JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY
© 2017 BY THE AMERICAN COLLEGE OF CARDIOLOGY FOUNDATION
PUBLISHED BY ELSEVIER

VOL. 70, NO. 24, 2017
ISSN 0735-1097/$36.00
https://doi.org/10.1016/j.jacc.2017.10.037

THE PRESENT AND FUTURE

TECHNOLOGY CORNER

Data Sharing and Cardiology

Platforms and Possibilities

Pranammya Dey,[a] Joseph S. Ross, MD, MHS,[b,c,d] Jessica D. Ritchie, MPH,[d] Nihar R. Desai, MD, MPH,[a,d] Sanjeev P. Bhavnani, MD,[e] Harlan M. Krumholz, MD, SM[a,c,d]

ABSTRACT

Sharing deidentified patient-level research data presents immense opportunities to all stakeholders involved in cardiology research and practice. Sharing data encourages the use of existing data for knowledge generation to improve practice, while also allowing for validation of disseminated research. In this review, we discuss key initiatives and platforms that have helped to accelerate progress toward greater sharing of data. These efforts are being prompted by government, universities, philanthropic sponsors of research, major industry players, and collaborations among some of these entities. As data sharing becomes a more common expectation, policy changes will be required to encourage and assist data generators with the process of sharing the data they create. Patients also will need access to their own data and to be empowered to share those data with researchers. Although medicine still lags behind other fields in achieving data sharing's full potential, cardiology research has the potential to lead the way. (J Am Coll Cardiol 2017;70:3018–25)

Data sharing in cardiology has progressed markedly in recent years. A 2016 study found patient-level data from 1 in 4 large cardiovascular trials conducted by major pharmaceutical companies to be available for sharing with outside investigators (1). In this paper, we examine the rationale for sharing clinical research data and discuss the major data sharing initiatives and platforms that are influencing cardiology research. We also present examples of how data sharing fits into the broader open-science movement and ultimately affects clinical care. This is particularly

From the [a]Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut; [b]Section of General Internal Medicine, Department of Medicine, Yale School of Medicine, New Haven, Connecticut; [c]Department of Health Policy and Management, Yale School of Public Health, New Haven, Connecticut; [d]Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, Connecticut; and the [e]Division of Cardiology, Scripps Clinic and Research Foundation, San Diego, California. Dr. Desai is supported by grant K12 HS023000-03 from the Agency for Healthcare Research and Quality, Rockville, Maryland. Drs. Ross, Desai, Krumholz, and Ms. Ritchie have received research support through Yale University from Janssen, the Pharmaceutical Companies of Johnson & Johnson (New Brunswick, New Jersey), and Medtronic (Minneapolis, Minneapolis) to develop methods of clinical trial data sharing. Drs. Ross and Desai, and Ms. Ritchie have received research support from the Blue Cross Blue Shield Association (Chicago, Illinois) to better understand medical technology evidence generation. Drs. Desai, Ross, and Krumholz have received research support from the Centers for Medicare & Medicaid Services (Baltimore, Maryland) to develop and maintain hospital performance measures that are used for public reporting. Drs. Ross and Krumholz have received research support, through Yale University, from Medtronic, Inc., and the Food and Drug Administration (Silver Spring, Maryland) to develop methods for post-market surveillance of medical devices. Dr. Ross has received support from the Laura and John Arnold Foundation (Houston, Texas) to support the Collaboration on Research Integrity and Transparency at Yale. Dr. Ross and Ms. Ritchie have received research support from the Food and Drug Administration to establish the Yale-Mayo Center for Excellence in Regulatory Science and Innovation. Dr. Krumholz is a recipient of research agreements from Medtronic, through Yale, to develop methods of clinical trial data sharing. Dr. Krumholz chairs a cardiac scientific advisory board for UnitedHealth; is a participant/participant representative of the IBM Watson Health Life Sciences Board; is a member of the advisory board for Element Science and the Physician Advisory Board for Aetna; and is the founder of Hugo, a personal health information platform. Dr. Bhavnani serves on the advisory boards of iVEDIX, WellSeek, and Misceo in scientific roles (non-remunerated); has received on behalf of Scripps Health a research and educational grant from the Qualcomm Foundation (San Diego, California). Mr. Dey has reported that he has no relationships relevant to the contents of this paper to disclose.

Manuscript received July 3, 2017; revised manuscript received October 16, 2017, accepted October 17, 2017.

