Validation of the Peripheral Ultrasound-guided Vascular Access Rating Scale

Primdahl, Stine C.; Weile, Jesper; Clemmesen, Louise; Madsen, Kristian R.; Subhi, Yousif; Petersen, Poul; Graumann, Ole

Published in:
Medicine (United States)

DOI:
10.1097/MD.0000000000009576

Publication date:
2018

Document version
Final published version

Document license
CC BY-ND

Citation for published version (APA):
Primdahl, S. C., Weile, J., Clemmesen, L., Madsen, K. R., Subhi, Y., Petersen, P., & Graumann, O. (2018). Validation of the Peripheral Ultrasound-guided Vascular Access Rating Scale. Medicine (United States), 97(2), [9576]. https://doi.org/10.1097/MD.0000000000009576

Terms of use
This work is brought to you by the University of Southern Denmark through the SDU Research Portal. Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

• You may download this work for personal use only.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk
Validation of the Peripheral Ultrasound-guided Vascular Access Rating Scale

Stine C. Primdahl, MD\(^{a,b,*}\), Jesper Weile, MD\(^{a,c}\), Louise Clemmesen, MD\(^{d}\), Kristian R. Madsen, MD\(^{e}\), Yousif Subhi, MD\(^{d,f}\), Poul Petersen, MD, PhD\(^{g}\), Ole Graumann, MD, PhD\(^{h,i,j}\)

Abstract

Evidence-based standards in proficiency are needed for ultrasound-guided peripheral intravenous access. In this study, we explored the validity of the Peripheral Ultrasound-Guided Vascular Access (P-UGVA) Rating Scale.

We recruited 3 groups of physicians (5 novices, 5 intermediates, and 5 experts) of increasing proficiency in peripheral ultrasound-guided intravenous access. All participants performed 3 peripheral ultrasound-guided intravenous accesses on three different patients. Performance was video-recorded by 3 cameras and the ultrasound image. Synchronized and anonymized split-screen film clips were rated using the P-UGVA rating scale by 2 assessors, which also assessed overall performance on a 1–5 Likert-scale. Evidence of validity was explored using the contemporary validity framework by Messick (content, response process, internal structure, relations to other variables, and consequences).

Content and response process was ensured in the development of the rating scale and validity study. Internal consistency of the P-UGVA rating scale was excellent and sufficient high for certification purposes (Cronbach’s alpha = 0.91). Proficiency groups were successfully discriminated by the UPGIVA rating scale (\(P = .029\), one-way ANOVA), and the P-UGVA rating scale scores also correlated strongly with the overall performance evaluations (\(\rho = 0.87\), \(P < .001\), Pearson correlation). We calculated a pass/fail score of 29, which lead to a theoretical false positive rate of 26.5% and false negative rate of 8.5%.

We present validity evidence for the P-UGVA rating scale and an evidence-based standard in proficiency for ultrasound-guided peripheral intravenous access.

Abbreviation: P-UGVA = peripheral ultrasound-guided vascular access.

Keywords: administration, education, intravenous, medical, ultrasonography

1. Introduction

Most hospitalized patients need a vascular access for treatment and diagnostic purposes.\(^{(1)}\) Ultrasound guidance is shown to reduce needle insertion attempts, reduce complication rates, and increase successful catheter placements when compared with the landmark technique. Consequently, the ultrasound-guided procedure improves patient satisfaction and reduces the requirement for central line placements.\(^{(2,3)}\) The benefits of ultrasound-guided peripheral intravenous access seem crystal clear clinically and from the patient’s perspective, but from an implementation perspective it is still unclear how to sufficiently augment clinicians level of proficiency.

An important detail about ultrasound-guided procedures is that they are highly operator-dependent\(^{(8,9)}\) and a sufficient competence level of the practitioner must be ensured to guarantee patient safety.\(^{(9)}\) Requirements of a minimum number of performed procedures or a certain time of experience are inadequate to evaluate proficiency, for example, because of individual variations in the practitioners’ learning curves.\(^{(10)}\) Evaluations for a range of clinical procedures are increasingly focused on actual performance rather than pseudomeasures of proficiency such as training time. This mastery learning approach requires that objective standards can be set for the procedure in question, after which the learner can focus on training until the set standard is reached. Accumulating scientific evidence supports the benefit of this approach.\(^{(11)}\)

We are currently challenged by no clear evidence-based standard in proficiency of ultrasound-guided intravenous access. In a Delphi consensus study, we recently developed a global rating scale for competence assessment of Peripheral Ultrasound-Guided Vascular
Access (P-UGVA). It is important when implementing such measures to ensure validity towards the construct—otherwise the concept of training towards a set level of standard falls flat as the standard becomes irrelevant.

