SPECIAL ARTICLE

The ClinicalTrials.gov Results Database — Update and Key Issues

Deborah A. Zarin, M.D., Tony Tse, Ph.D., Rebecca J. Williams, Pharm.D., M.P.H., Robert M. Califf, M.D., and Nicholas C. Ide, M.S.

ABSTRACT

BACKGROUND

The ClinicalTrials.gov trial registry was expanded in 2008 to include a database for reporting summary results. We summarize the structure and contents of the results database, provide an update of relevant policies, and show how the data can be used to gain insight into the state of clinical research.

METHODS

We analyzed ClinicalTrials.gov data that were publicly available between September 2009 and September 2010.

RESULTS

As of September 27, 2010, ClinicalTrials.gov received approximately 330 new and 2000 revised registrations each week, along with 30 new and 80 revised results submissions. We characterized the 79,413 registry and 2178 results of trial records available as of September 2010. From a sample cohort of results records, 78 of 150 (52%) had associated publications within 2 years after posting. Of results records available publicly, 20% reported more than two primary outcome measures and 5% reported more than five. Of a sample of 100 registry record outcome measures, 61% lacked specificity in describing the metric used in the planned analysis. In a sample of 700 results records, the mean number of different analysis populations per study group was 2.5 (median, 1; range, 1 to 25). Of these trials, 24% reported results for 90% or less of their participants.

CONCLUSIONS

ClinicalTrials.gov provides access to study results not otherwise available to the public. Although the database allows examination of various aspects of ongoing and completed clinical trials, its ultimate usefulness depends on the research community to submit accurate, informative data.

From the Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Bethesda, MD (D.A.Z., T.T., R.J.W., N.C.I.); and Duke Translational Medicine Institute, Duke University, Durham, NC (R.M.C.). Address reprint requests to Dr. Zarin at the National Library of Medicine, Bldg. 38A, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD 20894, or at dzarin@mail.nih.gov.

N Engl J Med 2011;364:852-60.

The ClinicalTrials.gov trial registry was launched more than a decade ago. Since that time, it has been evolving in response to various policy initiatives. The registry now contains information on more than 100,000 clinical studies and has emerged as a key element of many public health policy initiatives aimed at improving the clinical research enterprise. In 2008, a database for reporting summary results was added to the registry. In this article, we present an update on relevant policies, summarize the structure and contents of the results database, and show how ClinicalTrials.gov data can be used to gain insight into the state of clinical research.

KEY TRIAL-REPORTING POLICIES

Section 801 of the Food and Drug Administration Amendments Act (FDAAA)[1] expanded the legal requirements for trial reporting at ClinicalTrials.gov. It was passed into law amid concerns about ethical and scientific issues affecting the design, conduct, and reporting of clinical trials,[2] including the suppression and selective reporting of results based on the interests of sponsors,[3] unacknowledged alterations of prespecified outcome measures,[4] "offshoring" of human-subjects research,[5] and failure to report relevant adverse events.[6] Among other things, the FDAAA mandates the submission of summary results data for certain trials of drugs, biologics, and devices to ClinicalTrials.gov, whether the results are published or not,[7] and imposes substantial penalties for noncompliance. The law's scope is not limited to industry-sponsored trials intended to support marketing applications but includes studies not intended to inform FDA action (e.g., comparative-effectiveness trials of approved drugs or devices), regardless of sponsorship. Table 1 summarizes the scope of key reporting requirements of the FDAAA and two other policies: the registration policy of the International Committee of Medical Journal Editors[8] and regulations being implemented by the European Medicines Agency for registration and results reporting of clinical drug trials conducted in the European Union.[9,10]

DESCRIPTION OF CLINICALTRIALS.GOV

Data in ClinicalTrials.gov are self-reported by trial sponsors or investigators by means of a Web-based system.[7] Registration information is generally reported at trial inception. Each record contains a set of mandatory data elements that describe the study's purpose, recruitment status, design, eligibility criteria, and locations, as well as other key protocol details.[11] Additional information may be provided with the use of optional data elements. Before public posting, ClinicalTrials.gov conducts a quality review of the submitted information. Each trial (regardless of the number of study sites) is represented by a single record, which is assigned a unique identifier (i.e., NCT number). Each record is expected to be corrected or updated throughout the trial's

Table 1. Scope of Interventional Studies Covered by Major Reporting Policies.*

Policy Requirements†	Registration	Results Reporting
FDAAA[1]	Interventional studies of drugs, biologics, or devices (whether or not approved for marketing); phases 2 through 4; at least one U.S. site or IND or IDE	Same as registration scope, but interventional studies of drugs, biologics, or devices only after FDA-approved for any use
ICMJE[8]	Interventional studies of any intervention type, phase, or geographic location	Not applicable
EMA[9,10]	Interventional studies of drugs and biologics (whether or not approved for marketing); phase 1 (pediatrics only); phases 2 through 4; at least one European Union site	Same as registration scope

* For complete descriptions of policy requirements, see the references cited. EMA denotes European Medicines Agency, FDAAA Food and Drug Administration Amendments Act, ICMJE International Committee of Medical Journal Editors, IDE investigational device exemption, and IND investigational new drug application.

