

## CLINICAL INVESTIGATION

# Validity and reliability of an objective structured assessment tool for performance of ultrasound-guided regional anaesthesia

A. Shafqat<sup>1,\*</sup>, M. Rafi<sup>2</sup>, V. Thanawala<sup>2</sup>, N. M. Bedford<sup>2</sup>, J. G. Hardman<sup>1</sup> and R. A. McCahon<sup>2</sup>

<sup>1</sup>Anaesthesia & Critical Care, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, UK and <sup>2</sup>Department of Anaesthesia, Queen's Medical Centre, Nottingham University Hospitals NHS Trust, Nottingham, UK

\*Corresponding author. E-mail: [atif.shafqat@nhs.net](mailto:atif.shafqat@nhs.net)

## Abstract

**Background:** We examined the validity and reliability of the previously developed criterion-referenced assessment checklist (AC) and global rating scale (GRS) to assess performance in ultrasound-guided regional anaesthesia (UGRA).

**Methods:** Twenty-one anaesthetists' single, real-time UGRA procedures (total: 21 blocks) were assessed using a 22-item AC and a 9-item GRS scored on 3-point and 5-point Likert scales, respectively. We used one-way analysis of variance to compare the assessment scores between three groups (Group 1:  $\leq 30$  blocks in the preceding year; Group 2: 31–100; and Group 3:  $> 100$ ). The concurrent validity was evaluated using Pearson's correlation ( $r$ ). We calculated Type A intra-class correlation coefficient using an absolute-agreement definition in two-way random effects model, and inter-rater reliability using an absolute agreement between raters. The inter-item consistency was assessed by Cronbach's  $\alpha$ .

**Results:** The greater UGRA experience in the preceding year was associated with better AC [ $F(2, 18) 12.01$ ;  $P < 0.001$ ] and GRS [ $F(2, 18) 7.44$ ;  $P = 0.004$ ] scores. There was a strong correlation between the mean AC and GRS scores [ $r = 0.73$  ( $P < 0.001$ )], and a strong inter-item consistency for AC ( $\alpha = 0.94$ ) and GRS ( $\alpha = 0.83$ ). The intra-class correlation coefficient (95% confidence interval) and inter-rater reliability (95% confidence interval) for AC were 0.96 (0.95–0.96) and 0.91 (0.88–0.95), respectively, and 0.93 (0.90–0.94) and 0.80 (0.74–0.86) for GRS.

**Conclusions:** Both assessments differentiated between individuals who had performed fewer ( $\leq 30$ ) and many ( $> 100$ ) blocks in the preceding year, supporting construct validity. It also established concurrent validity and overall reliability. We recommend that both tools can be used in UGRA assessment.

**Keywords:** anaesthetists; checklist; educational assessment; reproducibility of results; ultrasound

**Editor's key points**

- Ultrasound-guided regional anaesthesia is a complex skill that requires reliable assessment tools for use in training.
- This study critically evaluated the validity and reliability of a previously developed assessment checklist and global rating scale for performance assessment and training.
- These tools may be used for effective feedback, with strong inter-rater agreement and inter-item consistency.

Reduced clinical opportunities during training, an increased focus on optimal patient safety, and greater public accountability have led to the need for an objective assessment of procedural skills in medicine.<sup>1,2</sup> The assessment of expertise in medicine may be formative (developmental) or summative (pass/fail). Assessments assist practitioners towards expert practice whilst protecting patients by ensuring that safe, acceptable standards of practice are maintained. Assessments must be sufficiently valid and reliable to withstand scrutiny and challenge from the learner and patient groups; they must be credible and consistent in order that they have value and meaning.<sup>3,4</sup> After the publication of recommendations for training in ultrasound-guided regional anaesthesia (UGRA),<sup>5</sup> a group of 18 UGRA experts used a modified Delphi technique to develop a criterion-referenced assessment checklist (AC) and the global rating scale (GRS) to assess the technical and non-technical aspects of UGRA performance.<sup>5,6</sup> However, the authors stated that future work should concentrate on establishing further evidence to support the validity and reliability of these assessments. Therefore, we examined the ability of the AC and the GRS to quantify the level of expertise in UGRA in anaesthetists (construct validity). We also examined the degree of inter-rater agreement and consistency of each assessment tool, and finally, the strength of agreement between the two assessments (concurrent validity).

