Radiography 25 (2019) 260–268

Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Review article

# Role of the virtual reality simulator (ScanTrainer) as a multidisciplinary training tool in transvaginal ultrasound: A systematic review and narrative synthesis

M. Almestehi [a, b, *], W. Alomaim [a], L. Rainford [a], D. Stokes [c], M. Stanton [a], M. Moran [a]

[a] Radiography and Diagnostic Imaging, School of Medicine, University College Dublin, Belfield, Dublin 4, Ireland
[b] Diagnostic Imaging, College of Applied Studies and Community Service, King Saud University, Riyadh, Saudi Arabia
[c] Health Sciences Library, University College Dublin, Belfield, Dublin 4, Ireland

ARTICLE INFO

*Article history:*
Received 21 October 2018
Received in revised form
6 December 2018
Accepted 15 December 2018
Available online 4 January 2019

*Keywords:*
Transvaginal ultrasound
Narrative synthesis
ScanTrainer
Simulation
Systematic review
Clinical skills

ABSTRACT

*Objectives:* This systematic review investigates the role of the ScanTrainer as a virtual reality training simulator and its impact on the scanning skills in transvaginal ultrasound of novice ultrasound practitioners.

*Key findings:* After searching ten databases for studies incorporating the simulator as a part of the learning/training process, ten out of 684 studies met the eligibility criteria and were included in the review. The analysis of the textual findings using narrative synthesis approach resulted in four themes: Validation (assessment of the validity of the simulator), Learning (using the simulator as a learning tool), Perspectives (the perceptions of participants trained on the simulator), and Transferable skills (skills developed on the simulator can transfer to clinical practice).

*Conclusion:* Although literature indicates that the simulator is valuable as a training/learning tool, there is insufficient evidence of measurable effects on clinical practice of simulator usage by different healthcare professions.



## Introduction

Ultrasonography is increasingly used as the primary diagnostic tool in Obstetrics and Gynaecology (OB-GYN) and it is considered safe, portable and affordable which makes it broadly available within the clinical environment.[1,2] Unlike X-ray or magnetic resonance imaging (MRI), practice in ultrasound is highly operator-dependent and requires various complex skills.[3,4] The UK based Consortium for the Accreditation of Sonographic Education (CASE) states that ultrasound scanning and interpretation is entirely dependent on the practitioner skills at the time of the scan and any deficiency within the scan cannot be resolved later by involving another more skilled practitioner.[4] It is common practice that ultrasound scanning is performed by practitioners with a variety of primary professional backgrounds, e.g. radiography, sonography, radiology, midwifery, obstetrics or gynecology.[5] Major factors contributing to competency in ultrasound scanning have been identified as the duration of training and the number and variety of scans performed.[1,2,6] For example, the Royal College of Radiologists (RCR) recommends that to be competent for independent scanning in gynecological ultrasound requires the completion of supervised practice for at least one year with a minimum one session per week and three to eight ultrasound examinations per session.[7] However, developing skills in ultrasound has become a considerable issue in the recent years because of the limited places that can be offered for clinical training.[8] The National Health Service (NHS) has reported higher service demand in England since 2012 with a 16% increase in emergency admissions, and a 23% increase in non-obstetric ultrasound examinations.[9] These demands on the sonographers within busy departments are impacting upon the time available for training students.[10,11] Training is further complicated in OB-GYN imaging as the majority of the ultrasound scans that are performed for gynecological diagnosis and for early pregnancy assessments are transvaginal scans (TVS). The intimate nature of this procedure leads to patients being reluctant to allow novice practitioners to perform training scans, and as a result, developing TVS skills for many ultrasound learners can be especially challenging.[12]

* Corresponding author. Radiography and Diagnostic Imaging, School of Medicine, University College Dublin, Belfield, Dublin 4, Ireland.
*E-mail address:* may.almestehi@hotmail.com (M. Almestehi).

https://doi.org/10.1016/j.radi.2018.12.009
1078-8174/

Thus, virtual reality simulation may be an alternative option that could help to provide an appropriate environment for TVS training.

The Medaphor-ScanTrainer is an example of a virtual-reality simulator that has been used widely as a training tool in transvaginal ultrasound.[13] In medical education, simulators must promote a safe self-study environment that provides performance measures (metrics) and useful feedback to learners to support the experiential learning, where the learner interacts with the learning process.[2,14] The ScanTrainer has been found to promote this environment for students and contains automatic feedback on the performed tasks. The principle aim of the simulator is to improve competency in performing transvaginal ultrasound.[13] However, to date, there is no review investigating to what extent transvaginal ultrasound skills in clinical practice can be developed/enhanced using VR simulation.

This systematic review explores the influence of VR-TVS ScanTrainer on TVS skills based on summarizing all peer-reviewed studies published in this field. The PICO model[9–11] (Table 1) was used to develop the central question: What is the impact of simulation-based training (the ScanTrainer) on TVS skills of novice practitioners in ultrasound?

## Methods

The protocols of this systematic review were registered with PROSPERO (International Prospective Register of Systematic Reviews) with registration number: CRD42017065776.[15] The majority of the aspects of this review were reported following PRISMA guidelines.[16]

### *Eligibility criteria, information sources, search strategy*

Only original research, peer-reviewed quantitative or qualitative studies written in English were included. Studies published as oral or poster abstracts without full texts were excluded, as were studies which discussed outcomes other than TVS skills or those which did not use the ScanTrainer simulator as a part of intervention (Table 1). An electronic search of 10 databases (see appendix A) was conducted with no date restriction using the search term (transvaginal OR trans-vaginal OR endo-vaginal OR OBGYN OR obgyn OR obstetric OR gynecology) AND (ultrasound OR (sonography or "ultrasonography") OR (ultrasonography or "echography" or "sonography")) AND (simulation OR simulator). This search was completed in October 2017 through available access to the Library of University College Dublin.

