Reviewer’s report

Title: Data driven subtypes of major depressive disorder

Version: 1 Date: 10 August 2012

Reviewer: Tim Croudace

Reviewer’s report:

I agree with the four sentences that provide the Background for this paper. I think it is definitely true that bottom-up approaches from signs and symptoms are warranted in the main-stay of psychiatric research. This angle characterizes much of the motivation behind my own work and my motivation in this area.

It seems sensible to attempt a systematic review, to ASSEMBLE papers that have applied different techniques to relevant data. This paper succeeds in this respect for one domain of morbidity (depression criteria of DSM-IV). Although I think it is useful to document what studies were found and the types of (latent structure) analysis that were performed, it remains a difficult issue for me to accept that comparing results across studies in this manner provides enough useful information. I think this is really an area where, within each study, with high quality data, more can be learned when a particular study compares and contrasts methods on the same dataset. I only wish more literature accepted that challenge, rather than pushing a set of results from one modelling perspective. I think this paper is of interest therefore, but ultimately slightly flawed in the ability to succeed in its aims. I enjoyed reading it, and suspect others will too, but I think that the results reported are just an arbitrary set of those published, from amongst those that could have been performed. We need more papers applying more types of analyses to the same data, in my view. If this paper were to achieve that, then it would be a valued contribution, changing the nature of some of the to-be-published work in this field in the future. The most challenging area is how to appraise the more complex "hybrid" type models, and also to work with different measurement levels for the original symptom ratings/responses. This represents a thoughtful account of important avenues for research though, and offers some accurate arguments and perspectives for a general medical readership (of this journal). I think that the conclusions are accurate and fair.

Major Compulsory Revisions
1 Could some more suggestions be made as to how to address patient heterogeneity in somatic symptoms.

2 Could the authors forecast whether it is ever likely that a mega-analysis could be performed using IRT linking of some of the questionnaires, for example.

3 Overall, I think that the paper needs to reflect both factor analysis terminology and that of item response theory, since they are essentially the same model. Is enough mention made of multidimensional item response theory, or mixture
models (IRT-based)?

4 It is very important to specify exactly which studies used binary and ordinal data as input to their analyses (vs any continuous).

Minor Essential Revisions

5 The introduction does not really give any representation to the literature(s) on discrete latent trait models, or latent class models with random effects - both of which might be relevant.

Discretionary Revisions

6 The authors might like to consider related perspectives published in articles by David Cole et al. (for adolescent depression).

7 It is likely that related studies outside of this systematic review remit are actually relevant to the direction in which the authors would like the field to travel. If some indication of those could be introduced, then I think it would make the discussion more rounded (and inclusive), in some way.

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests.