Received: 7 June 2018 | Revised: 3 September 2018 | Accepted: 3 September 2018

DOI: 10.1111/aas.13276

ORIGINAL ARTICLE

acta Anaesthesiologica Scandinavica

Remote real-time supervision via tele-ultrasound in focused cardiac ultrasound: A single-blinded cluster randomized controlled trial

Stig H. Jensen[1,2] | Jesper Weile[1,2] | Rasmus Aagaard[1,3] | Kåre M. Hansen[4] | Troels B. Jensen[5] | Morten C. Petersen[2] | Jacob J. Jensen[2] | Poul Petersen[2] | Hans Kirkegaard[1]

[1]Research Center for Emergency Medicine, Department of Clinical Medicine, Aarhus University, Aarhus N, Denmark

[2]Emergency Department Regional, Hospital West Jutland, Herning, Denmark

[3]Department of Internal Medicine, Regional Hospital Randers, Randers, Denmark

[4]Department of Anesthesiology and Intensive Care, Aalborg University Hospital, Aalborg, Denmark

[5]Department of Anesthesiology, Regional Hospital West Jutland, Herning, Denmark

Correspondence
Stig H. Jensen, Research Center for Emergency Medicine, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, J103 8200, Aarhus N, Denmark.
Email: stig.holm.jensen@post.au.dk

Funding information
The Holger & Ruth Hesse's Mindefond and the Knud og Edith Eriksens Mindefond supported this work. The funders had no role in the design or conduct of the study; collection, management, analyses, or interpretation of the data; preparation, review, or approval of the manuscript; nor the decision to submit the manuscript for publication.

Background: Supervision via tele-ultrasound presents a remedy for lacking on-site supervision in focused cardiac ultrasound, but knowledge of its impact is largely absent. We aimed to investigate tele-supervised physicians' cine-loop quality compared to that of non-supervised physicians and compared to that of experts.

Methods: We conducted a single-blinded cluster randomized controlled trial in an emergency department in western Denmark. Physicians with basic ultrasound competence scanned admitted patients twice. The first scan was non-supervised, and the second was non-supervised (control) or tele-supervised (intervention). Finally, experts in focused cardiac ultrasound scanned the same patient. Two blinded observers graded cine-loops recorded from all scans on a 1–5 scale. The outcome was the mean summarized scan gradings compared with a linear mixed-effects model.

Results: In each group, 10 physicians scanned 44 patients. From the mean summarized gradings, on a scale from 4 to 20, the second non-supervised scan grading was 10.9 (95% CI 10.2-11.7), whereas the tele-supervised grading was 12.6 (95% CI: 11.8-13.3). From the first to the second scan, tele-supervised physicians moved 9% (1.09; 95% CI: 1.00-1.19; $P = 0.041$) closer to the experts' quality than the non-supervised physicians.

Conclusion: Tele-supervised physicians performed scans of better quality than non-supervised physicians. The present study supports the use of tele-supervision for physicians with basic focused ultrasound competence in a setting where on-site supervision is unavailable.

1 | INTRODUCTION

Focused cardiac ultrasound (FOCUS) is a recommended approach in the evaluation of patients with dyspnea, chest pain, shock, cardiac trauma, and cardiac arrest.[1] Appropriate training and quality assurance are required for point-of-care US in general because its usage and quality are highly operator dependent.[2] Despite clinical recommendations and the advancement of small, mobile, and cheaper US systems, the usage of FOCUS remains unequally distributed between academic and non-academic facilities.[3] In Denmark, the unequal usage of FOCUS in emergency departments (EDs) also occurs among specialties, where physicians from other departments mainly perform FOCUS in the ED.[4] The major barriers curbing the dissemination are the lack of supervision, training, and competence retention.[3,4]

Supervision via audiovisual communication, known as tele-ultrasound (tele-US), serves as a technical remedy with the ability to provide training and supervised learning in remote locations where no supervisors are otherwise available. However, knowledge of the quality impact of supervision via tele-US is lacking.

Therefore, in this first randomized in-hospital tele-US project of FOCUS, we aimed to investigate the cine-loop quality when comparing tele-supervised physicians with non-supervised physicians and with experts.

Editorial Comment

Rapid and focused ultrasound in the Emergency department is believed to be diagnostically useful in patients with cardiopulmonary dysfunction or injury, but some training is needed in order to establish diagnoses with confidence. This trial demonstrated that skilled supervision of novice ultrasonographers, even by distance, led to improvement in the quality of their ultrasound diagnostic imaging.