**Table 1**
Selecting criteria using PICO model.

| Research question: what is the impact of simulation-based training on TVS skills of novice practitioners in ultrasound? | |
|---|---|
| P (population) | Any participant in the health domain (e.g., sonographers, students, trainees, novices, obstetricians, gynecologists, residents, physicians, radiographers, midwives, fertility specialists) |
| I (intervention) | Practice on computer-based virtual reality simulation in transvaginal ultrasound (ScanTrainer) |
| C (comparator) | All studies on the role of TVS ScanTrainer in education/training included whether they involved comparisons with control groups (clinical-based practice) or not |
| O (outcome) | All outcomes that influence the TVS skills (e.g. validity of the simulator, learning curves and transferable skills) |

| | Inclusion | Exclusion |
|---|---|---|
| Type of study | Full text original research, peer reviewed studies | Non-original research, abstracts, reviews, editorials |
| Study design | All designs | Economic studies |
| Simulation | VR-TVS simulation (ScanTrainer) | Studies that do not use the ScanTrainer |

### *Study selection*

The titles of the search results were screened, and the relative results were saved in Endnote X8 reference tool. Two reviewers (R1) and (R2) individually screened the abstracts of the retrieved studies based on the inclusion and exclusion criteria. The primary reviewer (R1) obtained the full-text articles of all abstracts that were considered relevant. Non-English studies and papers of oral/poster abstracts were excluded. The two reviewers (R1) and (R2) individually screened the full-text articles and selected the studies based on the eligibility criteria (Table 1). To minimize any potential bias, both reviewers resolved any conflict through discussion and reached consensus on eligible studies for the review.

### *Data extraction*

Data collection was performed using a data extraction sheet, which was used to record study design, participant characteristics, intervention, findings, and outcomes. The sheet was piloted by firstly generating and completing the sheet, and then reviewing the extracted data for its appropriateness in answering the review question; and finally making a few minor alterations to ensure the suitability of the sheet for the main phase of data extraction. Once this was done data was extracted from the selected studies and transcribed onto the sheet, and finally reviewed to eliminate any potential error that could occur during the extraction process.

### *Assessment of risk of bias*

Two assessment tools were used in this review to assess the risk of bias within the selected studies. The Effective Public Health Practice Project (EPHPP) was used to appraise quantitative studies, and the Critical Appraisal Skills Programme (CASP) was employed to assess qualitative studies.[17,18] The strength of the qualitative studies was classified as strong (more than 75% of criteria), moderate (50%–75%) and weak (less than 50%). The quality of quantitative studies was classified into strong, moderate and weak based on the EPHPP scoring system. The level of evidence was classified into A as strong evidence (consistent findings from multiple strong/high quality studies), B as moderate (consistent findings from multiple moderate quality studies), C as weak (consistent findings from multiple weak/low quality studies), D as uncertain evidence (only one study provides the evidence, or inconsistent findings from multiple study) and E as no evidence (no study).[19–21]

### *Data synthesis*

As the included studies varied in design and outcome, a narrative synthesis was used to analyze the collected data.[22] A narrative synthesis is "a form of storytelling", an approach that relies on using words and texts to summarize and explain findings of studies included in a systematic review.[22] This approach is recommended by the Cochrane Consumers and Communication Review Group (CC & CRG) if the review is heterogeneous in nature,[23] meaning that the meta-analysis or any other synthesis, e.g., meta-ethnography for qualitative studies, cannot be used because of the variety of the methodological designs of the review studies.[22] The data synthesis in this review followed the steps of the narrative synthesis which as recommended in CC & CRG guidelines for developing a preliminary synthesis, exploring relationships in the data and assessing the strength of results.[22,23] The step of developing a theory of how the intervention works was not applicable because of the exploratory nature of the synthesis process. The stage of preliminary synthesis involved a thematic analysis of the textual descriptions that were extracted from the selected studies. A framework was provided at the end of this review to summarize the final results of the synthesis process and provide the relationships between the studies (Fig. 2).

## Results

### *Study selection*

The systematic study process is summarized in Fig. 1.

### *Study characteristics*

The articles included in this review (Table 2) are three randomized controlled trials (RCT),[24–26] one mixed method study,[12] one experimental setup study,[27] one associational study,[28] one cross-sectional survey,[29] one cohort study[30] and two qualitative studies.[31,32]

### *Risk of bias of included studies*

Overall the methodological quality of the studies was strong, only one study categorized as weak, and three studies were moderate in strength. All the quantitative studies showed a moderate blinding (Table 3), as the outcomes' assessors within the studies were blinded to the intervention; however, the blinding of participants to the research questions was not discussed. All studies used valid methods to assess their outcomes, and most of the studies reported number and reasons for withdrawals of participants. The qualitative studies were moderate in strength as the credibility of findings was not clarified (Table 4).

### *Synthesis of results*

Four main themes (Fig. 2) emerged from the analyzed data which were Validation (assessment of the validity of the simulator), Learning (using the simulator as a learning tool), Perspectives (the perceptions of participants trained on the simulator), and Transferable skills (skills developed on the simulator can transfer to clinical practice).

