



● *Original Contribution*

IMMERSIVE VIRTUAL REALITY IN BASIC POINT-OF-CARE ULTRASOUND TRAINING: A RANDOMIZED CONTROLLED TRIAL

NANNA L. ANDERSEN,^{*} RUNE O. JENSEN,[†] LARS KONGE,^{‡,§} CHRISTIAN B. LAURSEN,^{¶,||}
CASPER FALSTER,^{¶,||} NIELS JACOBSEN,^{¶,||} MOHAMMAD T. ELHAKIM,^{‡,#} JONAS A. BOJSEN,^{‡,#}
MINNA RIISHEDE,^{**} MAJA L. FRANSEN,[#] BENJAMIN S.B. RASMUSSEN,^{‡,#} STEFAN POSTH,^{††}
LINUS SANT,[#] and OLE GRAUMANN^{‡,#}

^{*} Faculty of Health Sciences, Medicine, University of Southern Denmark, Odense, Denmark; [†] Research and Innovation Unit of Radiology, University of Southern Denmark, Odense, Denmark; [‡] Institute of Clinical Research, University of Southern Denmark, Odense, Denmark; [§] Copenhagen Academy for Medical Education and Simulation (CAMES), Copenhagen University, Copenhagen, Denmark; [¶] Department of Respiratory Medicine, Odense University Hospital, Odense, Denmark; ^{||} Odense Respiratory Research Unit (ODIN), Department of Clinical Research, University of Southern Denmark, Odense, Denmark; [#] Department of Radiology, Odense University Hospital, Odense, Denmark; ^{**} General practice, Region of Southern Denmark, Denmark; and ^{††} Department of Emergency Medicine, Odense University Hospital, Odense, Denmark

(Received 14 January 2022; revised 5 August 2022; in final form 22 August 2022)

Abstract—This study was aimed at comparing the learning efficacy of a traditional instructor-led lesson with that of a completely virtual, self-directed lesson in immersive virtual reality (IVR) in teaching basic point-of-care ultrasound (PoCUS) skills. We conducted a blinded, non-inferiority, parallel-group, randomized controlled trial in which final-year medical students were randomized to an instructor-led (n = 53) or IVR (n = 51) lesson. Participants' learning efficacy was evaluated by blinded assessors, who rated each participant's performance using the Objective Structured Assessment of Ultrasound Skills (OSAUS) assessment tool.

The mean total scores for participants were 11.0 points (95% confidence interval: 9.8–12.2) for the instructor-led lesson and 10.3 points (95% confidence interval: 9.0–11.5) for the IVR lesson. No significant differences were observed between the groups with respect to total score ($p = 0.36$) or subgroup objectives of the OSAUS score ($p = 0.34$ for familiarity, $p = 0.45$ for image optimization, $p = 0.96$ for systematic approach and $p = 0.07$ for interpretation). Maintenance costs for both courses were estimated at 400 euros each. Startup costs for the instructor-led course were estimated 16 times higher than those for the IVR course. The learning efficacy of an instructor-led lesson on basic US did not differ significantly from that of a self-directed lesson in IVR, as assessed using the OSAUS. The results suggest that IVR could be an equivalent alternative to instructor-led lessons in future basic US courses, but further research is warranted to clarify the role of IVR in PoCUS courses. (E-mail: nanselind@hotmail.com) © 2022 The Author(s). Published by Elsevier Inc. on behalf of World Federation for Ultrasound in Medicine & Biology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Key Words: Ultrasound education, Immersive virtual reality, Medical students.

INTRODUCTION

Point-of-care ultrasound (PoCUS) is performed bedside and yields dynamic imaging that can be easily correlated with the patient's symptoms. PoCUS is used in several specialties for procedural, diagnostic and screening applications and is easily repeatable (Moore and Copel 2011). The recent development of portable ultrasound

(US) scanners has made low-cost clinical imaging increasingly accessible, making it easier for clinicians as well as non-physician health care personnel to perform bedside PoCUS with high-quality imaging in new settings (Nielsen et al. 2019; Pietersen et al. 2021). Hands-on training and proper supervision are required to build the competencies of non-radiologists in obtaining and interpreting US images to aid their clinical decision making. Therefore, introducing basic US early in a physician's career seems desirable. Several studies on US courses for undergraduate medical students have already

Address correspondence to: Nanna L. Andersen, Kloevervaenget 10, Entrance 112, 3rd floor, 5000 Odense C, Denmark. E-mail: nanselind@hotmail.com

been performed, but the training methods, outcome measures and assessment criteria vary widely (Amini *et al.* 2015; Skalski *et al.* 2015; Pietersen *et al.* 2018; Tarique *et al.* 2018; Kahr Rasmussen *et al.* 2019; Carstensen *et al.* 2022).