2 | METHODS

2.1 | Design

This single-blinded cluster randomized controlled trial took place in the emergency department of the Regional Hospital West Jutland in the Central Denmark Region (CDR), an ED receiving approximately 40 000 patients every year.

The study was exempt from the informed consent requirements by the Regional Ethics Committee, CDR (inquiry 153/2016). The Danish Data Protection Agency approved data handling (case no. 1-16-02-175-16). The primary investigator contacted admitted patients, informed them about the study, and asked for their consent. The study anonymized all data. Reporting the study takes place in accordance with the "CONSORT 2010 statement: extension to cluster randomized trials".[5]

2.2 | Participants

Inclusion criteria for physicians were: (a) first postgraduate clinical employment, (b) employment in the ED of the Regional Hospital West Jutland, CDR, (c) certification from a 2-day US course (undertaken independently from this project), (d) a minimum experience of 10 FOCUS examinations, and (e) a minimum total experience of 60 point-of-care US examinations. The 2-day US course includes e-learning and hands-on-training in basic US physics, focused assessment with sonography for trauma, peripheral US-guided vascular access, FOCUS, and lung and abdominal point-of-care US.

Patients included in the study were a convenience sample (aged above 17 years) recruited in the ED when physicians, supervisors, and experts were available. Exclusion criteria were: (a) Glasgow Coma Score below 15, (b) isolation, (c) unable to give oral consent, (d) unstable vital signs, (e) life-threatening injuries, and (f) open wounds at the topical landmarks of FOCUS.

We aimed to include 10 physicians in each group and strived to have each physician scan five patients.

2.3 | Ultrasound examinations

Physicians scanned all patients twice: first, non-supervised, then either non-supervised (control) or tele-supervised (intervention). As a gold standard, an expert with US experience above 500 cardiac examinations scanned all patients once. Thereby all patients underwent three scans.

All examinations included four cardiac views: subcostal four-chamber (S4C), parasternal long (PLAX)- and short-axis (PSAX), and apical four-chamber (A4C) views. Both physicians and experts received instructions to store at least one cine-loop from each view. After each examination, physicians and experts individually reviewed stored cine-loops and selected the best from each view to keep for study analysis.

2.4 | Supervision via tele-US

Designing the tele-US setup highly prioritized commercially available and low-cost equipment. The final setup is schematized in Figure 1. It consisted of a US system (Vivid S6; GE Healthcare, Chicago, IL, USA), a video grabber (DVI2USB 3.0; Epiphan Video, Ottawa, ON, Canada), two web cameras (Logitech HD Pro C920 and Logitech HD Webcam C525; Logitech, Romanel-sur-Morges, Switzerland), a headset (Major II Bluetooth; Marshall, Bletchley, Milton Keynes, UK), and two laptop computers (on-site and remote). We integrated the video inputs from the US system and the two web cameras into webcam software (Manycam Studio Version; Visicom Media Inc, Brossard, QC, Canada). Online transmission took place via VoIP software (Skype; Skype Technologies, Luxembourg, Luxembourg).

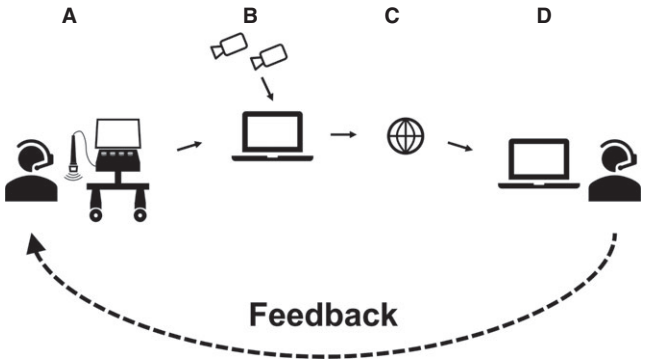

FIGURE 1 Tele-ultrasound setup: A, physician and ultrasound system; B, web cameras and on-site laptop; C, internet connection; D, remote laptop and supervisor

