Portable Ultrasound Device Usage and Learning Outcomes Among Internal Medicine Trainees: A Parallel-Group Randomized Trial

Andre Kumar, MD, MEd*, Yingjie Weng†, MHS, Libo Wang, MD‡, Jason Bentley, PhD‡
Marta Almli, MD, JD§, Jason Horne, MD‡, Ronald Witteles, MD‡, Neera Ahuja, MD‡, John Kugler, MD‡

1Department of Medicine, Stanford University School of Medicine, Stanford, California; 2Quantitative Science Unit, Stanford University School of Medicine, Stanford, California.

BACKGROUND: Little is known about how to effectively train residents with point-of-care ultrasonography (POCUS) despite increasing usage.

OBJECTIVE: This study aimed to assess whether handheld ultrasound devices (HUDs), alongside a year-long lecture series, improved trainee image interpretation skills with POCUS.

METHODS: Internal medicine intern physicians (N = 149) at a single academic institution from 2016 to 2018 participated in the study. The 2017 interns (n = 47) were randomized 1:1 to receive personal HUDs (n = 24) for patient care vs no-HUDs (n = 23). All 2017 interns received a repeated lecture series regarding cardiac, thoracic, and abdominal POCUS. Interns were assessed on their ability to interpret POCUS images of normal/abnormal findings. The primary outcome was the difference in end-of-the-year assessment scores between interns randomized to receive HUDs vs no-HUDs. Secondary outcomes included trainee scores after repeating lectures and confidence with POCUS. Intern scores were also compared with historical (2016, N = 50) and contemporaneous (2018, N = 52) controls who received no lectures.

RESULTS: Interns randomized to HUDs did not have significantly higher image interpretation scores (median HUD score: 0.84 vs no-HUD score: 0.84; P = .86). However, HUD interns felt more confident in their abilities. The 2017 cohort had higher scores (median 0.84), compared with the 2016 historical control (median 0.71; P = .001) and 2018 contemporaneous control (median 0.48; P < .001). Assessment scores improved after first-time exposure to the lecture series, while repeated lectures did not improve scores.

CONCLUSIONS: Despite feeling more confident, personalized HUDs did not improve interns’ POCUS-related knowledge or interpretive ability. Repeated lecture exposure without further opportunities for deliberate practice may not be beneficial for mastering POCUS.

Journal of Hospital Medicine 2020;15:154-159. © 2020 Society of Hospital Medicine

Point-of-care ultrasonography (POCUS) can transform healthcare delivery through its diagnostic and therapeutic expediency.1 POCUS has been shown to bolster diagnostic accuracy, reduce procedural complications, decrease inpatient length of stay, and improve patient satisfaction by encouraging the physician to be present at the bedside.2-8

POCUS has become widespread across a variety of clinical settings as more investigations have demonstrated its positive impact on patient care.1,9-12 This includes the use of POCUS by trainees, who are now utilizing this technology as part of their assessments of patients.13,14 However, trainees may be performing these examinations with minimal oversight, and outside of emergency medicine, there are few guidelines on how to effectively teach POCUS or measure competency.13,14 While POCUS is rapidly becoming a part of inpatient care, teaching physicians may have little experience in ultrasound or the expertise to adequately supervise trainees.14 There is a growing need to study what trainees can learn and how this knowledge is acquired.

Previous investigations have demonstrated that inexperienced users can be taught to use POCUS to identify a variety of pathological states.2-3,15-23 Most of these curricula used a single lecture series as their pedagogical vehicle, and they variably included junior medical trainees. More importantly, the investigations did not explore whether personal access to handheld ultrasound devices (HUDs) improved learning. In theory, improved access to POCUS devices increases opportunities for authentic and deliberate practice, which may be needed to improve trainee skill with POCUS beyond the classroom setting.14

This study aimed to address several ongoing gaps in knowledge related to learning POCUS. First, we hypothesized that personal HUD access would improve trainees’ POCUS-related knowledge and interpretive ability as a result of increased practice opportunities. Second, we hypothesized that trainees who
receive personal access to HUDs would be more likely to perform POCUS examinations and feel more confident in their interpretations. Finally, we hypothesized that repeated exposure to POCUS-related lectures would result in greater improvements in knowledge as compared with a single lecture series.

METHODS
Participants and Setting
The 2017 intern class \((n = 47)\) at an academic internal medicine residency program participated in the study. Control data were obtained from the 2016 intern class (historical control; \(n = 50\)) and the 2018 intern class (contemporaneous control; \(n = 52\)). The Stanford University Institutional Review Board approved this study.

