

High impact factors are meant to represent strong citation rates, but these journal impact factors are more effective at predicting a paper's retraction rate.

by Blog Admin

December 19, 2011

Journal ranking schemes may seem useful, but Björn Brembs discusses how the Thompson Reuters Impact Factor appears to be a reliable predictor of the number of retractions, rather than citations a given paper will receive. Should academics think twice about the benefits of publishing in a 'high impact' journal?

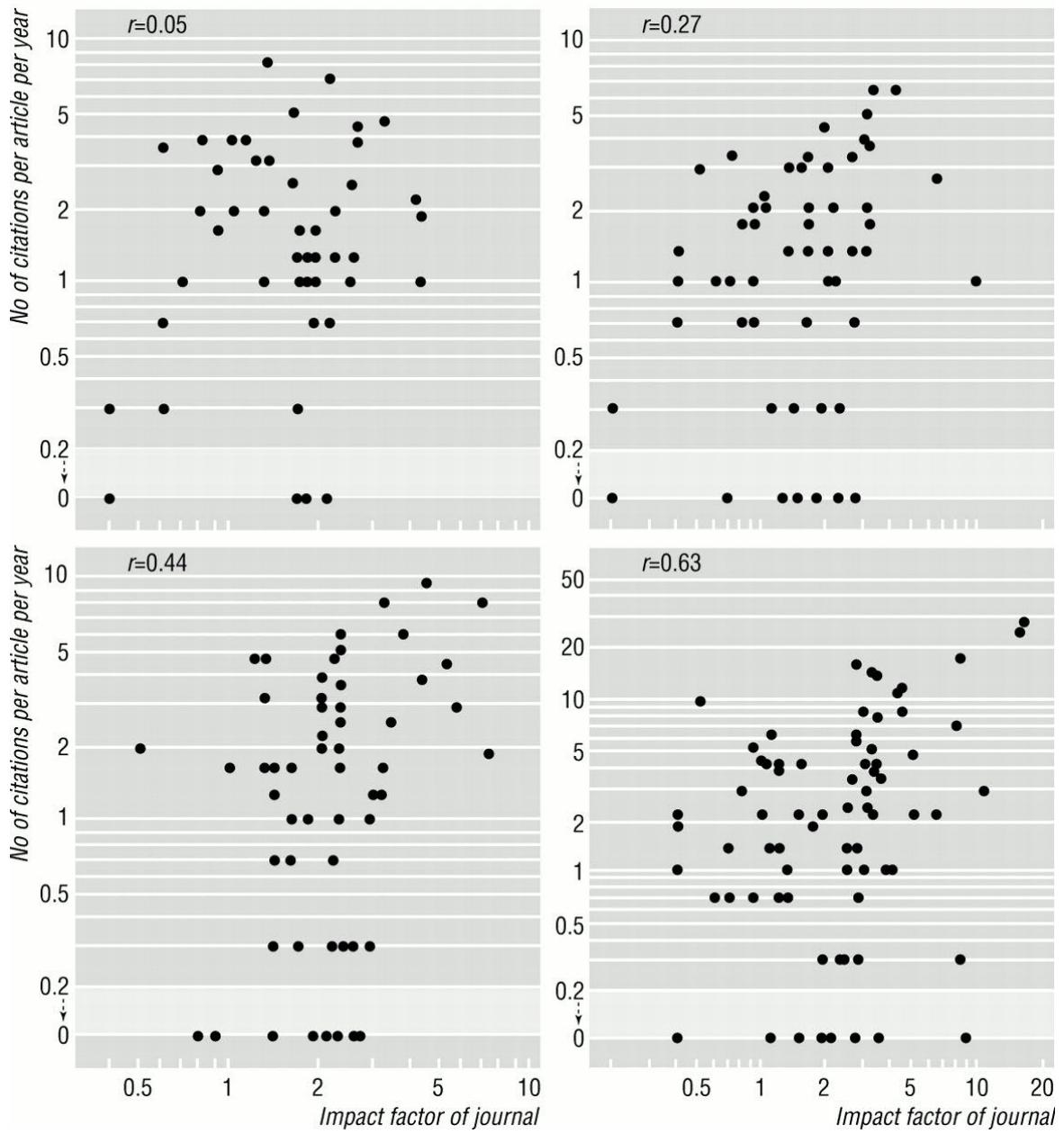
With more than 24,000 scholarly journals in which some piece of relevant research may be published, a ranking scheme seems like a boon: one only needs to read articles from a small, high-ranking subset of journals and safely disregard the [low-level chaff](#). At least this is how one might describe the development of journal ranks in the 1960s and 70s, when scores of new journals began to proliferate.



