ, 336*(7650), 924–926. doi:10.1136/bmj.39489.470347.AD

Healthy Simulation. (2014). Medical Simulation. Retrieved from http://healthysimulation.com/medical-simulation/

Hebbar, K. B., Cunningham, C., McCracken, C., Kamat, P., & Fortenberry, J. D. (2015). Simulation-based paediatric intensive care unit central venous line maintenance bundle training. *Intensive and Critical Care Nursing, 31*(1), 44–50. doi:http://dx.doi.org/10.1016/j.iccn.2014.10.003

Higgins, J. P. T., Altman, D. G., & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias. In J. P. T. Higgins & S. Green (Eds.), *Cochrane Handbook for systematic reviews of interventions version 5.1.0*: The Cochrane Collaboration 2011.

Higgins, J. P. T., & Deeks, J. J. (2011). Selecting studies and collecting data. In J. P. T. Higgins & S. Green (Eds.), *Cochrane Handbook for systematic reviews of interventions version 5.1.0*: The Cochrane Collaboration 2011.

Ilgen, J. S., Sherbino, J., & Cook, D. A. (2013). Technology-enhanced simulation in emergency medicine: a systematic review and meta-analysis. *Acad Emerg Med, 20*(2), 117–127. doi:10.1111/acem.12076

Institute for Healthcare Improvement. (2015). Patient Safety. Retrieved from http://www.ihi.org/Topics/PatientSafety/Pages/default.aspx

Ivers, N., Jamtvedt, G., Flottorp, S., Young, J. M., Odgaard-Jensen, J., French, S. D., Oxman, A. D. (2012). Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database of Systematic Reviews* (6). doi:10.1002/14651858.CD000259.pub3.

Jansson, M., Kääriäinen, M., & Kyngäs, H. (2012). Effectiveness of Simulation-Based Education in Critical Care Nurses' Continuing Education: A Systematic Review. *Clinical Simulation In Nursing, 9*(9), e355–e360. doi:10.1016/j.ecns.2012.07.003

Jansson, M. M., Syrjälä, H. P., Ohtonen, P. P., Meriläinen, M. H., Kyngäs, H. A., & Ala-Kokko, T. I. (2016). Randomized, controlled trial of the effectiveness of simulation education: A 24-month follow-up study in a clinical setting. *American Journal of Infection Control, 44*(4), 387–393. doi:http://dx.doi.org/10.1016/j.ajic.2015.10.035

Johnson, D., Corrigan, T., Gulickson, G., Holshouser, E., & Johnson, S. (2012). The effects of a human patient simulator vs. a CD-ROM on performance. *Military medicine,* (10), 1131–1135. Retrieved from http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/161/CN-00840161/frame.html

Keleekai, N. L., Schuster, C. A., Murray, C. L., King, M. A., Stahl, B. R., Labrozzi, L. J., . . . Glover, K. R. (2016). Improving Nurses' Peripheral Intravenous Catheter Insertion Knowledge, Confidence, and Skills Using a Simulation-Based Blended Learning Program: A Randomized Trial. Simulation in Healthcare, 11(6), 376–384. doi:10.1097/sih.0000000000000186

Lendahls, L., & Oscarsson, M. G. (2017). Midwifery students' experiences of simulation- and skills training. Nurse Education Today, 50, 12–16. doi:10.1016/j.nedt.2016.12.005

Liaw, S. Y., Wong, L. F., Chan, S. W.-C., Ho, J. T. Y., Mordiffi, S. Z., Ang, S. B. L., . . . Ang, E. N. K. (2015). Designing and Evaluating an Interactive Multimedia Web-Based Simulation for Developing Nurses' Competencies in Acute Nursing Care: Randomized Controlled Trial. *Journal of Medical Internet Research, 17*(1), e5. doi:10.2196/jmir.3853

Maneval, R., Fowler, K. A., Kays, J. A., Boyd, T. M., Shuey, J., Harne-Britner, S., & Mastrine, C. (2012). The effect of high-fidelity patient simulation on the critical thinking and clinical decision-making skills of new graduate nurses. *Journal of Continuing Education in Nursing, 43*(3), 125–134. doi:10.3928/00220124-20111101-02

Monsieurs, K. G. a., De Regge, M. b., Schelfout, S. a., D'Hondt, F. a., Mpotos, N. a., Valcke, M. c., & Calle, P. A. a. (2012). Efficacy of a self-learning station for basic life support refresher training in a hospital: a randomized controlled trial. *European Journal of Emergency Medicine, 19*(4), 214–219.

NICE. (2014). What we do. Retrieved from https://www.nice.org.uk/about/what-we-do

Patel, V. L., Kannampallil, T. G., & Shortliffe, E. H. (2015). Role of cognition in generating and mitigating clinical errors. *BMJ Quality and Safety*; 0:1–7. doi:10.1136/bmjqs-2014-003482. Retrieved from http://qualitysafety.bmj.com/content/early/2015/05/02/bmjqs-2014-003482.abstract

Rutherford-Hemming, T., Kelsey, N. C., Grenig, D. L., Feliciano, M., Simko, L., & Henrich, C. M. (2016). Multisite Single-Blinded Randomized Control Study of Transfer and Retention of Knowledge and Skill Between Nurses Using Simulation and Online Self-Study Module. Simulation in Healthcare, 11(4), 264–270. doi:10.1097/sih.0000000000000168

Salas, E., Paige, J. T., & Rosen, M. A. (2013). Creating new realities in healthcare: the status of simulation-based training as a patient safety improvement strategy. *BMJ Quality and Safety, 22*, 449–452.