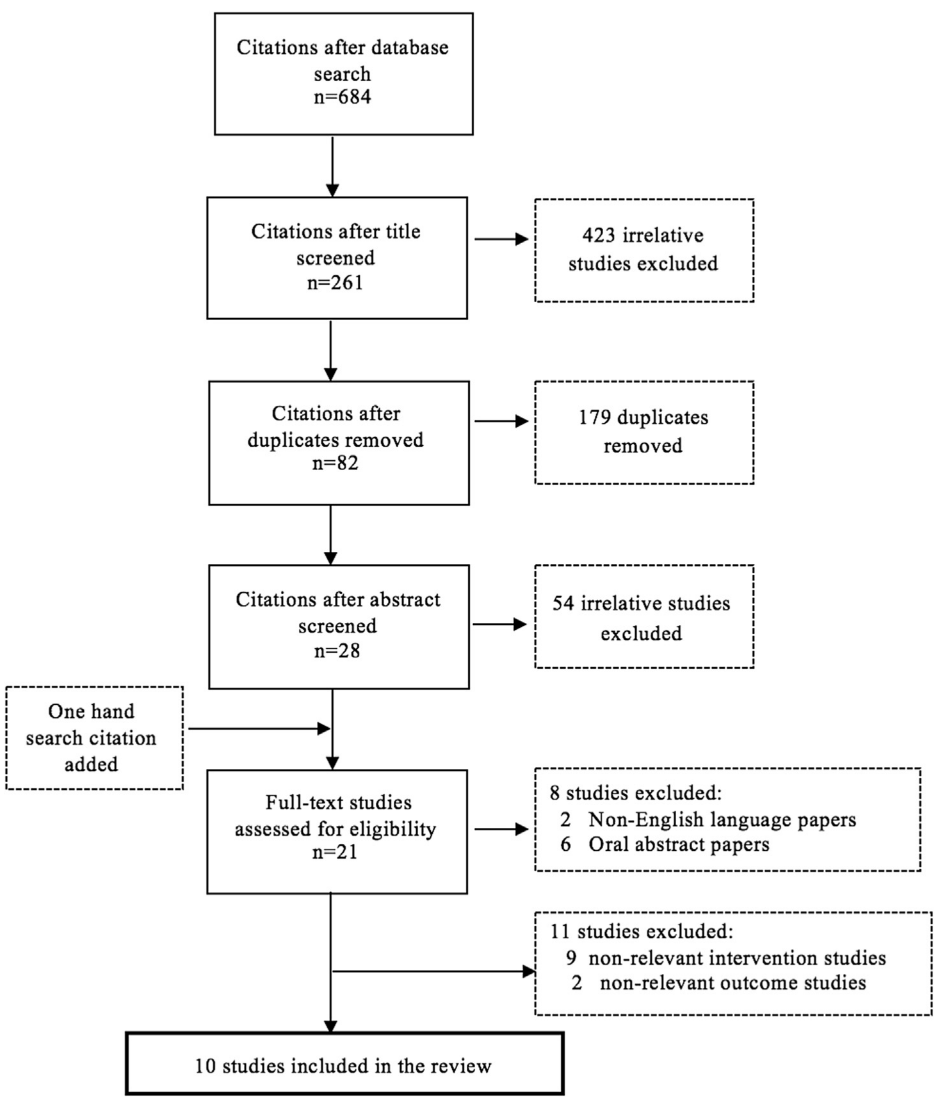


**Figure 1.** Flowchart of search results summarizing the process of selecting studies included in the systematic review.

**Table 2**
Characteristics and qualities of the review studies.

| Study | Country | Design | Sample | Main Theme | Quality |
|---|---|---|---|---|---|
| Tolsgaard (2017)[24] | DK | RCT | 52 residents | Transferable skills | S |
| Madsen (2017)[28] | DK | Associational study | 20 midwives | Learning | S |
| Al-Memar (2017)[30] | UK | Cohort | 24 doctors | Validation | M |
| Patel (2016)[29] | UK | Cross-sectional survey | 70 trainees | Perspectives | W |
| Tolsgaard (2015)[25] | DK | RCT | 26 residents | Transferable skills | S |
| Gibbs (2015)[32] | UK | Qualitative (interviews & thematic analysis) | 25 students | Perspectives | M |
| Chao (2015)[26] | FR | RCT | 34 residents | Learning | S |
| Gibbs (2014)[31] | UK | Qualitative (interviews & thematic analysis) | 12 students | Perspectives | M |
| Madsen (2014)[27] | DK | Cohort (experimental) | 28 doctors | Validation | S |
| Williams (2013)[12] | UK | Mixed method (triangulation) | 9 doctors | Perspectives | S |

DK: Denmark; UK: United Kingdom; FR: France; RCT: randomized controlled trial; S: strong; M: moderate; W: weak.

**Table 3**
EPHPP appraisal results of the quantitative studies.

| Study | Selection bias | Study design | Confounders | Blinding | Data collection method | Withdrawals | Overall |
|---|---|---|---|---|---|---|---|
| Tolsgaard (2017)[24] | 1 | 1 | 1 | 2 | 1 | 1 | S |
| Madsen (2017)[28] | 2 | 2 | 1 | 2 | 1 | 1 | S |
| Al-Memar (2017)[30] | 2 | 2 | 1 | 2 | 3 | 1 | M |
| Patel (2016)[29] | 3 | 2 | 1 | 2 | 1 | 3 | W |
| Tolsgaard (2015)[25] | 1 | 1 | 1 | 2 | 1 | 2 | S |
| Chao (2015)[26] | 2 | 1 | 1 | 2 | 1 | 1 | S |
| Madsen (2014)[27] | 2 | 2 | 1 | 2 | 1 | 1 | S |
| Williams (2013)[12] | 2 | 1 | 1 | 2 | 2 | 1 | S |

EPHPP rating. 1: strong; 2: moderate; 3: weak; S: strong (no weak rating); M: moderate (one weak rating); W: weak (two or more weak ratings).

**Table 4**
CASP appraisal results of the qualitative studies.

| Study | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Total score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Gibbs (2015)[32] | Yes | Yes | Yes | Yes | Unclear | Unclear | Yes | Unclear[a] | Unclear[b] | Yes | 6 |
| Gibbs (2014)[31] | Yes | Yes | Yes | Yes | Unclear | Unclear | Yes | Unclear[a] | Unclear[b] | Yes | 6 |
| Williams (2013)[12] | Yes | Unclear | Yes | Yes | Unclear | Unclear | Yes | Unclear[a] | Unclear[b] | Yes | 5 |

CASP appraisal questions: 1. Was there a clear statement of the aims of the research? 2. Is a qualitative methodology appropriate? 3. Was the research design appropriate to address the aims of the research? 4. Was the recruitment strategy appropriate to the aims of the research? 5. Was the data collected in a way that addressed the research issue? 6. Has the relationship between researcher and participants been adequately considered? 7. Have ethical issues been taken into consideration? 8. Was the data analysis sufficiently rigorous? 9. Is there a clear statement of findings? 10. Is the research valuable?[18]

[a] The analysis process not in-depth described.
[b] The credibility of findings not discussed.

