Irreproducible results: a response to Thomson Scientific

Mike Rossner,1 Heather Van Epps,2 and Emma Hill3

1Executive Director, The Rockefeller University Press
2Executive Editor, The Journal of Experimental Medicine
3Executive Editor, The Journal of Cell Biology

Thomson Scientific has posted a response (1) to our editorial on the reliability of their impact factor data (2). In it, they claim that our interpretation of the communication between our office and their Research Services Group was “misleading and inaccurate”. We have already published some excerpts from these communications in our previous editorial. For propriety’s sake, however, we have refrained from publishing internal Thomson Scientific e-mails, sent to us accidentally, which substantiate our claim that they could not provide us with the original data underlying the published 2006 impact factor calculations.

Although Thomson Scientific’s assertion that they do not have two separate databases may be correct, it is clear from their response that different groups within the corporation apply different filters to the data in their database, one of which removes erroneous records. Why this filter is not used for the published impact factors is still unclear.

Impact factors are determined from a dataset produced by searching the Thomson Scientific database using specific parameters. As previously stated, our aim was to purchase that dataset for a few journals. Even if those results were for some reason not stored by Thomson Scientific, it is inconceivable to us that they cannot run the same search over the same database to produce the same dataset. The citation data for a given year should be static. In essence, Thomson Scientific is saying that they cannot repeat the experiment, which would be grounds for rejection of a manuscript submitted to any scientific journal.

Continued on next page
Thomson Scientific argues that we did not inform them of the methodologies we would apply to the data when we purchased it. This is like asking someone who is buying a dictionary what words they intend to look up. In fact, our methodology was the same as theirs: a simple addition of the citation numbers divided by the number of citable articles.

We will not refute other points made by Thomson Scientific in their rebuttal, as others have already done so to some extent (see box). Instead we close this discussion with a plea to our fellow publishers to make their citation data available in a publicly accessible database, and thus free this important information from Thomson Scientific’s (and other companies’) proprietary stranglehold.

References
1. Pendlebury, D.A. 2007. Article Titled “Show me the Data”, Journal of Cell Biology, Vol. 179, No. 6, 1091–1092, 17 December 2007 (doi: 10.1083/jcb.200711140) is Misleading and Inaccurate. http://scientific.thomson.com/citationimpactforum/8427045/ (accessed January 4, 2008).

2. Rossner, M., E. Hill, and H. Van Epps. 2007. Show me the data. J. Cell Biol. 179: 1091–1092. doi:10.1083/jcb.200711140.