correspondence

Mobility is not the only way forward

In European research today, mobility has become firmly entrenched as a prerequisite for a successful scientific career; however, the reasons behind this emphasis on scientific nomadism are far from obvious. It is all too common for science policy-makers and funding bodies to extol the seemingly endless virtues of mobility, but it nearly amounts to heresy to ponder the possibility that a move abroad might also have drawbacks. It was therefore a breath of fresh air—to those of us who did not change countries the last time we changed jobs—to read Frank Gannon's recent editorial (Gannon F (2007) The downsides of mobility. *EMBO Rep* **8:** 201).

With a healthy dose of realism, Gannon lays out most of the disadvantages of moving, which, unfortunately, are as obvious as they are neglected. These range from losing out in pension schemes and property markets to wasting time and energy on readjusting to foreign administrative systems and social environments. There are still more factors that should dampen the current unrestrained enthusiasm for mobility: possibly serious language-related difficulties for researchers and their families, loss of support from relatives in raising one's children and the substantial costs of moving to a foreign country, which are often not reimbursed. Together, these hindrances can be dramatically vexing to the travelling scientist, to the point of seriously impeding his or her research productivity and career development—the officially proclaimed *raison d'être* for mobility.

All this would be acceptable if researchers were free to choose whether or not to move. However, this choice is not really free, as mobility is typically either a major advantage or a crucial requirement for competing for scarce funding in science. Therefore, a brilliant young scientist moving from, say, a PhD position in Edinburgh to a postdoctoral position in Cambridge could lose out against an average researcher relocating from Belgium to Luxembourg, simply because he or she will not be eligible for some of the most prestigious fellowships in science.

This perplexingly automatic promotion of mobility has left some researchers feeling as if the system were designed to minimize the opportunity for stable family and social ties, to the point where the best choice left to scientists is to devote themselves exclusively to their research. The real motives are more likely to be political—linked to the fostering of a cross-border European research culture—which might look good on the glossy pages of public relations brochures but does not always work well in or outside the laboratory. Perhaps, as Gannon suggests, mobility has simply become a fashionable buzzword in scientific policy circles, all at the expense of researchers.

Naturally, none of the above is meant to suggest that mobility is harmful or useless. There is no doubt that many scientific careers have been, and will continue to be, greatly boosted by moves to different research environments in new countries. What is at issue, however, is the mechanical endorsement of mobility as the only acceptable way forward for researchers, without any rational justification for why scientific progress is impossible without changing country.

The good news is that we are starting to see some steps toward displacing the mobility prerequisite. For example, the European Molecular Biology Organization (EMBO) has recently changed the rules for its fellowships so as to soften the requirement for mobility, and the newly created European Research Council (ERC) Starting Grants explicitly set out 'scientific excellence' as the sole selection criterion. That good science should be the only factor in distributing research funds sounds perfectly logical, but, alas, not perfectly typical in the mobility-obsessed world of scientific funding bodies. We can only hope that once others follow the positive examples of EMBO and the ERC, the old mobility reflexes of selection panels will dissipate. Let us hope, too, that they will not be replaced by an insurmountable bias for 'interdisciplinarity', 'translational research' or whatever happens to be the trendiest fad of tomorrow.

Boyan K. Garvalov is at the Max Planck Institute of Neurobiology, Martinsried, Germany.
E-mail: garvalov@neuro.mpg.de

How many scientific papers should be retracted?

Published scholarly articles commonly contain imperfections: punctuation errors, imprecise wording and occasionally more substantial flaws in scientific methodology, such as mistakes in experimental design, execution errors and even misconduct (Martinson *et al*, 2005). These imperfections are similar to manufacturing defects in man-made machines: most are not dangerous but a small minority have the potential to cause a disaster (Wohn & Normile, 2006; Stewart & Feder, 1987). Retracting a published scientific article is the academic counterpart of recalling a flawed industrial product (Budd *et al*, 1998).

However, not all articles that should be retracted are retracted. This is because the quality of a scientific article depends, among other things, on the effort and time invested in quality control. Mechanical micro-fractures in turbojet components are detected more readily than those in sculptures, as airplane parts are typically subjected to much more rigorous testing for mechanical integrity. Similarly, articles published in more prominent scientific journals receive increased attention and a concomitant increase in the level of scrutiny. This therefore raises the question of how many articles would have to be retracted if the highest standards of screening were universally applied to all journals.

PubMed provides us with a 'paleontological' record of articles published in 4,348 journals with a known impact factor (IF). Of

the 9,398,715 articles published between 1950 and 2004, 596 were retracted. This wave of retraction hits high-impact journals significantly harder than lower-impact journals (Fig 1A), suggesting that high-impact journals are either more prone to publishing flawed manuscripts or scrutinized much more rigorously than low-impact journals.

Here, we introduce a four-parameter stochastic model of the publication process (Fig 1C), which allows us to investigate both possibilities (see Supplementary Information for full details). The model describes the two sides of the publication–retraction process: the rigour of a journal in accepting a smaller fraction of flawed manuscripts (quality parameters α and θ), and post-publication scrutiny on the part of the scientific community (scrutiny parameters β and τ). We need four parameters to account for the possibility of both IF-independent and IF-dependent changes in quality and scrutiny (see Fig 1C, top). For each journal we compute a normalized impact r (a journal-specific IF divided by the maximum IF value in our collection of journals), and use it to define the strength of IF-dependent quality (r^α) and IF-dependent scrutiny (r^β).

