Feasibility and Clinical Implications of Remotely Guided Ultrasound Examinations

Quinn Dufurrena; Kazi Imran Ullah; Erin Taub; Connor Leszczuk; Sahar Ahmad

BACKGROUND: Remotely guided ultrasound (US) examinations carried out by nonmedical personnel (novices) have been shown to produce clinically useful examinations, at least in small pilot studies. Comparison of the quality of such exams to those carried out by trained medical professionals is lacking in the literature. This study compared the objective quality and clinical utility of cardiac and pulmonary US examinations carried out by novices and trained physicians.

METHODS: Cardiac and pulmonary US examinations were carried out by novices under remote guidance by an US expert and independently by US trained physicians. Exams were blindly evaluated by US experts for both a task-based objective score as well as a subjective assessment of clinical utility.

RESULTS: Participating in the study were 16 novices and 9 physicians. Novices took longer to complete the US exams (median 641.5 s vs. 256 s). For the objective component, novices scored higher in exams evaluating for pneumothorax (100% vs. 87.5%). For the subjective component, novices more often obtained clinically useful exams in the assessment of cardiac regional wall motion abnormalities (56.3% vs. 11.1%). No other comparisons yielded statistically significant differences between the two groups. Both groups had generally higher scores for pulmonary examinations compared to cardiac. There was variability in the quality of exams carried out by novices depending on their expert guide.

CONCLUSION: Remotely guided novices are able to carry out cardiac and pulmonary US examinations with similar, if not better, technical proficiency and clinical utility as US trained physicians, though they take longer to do so.

Dufurrena Q, Ullah KI, Taub E, Leszczuk C, Ahmad S. ***Feasibility and clinical implications of remotely guided ultrasound examinations*****. Aerosp Med Hum Perform. 2020; 91(7):592–596.**

The use of ultrasound (US) continues to increase in various medical specialties, including aerospace medicine. From its origins in radiology, the technology has been taken out of the traditional healthcare setting and into the austere environments encountered in fields such as combat medicine, global health, and aerospace medicine.[3] Additionally, US is used routinely by emergency medicine and critical care specialists at the bedside in what has been coined point of care US (POCUS).[7] POCUS has become an increasingly important tool in the evaluation of critical conditions such as shock, respiratory failure, and abdominal emergencies.[8] Compared to other imaging modalities such as roentgenology, computed tomography, and magnetic resonance imaging, POCUS has the benefits of portability, the absence of radiation, ease of education for medical professionals, and perhaps most importantly, enables clinical decision making based on real-time results.

In conjunction with the increased availability of modern digital communication networks, US images can now be transmitted from remote environments, whether it be Antarctica or low-Earth orbit, to appropriate consultants or remotely located providers. This has enabled non-experts and even nonmedically trained personnel to capture and transmit US data to distantly located experts who can then assist in making critical medical decisions. The feasibility of utilizing remotely guided POCUS exams has been explored in the aerospace medicine literature in multiple studies. A portable US has been taken to Mt. Everest, where novice operators were guided by an expert to carry out POCUS exams of the chest to evaluate for signs of

From the Department of Internal Medicine and the Department of Pulmonary, Critical Care and Sleep Medicine, Stony Brook University Hospital, Stony Brook, NY, USA; the Department of Internal Medicine, Montefiore Medical Center, Bronx, NY, USA; and the Renaissance School of Medicine at Stony Brook University, Stony Brook, NY, USA.

This manuscript was received for review in October 2019. It was accepted for publication in April 2020.

Address correspondence to: Sahar Ahmad, M.D., Division of Pulmonary, Critical Care and Sleep Medicine, Stony Brook University Hospital, HSC T17-040, Stony Brook, NY, USA; Sahar.ahmad@stonybrookmedicine.edu.