Virtual reality (VR) is advancing as an educational tool in both pre- and postgraduate medical training (Jensen and Konradsen 2018; Hooper *et al.* 2019; Weiner *et al.* 2019; Tursø-Finnich *et al.* 2022). Immersive VR (IVR) involves use of a head-mounted device (HMD) that allows the user to observe and move around in a simulated, virtual 3-D environment, while controllers or hand-tracking allow the user to interact with it. IVR offers an engaging, risk-free learning environment that does not require the presence of busy faculty, and it gives the user the opportunity to practice scenarios as well as technical skills (Latham *et al.* 2019). With the launch of the Oculus Rift HMDs in 2015 by Oculus, the technology has become more widely available and could potentially be used in US training for medical students (Pottle 2019; Saldana *et al.* 2020). Although only a few studies have explored the use of IVR in US education (Hu *et al.* 2020; Andersen *et al.* 2021; Rosenfeldt Nielsen *et al.* 2021), IVR could reduce the number of teaching hours and costs in future US training. Furthermore, IVR minimizes face-to-face contact, which is favorable in times when education is challenged by infectious disease outbreaks, as has been the case with the COVID-19 pandemic (Nayahangan *et al.* 2021). Even if IVR could be cost-effective when compared with more classical teaching methods and hands-on training, it is paramount that the learning efficacy is not inferior to that of the more classic methods.

The aim of this randomized controlled trial was to compare the learning efficacy of a traditional instructor-led lesson with that of a completely virtual, self-directed lesson in IVR applied to a course on basic US skills for final-year medical students. In this study, learning efficacy was defined as “the degree of confidence in application of skills and knowledge taught at a training session” (Srivastava *et al.* 2019), specifically evaluated through the internationally approved Objective Structured Assessment of Ultrasound Skills (OSAUS) (Tolsgaard *et al.* 2013).

METHODS

Study design

This study was a blinded, non-inferiority, parallel-group, randomized controlled trial. The allocation ratio was 1:1.

Participants and setting

A total of 110 final-year medical students attended a PoCUS course. Participation in the US course and the

subsequent assessment was mandatory, while enrollment in our study was voluntary. The study was reviewed by the Committee for Ethics in Research at the Research and Innovation Unit of Radiology, University of Southern Denmark, which concluded that the project included no biomedical intervention or other action that imposed a risk for participants. Thus, the project was not notifiable to the Research Ethics Committee System.

All students were assigned a personal but anonymized study number on their enrollment at the university. This number was used for identification of the participants in this study. As preparation, all participants attended a 45-minute virtual lecture on basic US techniques. The learning goals were the use of transducers, knobology and optimization of image quality. The students attended their first hands-on lesson in basic US the day after the lecture. Groups were trained in separate classrooms and kept apart from each other before the assessment. Figure 1 is an overview of the course and the subsequent assessment.

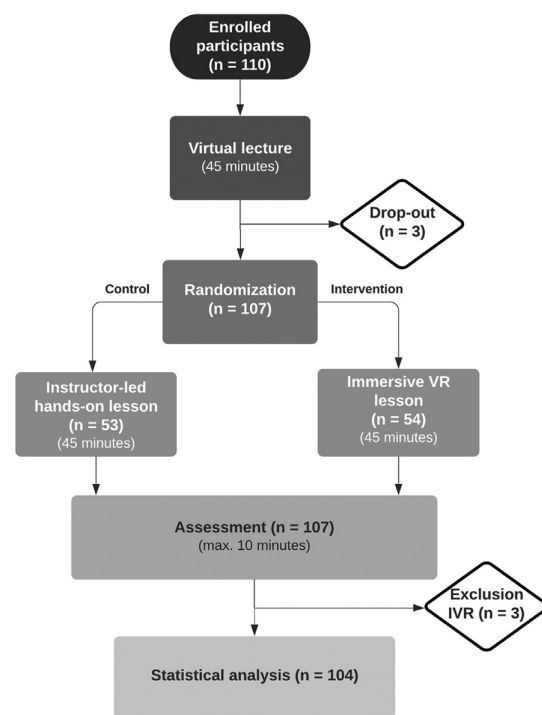


Fig. 1. Flowchart of the ultrasound course. The 110 enrolled participants attended a virtual lecture before randomization to instructor-led ($n = 53$) or immersive virtual reality (IVR) training ($n = 54$). Afterward, all participants' basic ultrasound skills were assessed by medical doctors with ultrasound experience, and hereof, 104 assessments were included in the statistical analysis.

Randomization

Before the lessons, participants were randomized to either the instructor-led group or the IVR group by unbiased university administrative staff.