This probabilistic description allows us to distinguish between the two scenarios described above for the higher incidence of retractions in high-impact journals. For all values of τ tested, we find that the posterior mode of the IF-dependent quality parameter (α) is close to 0 (Fig 1B), indicating a nearly uniform and IF-independent rigour in pre-publication quality control. Conversely, the posterior mode of the IF-dependent scrutiny parameter (β) is essentially 1 (Fig 1B), which translates into a linear dependence between the IF of the journal and the rigour of post-publication scrutiny. Our data therefore suggest that high-impact journals are similar to their lower-impact peers in pre-publication scrutiny, but are much more meticulously tested after publication.

Our model also allows us to estimate the number of papers that should have been retracted under all plausible explanations of reality permitted by the model (Fig 1C). We estimate the number of articles published between 1950 and 2004 that ought to be retracted to be more than 100,000 under the more pessimistic scenario ($\tau = 0.1$; red, Fig 1C), and greater than 10,000 under the most optimistic scenario ($\tau = 1$; green, Fig 1C). The gap between retractable and retracted papers is much wider for the lower-impact journals (Fig 1C). For example, for *Nature* (1999–2004 average IF = 29.5), with

Fig 1 | Dataset, model and estimation of the number of flawed articles in scientific literature. (**A**) Higher- and lower-impact factor (IF) journals have significantly more than expected and less than expected retracted articles, respectively. Each sphere represents the set of articles within the same IF range; the volume of the sphere is proportional to the set size and its colour represents the middle-of-the-bin IF value. The expected number of retractions is calculated under the assumption that all retractions are uniformly distributed among articles and journals. The red line indicates a hypothetical ideal correlation between the observed and the expected numbers of retractions. (**B**) and (**C**) explain the four-parameter graphical model describing our hypothetical stochastic publication–retraction process. (**B**) Estimated posterior distribution of parameter values for several values of impact-independent scrutiny. (**C**) Outline of the stochastic graphical model (top) and the posterior mean estimates of the number of articles that should be retracted (with 95% credible interval) plotted against different values of IF. Posterior distributions of estimated number of retractable articles: red- and green-coloured distributions correspond to $\tau = 0.1$ and $\tau = 1$, respectively; horizontal black solid lines indicate the actual number of retracted articles for individual IF bins and journals. The contour distributions represent individual journals, whereas the solid distributions correspond to the whole PubMed corpus binned by the IF value.

$\tau = 1$ (optimistic), we estimate that 45–67 articles should have been retracted, whereas only 30 actually were retracted. *Science* (IF = 26.7) and *Biochemistry* (IF = 4.1) have nearly identical numbers of papers published, but the actual numbers of retracted papers for these two journals are 45 and 5 respectively.

Our analysis indicates that although high-impact journals tend to have fewer undetected flawed articles than their lower-impact peers, even the most vigilant journals potentially host papers that should be retracted. However, the positive relationship between visibility of research and post-publication scrutiny suggests that the technical and sociological progress in information dissemination—the internet, omnipresent electronic publishing and the open access initiative—inadvertently improves the self-correction of science by making scientific publications more visible and accessible.

Supplementary information is available at EMBO *reports* online.
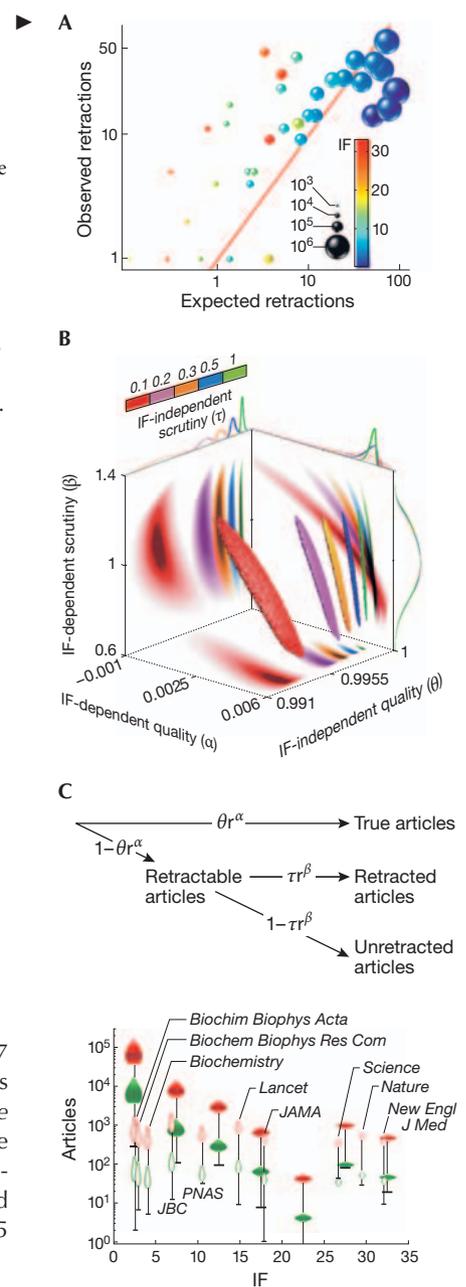
REFERENCES

Budd JM, Sievert M, Schultz TR (1998) Phenomena of retraction: reasons for retraction and citations to the publications. *JAMA* **280**: 296–297

Martinson BC, Anderson MS, de Vries R (2005) Scientists behaving badly. *Nature* **435**: 737–738

Stewart WW, Feder N (1987) The integrity of the scientific literature. *Nature* **325**: 207–214

Wohn DY, Normile D (2006) Korean cloning scandal. *Science* **312**: 980–981

Murat Cokol, Ivan Iossifov, Raul Rodriguez-Esteban & Andrey Rzhetsky are at Columbia University in New York, NY, USA.
E-mail: murat.cokol@dbmi.columbia.edu