



Published in final edited form as:

*Circulation*. 2018 February 27; 137(9): 938–947. doi:10.1161/CIRCULATIONAHA.117.031883.

## Availability and Use of Shared Data from Cardiometabolic Clinical Trials

Muthiah Vaduganathan, MD MPH<sup>1</sup>, Amulya Nagarur, MD<sup>2</sup>, Arman Qamar, MD<sup>1</sup>, Ravi B. Patel, MD<sup>3</sup>, Ann Marie Navar, MD PhD<sup>4</sup>, Eric D. Peterson, MD MPH<sup>4</sup>, Deepak L. Bhatt, MD MPH<sup>1</sup>, Gregg C. Fonarow, MD<sup>5</sup>, Clyde W. Yancy, MD MSc<sup>3</sup>, and Javed Butler, MD MPH MBA<sup>6</sup>

<sup>1</sup>Brigham and Women's Hospital Heart & Vascular Center and Harvard Medical School, Boston, MA, USA

<sup>2</sup>Department of Medicine, Massachusetts General Hospital, Boston, MA, USA

<sup>3</sup>Division of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

<sup>4</sup>Duke Clinical Research Institute and Division of Cardiology, Duke University Medical Center, Durham, NC, USA

<sup>5</sup>Ahmanson-UCLA Cardiomyopathy Center, University of California Los Angeles, Los Angeles, CA, USA

<sup>6</sup>Division of Cardiology, Stony Brook University, Stony Brook, NY, USA

### Abstract

**Background**—Sharing of patient-level clinical trial data has been widely endorsed. Little is known about how extensively these data have been utilized for cardiometabolic diseases. We sought to evaluate the availability and use of shared data from cardiometabolic clinical trials.

**Methods**—We extracted data from [ClinicalStudyDataRequest.com](https://clinicalstudydatarequest.com) (CSDR), a large, multi-sponsor, data sharing platform hosting individual patient-level data from completed studies sponsored by 13 pharmaceutical companies.

**Results**—From January 2013 to May 2017, the platform had data from 3,374 clinical trials of which 537 (16%) evaluated cardiometabolic therapeutics (phase 1, 36%; phase 2, 17%; phase 2/3, 1%; phase 3, 42%; phase 4, 4%). These covered 74 therapies and 398,925 patients. Diabetes mellitus (60%) and hypertension (15%) were the most common study topics. Median time from study completion to data availability was 79 months. As of May 2017, CSDR had received 318 submitted proposals, of which 163 had signed data sharing agreements. Thirty of these proposals were related to cardiometabolic therapies and requested data from 79 unique studies (15% of all trials, 29% of phase 3/4 trials). Most data requesters of cardiometabolic clinical trial data were from academic centers in North America and Western Europe (96%), and half the proposals were unfunded. Most proposals were for secondary hypothesis-generating questions with only 1

proposed reanalysis of the original study primary hypothesis. To date, 3 peer-reviewed papers have been published after a median of 19 (interquartile range 9 to 32) months from the data sharing agreement.

**Conclusions**—Despite availability of data from over 500 cardiometabolic trials in a multi-sponsor, data sharing platform, only 15% of these trials, and 29% of phase 3/4 trials, have been accessed by investigators thus far and a negligible minority of analyses have reached publication.

### Keywords

clinical trial; outcome and process assessment; research methodology

### Journal Subject Terms

Clinical Studies; Quality and Outcomes

---

## INTRODUCTION

Clinical trials represent the gold standard to test emerging cardiometabolic therapeutics and form the basis of most regulatory approvals.<sup>1</sup> These studies are increasingly becoming larger, costlier, and more complex to conduct.<sup>2</sup> Sharing participant-level data after trial completion is proposed as a mechanism to broaden their scientific impact and maximize return on investment. Responsible data sharing promises to enhance the individual contributions of research participants and may confirm study reproducibility and validity. Efforts to advance data sharing have accrued a broad base of support from governmental officials,<sup>3</sup> journal editors,<sup>4, 5</sup> charitable foundations,<sup>6, 7</sup> regulatory bodies,<sup>8</sup> the pharmaceutical industry,<sup>9</sup> and clinical trialists.<sup>10</sup> ACCESS CV (Academic Research Organization Consortium for Continuing Evaluation of Scientific Studies — Cardiovascular),<sup>11</sup> a 31-member panel of cardiovascular clinical trialists, was recently formed to operationalize these calls for data sharing in cardiology.

Despite this progress to improve data transparency and access, effective implementation and the mechanics of data sharing require further attention. Several existing data sharing initiatives have provided early insights into patterns of utilization of aggregate data across medical disciplines (Table 1),<sup>8, 12–16</sup> and suggest that these data are underutilized. Specialty-specific utilization, including use of data from cardiometabolic trials, is less clear. Of the pioneering industry-based data sharing platforms, [ClinicalStudyDataRequest.com](https://www.clinicalstudydatarequest.com) (CSDR)<sup>17</sup> is one of the oldest and largest, and hosts the greatest number of cardiometabolic trials from various industry sponsors. We provide an assessment of the availability and use of shared data from cardiometabolic clinical trials hosted by the CSDR multi-sponsor data repository. We hypothesized that similar to the experience with general medical clinical trials, demand for access to cardiometabolic trials would be limited.

## METHODS

### ClinicalStudyDataRequest.com

We extracted data from CSDR,<sup>17</sup> a publicly-available, online, multi-sponsor, data sharing platform, which is hosted by ideaPoint, Inc. (Boston, MA) and has been available since January 2013. Thirteen pharmaceutical companies, including Astellas, Bayer, Boehringer Ingelheim, Daiichi Sankyo, Eisai, GlaxoSmithKline, Lilly, Novartis, Roche, Sanofi, Takeda, UCB, and ViiV Healthcare, deposit deidentified, individual patient-level data into CSDR. Key eligibility criteria for listing, exempted or unshared studies, conditions for data access, and details regarding datasets and metadata, are summarized in Table 2 and expanded online.<sup>17</sup>

### Procedures for Data Requests

The procedures for data requests have been described previously<sup>18</sup> and summarized in Figure 1. In brief, data requestors first submit a proposal related to one or more hosted studies via a secure, web-based portal. Specific enquiries can also be submitted to sponsors regarding the availability of data from studies not hosted by CSDR.

Next, an independent review panel, managed by the Wellcome Trust (as of March 2015) reviews each proposal for completeness, scientific merit, the ability of the research plan to achieve the stated aims, and the qualifications and conflicts of interest of research team. The panel then reaches a decision and the data requesters sign a data use agreement.