Sampson, M., McGowan, J., Cogo, E., Grimshaw, J., Moher, D., & Lefebvre, C. (2009). An evidence-based practice guideline for the peer review of electronic search strategies. *Journal of Clinical Epidemiology, 62*(9), 944–952. doi:10.1016/j.jclinepi.2008.10.012

Schneider, P. J., Pedersen, C. A., Montanya, K. R., Curran, C. R., Harpe, S. E., Bohenek, W., . . . Wellman, K. E. (2006). Improving the safety of medication administration using an interactive CD-ROM program. American Journal of Health-System Pharmacy, 63(1), 59–64. doi:10.2146/ajhp040609

Simonsen, B. O., Daehlin, G. K., Johansson, I., & Farup, P. G. (2014). Improvement of drug dose calculations by classroom teaching or e-learning: a randomised controlled trial in nurses. BMJ Open, 4(10). doi:10.1136/bmjopen-2014-006025

Society for Simulation in Healthcare. (2014). Retrieved from http://www.ssih.org/

The International Nursing Association for Clinical Simulation & Learning. (2013). Retrieved from http://www.inacsl.org/i4a/pages/index.cfm?pageid=1

The Nordic Cochrane Centre. (2014). Review Manager (RevMan) (Version 5.3). The Cochrane Collaboration. Retrieved from http://tech.cochrane.org/revman

Weiner, G. M., Menghini, K., Zaichkin, J., Caid, A. E., Jacoby, C. J., & Simon, W. M. (2011). Self-directed Versus Traditional Classroom Training for Neonatal Resuscitation. *Pediatrics, 127*(4), 713–719. doi:10.1542/pedS.2010-2829

Figure A Flow-chart screening and inclusion of studies

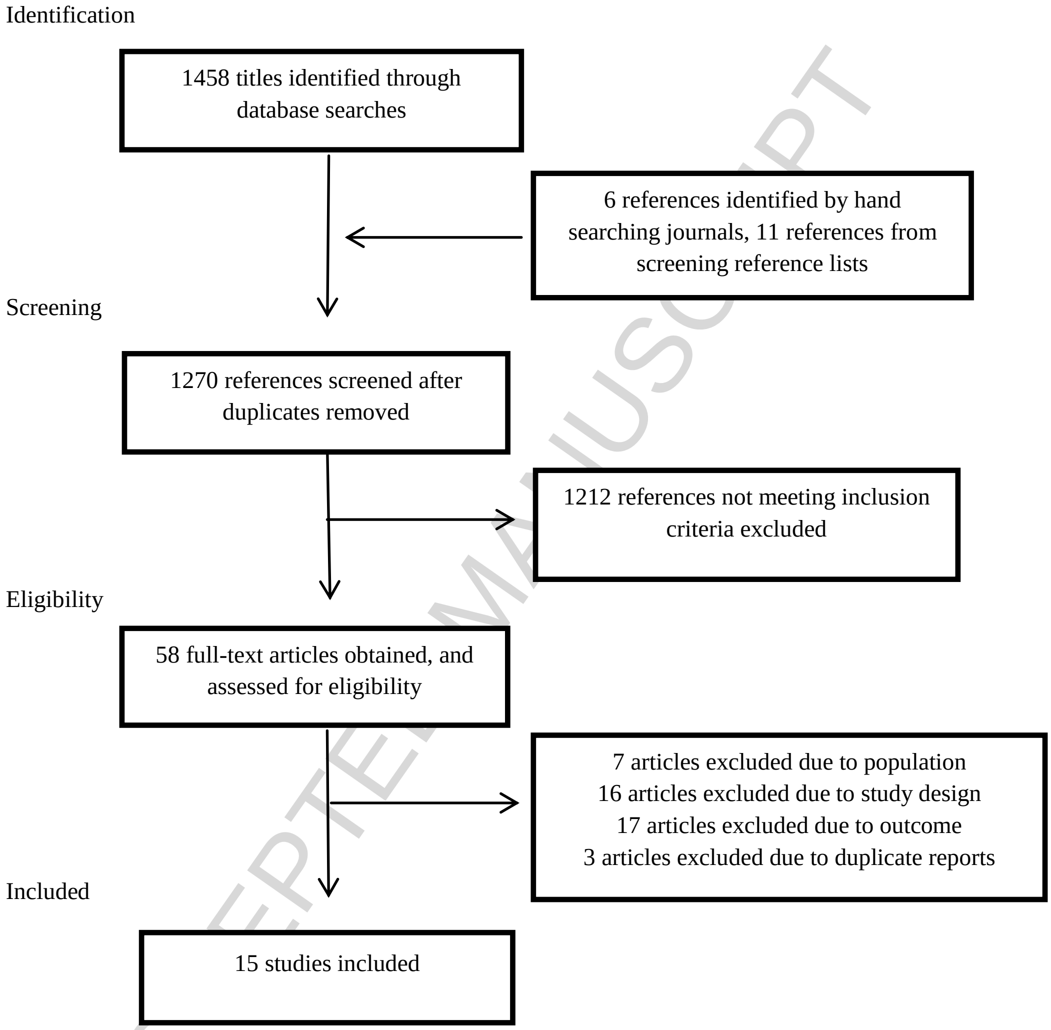
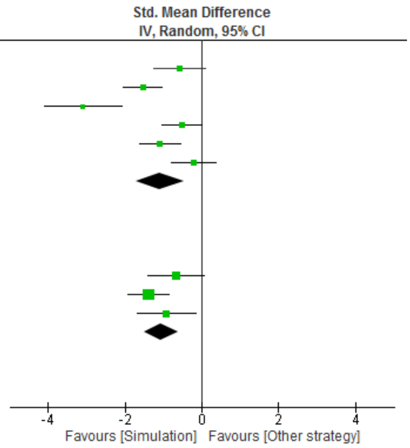