Study Design
The 2017 intern class \((n = 47)\) received POCUS didactics from June 2017 to June 2018. To evaluate if increased access to HUDs improved learning outcomes, the 2017 interns were randomized 1:1 to receive their own personal HUD that could be used for patient care and/or self-directed learning \((n = 24)\) vs no-HUD \((n = 23)\) (Figure). Learning outcomes were assessed over the course of 1 year (see “Outcomes” below) and were compared with the 2016 and 2018 controls. The 2016 intern class had completed a year of training but had not received formalized POCUS didactics (historical control), whereas the 2018 intern class was assessed at the beginning of their year (contemporaneous control; Figure). In order to make comparisons based on intern experience, baseline data for the 2017 intern class were compared with the 2018 intern class, whereas end-of-study data for 2017 interns were compared with 2016 interns.

Outcomes
The primary outcome was the difference in assessment scores at the end of the study period between interns randomized to receive a HUD and those who were not. Secondary outcomes included differences in HUD usage rates, lecture attendance, and assessment scores. To assess whether repeated lecture exposure resulted in greater amounts of learning, this study evaluated for assessment score improvements after each lecture block. Finally, trainee attitudes toward POCUS and their confidence in their interpretative ability were measured at the beginning and end of the study period.

Curriculum Implementation
The lectures were administered as once-weekly didactics of 1-hour duration to interns rotating on the inpatient ward rotation. This rotation is 4 weeks long, and each intern will experience the rotation two to four times per year. Each lecture contained two parts: (1) 20-30 minutes of didactics via Microsoft PowerPoint\textsuperscript{TM} and (2) 30-40 minutes of supervised practice using HUDs on standardized patients. Four lectures were given each month: (1) introduction to POCUS and ultrasound physics, (2) thoracic/lung ultrasound, (3) echocardiography, and (4) abdominal POCUS. The lectures consisted of contrasting cases of normal/abnormal videos and clinical vignettes. These four lectures were repeated each month as new interns rotated on service. Some interns experienced the same content multiple times, which was intentional in order to assess their rates of learning over time. Lecture contents were based on previously published guidelines and expert consensus for teaching POCUS in internal medicine.\textsuperscript{13, 24-26} Content from the Accreditation Council for Graduate Medical Education (ACGME) and
the American College of Emergency Physicians (ACEP) was also incorporated because these organizations had published relevant guidelines for teaching POCUS.13,26 Further development of the lectures occurred through review of previously described POCUS-relevant curricula.27-32

Handheld Ultrasound Devices
This study used the Philips LumifyTM, a United States Food and Drug Administration–approved device. Interns randomized to HUDs received their own device at the start of the rotation. It was at their discretion to use the device outside of the course. All devices were approved for patient use and were encrypted in compliance with our information security office. For privacy reasons, any saved patient images were not reviewed by the researchers. Interns were encouraged to share their findings with supervising physicians during rounds, but actual oversight was not measured. Interns not randomized to HUDs could access a single community device that was shared among all residents and fellows in the hospital. Interns reported the average number of POCUS examinations performed each week via a survey sent during the last week of the rotation.

Assessment Design and Implementation
Assessments evaluating trainee knowledge were administered before, during, and after the study period (Figure). For the 2017 cohort, assessments were also administered at the start and end of the ward month to track knowledge acquisition. Assessment contents were selected from POCUS guidelines for internal medicine and adaptation of the ACGME and ACEP guidelines.13,24,26 Additional content was obtained from major society POCUS tutorials and deidentified images collected by the study authors.13,24,33 In keeping with previously described methodology, the images were shown for approximately 12 seconds, followed by five additional seconds to allow the learner to answer the question.32 Final assessment contents were determined by the authors using the Delphi method.34 A sample assessment can be found in the Appendix Material.

Surveys
Surveys were administered alongside the assessments to the 2016-2018 intern classes. These surveys assessed trainee attitudes toward POCUS and were based on previously validated assessments.27,28,30 Attitudes were measured using 5-point Likert scales.

Statistical Analysis
For the primary outcome, we performed generalized binomial mixed-effect regressions using the survey periods, randomization group, and the interaction of the two as independent variables after adjusting for attendance and controlling of intra-intern correlations. The bivariate unadjusted analysis was performed to display the distribution of overall correctness on the assessments. Wilcoxon signed rank test was used to determine score significance for dependent score variables (R-Statistical Programming Language, Vienna, Austria).