DOI: https://doi.org/10.3357/AMHP.5527.2020

high-altitude pulmonary edema.[9] An astronaut aboard the International Space Station (ISS) carried out a focused assessment with sonography for trauma exam on themselves under guidance by a radiologist at mission control with satisfactory results.[10] Other studies carried out in space have shown the feasibility of using US to evaluate for trauma of the eye, shoulder injury, and even comprehensive echocardiograms.[1,2,4] Remotely guided US has even been suggested as a modality to carry out lumbar punctures aboard the ISS in order to better study spaceflight associated neuro-ocular syndrome.[6]

While the feasibility of remotely guided POCUS imaging has been established, objective comparison to imaging obtained by medical professionals has yet to be carried out. Without such a comparison it is difficult to describe the reliability and overall value of remotely guided POCUS. This is the objective of our work. In this study, nonmedical personnel, with minimal to no medical training, carried out limited cardiac and pulmonary POCUS exams under the guidance of a remotely positioned physician with POCUS expertise. The same exams were then carried out independently by physicians with POCUS training (although not US experts) and scores between the two groups, both objective and subjective, were statistically compared.

METHODS

Subjects

A convenience sample of 16 novices (10 women, 6 men; ages 19–21; 0 yr US training; 0 yr medical training) and 9 trained physicians (2 women, 7 men; ages 26–32, 2–5 yr US training; 2–5 yr post graduate medical training) were recruited for the study. Standardized patients (SPs) from the institution's 'simulation center' were recruited to serve as subjects for the US examinations. Novices were recruited as volunteers from the institution's affiliated university campus. Inclusion criteria for novices included less than 6 mo of formal medical training of any type and no experience with clinical US. Trained physician subjects were recruited from the institution's internal medicine residency and pulmonary-critical care fellowship graduate medical training programs. Inclusion criteria for physicians included formal US training during their medical training and between 1–3 yr experience in performing POCUS for medical decision-making. POCUS experts provided remote guidance to novices and evaluated image quality from both the novice and physician groups. POCUS expertise was defined as attending physician level of training (>3 yr post graduate), >1-yr training in critical care, and with at least 3 yr of clinical POCUS experience. Experts who guided novices were not involved in the scoring of images. The study was approved by the university's institutional review board and all participants provided written consent before participating.

Procedure

All POCUS exams for the study were carried out at the institution's 'simulation center', which is designed for observation and training of medical students and postgraduate trainees. Sonosite M-Turbo portable US machines with p21 transducers (Fujifilm, Bothell, WA) were used by both the novice and physician groups. A webcam was directed at the SPs, allowing visualization of the anatomical areas of the SP being examined, the US screen, and the US operator's hand and probe positioning. Video from the webcam, as well as real-time imaging from the US machine screen, were transmitted to the US expert, located in a separate room, via Skype (Microsoft, Palo Alto, CA) and internal closed-circuit technology, respectively (**Fig. 1**). Additionally, two-way audio communication was used between the novices and the guiding expert.

Novices were given a brief orientation to their task and a conceptual introduction to US technology. They were then instructed to carry out the exams by one of the two guiding POCUS experts. The expert instructed the novices on all necessary steps to carry out the exams, including transducer probe and mode selection, probe placement, gain adjustment and video clip storage, and other technical tasks. Exams assessing for pulmonary edema, pneumothorax, pleural effusions, pericardial effusion, left ventricular (LV) dysfunction, and LV regional wall motion abnormalities (RWMA) were carried out. The existing aerospace medicine literature has highlighted several of these conditions, from pulmonary edema in high-altitude pulmonary edema to pneumothorax and pericardial effusion seen in trauma, as important conditions that can occur in the aerospace medicine environment and can be diagnosed with POCUS.[5,11,12] The following POCUS views were carried out to evaluate for the above conditions: bilateral anterior lung fields (one image per anatomical side), bilateral pleural