## Discussion

The search strategy in this review resulted in 10 studies that evaluated the usage of the ScanTrainer as a training tool in transvaginal ultrasound. All studies were published between 2013 and 2017 and conducted in three countries which are the UK, Denmark, and France. The populations of the studies were OB/GYN doctors, residents, midwives, medical trainees, and ultrasound program students; with sample sizes ranging from 9 to 70 participants (Table 3). The outcomes of all ten studies were analyzed using thematic analysis which resulted in four themes: Validation, Learning, Perspectives, and Transferable skills.

### *Validation*

Simulation validity in medical education refers to the actuality of whether the simulator teaches what it is intended to teach. Face validity measures the realism of the simulator; content validity measures the suitability of the simulator as a learning modality.[33] Criterion validity compares the performance results from the simulator with those from another technique, such as clinical performance.[33,34] Construct validity measures the ability of the simulator to distinguish between expert and non-expert users.[33] The virtual reality simulators should demonstrate, at least, face, content and construct validity to be considered as valid training/learning tools for clinical practice.[33,34] Regarding the TVS ScanTrainer, three studies[27,28,30] showed strong evidence that the simulator was valid as a training tool in transvaginal ultrasound and it might represent the reality of TV scanning. Al-Memar's study[30] supported the face and content validity of the simulator relying on collecting satisfaction rates of 8 ultrasound experts (Gynecology doctors who performed more than 500 scans) on using the simulator as a training tool for TVS. The rating was obtained using a questionnaire following completion of two modules on the simulator (on basic gynecology and early pregnancy). The rating was based on how much the experts were satisfied with the statement that scanning on the simulator resembles scanning in reality (face validity) and on the ability of the simulator to teach TVS skills (content validity). The majority of the experts were satisfied with the realism of the simulator (88%; n = 7) on basic gynecology and (75%; n = 6) on early pregnancy. Similarly, 88% (n = 7) were satisfied with the usefulness of the simulator content to learn TVS skills. It could be argued that the number of experts in this study is relatively small, nevertheless, with searching the literature, no guideline was found that describes the required criteria to validate a

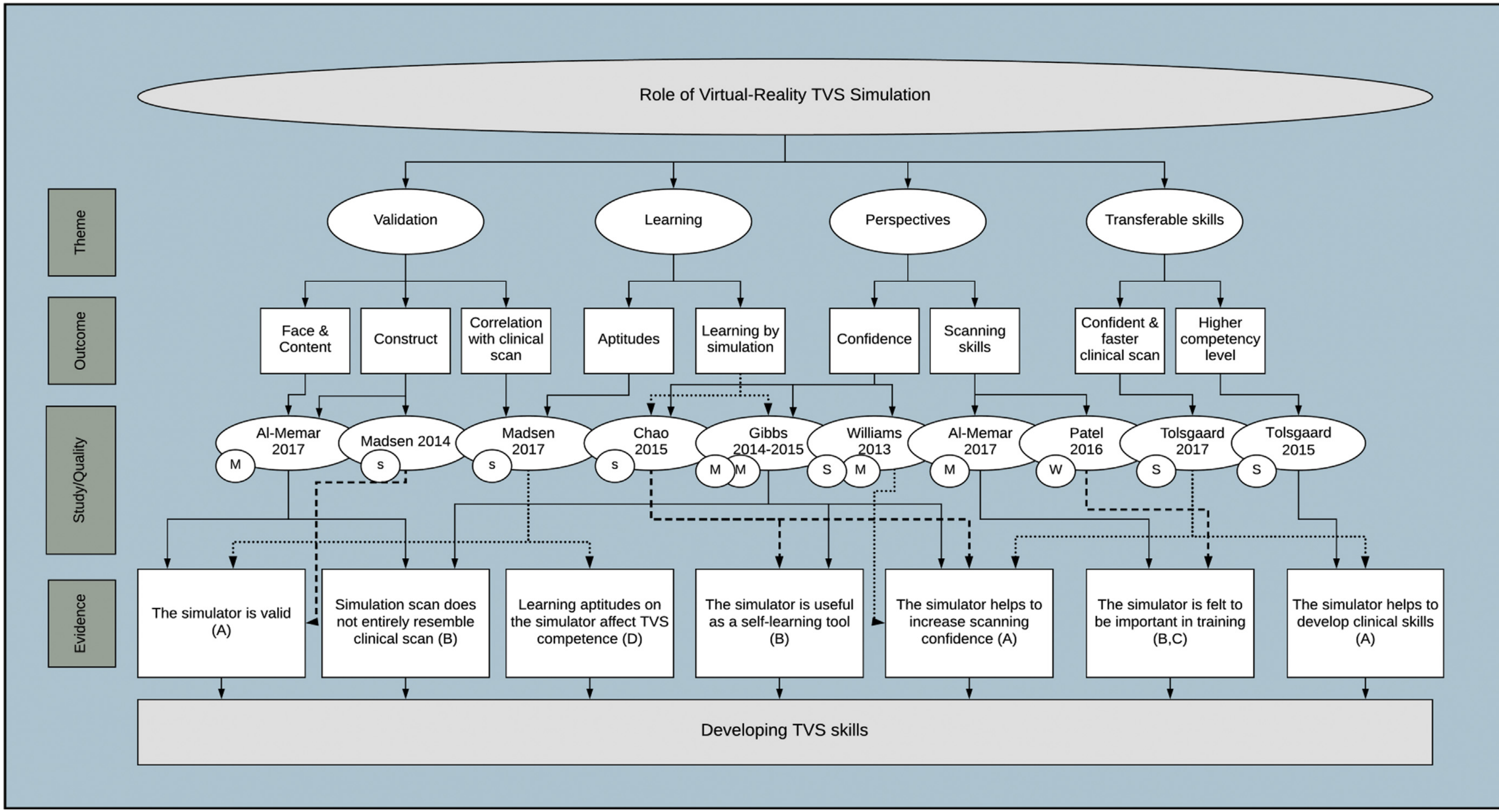


**Figure 2.** A summary of the systematic review showing the evidence synthesized from the selected studies. S: strong study; M: moderate study; W: weak study; A: evidence from multiple strong studies; B: evidence from multiple moderate studies; C: evidence from multiple weak studies, D: evidence from one study.

