Reviewer’s report

Title: Dealing with Missing Data in a Multi-question Depression Scale: A Comparison of Imputation Methods

Version: Date: 5 October 2006

Reviewer: Paula Diehr

Reviewer’s report:

General

The paper is now much clearer and easy to understand.

the abstract is still unclear. "simulated by randomly" and "non-random missing simulation" are not the correct terms and will confuse readers. I think the terms they need are "missing completely at random" and "missing at random". These need to be written exactly this way, even though the terms are stylistically ugly.

On page 6, I don't think the first method called MAR is actually MAR because it does not depend on any patient characteristics. It is just another MCAR simulation. The probability of being missing is just twice as high for question 6 as for the others. MAR means missingness depends on an observed patient characteristic.

The authors should consult with a statistician with expertise in the voluminous missing data lecture to get this nomenclature correct. It is tricky.

"non-random simulation" on page 7 should be "missing not at random". These terms must be used precisely.

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

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

Discretionary Revisions (which the author can choose to ignore)

What next?: Accept after minor essential revisions

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

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

I declare that I have no competing interests.