Instructor-led lesson

Participants were subdivided into teams of a maximum of 14 people by university administrative staff, and each team participated in an instructor-led lesson, with one instructor teaching all teams separately throughout the day. The lesson included approximately 10 minutes of traditional class-based teaching, in which the instructor explained the following learning points: how to turn on an US machine, how to choose the correct probe and preset, how to apply gel to the probe, orientation on the screen when using the probe, correct positioning of the patient and the examiner and US knobology. For the remainder of the lesson, the participants practiced scanning a gelatin phantom, inspired by Richardson et al. (2015). The gelatin solution was mixed with cornstarch and food coloring to make it non-transparent and then poured into a box measuring $19 \times 7 \times 7$ cm. An oblong water balloon filled with water was embedded in the gelatin to simulate a deep-lying vein. The participants shared two Philips Epiq Elite US machines (Philips, Amsterdam, The Netherlands) for approximately 30 minutes before moving on to the assessment. The lesson was then repeated for the next team.

Immersive virtual reality lesson

Participants were subdivided into teams with a maximum of 14 people who participated in the IVR lesson. Technical support staff were present and were allowed to help participants with purely technical issues, should any arise.

Two participants shared one Oculus Quest 2 headset and the associated controllers (Facebook Technologies, LLC, Irvine, USA): While one participant used the headset and followed the instructions, the other watched the simulation in real time on a Samsung Galaxy Tab A tablet (Samsung, Seoul, South Korea). The software used for the simulation was made by VitaSim (VitaSim ApS, Odense, Denmark).

In the first 10 minutes of the lesson, the participant wearing the headset watched a 3-D recording of an avatar operating the virtual US machine. In the 3-D recording, the above-mentioned learning points for basic US were explained. The virtual avatar also demonstrated how to scan a virtual replica of the gelatin phantom used in the instructor-led lesson (as described above). Meanwhile, the participant was able to walk around in the virtual classroom. The 3-D recording, which included movement, speech and interactions, had been recorded beforehand by the same instructor who taught the

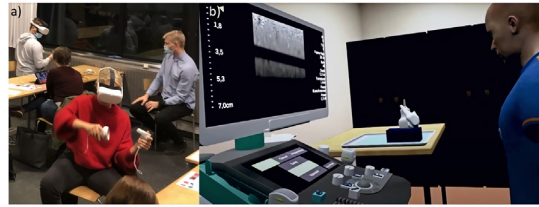


Fig. 2. (a) Setup for the immersive virtual reality lesson. (b) The virtual instructor is scanning a replica of the phantom from the instructor-led lesson.

instructor-led lesson. After the presentation, the participants had the opportunity to use the virtual US machine and scan the virtual phantom in the same way they would have done in a real-world setting. If necessary, the participant could also re-watch the 3-D recording. The participant had a maximum of 10 minutes to practice with the virtual US machine before switching with the next participant. Figure 2 illustrates the setup for the IVR lesson and simulation.

Pre-assessment preparation

At the end of the US lesson and preceding the assessment, the participants watched a short video on knobology for the US machine they would use during the assessment: a tablet with the Lumify application installed and connected to a L12-4 linear Lumify probe (Philips). The US scanner used during the assessment intentionally differed from the ones used during the lesson, as learning fundamental US knobology was more important than knowing the exact location of buttons on the specific US scanners. After watching the video, the participants completed a questionnaire on baseline information that included age, sex, previous US experience and self-assessed experience level in US. No mandatory US courses had formerly been a part of the curriculum of the university. Therefore, some of the students' former US experience might have originated from master's degree programs, spare-time jobs within public health care, voluntary informal US courses for medical students or clinical training at a hospital in the region, as written on the questionnaires.

Lastly, the students were randomly assigned to one of two separate assessment rooms.

Assessment

The US skill level of all participants was evaluated on a one-by-one basis by blinded assessors. The assessment was set up as a task-based exam. Participants were instructed (in writing) to scan a gelatin phantom, identify three olives embedded in it and measure their sizes in three dimensions. Each participant was equipped with a



Fig. 3. While a participant scans a gelatin phantom using the Lumify ultrasound scanner, a medical doctor with ultrasound experience assesses the participant's skills in basic ultrasound using the Objective Structured Assessment of Ultrasound Skills score.

Lumify US scanner and given 10 minutes to solve the task. The phantom for the assessment was also inspired by Richardson *et al.* (2015) but was made in a box measuring $10 \times 7 \times 7$ cm (Richardson *et al.* 2015). The correct measurements of the olives were known only by the assessors. The eight assessors were each allocated to their own stand equipped with a Lumify setup, a phantom and a paper describing the task. While the participants scanned the phantom, an assessor observed them and scored their performance on a standardized chart. Figure 3 illustrates the setup of the assessment.