RESULTS
Baseline Characteristics
There were 149 interns who participated in this study (Figure). Assessment/survey completion rates were as follows: 2016 control: 68.0%; 2017 preintervention: 97.9%; 2017 postintervention: 89.4%; and 2018 control: 100%. The 2017 interns reported similar amounts of prior POCUS exposure in medical school (Table 1).

Primary Outcome: Assessment Scores (HUD vs no HUD)
There were no significant differences in assessment scores at the end of the study between interns randomized to personal HUD access vs those to no-HUD access (Table 1). HUD interns reported performing POCUS assessments on patients a mean 6.8 (standard deviation [SD] 2.2) times per week vs 6.4 (SD 2.9) times per week in the no-HUD arm (P = .66). The mean lecture attendance was 75.0% and did not significantly differ between the HUD arms (Table 1).

Secondary Outcomes
Impact of Repeating Lectures
The 2017 interns demonstrated significant increases in pre-block vs postblock assessment scores after first-time exposure to the lectures (median preblock score 0.61 [interquartile range (IQR), 0.53-0.70] vs postblock score 0.81 [IQR, 0.72-0.86]; P < .001; Table 2). However, intern performance on the preblock vs postblock assessments after second-time exposure to the curriculum failed to improve (median second preblock score 0.78 [IQR, 0.69-0.83] vs postblock score 0.81 [IQR, 0.64-0.89]; P = .94). Intern performance on individual domains of knowledge for each block is listed in Appendix Table 1.

Intervention Performance vs Controls
The 2016 historical control had significantly higher scores compared with the 2017 preintervention group (P < .001; Appendix Table 2). The year-long lecture series resulted in significant increases in median scores for the 2017 group (median pre-intervention score 0.55 [IQR, 0.41-0.61] vs median postintervention score 0.84 [IQR, 0.71-0.90]; P = .006; Appendix Table 1). At the end of the study, the 2017 postintervention scores were significantly higher across multiple knowledge domains compared with the 2016 historical control (Appendix Table 2).

Survey Results
Notably, the 2017 intern class at the end of the intervention did not have significantly different assessment scores for several disease-specific domains, compared with the 2016 control (Appendix Table 2). Nonetheless, the 2017 intern class reported higher levels of confidence in these same domains despite similar scores (Supplementary Figure). The HUD group seldomly cited a lack of confidence in their abilities as a barrier to performing POCUS examinations (17.6%), compared with the no-HUD group (50.0%), despite nearly identical assessment scores between the two groups (Table 1).
DISCUSSION

Previous guidelines have recommended increased HUD access for learners,13,24,35,36 but there have been few investigations that have evaluated the impact of such access on learning POCUS. One previous investigation found that hospitalists who carried HUDs were more likely to identify heart failure on bedside examination.37 In contrast, our study found no improvement in interpretative ability when randomizing interns to carry HUDs for patient care. Notably, interns did not perform more POCUS examinations when given HUDs. We offer several explanations for this finding. First, time-motion studies have demonstrated that internal medicine interns spend less than 15% of their time toward direct patient care.38 It is possible that the demands of being an intern impeded their ability to perform more POCUS examinations on their patients, regardless of HUD access. Alternatively, the interns randomized to no personal access may have used the community device more frequently as a result of the lecture series. Given the cost of HUDs, further studies are needed to assess the degree to which HUD access will improve trainee interpretive ability, especially as more training programs consider the creation of ultrasound curricula.10,11,24,39,40

This study was unique because it followed interns over a year-long course that repeated the same material to assess rates of learning with repeated exposure. Learners improved their scores after the first, but not second, block. Furthermore, the median scores were nearly identical between the first postblock assessment and second preblock assessment (0.81 vs 0.78), suggesting that knowledge was retained between blocks. Together, these findings suggest there may be limitations of traditional lectures that use standardized patient models for practice. Supplementary pedagogies, such as in-the-moment feedback with actual patients, may be needed to promote mastery.14,35

Despite no formal curriculum, the 2016 intern class (historical control) had learned POCUS to some degree based on their higher assessment scores compared with the 2017 intern class during the preintervention period. Such learning may be informal, and yet, trainees may feel confident in making clinical decisions without formalized training, accreditation, or oversight. As suggested by this study, adding regular didactics or giving trainees HUDs may not immediately solve this issue. For assessment items in which the 2017 interns did not significantly differ from the controls, they nonetheless reported higher confidence in their abilities. Similarly, interns randomized to HUDs less frequently cited a lack of confidence in their abilities, despite similar scores to the no-HUD group. Such confidence may be incongruent with their actual knowledge or ability to safely use POCUS. This phenomenon of misplaced confidence