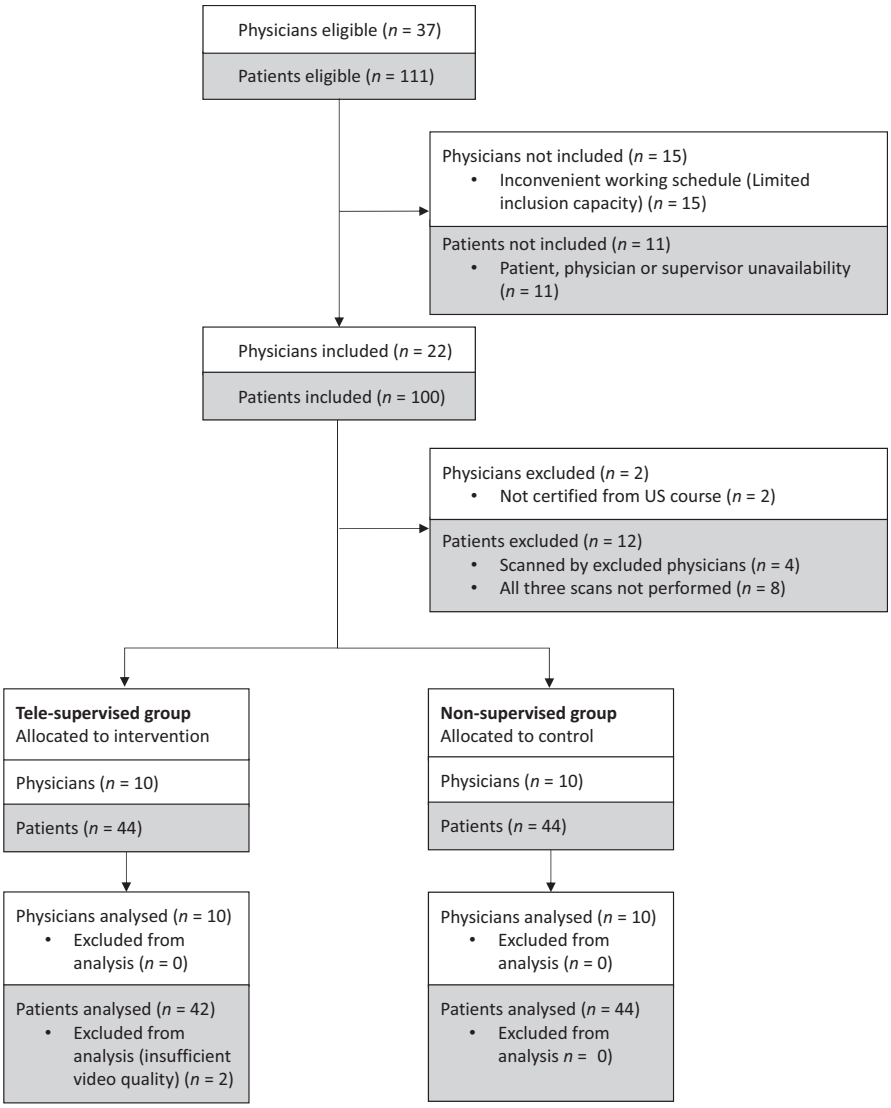

FIGURE 2 Participant flow. Screened, excluded, included, and analyzed patients in the study

Supervisors instructed the physicians to move the transducer slowly and to begin with the S4C view, followed by the PLAX, PSAX, and A4C views. Prior to study inclusion, both physicians and supervisors received instructions in a shared terminology for transducer movements, namely, slide, rotate, tilt, and rock.

2.5 | Outcomes

The outcome was the quality of all selected cine-loops graded on a scale from 1 to 5 (1 = no meaningful images; 2 = poor, not sufficient for interpretation; 3 = good, acceptable for interpretation; 4 = excellent, minor suggestions for improvement; 5 = outstanding, no suggestions for improvement).[6,7] In cases of windows left out, researchers reported cine-loops as missing. Two observers with US experience above 500 cardiac examinations graded all cine-loops that the physicians selected.

The study recorded the following baseline characteristics: age, gender, body mass index, previous cardiac- and lung-specific medical history, World Health Organization (WHO) performance score status, and present cardiac symptoms (pitting edema, chest pain, palpitations). Physicians also reported the number of FOCUS and point-of-care US examinations they had performed before inclusion.

Conducting a training session in scale usage assured interobserver agreement. Scale training comprised 51 pilot cine-loops distributed evenly among all four cardiac views. The two observers independently graded the 51 pilot cine-loops. After primary grading, the observers and two authors met online and reviewed all pilot cine-loops. Those involved discussed any disagreements until they obtained agreement. Thereafter, the observers graded all the pilot cine-loops again. The analysis does not include pilot cine-loops.