Figure B Risk of Bias in included studies

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Arnold et al. 2013	?	+	−	−	−	?	+
Cioffi et al. 2005	+	+	?	?	−	?	+
Corbridge et al. 2010	?	−	?	?	+	?	+
De Regge et al. 2008	+	+	?	?	?	?	+
Hebbar et al. 2015	?	+	−	−	?	+	+
Jansson et al. 2016	?	+	?	?	−	?	?
Johnson et al. 2012	?	−	?	+	?	?	+
Keleekai et al. 2016	+	+	?	?	+	?	+
Liaw et al. 2015	?	+	+	+	?	?	+
Maneval et al. 2012	+	+	+	+	+	?	+
Monsieurs et al. 2012	+	+	−	−	+	?	+
Rutherford-Hemming et al. 2016	+	+	?	?	+	?	+
Schneider et al. 2006	+	+	+	+	+	?	+
Simonsen et al. 2014	+	+	?	+	+	?	+
Weiner et al. 2011	+	?	+	+	+	?	+

Figure C Meta-analysis simulation-based training versus other learning strategies, outcome skills

Study or Subgroup	Simulation training Mean	SD	Total	Other learning strategies Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Realistic simulation training								
Cioffi et al. 2005	-69	9	18	-62	14	18	16.4%	-0.58 [-1.25, 0.09]
Hebbar et al. 2015	-15.8	1.1	39	-13.2	2.1	40	17.8%	-1.53 [-2.03, -1.03]
Johnson et al. 2012	-89.84	18.7	19	-42.02	9.1	16	13.2%	-3.09 [-4.11, -2.08]
Keleekai et al. 2016	-77	21	30	-67.3	16.4	29	17.7%	-0.51 [-1.03, 0.01]
Rutherford-Hemming et al. 2016	-46.1	17.6	35	-27.5	15.9	29	17.6%	-1.09 [-1.62, -0.56]
Weiner et al. 2011	-24.5	3.6	23	-23.8	3.1	23	17.2%	-0.20 [-0.78, 0.37]
Subtotal (95% CI)			**164**			**155**	**100.0%**	**-1.09 [-1.72, -0.47]**
Heterogeneity: Tau² = 0.51; Chi² = 32.71, df = 5 (P < 0.00001); I² = 85%								
Test for overall effect: Z = 3.41 (P = 0.0007)								
1.1.2 Computer-based simulation								
Johnson et al. 2012	-49.4	12.2	14	-42.07	9.1	16	28.2%	-0.67 [-1.41, 0.07]
Liaw et al. 2015	-38.66	6.35	35	-30.41	5.31	32	45.4%	-1.39 [-1.92, -0.85]
Schneider et al. 2006	10.4	7.98	15	16.85	5.27	14	26.4%	-0.92 [-1.69, -0.15]
Subtotal (95% CI)			**64**			**62**	**100.0%**	**-1.06 [-1.50, -0.62]**
Heterogeneity: Tau² = 0.04; Chi² = 2.60, df = 2 (P = 0.27); I² = 23%								
Test for overall effect: Z = 4.73 (P < 0.00001)								



Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.93), I² = 0%

Table A Characteristics of included studies

Author; Year, Country	Participants, Setting, Design	Intervention	Comparison	Outcome measures
Arnold et al., 2013 USA	28 Graduated nurses. A critical care orientation programme in a large hospital. RCT	Didactic instructions for emergency response followed by a 3.5 hour full-scale (high-fidelity) simulation in a simulation centre designed as an intensive care unit, with a specialized manikin.	Low fidelity emergency response training 3.5 hour session for one group, and a 3.5 hour computer-based simulation training for the second control group.	Skills assessed in a simulation centre using a 12 item tool called ERPT two weeks post intervention. Knowledge measured with a 12 item written exam one week after the simulation session.
Cioffi et al., 2005 Australia	36 midwifery students. In a midwifery education at university level. RCT	Two low-fidelity simulation sessions using fellow student playing role as patient, followed by a critical reflection of the decision-making process.	Lectures as usual.	Skills and knowledge measured by a self-developed master sheet assessing the participants' performance.
Corbridge et al., 2010 USA	20 advanced practice nursing students	High-fidelity simulation participating in intubation, and adjusting ventilator settings.	Online Power-Point presentation.	Knowledge measured by a 12-iten multiple-choice test.
De Regge et al., 2008 Belgium	103 registered nurses from non-critical care hospital wards. RCT	Basic life support training with one instructor to one trainee (IT), session lasting 30 minutes.	Basic life support training with one instructor to six trainees (GT), session lasting 30 minutes.	Skills measured by data from the manikin immediately after and ten months after simulation training.
Hebbar et al., 2015 USA	76 registered nurses in a Paediatric Intensive-care Unit. RCT	Low-fidelity simulation using a simple manikin to train procedures with a central venous line every three, six and twelve months with an instructor.	No simulation strategy, didactic training followed by test and observation of dressing-change annually.	Skills measured by a 17-point CVL dressing change checklist utilized locally. Measured before and at three, six and twelve months post intervention.
Jansson et al., 2016 Finland	17 critical care nurses in an intensive care ward. RCT	High-fidelity simulation with a 20 minutes instruction, followed by a 10 minutes simulation session focusing on adhering to a ventilator-bundle in intubated patients. Experimental group received verbal feedback and participated in a 60 min structured debriefing.	Attended the same high-fidelity simulation training, without the verbal feedback and structured debriefing.	Measurement at baseline and after three months in a simulated setting, and six and 24 months post-intervention in a real clinical setting. Skills measured by an 86-item Ventilator Bundle Observation Schedule with a scale ranging 0-60 points. Knowledge measured with a 49 item Ventilator Bundle Questionnaire, ranging from 0-37 points.
Johnson et al., 2012 USA	49 anaesthesia students in a military nursing simulation centre. RCT	Group one trained high-fidelity simulation with an advanced manikin. Three scenarios, 30 minutes of lengths. The other group watched a CD-ROM containing a Power-Point presentation covering three types of combat injuries.	Control group receiving no simulation strategy, but lectures as usual.	Skills measured by a 120 criteria CP instrument relative to the care of patients with battlefield injuries. Tested pre- and 1 month post-intervention.

