

- Liu SV (2006g) Why are “top” journals often so unlucky in inflicting top scandals? *Sci Ethics* **1**: 49–52
- Liu SV (2006h) Citation should credit pioneering researchers and original works. *Sci Ethics* **1**: 94–96
- Liu SV (2006i) Lessons not learnt or not revealed. *Sci Ethics* **1**: 24–27
- Liu SV (2007a) Hwang’s retracted publication still contributes to *Science*’s impact factor. *Sci Ethics* **2**: 44–45
- Liu SV (2007b) Nilsson and *Science* owe public more answers than a simple retraction. *Sci Ethics* **2**: 41–43
- Liu SV (2007c) PNAS: A permanent niche for aggressive sin. *Sci Ethics* **2**: 7–20
- Liu SV (2007d) Respecting published answers for important questions on epigenetics. *Top Watch* **2**: 19–20
- Liu SV (2007e) Logical review casts doubts on several publications in top journals. *Top Watch* **2**: 23–24
- Liu SV (2007f) Liu’s message blocked by ‘top’ journals. *Sci Ethics* **2**: 5–6
- Liu SV (2007g) The misconduct of ‘top’ journals in exposing scientific misconduct. *Sci Ethics* **1**: 17–19
- Liu SV (2007h) *Nature Genetics*: a natural guard for flawed publication? *Sci Ethics* **2**: 21–35

Shi V. Liu is Editor of *Scientific Ethics*

(<http://im1.biz>).

E-mail: [svl8se@im1.biz](mailto:svl8se@im1.biz)

doi:10.1038/sj.embor.7401053

## Response by Cokol *et al*

In our recent article in *EMBO reports* (Cokol *et al*, 2007), we described a mathematical model and large-scale data analysis to investigate the hidden factors underlying article retraction in scientific journals. Our study showed that retracted articles tend to appear more frequently in journals with a higher impact factor (IF). We proposed two possible explanations: either high-IF journals publish flawed manuscripts at a higher rate, or articles appearing in these journals are subjected to a higher level of post-publication scrutiny. As the latter provided a better explanation of our model and the data at our disposal, we concluded that low- and high-IF journals publish flawed papers at similar rates.

In his commentary on our study (Liu, 2007), Shi V. Liu makes various assertions that we would like to address. First, that we assume “...low-IF journals probably receive less attention and are therefore likely to contain more articles that ought to be retracted, but instead go unnoticed.” He has mistaken our conclusion for an assumption. We used our model and analysis to conclude that low-IF journals contain more

hidden retractable papers than high-IF journals. This is a testable prediction based on data and not an assumption.

Second, Liu states that we think “...high IF means high quality.” Nowhere in our article do we state or use the assumption that IF is a measure of quality.

Third, Liu states that “...more publications should be retracted, but [...] most of these retractions would still come from high-IF journals”. This is a legitimate hypothesis that can be tested, but one that is not supported by modelling or the analysis of large data sets.

Fourth, he maintains that “‘top’ journals often offer much stronger protection for their publications” than other journals. ‘Top’ journals might indeed offer some protection for their publications, but this protection is, quite obviously, limited to rejecting criticism of prior work submitted to the same journal. Suggesting that *Nature* would suppress criticism of a paper published in *Science* would imply a conspiracy theory. Furthermore, stronger protection of highly visible criticized papers does not mean that they will fail less often, as it is expected that they will be attacked more often.

Fifth, Liu comments that, “[t]he reason for the higher incidence of ‘error-prone’ papers published by these [high-IF] journals is not due to high post-publication scrutiny, but rather to lower standards of the pre-publication review”. Again, this is a hypothesis that can be tested, but we suggest that the verity of the alternative hypotheses be tested by the analysis of real data rather than asserted *a priori*.

Sixth, Liu says, “[a]s a scientist I have to give the same scrutiny to any publication, regardless of where it is published. It is therefore wrong to conclude that low-IF journals have not received adequate scrutiny and thus not retracted enough articles”. We do respect Dr Liu’s scientific integrity; however, the rigour of post-publication scrutiny depends not only on the honesty and integrity of people who read the article, but also on the amount of time devoted to it and the total number of people with diverse backgrounds who read the text. It is rather hard to argue that the high-IF journals are not read by a larger number of people than the low-IF ones.

Seventh, Liu comments that, “retractions are [...] not stochastic events”. Stochastic modelling is common in science; it is applicable to the description of a wide range of real-life systems, from highly ordered/deterministic to chaotic/random. Therefore, one

should not confuse the probabilistic formalism that we use to describe the retraction process with the statement that retraction is completely stochastic.

Eighth, he argues, “[e]ven for a general journal, the citation and retraction rates vary greatly between disciplines. [...] Thus, the argument for a positive correlation between the IF value of a journal and its post-publication scrutiny level does not reflect the reality”. Scientific fields tend to vary significantly in terms of the community size. It is therefore not surprising that smaller communities produce fewer papers and have smaller average IFs. We fail to see how this invalidates our argument about the inferred correlation between IF and post-publication scrutiny.

Ninth, Liu concludes that, “[i]t is therefore not valid to assume that there should be a uniform retraction rate among disciplines and journals”. Again, he mistakes our conclusion for an assumption.

Finally, Liu criticizes that we used “flawed impact factors” that are “based—at least in part—on the high number of citations of their [high-IF journals] retracted papers.” IF is a well-defined quantity and although its interpretation can be logically flawed, a consistently computed IF itself cannot. IFs are cumulative indicators of the state of a scientific reference network, not unlike the temperature of a physical system. If we agree that to cite a paper, one needs to look at it and/or read it, it is not unreasonable to assume that journal-specific IF values are correlated with the number of scientists exposed to each article published in the journal. As even retracted papers need to be seen and/or read to be cited, IF is a perfectly appropriate measure of post-publication scrutiny for our analysis.

We would like to emphasize that our conclusions were derived through the probabilistic analysis of 9 million articles. Although our results are open to criticism, a fair criticism should include suggestions for alternative data analysis and interpretation, rather than just the claim that a competing hypothesis is correct.

### REFERENCES

- Cokol M, Iossifov I, Rodriguez-Esteban R, Rzhetsky A (2007) How many scientific papers should be retracted? *EMBO Rep* **8**: 422–423
- Liu SV (2007) Comment on the Correspondence by Cokol *et al*. *EMBO Rep* **8**: 792–793

**Murat Cokol<sup>†</sup>, Ivan Iossifov, Raul Rodriguez-Esteban and Andrey Rzhetsky are at Columbia University in New York, USA.**

<sup>†</sup>E-mail: [murat.cokol@dbmi.columbia.edu](mailto:murat.cokol@dbmi.columbia.edu)

doi:10.1038/sj.embor.7401054