2.6 | Randomization and allocation

We deterministically allocated the physicians, assigning the first 10 to the intervention group and the next 10 to the control group. When choosing among physicians, besides inclusion criteria, we only considered their working schedules. We expected a random distribution of physician and patient characteristics other than period of

TABLE 1 Baseline characteristics

	Non-supervised	Tele-supervised
Patients		
n (%)	44 (51.2)	42 (48.8)
Sex (male), n (%)	25 (56.8)	26 (61.9)
Age, mean (SD), y	58 (20)	55 (18)
Body mass index, mean (SD), kg/m^2	27.5 (4.4)	27.9 (6.1)
WHO performance status, n (%)		
0	35 (81.4)	32 (76.2)
1	6 (14.0)	6 (14.3)
2	1 (2.3)	4 (9.5)
4	1 (2.3)	0 (0.0)
Pitting edema (yes), n (%)	12 (27.3)	8 (19.0)
Chest pain, n (%)		
Present	8 (18.6)	2 (4.8)
During physical effort	3 (7.0)	1 (2.4)
No	32 (74.4)	38 (90.5)
Palpitations (yes), n (%)	7 (15.9)	2 (4.8)
COPD (yes), n (%)	4 (9.1)	2 (4.8)
Physicians		
n (%)	10 (50)	10 (50)
Total POCUS experience, median (IQR), scans	65 (61; 75)	70.5 (63; 80)
Total FOCUS experience, median (IQR), scans	14.5 (14; 17)	14.5 (12; 19)
Supervisors		
Supervisor 1, n (%)		18 (42.9)
Supervisor 2, n (%)		24 (57.1)

FOCUS, focused cardiac ultrasound; POCUS, point-of-care ultrasound.

inclusion and working schedules. We allocated supervisors and experts according to their working schedules.

2.7 | Blinding

Blinding took place in the following steps: (a) experts were blind to the preceding two scans; (b) observers were blind to the group, sonographer, and type of cardiac US view.

2.8 | Statistical analyses

The study presented baseline characteristics as means and standard deviations (SD) for parametric continuous variables, medians and interquartile ranges (IQR) for non-parametric continuous outcomes, and numbers and percentages for categorical variables. Comparisons between groups were computed by using a one-way analysis of variance model, Kruskal-Wallis one-way analysis of variance, or Fisher's exact test, respectively.

Data for cine-loop quality comparison were modelled with a linear mixed-effects model. The means of observers' gradings were summed according to each scan, thereby including the mean grading of the four cardiac views (summed scan gradings). By using expert scans as references, the summed scan grading differences were compared groupwise with calculation of both the absolute difference and ratio of differences. The coefficient of intra-cluster correlation (ICC) at both the physician level and the patient and physician level combined are reported.

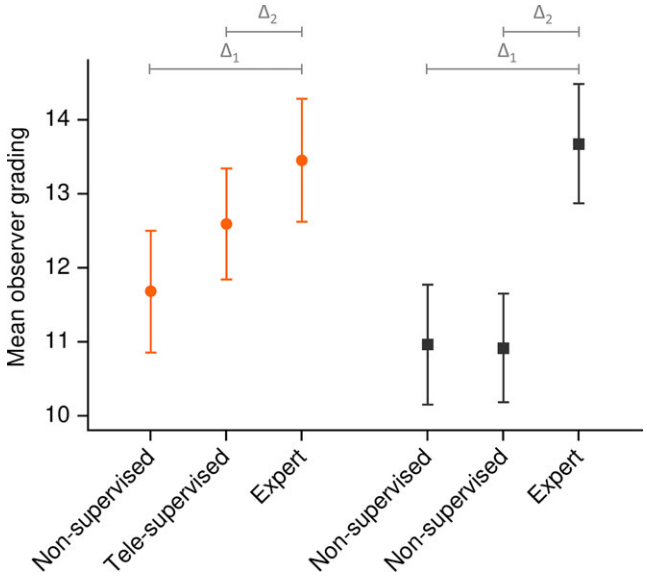

FIGURE 3 Summed scan gradings. Observers' mean gradings summed according to each scan, thereby including the mean grading of the four cardiac views, categorized by group. Values are results of the linear mixed-effects model presented as mean and 95% confidence intervals. Circles represent intervention group (44 patients scanned three times); squares represent control group [Colour figure can be viewed at wileyonlinelibrary.com]

The percentage of agreement and Cohen's kappa (including weighted kappa, $1-|i-j|/(k-1)$) of the two blinded observers' gradings were calculated for both the training and the final gradings. For interpretation of relative strength of agreement, the categorization by Landis and Koch was applied.[8]

All results were computed using STATA 14 (Statacorp, Houston, TX, USA).

