KI REPORTS
KIReports.org

Toward Transparency in Nephrology Research

Anna Francis[1,2] and Andrew Mallett[1,3,4]

[1]Faculty of Medicine, The University of Queensland, Queensland, Australia; [2]Child and Adolescent Renal Service, Queensland Children's Hospital, Queensland, Australia; [3]Department of Renal Medicine, Royal Brisbane and Women's Hospital, Queensland, Australia; and [4]Institute for Molecular Bioscience, The University of Queensland, Queensland, Australia

Kidney Int Rep (2020) **5,** 118–120; https://doi.org/10.1016/j.ekir.2019.11.019

See Clinical Research on Page 173

Transparency and reproducibility of nephrology research is an important and emerging topic. Research findings must be reproducible in order to demonstrate their reliability and translatability. In 1 famous case, an industry study of the effectiveness of an antidepressant in adolescents reported positive and safe results, but when reanalyzed by another team, serious safety concerns and no efficacy were found.[1] In order for other teams to reproduce results, the whole scientific process and endeavor needs to be transparent, to allow other teams to apply the same methodology. Transparency of methodology also fosters collaboration and scientific progress, as the easy transmission of knowledge allows other teams to extend on known approaches. Limitations and barriers to the widespread adoption of a transparent and reproducible approach include protecting patient confidentiality, logistical difficulties in data sharing, intellectual property and the understandable desire of some teams who set up large studies to have exclusive access to the results.

It is routine practice in other fields of science for data and analysis code to be uploaded to freely accessible online repositories such as GitHub (https://github.com) so that other teams can interrogate and reproduce the work. Within medicine, there is growing support for transparent and reproducible research, although the conversation is not without controversy.[2,3] The National Institutes of Health (NIH) have been vocal in their support of data sharing within medical research, stating that "data sharing is essential for expedited translation of research results into knowledge, products, and procedures to improve human health."[S1] Large NIH-funded studies are mostly expected to deposit deidentified data within an open repository, after the main findings of the study are published. The International Committee of Medical Journal Editors (ICMJE) have stated that "it is an ethical obligation to responsibly share data generated by interventional clinical trials because participants have put themselves at risk,"[S2] and have proposed that randomized trial data would be shared soon after publication. After strong debate in the medical community[2,3] they downgraded their recommendations from mandated data sharing to a requiring a data sharing plan.

A 2018 study assessed the effectiveness of data sharing after 2 major journals (*BMJ* and *PLOS Medicine*) instituted a data sharing policy for randomized trials.[4] The authors assessed 37 studies published in the 2 journals after the data sharing policy was introduced. Only 46% had accessible data. Of the studies with accessible data, 82% were fully reproducible in their primary outcomes. There were 2 studies in which the authors found minor errors but similar final outcomes, and 1 study did not have sufficient published methodology to replicate results. The authors found data sharing to be hampered by difficulties such as being unable to contact corresponding authors, refusal of data sharing requests, or receiving datasets that the original authors had not had the time or resources to format so as to be easily interpretable to an outsider. Data sharing can be taxing for individual researchers. Major request sites, such as Clinical Study Data Request (https://clinicalstudydatarequest.com) is a secure access site where researchers can request data and analyze it within the secure online environment. They cannot download the data, and, after a set time period, they lose access to it. However, for small players, the costs can be prohibitive, with institutions needing to pay tens of thousands of dollars per year to host and access data.[3]

Beyond these logistical difficulties, there are many potential constraints regarding data sharing within medical research. The most

Correspondence: Anna Francis, Queensland Children's Hospital, Stanley Street, South Brisbane 4101, Queensland, Australia. E-mail: anna.francis@health.qld.gov.au

important is the need to protect patient confidentiality at all times. It is imperative that individuals not be identifiable, which could place them at risk for being denied employment, denied insurance, or being otherwise targeted by individuals or companies. Data sharing agreements are important to protect the data. Agreements can limit data sharing to restrict the transfer of data to approved applicants and to ensure that data are used only for research purposes. Restrictions placed by federal, state, or jurisdictional law and by institutional or ethical review boards must also be adhered to. In addition, many research studies are funded in full or part by industry, and proprietary data, intellectual property, and patent issues can arise with data sharing. We also must acknowledge the hesitation of many researchers, who have put many years of effort into setting up clinical trials, in handing data over to other researchers soon after publication of the primary results.[2]

