Reviewer's report

**Title:** Psychological treatment of depression: A meta-analytic database of randomized studies

**Version:** 1  **Date:** 12 November 2007

**Reviewer:** Bruce E Wampold

**Reviewer's report:**

**General**

The compilation of RCTs of depression is an extremely valuable resource to those who wish to conduct meta-analyses of various hypotheses relative to the treatment of depression. As such, an up-to-date registry of trials and their results will be a boon to synthesis in this area.

The method of collecting the studies is very thorough; there should be few concerns about the thoroughness of the search. The nature of an on-line resource is that those who are aware of a trial not included can inform the authors and the data-base can be updated. Similarly, the inclusion of a trial can be questioned and, if the authors concur, the study can be removed or the result recalculated. Of course, such adaptations assume that the authors are responsive, both in terms of timeliness and in terms of fair evaluations of the trials. Given the authors’ diligence, these assumptions seem reasonable.

-------------------------------------------------------------------------------------------------------------------------------

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

There is an issue however that raises concerns. The authors are not only providing a registry of trials, but they are providing the effect sizes for comparisons within those trials. Users of this registry will most likely use the effect sizes provided, without checking—in this way, the authors have done a great service. Consequently, there should be no ambiguity with regard to the accuracy of the effect sizes. Anyone who has supervised meta-analytic projects knows well that not infrequently inferences must be made to extract the information needed to calculate effects—e.g., two subgroups may need to be pooled to determine means and standard deviations of a treatment. Even in the absence of such calculations, errors appear. In a meta-analysis care is taken, but the implications are relatively small in comparison of an incorrect effect size appearing in this registry, as the error will be replicated numerous times. It is necessary to have independent rater determine effect sizes for this registry and all discrepancies checked and resolved. The method employed to assure accuracy must be reported. Perusal of the table reveals some errors; for example, the description of the contrasts between groups does not always correspond to the numbers of the treatment.
Now there is a more formidable problem that has to do with trials with multiple measures of depression. The authors report that they calculate and reported the “mean” effect size. (It is unfortunate that they refer to other papers for how this was done—because this is the manuscript that reports a description of the registry, it is absolutely necessary to report details in this report). The essential problem is multiple measures of depression are highly correlated and any aggregation of effects within studies must account for that dependency. The issue is one primarily of standard error of the estimate of the effect size—multiple measures must not be treated as independent sources of information (determining a standard error that assumes that they uncorrelated) or that they are perfectly correlated. The latter is the case when the standard error is determined solely on the bases of the sample size and the mean of the various depression measures. (which is the case here). The supplemental materials do not report standard errors, so users will need to calculate that—and here the only way to do this is to base it sample size and estimates of the effect, thereby assuming that the dependent measures within studies are perfectly correlated (i.e., using multiple measures does not reduce the standard error). The issue of multiple measures should be addressed and described in the manuscript.

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

There is a confusion in table. Effects sizes are reported for treatment versus comparison and then for other comparisons. It is not always clear what group represents the control. Superior would be to have one column and report effects for all pairwise comparisons.

Discretionary Revisions (which the author can choose to ignore)

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

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

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