In this study, we aimed to explore the validity of the P-UGVA rating scale using the contemporary validity framework by Messick. The framework consists of 5 sources of validity (content, response process, internal structure, relations to other variables, and consequences) that should be explored in validity studies. In addition, we aimed to establish an evidence-based pass/fail level, which can be used as a standard for proficiency in ultrasound-guided intravenous access.

2. Methods

2.1. Study design

This prospective validity study of P-UGVA rating scale on physicians was conducted in the Emergency Department at the Regional Hospital West Jutland, Herning, Denmark from April 2015 to December 2015. The study was registered with the Central Denmark Region Committees on Health Research Ethics and a waiver was granted for performing the study (cf. inquiry 105/2014). Oral and written consent were obtained from all patients before entering the study. All aspects of the study followed the ethical principles for research as stated in the Declaration of Helsinki.

2.2. Participants

We recruited 3 groups of physicians (novices, intermediates, and experts) in relation to their competence in ultrasound-guided peripheral intravenous access. A novice was defined as a physician who had placed ≤ 10 ultrasound-guided peripheral intravenous accesses, an intermediate had placed between 10 to 25, and an expert had placed ≥ 25. All physicians were recruited from the Emergency Department and the Department of Anaesthesiology at Regional Hospital West Jutland, Herning, Denmark and all had experience with the conventional landmark technique. The novice group and the intermediate group consisted of medical interns, whereas the expert group had senior residents and specialist physicians. Power calculation was not possible as no previous validity studies on the UGPIVA rating scale exist. We aimed to recruit 5 participants in each of the 3 groups totalling 15 physicians for the study.

2.3. Study setting

All participants performed three ultrasound-guided peripheral intravenous accesses on three different patients. Participating patients were all in need of a peripheral intravenous access because of their medical condition. The patient was placed in a bed before the arrival of the physician. A table drawer was prepared with the procedural items (gloves, a tourniquet, sterile gel, alcohol swabs, an IV catheter, a fixation patch, a cotton pad, and a needles bucket). The table top itself was empty, thereby yielding the physicians option to individualize preparation for the procedure. Before the procedure, the table was placed near the foot end of the patient’s bed along with the ultrasound scanner, a chair, and a trashcan. When the physician arrived to the room he or she was asked to perform the ultrasound-guided procedure and to place the equipment, as he or she found best. The Edge ultrasound system with a HFL50x linear array transducer 6–15MHz (SonoSite Inc., Bothell, WA) was used. The physicians were filmed from three angles with 3 video cameras during the procedures. One GoPro 3+ Black Edition (GoPro Inc., San Mateo, CA) camera recorded from above the patient’s arm in which the vascular access was placed and two Canon Legria HFR506 camcorders (Canon Inc., Tokyo, Japan) recorded the arm from each side of the patient (Fig. 1). These 3 angles enabled a total overview of the room and the physician’s
2.4. Assessment of video-recorded performance

All recordings were synchronized and placed in one split-screen image frame, so the procedure was visible from different angles at the same time. All post production editing was performed using Final Cut Pro X version 10.2 (Apple Inc., Cupertino, CA) by one author (SCP). When the procedure was finalized, the film clips from the 3 different angles and the corresponding ultrasound film were all stored on an external hard drive as mpeg4 format files.