Author; Year, Country	Participants, Setting, Design	Intervention	Comparison	Outcome measures
Keleekai et al., 2016 USA	59 registered nurses at surgical units in a non-university teaching hospital. RCT	Online learning course followed by 8 hours simulation-based live course teaching peripheral intravenous catheter insertion (PIVC)	Standard teaching of PIVC. After completion of test at time 2, control group attended intervention	Skills and knowledge measured pre- and post-intervention, and after group-crossover. Skills measured with PIVC insertion skills checklist, a 28 item checklist. Knowledge measured with PIVC insertion knowledge checklist, a 14 item, 22-point tool.
Liaw et al., 2015 Singapore	67 registered nurses with less than five years' experience working in different hospital wards. RCT	High-fidelity simulation scenario as a pre-test where the participant should assess and manage a deteriorating patient. Simulation session lasted for 15 minutes. Immediately after, they independently undertook a 3 hours web-based simulation.	Attended the same high-fidelity scenario pre-test, without the web-based simulation.	Skills measured pre- and one week post-intervention in the simulation laboratory. The session was video-recorded and the researchers evaluated using a Rescuing a Patient in Deterioration Situations tool, an 18-item tool, ranging from 0-45 points.
Maneval et al., 2012 USA	26 new graduate nurses, in a hospital setting in non-specialised wards. RCT	Standard orientation programme plus six high-fidelity patient simulation experiences with the Laerdal Sim-Man®.	Standard orientation programme.	Knowledge measured with The health sciences reasoning test, a 33-item multiple-choice test.
Monsieurs et al., 2012 Belgium	90 registered nurses, in a hospital setting. RCT	Self-learning training (SL) watching a short video and training with Resusci Anne Skills Station (Laerdal, Norway) installed in a small room and available 24/7.	Instructor lead training (IL) traditional basic life-support course with a certified BLS instructor. Group of six participants, and duration of session 120 minutes.	Skills measured with data gathered from the manikin.
Rutherford-Hemming et al., 2016 USA	64 registered nurses in post-partum or birthing center in hospital. RCT	Low-fidelity simulation with actors trained as standardized patients.	Online self-study module with Power-Pint slides containing pictures	Skills measured with The Performance Observation Measurement Tool, a 14-item instrument. Knowledge measured with Neurologic Knowledge Assessment, a 12-item tool.
Schneider et al., 2006 USA	30 registered nurses in medical or surgical hospital units. RCT	Interactive CD-rom titled "basic medication administration"	No additional training	Skills observed trough observation of medication errors in the ward
Simonsen et al., 2014 Norway	141 nurses from hospitals and primary care. RCT	2-day self-directed interactive internet-based e-learning course on medication administration	1-day conventional classroom, and 1-day self-study	Knowledge measured with MCQ test in drug-dose calculation, 14 questions
Weiner et al., 2011 USA	46 registered nurses, working in a postpartum ward RCT	Self-directed with textbook and access to simulation room before a 90 min. simulation session with instructor.	Traditional classroom training, 6 hours of lectures and instructor-directed skill stations.	Skills measured by data from manikin and knowledge measured by written tests.

Table B Summary of findings effect of simulation-based training compared to other learning strategies