3 | RESULTS

3.1 | Patients

Data collection took place from October 2016 to July 2017. At the primary investigator's discretion, we screened 111 patients yet excluded 11 due to patient, physician, expert, or supervisor unavailability. Subsequently, we excluded two of the physicians originally included (also including four of their patients); as it turned out, they were not US course certified. We excluded eight patients since they lacked one or more of the three scans. In total, we allocated 88 patients, 44 into each group. In the intervention group, we excluded two patients during analysis due to a cine-loop recording error that

TABLE 2 Summed scan grading differences

Total score	Tele-supervised (Mean, 95% CI)	Non-supervised (Mean, 95% CI)	Absolute difference (95% CI)	*P*-value	Ratio of differences (95% CI)	*P*-value
Δ_1 (expert—non-supervised)	1.77 (1.06; 2.49)	2.72 (2.02; 3.42)	−0.94 (−1.94; 0.06)	0.065	0.92 (0.85; 1.00)	0.062
Δ_2 (expert—tele/non-supervised)	0.86 (0.23; 1.48)	2.76 (2.15; 3.37)	−1.90 (−2.78; −1.03)	<0.001	0.84 (0.78; 0.91)	<0.001
$\Delta_1 - \Delta_2$	0.92 (0.29; 1.54)	−0.05 (−0.66; 0.57)	0.96 (0.08; 1.84)	0.032	1.09 (1.00; 1.19)	0.041

Control and intervention comparison by subgroup subtractions (as illustrated with delta-signs in Figure 3); presented as absolute difference and ratio of differences.

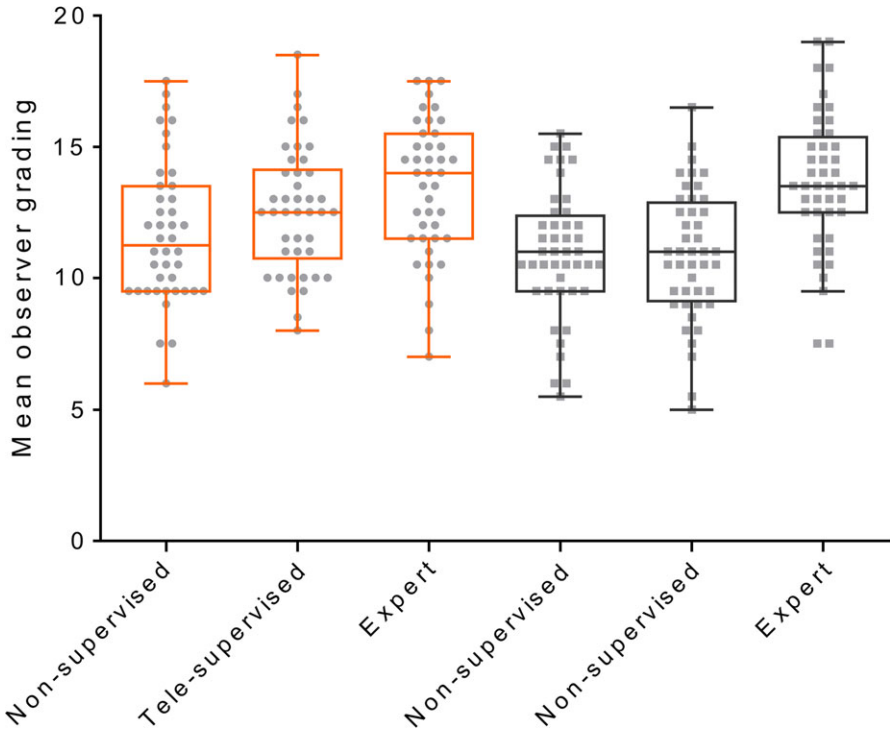

FIGURE 4 Original data. Observers' mean gradings summed according to each scan, thereby including the mean grading of the four cardiac views, categorized by group. Original data are presented as dots overlaid by box plots (median, first and third percentiles, and whiskers, 3⁄2 of first and third percentile). Circles represent intervention group (44 patients scanned three times); squares represent control group [Colour figure can be viewed at wile yonlinelibrary.com]

degraded video quality and made observer grading impossible (Figure 2). Baseline characteristics were comparable between groups (Table 1).