Access to deidentified data and meta-data is then granted through a secure enclave with in-built SAS (SAS Institute, Cary, NC) and R (R Foundation) statistical software for 12–24 months. Up to 5 statistical software licenses to analyze shared data are supported by the involved study sponsors. The private user interface is password-protected and only accessible to data requesters. Data elements are fully anonymized with technical safeguards in place to prevent researchers from downloading original patient-level data. The analysis system further allows data requesters to combine study data from multiple sponsors.

The entire data sharing process via CSDR is tracked and transparently displayed online. Number of requests for data access of listed and unlisted studies, together with final decisions from the independent review panel, are presented. A lay summary and key study details of each approved proposal are made publicly available after data use agreements are signed. Data requesters are expected to post a summary of their research plan on a registry or website within 1 year, and to submit their findings for peer-reviewed publication. Protocols and expectations for reviewing manuscripts prior to submission are sponsor-specific. After publication, the citation and statistical analysis plan are also posted on CSDR.

### Data Extraction, Linked Sources, and Statistical Methods

We queried ClinicalTrials.gov to obtain a crude estimate of total trial portfolio per sponsor including number of unshared trials. We identified studies registered with ClinicalTrials.gov supported by each sponsor by employing the following limits: updated through May 2017, interventional study design, closed enrollment, and trial phase (II–IV).

We detailed category of study identified by each sponsor. We identified cardiometabolic trials as those evaluating therapies targeting established cardiovascular disease (atrial fibrillation/atrial flutter, coronary artery disease, heart failure, peripheral vascular disease, stroke, venous thromboembolism) or cardiometabolic risk factors (diabetes mellitus, dyslipidemia, hypertension, metabolic syndrome, obesity). Trial size, phase, drug or drug combination, and date of data availability were also documented. We then linked each cardiometabolic study with its corresponding ClinicalTrials.gov and sponsor entry to estimate time from study completion to data availability.

We reviewed all approved proposals from the inception of the platform through May 2017. Key characteristics of lead data requesters were detailed by reviewing their CSDR affiliation information and online faculty profile. For each cardiometabolic proposal, we applied the search terms of the first and last name of the submitting investigator and key words from the title of the proposal in PubMed/MEDLINE to determine publication status. All queries were performed through May 2017. Only publicly-available, trial-level data were accessed and thus this study was not submitted for institutional review board approval and individual trial participants were not contacted for informed consent.

## RESULTS

### Shared Data Availability on CSDR

From January 2013 to May 2017, the platform had data from 3,374 clinical trials of which 537 (16%) evaluated cardiometabolic therapeutics (phase 1, 36%; phase 2, 17%; phase 2/3, 1%; phase 3, 42%; phase 4, 4%). These covered 74 therapies and 398,925 patients (Figure 2). Diabetes mellitus (60%) and hypertension (15%) were the most common study topics (Figure 3).

Median time from study completion to data availability was 79 (interquartile range 52 to 108) months with a range of 5 to 211 months (Figure 4). When examining only trials made available in 2016 and 2017, time to data availability was slightly shorter (median 65 [interquartile range 40 to 86] months; range 5 to 187 months).

### Sponsor-Specific Data Sharing

Most industry sponsors required the original study to be accepted or published at the time of data sharing, and sponsors variably required study drugs to have been approved by either or both the Food and Drug Administration and the European Medicines Agency (Table 2). Common reasons for study exemption include any factors that posed challenges to fully anonymizing data (small, single-center experiences, studies of rare diseases, genomic data), non-English studies, practical constraints related to trial size, certain legal provisions, or threats to commercial/intellectual property. Sponsors shared between 1% and 95% of their total estimated portfolio from ClinicalTrials.gov on the CSDR platform. GlaxoSmithKline shared data from the highest number of cardiometabolic trials (n=234) and total trials (n=2027, 77% of the total estimated trial portfolio).

## Cardiometabolic Proposals and Data Use

As of May 2017, CSDR had received 318 submitted proposals, of which 235 met initial processing requirements. Of these, 5 remain in process, 6 were withdrawn by the requestor, and 26 were rejected or required revision. As such, 198 were approved (with or without conditions) and 163 (51%) had signed data use agreements.

Thirty of these proposals were related to cardiometabolic therapies and requested data from 79 unique studies (representing 15% of available cardiometabolic trials; Figure 2). These studies were primarily phase 3 (n=57 trials, 25% of available phase 3 cardiometabolic studies) and phase 4 (n=14 trials, 61% of available phase 4 cardiometabolic studies). The median number of trials requested by each proposal was 1 (range 1 to 45). Most proposals requested data from a single sponsor, while 5 requested data from more than 1. Five studies were requested more than once (range 2 to 6 times requested). The most common topics of requested studies were diabetes (72%), venous thromboembolism (8%), and atrial fibrillation/atrial flutter (8%). Half of the proposals did not specify a funding source for the analysis.

Most proposals focused on statistical or research methodology (n=6), risk prediction (n=6), or meta-analyses/systematic reviews (n=4). Other proposal objectives, including translation of clinical trial findings to real-world settings (n=3), subgroup analyses (n=3), predictors of response (n=2), or disease characterization (n=2), were less common. Only 1 proposal intended to reanalyze the original study primary hypothesis.

## Characteristics of Data Requesters

Four investigators were the lead researchers on more than 1 proposal, such that there were 26 unique data requesters of cardiometabolic studies. The majority (85%) were men. Data access requests spanned 6 countries. Twenty five of the 26 investigators were based in North America or Western Europe and were primarily affiliated with an academic medical center. A single data requester worked for a pharmaceutical company. Thirty-eight percent of data requesters were specialists in cardiology, hypertension, or diabetes, while the remaining worked in areas outside cardiometabolic health, including epidemiology, statistics, health services, and public health.

## Publication Status

To date, 3 (10%) of the 30 proposals had accompanying peer-reviewed publications<sup>19–21</sup> at a median of 19 (interquartile range 9 to 32) months from completion of the data use agreement. Kent et al. assessed the relationship between baseline risk and absolute treatment effects across 32 large trials (only 1 of which was requested through CSDR).<sup>20</sup> Hilkens et al. performed an exploratory analysis defining risk of intracerebral hemorrhage with varying systemic blood pressures in patients with recent ischemic stroke.<sup>19</sup> Walker et al. conducted a systematic review of 12 studies (only 1 of which was requested through CSDR) on the efficacy and safety of dipeptidyl peptidase-4 inhibitors in patients with diabetes mellitus and chronic kidney disease.<sup>21</sup>

## DISCUSSION

This interim analysis of cardiometabolic clinical trials hosted on a large, multi-sponsor data sharing platform highlights several important findings: 1) although individual-patient level data from >500 cardiometabolic trials are already available, only 15% have been accessed to date (~4.5 years after the platform's inception); 2) requests for data access were commonly unfunded and come from a small number of investigators of restricted demographics (the vast majority from North America/Western Europe and academic medical centers); 3) data access requests often focus on new hypothesis-generating questions, and rarely attempt to validate or refute original study findings; and 4) few publications have resulted a median of >18 months after data access.