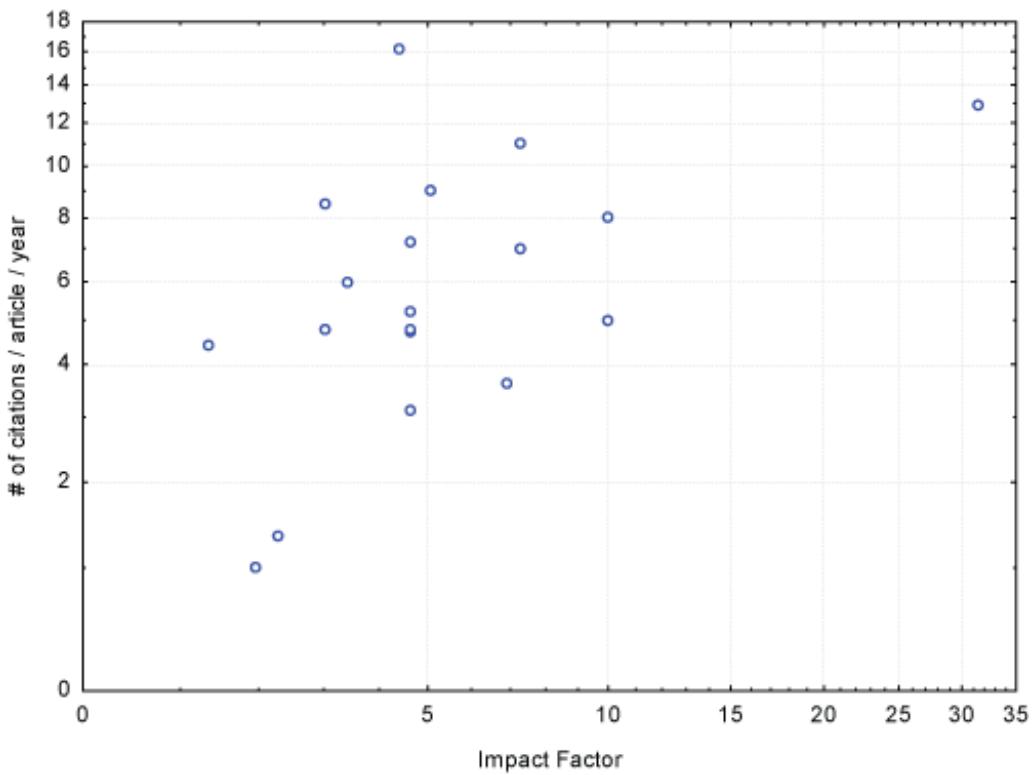
Today, however, journal rank is used for much more than just filtering the paper deluge. Among the half-dozen or so ranking schemes, one *de facto* monopolist has emerged which dictates journal rank: [Thomson Reuters' Impact Factor](#) (IF). At many scientific institutions, funders and governing bodies are using the IF to rank the content of the journals as well: if it has been published in a high-ranking journal, it must be good science, or so the seemingly plausible argument goes. Thus, today, scientific careers are made and broken by the editors at high-ranking journals.

As a scientist today, it is very difficult to find employment if you cannot sport publications in high-ranking journals. In the increasing competition for the coveted spots, it is starting to be difficult to find employment with only few papers in high-ranking journals: a consistent record of 'high-impact' publications is required if you want science to be able to put food on your table. Subjective impressions appear to support this intuitive notion: isn't a lot of great research published in *Science* and *Nature* while we so often find horrible work published in little-known journals? Isn't it a good thing that in times of shrinking budgets we only allow the very best scientists to continue spending taxpayer funds?