Outcome	Results Intervention (I)	Control (C)	Effect sizes (95% CI)	Number of studies (Participants)	Quality of evidence (GRADE)
Skills	Mean 69 (SD 9)	Mean 62 (SD 14)	Realistic simulation SMD -1.09 (-1.72 to -0.47)	6 (319)	⊕⊕◯◯ LOW [1,2,3,4]
	Mean 15.8 (SD 1.1)	Mean 13.2 (SD 2.1)			
	Mean 89.84 (SD 18.7)	Mean 42.02 (SD 9.1)			
	Mean 77 (SD 21)	Mean 67.3 (SD 16.4)			
	Mean 46.1 (SD 17.6)	Mean 27.5 (SD 15.9)			
	Mean 24.5 (SD 3.6)	Mean 23.8 (SD 3.1)			
	Mean 49.4 (SD 12.2)	Mean 42.07 (SD 9.1)	Computer-based SMD -1.06 (-1.50 to -0.62)	3 (126)	⊕⊕◯◯ LOW [3,5]
	Mean 38.66 (SD 6.35)	Mean 30.41 (SD 5.31)			
	Mean 10.4 (SD 7.98)	Mean 16.85 (SD 5.27)			
Knowledge	Mean 12 (SD 9)	Mean 14 (SD 12)	SMD -0.18 (CI -0.84 to 0.47)	6 (251)	⊕⊕◯◯ LOW [5,6,7]
	Mean 9.2 (SD 1.3)	Mean 9.1 (SD 1.7)	SMD -0.06 (CI -0.94 to 0.81)		
	Mean 17.4 (SD 1.9)	Mean 13.9 (SD 2.2)	SMD -1.68 (CI -2.28 to -1.08)		
	Mean 51.5 (SD 4)	Mean 49.8 (SD 3.1)	SMD -0.46 (CI -1.24 to 0.32)		
	Mean 67.7 (SD 15.9)	Mean 66.4 (SD 16)	SMD -0.08 (CI -0.57 to 0.41)		
	Mean 22.08 (SD 2.84)	Mean 21.69 (2.25)	SMD -0.15 (CI -0.73 to 0.43)		

SD = Standard deviation, CI = Confidence Interval, SMD = Standardized Mean Difference

GRADE Working Group grades of evidence

High quality: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Footnotes:

1. One of the included studies had unclear risk of bias on randomisation
2. Two of the included studies had high risk of bias, and one study had unclear risk regarding blinding
3. One of the included studies had high risk of bias, and two had unclear risk regarding incomplete outcome data
4. The results differ from significant to not significant
5. One of the studies had unclear risk of bias on allocation concealment
6. Four of the studies had unclear risk of bias on blinding
7. One of the studies had high risk of bias due to no reported baseline characteristics

Table C Summary of findings effect of high-fidelity simulation compared to other simulation strategy

Outcome	Results		Effect measure (95% CI)	Number of studies (Participants)	Quality of evidence (GRADE)
	Intervention (I)	**Control (C)**			
Skills	Mean 89.84 (SD 18.7)	Mean 49.4 (SD 12.2)	SMD -2.42 (-3.35 to -1.49)	1 (46)	⊕⊕◯◯ LOW[1,2,3,5]
Knowledge	*High-fidelity versus low fidelity*			1 (26)	⊕⊕◯◯ LOW [4,5]
	Mean 76 (SD 8)	Mean 67 (SD 9)	SMD -1.01 (-1.98 to -0.04)		
	High-fidelity versus computer-based				
	Mean 76 (SD 8)	Mean 80 (SD 9)	SMD 0.45 (-0.46 to 1.37)		
	Low-fidelity versus computer-based				
	Mean 67 (SD 9)	Mean 80 (SD 9)	SMD 1.38 (0.32 to 2.43)		

SD = Standard deviation, CI = Confidence Interval, SMD = Standardised Mean Difference

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect
Footnotes:

1. The study has high risk of bias on allocation concealment
2. The study has unclear risk of bias on blinding
3. One study has unclear risk of bias on incomplete outcome data
4. The study has high risk of bias on blinding and incomplete outcome data, and unclear risk of bias on allocation concealment
5. The study has unclear risk of bias on random sequence generation

Table D Summary of findings effect of different organisation of simulation training

Outcome	Results Intervention (I)	 Control (C)	Effect measure (95% CI)	Number of studies (Participants)	Quality of evidence (GRADE)
Skills				3 (210)	⊕⊕◯◯ LOW [1,2,3,4]
	Ventilations with correct volume Median 4 (0-7)	Median 5 (1-7)			
	Compressions with correct depth Median 59 (9-86)	Median 33 (8-82)			
	Mean compression depth 38-51 mm Events 13/45	Events 25/45	RR 0.52 (0.31 to 0.88)		
	Mean compression depth over 38 mm Events 17/45	Events 27/45	RR 0.63 (0.40 to 0.98)		
	Mean compression rate 80 – 120 pr. Min Events 29/45	Events 31/45	RR 0.94 (0.70 to 1.25)		
	Ventilation volume 400-1000 ml Events 24/45	Events 21/45	RR 1.14 (0.75 to 1.73)		
	Any incomplete release ≥ 5mm Events 23/45	Events 21/45	RR 1.10 (0.72 to 1.67)		
	Mean 35.27 (SD 14)	Mean 35,2 (SD 9)	SMD -0.01 (-1.06 to 1.05)		
Knowledge	Mean 23.6 (SD 12)	Mean 21.2 (SD 12)	SMD -0.19 (-1.26 to 0.89)	1 (17)	⊕⊕◯◯ LOW [1,3,5,6]

SD = Standard deviation, CI = Confidence Interval, RR = Risk Ratio

GRADE Working Group grades of evidence

High quality: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Footnotes:

1. The study has unclear risk of bias on blinding
2. The study has unclear risk of bias on incomplete outcome data
3. The study has unclear risk of bias on selective reporting
4. The results of the study are not conclusive because they are in favour of both the experimental and control group. The results are not significant.
5. The study has unclear risk of bias on allocation concealment
6. The study has unclear risk of bias due to high drop-out.

