Data sharing requirements: perspectives from three authors

DATA SHARING: THE GOOD, THE BAD, AND THE UGLY

Kurt T. Barnhart, M.D., M.S.C.E.

The International Committee of Medical Journal Editors (ICMJE) believes there is an ethical obligation to responsibility share data generated by interventional clinical trials (1). The goal of this opinion is that sharing of de-identified individual participant data will become the norm. The World Health Organization has articulated that best practice for clinical trials include: universal prospective registration, public disclosure of results, and data sharing. Achieving these goals will maximize generation of high quality data resulting in the practice of evidence based medicine, while concomitantly maximizing the ethical obligation to trial participants who put themselves at risk. The research community has started to adopt prospective clinical trail registration, while uptake is yet to be universal. Achieving the other goals, including data sharing, is both good and bad, and assuredly will be ugly.

There are obvious benefits to data sharing. The availability of scrutiny of data by peers, or publicly, will increase transparency. Transparency will minimize false or overreaching conclusions, as well as potential of suppression of "undesirable" findings. Sharing data will also maximize use of data allowing secondary analysis, addressing additional questions, generating additional hypotheses and aggregation with other data. There is clearly inherent value in pooling data, when possible, allowing potential increase in power, assessment of subgroups, or validation of findings.

Drawbacks in data sharing include how to appropriately give scholarly credit to those who generate and share the data, as well as those who use the data for secondary purposes. The conception, conduct, and analysis of a clinical trial is very expensive and time consuming. Despite this effort, it is often difficult to distinguish impact of manuscripts from the primary data from manuscripts based on reanalysis or aggregation. If academic productivity is measured by the number of scholar manuscripts (as it often is), it will be far easier to develop a career in secondary analysis compared to primary data generation.

The logistics of data sharing will be ugly. Many aspects will have to be resolved. Examples include how to ensure transparent data request, approval and prioritization. How will data be archived and at whose expense? What data security is necessary to protect data? While data will be de-identified, there is always the potential of unblinding a subjects' identity based on data. This is especially true if genetic data is included. There have been instances when individuals have claimed they were identifiable based on public sharing of genetic data.

While these important issues are debated and practice evolves, clinical trialists should start to consider data sharing plans now. While there is consensus that data should be shared, there is no consensus on how much data should be shared or how. A data sharing plan should indicate what particular data will be shared (i.e. individual patient data, data dictionary, protocol, data analysis plan, primary and secondary analysis, genetic data), with whom it can be shared, when it will be available, for how long and what is the mechanism. A data sharing plan is already a requirement of National Institutes of Health (NIH) grant application. Data sharing statements may soon be required at the time of manuscript submission. While data sharing is not yet mandated, in the wild west that is publication peer review data, consideration of data sharing statements are starting to influence editorial decisions.

THE TIMES (AND THE DATA) THEY ARE A-CHANGIN'…FOR THE BETTER

Richard S. Legro, M.D.

When Bob Dylan came out with his eponymous album of this section's title in 1964, did anyone see him one day winning a Nobel Prize? Similarly when the ICMJE, to which *Fertility and Sterility* and most other high impact journals belong, announced in 2004 (2) the mandatory registration of all clinical trials prior to commencement, did anyone foresee the radical transformation in clinical trial reporting and oversight that would follow. Now the most recent 2017 ICMJE announcement regarding clinical trials requires a report in the manuscript of a data sharing plan (1). This plan should include if and how data will be shared, though it does not (yet) mandate data sharing. Let me, in this Inkling, review the impact of the initial 2004 ICMJE mandate, discuss the implementation of the 2017 mandate and speculate about the future of data sharing.

The initial 2004 trial registration mandate resulted from the fact that there was selective reporting of clinical trial results. Trials in which a drug or interventions succeeded were reported, but trials with negative or harmful results never saw the light of day. While it is popular to identify Big Pharma as the primary culprit, there are plenty of instances of investigator-initiated and NIH funded trials where only the positive survived to publication and the negative were buried (3). Without transparency of clinical trial registration prior to initiation that identifies among other criteria, the primary and secondary outcomes, there was also post hoc cherry picking of outcomes. Surrogate markers that were an afterthought of the trial were elevated to primary outcomes if no major health benefits were noted. I think metformin has been highlighted in the title of multiple manuscripts to improve just about every surrogate marker of inflammation in women with polycystic ovary syndrome. My response: why would anyone do such an intensive expensive trial for such a clinically irrelevant outcome as change in circulating PAI-1, CRP, IL-6, etc.? Failure to account for multiple hypothesis testing leads to an increased incidence of type 1 errors. While the uptake of clinical trial registration was gradual, it is now rare, if not exceptional to see a clinical trial reported that was not registered a priori. As an Associate Editor of this *journal*, I can say that while submission of such a manuscript still occurs, they are not sent out for review as they do not meet the 2004 ICMJE mandate for clinical trial registration. The times they are a-changin.'

