Published in final edited form as:

*JAMA*. 2016 March 22; 315(12): 1283–1284. doi:10.1001/jama.2016.2374.

# Use of Open Access Platforms for Clinical Trial Data

**Dr. Ann Marie Navar, MD, PhD**, **Dr. Michael J. Pencina, PhD**, **Dr. Jennifer A. Rymer, MD**, **Dr. Darcy M. Louzao, PhD**, and **Dr. Eric D. Peterson, MD, MPH**

Duke Clinical Research Institute, Duke University Medical, Center, Durham, NC

Concerns over bias in clinical trial reporting have stimulated calls for more open data sharing.[1] In response, multiple pharmaceutical companies have created mechanisms for investigators to access patient-level clinical trials data. However, if and how these shared clinical trial data are being used is unknown.

## Methods

We evaluated how many clinical trials were publicly available to investigators through 3 open access platforms: ClinicalStudyDataRequest.com, Yale University Open Data Access Project (YODA) and the Supporting Open Access for Researchers (SOAR) initiative. [2,34] Sponsors depositing data in these platforms include: GlaxoSmithKline (GSK), Astellas, Boehringer Ingelheim, Eisai, Lilly, Novartis, Roche, Sanofi, Takeda, Union Chimique Belge, ViiV Healthcare, Johnson & Johnson, Medtronic, and Bristol-Myers Squibb. Company policies on what trials are shared vary and are available online, but most include all trials within certain date ranges after regulatory review and publication of results. GSK, for example, shares all clinical studies after 2000 after the medicine has been approved or development terminated and the study is accepted for publication except for rare diseases and healthcare products.

Investigators submit research a proposal to the platforms, which is first reviewed administratively for the availability of the trial(s) requested and completeness of the application. Next, the proposal is reviewed by a review panel. At ClinicalStudyDataRequest.com, this panel is comprised of an independent group of experts. At YODA and SOAR, this panel is made up of members of the YODA and SOAR platform. The panel then rejects or approves the proposal based on scientific merit and adequacy of the research design to achieve scientific objectives. A data sharing agreement is then created. Details of approved proposals with data sharing agreements in place are available on the ClinicalStudyDataRequest and YODA websites, and were obtained directly from SOAR. Each platform requires investigators to report any resulting publications.

We reviewed all proposals with data use agreements in place since inception of each platform (first in 2013) and December 31, 2015, the characteristics of accepted proposals, and reported publications. We classified the main objective of the analysis based on review of the analytic plan and study design.

Address for correspondence: Ann Marie Navar, MD, PhD; Duke Clinical Research Institute, 2400 Pratt St. #7521, Durham, NC, USA 27705; Tel: 919-668-8666; Fax: 919-668-7089; navar006@dm.duke.edu.

## Results

A total of 3140 clinical trials were available in the platforms; 44.3% were phase 3, 23.7% phase 1, 18.1% phase 2, and 13.9% phase 4. Of the 234 proposals submitted, 177 had been processed and met initial requirements (e.g. requested data available, proposal complete). The review panel rejected 12 of 177 (6.8%), 10 were under review, and 4 had been withdrawn. Of the 154 proposals approved, data sharing agreements were completed for 113, including requests from 17 countries, most from the United States (n=61, 54.0%). Most studies were not directly funded (n=77, 68.1%).

The median number of trials requested by each proposal was 2 (range 1–59, inter-quartile range 1–6). Most proposals requested trials from a single sponsor, only 29 proposals (8%) requested data from trials conducted by more than one sponsor. Only 505 unique trials (15.7% of all available trials), including 349 phase 3 trials (25.1% of all phase 3 trials). Analytic goals of these proposals varied. Secondary analyses of a trial's treatment effect were most common (n=50, 44.3%), followed by analyses of the disease state itself (n=31, 27.4%, Table). Validation of the study primary endpoint was rare (n=5, 4.4%). Only one proposal led to publication, a re-analysis of the effectiveness of paroxetine and imiprimane in adolescents.[5]

## Discussion

Although over 3000 trials are available to investigators through open data platforms, only 15.7% have been requested by a limited number of researchers. Most proposals focused on non-prespecified subgroups or predictors of response rather than validation of study results.

Reasons for underutilization of clinical trials data may include lack of knowledge about these resources, possibly due to lack of publication of results from initial studies, or lack of funding to support the analyses. Incentive for validation studies may be limited as confirmatory re-analyses are less likely to be published. However, the one publication using these data was a validation study that found contradictory results from the initial article.[5]

This study has limitations. We focused on three platforms and were unable to obtain data from individual companies that share data such as Merck or Pfizer. Next, details on rejected studies were unavailable. Certain information in the proposals may be incomplete, such as funding or specific analytic plans. There may be a lag between publication of a report and posting on the platform website, and we did not independently search for publications.

