EUROPEAN
UROLOGY

European Association of Urology

Platinum Opinion

Innovations in Statistical Review at *European Urology*

Melissa Assel [a], Daniel D. Sjoberg [a], James W.F. Catto [b], Andrew J. Vickers [a,*]

[a] Department of Epidemiology & Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA; [b] Academic Urology Unit, University of Sheffield, Sheffield, UK

It has repeatedly been demonstrated that the quality of statistics in the clinical research literature is very poor. As long ago as 1994, when statistician Doug Altman declared that poor quality medical research was a "scandal", he was able to cite studies documenting a high prevalence of statistical errors in clinical research [1]. Specifically in urology, Scales et al. [2] reviewed all 83 papers with statistical testing that were published in one of four urology journals in August 2005, and reported that 71% included at least one error. Imagine trying to explain that to, for example, a patient with bladder cancer, admitting that we cannot be sure of the next treatment steps because nearly three out of every four research papers have a statistical flaw.

Almost 20 yr after Altman's rallying cry, we at *European Urology* decided that a new approach was needed. In early 2013, statisticians were invited to join the editorial group, and we established three principles for improving the quality of statistics in papers published in our journal: (1) we would develop systematic guidance for authors based on a review of common errors in urology research; (2) all papers with substantive statistics that were under serious consideration for publication in *European Urology* would be reviewed by a statistician; and (3) we would continue to innovate with respect to statistical reporting.

Systematic guidance to authors was based on a review of close to 50 recent manuscripts published in or submitted to *European Urology* to identify common errors. In keeping with prior authors, we found a high prevalence of errors: indeed, it took until the 15th paper reviewed before we found a paper without an obvious error. The guidance was published in 2015 in *European Urology* [3]. This detailed guidance—amounting to seven journal pages—is quite distinct from the brief bullet points found in typical instructions to authors such as "It should be clear which statistical test is associated with each p value reported", mixed in with guidance on word counts and font sizes.

Those guidelines serve as the current basis for statistical peer review of *European Urology* manuscripts. Papers with substantive statistics that remain under consideration after an initial round of review and revision are sent to one of three statisticians for statistical evaluation after the first revision. In the years leading up to 2013, only approximately 2–3% of *European Urology* manuscripts were evaluated by a statistician. Since our implementation of the systematic statistical review in October 2013, 430 manuscripts have been reviewed by a statistical editor (A.J.V., D.D.S., or M.A.). An audit in 2018 found that >95% of papers with substantive statistics were subject to statistical review. Such a review goes beyond merely referring to the guidelines and evaluating the statistical methods and results; it includes feedback on how results should be presented and suggestions to improve the communication of concepts in figures and tables. Furthermore, statistical reviewers assess whether the interpretation of the results, including the discussion and conclusions sections, is consistent with the analytic approach and findings: understanding what numbers do and do not mean is, after all, a key part of statistical training and practice. As an example, a typical comment on the conclusions section is that causal language should be avoided if all that is reported is a statistical association.

One critical aspect of statistical peer review at *European Urology* is that we view our role not as traditional peer reviewers—providing the editor with a recommendation for or against publication—but primarily as one that supports authors by improving the quality of papers. It is rare (no more than 2 or 3 papers per year) that we reject a paper for publication because of irreconcilable statistical issues; our day-to-day job is to suggest to authors how they could

* Corresponding author. Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. Tel. +1 646-888-8233.
E-mail address: vickersa@mskcc.org (A.J. Vickers).

improve the statistical analysis or presentation. On several occasions, we have reached out to authors by phone to help them understand statistical issues and improve their manuscript. Statistical peer review is not a guarantee that published papers are completely free from statistical errors. However, we do catch many errors before publication and our review ensures that certain basic standards are upheld.

Our third principle of statistical review at *European Urology* is that we will continue to innovate. A recent example is our interest in the programming code underpinning statistical analysis. Starting in mid-2016, submitting authors were asked whether they used statistical code and, if so, whether they would be willing to submit it for archiving if their paper were accepted. Authors were informed that their response to this question would not impact their chances of acceptance. We published the initial findings from our experience in the *Annals of Internal Medicine* [4], reporting that use of code was incomplete (more than one-third of papers with nontrivial statistical analyses did not use code) and that no submitted code satisfied the three basic principles of good software: annotation, avoidance of repetition, and formatting for presentation.

We hope to encourage other journals to pursue the use of detailed statistical guidelines, routine statistical review, and innovations such as submission of programming code. To ensure the reliability of published works, it is essential to improve the quality of statistics reported throughout the medical literature. This is best done by including statisticians in the review process and giving detailed statistical guidance to prospective authors.

References

[1] Altman DG. The scandal of poor medical research. BMJ 1994;308:283–4.

[2] Scales Jr CD, Norris RD, Peterson BL, Preminger GM, Dahm P. Clinical research and statistical methods in the urology literature. J Urol 2005;174:1374–9.

[3] Vickers AJ, Sjoberg DD. Guidelines for reporting of statistics in European Urology. Eur Urol 2015;67:181–7.

[4] Assel M, Vickers AJ. Statistical code for clinical research papers in a high-impact specialist medical journal. Ann Intern Med 2018;168:832–3.