**Methods**

We requested an ethics review by the University of Nottingham Medical School Research Ethics Committee, which approved the study (approval reference: K09052013LT 13053 SCS Anaesthesia). Anaesthetists working at the Nottingham University Hospitals NHS Trust were invited to participate in the study via e-mail. A participant information sheet was forwarded to those who expressed an interest and a written informed consent was gained in advance of any study activity. The patients of the participating anaesthetists were also given an information leaflet before their surgery, and their written informed consent was sought in the morning of their surgery. Each participating anaesthetist was given a participant identification number before the commencement of the study.

This dual-site, blinded observational study was conducted concurrently at the Queen's Medical Centre and the City Hospital campuses of the Nottingham University Hospitals NHS Trust. Anaesthetists were eligible for study inclusion if they planned to perform an ultrasound-guided nerve or plexus block as part of their usual management of a patient, and the patient had agreed to take part. The exclusion criteria included anaesthetists or patients who did not wish to participate and

patients who did not require UGRA. The clinical decision to perform UGRA was taken in all cases by the attending anaesthetist.

Before the commencement of the UGRA procedure, each participant completed a self-reported questionnaire with regard to the number of ultrasound-guided nerve blocks they had completed in the preceding year. To minimise observer bias, both the investigators were kept blinded from the completed self-reported questionnaire, which was submitted to them in a sealed envelope. Subsequently, two anaesthetist investigators (A.S. and M.R.) observed the participants together and used AC and GRS to assess independently the UGRA performance by participants during routine operating lists. Assessment occurred in real time during the performance of a single UGRA procedure by each participant, began with the initial preparation and set-up of equipment, and ended at completion of the procedure. The two investigators completed the assessments simultaneously and did not influence the clinical practice of the participants in any way.

A.S. and M.R. had been trained to use both assessment tools (AC and GRS) before the study commencement. This involved a week of practice assessment sessions (5 half-days) with a facilitated debriefing from the research team, so that both assessors were familiar with the assessment tools and that they had a shared understanding of the UGRA performance. In brief, the AC comprises 22 items scored on a 3-point Likert scale [not performed (0); poorly performed (1); well performed (2)] (Appendix A), whereas the GRS consists of nine categories scored on a 5-point Likert scale with descriptive anchors of performance to assist in scoring (Appendix B).<sup>6</sup> One of the categories of GRS that is the item 'overall performance' was excluded from the calculation of GRS score. In addition to that, we did not record a 'pass/fail' assessment.

**Statistics**

In line with previous studies, we estimated that we would need to recruit between 20 and 40 participant anaesthetists.<sup>7–11</sup> For the purpose of analysis, we arbitrarily allocated all the participants to one of three groups, based on the self-reported questionnaire with regard to the number of ultrasound-guided nerve blocks they had completed in the preceding year (Group 1:  $\leq 30$ ; Group 2: 31–100; and Group 3:  $>100$ ).

The statistical analysis used STATA/IC version 10.0 (Stata-Corp, College Station, TX, USA). Normality of data was assessed by histogram and the Shapiro–Wilk and skewness/kurtosis tests. To test whether higher total assessment scores were associated with a greater number of ultrasound-guided nerve blocks in the preceding year, we used a one-way analysis of variance (ANOVA) to compare the AC and GRS scores for Groups 1–3. Where a significant difference was identified, we performed appropriate *post hoc* comparisons with a Bonferroni adjustment to adjust for multiple comparisons.

We performed an exploratory analysis of the relationships between the values of AC score, GRS score, response to GRS item 'overall performance', and number of blocks in the preceding year by calculating the Spearman correlation coefficient  $\rho$ . Similarly, we evaluated the concurrent validity of the assessment tools by calculating the Pearson correlation coefficient ( $r$ ). In all analyses, we used a two-tailed *P*-value less than 0.05 to indicate statistical significance.

To assess the inter-rater agreement, we calculated the Type A intra-class correlation coefficient (ICC) using an absolute-agreement definition in a two-way random effects model.

Download English Version:

<https://daneshyari.com/en/article/10224683>

Download Persian Version:

<https://daneshyari.com/article/10224683>

[Daneshyari.com](https://daneshyari.com)