2.5. Assessors

The 2 assessors (authors LC and KRM) are both specialists in anaesthesiology and skilled experts in ultrasound-guided intervention with each more than 10 years of experience with peripheral ultrasound-guided intravenous access. The assessors are not affiliated to the department or the hospital where the physicians were recruited for the study. Initially, the 2 assessors were instructed in how to rate all the physicians’ performances according to the scale, and individually rated 3 pilot films. All ratings were made using the P-UGVA rating scale. In addition, the assessors were asked to rate the physician’s overall performance in each film on a 5-point Likert scale. The pilot films showed the performance of a novice, an expert, and an intermediate. After the assessors had rated the test films, the authors’ met with the 2 assessors to discuss the results. Disagreements between the 2 assessors were discussed until agreement was obtained on all elements. Within 4 following days, the assessors rated the same 3 pilot films again and their ratings were approximately alike. Results from the pilot films are not included in the analyses. Afterwards, assessors were allowed to rate the films starting from September 2015 and the last film was rated in December 2015.

2.6. Evidence of validity

We used the contemporary validity framework as introduced by Messick, which describes 5 sources of validity.

Content, which describes the relevance of the measure to the construct. This source of validity was ensured by developing the P-UGVA global rating scale through a Delphi consensus approach. The Delphi consensus technique is an anonymous structured approach used to obtain consensus among experts through a number of Delphi rounds.

Response process, which is to eliminate or control potential sources of bias. This was addressed by performing a pilot study to detect any potential sources of bias in the data collection process. Participants where blinded to the content of the scale they were assessed by. All videos were anonymized before exposure to the assessors. We trained the assessors using pilot videos.

Internal structure, which is the degree to which items measuring similar constructs produce similar results. This internal consistency was explored using Cronbach’s alpha.

Relations to other variables, which is how well global rating scores correlate to other variables of proficiency such as experience level. This part was evaluated by comparing P-UGVA rating scale scores between the groups of predefined competence levels. In addition, we evaluated whether the P-UGVA rating scale scores were representative to the assessors’ general impression by correlating the P-UGVA rating scale scores to the corresponding overall performance scores.

Consequences, which is the social implications of obtaining a certain score. For this part, we established a pass/fail level using
the contrasting groups’ method. Using the pass/fail level, we explored the consequences of this standard in terms of false positives (% of novices that will pass) and false negatives (% of experienced that will fail).

2.7. Data analysis and statistics

For each participant, we calculated a mean score from the 3 attempts. This approach allows for fluctuation in the performance, which we hypothesized would be the case when we include nonselected patient cases. Internal consistency was explored by calculating Cronbach’s alpha. For the relations to other variables and consequences analyses, we calculated the mean of the scores from the assessors. Relation of P-UGVA rating scale scores to predefined competence levels were explored using a 1-way analysis of variance (ANOVA). Correlation between P-UGVA rating scale scores and overall performance scores were made using the Pearson’s correlation. Contrasting groups’ standard setting method was used to calculate a pass/fail score. This score is identified as the intersect between distribution curves of 2 groups of interest (novices and experts). We calculated false positives and false negatives using the cumulative distribution function for each curve (to the right and to the left of the point of intersect, respectively). Because of the small numbers in validation studies, we did sensitivity analyses, where consequences analyses were repeated after excluding participants one by one; however, because of the large number of analyses and results only significant changes to the direction of the results will be noted. All statistical analyses were made using SPSS version 23.0.0.0 (IBM Corp., Armonk, NY). P-values below 0.05 were interpret as statistical significance.

3. Results

All 15 physicians completed three P-UGVA rating scale procedures. Scores were obtained from 42 out of the 45 (93%) performed procedures. Three films—2 from the novice group and 1 from the expert group—were left out, because of one missing film clip in each case. The statistical analyses were based on scores from 13 novice performances, 15 intermediate performances, and 14 expert performances (Fig. 3).

3.1. Content

The global P-UGVA rating scale was developed by presenting Danish ultrasound experts from the specialties anesthesiology, emergency medicine and radiology with a number of key elements on the P-UGVA procedure. The experts determined the importance of each key element on a 5-point Likert scale. The key elements were modified and new key elements were suggested until all ultrasound experts found all elements important or extremely important and no new key elements were suggested. Details about the Delphi study for developing the P-UGVA rating scale is available in a previous report. The final rating scale consists of eight key elements: preparation of utensils, ergonomics, preparation of the ultrasound device, identification of blood vessels, anatomy, hygiene, coordination of the needle, and completion of the procedure (Table 1).

3.2. Response process

We did pilot investigations to address potential issues in the response process (see details in the Methods section). After the assessor training, the assessors were in agreement when they rerated the pilot videos.

