EDITORIAL

Statistical reviewing: constructive criticism towards reproducible research

Georg Heinze, Statistical Editor

Scientific evidence has always been closely connected to empirical research results. Statistics help to concisely summarize research findings and to draw scientifically defendable conclusions from them. Advances in statistical methodology have also paved the way to new opportunities in medical research, for example by providing methods of analysis of complex study designs or of high-dimensional molecular data. However, the growing complexity of these methods has led to excessive demands to reviewers of medical papers. Therefore, many medical journals nowadays rely on the expertise of statistical reviewers, and statistical reviewing has also become an integral part of the reviewing process of Transplant International. By implementing a Statistical Editor for this journal, Associate Editors and Editors can now request an expert review that specifically addresses questions of study designs, of selecting the most appropriate statistical methods, of describing them concisely, of plausibility and consistency of the results and of agreement of the conclusions drawn in a manuscript with the presented empirical results.

In statistical reviewing, it is important to point early at possible design failures of studies, for example to the likelihood of a severe selection bias, which would invalidate any conclusions. Usually such design errors cannot be corrected by any statistical method and are the most common statistical cause of rejection. In evaluating the choice of statistical methods for a particular analysis, our concern is not elegancy but rather robustness; that is, results should not be crucially depending on particular, often untestable, assumptions to hold. Description of statistical methods should always be concise, but with a minimum level of accuracy that would allow other analysts to reproduce the results of a manuscript with the same data. This requirement is in line with the initiative for reproducible research. It may sometimes require to prepare a supplementary document accompanying the main manuscript, which contains a comprehensive description of statistical methodology and sensitivity analyses. Inconsistency of results is often simple to reveal, for example by comparing *P*-values

and confidence intervals or by graphs obviously not agreeing with the reported statistical measures. Many cases of consistency checking, however, are trickier and require more experience with statistical methodology and medical data. The most common violation of the last point outlined above, agreement of conclusions with results, occurs if authors conclude 'equality' from 'non-significance', a misinterpretation that origins in a large type II (false negative) error probability typically inversely associated with sample size. The opposite case, falsely concluding a difference from an apparent significance, often results from conducting too many statistical tests in a study with no proper multiplicity correction, or, more subtly, from making too many data-driven decisions in a data analysis.

Our statistical reviews are always conducted under the paradigm of constructiveness. This means that where possible, we aim at helping the authors to improve the quality of their analysis, presentation and conclusions. It should also be added that the relevance of the medical research question of a study is not evaluated in a statistical review. Therefore, statistical reviewing is not dazzled by a particularly late-breaking research topic, and sometimes detects errors in a study report that would be left uncovered otherwise.

To help authors preventing the above-mentioned statistical issues already when preparing a manuscript, the Editors and I have compiled a list of ten important points to consider. The list has been made available at

doi:10.1111/tri.12758

the journal website as part of the author guidelines. We are confident that these simple rules will help potential authors to improve the quality and intelligibility of their research reports, to the benefit of authors, reviewers and readers of Transplant International.

Georg Heinze, PhD, associate professor, is biostatistician and head of the Section for Clinical Biometrics at the Medical University of Vienna. He provides regular consulting services to clinical researchers and has published >200 scientific contributions as main or co-author, many of them arising from international research projects. He regularly teaches medical biostatistics and statistical consulting at the Medical University of Vienna and supervises several postdoc, PhD and master students. He has served as reviewer for many leading medical and statistical journals. His own research focuses on biostatistics, regression modelling, observational studies and survival analysis.