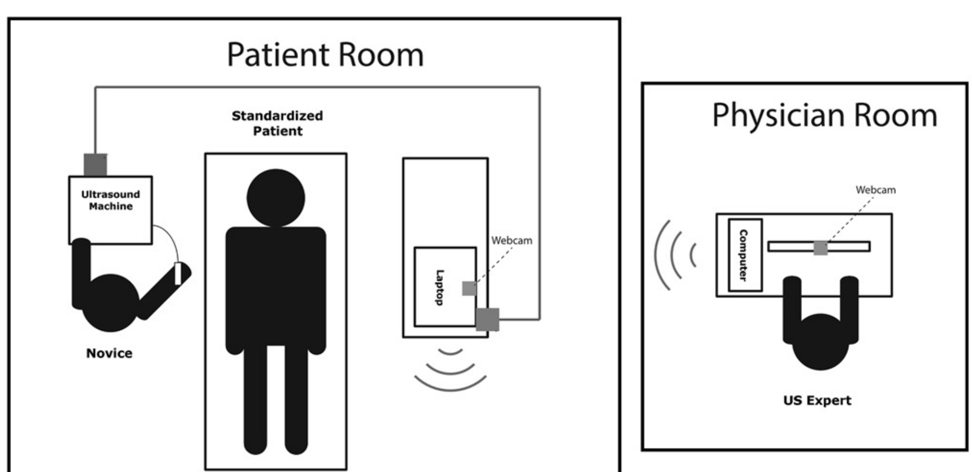

Fig. 1. Experimental schematic for the novice group.

spaces, parasternal long axis (PSLA), and parasternal short axis (PSSA).

The physician group carried out the same US exams, though independently without guidance from an US expert. They were given only written instructions to carry out the particular POCUS views, the conditions they were assessing for, and which US 'signs' to evaluate for. For example, the trained physicians were instructed to carry out an 'Anterior Lung Field' exam, to assess for 'pulmonary edema,' and to evaluate for 'A-lines.' One of the guiding US experts carried out the same US exam on select standardized patients to serve as a 'gold standard' and to ensure that images were feasible on the SPs. Time to complete each exam was recorded.

Video clips of the listed views were saved and later evaluated by two POCUS experts not otherwise involved in study. These evaluating experts were blinded to the operator's level of training and were given specific instructions to carry out systematic assessments of each POCUS exam. Each exam received an objective and subjective score.

The objective component was based on the Quality of Ultrasound Imaging and Competence Score, which has been validated in the literature for scoring abdominal ultrasound.[13] This score included a binary (yes/no) score to determine if specific technical tasks were carried out appropriately, such as correct probe selection, appropriate gain, and correct placement of probe, among other technical features of the exam. Objective scores were reported as a percentage of the maximum possible score achievable. The subjective component used a Likert scale to qualitatively assess the clinical usefulness of the exams. An example of such is 'With this exam I am able to confidently rule out pulmonary edema.' Scores ranged from 'Strongly Disagree' to 'Strongly Agree' and were converted numerically for statistical analysis (1–5, respectively) (**Fig. 2**). Exams that were scored as either 'Agree' or 'Strongly Agree' (>3) were deemed to be 'clinically useful.' These distinctions of clinical utility were determined by internal expert consensus as no existing guidelines were found in the literature to guide this novel study. Additionally, to evaluate potential differences in the expert guiders' abilities, average scores from their respective novice participants were compared.

Task-Specific Checklist	Yes	No
Appropriate Machine Settings Used?		
Appropriate Gain Used?		
Appropriate Depth Used?		
Visualizes right ventricle clearly?		
Visualizes left ventricle clearly?		
Visualizes all left ventricular walls clearly?		
Overall Image Quality Acceptable?		

With This Exam I Am Able To Confidently Rule Out The Following:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Large/Clinically Significant Pericardial Effusion					
Severe Left Ventricular Dysfunction					
Regional Wall Motion Abnormalities					

Fig. 2. Scoring template examples: objective/task based, subjective/clinical utility.