### TABLE 1. Participation and Performance of the 2017 Intern Class

<table>
<thead>
<tr>
<th>2017 Intern Class Overall</th>
<th>HUD</th>
<th>No HUD</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Participants</td>
<td>47</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>Preintervention Quiz Response (%)</td>
<td>46</td>
<td>24 (100)</td>
<td>22 (95.7)</td>
</tr>
<tr>
<td>Postintervention Quiz Response (%)</td>
<td>42</td>
<td>21 (88.0)</td>
<td>21 (91.3)</td>
</tr>
<tr>
<td>Participated in a POCUS Course in Medical School (%)</td>
<td>10 (21.3)</td>
<td>6 (25)</td>
<td>4 (17.1)</td>
</tr>
<tr>
<td>Number of Lectures Attended (%)</td>
<td>N = 84</td>
<td>N = 44</td>
<td>N = 40</td>
</tr>
<tr>
<td>Introduction</td>
<td>55 (65.5)</td>
<td>29 (65.9)</td>
<td>26 (65.0)</td>
</tr>
<tr>
<td>Abdominal</td>
<td>55 (65.5)</td>
<td>27 (61.4)</td>
<td>28 (70.0)</td>
</tr>
<tr>
<td>Thoracic</td>
<td>68 (81.0)</td>
<td>37 (84.1)</td>
<td>31 (77.5)</td>
</tr>
<tr>
<td>Cardiac</td>
<td>51 (60.7)</td>
<td>27 (61.4)</td>
<td>24 (60.0)</td>
</tr>
<tr>
<td>Mean Number of Blocks per Intern</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Mean Times per Week Performing POCUS (SD)</td>
<td>6.57 (2.62)</td>
<td>6.80 (2.21)</td>
<td>6.40 (2.93)</td>
</tr>
<tr>
<td>Reported a Lack of Confidence in Interpreting POCUS Images</td>
<td>35.9%</td>
<td>17.6%</td>
<td>50.0%</td>
</tr>
<tr>
<td>Median Knowledge Domain Score [IQR]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.86 [0.76-0.91]</td>
<td>0.86 [0.79-0.92]</td>
<td>0.86 [0.75-0.89]</td>
</tr>
<tr>
<td>Lung</td>
<td>0.80 [0.60-1.00]</td>
<td>0.80 [0.60-1.00]</td>
<td>0.80 [0.60-1.00]</td>
</tr>
<tr>
<td>Echo</td>
<td>0.89 [0.74-0.94]</td>
<td>0.91 [0.82-0.91]</td>
<td>0.86 [0.73-0.95]</td>
</tr>
<tr>
<td>Abdominal</td>
<td>0.89 [0.78-0.89]</td>
<td>0.78 [0.78-0.89]</td>
<td>0.89 [0.78-0.89]</td>
</tr>
<tr>
<td>Volume Assessment</td>
<td>0.75 [0.50-0.75]</td>
<td>0.75 [0.50-0.75]</td>
<td>0.75 [0.50-0.75]</td>
</tr>
<tr>
<td>Anatomical Identification</td>
<td>0.91 [0.86-0.95]</td>
<td>0.91 [0.86-0.95]</td>
<td>0.91 [0.82-0.95]</td>
</tr>
<tr>
<td>Normal Findings</td>
<td>0.70 [0.45-0.80]</td>
<td>0.60 [0.40-0.80]</td>
<td>0.80 [0.60-0.80]</td>
</tr>
<tr>
<td>Abnormal Findings</td>
<td>0.85 [0.69-0.92]</td>
<td>0.85 [0.77-0.92]</td>
<td>0.85 [0.69-0.92]</td>
</tr>
</tbody>
</table>

Abbreviations: HUD, handheld ultrasound device; IQR, interquartile range; POCUS, point-of-care ultrasound; SD, standard deviation.
TABLE 2. Outcomes of Learning with Repeated Lecturesa

