SPECIAL COMMUNICATION

Researchers' Experience with Clinical Data Sharing

Geoffrey A. Mospan, PharmD, BCPS, and Kurt A. Wargo, PharmD, FCCP, BCPS, (AQ-ID)

The topic of transparency in industry-sponsored clinical trials has gathered the attention of researchers in medicine. Patient-level data from recently completed clinical trials is now available for investigators to reanalyze or perform new analyses. In this Special Communication, the authors discuss their experience using this type of research and provide recommendations for success. (J Am Board Fam Med 2016; 29:805–807.)

Keywords: **Industry, Medicine, Research, Research Personnel**

Introduction to Industry-Sponsored Clinical Trial Data

Clinical trial data sharing has recently become a topic of interest among many researchers. The International Committee of Medical Journal Editors (ICMJE) proposed a new policy in order for the data from all published clinical trials to be reported because participants in these trials put themselves at risk.[1] This deidentified, individual patient data (IPD) should be shared within 6 months of trial publication and a plan for data sharing should be included with the initial clinical trial registration.[1] If this policy is adopted, authors will be required to provide IPD to publish their findings in an ICMJE member journal. The goals of this data sharing will be to "increase trust in the conclusions drawn from trials, enable independent confirmation of the results, and foster the development and testing of new hypotheses."[1] In fact, it is encouraged and expected that researchers granted access to the shared IPD will publish the results of their analysis.[2] ICMJE received a considerable amount of feedback on the proposal, and has not yet published an official policy statement.

In May 2013, GlaxoSmithKline (GSK) made available data from nearly 1000 clinical trials they sponsored since 2007.[3] Since the availability of this data, 58 proposals to gain access to the data of clinical trials had been made, with only 13 researchers using the data to perform research.[3] Unfortunately, none of the data provided by GSK have yet led to peer-reviewed publications.[3]

Here we describe our experiences using the clinical data provided through this initiative, and developed a list of recommendations in hopes that it will help others.

Steps for Successful Clinical Trial Database Research

Like any other avenue of research, the first step in the process is to formulate and identify the question or problem. Given the nature of using clinical trial data for research, the hypothesis is likely to develop after an investigator ponders a clinical question or reads a trial in a medical journal. Furthermore, in the United States, industry-sponsored clinical trials can be located in a public registry at www.clinicaltrials.gov. This is how our journey using clinical trial data began.[4] We are often asked about appropriate duration of antibiotic use in male patients with urinary tract infections (UTIs). After performing a literature search on the topic, we came to a trial that used a shorter course of levofloxacin in treatment of UTI. However, this study combined the results of female and male patients in the dataset. Therefore, by obtaining the clinical trial data, we were able to perform a post-hoc, subgroup analysis in an effort to determine whether

This article was externally peer reviewed.
Submitted 17 June 2016; revised 26 August 2016; accepted 30 August 2016.
From the Wingate University School of Pharmacy, Wingate, NC (GAM); Wingate University School of Pharmacy, Hendersonville, NC (KAW).
Funding: none.
Conflict of interest: none declared.
Corresponding author: Geoffrey Mospan, PharmD, BCPS, 515 N. Main Street, Wingate, NC 28174 (E-mail: g.mospan@wingate.edu).

a difference in cure rates existed between males and females.

The next step was the largest hurdle in this research process, and that was to determine whether the data were obtainable. There are 2 main platforms currently where clinical trial data inquiries can be made: through the Yale University Open Data Access (YODA) Project[5] and ClinicalStudyDataRequest.com. The authors of this editorial used the former platform. On these platforms, a list of participating pharmaceutical companies and the available trials can be found. Of note, if a trial is not listed, a request or inquiry can be made to determine its' availability. In addition, if a pharmaceutical company does not appear on either of these 2 platforms, they may provide clinical trial data through their Web site.

After confirming that the data from the clinical trial were available, we submitted a proposal to the YODA Project that included information on methods, hypothesis, objectives, and statistical analysis. Based on our understanding, the review process for obtaining data are not synonymous with an institutional review board (IRB), and we concomitantly submitted an application to our local IRB.