3.2 | Interobserver reliability

Gradings of all cine-loops showed a 50% agreement (κ 0.32); in comparison, the weighted agreement was 87% (κ 0.53), interpreted as moderate in concordance with Landis & Koch.[8]

3.3 | Outcome

The mean of non-supervised summed scan gradings was 11.7 (95% CI: 10.9-12.5) in the intervention group and 11.0 (95% CI: 10.2; 11.8) in the control group, when applying the linear mixed-effects model. During tele-supervision, the intervention group had a mean grading of 12.6 (95% CI: 11.8-13.3). When repeating the non-supervised scan, the control group had a mean grading of 10.9 (95% CI 10.2-11.7). Expert scans were 13.5 (95% CI: 12.6-14.3) and 13.7 (95% CI: 12.9-14.5) in the intervention and control group, respectively (Figure 3). The coefficient of intra-cluster correlation (ICC) was 0.03 at the physician level and 0.63 for the patient and physician level combined.

When inspecting expert vs physician scoring differences, the tele-supervised physicians came closer to the expert level than non-supervised physicians. This difference was statistically significant with an absolute difference of 0.96 (95% CI: 0.08-1.84; Table 2).

For complete overview of original data according to each scan categorized by group, a Tukey plot is shown in Figure 4.

4 | DISCUSSION

We conducted the first randomized in-hospital tele-US project of FOCUS. Tele-supervised physicians performed cine-loops with a significantly higher quality than comparable non-supervised physicians did. Moreover, the tele-supervised physicians produced cine-loops of a quality closer to that of experts than the non-supervised physicians.

The tele-supervised physicians' mean of summed scan gradings improved, implying that they did not have a fully saturated learning potential before supervision via tele-US. We believe that the US competency of the physicians included in the present study is widely generalizable to all physicians certified from comparable basic US courses.

Previous studies investigating feasibility of supervision via tele-US have shown positive results regarding technical and user satisfaction aspects.[9-11] Taken together with our results, this indicates that tele-US can serve as a favorable means to ensure competency transfer from controlled course settings into everyday clinical practice.

Few other studies have investigated the diagnostic effect of remote real-time supervision. Kim et al[12] found an increase in diagnostic confidence of tele-supervised examinations when compared to non-supervised pediatric acute appendicitis scans that EM residents with 1-2 years' experience with US performed. Grant et al conducted another prospective tele-US study of the remote diagnosis of congenital heart disease. They found a significant improvement in the diagnostic accuracy of pediatricians' tele-supervised echocardiograms compared with the same pediatricians' non-supervised scans.[13] Together, Kim et al, Grant et al, and our study indicate that tele-US is beneficially applicable to both focused and comprehensive US, and that a wide spectrum of users may benefit from tele-supervision.

The present study showed that the scan quality of the first and second non-supervised scans in the control group did not differ significantly in the summed scan gradings nor when computing differences to expert scans. This indicates non-supervised physicians repeating their own scans do not improve the quality of the cine-loops produced. This occurs neither instantly via repetition nor subsequently when scanning the next patient. Generally, in both groups, non-supervised physicians were unable to present cine-loops with a summed scan grading above 12. This underlines the overall importance of supervised learning within US and exhibits a gap in point-of-care US research. Future investigations should aim to clarify the appropriate amount of education needed to obtain basic FOCUS competencies in the clinic.

Our study has some limitations. First, the small number of included physicians potentially lowers the generalizability of our findings. Second, the study design introduced data clustering; however, the statistical analysis considered this.

Third, this was a short-term temporary implementation of tele-US. It is not clear whether physicians retain the skills in the long term. Fourth, we used a non-validated global assessment scale instead of validated checklists or a validated generic assessment scale.[14-16] Global rating scales generally capture more nuances and complementary perspectives than checklists if raters receive sufficient training.[17] Therefore, we included few raters with considerable insight into the assessed objective and facilitated assessment training and shared discussion. A comprehensive generic assessment scale, the Objective Structured Assessment of Ultrasound Skills scale, integrates all learning objectives in US. However, we considered this too excessive for our acquisition purpose only.[14]

Finally, we left out a closely related study arm: on-site–supervised scans. The premise of the study was that on-site supervision is lacking in nonteaching hospitals, and thus, we considered non-supervision the most relevant control.