### Data Sharing in Cardiology

The pharmaceutical industry has already pioneered efforts to expand data access and transparency. Although CSDR represents the most comprehensive data sharing platform of industry-supported clinical trials, we estimate that only a fraction (~30%), varying significantly by sponsor, of total trial portfolios are available for access and only become accessible >5 years after trial completion, even for studies added in the last 2 years. A prior analysis consistently reported that 25% of large, industry-sponsored advanced-phase cardiovascular trials had available individual patient-level data.<sup>22</sup> Indeed, a recent systematic audit demonstrated that commitments and policies to individual-level patient data access are highly variable across major pharmaceutical companies.<sup>23</sup>

As we prepare for more widespread and routine data sharing on shorter timelines in cardiology,<sup>11</sup> examining these existing initiatives may allow anticipation of barriers to effective system implementation and shared data consumption. This analysis embedded within CSDR demonstrated relatively sparse utilization of these cardiometabolic clinical trial data, consistent with prior published reports of earlier analyses of shared general medical clinical trial data hosted on multiple industry-supported open-access platforms.<sup>13–16</sup> A previous report provided an overview of 3 data sharing platforms of industry-sponsored trials through the end of 2015 and demonstrated that only ~15% of studies had been requested.<sup>14</sup> In the current analysis, we specifically look at availability and use of cardiometabolic trials through mid-2017 on the largest of these platforms, CSDR, and found similar overall demand for access. We further linked each hosted cardiometabolic study with its corresponding ClinicalTrials.gov and sponsor entry (for key trial characteristics) and linked each approved proposal with a PubMed/MEDLINE query (to determine publication status).

The National Heart, Lung, and Blood Institute (NHLBI) data repository,<sup>12</sup> which is coordinated by Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC), has showed greater demand for data reuse of its large clinical trials and observational studies. Although the data repository only hosts data from 100 clinical trials, over 800 data requests have been received from 2000 to 2016, especially for larger, more recent, cardiovascular treatment and prevention trials.<sup>12</sup>

## Reasons for Data Requests

Consistent with prior experiences,<sup>13–16</sup> most requests for cardiometabolic trial data focused on new hypothesis-generating questions or exploratory analyses. Only 1 cardiometabolic proposal in CSDR requested data for reanalysis (of the COPERNICUS [Carvedilol Prospective Randomized Cumulative Survival] trial),<sup>24</sup> perhaps due to low perceived publication value of analyses confirming previously published clinical trial findings. Reanalysis of the original study hypotheses, which was infrequent in this CSDR experience, may theoretically validate<sup>25</sup> or refute the original trial's findings.<sup>26</sup> In a systematic review of 37 reanalyses of published studies, more than one third led to data interpretations that diverged from the original studies' conclusions.<sup>27</sup> In addition, platforms such as CSDR enable data requests from more than 1 sponsor, but multi-sponsor proposals were infrequent in our experience.

## Barriers to Data Sharing and Usability

Although the collective goals of data sharing are to advance science, maximize the return of patient participation, and ultimately improve public health, several practical considerations need to be addressed prior to realizing this promise and potential. Many hurdles remain in transforming existing platforms of data access into integrated systems that promote data usability, utility, and productivity.<sup>7, 28</sup>

Building and maintaining high-quality data repositories is cost- and resource-intensive. Costs incurred may depend on the specific data sharing model, the size and complexity of shared data, and the structure of the user interface. However, the following 4 elements that contribute to cost are likely common to any viable data sharing platform: 1) infrastructure and maintenance; 2) data standardization and quality control; 3) human resources for technical expertise and administration; and 4) opportunity costs and potential loss in investment to other research activities.<sup>29</sup> Upfront resource investment into building sustainable and comprehensive data sharing platforms with standardized data elements and user-friendly interfaces may enhance the quality, accessibility, and usability of shared data, but may be costly and financially untenable.

The actual costs of data sharing are difficult to estimate. Costs to support data sharing models that provide limited or open access to data stored on digital repositories, such as Dryad,<sup>30</sup> depend on the size of shared data, and may still be costly for larger datasets, such as genomic data,<sup>31</sup> due to overage fees incurred after a certain size limit. Costs and resources required for models facilitating extensive and comprehensive data sharing may be more substantial. Efforts to prepare data and meta-data for broad sharing in the NHLBI data repository were estimated as ranging from 85 to 350 full-time equivalent hours per study.<sup>12</sup> The initial costs required to establish an online data sharing system for 2 large cancer screening trials supported by the National Cancer Institute was estimated at ~\$300,000, with an additional ~\$26,000 needed for monthly support and maintenance.<sup>32</sup> The Alzheimer's Disease Neuroimaging Initiative, a disease-specific data sharing platform that maintains secure, standardized, and comprehensive data, also provided early estimates of the practical costs of data sharing. Data sharing efforts were estimated to account for 10–15% of the \$130 million total project costs and occupy 15% of the time of the primary project investigators.<sup>29</sup>



Whether the clinical trial data sharing enterprise should be financed by sponsors or data requesters remains to be determined.<sup>33</sup>

Secondary data analyses also necessitate funding and analytic resources, which may explain the demographic predilection of data requesters. Data requesters of cardiometabolic studies in CSDR were primarily from academic medical centers in North America or Western Europe. Consistently, 88% of data requests from the NHLBI data repository originated from the US or Canada.<sup>12</sup> This initial data utilization pattern supports general concerns regarding the preferential shared data use in high- compared with low-income countries.<sup>7</sup> We found that half the proposals disclosed no specific external funding, which may preclude timely completion of data analyses and hinder ultimate publication. Limited funding and support were commonly cited factors in a cross-sectional, web-based survey of BioLINCC users as reasons delaying completion of analyses and publication.<sup>34</sup>

The limited requests of industry-sponsored trials may be related to lack of knowledge of data availability by the general cardiovascular research community. Indeed, when data sharing efforts are actively advertised and promoted, interest in data access appears to be high. For instance, in the SPRINT (Systolic Blood Pressure Intervention Trial) Data Analysis Challenge,<sup>35</sup> which represented a collaboration between the *New England Journal of Medicine*, the NHLBI, and the SPRINT Data Coordinating Center, 143 complete entries were received from 26 countries over a short duration. It is encouraging that data requesters in our CSDR-based experience carried diverse backgrounds with nearly two-thirds in fields outside cardiometabolic health. These shared data sources represent important opportunities in the development of research careers across disciplines.

