

Reviewer's report

Title: Inaccurate rounding of statistics in medical papers

Version: 1 **Date:** 4 February 2004

Reviewer: B D D McCullough

Reviewer's report:

General

The authors have identified and documented a serious problem.

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)

I cannot refer to page numbers because the pages are not numbered.
The authors will have to search for the corrections.

The authors have some difficulties with English, which made it difficult for me to understand what they meant. I have endeavored to make clear their intent in the following changes.

Abstract: ``One or more of the errors" ==> ``At least one such error"

Background: ``no literature proving" ==> ``no literature verifying"

Methods: ``accuracy of figures consisting of" ==> ``accuracy of results consisting of"

``We cannot check figures consisting" ==> ``We cannot check results consisting"

``We checked all the statistical figures (consisting" ==> ``We checked all the statistical results (consisting"

``in the additional files" ==> ``in the supplementary files"

``We only considered an error" ==> ``We only determined that a result was in error"

``For instance, the result of" ==> ``For instance, the p-value reported in"

``suggesting a transcription error" ==> ``suggesting a reporting error"

Results and discussion: "One of more of these errors" ==> "At least one error"

Additionally, there is one factual error and one statistical error to be fixed.

The factual error is when they write "there is apparently no literature...." but there is such a literature. See, for example, the classic article "On Rounding Percentage's" by Diaconis and Freedman, {it Journal of the American Statistical Association}, { f 74}(366) Jun 1979, pp. 359-364. You should read this article and see if it gives you ideas to make your paper better. At least cite the article.

The statistical error has to do with the application of Benford's Law. I suspect that your application of Benford's Law is wrong. The first digit is not uniform. The second digit is not uniform. I think even the third digit is not uniform. Uniformity is not expected until at least the fourth digit. Now, if all the test statistics / p-values that you analyzed are only reported to two or three digits, then the last digit should not be uniform. You need to check this. If I am right, then you should delete your sentences about Benford's Law.

However, you can replace it with a nice test. If rounding is done correctly, then the last digit of the 610 test statistics and p-values should be 50% even and 50% odd. Check this.

Discretionary Revisions (which the author can choose to ignore)

Results and discussion: "These kinds of errors are probably relatively unimportant.." If this is true, then why should this paper be published? These kinds of errors are indicative of gross carelessness. A researcher who cannot be trusted to do something simple, like round correctly, cannot be trusted to do more complicated things, like design an experiment. You should make this point strongly!

Recommendations: In number three, the researchers should make their data AND code freely available as proof that they have no errors to hide. Data alone are not sufficient to replicate results, as many details are hidden in the code. This has been shown in the economics literature by

Dewald, William G., Jerry Thursby and Richard G. Anderson. (1986) "Replication in Empirical Economics: The {it Journal of Money, Credit and Banking} Project," {it American Economic Review}, 76, 587-603

and

McCullough, B. D., Kerry Anne McGeary and Teresa D. Harrison (2004), "Lessons from the JMCB Archive," {it Journal of Money, Credit and

Banking} (to appear)

What next?: Accept after minor essential revisions

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: No

Declaration of competing interests:

None.