Clinical Science (2021) **135** 101–103
https://doi.org/10.1042/CS20201298

PORTLAND PRESS

Commentary

Time to *fiddle* with your unpublished data

Michael J. Ryan[1,2] **and Rhian M. Touyz**[3]

[1]Department of Physiology, University of Mississippi Medical Center, Jackson, MS, U.S.A.; [2]G.V. (Sonny) Montgomery Veterans Affairs Hospital, Jackson, MS, U.S.A.; [3]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, U.K.

Correspondence: Michael J. Ryan (mjryan@umc.edu) or Rhian M. Touyz (rhian.touyz@glasgow.ac.uk)

Most scientific investigators conduct well-designed and controlled preclinical experiments generating data that are difficult to explain, contrast with existing scientific dogma, or represent a perceived negative result. It is common for these findings to remain hidden away in a drawer from the greater scientific community. However, these unseen results can lead to publication bias, have the potential to significantly advance scientific disciplines if they are published, and can help investigators avoid repeating experiments that have already been done, thus saving money and time. Moreover, these unexpected data may actually have significance if re-interpreted leading to new hypotheses. This editorial commentary highlights a novel user-friendly tool developed by Bernard and colleagues (REF) to help investigators determine appropriate options for disseminating unpublished data in order to make them available to the broader scientific community. In addition, this commentary serves as an announcement for an upcoming special call for papers on meta-research to be published in *Clinical Science*. Meta-research is the evaluation and study of existing scientific literature and data. It is an evolving field dedicated to improving rigor and reproducibility in science, an endeavor to which *Clinical Science* and Portland Press are committed.

In 2012, Begley and Ellis [1] published an article highlighting the critical need for dramatic improvements in the rigor and reproducibility of preclinical research. The authors attempted to replicate the key findings of 53 landmark papers in the field of cancer therapeutics. Of those 53 studies, only 6 were replicated. It was noted that the failure of preclinical studies to translate into meaningful therapeutic targets was based on several factors. The most common factors cited included the infrequent use of randomization and blinding in experimental design, and the selective inclusion of data that supports the hypothesis while omitting data that may have enhanced the interpretation, or led to a different conclusion.

It should not come as a surprise that the issue of reproducibility in preclinical studies reaches far beyond the cancer therapeutics field. For example, a similarly impactful article by Prinz et al. [2] focused on the reproducibility of preclinical work in the areas of oncology, women's health, and cardiovascular research. The authors described how pharmaceutical companies, in this case Bayer Healthcare (Berlin, Germany), commonly have internal programs designed to test and validate published preclinical scientific data. This is a way to ensure that the company is making sound decisions related to therapeutic targets to pursue. During an assessment of the internal program at Bayer Healthcare, 23 participating scientists completed a questionnaire of approximately 67 validation projects that were attempted in the three areas described above. The results of the questionnaire noted that only 20–25% of the preclinical studies upon which the validation experiments were based, completely matched the published result. It was postulated that the low rate of verification was commonly related to the incorrect use of statistical tests in the published preclinical studies, the low sample size used in the papers, and the general pressure under which scientists operate to publish only positive results.

The significant pressure to publish positive results that are perceived to be high impact is brought to bear both by academic institutions and funding agencies that place high value on, and therefore incentivize, these types of results. For example, high-impact papers in high-impact journals are used by institutions to evaluate candidates who are eligible for promotion and/or tenure. Similarly, funding agencies seek to reward the most innovative and impactful work, and peer reviewer panels understandably heavily weigh

Received: 07 December 2020
Revised: 16 December 2020
Accepted: 17 December 2020

Version of Record published:
06 January 2021

Clinical Science (2021) **135** 101–103
https://doi.org/10.1042/CS20201298

the publication impact of grant submitters in their evaluations. Although this competitive environment has advantages to drive science forward, there are key limitations of the current system that may be slowing scientific progress. Chief among these, is the perception that negative results are not welcome in the scientific literature.

Fortunately, the attitude toward publishing these 'hidden' datasets are changing within the scientific community, and there is a concerted effort to improve rigor and reproducibility in science. For example, there are now incentives offered by scientific societies that give monetary awards for the publication of high quality negative data, or data that contrast with the existing literature [3]. Efforts to improve in this area are also taking center stage at major funding agencies such as the National Institutes of Health in the United States that outlined plans to enhance rigor and reproducibility in science [4]. The NIH noted that irreproducibility in science does not necessarily occur as a result of academic misconduct. Rather, it often results from poor or inconsistent educational training, the lack of detailed methods in the published literature, and the failure of experimental designs to consider blinding, randomization, sample size, and sex as a biological variable. Importantly, the NIH also cited the potential detriment of not having unpublished datasets and negative results widely available to the scientific community.