Statistical Analysis

All analyses were carried out using SAS v9.4 (SAS Institute Inc., Cary, NC). The scores given by the two evaluating experts were averaged for each individual examination. Median scores as well as interquartile range for each group were reported. Objective scores were analyzed as continuous variables, while subjective scores were analyzed as categorical variables. All continuous variables were not normally distributed and therefore the Wilcoxon rank sum was used. Due to the small sample size, the Fisher exact test was used for categorical variables. Separate POCUS exams for the same clinical conditions (e.g., left pleural space and right pleural space) were combined and reported as an aggregate score. Scores from the individual pulmonary and cardiac examinations were combined to generate a total score for each organ system. Time to completion of the entire set of POCUS exams were compared between the groups. Exam scores of a single guiding US expert, performed to establish a 'gold standard' comparator, are reported. No statistical analysis was carried out between the scores of the US expert and the other groups due to the small sample size of the expert's results, and as this was not the primary focus of this study. P-values $\leq$ 0.05 were considered significant.

RESULTS

Novices required a significantly longer amount of time to carry out the included POCUS exams when compared to the physician group (median 641.5 s vs. 256 s, $P < 0.0001$). Median time to completion for the US expert was 209.5 s. For the objective component, novices achieved significantly higher scores for pneumothorax evaluation (100% vs. 87.5%, $P = 0.0326$).

No other comparisons between the objective scores achieved statistical significance, though novice scores trended higher for several examinations (lung parenchyma, PSSA, overall lung, and cardiac). Trained physicians' scores trended higher in assessments of the pleural spaces and PSLA, though without statistical significance. Generally, the US expert scored very high objective scores, with median scores greater than 90% for all pulmonary exams, and the lowest score for PSSA (median 64.3%) (**Table I**).

For the subjective score, novices achieved a clinically useful exam more often than trained physicians in the evaluation of RWMA (56.3% vs. 11.1%, $P = 0.0405$). Similar to the objective scores, novice scores trended higher for several examinations without statistical significance (pneumothorax, pulmonary edema, LV dysfunction, and overall pulmonary and cardiac exams). Trained physician scores trended higher for assessment

Table I. Objective Scores and % of Maximum Score (Interquartile Range).

EXAM	NOVICES ($N = 16$)	TRAINED PHYSICIANS ($N = 9$)	EXPERT ($N = 1$)	*P*-VALUE
Time to completion (seconds)	641.5 (109)	256 (167)	209.5 (23.0)	<0.0001*
Evaluation for pneumothorax	100.0 (6.3)	87.5 (20.8)	97.9 (4.2)	0.0326*
Evaluation of parenchymal pattern	89.6 (25.0)	79.2 (20.8)	100.0 (0.0)	0.5486
Evaluation of pleural space	72.9 (18.8)	75.0 (25.0)	91.7 (8.3)	0.3189
Combined lung	85.4 (19.5)	81.3 (15.3)	96.5 (1.4)	0.4102
Parasternal long axis	62.50 (46.9)	68.8 (25.0)	81.3 (0.0)	0.4596
Parasternal short axis	64.3 (42.9)	57.1 (28.6)	64.3 (14.3)	0.4406
Combined cardiac	66.5 (37.5)	62.9 (8.9)	72.8 (7.1)	0.9774

* $P < 0.05$; statistical comparison between Novices and Trained Physicians only, not Expert.

of pleural and pericardial effusions, though again, results were not statistically significant. Again, the US expert tended to score high in the subjective component as well, achieving clinically useful exams 100% of the time for all but two evaluations (LV dysfunction and pleural effusion) (**Table II**).

Scores achieved by novices varied significantly based on who their guiding expert was, with those under guidance of Expert 1 achieving higher scores compared to novices guided by Expert 2 (median objective score 87.4% vs. 67.3%, $P < 0.01$; median converted subjective score 3.3 vs. 2.1, $P < 0.01$).