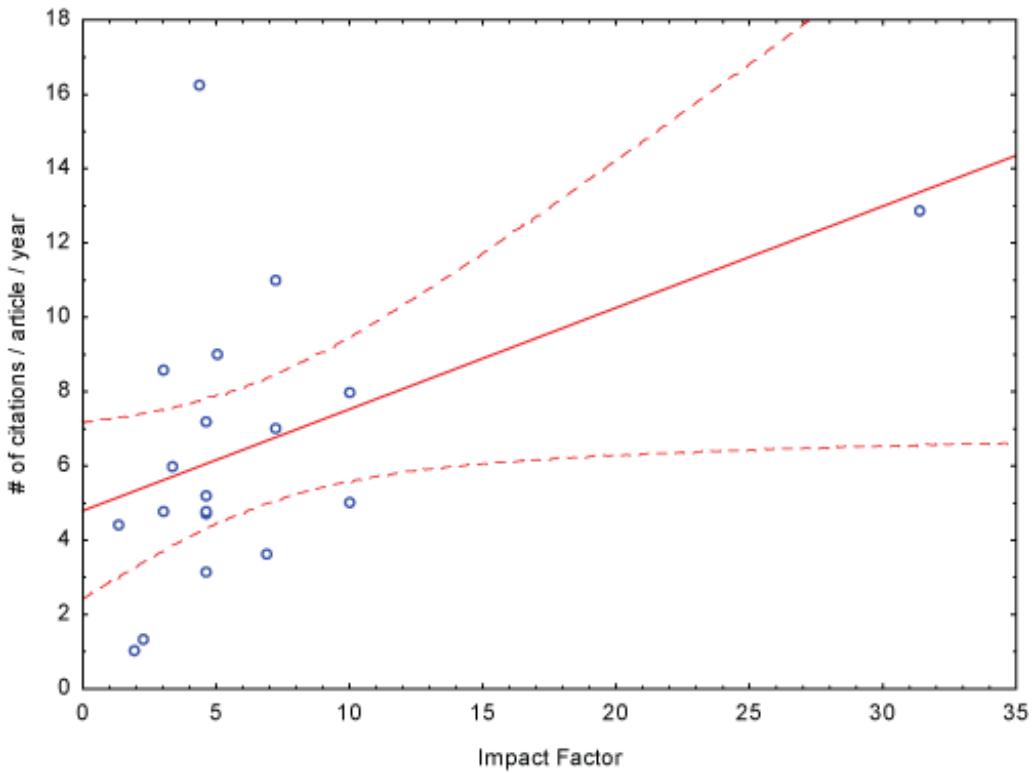
In any area of science, intuition, plausibility and subjective judgment may be grounds for designing experiments, but scientists all the time have to change their subjective judgment or discard their pet hypothesis if the data don't support them. Data, not subjective judgment is the basis for sound science. In this vein, many tests of the predictive power of journal rank have been carried out. One of them was published in 1997 in the [British Medical Journal](#). One figure in this paper shows four examples of researchers whose publications had been plotted with their annual citations against the IF:



As can be seen from the R values at the top left of each graph, the correlation between impact factor and actual citations is not all that great. However, these are only four examples. Maybe this would be different for other researchers? In the absence of any easily available dataset where IFs and citations of individuals are compiled, I just took my own publications (according to [Google Citations](#)), looked up the current IFs and plotted them in the same way:



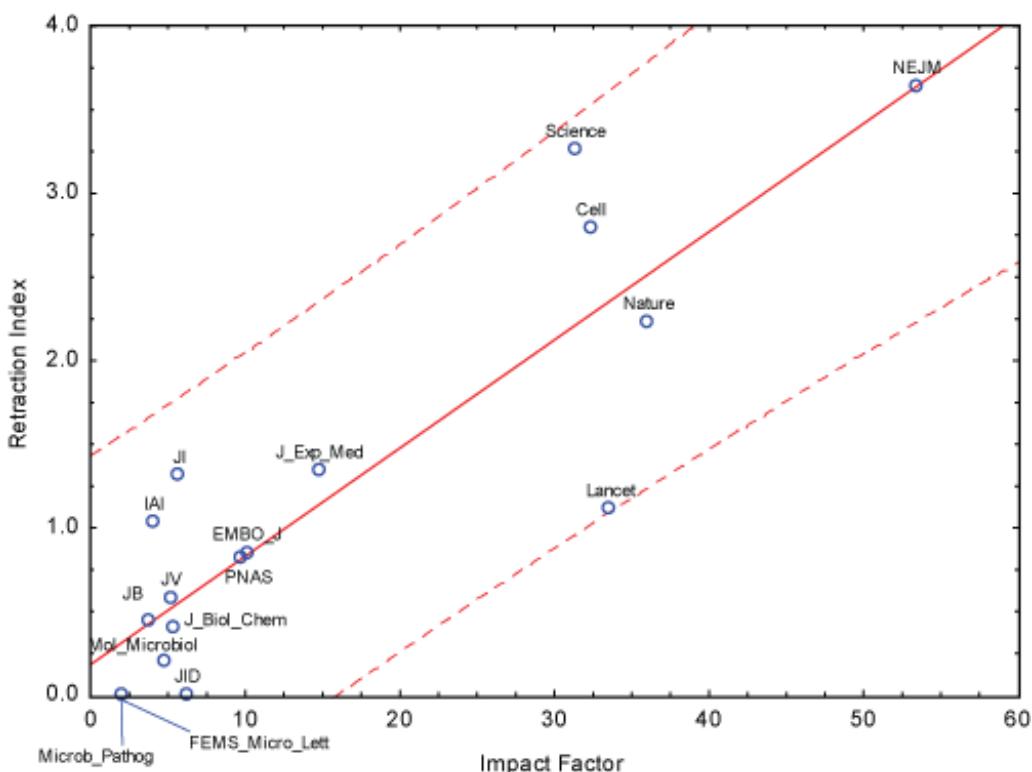
The R value for this correlation is 0.55, so pretty much in the range of the published values. I have also done a linear regression on these data, which provides me with a somewhat more meaningful metric, the “R squared” value, or the [Coefficient of Determination](#). For my data, the adjusted value for this coefficient is less than 0.3, a very weak measure of a correlation, suggesting that the predictive power between IF and citations, at least for my publications, is not very strong, despite it being statistically significant ($p<0.004$). Here is a linear plot of the same data as above:



Our *Science* paper stands out as an outlier on the far right and often such outliers tend to artificially skew regressions. Confirming the interpretations so far, removing the *Science* paper from the analysis renders the regression non-significant (adjusted R²: 0.016, p=0.275). Thus, with the available data (to me) so far, there seems to be little reason to expect highly-cited research in high-ranking journals. In fact, our most frequently cited paper is smack in the middle of the IF scale.

More recently, there was [another publication](#) assessing the predictive power of journal rank. This time, the authors built on the notion that the pressure to publish in high-ranking journals. If your livelihood depends on this *Science/Nature* paper, doesn't the pressure increase to maybe forget this one crucial control experiment, or leave out some data points that don't quite make the story look so nice? After all, you know your results are solid, it's only cosmetics which are required to make it a top-notch

publication! Of course, in science there never is certainty, so such behavior will decrease the reliability of the scientific reports being published. And indeed, [together with the decrease in tenured positions, the number of retractions has increased](#) at about 400-fold the rate of publication increase. The authors of this study, Fang and Casadevall, were so nice to provide me with access to their data so I could compile the same kind of regression analysis I did for my own publications:



This already looks like a much stronger correlation than the one between IF and citations. How do the critical values measure up? The regression is highly significant at $p<0.000003$, with a coefficient of determination at a whopping 0.77. Thus, at least with the current data, IF indeed seems to be a more reliable predictor of retractions than of actual citations. How can this be, given that the IF is supposed to be a measure of citation rate for each journal? There are many reasons why this argument falls flat, but here are the three most egregious ones:

- The IF is negotiable and doesn't reflect actual citation counts ([source](#))
- The IF cannot be reproduced, even if it reflected actual citations ([source](#))
- The IF is not statistically sound, even if it were reproducible and reflected actual citations ([source](#))

In other words, all the organizations that require scientists to publish in 'high-impact journals' at the same time require them to publish in 'high-retraction journals'. I wonder if requiring publication in high-retraction journals can be good for science?

This post was originally posted on Bjoern Brembs' [blog](#).

Related posts:

1. [Academics shouldn't be afraid that their work may not be being cited as much as they would like: citation rates vary widely across disciplines](#)
2. [Google Scholar Citations: a way for academics to compute citation metrics and track them over time](#)
3. [Impact is a strong weapon for making an evidence-based case for enhanced research support but a state-of-the-art approach to measurement is needed](#)
4. [Universities need proactive and imaginative strategies to communicate their research and to achieve high impact scores](#)
5. [There is a pathetic lack of functionality in scholarly publishing. We must end for-profit publishing and allow libraries to make available the works of their scholars for all](#)