Although CSDR employs a “learned-intermediary”<sup>36</sup> or “gatekeeper”<sup>37</sup> model for data sharing, which leverages an independent review board that handles and transparently documents data access decisions, the optimal model(s) for data sharing are yet to be determined. Ongoing efforts to merge isolated shared data silos into an integrated, secure,<sup>38</sup> global clinical trial platform<sup>39</sup> with standardized data elements<sup>40</sup> are currently underway. Consortia, such as ACCESS CV, will need to tackle other unresolved issues including appropriate incentives and credit for data generators,<sup>41</sup> a reasonable timeline for proprietary data use prior to public release, and mechanisms and structures to ensure compliance.

### Study Limitations

This data sharing report is subject to certain limitations. There are a number of existing industry-based data sharing mechanisms including CSDR, the Supporting Open Access to Researchers initiative, the Yale Open Data Access project, and direct-to-sponsor models. We restricted our analysis to CSDR as other industry-sponsored platforms hosted few cardiometabolic trials. Only proposals with active data use agreements were accessible via CSDR, thus we were not able to analyze all data requests prior to screening and review. We did not have access to the timeline of the approval process within CSDR to better understand the efficiency of the system. Although we applied inclusive search criteria in PubMed/MEDLINE, it is possible we missed publications generated from these analyses. Our estimate of total trial portfolio using ClinicalTrials.gov was crude and may overestimate



eligible trials as the query was not limited specific time windows or drug approval status and did not account for unregistered trials or those registered elsewhere.

## Conclusions

Although data from over 500 cardiometabolic trials have been made available on a large, multi-sponsor data sharing platform, only 15% of these trials have been accessed by investigators thus far and few analyses have reached publication. Most proposals evaluated hypothesis-generating or exploratory aims, while validation studies were rarely conducted. Barriers to utilization, optimal timeframe to availability, and funding mechanisms for shared data of cardiometabolic clinical trials need clarification.

## Acknowledgments

None

### SOURCES OF FUNDING

None

### DISCLOSURES

**Drs. Muthiah Vaduganathan** and **Arman Qamar** are supported by the NHLBI T32 postdoctoral training grant (T32HL007604).

**Dr. Ann Marie Navar** has received research support from Sanofi, Regeneron, and Amgen and has served as a consultant or on the advisory board of Sanofi and Amgen. She is also supported by the NHLBI (K01 HL13341).

**Dr. Eric D. Peterson** has received research support from Janssen Scientific Affairs and Bayer.

**Dr. Deepak L. Bhatt** discloses the following relationships - Advisory Board: Cardax, Elsevier Practice Update Cardiology, Medscape Cardiology, Regado Biosciences; Board of Directors: Boston VA Research Institute, Society of Cardiovascular Patient Care; Chair: American Heart Association Quality Oversight Committee; Data Monitoring Committees: Cleveland Clinic, Duke Clinical Research Institute, Harvard Clinical Research Institute, Mayo Clinic, Mount Sinai School of Medicine, Population Health Research Institute; Honoraria: American College of Cardiology (Senior Associate Editor, Clinical Trials and News, ACC.org), Belvoir Publications (Editor in Chief, Harvard Heart Letter), Duke Clinical Research Institute (clinical trial steering committees), Harvard Clinical Research Institute (clinical trial steering committee), HMP Communications (Editor in Chief, Journal of Invasive Cardiology), Journal of the American College of Cardiology (Guest Editor; Associate Editor), Population Health Research Institute (clinical trial steering committee), Slack Publications (Chief Medical Editor, Cardiology Today's Intervention), Society of Cardiovascular Patient Care (Secretary/Treasurer), WebMD (CME steering committees); Other: Clinical Cardiology (Deputy Editor), NCDR-ACTION Registry Steering Committee (Chair), VA CART Research and Publications Committee (Chair); Research Funding: Amarin, Amgen, AstraZeneca, Bristol-Myers Squibb, Chiesi, Eisai, Ethicon, Forest Laboratories, Ironwood, Ischemix, Lilly, Medtronic, Pfizer, Roche, Sanofi Aventis, The Medicines Company; Royalties: Elsevier (Editor, Cardiovascular Intervention: A Companion to Braunwald's Heart Disease); Site Co-Investigator: Biotronik, Boston Scientific, St. Jude Medical (now Abbott); Trustee: American College of Cardiology; Unfunded Research: FlowCo, Merck, PLx Pharma, Takeda.

**Dr. Gregg C. Fonarow** reports significant consulting for Novartis, and modest consulting for Amgen, Janssen, Medtronic, and St Jude Medical.

**Dr. Javed Butler** has received research support from the NIH and European Union; and has been a consultant for Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Janssen, Novartis, Relypsa, ZS Pharma, Medtronic, Merck, and CVRx.

**Drs. Amulya Nagarur, Ravi B. Patel, and Clyde W. Yancy** report no relationships relevant to the contents of this manuscript to disclose.

