Reviewer’s report

Title: Selecting Optimal Screening Items for Delirium: An Application of Item Response Theory

Version: 2 Date: 13 August 2012

Reviewer: Claudia Spies

Reviewer’s report:

Yang and Colleagues performed a complex and very promising analysis of a database of screening data from a multicenter trial in order to identify indicators for delirium according to the CAM features. Generally speaking this study is of high interest for delirium research, -the methodology is sophisticated but very difficult to follow. For readers unfamiliar with statistical methods such as ‘permuted parallel analysis’ (page 14, line 281) the results e.g. presented in Table 2 of the manuscript are very hard to follow. The overall analysis presented by the authors is a mixture of theory- and data-driven analyses and expert consensus decisions (e.g. lines 303-304: “We considered the preponderance of the evidence in making dimensionality decisions, together with the input of the CEP on the interpretation of secondary factors.”). It remains unclear from the description whether this approach is common praxis or was developed by the authors. If it was common praxis authors should provide the rational for this. We would suggest asking a separate statistical review by a statistician familiar with factor analyses and item response theory to judge the appropriateness of the statistical methods used.

It remains further unclear how the authors considered potential differences in age and gender. One may image that some of the source and indicator items (lines 235-237) were gender-specific or age-specific and it is unclear if this was considered by the authors. E.g. Jane and colleagues identified a gender bias in diagnostic criteria for personality disorders in DSM-IV criteria (Jane JS, Oltmanns TF, South SC, Turkheimer E. Gender bias in diagnostic criteria for personality disorders: an item response theory analysis. J Abnorm Psychol. 2007 Feb;116(1):166-75). In the actual dataset more than 2/3 of the participants were women and it remains unclear whether this fact was adequately considered by the authors.

From a clinical point of view a database with more than 4500 cases does seem appropriate for the overall task of this study.

However, there are some aspects worth noting.

Major comments:

Was Delirium determined according to DSM 4 criteria? Or according to ICD 10 criteria?

Both are the current “Gold Standards”. If the CAM was used as a reference
standard does this not exclude part of the delirious patients?? Or are DSM 4 and ICD 10 and CAM equal? Please explain in detail and add in detail to the limitations section

Subsyndromal / pre-delirium as well as severity of delirium is an issue, IRT does seem to be a suitable method to encompass a less simplistic definition of delirium than the CAM is offering. Please explain how IRT can be used to define Delirium in an ordinal approach.

P.11 l 210-211 Of the 4744 Patients, how many displayed Delirium? How was the consent handled in delirious Patients?
P.11 l.224 How, was the consent handled in Patients with low MMSE values? This needs to be explained

Minor comments:
P.5 l.80 Please refer to both current “Gold Standards” not only DSM 4!
P.5, l. 83 If I am not mistaken DSM 4 as well as ICD 10 not only capture but define Delirium criteria? Please explain at this point how ICD-10 delirium positive patients fit into the picture?
P.5, l. 84 Why are DSM 4 criteria challenging to apply? The 4 CAM criteria do not seem to be easier to apply, please explain.
P.5 l.87 – 94 What are the authors trying to explain to the reader with this passage?
The CAM is often used? That is surely not a quality indicator…
The CAM is more often used than the Gold Standards? Medically, speaking this is a frightening thought… Or do CAM positive patients reflect 100% of the DSM -4 and ICD 10 positive patients?
P.5/6 l 95-101 Are there other Delirium instruments that are easier to apply? CAM –ICU takes 90 sec in average, ICDSC takes in average 2 minutes, How about the MDAS, DRS R98, Nu-DSESC etc. they all seem to take much less time with excellent characteristics. Please explain…
P.6 l.92-99 I would assume that the binary delirium definition of the CAM is a limitation to the research question / goal at hand. Wouldn`t a severity instrument such as the MDAS or the DRS-R-98 be more helpful in this matter?
P.10 l.97 Again, did the Authors use the CAM Definition or the DSM 4 Definition for Delirium.