Platforms designed to provide access to individual patient data have developed in response to calls for increased transparency of clinical trial data. Yet early use has been limited. Availability of shared clinical trial data should be promoted to researchers and use of individual patient data for validation studies encouraged.

## References

1. Krumholz HM, Peterson ED. Open access to clinical trials data. JAMA. 2014; 312(10):1002–1003. [PubMed: 25203080]
2. The YODA Project. [Accessed January 26, 2016] The YODA Project web site. yoda.yale.edu

3. ClinicalStudyDataRequest.com. [Accessed January 26, 2016] ClinicalStudyDataRequest.com web site

4. Pencina MJ, Louzao DM, McCourt BJ, et al. Supporting open access to clinical trials data for researchers: The Duke Clinical Research Institute–Bristol-Myers Squibb Supporting Open Access to Researchers Initiative. Am Heart J. 2016; 172:64–9. [PubMed: 26856217]

5. Le Noury J, Nardo JM, Healy D, Jureidini J, Raven M, Tufanaru C, Abi-Jaoude E. Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. BMJ. 2015; 351:h4320. [PubMed: 26376805]

**Table**

Analytic Goal of Planned Studies with Representative Examples

| Analytic Focus | N | Representative Example |
|---|---|---|
| Secondary analysis of treatment effect | 50 | |
| Predictors of response | 18 | Can Machine Learning Algorithms Using General Labs Predict Clinical Remission and Mucosal Healing for Patients on Vedolizumab? |
| Subgroup analysis | 10 | Low-dose vs. Standard-dose Unfractionated Heparin for Percutaneous Coronary Intervention in Acute Coronary Syndromes Treated with Fondaparinux in Patients with Peripheral Arterial Disease |
| Safety/side effect prediction | 11 | Review of Fever and Febrile Convulsion Rates in Children after Trivalent Influenza Vaccine |
| Pharmacokinetic evaluation | 6 | Development of Population Pharmacokinetic/Pharmacodynamic Models for Pazopanib within the EuroTARGET Project |
| Impact on biomarker | 2 | The Effects of Maraviroc versus Efavirenz in Combination with Zidovudine/Abacavir on the CD4/CD8 Ratio in Treatment-naïve HIV-infected Individuals |
| Off-target treatment effect | 2 | Efficacy of Belimumab on Mucocutaneous Involvement in Patients with Systemic Lupus Erythematosus (MUCOBEL) |
| Duration of response | 1 | Statistical Methods for Estimating the Duration of Protection of Malaria Vaccines: Secondary Analysis of Data from the RTSS/AS02A Trial in Mozambique |
| Secondary disease state analysis | 31 | |
| Predictors of disease progression or severity | 11 | Incidence Rate and Risk Factors Associated with Hepatocellular Carcinoma Development in Patients with Chronic Hepatitis C and Severe Thrombocytopenia: the Lesson from ENABLE Studies |
| Complications of disease or procedure | 6 | Relationship of Activated Clotting Time and Bleeding-related Outcomes in the FUTURA OASIS-8 Trial |
| Alternate outcome measures/disease classification schemes | 10 | Developing New Outcome Measures for SLE by De-convoluting Available Datasets |
| Disease characterization | 4 | Cytokine Patterns in Anti-viral Disease Vaccines: A Knowledge-driven Approach |
| Placebo or placebo effect | 4 | Predictors of Placebo Response in Pediatric Patients with Generalized Anxiety Disorder |
| Re-evaluation of primary endpoint in study population | 5 | COPERNICUS Re-analysis |
| Meta-analysis of treatment effect | 12 | Assessing the Comparative Efficacy of Therapeutics for Chronic Obstructive Pulmonary Disease: A Meta-analysis of Randomized Trials |
| Statistical or trial methodology | 11 | Assessing and Reporting Heterogeneity of Treatment Effect in Randomized Clinical Trials |

Abbreviations: COPERNICUS, Carvedilol Prospective Randomized Cumulative Survival; ENABLE, Eltrombopag Therapy for HCV-Related Thrombocytopenia; EuroTARGET, Targeted Therapy in Renal Cell Cancer: Genetic and Tumour Related Biomarkers for Response and Toxicity; FUTURE OASIS-8, Fondaparinux Trial With Unfractionated Heparin During Revascularization in Acute Coronary Syndromes; HIV, human immunodeficiency virus; SLE, systemic lupus erythematosus