DISCUSSION

This study found that nonmedical personnel can carry out remotely guided cardiac and pulmonary POCUS examinations with similar technical proficiency and clinical utility as nonexpert US trained physicians. Scores, both in objective and subjective components, showed few statistically significant differences between the two groups and, in instances where there were differences, the novice group scored higher than the trained physicians. Though novices did require a significantly longer amount of time to complete the study, this difference is not clinically significant. The median time used by the novices is reasonable for such an extensive cardiopulmonary POCUS examination and, in a true emergency, a more limited POCUS exam could be carried out at the discretion of the guiding expert. While these findings indicate that the presence of a physician may not be needed for diagnostic purposes, this does not obviate the advantage of having a medical professional present. If a critical medical condition were identified by POCUS, such as tension pneumothorax or pericardial tamponade, the ability to intervene would be severely limited in the absence of a medical professional. In the most emergent of cases, however, remotely guided interventions may also be feasible.

Across all groups, the highest scores were noted in the pulmonary-specific POCUS exams. This finding is explained, at least in part, by the relative difficulty of carrying out cardiac POCUS exams. Variability of body habitus between individuals, the difficulty of finding a representative two-dimensional image of a three-dimensional organ, and the need to maintain probe placement to evaluate movement of the heart structures render cardiac US examinations technically difficult. Within the cardiac examinations, assessment for LV dysfunction and RWMA are dependent on the evaluation of multiple structures throughout systole and diastole, while pericardial effusion can be more easily ruled out by simply visualizing the absence of fluid in the pericardial space. It is not surprising, therefore, that assessment of LV dysfunction and RWMA were among the lowest scored exams, including at the expert level, while assessment for pericardial effusion was more consistently achieved. Our findings indicate that while remotely guided novices may be able to reliably rule out conditions such as pulmonary edema and pneumothorax, examinations for others such as pleural effusion and cardiac conditions should be interpreted with caution.

The difference in scores seen within the novice group based on their expert guider is an important finding. This highlights that even among US experts, there are differences in the ability to convey this expertise and guide novices to achieve useful POCUS imaging. Among the experts in our study, Expert 1 had more experience in US education compared to Expert 2, which likely explains the difference in scores. This finding indicates that additional training of US experts in

Table II. Subjective Scores and Number (%) of Exams in Each Group Determined to be Clinically Useful, i.e., Scored Either 'Agree' or 'Strongly Agree' to Question 'With This Video I Am Able to Confidently Rule Out…'.

EXAM	NOVICES ($N = 16$)	TRAINED PHYSICIANS ($N = 9$)	EXPERT ($N = 1$)	*P*-VALUE
Pneumothorax	12 (80.0)	4 (44.4)	2 (100.0)	0.0994
Pulmonary edema	11 (68.8)	6 (66.7)	2 (100.0)	1.0000
Pleural effusion	4 (25.0)	4 (44.4)	1 (50.0)	0.3942
Combined lung	7 (43.8)	3 (33.3)	2 (100.0)	0.6913
Pericardial effusion	8 (50.0)	5 (55.6)	2 (100.0)	1.0000
LV dysfunction	4 (25.0)	0 (0.0)	1 (50.0)	0.2601
RWMA	9 (56.3)	1 (11.1)	2 (100.0)	0.0405*
Combined cardiac	8 (50.0)	1 (11.1)	2 (100.0)	0.0875

* $P < 0.05$. Statistical comparison between Novices and Trained Physicians only, not Expert.

remote guidance may be needed if the technique is to be more widely implemented.

Limitations of this study included a relatively small sample size and a paucity of standardized, validated scoring systems for POCUS in the literature. The inclusion of the guiding US expert's scores helped to mitigate this lack of a validated scoring system. The high scores achieved by the expert for most exams indicate that all POCUS views, with perhaps the exception of PSSA, were obtainable on the SPs. An additional limitation is the use of multiple SPs, which was required due to logistical constraints. A total of four SPs were used and, while similar in age and body habitus, they were not identical anatomically, making strict comparison of scores across SPs imperfect.