5 | CONCLUSION

Tele-supervised physicians performed scans of better quality than non-supervised physicians. The present study supports the use of tele-supervision for physicians with basic focused US competence in a setting where on-site supervision is unavailable.

ACKNOWLEDGEMENTS

The authors would like to thank the Emergency Department, Regional Hospital West Jutland, CDR and all participating physicians and patients. The authors especially want to thank Stine Primdahl, Professor Bo Løfgren, and Associate Professor Lars Bolvig for participation in project discussions, and statistician Aparna Udupi for statistical consultancy.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest. No company sponsored this research, and the authors have full control of all data.

ORCID

Stig H. Jensen http://orcid.org/0000-0003-3815-2141

REFERENCES

1. Labovitz AJ, Noble VE, Bierig M, et al. Focused cardiac ultrasound in the emergent setting: a consensus statement of the American Society of Echocardiography and American College of Emergency Physicians. *J Am Soc Echocardiogr*. 2010;23:1225-1230.
2. Moore CL, Copel JA. Point-of-care ultrasonography. *N Engl J Med*. 2011;364:749-757.
3. Micks T, Sue K, Rogers P. Barriers to point-of-care ultrasound use in rural emergency departments. *CJEM*. 2016;18:475-479.
4. Nielsen K, Lauridsen JR, Laursen CB, Brabrand M. Physicians using ultrasound in Danish emergency departments are mostly summoned specialists. *Scand J Trauma Resusc Emerg Med*. 2015;23:51.
5. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345:e5661.
6. Henwood PC, Mackenzie DC, Liteplo AS, et al. Point-of-care ultrasound use, accuracy, and impact on clinical decision making in Rwanda hospitals. *J Ultrasound Med*. 2017;36:1189-1194.
7. Jain D, Grejs AM, Bhavsar R, Bang U, Sloth E, Juhl-Olsen P. Focused cardiac ultrasound is feasible in parturients; a prospective observational study. *Acta Anaesthesiol Scand*. 2017;61:1105-1113.
8. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363-374.
9. Dyer D, Cusden J, Turner C, et al. The clinical and technical evaluation of a remote telementored telesonography system during the

acute resuscitation and transfer of the injured patient. *J Trauma*. 2008;65:1209-1216.
10. Liteplo AS, Noble VE, Attwood B. Real-time video streaming of sonographic clips using domestic internet networks and free video-conferencing software. *J Ultrasound Med*. 2011;30:1459-1466.
11. Zennaro F, Neri E, Nappi F, et al. Real-time tele-mentored low cost "point-of-care US" in the hands of paediatricians in the emergency department: diagnostic accuracy compared to expert radiologists. *PLoS One*. 2016;11:e0164539.
12. Kim C, Kang BS, Choi HJ, Lim TH, Oh J, Chee Y. Clinical application of real-time tele-ultrasonography in diagnosing pediatric acute appendicitis in the ED. *Am J Emerg Med*. 2015;33:1354-1359.
13. Grant B, Morgan GJ, McCrossan BA, et al. Remote diagnosis of congenital heart disease: the impact of telemedicine. *Arch Dis Child*. 2010;95:276-280.
14. Tolsgaard MG, Todsen T, Sorensen JL, et al. International multispecialty consensus on how to evaluate ultrasound competence: a Delphi consensus survey. *PLoS One*. 2013;8:e57687.
15. Nielsen DG, Gotzsche O, Eika B. Objective structured assessment of technical competence in transthoracic echocardiography: a validity study in a standardised setting. *BMC Med Educ*. 2013;13:47.
16. Gaudet J, Waechter J, McLaughlin K, et al. Focused critical care echocardiography: development and evaluation of an image acquisition assessment tool. *Crit Care Med*. 2016;44:e329-e335.
17. Ilgen JS, Ma I, Hatala R, Cook DA. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Med Educ*. 2015;49:161-173.

How to cite this article: Jensen SH, Weile J, Aagaard R, et al. Remote real-time supervision via tele-ultrasound in focused cardiac ultrasound: A single-blinded cluster randomized controlled trial. *Acta Anaesthesiol Scand*. 2019;63:403–409. https://doi.org/10.1111/aas.13276