## References

1. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. *JAMA*. 2014; 311:368–377. [PubMed: 24449315]
2. Butler J, Tahhan AS, Georgiopoulou VV, Kelkar A, Lee M, Khan B, Peterson E, Fonarow GC, Kalogeropoulos AP, Gheorghiade M. Trends in characteristics of cardiovascular clinical trials 2001–2012. *Am Heart J*. 2015; 170:263–272. [PubMed: 26299223]
3. Warren E. Strengthening research through data sharing. *N Engl J Med*. 2016; 375:401–3. [PubMed: 27518656]
4. Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, Fletcher J, Frizelle FA, Groves T, Haileamlak A, James A, Laine C, Peiperl L, Pinborg A, Sahni P, Wu S. Sharing clinical trial data—a proposal from the International Committee of Medical Journal Editors. *N Engl J Med*. 2016; 374:384–386. [PubMed: 26786954]
5. Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, Hong ST, Haileamlak A, Gollogly L, Godlee F, Frizelle FA, Florenzano F, Drazen JM, Bauchner H, Baethge C, Backus J. Data sharing statements for clinical trials - a requirement of the International Committee of Medical Journal Editors. *N Engl J Med*. 2017; 376:2277–2279. [PubMed: 28581902]
6. Wellcome Trust. Sharing public health data: a code of conduct. Accessed October 20<sup>th</sup>, 2017. <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Public-health-and-epidemiology/index.htm>
7. Pisani E, Aaby P, Breugelmans JG, Carr D, Groves T, Helinski M, Kamuya D, Kern S, Littler K, Marsh V, Mboup S, Merson L, Sankoh O, Serafini M, Schneider M, Schoenenberger V, Guerin PJ. Beyond open data: realising the health benefits of sharing data. *BMJ*. 2016; 355:i5295. [PubMed: 27758792]
8. Bonini S, Eichler HG, Wathion N, Rasi G. Transparency and the European Medicines Agency—sharing of clinical trial data. *N Engl J Med*. 2014; 371:2452–2455. [PubMed: 25539105]
9. Pharmaceutical Research and Manufacturers of America, European Federation of Pharmaceutical Industries and Associations. Principles for responsible clinical trial data sharing. Accessed October 20<sup>th</sup>, 2017. Published July 18<sup>th</sup>, 2013. <http://phrma.org/sites/default/files/pdf/PhRMAPrinciplesForResponsibleClinicalTrialDataSharing.pdf>
10. Rathi VK, Strait KM, Gross CP, Hrynaszkiewicz I, Joffe S, Krumholz HM, Dzara K, Ross JS. Predictors of clinical trial data sharing: exploratory analysis of a cross-sectional survey. *Trials*. 2014; 15:384. [PubMed: 25277128]
11. ACCESS CV. Patel MR, Armstrong PW, Bhatt DL, Braunwald E, Camm AJ, Fox KA, Harrington RA, Hiatt WR, James SK, Kirtane AJ, Leon MB, Lincoff AM, Mahaffey KW, Mauri L, Mehran R, Mehta SR, Montalescot G, Nicholls SJ, Perkovic V, Peterson ED, Pocock SJ, Roe MT, Sabatine MS, Sekeres M, Solomon SD, Steg G, Stone GW, Van de Werf F, Wallentin L, White HD, Gibson M. Sharing data from cardiovascular clinical trials—a proposal. *N Engl J Med*. 2016; 375:407–409. [PubMed: 27518659]
12. Coady SA, Mensah GA, Wagner EL, Goldfarb ME, Hitchcock DM, Giffen CA. Use of the National Heart, Lung, and Blood Institute data repository. *N Engl J Med*. 2017; 376:1849–1858. [PubMed: 28402243]
13. Krumholz HM, Waldstreicher J. The Yale Open Data Access (YODA) Project—a mechanism for data sharing. *N Engl J Med*. 2016; 375:403–405. [PubMed: 27518657]
14. Navar AM, Pencina MJ, Rymer JA, Louzao DM, Peterson ED. Use of open access platforms for clinical trial data. *JAMA*. 2016; 315:1283–1284. [PubMed: 27002452]
15. Strom BL, Buyse M, Hughes J, Knoppers BM. Data sharing, year 1—access to data from industry-sponsored clinical trials. *N Engl J Med*. 2014; 371:2052–2054. [PubMed: 25317745]
16. Strom BL, Buyse ME, Hughes J, Knoppers BM. Data sharing - is the juice worth the squeeze? *N Engl J Med*. 2016; 375:1608–1609. [PubMed: 27783903]
17. ideaPoint, Inc. ClinicalStudyDataRequest.com online data sharing platform. Accessed October 20<sup>th</sup>, 2017, <https://clinicalstudydatarequest.com/>

18. Nisen P, Rockhold F. Access to patient-level data from GlaxoSmithKline clinical trials. *N Engl J Med*. 2013; 369:475–478. [PubMed: 23902490]
19. Hilken NA, Greving JP, Algra A, Klijn CJ. Blood pressure levels and the risk of intracerebral hemorrhage after ischemic stroke. *Neurology*. 2017; 88:177–181. [PubMed: 27927936]
20. Kent DM, Nelson J, Dahabreh IJ, Rothwell PM, Altman DG, Hayward RA. Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials. *Int J Epidemiol*. 2016; 45:2075–2088. [PubMed: 27375287]
21. Walker SR, Komenda P, Khojah S, Al-Tuwaijri W, MacDonald K, Hiebert B, Tangri N, Nadurak SWD, Ferguson TW, Rigatto C, Tangri N. Dipeptidyl peptidase-4 inhibitors in chronic kidney disease: a systematic review of randomized clinical trials. *Nephron*. 2017; 136:85–94. [PubMed: 28178698]
22. Murugiah K, Ritchie JD, Desai NR, Ross JS, Krumholz HM. Availability of clinical trial data from industry-sponsored cardiovascular trials. *J Am Heart Assoc*. 2016; 5:e003307. [PubMed: 27098969]
23. Goldacre B, Lane S, Mahtani KR, Heneghan C, Onakpoya I, Bushfield I, Smeeth L. Pharmaceutical companies' policies on access to trial data, results, and methods: audit study. *BMJ*. 2017; 358:j3334. [PubMed: 28747301]
24. Packer M, Coats AJ, Fowler MB, Katus HA, Krum H, Mohacsi P, Rouleau JL, Tendera M, Castaigne A, Roecker EB, Schultz MK, DeMets DL. Effect of carvedilol on survival in severe chronic heart failure. *N Engl J Med*. 2001; 344:1651–1658. [PubMed: 11386263]
25. Patel A, Pieper K, Myburgh JA, Perkovic V, Finfer S, Yang Q, Li Q, Billot L. Reanalysis of the Crystalloid versus Hydroxyethyl Starch Trial (CHEST). *N Engl J Med*. 2017; 377:298–300. [PubMed: 28723325]
26. Christakis DA, Zimmerman FJ. Rethinking reanalysis. *JAMA*. 2013; 310:2499–2500. [PubMed: 24346985]
27. Ebrahim S, Sohani ZN, Montoya L, Agarwal A, Thorlund K, Mills EJ, Ioannidis JP. Reanalyses of randomized clinical trial data. *JAMA*. 2014; 312:1024–1032. [PubMed: 25203082]
28. Merson L, Gaye O, Guerin PJ. Avoiding data dumpsters—toward equitable and useful data sharing. *N Engl J Med*. 2016; 374:2414–2415. [PubMed: 27168351]
29. Wilhelm EE, Oster E, Shoulson I. Approaches and costs for sharing clinical research data. *JAMA*. 2014; 311:1201–1202. [PubMed: 24556937]
30. Goodhill GJ. Open access: practical costs of data sharing. *Nature*. 2014; 509:33.
31. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, Staudt LM. Toward a shared vision for cancer genomic data. *N Engl J Med*. 2016; 375:1109–1112. [PubMed: 27653561]
32. Zhu CS, Pinsky PF, Moler JE, Kukwa A, Mabie J, Rathmell JM, Riley T, Prorok PC, Berg CD. Data sharing in clinical trials: An experience with two large cancer screening trials. *PLoS Med*. 2017; 14:e1002304. [PubMed: 28542201]
33. Devereaux PJ, Guyatt G, Gerstein H, Connolly S, Yusuf S. Toward fairness in data sharing. *N Engl J Med*. 2016; 375:405–407. [PubMed: 27518658]
34. Ross JS, Ritchie JD, Finn E, Desai NR, Lehman RL, Krumholz HM, Gross CP. Data sharing through an NIH central database repository: a cross-sectional survey of BioLINCC users. *BMJ Open*. 2016; 6:e012769.
35. Burns NS, Miller PW. Learning what we didn't know - the SPRINT Data Analysis Challenge. *N Engl J Med*. 2017; 376:2205–2207. [PubMed: 28445656]
36. Mello MM, Francer JK, Wilenzick M, Teden P, Bierer BE, Barnes M. Preparing for responsible sharing of clinical trial data. *N Engl J Med*. 2013; 369:1651–1658. [PubMed: 24144394]
37. Rockhold F, Nisen P, Freeman A. Data sharing at a crossroads. *N Engl J Med*. 2016; 375:1115–1117. [PubMed: 27653563]
38. Taitsman JK, Grimm CM, Agrawal S. Protecting patient privacy and data security. *N Engl J Med*. 2013; 368:977–979. [PubMed: 23444980]
39. Bierer BE, Li R, Barnes M, Sim I. A global, neutral platform for sharing trial data. *N Engl J Med*. 2016; 374:2411–2413. [PubMed: 27168194]