3.3. Internal structure

Internal consistency of the P-UGVA rating scale was excellent and sufficiently high for certification purposes (Cronbach’s alpha = 0.91).

3.4. Relations to other variables

The P-UGVA rating scale was able to discriminate between the three groups (Fig. 4) (P=.029, one-way ANOVA). Novices received a mean of 25.3 (SD: 6.0) points, intermediates a mean of 30.7 (SD: 2.1) points, and experts a mean of 33.0 (SD: 2.9) points. In a post-hoc analysis where we looked at comparisons between the groups, this trend of increasing P-UGVA rating scale score with proficiency level was strongest between the novices and the experts where it reached a level of statistical
There was a strong correlation between the P-UGVA scores and the overall performance scores (Fig. 4) ($\rho = 0.87$, $P < .001$, Pearson correlation).

### 3.5. Consequences

We calculated a pass/fail score of 29 (Fig. 5). Using the distribution curves, this would theoretically lead to a false positive rate of 26.5% and a false negative rate of 8.5%.
Sensitivity analyses suggested that one novice was particularly talented and received a high score, and that considering that individual as an outlier and excluding from the analyses would not change the pass/fail score, but would lower the theoretical false positive rate to 12.5% and false negative rate to 6.5%.

4. Discussion

Overall, we found strong evidence of validity for P-UGVA rating scale from all 5 sources in the contemporary framework. Alignment of content was ensured through development of the scale using a Delphi-process on input from experts. Response process was ensured through pilot studies and training of the assessors. Internal structure showed excellent reliability at a level considered sufficiently high for certification (Cronbach’s alpha > 0.8). The P-UGVA rating scale scores were related to other variables, as experts significantly outperformed novices, and scores correlated strongly with the overall performance assessments. Finally, consequences of a pass/fail standard were explored, which suggested acceptable false positive and negative rates.

Rice et al. have in a previous study developed a simulation-based assessment tool, a checklist, for the placement of ultrasound-guided peripheral Intravenous access. A simulation based tool permits the assessors to make standardized observations; it is however limited by not portraying a real life clinical setting, where placement difficulties differ due to factors as the patients’ clinical state, depth and diameter of veins, and so on. Our study is stronger when compared to the study of Rice et al. by portraying the clinical setting, and also by developing the content of the scale through the Delphi technique. Further, we developed a global rating scale as it is regarded superior to checklists. The P-UGVA rating scale makes it possible for the instructor to give the performer feedback on the specific items that the performer needs to improve. This makes the P-UGVA rating scale a necessity in training sessions. We ensured the validity of the P-UGVA rating scale by use of the contemporary validity framework as introduced by Messick. The contemporary framework is regarded superior to the traditional framework when testing the quality of an assessment instrument and therefore strengthen our study further.

This study explored validity of the P-UGVA rating scale in a real-life clinical setting with several cameras, synchronized ultrasound images, anonymized film clips, and multiple procedures and assessors. These are strengths of this study, but can also be considered as limitations that should be noted in relation to local implementation of the P-UGVA rating scale. Our assessors were able to go back or pause in the film to ensure correct assessment. This is not possible with onsite assessment and further complicates the issue by nonblinded assessors, which previous studies find may significantly influence the results because of a preconception of proficiency from factors such as the age, sex, or the title of the evaluated physician. Categorization of participants into different levels based on experience was challenging. This was somewhat in conflict with the overall statement of the study that ultrasound learning curves are heterogeneous and level of expertise was not based on the number of procedures performed. However, there was large difference in experience between the three groups and thereby risk of misclassification was limited. The presented evaluation required 2 assessors and 3 procedures, and studies are needed to explore the impact on validity sources by changing the numbers of assessors and procedures. It is practically and economically desirable with fewest possible assessors and procedures, but this must be balanced with an acceptable level of reliability.

In conclusion, we find evidence of validity for the P-UGVA rating scale and provide an evidence-based standard for proficiency in ultrasound-guided peripheral intravenous access. Future studies will determine its final role in the training and certification of physicians.