To objectively assess the participants, the assessors used a quantitative score derived from the OSAUS score (Tolsgaard *et al.* 2013). Objectives in the OSAUS score concerning diagnostic and clinical decision making were omitted. The four OSAUS objectives chosen were familiarity with the US equipment, optimization of the US image, systematic approach, and interpretation of the US image. This enabled each participant to score from 4 to 20 points (maximum: 5 points per objective). On the OSAUS grading scale, a short description of specific grading criteria was elaborated; for example, for optimization of the US image “1: fails to optimize images, 3: competent image optimization but not done consistently, 5: consistent optimization of images.” The remaining criteria are elaborated on the original OSAUS table (Tolsgaard *et al.* 2013). Time to complete the task was not part of the objectives for assessing the students as showing and assessing their fundamental US skills were more important with respect to basic US skills than completion time.

The eight blinded assessors were physicians with experience in PoCUS. To ensure that all assessors were sufficiently prepared to rate the participants, they had practiced the assessment beforehand by watching three demonstration videos of students at different skill levels

(low, moderate, high) attempting to solve the task of scanning a phantom using the Lumify setup, similar to the real assessment. Their individual assessments were reviewed prior to study start and deemed acceptable regarding interrater variability, although no κ value was calculated. While watching the video, the assessors practiced using the above-mentioned OSAUS score. The assessors were instructed not to interfere during the assessment, but to provide verbal feedback to the participants afterward. The participants were not informed of the results of the assessment.

Outcomes

The primary outcome was the average number of points based on the OSAUS score. The secondary outcome was the average number of points for each objective in the OSAUS and the course expenses.

Sample size

To estimate the difference that could be detected between the two groups, a non-inferiority power calculation was made. A standard deviation of 20% corresponded to 0.8–4 points on the OSAUS score, depending on the number of points the participant obtained (4–20 points in total). Furthermore, the power was set to 80% with a 5% level of significance (equal to $p < 0.05$). Using the aforementioned parameters, a sample size of 110 participants, 55 in each group based on an allocation ratio of 1, was sufficient to detect a significant difference of approximately 10% between the groups.

Statistical methods

The results from the assessment were collected as absolute numbers. Means and standard deviations were calculated as descriptive parameters. The Shapiro–Wilk test was used to test for normal distribution. Boxplots were made to show means, medians, quartiles and spreads for the average total OSAUS scores, as illustrated in Figure 4. A two-sample *t*-test was used to test for significance in difference in OSAUS scores between the groups. Median, interquartile range (IQR) and Wilcoxon signed-rank test would be used to test for significance if any data were not normally distributed. A *p* value < 0.05 was considered to indicate statistical significance. Stata IC version 16.1 (StataCorp LLC, College Station, TX, USA) was used for statistical analysis.

RESULTS

Baseline characteristics of the participants are outlined in Table 1. Of 110 enrolled participants, 3 dropped out before randomization because of inability to attend the subsequent hands-on training. The remaining 107 participants were enrolled in the hands-on training of the

Table 1. Baseline information regarding the final-year medical students participating in the randomized controlled trial for point-of-care ultrasound teaching (n = 104)

Baseline information	Instructor-led group (n = 53)	Immersive virtual reality group (n = 51)
Age, years (mean, range)	27.0, 24–32	26.8, 23–31
Sex		
Male	26	20
Female	26	31
Not disclosed	1	0
Previous participation in an ultrasound course	13 (24.5%)	12 (23.5%)
For yes, number of hours (mean; range), $p = 0.37$	5.3; 1–20	8.4; 1–30
Self-assessed experience level		
None	33 (62%)	30 (60%)
Little	16 (30%)	19 (38%)
Some	4 (7.5%)	1 (2%)
Very	0 (0%)	0 (0%)

US course and completed both their respective lessons and the assessment. Three assessments were discarded: One participant enrolled in the course the day before data collection and was insufficiently prepared for the course. Assessments of two other participants were discarded because of a mismatch in study numbers at the assessment: the same study number had been written on two different assessment pages. Thus, a total of 104 participants were included in data analysis: 53 in the instructor-led group and 51 in the IVR group. The study flowchart is provided in Figure 1.

Efficacy of the learning methods

Data from the assessment were considered normally distributed as the Shapiro–Wilk tests were non-significant ($p > 0.05$), except for the IVR group's interpretation score ($p < 0.05$).

