Reviewer’s report

Title: The accuracy of the MMSE in detecting cognitive impairment when administered by general practitioners

Version: 1 Date: 21 September 2007

Reviewer: Hein van Hout

Reviewer's report:

General
Pezzotti et al. evaluated the accuracy of MMSE scores obtained by GPs by comparing them to scores obtained by Alzheimer’s Evaluation Units (UVA).

MMSE scores of 59 GPs with those of Alzheimer evaluation unit(s?) on 317 patients

Only suspected previously undiagnosed patients with MMSE <25 were referred.

Was there one UVA or more? If there were more do they function the same?

What did they do to standardise the scores? In the Netherlands a consensus version of the MMSE was made including the interpretation of the scores. The total score can differ easily 3 or 4 points if the assessors use different interpretations. For example, if you are strict on the date, one day wrong is a fault (while in our consensus up to three days is considered a good answer).

UVA scorers were systematically higher. The authors mention this and find several associations but do not really explain this difference. Was this the case for all participating UVA’s?

Strangely the difference between the standardised scores were much smaller, how do the authors explain this?

There is likely to be verification bias:
The difference was (partly?) explained by ‘better patients’. Borderline scoring patients (23 or 24) according to the GPs have more chance to receive higher UVA scores than patients scoring far under the cut off, simply due to measurement error and daily variation. To check this I would suggest that the authors do a sensitivity analysis excluding patients scoring around the cut off, for example by taking the standard error of the mean of the GPs scores under 24 as extra limit.

In fact the difference was largely explained by patients with MCI and patient with no impairment, which confirms this explanation.

Remarkably the population had an average score of about 16. I assume that the
diagnoses were seriously delayed otherwise I would not expect such low average scores. Can the authors comment on this?

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

ROC score for UVA has no value as their MMSE assessment was used to make the clinical diagnosis. It was part of the assessment and interpretation, consequently it has a good ROC score. The same is true for table 2 with the accuracy measures, it may hold for the GPs but it is ridiculous for the UVA. In addition, we do not know whether the UVA was blinded for the GPs MMSE score, that would additionally compromise the accuracy measures on the GP.

The ROC value for the GPs may make sense although there is the problem of the selected sample causing verification bias, and whether the GPs’ MMSE scores were blinded to the UVA team.

Except for these limitations which should be discussed, for the Discussion section I miss a wider perspective in what the role of GPs can be in timely diagnostics.

Also it would be interested to see a the problem of taboo being mentioned. I understand that, as in most countries (including Italy?) Alzheimer disease is still taboo for many patients, families and professionals. If you decide to screen and refer all suspected patients, as the authors did, it means that the GPs did not have trouble to test their patients and confront them with a referral to an Alzheimer unit. This is truly remarkable and the numbers of successful referral are indeed impressive as only 28 persons out of 397 were not tested by the GP, and only 15 out of 397 were not referred. Do the authors have any (anecdotal) information about this? See dfor further reading papers of H. van Hout or S. Iliffe.

Another point for the discussion is the timely diagnosis and ditto screening. In my opinion MMSE screening would makes sense especially for borderline persons, as for persons below 18 it is clear for anybody that cognitive impairment is a serious problem. So it raises the question whether these persons were in fact detected timely.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Another discussion point is that the MMSE is a little long for busy general practices (at least in the Netherlands, Australia and the UK). The authors could comment on a variety of briefer tests (i.e. MIS, MISplus, MINIcog, GPCOG) which are probably more attractive for use general practice (see Brodaty 2006).
Discretionary Revisions (which the author can choose to ignore)

Instead of their figure 1 the authors may consider a bland-altman plot to picture the MMSE differences between GPs and UVA. Such a plot could additionally show whether the difference was systematic or in certain score ranges only.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of limited interest

Quality of written English: Acceptable

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

Declaration of competing interests:

'I declare that I have no competing interests