Reviewer’s report

Title: A simple, early assessment predicts upper extremity function after stroke - a prospective cohort study

Version: 2 Date: 19 March 2015

Reviewer: Cathy Stinear

Reviewer’s report:

This is an interesting and well-written study, with a relatively large number of patients. It has the potential to make a useful contribution to international efforts in the development of simple, accurate prognostic tools for stroke rehabilitation.

Minor essential revisions
1. FMA-UE is usually considered a measure of impairment, while ARAT is usually considered a measure of function. Please clarify throughout the manuscript.

Major compulsory revisions
1. This is not a prospective study, in that the process for selecting the ARAT items for the ARAT-2 was not complete until all the data had been collected and analysed. To be prospective, the study would need to test the predictive power of ARAT-2 from the outset. Please revise the title and manuscript accordingly.
2. It’s unclear how predicting a binary outcome on the FM scale (which in turn relates to the ability to drink from a cup) solves the problem identified in the Introduction, line 96: “However, in these [previous] studies … the outcome does not tell us if this predicted motor function can be useful for daily activities”. Drinking from a cup is only one of many important daily activities, and it doesn’t require much in the way of individuated finger movements, which are essential for other important daily activities.
3. The FMA-UE scores were binarised as below 32 points and 32 points or more, and this cut-off point was established in the authors’ previously published study. It seems that some of the 30 patients in the previous study might also have been part of the present study though, as both studies report on participants in the SALGOT study. Ideally, there should be no overlap, if the results from the first study of 30 patients are being used to rationalise the choice of cut-off in the second study.
4. The previous study only included patients who could perform the drinking task with their affected upper limb. These patients had FMA-UE scores ranging from 32 to 64, and ARAT scores ranging from 24 to 54. Patients who couldn’t perform the task were excluded, so it’s impossible to know whether their scores exceeded 32 on the FMA-UE or 24 on the ARAT. The choice of cut-off score would have more support if it was shown to clearly distinguish between those who can perform the task, and those who can’t. This is partly addressed in Line 135
onwards, where the accuracy of the FMA-UE cut-off in correctly classifying patients’ ability to “perform a drinking task” was evaluated at 3 time points during the study. However, these data and analyses don’t seem to be reported, and there are no details provided on the drinking task. What constitutes successful performance?

5. More generally, it seems that this study has used a drinking task to represent meaningful upper limb function, used a FMA-UE score to predict who can perform the drinking task, then used two items from the ARAT scale to predict who will exceed the FMA-UE score cut-off. The relationship between ARAT-2 and the drinking task is therefore somewhat indirect, via the FMA-UE score. A more robust test of the ARAT-2 would be to see whether it directly predicts which patients can perform a (defined) drinking task, or preferably more than one functionally meaningful tasks.

6. The results and discussion describe the ARAT-2 as able to correctly predict whether patients can perform a drinking task. It’s important to note that the ARAT-2 was used to predict whether patients would exceed the FMA-UE score cut-off. This isn’t quite the same thing as being able to perform a drinking task (which is not described).

7. It’s perhaps not surprising that the predictions made at 3 days were most accurate for FMA-UE score at 10 days. At 3 days it is probably more clinically relevant to be able to make accurate predictions for time points later in recovery.

8. Was a standardised approach used for the clinical scales, such as that outlined by Yozbatiran et al. for the ARAT (NNR, 2008)? How many raters administered the clinical scales and were they trained for consistency?

9. How does predicting whether a patient will achieve one of two outcomes on the FM scale help therapy teams and patients? How might this information be used by therapy teams?

10. What was the therapy dose completed by the patients in this study? How might variations in therapy dose affect their outcomes, and the accuracy of the predictions? Or does ARAT-2 make predictions that are accurate for any therapy dose? What might this mean for the role of therapy?

11. The discussion notes (line 235) that accurate predictions are more difficult for patients with higher levels of initial impairment. This is true of prediction tools that rely solely on clinical assessment. In contrast, accuracy for these patients can be higher when using measures of corticomotor pathway integrity. Simple clinical assessment is unable to detect intact, but not yet functional, corticomotor connections that can support subsequent recovery. This point could be usefully added to the possible reasons outlined in line 236 onwards.

12. It would be helpful for the reader if the description of the third approach (line 260) was more informative, and identified as the algorithm from reference 13, in place of “objective evaluation”. This will help the reader make the link back to the previous points in the discussion.

13. Line 288 comments on the need for early evaluation – at days 3 and 10. The following sentence mentions assessments at 1 month. The next paragraph
continues in this vein. However it’s not clear in the method that assessments at 1 month were used to make predictions at 12 months, as the focus is on predictions made from assessments at 3 days.

14. It would be good to identify the need for validation of the ARAT-2 in a new cohort of patients, in a prospective study.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.