† ClinicalTrials.gov allows the reporting of interventional and observational studies that are in conformance with any applicable human subject or ethics review regulations (or equivalent) and any applicable regulations of the national (or regional) health authority (or equivalent).

life cycle, and all changes are tracked on a public archive site that is accessible from each record (through a "History of Changes" link). Summary results data are entered in the results database after a trial is completed or terminated (Table 2). Once posted, results records are displayed with corresponding registry (summary protocol) information for each study. Resources and links to additional information are inserted by the National Library of Medicine to enhance the overall usefulness of the database. ClinicalTrials.gov is designed to benefit the general public by expanding access to trial information, but different parts of the database are likely to be of more or less direct use to different audiences.

QUALITY ASSURANCE

ClinicalTrials.gov uses automated business rules to alert data providers when required information is missing or when certain data elements are internally inconsistent. After passing automated validation, all submissions are individually reviewed before public posting to assess whether entries are complete, informative, internally consistent, and not obviously invalid; specific criteria for this assessment are described on the Web site.[15] Although the review of summary protocol information is generally straightforward, that of results submissions is more complex. The goal, at a minimum, is to determine whether entries provide an accurate depiction of the study design and whether the results can be understood by an educated reader of the medical literature. Some invalid data can be detected by ClinicalTrials.gov staff; however, other data cannot be verified because ClinicalTrials.gov does not have an independent source of study data (e.g., "624 years" is clearly an invalid results entry for mean age, whereas "62.4 years" may or may not be the true mean age). Submissions are not posted on the public site until quality requirements are met; if any important problems are detected (Table 3), results records are returned to the data providers for revision. However, individual record review has inherent limitations, and posting does not guarantee that the record is fully compliant with either ClinicalTrials.gov or legal requirements.

RELATION TO PUBLICATION

ClinicalTrials.gov is designed to complement, not replace, journal publication. The results database provides public access to a complete set of summary results in a structured system that supports search and analysis. These data are primarily tabular in format and lack significant narrative portions. The database facilitates identification of acts of omission (e.g., incomplete reporting of outcome measures) and acts of commission (e.g., unacknowledged changes to prespecified outcome measures). Journals select research articles for publication on the basis of their target audiences, and the articles supplement reported data

Table 2. Summary Objectives and Description of Requirements for the ClinicalTrials.gov Results Database.

Objectives
Satisfy legal requirements
Promote objective, standardized reporting by capturing key trial features in the form of tabular data while minimizing potentially subjective narrative text
Facilitate "good reporting practices," including accommodation of publishing[12] and regulatory[13] guidelines
Provide structured data entry to ensure complete reporting, efficient quality review, and consistent display of both required and voluntary data elements
Support detailed searches with the use of the database structure and other National Library of Medicine functions[14]
Description of scientific modules (in tabular format)
Participant flow: Progress of research participants through each stage of a trial according to group, including the number of participants who dropped out of the clinical trial
Baseline characteristics: Demographic and baseline data for the entire trial population and for each group
Outcome measures and statistical analyses: Aggregate results data for each primary and secondary outcome measure according to group; statistical analyses as appropriate
Adverse events: List of all serious adverse events; list of other (not including serious) adverse events in each group that exceed a frequency threshold of 5% within any group; both lists include adverse events, whether anticipated or unanticipated, and grouped by organ system
Administrative information
Key dates and contact information
Description of agreements, if any, between the sponsor and the principal investigator that would restrict dissemination of results by the principal investigator

Table 3. ClinicalTrials.gov Quality Review Criteria.