Appendix A Search strategy MEDLINE (OVID)

Search date: 15.12.16

Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE 1946 to present

▼ **Search History** (17) View Sa

# ▲	Searches	Results	Type	Actions	Annotations
1	exp Nurses/	75741	Advanced	Display Results | More ▾	
2	(nurse adj1 (clinician* or anesthetist* or anaesthetist* or practitioner)).tw.	5946	Advanced	Display Results | More ▾	
3	((registered or acute or intensive or critical) adj3 nurse*).tw.	15197	Advanced	Display Results | More ▾	
4	1 or 2 or 3	89790	Advanced	Display Results | More ▾	
5	Patient Simulation/	3741	Advanced	Display Results | More ▾	
6	Computer Simulation/	155300	Advanced	Display Results | More ▾	
7	Manikins/	3779	Advanced	Display Results | More ▾	
8	(manikin* or mannequin*).tw.	3170	Advanced	Display Results | More ▾	
9	((virtual or technology) adj3 simulation*).tw.	1495	Advanced	Display Results | More ▾	
10	simulat*.tw.	356083	Advanced	Display Results | More ▾	
11	5 or 6 or 7 or 8 or 9 or 10	442300	Advanced	Display Results | More ▾	
12	exp Professional Competence/	93525	Advanced	Display Results | More ▾	
13	skill*.tw.	138033	Advanced	Display Results | More ▾	
14	expertice.tw.	6	Advanced	Display Results | More ▾	
15	((nursing or technical) adj2 competence*).tw.	891	Advanced	Display Results | More ▾	
16	12 or 13 or 14 or 15	211561	Advanced	Display Results | More ▾	
17	4 and 11 and 16	330	Advanced	Display Results | More ▾	

Appendix B Excluded studies

Study	Excluded due to
Andrighetti, T. P., Knestrick, J. M., Marowitz, A., Martin, C. & Engstrom, J. L. (2012) Shoulder dystocia and postpartum hemorrhage simulations: student confidence in managing these complications. *Journal of midwifery & women's health* (1), s. 55–60.	Excluded due to study design, not randomised.
De Regge, M., Monsieurs, K. G., Valcke, M. & Calle, P. A. (2012) Training nurses in a self-learning station for resuscitation: factors contributing to success or failure. *Journal of Emergency Nursing*, 38 (4), s. 386–91.	Duplicate report from the already included study of Monsieurs et al. (2012).
Drake, S. A., Langford, R., Young, A., & Ayers, C. (2015). Forensic nursing science knowledge and competency: the use of simulation. *Critical Care Nursing Quarterly, 38*(1), 81–88.	Excluded due to reporting of outcome measure not distinguished between nurse and student.
Ford, D.G., Seybert, A. L., Smithburger, P. L., Kobulinsky, L. R., Samosky, J. T. & Kane-Gill, S. L. (2010) Impact of simulation-based learning on medication error rates in critically ill patients. *Intensive Care Medicine*, 36, s. 1526–1531.	Excluded due to study design, not randomised.
Freeland, T. R., Pathak, S., Garrett, R. R., Anderson, J. A., & Daniels, S. K. (2016). Using Medical Mannequins to Train Nurses in Stroke Swallowing Screening. *Dysphagia, 31*(1), 104-110. doi:10.1007/s00455-015-9666-6	Excluded due to study design, not randomised.
Gundrosen, S., Solligard, E., & Aadahl, P. (2014). Team competence among nurses in an intensive care unit: the feasibility of in situ simulation and assessing non-technical skills. *Intensive & Critical Care Nursing, 30*(6), 312–317.	Excluded due to Outcome measure. Do not measure effect of simulation-training.
Harvey, E. M., Echols, S. R., Clark, R. & Lee, E. (2014) Comparison of Two TeamSTEPPS® Training Methods on Nurse Failure-to-Rescue Performance. *Clinical Simulation In Nursing*, 10 (2), s. e57–e64.	Excluded due to study design. A quasi-experimental design, not randomised.
Jansson, M. M., Ala-Kokko, T. I., Ohtonen, P. P., Merilainen, M. H., Syrjala, H. P., & Kyngas, H. A. (2014). Human patient simulation education in the nursing management of patients requiring mechanical ventilation: a randomized, controlled trial. *American Journal of Infection Control, 42*(3), 271–276.	Duplicate report from the already included study of Jansson et al. (2016).
Johnson, D. & Johnson, S. (2014) The Effects of Using a Human Patient Simulator Compared to a CD-ROM in Teaching Critical Thinking and Performance. *US Army Medical Department Journal*, s. 59–64.	Duplicate Report from the already included study Johnson et al. (2012).
Pozner, C. N., Almozlino, A., Elmer, J., Poole, S., McNamara, D. A. & Barash, D. (2011) Cardiopulmonary resuscitation feedback improves the quality of chest compression provided by hospital health care professionals. *American Journal of Emergency Medicine*, 29 (6), s. 618–625.	Excluded due to Outcome measure. Do not measure effect of simulation training.
Szlachta, J. (2013) Peer Instruction of First-Year Nurse Anesthetist students: A Pilot Study of a Strategy to Use Liited Faculty Resources and Promote Learning. *Journal of Nurse Education*, 52 (6), s. 355–359.	Excluded due to outcome measure. Do measure feedback from a fellow student, rather than the effect of simulation training.

