

Appreciating data: warts, wrinkles and all

In the glitzy world of Hollywood and Bollywood, each year sees the development of more extravagant digital special effects. Many productions have long since broken the constraints imposed by physics and biology and although the superhuman feats of modern action heroes rapidly become tedious, infinitely worse is the frequent subjugation of storylines to digital effects. Scientific publishing has not escaped the influence of image-processing technologies.

A whole new generation of scientists has known nothing but the magical world of Photoshop and regularly use tools with fashionable names such as 'clone' and 'healing'. Simultaneously, science output has grown more rapidly than either the availability of funding or places to publish — in fact, rising pressures to publish in prestigious journals has been driven by an increasingly competitive spirit among life science researchers encouraged by 'impact-factor-hungry' funding agencies. Add to these the accelerated speed of scientific discovery and one is left with a recipe for the frenetic research environment in many biological specialities. As discussed in our February editorial (*Nature Cell Biology*, **8**, 101 (2006)), some scientists succumb to unethical tactics, such as outright fraud, plagiarism or the misuse of confidential information. This is rare; more prevalent is sloppy experimentation or misrepresentation of data (*Nature* **435**, 737–738 (2005)). This can happen either at the level of data acquisition (data selection or questionable statistics) or, aided by easy to use digital toolboxes, 'data beautification' (image manipulation of properly acquired data to enhance, clarify or conform to the expected). A particularly egregious case in point (*Nature Cell Biol.* **5**, 320–329 (2003)) prompted us to investigate, in consultation with the community and other journals interested in the topic, what more could and should be done by publishers to discourage such practices.

As a first step, we have expanded our guidelines, based on community feedback, to explicitly delineate what information should be presented in a published paper. For now, we have focussed on two types of data that are subject to manipulation most often: gels and immunoblots, and light microscopy. Our new guidelines (http://www.nature.com/ncb/about/ed_policies/index.html) range from appropriate antibody controls to the type of metadata required for the proper interpretation of micrographs. We have also formalized our previous recommendations on best practice for image cropping and splicing (<http://www.nature.com/ncb/journal/v6/n4/full/ncb0404-275.html>).

There is a myth that editors only like clean data that show striking effects. What we actually like is solid data that provides striking conceptual advances. Effects may be small, but statistics and controls are needed to make them believable. Data should be clearly presented and concise, but not at the expense of important information. Let's celebrate real data — wrinkles, warts and all. We want to publish gritty documentary movies, not digitally beautified yarns!

We hope these guidelines will aid the publication of more informative datasets. Importantly, we reemphasize that neither the referees nor the editors are the data-police (see also *Nature Cell Biology*, **8**, 101 (2006)). Senior investigators and corresponding authors are responsible for assuring that data submitted for publication represents the experimental results accurately and fairly. We suggest that they are also responsible for ensuring that their students are educated in appropriate scientific conduct.

Nature Protocols

All *Nature* journals publish concise papers, without repetitive introductions and discussions but with just enough terse prose to contextualize the data and to allow clear presentation of the findings. The aim is to appeal to a broadly interested readership that does not have the time to wade through a dozen pages to extract the new findings. On the other hand, we aim to remain accessible to our diverse readership, which requires that data are framed by a sufficiently broad introduction and detailed discussion. At the same time, the data presented must be of the high quality and thoroughness rightly expected of a *Nature* journal, including exhaustive controls. Clearly, it is a challenge to successfully fulfil these divergent requirements and our editors work closely with authors to optimize each paper on a case-by-case basis. As discussed above, our new guidelines for data processing and presentation will maximize the information content and the scientific robustness of our papers.

A valid criticism remains: although our methods sections allow the accurate interpretation of the data, repeating experiments solely on the basis of the information provided can be cumbersome. Reproducibility is at the heart of the scientific validation process, therefore, a key requirement for publishing in a *Nature* journal is distribution of reagents and full protocols that allow independent validation of the data reported.

To facilitate this process, Nature Publishing Group launches *Nature Protocols* (www.natureprotocols.com) in mid March. The aim is to provide an online resource of a growing set of cutting-edge experimental protocols. Each protocol is described step-by-step in enough detail to allow repetition of the experiment. Following the full launch in June, the site will host both commissioned peer reviewed content, as well as freely accessible unvetted protocols and comments posted by the community. After mid March, we will be actively encouraging our authors to post protocols supporting the experiments published in this journal. These open access protocols will be linked with the online manuscript to provide a valuable extension to our papers. We hope that you will contribute; should you require further convincing, our editors or *Nature Protocols* will be happy to discuss.

We believe that our improved standards for data presentation and restrictions on image manipulations, and *Nature Protocols*, will increase the robustness, completeness and transparency of papers in *Nature Cell Biology*.