40. Kush R, Goldman M. Fostering responsible data sharing through standards. *N Engl J Med.* 2014; 370:2163–2165. [PubMed: 24897080]
41. Bierer BE, Crosas M, Pierce HH. Data authorship as an incentive to data sharing. *N Engl J Med.* 2017; 376:1684–1687. [PubMed: 28402238]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

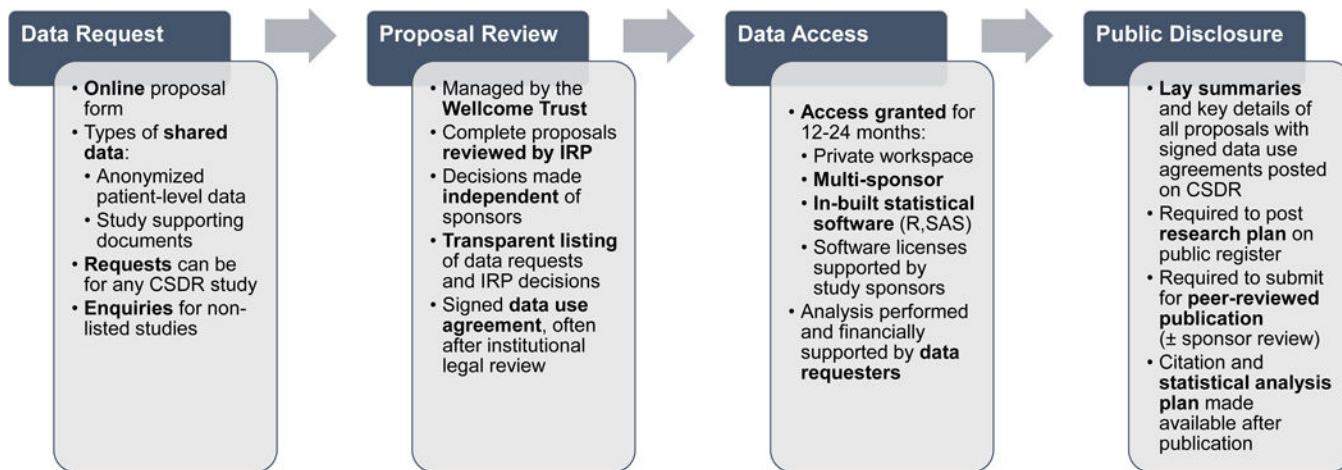
## CLINICAL PERSPECTIVES

### What is new?

- Data from over 500 cardiometabolic trials have been made available on a large, multi-sponsor data sharing platform.
- Median time from study completion to data availability was over 6 years.
- Most data requesters of cardiometabolic clinical trial data were from academic centers in North America and Western Europe, and half the proposals were unfunded.
- Only 15% of these trials have been accessed by investigators thus far and few findings have reached publication.
- Most requests for shared data access focus on new hypothesis-generating questions rather than validation of original study findings.

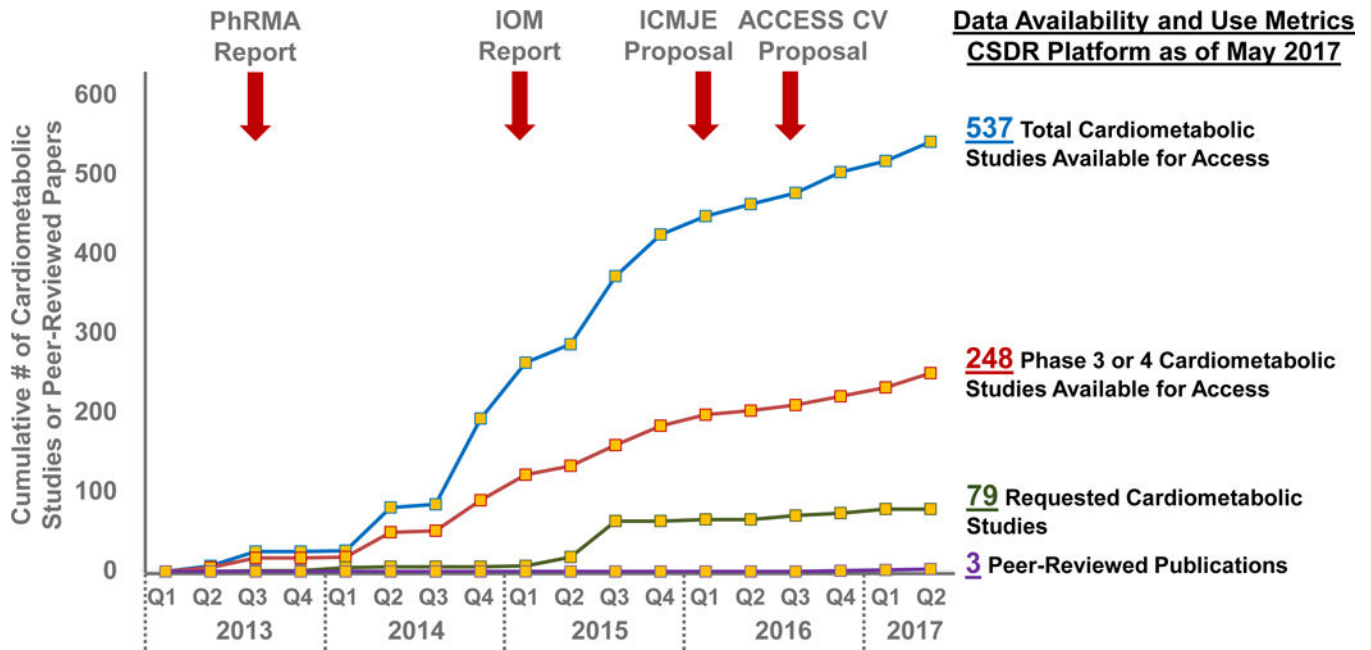
### What are the clinical implications?