simulator not to mention the size of the sample needed for optimum validation.[34]

The evidence, also, demonstrates the construct validity of the ScanTrainer. Al-Memar's study[30] indicated that the simulator could distinguish between different levels of competency in TVS. All the participants of this study (n = 24) performed the two modules on basic gynecology and early pregnancy on the simulator, and the assessment scores of these modules showed a statistically significant difference between experts and novices (P = 0.002 for gynecology, P = 0.03 for early pregnancy).[30] Similarly, the findings of Madsen's study[27] supported the construct validity of the simulator. This study recruited 16 medical students, who were novices in ultrasound, and 12 OB/GYN consultants; and both groups completed seven modules on the simulator. The sum scores of these modules showed significant difference (P < 0.001) between the students (43.8%, range 17.9–68.9%) and the consultants (82.8%, range 60.4–91.7%).[27]

A further study supported the validity of the modules of the simulator to be used as an assessment tool in TVS training.[28] Previous research in medical education has suggested that to validate the simulator as suitable for competency assessment, the practice performance on the simulator should predict or at least correlate with the performance in the clinical setting.[33] This evidence was found in Madsen's study[28] in which 20 midwives, with no experience in ultrasound, were recruited and completed the modules of the simulator, and then six of them continued to clinical training. The mean of the performance scores calculated for the midwives on the simulator setting showed a significant correlation with the clinical mean performance scores (Pearson correlation 0.81; P = 0.049), suggesting that the practice performance on the simulator may predict the performance in clinical training; however, it must be noted that small number of participants who were assessed in clinical training (n = 6). This evidence could validate using the simulator as a performance assessment tool; however, it requires confirmation in the future with another study involving a larger sample size since this study was presented as a pilot.

*Learning*

A moderate level of evidence was found in four studies that the simulator can be used as a learning tool that could help to develop TVS skills.[12,26,31,32] One of the studies showed that the participants who underwent 10 h of training on the simulator (n = 6 doctors) attained higher scores on the post-test simulator assessment than those who completed 10 h of clinical training (n = 5 doctors). Although the difference between the average scores of the two groups was not statistically significant on the post-test assessment (U = 13, P = 0.0556), this difference was closer to the alpha value of 0.05 than this in the pre-test assessment (U = 12, P = 0.6623).[12] Additional study found that the TVS learners (n = 16 residents) who performed a self-learning session on the simulator showed higher competence in acquiring simulation images than those who attended a traditional lecture (n = 18 residents) of how to use the simulator and to obtain sagittal and coronal images.[26] Similar results were found in the qualitative approach of both Gibbs' studies.[31,32] The results of interviews with participants (n = 25 students,[32] n = 12 students[31]) and mentors (n = 14)[32] showed that the students learned best by performing the simulator tasks rather than listening to lectures. For examples, one of the students said: "I found I remembered things I learnt on the simulator more than being told about them",[31] additionally, a mentor said "I spent less time this year teaching my student the basics of performing a scan. She'd managed to book time on the simulator before she scanned patients in the department".[32] These findings support the theory that the experiential learning is more efficient than the traditional teaching in meeting the educational objectives.[14]

Conversely, one of the studies suggests that the learning aptitudes could affect the TVS competence in clinical practice.[28] Madsen's study (n = 6 midwives who were assessed on patients) demonstrated that of the fast learners on the simulator, 50% (n = 3) of the participants who completed the simulator tasks efficiently in a shorter time than the others, provided better TVS performance in clinical settings than slow learners (P = 0.025).[28] These findings suggest that the simulator could potentially be used to predict the duration of training that the learner would need to improve his/her TVS skills in the clinical setting based on his/her learning aptitudes on the simulator. However, it must be highlighted that the findings were based on a small number of participants and further larger studies to confirm this finding are required.

*Perspectives*

Strong evidence was found in five studies that the simulator helped to increase confidence in TV scanning. In William's study,[12] all participants (n = 5 doctors) who experienced simulation practice felt confident in their first clinical scan. Also, in Chao's study,[26] all respondents (n = 29 residents) agreed that the simulation could increase scanning confidence. Both Gibbs' studies results,[31,32] showed that the simulation practice aids to build confidence in TVS. Tolsgaard's study[24] confirmed these opinions when the intervention group, who practiced on the simulator before starting the clinical training (n = 28 residents) attained higher confidence scores than the control group (n = 26 residents).[24]

Two studies demonstrated that the simulator could be useful in developing TVS skills based on the participants' perceptions.[29,30] In Al-Memar's study,[30] 83% of the participants (n = 24 Gyne doctors) believed that the VR simulation has a role in ultrasound training. However, Patel's study[29] reported that 46% of the participants (n = 32 trainees in OBGYN) agreed that the simulator was helpful in improving scanning skills, 31% were neutral, and 14% disagreed. The trainees in Patel's study varied in the stage of specialty training (ST), and the participants who were in the basic stage (ST1-2) and intermediate stage (ST3-5) found the simulator was useful; however, those of the advanced stage (ST6-7) felt neutral with the benefit of the simulator.[29] This result suggests that the simulator could be particularly useful for learners who have little experience in ultrasound scanning and those who have poorly developed psychomotor skills (i.e., hand-eye coordination skills[3]); however, for those who have mastered orientation and have gained prior experience in ultrasound the simulator could be less valuable.