The YODA Project conducts a blinded review of the proposal, without assistance from the sponsoring drug manufacturer.[6] After notice of approval, a Data Use Agreement (DUA) was signed between ourselves and the YODA Project. The DUA detailed requirements for us, as the investigators, including maintaining confidentiality, reporting results, and limiting analyses to those included in our research proposal. The individual patient data were deidentified and provided on a secure, remote desktop, the SAS Solutions on Demand Secure Portal. Inside this portal included numerous documents in a spreadsheet format with all the data collected in the original clinical trial. A detailed protocol from the original investigators was included on the secure portal. This provided valuable insight into the methods, definitions, and statistical analysis from the trial, more than what was published in the original manuscript.

During a poster session of our post-hoc data, we were asked about the "cleanliness of the data" using this type of research platform. The available data were in the final format needed for evaluation, decoding or decrypting were not necessary and results of the study objectives were clearly reported (ie, cured, failed, etc). To ensure that appropriate analyses were made, we replicated the results of the original trial before conducting our own study. In other words, using the data on the platform, we recalculated the cure rates in each group from the original trial and matched the data with the published results. This additional step was important to complete for quality assurance purposes.

As the researchers performing this secondary analysis, we believed that the process was appropriately transparent based on the information provided by the drug manufacturer. Any questions we might have had about data collection, specifications, or protocols were clearly delineated in the documents available to us.

Limitations of Using Clinical Trial Data

The largest time-consuming challenge we faced during the process was becoming familiar with the format of the clinical trial data. Each spreadsheet included headers and abbreviations that were delineated in a separate document. After identifying the necessary data and its' location within the portal, we were ready to perform statistical analyses. Therein lies another limitation; this could only take place on the remote desktop with the statistical software included, the clinical trial data could not be downloaded to a personal computer. Therefore, we were not able to use statistics software with which we were familiar; rather, we had to learn how to use the available software.

Pharmaceutical companies recently committed to providing clinical trial data to independent researchers and not all trials are available, especially older trials. Many pharmaceutical companies do not make their data available until after a certain time period after the trial has been performed and the drug has been approved in Europe and the United States.

The types of research that can be performed using clinical trial data are limited. Researchers may be able to obtain data from multiple clinical trials in an effort to perform a meta-analysis. However, the majority of research will likely be post hoc, subgroup analyses which, in turn, has limitations. Firm conclusions cannot be made based on having underpowered statistics and caution should be made when interpreting such studies. Nonetheless, subgroup analyses can provide new insight to clinical inquiries and become a stepping stone for future studies.

Conclusion

Using up-to-date literature to make informed decisions about the optimal care of patients is crucial to today's practitioners. These practitioners have historically taken a leap of faith in their reliance on investigators of published clinical trials to exhibit sound, and unbiased, research principles and ethics when conducting, analyzing, and publishing their research. The time has come when we no longer have to rely solely on their published data; rather, we can access this data and analyze it ourselves. Unfortunately, the difficulty with which this data can be accessed and analyzed will likely deter some practitioners and researchers from taking it on themselves to conduct these types of analyses.

Recommendations

- Formulate a hypothesis or question that data in the published clinical trial did not address.
- Ensure accessibility of individual patient data of the clinical trial from YODA or ClinicalStudyDataRequest.com.
- Submit request/proposal to study the data to the platform and to your institutional IRB.
- Read and understand the protocol of the original investigators, to ensure your accurate assessment of the data provided.
- Familiarize yourself with the format of the clinical trial data provided to understand the abbreviations that were used.
- Before conducting your own analysis, consider reproducing the results of the original investigators to ensure validity of the data.

References

1. Taichman DB, Backus J, Baethge C, et al. Sharing Clinical trial data—A proposal from the international committee of medical journal editors. N Engl J Med 201(4)6;374:384–6.
2. Principles for responsible clinical trial data sharing. Available from: http://transparency.efpia.eu/responsible-data-sharing. Accessed May 25, 2016.
3. Strom BL, Buyse M, Hughes J, Knoppers BM. Data sharing. Year 1—Access to data from industry-sponsored clinical trials. N Engl J Med 201(22)4;371:2052–4.
4. Mospan GA, Wargo KA. 5-day versus 10-day course of fluoroquinolones in outpatient males with a urinary tract infection (UTI). J Am Board Fam Med 2016;29:654–62.
5. The YODA Project home page. tp://yoda.yale.edu.
6. Krumholz HM, Waldstreicher J. The Yale Open Data Access (YODA) Project—A mechanism for data sharing. N Engl J Med 2016 Aug 4;375(5):403–5.