Quality Review Criterion	Description	Example	Comment
Lack of apparent validity	Data are not plausible on the basis of information provided	Outcome measure data: mean value of 263 hours of sleep per day	Measure of mean hours per day can have values only in the range of 0 to 24, so value of 263 is not valid
Meaningless entry	Information is too vague to permit interpretation of data	Outcome measure: description states "clinical evaluation of adverse events, laboratory parameters, and imaging"; data reported as 100 and 96 participants in each group	Data are uninformative; unclear what counts of 100 and 96 participants refer to; description of outcome measure not sufficient for an understanding of the specific outcome
Data mismatch	Data are not consistent with descriptive information	Outcome measure is described as "time to disease progression"; data reported as 42 and 21 participants in each group	A time-to-event measure requires a unit of time (e.g , days or months)
Internal inconsistency	Information in one section of record conflicts with or appears to be inconsistent with information in another section	Study type is "observational," but study title includes the word "randomized"	Randomized studies are interventional, not observational
Trial design unclear	Structure of tables and relevant group names and descriptions do not permit a reader to understand the overall trial design or do not accurately reflect the design	Results modules: participant flow and baseline characteristics entered as a two-group study with a total of 400 participants; outcomes entered for three comparison groups with 600 participants	If there is a third group, this should be reflected in the description of participant flow and baseline characteristics

with peer-reviewed discussions of background, rationale, context, and implications of findings. Journal editors who abide by the standards set by the International Committee of Medical Journal Editors recognize these complementary roles and consider manuscripts for publication even when the results of a trial have already been posted on ClinicalTrials.gov.[8]

DESCRIPTIVE DATA ABOUT TRIALS IN CLINICALTRIALS.GOV

Table 4 provides summary data on registry and results records for interventional studies that were publicly available on September 27, 2010. As of this date, approximately 330 new registrations and 2000 revised registrations had been submitted each week.

RESULTS DATABASE

All studies registered at ClinicalTrials.gov are eligible for results submission; however, submission of results is required only for trials covered by the FDAAA (Table 2). Approximately 30 new and 80 revised records had been received each week; we estimate that full compliance with the FDAAA would lead to results submission for approximately 40% of newly registered studies, or over 100 new records per week.

The results of 3284 registered trials had been submitted by 666 data providers. Of these trials, 2324 had been posted publicly; the remainder either were undergoing quality-assurance review by ClinicalTrials.gov staff or were returned for correction.

Of 2178 clinical trials with posted results records, 20% had more than two reported primary outcome measures and 5% had more than five. For some studies, posted results include more than 100 primary and secondary outcome measures. The FDAAA requires the reporting of all primary and secondary outcome measures, and ClinicalTrials.gov does not limit the number of primary and secondary outcome measures that can be listed. Other prespecified and post hoc outcome measures may also be listed.

Of the 2324 posted results entries, 14% were linked to a PubMed citation through an indexed NCT number[16]; other publications that may exist could be found only through focused PubMed searches. We randomly selected a sample of 150 posted results records in September 2009 and conducted manual searches in an attempt to identify all associated publications. Using all available data, we found that 38 of these studies (25%) had an associated citation in September 2009, and 78 (52%) by November 2010. Although

this percentage may continue to increase, it is unlikely that all outcomes from these studies will be published.[17]

SECONDARY FINDINGS FROM CLINICALTRIALS.GOV DATA

A growing number of researchers are using ClinicalTrials.gov data to examine various aspects of the clinical research enterprise. For example, recent studies evaluated registration records to analyze trends in the globalization of the clinical research enterprise,[5,18] the level of selective publication of study results,[19,20] and the degree of correspondence between registered and published outcome measures.[19-21] Scoggins and Patrick reviewed registration records to identify the types of trials for which patient-reported outcomes were

Table 4. Characteristics of Interventional Study Records Posted at ClinicalTrials.gov as of September 27, 2010.*

Variable	Registry Records (N=79,413)	Results Records (N=2178)
Lead sponsor class — no. of records (no. of sponsors)		
Industry	28,264 (2880)	1802 (200)
Nonindustry	51,149 (4372)	376 (196)
Recruitment status — no. of records		
Recruiting	22,696	0
Active, not recruiting	12,343	74
Completed	34,549	1883
Terminated†	3,551	221
Other‡	6,274	0
Intervention type — no. of records§		
Drug or biologic	56,580	1935
Medical procedure	9,636	69
Device	6,012	127
Other¶	16,771	185
Study phase — no. of records‖		
0 or 1	9,359	271
1–2 or 2	20,023	393
2–3 or 3	13,822	844
4	7,890	375
Intervention model — no. of records		
Parallel assignment	38,813	1321
Single group assignment	21,765	497
Crossover assignment	6,543	331
Factorial assignment	1,524	18
Missing data	10,768	11
Data monitoring committee — no. of records/total no. with responses (%)		359/1509 (24)
By study phase		
0 or 1		12/221 (5)
1–2 or 2		113/314 (36)
2–3 or 3		142/520 (27)
4		45/298 (15)
Phase not available		47/156 (30)
By enrollment**		

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.