*Transferable skills*

In two studies, Tolsgaard[24,25] reported findings that indicated the skills developed in simulation practice can transfer to clinical TVS performance which involved new residents, (n = 33)[25] and (n = 54),[24] as a part of OBGYN training. Both studies showed that practicing on simulation aids to improve clinical competence and increase scanning speed in transvaginal ultrasound.[24,25] Clinical competence was measured by a performance checklist tool in one study[25] and patients' ratings on perceived safety and discomfort in the other study.[24] The studies reported that the intervention group achieved 20.1% higher scores in the checklist tool (95% CI, 11.1–29.1),[25] 7.9% higher safety ratings (95% CI, 0.5–14.7)[24] and 18.5% decrease in discomfort ratings (95% CI, 10.7–25.5)[24] than the control group. Moreover, the study[24] showed that the duration of scans performed by the intervention group were reduced by 20% (1 min 32 s (95% CI), 7 s to 2 min 66 s) compared to the control group. Additionally, the results of Chao's study[26] supported this finding by illustrating that 86% of the intervention group reported that training on simulation would increase scanning speed.[26] These findings indicate that the skills which have been developed within simulation-based training may transfer effectively to clinical practice.

It is important to note that both Tolsgaard's studies[24,25] used two types of TVS simulation: VR-TVS simulator and a pelvic mannequin.[24,25] Therefore, the positive effects of this intervention cannot be assigned solely to the VR-TVS simulation. At the same time, Madsen's study showed that there was a significant correlation ($P = 0.049$) between the learning curves on the VR-TVS simulator and the learning curves on the clinical scans; however, no correlation was found between the learning curves on the mannequin-based simulator and the clinical learning curves.[28]

### *Impact of the simulator on the TVS skills*

Although the current literature shows positive insights on the usefulness of the ScanTrainer as a training tool (Fig. 2), it provides limited evidence on the actual impact of the simulator on developing TVS skills. Most of the studies reviewed followed subjective approaches to investigate the usefulness of the simulator in developing TVS skills, which were based on the perceptions of the learner or the trainee on the use of the simulator.[12,29–32] Only two studies in the current literature[24,25] investigated the efficiency of the simulation practice, on new residents, based on measurable outcomes such as duration of scans, number of supervised scans, number of repeated scans and competency scores on assessment scales.[35] Both studies suggested that the simulation-based practice has positive effects on the TVS skills and it might potentially reduce the duration of clinical training; however, the studies allowed the trainees to practice on phantoms using an ultrasound machine that was similar to that being used in the clinical setting, such an intervention could influence the final results, as has been discussed earlier.

The paucity of literature including objective and accurate measurement of the effect of using the ScanTrainer as a training tool in developing TVS skills has been identified. This may be due to the complexity of defining measurable skills in ultrasound imaging especially of those which could be improved by using the simulation-based practice. The primary purpose of the ScanTrainer is to develop and improve the psychomotor skills, which if enhanced could positively influence the practitioner's confidence and therefore improve the clinical practice. However, ultrasound skills are not limited to image acquisition; in fact, image interpretation, report writing, and communication skills are other skills that also considered essential in optimizing sonographer competency.[4] Such skills have not been adequately investigated in the literature and also require further investigation. The current literature predominantly focusses on medical residents,[24,25] and therefore, the generalizability of the results to other ultrasound practitioners such as radiographers and midwives requires further research.

## Strengths and limitations

The review was strengthened with the systematic protocols which were registered in PROSPERO and following the PRISMA guidelines. However, there are some limitations to acknowledge in the review. Although ten databases were searched for eligible studies, the search term was limited which could have reduced the rigor of this systematic review, and PICO criteria should be used to inform precisely the search terms. Only English studies were included, and no search was made on the grey literature. No third party was asked to resolve any difference of opinion with eligibility of articles for inclusion. The included studies were heterogeneous, and not all the participants were novices in ultrasound; however, this limitation was minimized by employing the narrative approach which involved discussion of the main themes which emerged from the studies to synthesize the final evidence.

## Conclusion

Limited peer-reviewed literature on the role of the transvaginal simulator as a multidisciplinary training tool in TV ultrasound was identified. This literature typically related to medical trainees and employed research methodologies which lacked consistency in metrics to measure skill acquisition upon use. A positive influence upon ultrasound training was indicated in the research performed to date; however, future studies should incorporate a broader spectrum of professionals to reflect the health professions routinely involved in obstetric and gynecological ultrasound scanning. Furthermore, the upcoming studies should be designed around measurable effects of simulation-based training on clinical practice and skill acquisition and should consider the prior level of experience and previous exposure to ultrasound.

## Funding

King Saud University, Ministry of Education, Saudi Arabia.

Royal Embassy of Saudi Arabia, Cultural Bureau in Dublin, Ireland.

## Conflict of interest

None

## Acknowledgment

The author acknowledges King Saud University which supported her study through the Royal Embassy of Saudi Arabia, Cultural Bureau in Dublin, under Grant IRKS1601.

