**Reviewer’s report**

**Title:** Validation of the Imperial College Surgical Assessment Device for spinal anesthesia.

**Version:** 0  **Date:** 12 May 2017

**Reviewer:** Matthew McEvoy

**Reviewer's report:**

Paper Title: Validation of the Imperial College Surgical Assessment Device (ICSAD) for spinal anesthesia.

Thank you for the opportunity to review the manuscript entitled "Validation of the Imperial College Surgical Assessment Device (ICSAD) for spinal anesthesia." The authors set out to assess the validity and reliability of the scores produced by the ICSAD for spinal procedures on a partial task training simulator. I commend the authors on the completion of this study, but I also have several major concerns at this time concerning this study and the presentation of the results. There are also numerous minor typographical and grammatical errors throughout the paper, but I will not address those currently.

1. The ICSAD assessment tool has been previously validated with other procedures in the past, including clinical performance of labor epidurals and supraclavicular blocks. Given the fact that each of these is more complex in nature than a spinal block and given that these were performed in clinical settings, why not perform this study in the clinical setting? I am currently not convinced by the explanation of the rationale in the Discussion that a partial task trainer was the proper study setting in order to add new knowledge to the literature.

2. Overall, the study is well-designed and well performed (given the limitation noted above). It shows a difference between trainees who have done no spinals and those who have done a few or many. The same might be achieved by knowing a trainee's procedure count, requiring no complex assessment. However, I do believe that the study makes the false assumption that time in training correlates with novice, intermediate, and expert technical skills. This is actually a false premise. It is generally true that skills increase with increased experience, but it is not true for all physicians, and certainly not true for technical skills where digital natives may have much better spatial reasoning and their learning curve may be significantly different than those who trained 10-20 years ago. To assess this, I would have preferred to see all participants perform 3-5 attempts. This could have assessed consistency in performance and whether there was a learning curve on the simulator (or in the clinical setting). Analysis of repeated performance over time (e.g. CUSUM) would make the conclusions more powerful. It should also be pointed out that in Dreyfus' model of skill
acquisition, the levels are novice, advanced beginner, competent, proficient, and expert. Intermediate is not a level, nor should it be the goal of competency assessment. Assessing competency is the goal of most of these assessments, and it ought to be as it can determine when clinical autonomy can be granted.

3. Building upon the prior point, given the fact that this report is not clinical and is relying upon the prior work of validation of the ICSAD tool, I would like to see an analysis of performance with a standard setting irrespective of the training level (setting of a minimum passing score). The study proposes that it distinguishes skill level, but it assumes the manner in which the groups are divided. Advancing the educational research literature will occur when competency-based assessment of success or failure (as mentioned in the Introduction) can be undertaken without reference to time in training or time past graduation, as some trainees gain skills quite quickly and some post-graduates have skills that fade. I would recommend that the authors consider assessing their data based upon the measured performance in relation to a determined standard, which in this case may need to be created (akin to Anghof or Hofstee methods).

4. The discussion section seems focused on GRS superiority over TSC modalities. This discussion may represent the authors' beliefs and previous literature, but does not seem truly germane to the study, especially when only one of many measures showed any difference between intermediate and expert participants. This would actually point to the fact that such a study would be of great value if done in patients, as the best metric of block success is whether the procedure was done in a safe manner, whether the patient was satisfied, and whether the block achieved the desired clinical outcome for surgical anesthesia. Assessment of motion and time of placement may focus on proficiency and expertise, but these may also be misleading.

5. Abstract and manuscript cite different Cohen Kappa values - 0.87 v. 0.76.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

Yes

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

Yes
Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

None

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal