

An update on data reporting standards

We discuss editorial policies that aim to facilitate transparency and reproducibility, and their impact on the research content published in *Nature Cell Biology*.

In May 2013, the Nature life science research journals including *Nature Cell Biology* implemented new measures to raise the standard of methodological and statistical reporting in our papers. A key aspect of this initiative was the introduction of a mandatory [reporting checklist](#) that catalogued details of statistical information, experimental design and reagents, replicability of experiments, and compliance with editorial policies. Although the checklist itself is not published, the details provided are included in the legends and/or methods of published papers. All papers sent out for review, including revisions, are accompanied by the reporting checklist and reviewers are asked to comment on the methodological details provided. Asking our authors to provide this information at the outset is an important step to ensure that concerns about details of experimental design, analysis and statistical data are raised early in the review process and addressed satisfactorily in revision. This not only pre-empts multiple review cycles to address these issues further along the review process, but also ensures that details about statistical tests and experimental design are visible and accessible to editors, referees and readers alike. To promote clarity and transparency in reporting data and experimental details, we removed the word limits on the methods section. Authors are also urged to use the [Protocol Exchange](#), an open access resource at Nature Publishing Group for depositing step-by-step protocols of complex experimental procedures to complement methods sections in published papers.

In addition to defining statistical measures, error bars, tests and probability values, authors are asked to provide a detailed description of the samples used to derive the statistics and state the number of times the findings were reproduced. We strongly encourage authors to provide the source data used to derive the graphical data presented in the figures; this is especially important for small sample sizes ($n \leq 5$). 43% of papers published between July 2013 and February 2014 provided source data (a non-mandated option available to authors for data of their selection) for a variety of experimental data, including quantification of fluorescence and immunoprecipitation data, RT-PCR data, growth data for tumours and mass spectrometry data. Deposition of source data underlying large data sets is [mandated](#) for certain types of data in community-endorsed repositories; authors are asked to provide an accession number to allow confidential access for reviewers during the review process. In addition to increasing data transparency, source data is particularly useful in cases where specialist readers may be able to reanalyse the data to derive their own conclusions or support their own ongoing research. However, in some cases, the numbers behind the graphs are not likely to be much more informative than plotting the individual data points. Thus, we continue to strongly encourage authors to show the full spread

of individual data points in the figures where possible, particularly when sample size is small. Although bar graphs are the norm for data representation in most cell biology papers, we are pleased to see a small proportion of papers beginning to explore alternative formats for data display, including plotting the individual data points (see for example: *Nat. Cell Biol.* **15**, 1351–1361 (2013); *Nat. Cell Biol.* **15**, 1294–1306 (2013)). We join our sister journal, *Nature Methods*, in urging our authors to use box plots when sample size is greater than 5, and invite readers to explore BoxPlotR, an online tool for generating box plots developed as a collaboration between Nature Publishing Group and the community (*Nat. Methods* **11**, 113; 2014).

As well as details of statistical tests and sample sizes, we also ask for information about the source of cell lines used in the study, and whether cell lines have been authenticated and tested for mycoplasma contamination. An audit of papers with data generated in cell lines, published in the journal between August and December 2013, revealed that testing for mycoplasma contamination is fairly common and reported in 81% of papers, whereas only 19% of published papers carried out cell line authentication. So, although cell line contamination, misidentification and genetic drift is recognized by the [National Institute of Health \(NIH\)](#) as an issue that could potentially impair efforts to reproduce findings and many institutions provide cell line authentication services in core facilities, it has yet to become a routine aspect of experimental design in cell biology laboratories. The [International Cell Line Authentication Committee \(ICLAC\)](#) which was established in 2012 to raise awareness of cell line contamination and misidentification and to promote authentication, provides a host of resources for researchers to incorporate authentication into research practise, including maintaining an extensive database of cell lines that are known to be cross-contaminated or misidentified. Although Nature journals do not mandate cell line authentication, we would encourage researchers to incorporate regular testing into their experimental design.

We have been gratified by the largely positive feedback that we have received from our readers, authors and referees on our efforts to raise reporting standards. Beyond aiding in the transparency and clarity of reporting, it is our hope that the checklist will also help raise awareness of commonly encountered issues related to experimental design, statistical description, data analysis and presentation. Fundamental topics in statistics are covered in a series of monthly columns in *Nature Methods* launched last year; we hope our readers will find these pieces to be a valuable resource. These guidelines were developed in consultation with the research community and will evolve with feedback from the community; thus, we would like to hear your thoughts about our data reporting standards at cellbio@nature.com.