Although initially the impetus to trial registration was primarily a carrot, i.e. possible publication of your work in a high impact journal, the failure to register and timely report the

results of a clinical trial on Clinicaltrials.gov now carries substantial penalties through the Department of Health and Human Services Final Rule, implemented in January of 2017 (with a similar NIH policy at the same time). The Final Rule requires a responsible party to both register the trial at Clinicaltrials.gov and to submit summary results information to ClinicalTrials.gov for any applicable clinical trial (within one year of completion), regardless of whether the drug, biological, or device products under study have been approved, licensed, or cleared for marketing by the Food and Drug Administration. Noncompliance can be noted on the clinical trial record at Clinicaltrials.gov and can in certain instances result in significant monetary penalties (on a daily basis until corrected) to the sponsor. Our academic health center has responded by requiring all registered clinical trials, whether NIH, industry or investigator-initiated to comply with the final rule for registration and reporting. The times they are a-changin.'

The impact of publishing a statement requiring a data sharing plan in a manuscript, especially when the statement can read, "Data will not be shared," seems minimal at first, *a la* the 2004 ICMJE announcement. But how will those investigators who refuse to share data be viewed by their peers and eventually by their peer reviewers and editors? If most or the best investigative teams are sharing their data, will not eventually peer pressure force the others to come on board? The NIH requires a formal data sharing plan (without the option of an opt out) for all clinical trial applications with direct costs over $500,000 per year. Industry and the NIH are making de-identified clinical trial data available right now. As an example, The Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) has created a Date and Specimen Repository (DASH; https://dash.nichd.nih.gov) for clinical studies, where these can be accessed by investigators throughout the world. NICHD-funded clinical trial networks, including the Reproductive Medicine Network, have already posted their data from specific trials for accession. To date there have been over 10,000 queries requesting data and/or specimens at DASH with 10% originating from abroad (Personal Communication, Diana Bianchi, Director NICHD, November 13, 2017). The times they are a-changin.'

The need for clinical trial data sharing is increasing in a world where emerging evidence based medicine data synthesis methods above and beyond conventional meta-analysis can weave together disparate clinical trial data to prioritize research and guide personalized patient care. These include network meta-analysis, a tool to rank the efficacy of various interventions where head to head comparisons are lacking. For instance a recent network meta-analysis we performed suggested the next trial of ovulation induction in women with polycystic ovary syndrome should be letrozole vs clomiphene plus metformin (4). Another emerging method is individual patient data (IPD) meta-analysis which uses original (de-identified) patient data to better identify responders and non-responders to treatments (as well as those experiencing adverse events) especially timely in this age of personalized medicine. Thus it does not take too much speculation to see the next steps in data sharing, i.e. accessible data through a clear mechanism, and more likely the immediate availability upon publication to all interested parties of de-identified clinical trial databases relating to the reported primary and secondary outcomes of clinical trials. This will likely first be piloted by the highest impact journals and then trickle down the hierarchy of journals. I hope that *Fertility and Sterility* will be high up the water chain. Clinical trial protocols, once highly protected documents, the intellectual property of clinical trialists, are now available as supplementary materials at many high impact journals, allowing assessment of rigor and easy replication of results. Replication is the hallmark of science. The times they are a-changin.'

The requirement to share all primary data from a clinical trial will ultimately reward those investigators, independent of name or nation or funding body from any categorical discrimination (think of investigators from the BRICS, the up and coming economies of the world (Brazil, Russia, India, China, and South Africa) and beyond and the more rigorous scrutiny their trials evoke). When the data are shared, the data will speak for itself. We are fortunately in a time of greater transparency of research coupled with increasing concerns about the lack of reproducibility of our research. One of the reasons for lack of reproducibility is the increasing recognition of fraudulent trials, where data are invented or massaged to reach a pre-determined outcome. Unfortunately according to the website Retraction Watch (http://retractionwatch.com/the-retraction-watch-leaderboard/), the largest mass perpetrators of fraudulent published data resulting in manuscript retraction are clinical trialists. Yoshitaka Fujii, a Japanese Anesthesiologist, has had to date 183 manuscripts retracted, followed by a German anesthesiologist Joachim Boldt with 96 (and yes, there are U.S. clinical investigators on the list and 30 of the top 31 are men!). Most, if not all of these retracted clinical trials by these investigators were never performed and the results were simply made up. These fraudulent trials are also present in our field of infertility research and one can argue that the double barrels of academic pressure and financial rewards for developing interventions that improve pregnancy rates is fertile soil for potential fraud and abuse of data. These trials in our field, if fraudulent or incorrectly analyzed, may also sway practice the wrong way if they are incorporated into meta-analyses. For instance, there have been conventional aggregate data meta-analyses of aspirin use in in vitro fertilization (IVF) showing a benefit on pregnancy rate (5). When an IPD meta-analysis of this topic was performed, only 6 of 10 eligible RCTs in the study could provide independent patient data (6). The conventional meta-analysis using the aggregate data from all 10 studies showed an effect favoring aspirin, whereas when only the aggregate data from the 6 studies providing IPD data were used the effect direction reversed against the use of aspirin in IVF. Such findings are sobering, and while full disclosure of clinical trial datasets has its own issues as noted by my fellow Inklings authors, the benefits of transparency of data both in primary and secondary analyses as well as in evidence based syntheses of clinical trials far outweigh the burdens of reporting. Following the stream, mandatory data sharing is likely in the near future and we should prepare for it. As Bob Dylan sang in 1964, "…the waters [a]round [us] have grown…" (and they are not retreating). He advised that if you wanted to survive, "…you better start swimming/or you'll sink like a stone," for not only have the times changed, they are still a-changing'- for the better.

Download English Version:

https://daneshyari.com/en/article/8779795

Download Persian Version:

https://daneshyari.com/article/8779795

Daneshyari.com