- As we prepare for more widespread data sharing in cardiology, this interim look at an existing data sharing initiative may allow anticipation of barriers to effective system implementation and shared data consumption.



**Figure 1. Mechanics of Data Sharing via the [ClinicalStudyDataRequest.com](https://ClinicalStudyDataRequest.com) (CSDR) Platform**

This flow diagram highlights the 4 major steps of this data sharing model including the initial request for access to clinical trial data, review by an Independent Review Panel (IRP), access to a multi-sponsor analysis system, and public dissemination of secondary research findings.

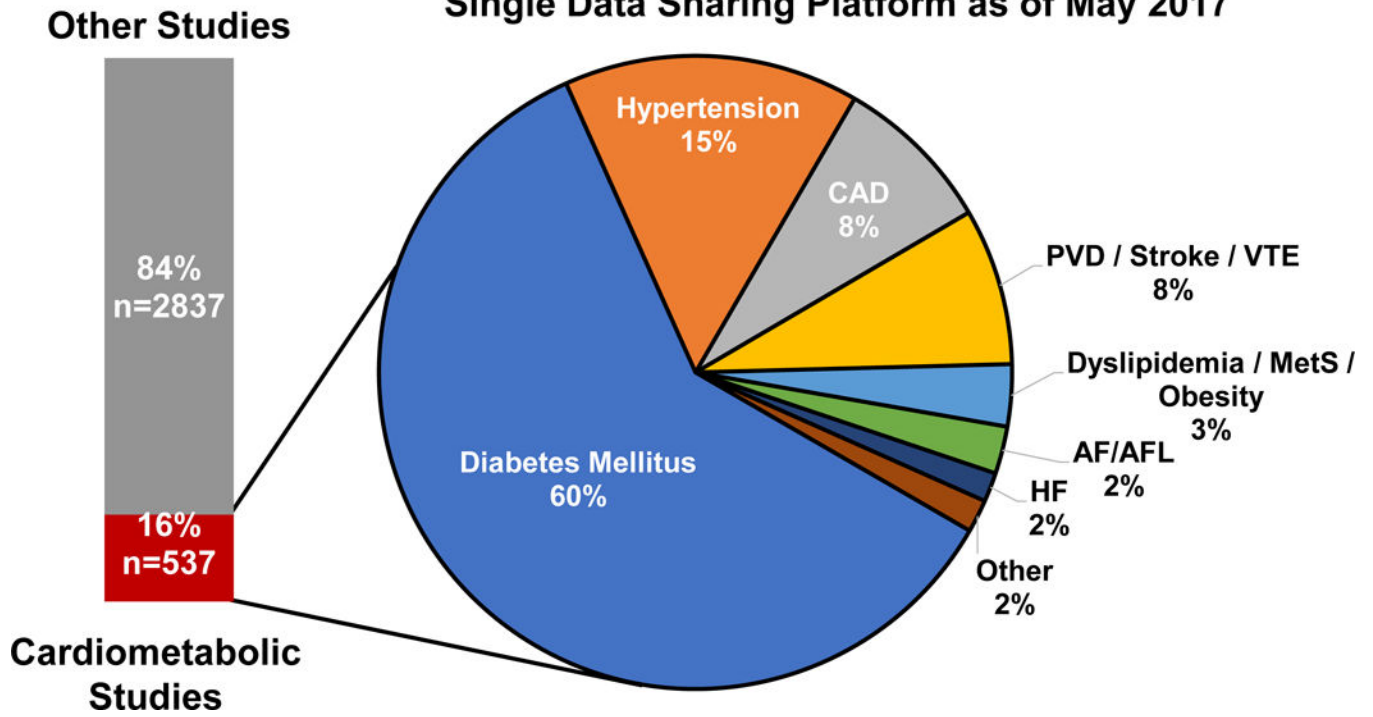


**Figure 2. Availability and Metrics of Use of Cardiometabolic Clinical Trial Data Hosted on the Multi-Sponsor Data Sharing Platform, [ClinicalStudyDataRequest.com](http://ClinicalStudyDataRequest.com)**

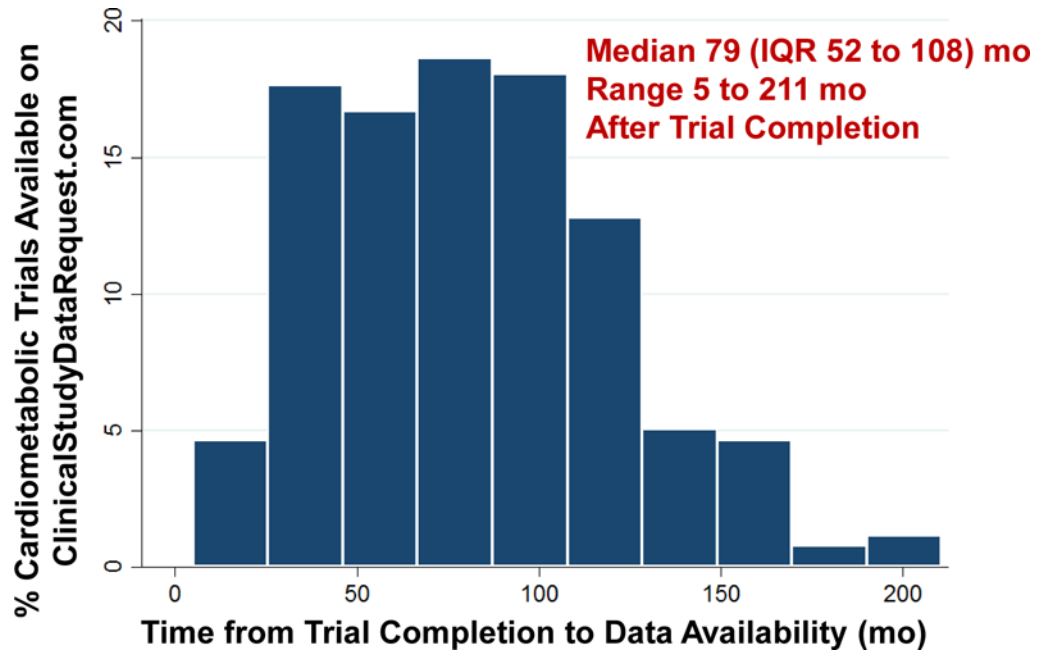
Landmark working group reports and proposals endorsing access to individual patient-level data have been highlighted (red arrows). The blue and red lines reflect availability of data from completed cardiometabolic clinical trials. The green line reflects dates that each cardiometabolic study was first requested through this platform with signed data sharing agreements. The purple line reflects peer-reviewed publications based on these shared data identified through PubMed/MEDLINE. Abbreviations: ACCESS CV = Academic Research Organization Consortium for Continuing Evaluation of Scientific Studies – Cardiovascular; CSDR = [ClinicalStudyDataRequest.com](http://ClinicalStudyDataRequest.com); ICMJE = International Committee of Medical Journal Editors; IOM = Institute of Medicine; PhRMA = Pharmaceutical Research and Manufacturers of America.



