Reviewer's report

**Title:** Screening of Depression In Primary Care Setting - Prevalence and Instrument

**Version:** 4  **Date:** 4 April 2014

**Reviewer:** Mark Haddad

**Reviewer's report:**

Major compulsory revisions.

The key issues that are problematic have been partially addressed.

Throughout the article, the authors do not make adequate distinction between prevalence studies using diagnostic measures and those using other approaches. For instance in the lengthy section noting prevalence of depression in nations and regions, a study is cited (Pakistan/PLOSOne) that uses an unvalidated tool via telephone interview - and the degree to which the resulting estimate relates to a valid prevalence is uncertain. It is likely that a number of the studies noted provide 'prevalence' estimates that utilise self-report measures.

An issue of central importance in prevalence studies - and in studies determining the appropriateness of screening measures - is diagnostic accuracy.

It would be helpful if the authors considered this central issue and provided some statement drawing a distinction between the findings they note that are derived from appropriately conducted epidemiological studies, and those based on self-report measures, or unvalidated survey tools.

2. Previous comments noted that the authors adopt a generally uncritical approach to the value of screening for depression; although some minor alterations have been made, this article still does not provide a balanced appraisal of the benefits and harms of screening - particularly the likely extent of false positives identified by PHQ screen (likely to be =>50%).

Consideration of the range of harms associated with inaccurate evaluation of condition status is +relevant to this study/ ethical dimensions. cf http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3291670/

In relation to this, the authors note that 'Criteria for diagnosis of Major Depressive Disorder based on DSM IV, was not used as gold standard or validation of PHQ-9 cutoff in this study, as the purpose of the study was to assess the value of screening instrument for depression in primary care setting..'

This needs to be reconsidered and possibly phrased in a different way - The limitation of not using a gold standard is that there is no data/findings concerning how accurate the screen test is within the population studied; the exercise of administering the PHQ has provided findings on people with clinical features of depression who may or may not have depressive disorder. Other well-conducted diagnostic accuracy studies indicate that around half (though this depends on
true prevalence) of the screen positives will not have depression.

assessing the 'value of a screening instrument for depression' - will crucially involve evaluating its diagnostic accuracy.

3. In relation to the statement (previously commented upon - at some length) 'According to 2001 Health Report of WHO, nearly 15% of patients with major depression have lifetime risk of committing suicide.' the authors may consider it unimportant but the repetition of inaccurate findings is unscientific, and the use of secondary sources to source epidemiological findings is weak academic practice. I have previously noted a range of high quality research findings pertinent to this, and encourage the authors to revise accordingly.

4- It may be useful to make explicit reference to appropriate guidelines for practice of observational studies e.g.
http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.0040297
or http://www.ncbi.nlm.nih.gov/books/NBK53279/

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

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'