In this issue of *Clinical Science*, Bernard et al. (REF) offer a new tool that is intended to help the scientific community overcome the stigma of negative data and make it more widely available to the scientific community. This change in mentality is critical for authors, reviewers, editors, and readers. The authors picked up on the concept that most investigators are likely to have significant datasets from well-designed studies that are simply 'filed' away never to be viewed by the scientific public. The common factors influencing an investigator's decision to not share data can be attributed to the misconception that results that do not achieve statistically significant differences are less relevant than those that do, the perception that results may be faulty if they do not match the published literature, small sample sizes in the experimental design making investigators hesitant about sharing the data, and selective reporting of data that is likely to support the hypothesis while being less likely to publish those that do not. The publication costs associated with publishing these data may also be a concern to investigators, and there is a need to educate the scientific community about the resources and repositories than can be used to overcome all of these challenges.

In the present paper, the authors present the **f**ile **d**rawer **d**ata **l**iberation **e**ffort (*fiddle*), an open source app that is described as a 'matchmaking' tool for investigators to determine different ways to publish data in both traditional journals and through less traditional means. *fiddle* is a user-friendly app that allows the investigator to select options such as the type of dataset, cost of publication, indexing of the final published work, whether or not the data are peer-reviewed, and the speed at which the investigator wants the dataset available to the public. App users can also approach *fiddle* by selecting scenarios that describe their specific needs. For example, 'My dataset is incomplete', or 'I don't have funding to pay for publication charges'. *fiddle* makes recommendations as to type of publication (i.e. data repository, micropublication, preprint etc.) that meets the needs selected by the investigator. The ultimate goal is to help investigators identify an outlet that will allow for the dissemination of data from properly designed and controlled experiments that may not otherwise be submitted for publication. *fiddle* very ably accomplishes this goal, and thus provides an important tool that can be used to advance the entire scientific enterprise.

As investigators begin to empty their literal and figurative data file drawers to bring new datasets into the public domain, an important step toward improving rigor and reproducibility in science will have been taken. With an influx of available data, the field of meta-research can evolve. Meta-research is the evaluation and study of existing scientific literature and data. It is an evolving field dedicated to improving rigor and reproducibility in science. The community of meta-researchers is small, but growing. The efforts of these researchers have already had an important impact on issues such as effective data presentation allowing for more careful data analysis by the scientific community [5], and their impact on improving preclinical and clinical research will continue to grow. *Clinical Science* and Portland Press are committed to the highest standards of scientific integrity. As a testament to this commitment, the journal will be announcing a call for meta-research papers to be published. *Clinical Science* looks forward to furthering the cause of making data available and ensuring that the published literature meets the standards that are needed to advance biomedical research.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work was supported by the Veteran's Administration Merit Award [grant number BX002604 (to M.J.R.)]; the UMMC-Department of Physiology and Biophysics [grant numbers P20GM104357, U54GM115428]; and the British Heart Foundation [grant numbers RE18/6/34217, CH/12/29762 (to R.M.T.)].

PORTLAND PRESS

Abbreviation

fiddle, file drawer data liberation effort.

References

1 Begley, C.G. and Ellis, L.M. (2012) Drug development: raise standards for preclinical cancer research. *Nature* **483**, 531–533, https://doi.org/10.1038/483531a

2 Prinz, F., Schlange, T. and Asadullah, K. (2011) Believe it or not: how much can we rely on published data on potential drug targets? *Nat. Rev. Drug Discov.* **10**, 712, https://doi.org/10.1038/nrd3439-c1

3 (2017) Rewarding negative results keeps science on track. *Nature* **551**, 414, https://doi.org/10.1038/d41586-017-07325-2

4 Collins, F.S. and Tabak, L.A. (2014) Policy: NIH plans to enhance reproducibility. *Nature* **505**, 612–613, https://doi.org/10.1038/505612a

5 Weissgerber, T.L., Milic, N.M., Winham, S.J. and Garovic, V.D. (2015) Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol* **13**, e1002128, https://doi.org/10.1371/journal.pbio.1002128