Within medicine, transparency and reproducibility have been best integrated into genomics research. NIH-funded genomics studies have data deposited into a federal repository, called dbGAP (database of Genotypes and Phenotypes). At a population level, our fundamental understanding of genetic variation and architecture hinges upon transparency of large-scale genomic data that have been acquired through global research initiatives and collaborations, such as is exemplified by GnomAD (https://gnomad.broadinstitute.org). Furthermore, common and transparent international guidelines for analysis of genetics variants relating to human phenotypes has led to similarly open and transparent archive deposition of such assessments of individual variants.[5] Adding further to critical transparency and reproducibility have been further initiatives to dynamically assess the evidence of gene—disease associations as these are rapidly evolving. Within nephrology, scalable benefits are being realized by scaled genomic data sharing,[6] although this is creating greater clarity around issues of interpretation, specifically phenotype assessment, utility, and evidence-based guidelines.[7,8]

Transparency extends beyond data sharing and includes reporting funding sources, conflicts of interest, preregistration of protocols, and detailed explanations of methodology, statistical analysis, and results. *Kidney International* and *Kidney International Reports* require the identification of funding sources and any conflicts of interest on the ICMJE Form for Disclosure of Potential Conflicts of Interest, as do most major nephrology journals. The Equator (Enhancing the QUAlity and Transparency Of health Research) Network was established to enhance the quality of health research by standardizing reporting guidelines for methodology, analyses, and results for major study types. An example reporting guideline is the Consolidated Standards of Reporting Trials (CONSORT) checklist for randomized trials.[S3]

Fladie *et al*. have evaluated the nephrology literature for indicators of transparency and reliability of research.[9] The authors have gone to lengths to make the process for this article transparent and reproducible. They randomly sampled 172 articles from English language nephrology over 5 years (2014—2018). They found that transparent and reproducible research practices are rare in nephrology; for example, only 2% of studies made the analysis script or protocol available, and less than one-fourth had data available. The authors suggest data sharing and the use of online repositories as ways forward to enhance reproducible and robust research.

Many of the markers of reliability applied by Fladie *et al*. usually apply only to randomized trials; however, in this study they have applied to a wide range of study types, including retrospective cohort studies. Some of the markers of transparency and reliability chosen by the authors include whether the study was preregistered and whether a protocol is available. Although this is the standard for randomized trials (only 17 of which are included in this study), many of the included studies were retrospective registry cohort studies, where pre-registering a protocol may be of less importance. Despite this, the level of data sharing or availability of analysis scripts or study protocols was low throughout the studies inspected, suggesting that nephrology research has a long way to go before it is transparent and reproducible.

Our clinical decision making relies on robust evidence. This evidence can be best generated by transparent and reproducible research. Achieving this goal involves working as a community to facilitate data sharing while ensuring patient confidentiality is protected, and undue burdens are not placed on the initial researchers. In addition, it is important that researchers be rewarded for their endeavors, especially in light of the enormous amount of work required to set up a clinical trial. Nephrology journals are increasingly requiring data sharing agreements, particularly for genomic research, which will hasten the widespread adoption of transparent and reproducible research practices and will strengthen the evidence base for our clinical practice.

DISCLOSURE

AM has received competitive research funding from Sanofi-Genzyme and has been an advisory board member for Otsuka (Australia). The other author declared no competing interests.

SUPPLEMENTARY MATERIAL

Supplementary File (Word)

Supplementary References.

REFERENCES

1. Birkenhager TK. Both paroxetine and imipramine appear to be ineffective in adolescents with major depression, furthermore doubts have risen about their safety. *Evid Based Med*. 2016;21:92.
2. Kalager M, Adami HO, Bretthauer M. Recognizing data generation. *N Engl J Med*. 2016;374:1898.
3. Rockhold F, Nisen P, Freeman A. Data sharing at a crossroads. *N Engl J Med*. 2016;375:1115–1117.
4. Naudet F, Sakarovitch C, Janiaud P, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in The BMJ and PLOS Medicine. *BMJ*. 2018;360:k400.
5. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–424.
6. Mallett AJ, McCarthy HJ, Ho G, et al. Massively parallel sequencing and targeted exomes in familial kidney disease can diagnose underlying genetic disorders. *Kidney Int*. 2017;92:1493–1506.
7. Snoek R, van Eerde AM, Knoers N. Importance of reliable variant calling and clear phenotyping when reporting on gene panel testing in renal disease. *Kidney Int*. 2017;92:1325–1327.
8. Groopman EE, Rasouly HM, Gharavi AG. Genomic medicine for kidney disease. *Nat Rev Nephrol*. 2018;14:83–104.
9. Fladie IA, Adewumi TM, Vo NH, et al. An evaluation of nephrology literature for transparency and reproducibility indicators: cross-sectional review. *Kidney Int Rep*. 2020;5:173–181.