The participants in the instructor-led group had a mean OSAUS score of 11.0 points (95% confidence interval [CI]: 9.8–12.2, standard deviation [SD] = 4.3)

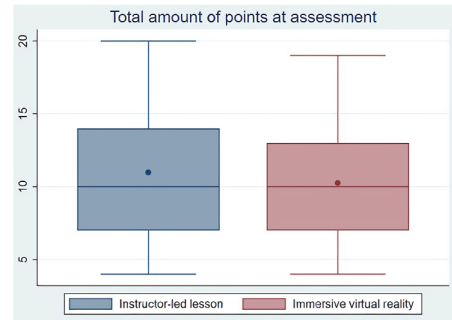


Fig. 4. Boxplots revealing medians and spreads in OSAUS score for the instructor-led lesson and the immersive virtual reality (IVR) lesson at the assessment. The dots represent means (11.0 points for instructor-led lesson and 10.3 points for IVR, respectively).

in total. The mean score for the IVR group was 10.3 points (95% CI: 9.0–11.5, SD = 4.4).

For interpretation of the findings, the instructor-led and IVR groups scored medians of 3 (IQR: 2–3) and 2 points (IQR: 2–3) ($p = 0.07$), respectively.

There were no significant differences between the groups for any of the subgroup objectives. The results of the two-sample t-tests and the Wilcoxon signed-rank test are shown in Table 2.

Course expenses

The running costs for the two lessons were estimated at 400 euros each. This included the phantoms and salary for the instructor of the instructor-led lessons and salaries for three technical support workers at the IVR lessons, respectively. All remaining materials were borrowed for both lessons.

The general startup costs for the IVR lesson were estimated at 6000 euros, which included 10 Oculus Quest 2 headsets with controllers, 10 Samsung Galaxy A6 tablets and a yearly software license to the VR simulation. The total expenses for an instructor-led lesson were estimated at approximately 100,000 euros and

Table 2. Differences in OSAUS score between the randomized groups

	1. Familiarity		2. Image optimization		3. Systematic approach		4. Interpretation		In total	
	IL (n = 53)	IVR (n = 51)	IL (n = 53)	IVR (n = 51)	IL (n = 53)	IVR (n = 51)	IL (n = 53)	IVR (n = 51)	IL (n = 53)	IVR (n = 51)
Mean/median (pts)	2.9	2.6	2.8	2.6	2.7	2.7	3	2	11.0	10.3
SE	0.2	0.2	0.2	0.2	0.2	0.2	—	—	0.6	0.6
SD	1.2	1.2	1.3	1.2	1.3	1.4	—	—	4.3	4.4
95% CI/IQR	(2.5–3.2)	(2.3–3.0)	(2.4–3.1)	(2.2–2.9)	(2.3–3.1)	(2.3–3.1)	(2–3)	(2–3)	(9.8–12.2)	(9.0–11.5)
Two-sample t-test or Wilcoxon signed-rank test	$t = 0.96, p = 0.34$		$t = 0.77, p = 0.45$		$t = 0.04, p = 0.96$		<u>$z = 1.83, p = 0.07$</u>		$t = 0.91, p = 0.36$	

CI = confidence interval; IL = instructor-led; IQR = interquartile range; IVR = immersive virtual reality; OSAUS = Objective Structured Assessment of Ultrasound Skills; SD = standard deviation; SE = standard error.

The OSAUS scores were obtained at an assessment directly after the instructor-led lesson and the self-directed immersive virtual reality lesson. Two-sample T-test is in bold, Wilcoxon signed-rank is underlined.

included two US machines of midrange quality, salary for the instructor and materials for the gelatin phantoms. Maintenance costs for the US machines are not included.

DISCUSSION

The results of this randomized controlled trial indicated that the learning efficacy of a self-directed IVR lesson on basic US was not inferior to that of an instructor-led lesson when assessed according to OSAUS test scores immediately following the training. To our knowledge, this study is the first randomized controlled trial to compare the efficacy of an instructor-led lesson with that of IVR in teaching basic US to medical students on a large scale.

A pilot study investigated the effect of IVR on teaching basic US skills to medical students compared with e-learning (Rosenfeldt Nielsen *et al.* 2021). The study was designed as a double-blind, parallel-group, block-randomized, controlled superiority trial wherein both groups went through non-supervised self-preparation via IVR or e-learning, followed by hands-on learning and an OSAUS test. The IVR group ($n = 11$) had significantly higher OSAUS scores than the e-learning group ($n = 9$) with a mean difference of 17 points, equaling 13.5%. The significantly higher OSAUS scores suggest that medical students could benefit more from IVR training than conventional e-learning when learning basic US without supervision.

