Reviewer’s report

Title: Recognizing Acute Delirium As part of your Routine [RADAR]: A validation study

Version: 4 Date: 25 December 2014

Reviewer: Claudia Spies

Reviewer’s report:

In their manuscript “Recognizing Acute Delirium As part of your Routine – A validation study”, Dr Voyer and colleagues studied a new screening tool for delirium (RADAR) which was designed by an expert panel to overcome three major problems of other tools they mentioned in their introduction: (a) the time effort it takes to perform daily delirium-screening; (b) the training efforts and the limited applicability for non-specialists and (c) the specific setting in which a score is validated and the problem of generalizing the findings to a wider patient collective.

In general, we appreciate the efforts the authors put on the increase of delirium screening in clinical routine. It is true that implementation rates for monitoring tools are still low in all medical settings, especially outside the ICU. Nevertheless the manuscript has major shortcomings and general issues:

MAJOR ISSUES (GENERAL):

The length of the sections should be reconsidered. We tried to make detailed recommendations for adaption.

The authors should be very precise about the facts they use to raise their clinical question: which authors/publications/references support the fact that the CAM/CAM-ICU takes 5-10 minutes in routine? Which other tests are there available? We would suggest a systematic literature search that addresses this question for the commonly used screening tools.

The other general issue is the reference-standard. It is understandable that research assistants did the CAM as a reference standard but the authors use DSM-IV-TR in their results as their reference. The DSM-IV-TR cannot be put on one level with the CAM as it is a test for delirium detection in clinical-routine as well.

MAJOR ISSUES (DETAIL)

Introduction/Background:

The introduction is very detailed and epic (>>1,000 words (!)). We suggest a clear and short introduction that outlines the problem and brings up a clear hypothesis and research question.

As the RADAR-development is part of the methodology, it should be considered to put this whole part into the methods section. The authors should exactly quote
which references underline that the CAM/CAM-ICU takes 5 to 10 minutes. Especially on the ICU the CAM-ICU is popular because of its’ fast applicability. Problems that evolve when solely using the CAM usually refer to the amount of training that is necessary to perform it with a high validity (especially in routine) and the problem that the CAM only allows a yes/no decision regarding delirium, whereas other scores (ICDSC, Nu-DESC etc.) also allow the detection of “subsyndromal” states (patients that have clinical signs for delirium do not fulfill all necessary criteria for delirium). Why do the authors only refer to the CAM?

Methodology:
The methodology is extremely detailed and might be more focused on the clear research questions that have been phrased in the introduction PLUS the development of the score (which is unknown to the potential reader).

Patient enrolment: We appreciate the author’s efforts to explain their inclusion process but it might increase the intelligibility to put this to a supplementary file and focus to the relevant points: consent was given either by the patient or his/her legal representative (if necessary).

The validation of a new test usually requires testing against the reference-standard. In case of delirium these are the DSM criteria (currently in the 5th edition). The CAM is a screening test as well. The authors should carefully describe why they have chosen the CAM as a reference standard for their setting although they doubt its’ generalizability in the introduction. Furthermore it is absolutely correct that the CAM has been designed as a structured test for checking DSM-III (three) (the authors write that on p. 7; l. 235) but afterwards the manuscript always refers to DSM-IV-text-revised (four, text revised). This is important for the entire manuscript: the authors use the CAM as the reference-standard NOT the DSM-IV-TR criteria. Only, a psychiatrist or special trained research staff should use these criteria.

The authors should explain the sentence “During the course of the study, RAs assessed samples of patients and residents simultaneously and independently for inter-rater reliability (n = 37/193; 19% of total sample). Percentage of agreement varied from 72% to 100%, while kappa values ranged from 0.30 to 1, depending of the item tested“ (p.7, ll. 246-248) in detail: how was the kappa-agreement calculated? Do the authors refer to different CAM-domains?

The authors mention the cut-off (1/3 RADAR-items positive) in the section “statistical analysis“ for the first time. Did they try other cut-offs as well? Why not opening a section where RADAR is explained in detail – why this important information to the supplementary file? (in short: how have items be chosen, how is it applied, what is the cut-off)

Results
The results section should me more focused on the main results. In general the length of the result section seems appropriate but the authors validated a score: The validation-data itself seem somewhere lost in the rest of information. Which
statistical test did the authors use for repeated measurements?

Discussion:
- The authors should also mention hyperactive and mixed delirium as RADAR items solely refer to hypoactive states.
- The authors should carefully discuss the 7-second result. The authors reveal that RADAR takes 7 seconds. How can that be the case? You should also mention the time that is necessary to gather the information to answer the questions? The 7 seconds seem rather the time it takes to make three crosses on a medical chart than the time it takes to acquire the clinical information that is necessary to answer these questions. We ask the authors to discuss this carefully.
- The authors should discuss why they did not use the reference-standard (either DSM-IV or 5) but the CAM for validation (it is not the same!).
- The authors did mention neither the testing for psychometric properties of the items, their development nor the evaluation of their cut-off. How did they manage that? Please, insert this into the main text. This should not be mentioned exclusively in the supplementary files.
- How sustainable was the “45-minutes-training-package” the authors mentioned?

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
'I declare that I have no competing interests'