Appendix C.1 GRADE Evidence profile

Question: Simulation-based training compared to other learning strategies for Nurses

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other consideration	High-Fidelity Simulation	Other learning strategies	Relative (95% CI)	Absolute (95% CI)		
Skills (assessed with: Scoring instruments)[1]												
6	randomised trials	serious [2,3,4]	serious [1]	not serious	not serious	none	164	155	-	SMD **1.09 SD lower** (-1.72 lower to -0.47 lower)	⨁⨁◯◯ LOW	CRITICAL
Knowledge (assessed with: Questionnaire or written test)												
6	randomised trials	serious [5,6,7]	serious [8]	not serious	not serious	none	One of the studies has significant results in favour intervention, the other five found no significant difference between groups, however results in favour intervention				⨁⨁◯◯ LOW	CRITICAL

Bibliography (systematic reviews):
MD – mean difference, RR – relative risk
CI: Confidence interval; **SMD:** Standardised mean difference

1. The results differ from significant to not significant
2. One of the included studies had unclear risk of bias on randomisation
3. Two of the included studies had high risk of bias, and one study had unclear risk regarding blinding
4. One of the included studies had high risk of bias, and two had unclear risk regarding incomplete outcome data
5. One of the studies had unclear risk of bias on allocation concealment
6. One of the studies had unclear risk of bias on blinding
7. One of the studies had high risk of bias due to no reported baseline characteristics
8. The results are both in favour of the intervention and the control group

Appendix C.2 GRADE Summary of findings table

Summary of findings:

High Fidelity Simulation compared to Other forms of simulation strategy for nurses

Patient or population: nurses
Intervention: High Fidelity Simulation
Comparison: Other forms of simulation strategy

Outcomes	Impact	№ of participants (Studies)	Quality of the evidence (GRADE)
Skills assessed with: Scoring forms	One study reports significant effect of High-fidelity Simulation, whereas one study lacks data from one of the comparison groups and therefore cannot report on the outcome. [4]	63 (2 RCTs)	⊕⊕◯◯ LOW [1 2 3 6]
Knowledge assessed with: Test	The study shows a better, but not significant result of High-fidelity Simulation on knowledge.	28 (1 RCT)	⊕⊕◯◯ LOW [5]

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio; **OR:** Odds ratio;

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

1. One study has high risk of bias on allocation concealment
2. One study has high risk of bias on blinding, whereas one study has unclear risk
3. One study has high risk of bias on incomplete outcome data, whereas one has unclear risk
4. Two studies show significant effect, one shows better but not significant difference, and one study shows no difference between the groups
5. The study has high risk of bias on blinding and incomplete outcome data, and unclear risk of bias on allocation concealment
6. Two studies has unclear risk of bias on random sequence generation

Appendix C.3 GRADE Summary of findings table

Summary of findings:

Different organisation of simulation training

Patient or population: basic life-support training for nurses

Setting:

Intervention: Individual Simulation Training or Voice feedback

Comparison: Group Training or No voice feedback

Outcomes	Impact	№ of participants (studies)	Quality of the evidence (GRADE)
Skills assessed with: Data collected from manikin, plus observation	The results are inconclusive. On some measurements, the intervention is in favour of the experimental group, whereas on other measurements the results are in favour of the control group.	210 (3 RCTs)	⊕⊕◯◯ LOW [1,2,3,4]
Knowledge	The results are not significant, and have a Broad CI, which crosses the zero-effect line. There is a high dropout, and small sample size.	17 (1 RCT)	⊕⊕◯◯ LOW [1,3,5]

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval

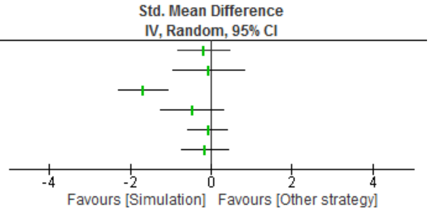
GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

1. The study has unclear risk of bias on blinding
2. The study has unclear risk of bias on incomplete outcome data
3. The study has unclear risk of bias on selective reporting
4. The results of the study are not conclusive because they are in favour of both the experimental and control group. The results are not significant.
5. The results of the study are not conclusive with a broad CI

Appendix D Forest Plots

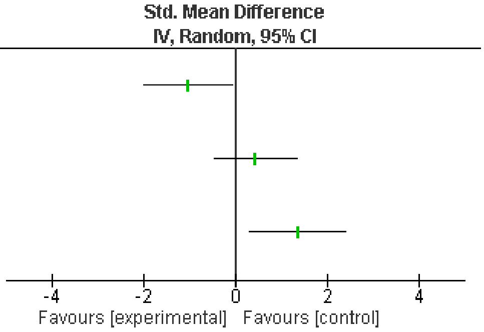
Forest Plot 1 simulation-based training versus other learning strategies, outcome knowledge

Study or Subgroup	Simulation training Mean	SD	Total	Other learning strategies Mean	SD	Total	Std. Mean Difference IV, Random, 95% CI
Cioffi et al. 2005	12	9	18	14	12	18	-0.18 [-0.84, 0.47]
Corbridge et al. 2010	-9.2	1.3	10	-9.1	1.7	10	-0.06 [-0.94, 0.81]
Keleekai et al. 2016	-17.4	1.9	30	-13.9	2.2	29	-1.68 [-2.28, -1.08]
Maneval et al. 2012	-51.5	4	13	-49.8	3.1	13	-0.46 [-1.24, 0.32]
Rutherford-Hemming et al. 2016	-67.7	15.9	35	-66.4	16	29	-0.08 [-0.57, 0.41]
Weiner et al. 2011	-22.08	2.84	23	-21.69	2.25	23	-0.15 [-0.73, 0.43]