An explorative pilot study investigated the effect of teaching US-guided peripheral venous cannulations through IVR (Andersen *et al.* 2021). The study was designed as a randomized controlled trial. After an e-learning session on basic US, medical students in the IVR group ($n = 10$) proceeded to additional IVR training on US-guided peripheral catheter placement, whereas students in the control group ($n = 9$) received no further training. The results of the subsequent evaluation revealed that in total, the proportion of successful cannulations was significantly higher in the IVR group than in the control group (73% vs. 22%). Similarly, the proportion of surface punctures correlated to successful cannulations was significantly higher in the IVR group than in the control group. Even though the participants in the IVR group received more training, the results of this pilot study suggest that IVR training could support an existing curriculum in teaching peripheral venous catheter placement to medical students. Future studies could investigate the effect of IVR training compared with phantom training to achieve additional comparable results.

The use of IVR to support the training of US skills and related anatomy for third-year medical students has been explored by Hu *et al.* (2020). The participants

attended a 6-h practical workshop including (i) an instructional lecture on basic concepts and handling of US equipment, (ii) capture of screenshots of relevant anatomical structures in either electronic atlases or IVR and (iii) practical US training. Their results indicated that participants in the intervention group ($n = 47$) had significantly higher scores on the US task test compared with the control group ($n = 54$). Furthermore, in 6 of 10 US tasks, the intervention group performed significantly better than the control group. A combined learning experience regarding US and anatomy in IVR could benefit students' learning and comprehension of the US image.

The results of the above-mentioned studies suggest that the effect of IVR training seems most optimal in synergy with other learning modalities. Although the results of this study indicate that the learning efficacies of the two teaching methods were equivalent, it is important to recognize that not all students benefit from a simulated learning experience. A systematic review found IVR to be useful in acquiring surgical skills regarding, for instance, laparoscopy, knot tying and open surgery (Tursø-Finnich *et al.* 2022). However, in some situations, IVR did not prove advantageous compared with less immersive setups or traditional instructions. In some studies, IVR was even found to be counterproductive. This was attributed mainly to reports of increased immersion, resulting in a high cognitive load during the learning sessions. These findings suggested that not all hands-on courses are necessarily interchangeable with IVR.

Future training versus course expenses

For this study, the course expenses for the instructor-led lesson were estimated to be equal to those of the IVR lesson. This further strengthens the notion of using IVR in basic US courses, particularly at universities with limited resources or personnel for US training. Other US simulator training has also been investigated and found to have positive results (Bentley *et al.* 2015; Le *et al.* 2019; Østergaard *et al.* 2019b). However, simulators for abdominal US lie in the price range of 20,000 to 90,000 euros, which could make simulator training more expensive than traditional US training (Østergaard *et al.* 2019a). In comparison, the HMDs used in this study are priced at 285 euros per set, and the applications for training seem inexhaustible. Meanwhile, implementing an IVR US course requires purchasing an appropriate number of headsets for the class.

As the IVR lesson was self-directed and instructions for the HMD and controllers were available, a physician or well-educated technical instructor is not necessarily required to supervise future lessons, which could reduce salary expenses.

Although general startup costs for an instructor-led lesson were estimated to be around 16 times higher than the expenses for an IVR lesson, it is important to note that all materials would be usable for several semesters of US training. Most medical schools or universities have already invested in US equipment for teaching or clinical purposes. This would significantly decrease the cost of an instructor-led lesson as well as the following US training sessions. The US machines could also be used in a ward, making it more relevant to compare the cost of software development and training equipment with the continued expense of hiring instructors. Alternatively, handheld US equipment, for instance Lumify or Butterfly iQ (Butterfly Network, Guilford, CT, USA), could be a cheaper and simpler solution for PoCUS training, ranging between 2300 and 8700 euros.

Furthermore, the IVR simulation used in this study is limited to basic US for the time being. Expanding IVR to include new modules, a full US course or training in other clinically relevant procedures without the need to buy new simulation equipment is yet to be investigated. New or alternative uses for the IVR headsets in the field of medical training could make the purchase a more favorable economic investment in the future, similar to purchasing US equipment.

Strengths and limitations

Having pre-defined learning goals for the lessons and using the same instructor to teach both lessons, despite one being virtual and not face-to-face, ensured uniformity in content and presentation. Furthermore, the use of the internationally validated OSAUS score to compare the groups' performances allowed for a quantitative outcome (Todsén et al. 2015). These factors increase the credibility and reproducibility of the current study. Additionally, the questionnaire on baseline information and clarification on the allocation of participants formerly entering any US courses supported homogeneity between the groups, thereby reducing the risk of confounding.

For our study, eight different doctors were recruited as assessors. This setup was chosen to avoid a bottleneck in the flow of students during the assessment. To minimize inter-observer variance, the assessors were trained in using the OSAUS by watching three videos, as mentioned under Methods.