In conclusion, this study showed that remotely guided novices are able to carry out cardiac and pulmonary POCUS examinations as well as, if not better than, trained physicians. These findings reaffirm what has been seen in previous studies in the literature and highlight the importance of this resource in austere environments where medical personnel may not be present. Limitations of this technique must be acknowledged, however, as shown by relatively lower scores in more difficult US examinations such as cardiac POCUS. Additional training, both for novices who may be called upon to use US and for guiding experts, would help to mitigate these limitations and further expand the utility of this resource in aerospace medicine.

ACKNOWLEDGMENTS

Financial Disclosure Statement: The authors have no competing interests to declare.

Authors and affiliations: Quinn Dufurrena, M.D., and Erin Taub, M.P.H., Department of Internal Medicine, and Sahar Ahmad, M.D., Department of Pulmonary, Critical Care and Sleep Medicine, Stony Brook University Hospital, Stony Brook, NY, USA; Kazi Imran Ullah, M.D., Department of Internal Medicine, Montefiore Medical Center, Bronx, NY, USA; and Connor Leszczuk, B.M., Renaissance School of Medicine at Stony Brook University, Stony Brook, NY, USA.

REFERENCES

1. Chiao L, Sharipov S, Sargsyan AE, Melton S, Hamilton DR, et al. Ocular examination for trauma; clinical ultrasound aboard the International Space Station. J Trauma. 2005; 58(5):885–889.
2. Fincke EM, Padalka G, Lee D, van Holsbeek M, Sargsyan AE, et al. Evaluation of shoulder integrity in space: first report of musculoskeletal US on the International Space Station. Radiology. 2005; 234(2):319–322.
3. Gharahbaghian L, Anderson KL, Lobo V, Huang R-W, Poffenberger CM, Nguyen PD. Point-of-care ultrasound in austere environments: a complete review of its utilization, pitfalls, and technique for common applications in austere settings. Emerg Med Clin North Am. 2017; 35(2):409–441.
4. Hamilton DR, Sargsyan AE, Martin DS, Garcia KM, Melton SL, et al. On-orbit prospective echocardiography on International Space Station crew. Echocardiography. 2011; 28(5):491–501.
5. Johansen BD, Blue RS, Castleberry TL, Antonsen EL, Vanderploeg JM. Point-of-care ultrasound for pulmonary concerns in remote spaceflight triage environments. Aerosp Med Hum Perform. 2018; 89(2):122–129.
6. Lerner DJ, Chima RS, Patel K, Parmet AJ. Ultrasound guided lumbar puncture and remote guidance for potential in-flight evaluation of VIIP/SANS. Aerosp Med Hum Perform. 2019; 90(1):58–62.
7. Moore CL, Copel JA. Point-of-care ultrasonography. N Engl J Med. 2011; 364(8):749–757.
8. Narasimhan M, Koenig SJ, Mayo PH. A whole-body approach to point of care ultrasound. Chest. 2016; 150(4):772–776.
9. Otto C, Hamilton DR, Levine BD, Hare C, Sargsyan AE, et al. Into thin air: extreme ultrasound on Mt Everest. Wilderness Environ Med. 2009; 20(3):283–289.
10. Sargsyan AE, Hamilton DR, Jones JA, Melton S, Whitson PA, et al. FAST at MACH 20: clinical ultrasound aboard the International Space Station. J Trauma. 2005; 58(1):35–39.
11. Wagner MS, Garcia K, Martin DS. Point-of-care ultrasound in aerospace medicine: known and potential applications. Aviat Space Environ Med. 2014; 85(7):730–739.
12. Yang W, Wang Y, Qiu Z, Huang X, Lv M, et al. Lung ultrasound is accurate for the diagnosis of high- altitude pulmonary edema: a prospective study. Can Respir J. 2018; 2018:5804942.
13. Ziesmann MT, Park J, Unger BJ, Kirkpatrick AW, Vergis A, et al. Validation of the quality of ultrasound imaging and competence (QUICk) score as an objective assessment tool for the FAST examination. J Trauma Acute Care Surg. 2015; 78(5):1008–1013.