## Appendix A

Search term & databases.

| Search Term | Database |
|---|---|
| (transvaginal OR trans-vaginal OR endo-vaginal OR OBGYN OR obgyn OR obstetric OR gynecology) AND (ultrasound OR (sonography or "ultrasonography") OR (ultrasonography or "echography" or "sonography")) AND (simulation OR simulator) | Medline (Ovid)<br>Medline (PubMed)<br>Sage<br>EMBASE<br>BESCO (Academic Search Complete)<br>BESCO (British Education Index)<br>EPPI (Centre Database of Education Research)<br>ERIC (ProQuest)<br>CINAHL Plus (EBSCO host)<br>Australian Education Index (ProQuest Dialog) |

## Appendix B

Results of data extraction (ordered from latest publications).

| Study | Intervention | Findings | Outcomes |
|---|---|---|---|
| Tolsgaard (2017)[24] | 4 weeks simulation-based training (VR & mannequin) | ➢ IG: 18.5% (95% CI, 10.7–25.5) lower patient discomfort scores,<br>➢ IG:7.9% (95% CI, 0.5–14.7) higher scores of patient perceiving safety,<br>➢ IG: less scan duration by an average of 1 min 32 s (95% CI, 7 s to 2 min 66 s) than control group.<br>➢ IG:11.1% (95% CI, 2.5–18.9) higher scores in patient confidence<br>➢ IG: odds of trainee supervision and repeated examination were reduced by 45.3% (95% CI, 33.5–55.1) | Duration of clinical scan<br>Patient care<br>Repeated examinations<br>Supervision |
| Madsen (2017)[28] | 1st VR simulation<br>2nd pelvic mannequin<br>3rd clinical practice | ➢ Mean performance scores on VR simulation were significantly correlated with clinical mean performance scores (Pearson correlation 0.81; P = 0.049)<br>➢ Fast learners on simulation demonstrated higher scores in clinical practice (P = 0.025).<br>➢ Significant difference between fast learners and slow learners in age (P = 0.002), and in clinical experience (P = 0.011). | Predicting US clinical performance using VR simulation<br>Learning aptitudes |
| Al-Memar (2017)[30] | VR-TVS simulation | ➢ At least 75% of participants believed simulation is important in ultrasound training<br>➢ Face validity: 75% satisfied on the realism of the target objects and on the image appearance<br>➢ 50% not satisfied on the instrument handling, and around 60% not satisfied on the force sensation of the probe.<br>➢ Content validity:88% of experts were satisfied on the usefulness of the contents on the simulation to help acquiring ambidexterity skills<br>➢ 100% of experts satisfied on the overall usefulness of content of the simulation in learning ultrasound skills.<br>➢ Significant difference between experts scores and novice scores on the simulation. In gynecology modules, mean of experts score = 40.8, and of novice = 33.3(P = 0.001), in early pregnancy mean of experts 17.8, mean of novice scores = 15.9 (P = 0.007).<br>➢ Screams due to over pressure of the transducer were significantly less with experts than with novices<br>➢ No significant difference between groups on the time taken on the assessments | Validation of the simulator |
| Patel (2016)[29] | Simulation training | ➢ VR-TVS simulation was available to 50% (n = 35) of respondents. 46% of them (n = 16) found simulation was helpful in improving scanning skills, (31%, n = 11) were neutral and (14%, n = 5) disagreed | Perceptions of trainees on using simulation |
| Tolsgaard (2015)[25] | 4 weeks simulation-based training (VR & mannequin | ➢ IG: 20.1% (95% CI, 11.1–29.1) higher competence scores | Clinical performance |
| Gibbs (2015)[32] | VR-TVS simulation | ➢ advantages (ability to repeat the examinations without pressure, having adequate time to learn anatomy and pathology, providing of better understanding of theoretical concepts)<br>➢ Limitations (lack of skills in relation to patient interactions, decision making, report writing) | Experience of students on the simulation practice |
| Chao (2015)[26] | VR-TVS simulation<br>40 min self-learning vs. 40 min. lecture on obtaining VR images | Mean score of IG (12; SEM, 0.8) was significantly higher than the mean score of CG (9; SEM, 1.0; P = 0.0302)<br>➢ 15 participants of CG and 14 participants of IG filled the questionnaire. 100% agreed on that simulation increased confidence in scanning.<br>➢ 94% CG & 86% IG thought simulation would increase scan speed<br>➢ 100% both thought simulation would increase capacity to produce images<br>➢ 100% found conventional teaching remained useful in addition to VR simulation | Impact of VR simulation training on obtaining TVS images |
| Madsen (2014)[27] | VR-TVS simulation | ➢ 62.9% of maximum score was the pass/fail level<br>➢ 88.4% of the maximum score was the proficiency level<br>➢ novices reached proficiency level in a median of 219 min (range, 150 –251 min) | Validation of the simulator metrics |
| Gibbs (2014)[31] | VR-TVS simulation | 4 themes:<br>➢ "Opportunities to build confidence in a non-threatening environment"<br>➢ "The enjoyment of simulation as a learning strategy"<br>➢ "How simulation supports the implementation of theory into practice"<br>➢ "Suggestions for improving the simulator experience"[31] | Students' experience on the simulation practice |
| Williams (2013)[12] | VR-TVS simulation | ➢ No statistically significant difference between the mean scores (post-test) of the groups (IG, CG) (P = 0.0556).<br>➢ 8/9 participants felt that simulation could be helpful in increasing confidence<br>➢ IG felt confident within first clinical scan | Simulation practice and confidence |

## References

1. Maul H, Scharf A, Baier P, Wüstemann M, Günter HH, Gebauer GG, et al. Ultrasound simulators: experience with the SonoTrainer and comparative review of other training systems. *Ultrasound Obstet Gynecol* 2004;**24**(5):581–5.
2. Sidhu HS, Olubaniyi BO, Bhatnagar G, Shuen V, Dubbins P, et al. Role of simulation-based education in ultrasound practice training. *J Ultrasound Med* 2012;**31**(5):785–91.
3. Nicholls D, Sweet L, Hyett J. Psychomotor skills in medical ultrasound imaging. *J Ultrasound Med* 2014;**33**(8):1349–52.