This study has limitations. Both lessons were limited by and built up around the classic 45-minute lesson structure provided by the university. This gave each subdivided team 45 minutes to learn about basic hands-on US before the assessment. As Stepan et al. (2017) mentioned in their article on the use of IVR in teaching neuroanatomy, a portion of the study time was spent getting acquainted with the HMD and controllers before

learning about the course material. Spending more time on the simulation or equipping every participant with an HMD, yet keeping the lesson within 45 min, could have increased the efficacy of the IVR lesson. Another limitation was the phantom of choice: three of the gelatin phantoms failed to maintain their structure throughout all assessments, which might have led to misinterpretation of the findings. The assessors were made aware that visualization of the olives embedded in the phantom could be more challenging for the last rounds of students and their assessment should not be affected by it.

In this study, the modified OSAUS score worked as a surrogate marker for learning efficacy. The included objectives helped the assessors focus on basic US skills at the assessment but limited the clinical relevance.

By quantifying the participants' performance, specific details on how the participants' US technique could improve for future training were not elaborated.

The results of one of the objectives, interpretation of the findings, entail risk of type 2 error. As the p value borders on the threshold for statistical significance ($p = 0.07$), the null hypothesis might have been falsely rejected. A larger study population could have minimized this risk. Furthermore, because of the three dropouts and three discarded assessments, the sample size of 110 participants was not reached as warranted, based on the power calculation.

Future research

Further investigations into the use of IVR in US education and training are needed. As the objective of this study focused only on basic US, the use of IVR in teaching PoCUS in general remains unexplored. Continuous development of IVR as an educational tool for medical procedures could make it a favorable supplement or even an alternative to current training in the future.

CONCLUSIONS

The learning efficacy of an instructor-led lesson on basic US did not differ significantly from that of a self-directed lesson in IVR, as assessed using the OSAUS. The results of this study suggest that IVR could be beneficial in future basic US courses, but this warrants further research to clarify whether IVR is suitable for PoCUS courses.

Conflict of interest disclosure—R.O.J. is CEO and co-founder of VitaSim, which provided the virtual reality software used in this study. R. O.J. did not assess any of the participants' ultrasound skills nor teach the ultrasound course. The remaining authors do not have any conflicts of interest to disclose.

