

Reviewer's report

Title: Reliability of journal impact factor rankings

Version: 1 **Date:** 5 June 2007

Reviewer: Eleanor Pullenayegum

Reviewer's report:

“Reliability of journal impact factor rankings” by Darren C Greenwood – Statistical Review

This is an interesting application of previous methods for obtaining confidence intervals around ranks to the ranking of impact factors.

Minor Essential Revisions.

1. The impact factor as defined (the number of citations a journal has received in the last complete year for articles published in the two preceding years, divided by the total number of articles the journal published in the two preceding years) is not measured subject to sampling error and hence known rather than estimated. There is thus no need to put a measure of uncertainty around it. The underlying rate at which any particular article in the journal is cited is, however, unknown and hence estimated subject to sampling error. Perhaps the wording throughout the manuscript could be tightened up to reflect this?
2. The proposed model treats the observed number of citations as Poisson with mean $l_i n_i$, and $\log(l_i) = l_i + n_i$. Am I correct in thinking that the total number of articles n_i is treated as fixed? Would the results change at all if the number of articles was treated as Poisson (also with random effects to model between-journal variation)? In addition, adopting a random effect for l_i models between-journal variability, but what about within-journal variability (i.e. some articles are more likely to be cited than others)? Is this accounted for in the model?
3. Under the proposed model, the estimated impact factors (i.e. l_i) are presumably not the same as the observed impact factors (O_i / n_i) because the introduction of the random effects results in some shrinkage? Could this be clarified in the text and in Figure 1?
4. (Last paragraph of Methods) Note that if $\log(l_i) = l_i + n_i$, then l_i is not the mean citation rate, because a log link has been used.
5. (Second paragraph of Results) To what extent are the observed results a product of the model (i.e. lognormal random effects)? Equivalently, how sensitive are the results to the model assumptions? How well does the model fit the data?
6. Figure 1: Is Figure 1 showing the observed impact factors, the mean values of l_i or some other summary statistic?
7. Figure 2: Similarly, is Figure 2 showing the mean ranks, the median ranks or the observed ranks?

Discretionary Revisions

8. Since you are using a Bayesian approach, perhaps you could identify this in the methods section of the abstract? (Using MCMC methods does not necessarily imply a Bayesian approach).

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: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

'I declare that I have no competing interests'