Acknowledgements

The authors would like to thank MD Associate Professor Bo Løfgren, MS Sinne Eika Rasmussen, and MS Mette Amalie Kristensen for lending us video cameras and equipment. The authors would also like to thank nurse and sales consultant Anne Mette Bove from Secma for lending us the Sonosite Edge for the study.

References

[1] Troianos CA, Hartman GS, Glas KE, et al. Guidelines for performing ultrasound guided vascular cannulation: recommendations of the American Society of Echocardiography and the Society of Cardiovascular Anesthesiologists. Anesth Analg 2012;114:46–72.
[2] Randolph AG, Cook DJ, Gonzales CA, et al. Ultrasound guidance for placement of central venous catheters: a meta-analysis of the literature. Crit Care Med 1996;24:2053–8.
[3] Hind D, Calvert N, McWilliams R, et al. Ultrasound locating devices for central venous cannulation: meta-analysis. BMJ 2003;327:361.
[4] Costantino TG, Parikh AK, Satz WA, et al. Ultrasonography-guided peripheral intravenous access versus traditional approaches in patients with difficult intravenous access. Ann Emerg Med 2005;46:436–61.
[5] Au AK, Rote MJ, Grzybowski RJ, et al. Decrease in central venous catheter placement due to use of ultrasound guidance for peripheral intravenous catheters. Am J Emerg Med 2012;30:1930–4.
[6] Shokoohi H, Boniface K, McCarthy M, et al. Ultrasound-guided peripheral intravenous access program is associated with a marked reduction in central venous catheter use in noncritically ill emergency department patients. Ann Emerg Med 2013;61:198–203.
[7] Karakitsos D, Labropoulos N, De Grooth E, et al. Real-time ultrasound-guided catheterisation of the internal jugular vein: a prospective
comparison with the landmark technique in critical care patients. Crit Care 2006;10:R162.

[8] Hertzberg BS, Kliewer MA, Bowie JD, et al. Physician training requirements in sonography: how many cases are needed for competence? Am J Roentgenol 2000;174:1221–7.

[9] Moore C, Copel J. Point-of-care ultrasonography. N Engl J Med 2011;749–57.

[10] Jang TB, Ruggeri W, Dyne P, et al. The learning curve of resident physicians using emergency ultrasonography for cholelithiasis and cholecystitis. Acad Emerg Med 2010;17:1247–52.

[11] Cook DA, Brydges R, Zendejas B, et al. Mastery learning for health professionals using technology-enhanced simulation: a systematic review and meta-analysis. Acad Med 2013;88:1178–86.

[12] Primdahl SC, Todsen T, Clemmesen L, et al. Rating scale for the assessment of competence in ultrasound-guided peripheral vascular access: a Delphi consensus study. J Vasc Access 2016;17:440–5.

[13] Ghaderi I, Manji F, Park YS, et al. Technical skills assessment toolbox: a review using the unitary framework of validity. Ann Surg 2015;261:251–62.

[14] Thomsen AS, Subhi Y, Kilgaard JF, et al. Update on simulation-based surgical training and assessment in ophthalmology: a systematic review. Ophthalmology 2015;122:1111–30.

[15] Subhi Y, Todsen T, Konge L. An integrable, web-based solution for easy assessment of video-recorded performances. Adv Med Educ Pract 2014;5:103–5.

[16] Graham B. Delphi as a method to establish consensus for diagnostic criteria. J Clin Epidemiol 2003;56:1150–6.

[17] Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs 2000;32:1008–15.

[18] Downing SM. Reliability: on the reproducibility of assessment data. Med Educ 2004;38:1006–12.

[19] Rice J, Crichlow A, Baker M, et al. An assessment tool for the placement of ultrasound-guided peripheral intravenous access. J Grad Med Educ 2016;8:202–7.

[20] Ilgen JS, Ma IW, Hatala R, et al. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. Med Educ 2015;49:161–73.

[21] Val Y, Vogt , Vanessa M, et al. Is a resident’s score on a videotaped objective structured assessment of technical skills affected by revealing the resident’s identity? Am J Obstet Gynecol 2003;189:688–91.

[22] Konge L, Vitmann P, Clemmentsen P, et al. Reliable and valid assessment of competence in endoscopic ultrasonography and fine-needle aspiration for mediastinal staging of non-small cell lung cancer. Endoscopy 2012;44:928–33.