### Cardiometabolic Study Data Availability from a Single Data Sharing Platform as of May 2017



**Figure 3. Breakdown by Study Area of Cardiometabolic Trials Available for Data Requests**  
 Diabetes mellitus (60%) and hypertension (15%) were the most common study topics.  
 Abbreviations: AF/AFL = atrial fibrillation/atrial flutter; CAD = coronary artery disease; HF = heart failure; MetS = metabolic syndrome; PVD = peripheral vascular disease; VTE = venous thromboembolism.



**Figure 4. How Long Does It Take for Data to Become Accessible?**

Time from trial completion (extracted from ClinicalTrials.gov) to individual patient-level data availability on the [ClinicalStudyDataRequest.com](https://clinicalstudydatarequest.com) platform. Abbreviations: IQR = interquartile range.

**Table 1**  
 Select Published Experiences of Existing Clinical Trial Data Sharing Platforms Across Medical Disciplines.

Platform	Reference	Time-Frame of Access	Type of Data Shared	Total # of Trials Hosted	Total # of Requests	Total # Granted or Approved
<b>Data Reports Submitted for Regulatory Approval</b>						
European Medicines Agency	Bonini et al. <sup>8</sup>	2010–2013	Any data submitted for marketing authorization of any medicinal product	–	750	480
<b>Federally-Funded Clinical Studies</b>						
NHLBI Data Repository	Coady et al. <sup>12</sup>	2000–2016	Patient-level data from large NHLBI-supported trials and observational studies	100	800–850	–
<b>Industry-Sponsored Clinical Studies</b>						
CSDR, SOAR Initiative, and YODA Project	Navar et al. <sup>14</sup>	2013–2015	Patient-level data from select clinical studies sponsored by 14 companies	3,255	234	154
CSDR	Strom et al. <sup>16</sup>	2013–2015	Patient-level data from select clinical studies sponsored by 13 companies	3,049	177	144
	Strom et al. <sup>15</sup>	2013–2014	Patient-level data from select clinical studies sponsored by 10 companies	>1,200	58	36
YODA Project	Krumholz et al. <sup>13</sup>	2014–2015	Patient-level data from select trials sponsored by Johnson & Johnson	123	29	29

Abbreviations: CSDR = ClinicalStudyDataRequest.com; NHLBI = National Heart, Lung, and Blood Institute; SOAR = Supporting Open Access to Researchers; YODA = The Yale Open Data Access.

**Table 2**  
Inclusion Criteria, Number of Trials with Available Individual Patient-Level Data, and Estimated Total Trial Portfolio by Industry Sponsor.

Sponsor	Inclusion Criteria for Data Sharing	# of Shared Cardiometabolic Trials	# of Shared Other Trials	Total Trials with Available Data	Crude Estimate of Trial Portfolio per Sponsor*	% of Estimated Trial Portfolio with Shared Data on CSDR
<b>Astellas</b>	Phase 1–4 interventional studies for approved products completed after 1/2010 Phase 1–4 interventional studies for compounds terminated after 6/2015	12	26	38	624	6
<b>Bayer</b>	Approved products (by both FDA and EMA) after 1/2014 Approved products (by only 1 agency) after 1/2014 if no plan for ongoing review or submission	0	8	8	763	1
<b>Boehringer Ingelheim</b>	–	75	262	337	1303	26
<b>Daiichi-Sankyo</b>	Phase 2–4 interventional studies for approved products after 1/2014 (by FDA and EMA)	16	0	16	165	10
<b>Eisai</b>	Phase 2 and 3 studies for approved products after 1/2014 (by FDA and EMA) accepted for publication Phase 4 published interventional studies for approved products after 1/2014 (by FDA and EMA)	0	7	7	251	3
<b>GSK</b>	Global interventional studies ongoing or started after 12/2000 All interventional studies started after 1/2013	234	1793	2027	2633	77
<b>Lilly</b>	Phase 2–4 interventional studies for FDA-approved products on/after 1999 Phase 2–4 global interventional studies for FDA- and EMA-approved products started after 1/2007 Phase 2–4 regional/local interventional studies for FDA- and EMA-approved products started after 1/2014	13	253	266	953	28
<b>Novartis</b>	Phase 2–3 studies for previously approved products (by FDA and EMA) being submitted for new indication as of 1/2014 with decision regarding original study publication Phase 2–3 studies for approved products prior to 1/2014 (case-by-case review) with decision regarding original study publication	3	65	68	1455	5
<b>Roche</b>	Phase 2 and 3 studies started after 1/1999 Phase 4 studies for approved or terminated products started after 1/1999	11	176	187	858	22
<b>Sanofi</b>	Phase 2–4 interventional clinical studies for approved products after 1/2014 with accepted manuscript of original study	27	7	34	1379	2
<b>Takeda</b>	Phase 1–4 intervention trials for approved products after 1/2005 Phase 1–4 interventional trials for products terminated on or after 1/2014	146	165	311	430	72
<b>UCB</b>	Pivotal studies for regulatory approval of certolizumab, lacosamide, rosiglitone, levetiracetam	0	34	34	324	10
<b>ViiV</b>	Phase 2–4 global interventional studies of HIV drugs	0	41	41	43	95
<b>All Sponsors</b>		<b>537</b>	<b>2,837</b>	<b>3,374</b>	<b>11,181</b>	<b>30</b>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

\* ClinicalTrials.gov was queried for completed, interventional trials sponsored by the company without limits to study completion/initiation dates, global or regional enrollment, drug approval status, or publication status.

Abbreviations: CSDR = [ClinicalStudyDataRequest.com](https://clinicalstudydatarequest.com); EMA = European Medicines Agency; FDA = Food and Drug Administration; GSK = GlaxoSmithKline; HIV = human immunodeficiency virus.