<table>
<thead>
<tr>
<th>Knowledge Domain</th>
<th>First Block</th>
<th>Second Block</th>
<th>P Value</th>
<th>First Block</th>
<th>Second Block</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before</td>
<td>After</td>
<td>P Value</td>
<td>Before</td>
<td>After</td>
<td>P Value</td>
</tr>
<tr>
<td>Number of Completed Assessments</td>
<td>36</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Overall Correctness [IQR]</td>
<td>0.61 [0.53-0.70]</td>
<td>0.81 [0.72-0.86]</td>
<td>&lt;.001</td>
<td>0.78 [0.69-0.83]</td>
<td>0.81 [0.64-0.89]</td>
<td>.94</td>
</tr>
<tr>
<td>Lung</td>
<td>0.20 [0.20-0.60]</td>
<td>0.80 [0.60-1.00]</td>
<td>&lt;.001</td>
<td>0.60 [0.40-0.80]</td>
<td>0.80 [0.00-1.00]</td>
<td>.57</td>
</tr>
<tr>
<td>Echo</td>
<td>0.64 [0.55-0.73]</td>
<td>0.75 [0.68-0.83]</td>
<td>.003</td>
<td>0.82 [0.73-0.91]</td>
<td>0.82 [0.64-0.91]</td>
<td>.40</td>
</tr>
<tr>
<td>Abdominal</td>
<td>0.72 [0.67-0.78]</td>
<td>0.89 [0.78-0.89]</td>
<td>.005</td>
<td>0.78 [0.78-0.89]</td>
<td>0.78 [0.78-0.89]</td>
<td>.47</td>
</tr>
<tr>
<td>Volume Assessment</td>
<td>0.25 [0.25-0.50]</td>
<td>0.75 [0.50-0.75]</td>
<td>.001</td>
<td>0.50 [0.25-0.75]</td>
<td>0.75 [0.00-0.75]</td>
<td>.61</td>
</tr>
<tr>
<td>Anatomical Identification</td>
<td>0.70 [0.55-0.78]</td>
<td>0.87 [0.77-0.91]</td>
<td>&lt;.001</td>
<td>0.87 [0.74-0.91]</td>
<td>0.87 [0.78-0.96]</td>
<td>.64</td>
</tr>
<tr>
<td>Normal Findings</td>
<td>0.40 [0.35-0.60]</td>
<td>0.60 [0.40-0.80]</td>
<td>.003</td>
<td>0.60 [0.60-0.60]</td>
<td>0.80 [0.20-0.80]</td>
<td>.86</td>
</tr>
<tr>
<td>Abnormal Findings</td>
<td>0.58 [0.46-0.71]</td>
<td>0.77 [0.62-0.85]</td>
<td>.001</td>
<td>0.77 [0.69-0.85]</td>
<td>0.77 [0.62-0.85]</td>
<td>.59</td>
</tr>
</tbody>
</table>

aAfter the second lecture series, the 2017 interns generally improved their median score postlecture vs prelecture series. However, after the second series, the interns had no difference in median score postlecture versus prelecture series. We included only interns that had both before-session and after-session assessments in the comparisons. Wilcoxon signed rank tests on dependent samples were performed to calculate the P values.

Abbreviation: IQR, interquartile range

is known as the Dunning–Kruger effect, and it may be common with ultrasound learning.41 While confidence can be part of a holistic definition of competency,13 these results raise the concern that trainees may have difficulty assessing their own competency level with POCUS.25

There are several limitations to this study. It was performed at a single institution with limited sample size. It examined only intern physicians because of funding constraints, which limits the generalizability of these findings among medical trainees. Technical ability assessments (including obtaining and interpreting images) were not included. We were unable to track the timing or location of the devices’ usage, and the interns’ self-reported usage rates may be subject to recall bias. To our knowledge, there were no significant lapses in device availability/functionality. Intern physicians in the HUD arm did not receive formal feedback on personally acquired patient images, which may have limited the intervention’s impact.

In conclusion, internal medicine interns who received personal HUDs were not better at recognizing normal/abnormal findings on image assessments, and they did not report performing more POCUS examinations. Since the minority of a trainee’s time is spent toward direct patient care, offering trainees HUDs without substantial guidance may not be enough to promote mastery. Notably, trainees who received HUDs felt more confident in their abilities, despite no objective increase in their actual skill. Finally, interns who received POCUS-related lectures experienced significant benefit upon first exposure to the material, while repeated exposures did not improve performance. Future investigations should stringently track trainee POCUS usage rates with HUDs and assess whether image acquisition ability improves as a result of personal access.

Disclosures. Dr. Kumar received a Stanford Seed Grant for Junior Faculty to purchase equipment used in the study. Dr. Witteles received honorarium from Pfizer and Alnylam Pharmaceuticals outside the submitted work. All other authors have nothing to disclose.

References