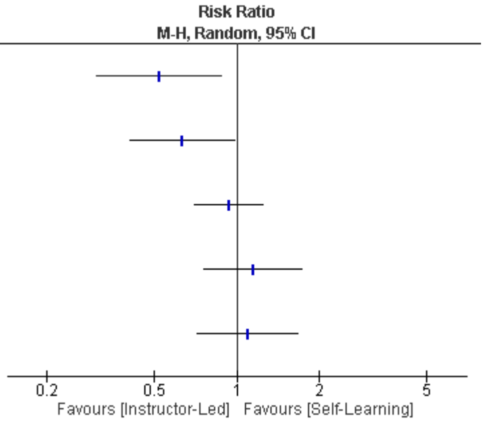
Forest Plot 2 comparison different types of simulation-based strategies, outcome knowledge

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Std. Mean Difference IV, Random, 95% CI
2.2.1 High fidelity simultaion versus low fidelity							
Arnold et al, 2013	-76	8	10	-67	9	9	-1.01 [-1.98, -0.04]
2.2.2 High fidelity siulation versus computer-based							
Arnold et al, 2013	-76	8	10	-80	9	9	0.45 [-0.46, 1.37]
2.2.3 Low fidelity simulation versus Computer-based siulation							
Arnold et al, 2013	-67	9	9	-80	9	9	1.38 [0.32, 2.43]



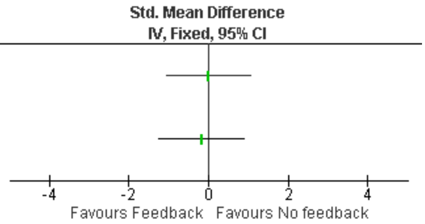
Forest Plot 3 self-learning versus instructor-led simulation training, outcome skills

Study or Subgroup	Self learning station Events	Total	Instructor-led training Events	Total	Risk Ratio M-H, Random, 95% CI
3.1.1 Mean compression depth 38-51mm					
Monsieurs et al, 2012	13	45	25	45	0.52 [0.31, 0.88]
3.1.2 Mean compression depth over 38 mm					
Monsieurs et al, 2012	17	45	27	45	0.63 [0.40, 0.98]
3.1.3 Mean compression rate 80-120 pr min					
Monsieurs et al, 2012	29	45	31	45	0.94 [0.70, 1.25]
3.1.4 Ventilation volume 400-1000 ml					
Monsieurs et al, 2012	24	45	21	45	1.14 [0.75, 1.73]
3.1.5 Any incomplete release ≥ 5 mm					
Monsieurs et al, 2012	23	45	21	45	1.10 [0.72, 1.67]



Forest Plot 4 feedback and debriefing versus no feedback and debriefing in high-fidelity simulation, outcomes skills and knowledge

	Feedback and debriefng			No feedback or debriefing			Std. Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**IV, Fixed, 95% CI**
3.9.1 Skills							
Jansson et.al., 2016	-35.27	14	11	-35.2	9	5	-0.01 [-1.06, 1.05]
3.9.2 Knowledge							
Jansson et.al., 2016	-23.6	12	10	-21.2	12	5	-0.19 [-1.26, 0.89]



Appendix E Search-terms PIO

Subject headings	Population	Intervention	Outcome
MeSH (MEDLINE, CDSR, DARE, HTA, CENTRAL, SveMed+)	Nurses Nurse Anesthetists Nurse Clinicians Nurse Practitioners	Computer Simulation Patient Simulation Problem-Based Learning Teaching Manikins Competency-based education.	Clinical Competence Professional competence
Emtree (Embase)	Nurse Registered nurse Nurse Anesthetists Nurse Clinicians Nurse Practitioner Clinical Nurse Specialist Acute nurse practitioner Emergency nurse practitioner	Simulation Simulator Computer simulation	Competence Clinical Competence Nursing competence Professional competence
CINAHL	Nurse Practitioners Nurses Nursing practice Clinical Nurse Specialists Registered nurses Critical care nursing Intensive care nursing	Simulations Computer simulation Patient simulation Models, anatomic Teaching methods, Clinical	Clinical competence Professional competence Nursing skills
ERIC	Nurses Nursing	Simulation Computer simulation	Competence Skill development
Text words – All databases (English)	• Nurse • Nursing • Acute care nurse • Intensive care nurse • Intensive care nursing • Registered nurse • Nurse Anaesthetists • Nurse Clinicians • Nurse Practitioners	• Simulation. • Virtual reality simulation. • Mannequin – based simulation. • Technology-enhanced simulation training. • Computer simulation. • Training. • Mannequin. • Manikin. • Patient simulation • Problem-Based Learning • Teaching • Competency-based education. • Competency based education.	• Skill. • Learning. • Clinical competence • Clinical skill • Professional competence • Technical expertise • Nursing skills
Text words SveMed+ (Norwegian)	• Sykepleier	• Simulering • Medisinsk simulering	• Kunnskap • Ferdighet

Highlights

- Simulation-based training is an effective strategy to improve nurses' skills
- Simulation can be an element in an tailored strategy to improve healthcare quality, and reduce errors
- There is a need of good quality RCTs to establish which simulation-strategy, and organisation is most effective