REFERENCES

- Amini R, Stolz LA, Gross A, O'Brien K, Panchal AR, Reilly K, Chan L, Drummond BS, Sanders A, Adhikari S. Theme-based teaching of point-of-care ultrasound in undergraduate medical education. *Intern Emerg Med* 2015;10:613–618.
- Andersen NL, Jensen RO, Posth S, Laursen CB, Jørgensen R, Graumann O. Teaching ultrasound-guided peripheral venous catheter placement through immersive virtual reality: An explorative pilot study. *Medicine (Baltimore)* 2021;100:e26394.
- Bentley S, Mudan G, Strother C, Wong N. Are live ultrasound models replaceable? Traditional versus simulated education module for FAST exam. *West J Emerg Med* 2015;16:818–822.
- Carstensen SMD, Velander MJ, Konge L, Østergaard M, Pfeiffer Jensen M, Just SA, Terslev L. Training and assessment of musculoskeletal ultrasound and injection skills—A systematic review. *Rheumatology* 2022;keac119.
- Hooper J, Tsiridis E, Feng JE, Schwarzkopf R, Waren D, Long WJ, Poultsides L, Macaulay W. NYU Virtual Reality Consortium. Virtual reality simulation facilitates resident training in total hip arthroplasty: A randomized controlled trial. *J Arthroplasty* 2019;34:2278–2283.
- Hu KC, Salcedo D, Kang YN, Lin CW, Hsu CW, Cheng CY, Suk FM, Huang WC. Impact of virtual reality anatomy training on ultrasound competency development: A randomized controlled trial. *PLoS One* 2020;15:e0242731.
- Jensen L, Konradsen F. A review of the use of virtual reality head-mounted displays in education and training. *Educ Inf Technol* 2018;23:1515–1529.
- Kahr Rasmussen N, Andersen TT, Carlsen J, Østergaard ML, Konge L, Albrecht-Beste E, Nielsen MB. Simulation-based training of ultrasound-guided procedures in radiology—A systematic review. *Ultraschall Med* 2019;40:584–602.
- Latham K, Kot P, Wariach A, Al-Jumeily D, Puthuran M, Chandran A. A review on the development of a virtual reality learning environment for medical simulation and training. *The Fourth International Conference on Applications and Systems of Visual Paradigms (VISUAL 2019)*. Rome, Italy. 30 June–4 July.
- Le CK, Lewis J, Steinmetz P, Dyachenko A, Oleskevich S. The use of ultrasound simulators to strengthen scanning skills in medical students: A randomized controlled trial. *J Ultrasound Med* 2019;38:1249–1257.
- Moore CL, Copel JA. Point-of-care ultrasonography. *N Engl J Med* 2011;364:749–757.
- Nayahangan LJ, Konge L, Russell L, Andersen S. Training and education of healthcare workers during viral epidemics: A systematic review. *BMJ Open* 2021;11:e044111.
- Nielsen MB, Cantisani V, Sidhu PS, Badea R, Batko T, Carlsen J, Claudon M, Ewertsen C, Garre C, Genov J, Gilja OH, Havre R, Kosiak M, Kosiak W, Pilcher J, Prosch H, Radzina M, Rafailidis V, Rykkje A, Serra A, Sotiriadis A, Østergaard M, Dietrich CF. The use of handheld ultrasound devices—An EFSUMB Position Paper. *Ultraschall Med Eur* 2019;40:30–39.
- Østergaard ML, Konge L, Kahr N, Albrecht-Beste E, Nielsen MB, Nielsen KR. Four virtual-reality simulators for diagnostic abdominal ultrasound training in radiology. *Diagnostics (Basel)* 2019;9:50. a.
- Østergaard ML, Rue Nielsen K, Albrecht-Beste E, Kjær Ersbøll A, Konge L, Bachmann Nielsen M. Simulator training improves ultrasound scanning performance on patients: A randomized controlled trial. *Eur. Radiol.* 2019;29:3210–3218 b.
- Pietersen PI, Madsen KR, Graumann O, Konge L, Nielsen BU, Laursen CB. Lung ultrasound training: A systematic review of published literature in clinical lung ultrasound training. *Crit Ultrasound J* 2018;10:23.
- Pietersen PI, Mikkelsen S, Lassen AT, Helmerik S, Jørgensen G, Nadim G, Christensen HM, Wittrock D, Laursen CB. Quality of focused thoracic ultrasound performed by emergency medical technicians and paramedics in a prehospital setting: A feasibility study. *Scand J Trauma Resusc Emerg Med* 2021;29:40.
- Pottle J. Virtual reality and the transformation of medical education. *Future Healthc J* 2019;6:181–185.
- Richardson C, Bernard S, Dinh VA. A cost-effective, gelatin-based phantom model for learning ultrasound-guided fine-needle aspiration procedures of the head and neck. *J Ultrasound Med* 2015;34:1479–1484.
- Rosenfeldt Nielsen M, Kristensen EQ, Jensen RO, Møllerup AM, Pfeiffer T, Graumann O. Clinical ultrasound education for medical students: Virtual reality versus e-learning, a randomized controlled pilot trial. *Ultrasound Q* 2021;37:292–296.
- Saldana D, Neureither M, Schmiesing A, Jahng E, Kysh L, Roll SC, Liew SL. Applications of head-mounted displays for virtual reality in adult physical rehabilitation: A scoping review. *Am J Occup Ther* 2020;74:7405205060p1-7405205060p15.
- Skalski JH, Elrashidi M, Reed DA, McDonald FS, Bhagra A. Using standardized patients to teach point-of-care ultrasound-guided physical examination skills to internal medicine residents. *J Grad Med Educ* 2015;7:95–97.
- Srivastava AP, Babu V, Shetye SK. Exploring students' learning efficacy: Predicting role of teachers' extra role behaviour and art-based innovation pedagogies. *Eur J Train Dev* 2019;13:842–859.
- Stepan K, Zeiger J, Hanchuk S, Del Signore A, Shrivastava R, Govindaraj S, Iloretta A. Immersive virtual reality as a teaching tool for neuroanatomy. *Int Forum Allergy Rhinol* 2017;7:1006–1013.
- Tarique U, Tang B, Singh M, Kulasegaram KM, Ailon J. Ultrasound curricula in undergraduate medical education: A scoping review. *J Ultrasound Med* 2018;37:69–82.
- Todsén T, Tolsgaard MG, Olsen BH, Henriksen BM, Hillingsø JG, Konge L, Jensen ML, Ringsted C. Reliable and valid assessment of point-of-care ultrasonography. *Ann. Surg* 2015;261:309–315.
- Tolsgaard MG, Todsén T, Sørensen JL, Ringsted C, Lorentzen T, Ottesen B, Tabor A. International multispecialty consensus on how to evaluate ultrasound competence: A Delphi consensus survey. *PLoS One* 2013;8:e57687.
- Tursø-Finnich T, Jensen RO, Jensen LX, Konge L, Thinggaard E. Virtual reality head-mounted displays in medical education—A systematic review. *Simul Healthc* 2022; doi: 10.1097/sih.0000000000000636 Published online February 8.
- Weiner E, Gordon J, Rudy S, McNew R. Expanding virtual reality to teach ultrasound skills to nurse practitioner students. *Stud Health Technol Inform* 2019;264:893–897.