

The long road to reproducibility

In the mission to reduce irreproducibility, true change can only come about if all stakeholders — researchers, institutions, funders and journals — join together with common purpose.

In May 2013, recognizing the challenges inherent to increasing reproducibility and rigour in research, the Nature journals took steps to raise the standards of methodological reporting in life science papers. A key feature of this undertaking was the creation of a multi-point reporting checklist, which consolidated methodological details common to a number of disciplines in the life sciences that were identified as contributing to irreproducibility. The checklist was presaged by a NINDS (National Institute of Neurological Disorders and Stroke) workshop in 2012 on poor reporting standards in preclinical studies; the discussions from this workshop culminated in recommendations for increasing reporting standards for animal experiments (*Nature* **490**, 187–191; 2012). *Nature Cell Biology* has, for over a decade, encouraged data transparency in the form of mandated inclusion of source data for gels and blots, and mandates the deposition of certain kinds of datasets in community-endorsed repositories. In May 2013, we extended this policy to encourage the provision of source data underlying graphical data. We have also long since supported the deposition of step-by-step protocols in the open repository Protocol Exchange that can then be linked to the original *Nature Cell Biology* publication. Building on these initiatives, the rollout of the reporting checklist was accompanied by removing the length limits on the Methods section in our papers to allow for detailed methodological descriptions that would facilitate reproducibility.

Since the advent of these efforts, we have discovered that clear reporting on statistics remains singularly challenging in a number of fields represented in the pages of *Nature Cell Biology*. Among the most common issues we encounter is a lack of clarity regarding sample size and the nature of the samples used to derive statistics — whether they are biological or technical replicates, and whether samples are derived from single experiments or aggregated across multiple independent experiments. Aggregation of samples across multiple independent experiments to derive statistics, while likely to be frowned upon by statisticians, is common practise across many fields covered within the journal. We therefore request that authors describe the samples used to derive statistics and the scope of the analysis in detail, as the statistics are difficult to interpret without precise information regarding sample size. We strongly discourage deriving statistics from technical replicates (as it is not meaningful for understanding biological variation), unless it is of value to highlight the low variability within an assay. If sample size is low or if experiments have been replicated a small number of times, authors are encouraged to show the full spread of the data by plotting the individual data points or by providing all data points underlying

graphical representations in a ‘source data’ supplementary table accompanying the paper. We have seen a steady improvement in reporting practise, with greater clarity in reporting statistics and an increasing number of papers now plotting individual data points for low sample sizes, as well as providing source data. These changes have been achieved through a sustained and concerted effort by the journal editors, authors and reviewers, and ultimately resulted in what we hope are better papers for authors, readers and the community at large. However, for fundamental change in statistical practise to take root, universities, institutes and funders need to ensure that better training is available for researchers at all stages of career progression.

Reagents — antibodies, cell lines, chemical probes and oligonucleotides, to list but a few — are another central aspect of research design plagued by issues of poor validation and inadequate reporting. Poor antibody validation and cell line misidentification are increasingly coming to the fore as crucial elements underlying irreproducibility, as funders and standards organisations draw attention to these issues. Research with misidentified cells and poorly characterized antibodies continues to pollute the literature, although the degree to which this is a serious problem is certainly variable between fields. Aiming to tackle the issue of cell line identity, we have been strengthening reporting standards by asking authors of all papers published in Nature journals from May 2015 onwards to state whether the cell lines used in their study are found in the database of misidentified cell lines maintained by the International Cell Line Authentication Committee (ICLAC) or in the National Centre for Biotechnology Information (NCBI) BioSample database. Although authors of all papers published in Nature journals are encouraged to provide this information, it is currently mandatory for cancer research papers, owing to the acknowledged issues with cell line identity in this area (*Nature* **520**, 264; 2015). We have long asked authors to provide details of antibody source and validation, and we will be following the ongoing discussions on how best to address the deep-rooted issues of insufficient antibody characterization.

In 2014, the National Institutes of Health (NIH), together with Nature Publishing Group and *Science*, convened a workshop on reproducibility and rigour that put forth a set of recommendations to further reproducibility, robustness and transparency in research. Many of these guidelines — which include emphasis on increasing statistical rigour, transparency in reporting, availability of data and materials, best practise guidelines for presentation of digital data, and validation of reagents — are enshrined in the Nature journal reporting checklist. They have also been endorsed by numerous other journals, associations and professional societies. Notably, the recommendations also urge journals to adopt a reporting checklist enumerating aspects of reporting standards, as currently practised by the Nature journals. To our knowledge, very few cell and molecular biology journals have adopted checklists. Regardless of whether this is the most optimal solution by which journals can enhance reporting standards, it is clear that a multipronged effort is needed from all quarters of the research community. Encouragingly, academic societies and research

institutes, which are engines for promoting awareness and driving a change in practise within the research communities, are beginning to engage with the issues of reproducibility.

The challenges presented by irreproducibility are vast and seemingly intractable, seeing as they are so entrenched in research practises. The Nature journals have covered the topic of irreproducibility from a variety of perspectives (<http://www.nature.com/news/reproducibility-1.17552>). Change will only come about through the concerted effort and engagement of all relevant stakeholders: institutions raising awareness and training students, post-doctoral researchers and principal investigators in best practise, as well as funders and journals implementing efforts to raise reporting standards.

Farewell to our chief

This year's last issue of *Nature Cell Biology* is also the last for Sowmya Swaminathan after nearly 13 years at the journal and 6 years as its chief editor. Under Sowmya's able leadership, the journal has evolved

both in terms of content and editorial process. Being receptive to the concerns of the scientific community, she has implemented procedures to improve both the author and reader experience. Among these is the limit we now set on supplementary material published with our papers. She has also taken a strong role in facilitating the transfer of manuscripts between the Nature research journals, in particular to *Nature Communications*, and more recently in developing ways to provide more specific editorial guidance to authors when invited to revise their paper. Sowmya has also been championing the efforts to tackle the issues discussed above regarding data reporting.

Beyond her contributions to the journal, Sowmya has been a fantastic mentor to the *Nature Cell Biology* editors, always available to provide guidance and support. The team will miss her thoughtfulness, energy and ability to inspire, but looks forward to continue working with Sowmya in her new position as head of editorial policy for the Nature journals, and wishes her all the best in this new role.

We also take this opportunity to welcome our new chief editor, Alexia-Ileana Zaromytidou, who is well known to the cancer, cytoskeleton and cell adhesion and migration communities.