Listen to this manuscript's audio summary by *JACC* Editor-in-Chief Dr. Valentin Fuster.

important in cardiology, given its history of generating evidence for knowledge creation and secondary data analysis.

BACKGROUND

Several articles have argued that medicine would benefit from greater data sharing (2–5). Two central observations set the rationale for this new research paradigm.

1. Much data generated in clinical trials are kept out of public view. A recent study showed that nonpublication of clinical trials hovers around 50% across major academic institutions in the United States (6). The studies are not missing at random. Research shows that trials are more likely to be published if they yield positive or statistically significant outcomes (7). Withholding data, positive or negative, can cause harm. In a famous example, Merck held unpublished data showing that Vioxx (rofecoxib) likely increased the risk of acute myocardial infarction; it nevertheless took years for the drug to be withdrawn from the market, exposing numerous patients to potentially unnecessary risk (2).
2. Clinical trial data are often used inefficiently, with little opportunity for independent validation. Even when clinical trial results are published, the underlying patient-level data often remain unavailable. Many questions that could be asked using the data remain unaddressed, including secondary research questions and examination of rare outcomes not reported in the main publications. As a result, the expense and effort of creating the data resource may produce suboptimal yield (8). In addition, when independent scientists cannot view and analyze the data, they cannot verify and replicate the results (9).

Data sharing in medicine lags behind that found in other scientific disciplines. Physicists can access shared data from the Large Hadron Collider, astronomers from the Hubble Space Telescope, and geneticists from the Human Genome Project (10). In contrast, clinician-researchers often cannot answer investigative questions because they lack access to existing clinical trial data.

To address this concern, the Institute of Medicine (now the National Academy of Medicine) in 2015 called for researchers to share the "full analyzable data set with metadata" within 18 months of finishing a study, thereby fostering a culture in which data sharing is the standard (11). Senator Elizabeth Warren has called for making data sharing "a condition of publication in major medical journals" (12). Former Vice President Joe Biden endorsed data sharing in the Cancer Moonshot initiative, and the 21st Century Cures Act, which funds the Moonshot, empowers the Director of the National Institutes of Health (NIH) to require sharing of data from NIH-supported research (13,14). Through groups such as the Academic Research Organization Consortium for Continuing Evaluation of Scientific Studies–Cardiovascular, trialists are coming together in advocacy of sharing and are calling for standards (15). Thus, the move toward the sharing of clinical trial data is advancing.

INITIATIVES FOR DATA SHARING AND REPORTING OF RESULTS

Several organizations have sought to promote clinical trial data sharing, even to require it as a condition for funding. The reporting of results is also increasingly being encouraged. We explain these processes and highlight notable examples in the following section.

RESULTS REPORTING: FOOD AND DRUG ADMINISTRATION AMENDMENTS ACT OF 2007 AND NIH. Unlike full data sharing, which involves releasing the underlying deidentified patient-level data, results reporting refers to releasing summary results from clinical trials. The Food and Drug Administration (FDA) Amendments Act increased the scope of requirements for clinical trial sponsors of FDA-regulated products to register studies and report results at ClinicalTrials.gov (16). The Department of Health and Human Services Final Ruling on the FDA Amendments Act, effective January 2017, includes (among other features) a results-reporting requirement for phase II and phase III trials of products that have not gained approval (17).

This policy regarding FDA-regulated products complements new guidance from the NIH. As of January 2017, the NIH expects all trials that it partly or fully funds to be registered and to report summary results at ClinicalTrials.gov (18). Because many trial results previously went unreported, this new policy clarifies existing statutory ambiguities so that researchers and funders know what specific information to submit for compliance (16). The NIH will also withhold funding for clinical trials from institutions that fail to meet registry and reporting requirements, with an option to subject researchers and trial sponsors to monetary penalties (19). The new NIH policy requires results reporting for clinical trials at all stages, including phase I safety trials, and submission of original protocols and statistical analysis plans (17).

ABBREVIATIONS AND ACRONYMS

BioLINCC = Biologic Specimen and Data Repository Information Coordinating Center

EHR = electronic health record

FDA = Food and Drug Administration

ICMJE = International Committee of Medical Journal Editors

NHLBI = National Heart, Lung, and Blood Institute

NIH = National Institutes of Health

PCORI = Patient-Centered Outcomes Research Institute

PMI = Precision Medicine Initiative

S4S = Sync for Science

YODA = Yale University Open Data Access