4. Harrison G, Dolbear G. *Standards for sonographic education*. Consortium for the Accreditation of Sonographic Education; 2018.
5. SCoR and BMUS. *Guidelines for professional ultrasound practice*. 2017.
6. ISUOG. ISUOG Education Committee recommendations for basic training in obstetric and gynecological ultrasound: ISUOG Recommendations. *Ultrasound Obstet Gynecol* 2014;**43**(1):113–6.
7. RCR. *Ultrasound training recommendations for medical and surgical specialties*. 3rd ed. The Royal College of Radiologists; 2016.
8. Parker P, Harrison G. Educating the future sonographic workforce: membership survey report from the British Medical Ultrasound Society. *Ultrasound* 2015;**23**(4):231–41.
9. Baker C. *NHS key Statistics: England, May 2018*. House of Commons Library; 2018.
10. Gould DA, Reekers JA, Kessel DO, Chalmers NC, Sapoval M, Patel AA, et al. Simulation devices in interventional radiology: caveat emptor. *Cardiovasc Interv Radiol* 2006;**29**(1):4–6.
11. Alfred A, Katherine KM, Carol BB, Trish C, Lori C, Peter MD, et al. Obstetric and gynecologic ultrasound curriculum and competency assessment in residency training programs: consensus report. *J Ultrasound Med* 2018;**37**(1):19–50.
12. Williams CJ, Edie JC, Mulloy B, Flinton MD, Harrison G, et al. Transvaginal ultrasound simulation and its effect on trainee confidence levels: a replacement for initial clinical training? *Ultrasound* 2013;**21**(2):50–6.
13. Chalouhi G, Bernardi V, Ville Y. Ultrasound simulators in obstetrics and gynecology: state of the art. *Ultrasound Obstet Gynecol* 2015;**46**(3):255–63.
14. Fanning RM, Gaba DM. The role of debriefing in simulation-based learning. *Simulat Healthc J Soc Med Simulat* 2007;**2**(2):115–25.
15. Almestehi M, Alomaim W, Stanton M, Stokes D, Moran M. *The role of transvaginal ultrasound simulation in education*. 2017 [cited 2017]; Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017065776.
16. Moher D, Altman DG, Liberati A, Tetzlaff J, Takkouche B, Norman G. PRISMA statement. *Epidemiology* 2011;**22**(1). 128-128.
17. EPHPP, Effective Public Health Practice Project. *McMaster evidence review & synthesis centre*. 2017. Available from: https://merst.ca/ephpp/Quality [Accessed 2017].
18. CASP, Critical Apprasial Skills Programme. *CASP (qualitative research) checklist*. 2017 [online] Available at: www.casp-uk.net [Accessed 2017].
19. Owens DK, Lohr NK, Atkins D, Treadwell JR, Reston JT, Bass EB, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—agency for healthcare research and quality and the effective health-care program. *J Clin Epidemiol* 2010;**63**(5):513–23.
20. Tulder vMW, Furlan A, Bombardier C, Bouter LM, Editorial Board Cochrane LMC, G. Editorial Board of the Cochrane Collaboration Back Review, et al. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back review group. *Spine* 2003;**28**(12):1290–9.
21. Group, G.W., Grading of recommendations assessment, development and evaluation (GRADE). [Access date, 2012], vol. 23.
22. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. *A product from the ESRC methods programme Version* 2006;**1**:b92.
23. Ryan R. *Cochrane Consumers and communication review group: data synthesis and analysis*. Cochrane Consumers and Communication Review Group; 2013. Available at: http://cccrg/. cochrane. org [Accessed 2016].
24. Tolsgaard MG, Ringsted C, Rosthoj S, Norgaard L, Moller L, Freiesleben NC, et al. The effects of simulation-based transvaginal ultrasound training on quality and efficiency of Care: a Multicenter Single-blind randomized trial. *Ann Surg* 2017;**265**(3):630–7.
25. Tolsgaard MG, Ringsted C, Dreisler E, Norgaard LN, Petersen JH, Madsen ME, et al. Sustained effect of simulation-based ultrasound training on clinical performance: a randomized trial. *Ultrasound Obstet Gynecol* 2015;**46**(3):312–8.
26. Chao C, Chalouhi GE, Bouhanna P, Ville Y, Dommergues M. Randomized clinical trial of virtual reality simulation training for transvaginal gynecologic ultrasound skills. *J Ultrasound Med* 2015;**34**(9):1663–7.
27. Madsen ME, Konge L, Norgaard LN, Tabor A, Ringsted C, Klemmensen AK, et al. Assessment of performance measures and learning curves for use of a virtual-reality ultrasound simulator in transvaginal ultrasound examination. *Ultrasound Obstet Gynecol* 2014;**44**(6):693–9.
28. Madsen ME, Norgaard LN, Tabor A, Konge L, Ringsted C, Tolsgaard MG. The predictive value of ultrasound learning curves across simulated and clinical settings. *J Ultrasound Med* 2017;**36**(1):201–8.
29. Patel H, Chandrasekaran D, Myriokefalitaki E, Gebeh A, Jones K, Jeve YB, et al. The role of ultrasound simulation in obstetrics and gynecology training: a UK trainees' Perspective. *Simulat Healthc J Soc Med Simulat* 2016;**11**(5):340–4.
30. Al-Memar M, Saso S, Bobdiwala S, Ameye L, Stalder C, Joash K, et al. Validation of a virtual reality simulator for the use of transvaginal ultrasonography in gynaecology and early pregnancy. *Aust J Ultrasound Med* 2017;**20**(3):97–105.
31. Gibbs V. An investigation into sonography student experiences of simulation teaching and learning in the acquisition of clinical skills. *Ultrasound* 2014;**22**(3):173–8.
32. Gibbs V. The role of ultrasound simulators in education: an investigation into sonography student experiences and clinical mentor perceptions. *Ultrasound* 2015;**23**(4):204–11.
33. McDougall EM. Validation of surgical simulators. *J Endourol* 2007;**21**(3):244–7.
34. Schout BMA, Hendrikx AJM, Scheele F, Bemelmans BLH, Scherpbier AJJA. Validation and implementation of surgical simulators: a critical review of present, past, and future. *Surg Endosc* 2010;**24**(3):536–46.
35. Tolsgaard MG, Todsen T, Sorensen JL, Ringsted C, Lorentzen T, Ottesen B, et al. International Multispecialty consensus on how to evaluate ultrasound competence: a Delphi consensus survey. *PLoS One* 2013;**8**(2):e57687.