



HHS Public Access

Author manuscript

Ann Intern Med. Author manuscript; available in PMC 2017 March 20.

Published in final edited form as:

Ann Intern Med. 2016 September 20; 165(6): 421–430. doi:10.7326/M15-2658.

ClinicalTrials.gov and Drugs@FDA: A comparison of results reporting for new drug approval trials

Lisa M. Schwartz, MD, MS, Steven Woloshin, MD, MS, Eugene Zheng, BA, Tony Tse, PhD, and Deborah A. Zarin, MD

The Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, NH (LMS, SW), Informulary, Inc (EZ) and the National Library of Medicine, Bethesda, MD (TT, DAZ)

Abstract

Background—Pharmaceutical companies and other trial sponsors must submit certain trial results to ClinicalTrials.gov. The validity of these results is unclear.

Purpose—To validate results posted on ClinicalTrials.gov against publicly-available FDA reviews on Drugs@FDA.

Data sources—ClinicalTrials.gov (registry and results database) and Drugs@FDA (medical/statistical reviews).

Study selection—100 parallel-group, randomized trials for new drug approvals (1/2013 – 7/2014) with results posted on ClinicalTrials.gov (3/15/2015).

Data extraction—Two assessors systematically extracted, and another verified, trial design, primary/secondary outcomes, adverse events, and deaths.

Results—The 100 trials were mostly phase 3 (90%) double-blind (92%), placebo-controlled (73%), representing 32 drugs from 24 companies. Of 137 primary outcomes from ClinicalTrials.gov, 134 (98%) had corresponding data in Drugs@FDA, 130 (95%) had concordant definitions, and 107 (78%) had concordant results; most differences were nominal (i.e. relative difference < 10%). Of 100 trials, primary outcome results in 14 could not be validated. Of 1,927 secondary outcomes from ClinicalTrials.gov, 1,061 (55%) definitions could be validated and 367 (19%) had results. Of 96 trials with 1 serious adverse event in either source, 14 could be compared and 7 were discordant. Of 62 trials with 1 death in either source, 25 could be compared and 17 were discordant.

Limitations—Unknown generalizability to uncontrolled or crossover trial results.

Address for reprint requests: Steven Woloshin, MD, The Dartmouth Institute, Level 5, WTRB, 1 Medical Center Drive, Lebanon, NH 03756.

Author mailing addresses: Lisa M. Schwartz, MD, MS, The Dartmouth Institute, Level 5, WTRB, 1 Medical Center Drive, Lebanon, NH 03756

Steven Woloshin, MD, MS, The Dartmouth Institute, Level 5, WTRB, 1 Medical Center Drive, Lebanon, NH 03756

Eugene Zheng, BA, 2900 University Ave SE, Apt 306, Minneapolis, MN 55414

Tony Tse, PhD, National Library of Medicine, Bethesda, MD, 8600 Rockville Pike, Bethesda, MD 20894

Deborah A. Zarin, MD, National Library of Medicine, Bethesda, MD, 8600 Rockville Pike, Bethesda, MD 20894

Conflict of Interest Statement Drs. Schwartz and Woloshin are co-founders and shareholders in Informulary, Inc., a company that provides data about the benefits, harms and uncertainties of prescription drugs. EZ was paid by Informulary for work on this study. Drs. Zarin and Tse are Director and Program Analyst, respectively, at ClinicalTrials.gov.

Conclusion—Primary outcome definitions and results were largely concordant between ClinicalTrials.gov and Drugs@FDA. Half of secondary outcomes could not be validated because Drugs@FDA only includes “key outcomes” for regulatory decision-making; nor could serious adverse events and deaths because Drugs@FDA frequently only includes results aggregated across multiple trials.

Sponsors are required by Federal law to submit summary results of applicable clinical trials (including trials beyond Phase 1 supporting Food and Drug Administration (FDA) new drug approvals) to ClinicalTrials.gov for public posting. (1) Submissions consist of minimum “basic results” data elements in a tabular format, including results for all primary and secondary outcomes prespecified in the study protocol and all anticipated and unanticipated serious adverse events observed during the trial. (2) This law also requires ClinicalTrials.gov to assess ways of verifying the accuracy of sponsor-submitted results information, including using public sources such as FDA advisory committee summary documents and FDA action package approval documents.(3) Although ClinicalTrials.gov currently assesses internal consistency through quality checks(4), the validity of posted results can only be assessed by comparing submitted data with external reference standards. Recent studies comparing ClinicalTrials.gov data with peer-reviewed journal publications suggest that discrepancies in reported primary and secondary outcomes, numerical results, and adverse events, are relatively common, although which source is more likely to be correct is unclear. (5–7)

FDA drug approval packages may represent a better “reference standard” than publications for validating results posted on ClinicalTrials.gov because journal editors and peer-reviewers typically lack access to individual participant data. Consequently, investigators can choose to publish outcomes based largely on statistical significance or other criteria.(8–10) In contrast, FDA statisticians, who have access to individual participant data, can independently analyze sponsor-submitted trial results based on what they believe are the best statistical practices. (11, 12). Independent analysis of individual participant data from a trial can yield treatment effects that range in direction, magnitude, and statistical significance, based on which particular outcome is selected and how it is analyzed (e.g., discretion in selecting measurement populations, such as intention-to-treat vs. per-protocol population; accounting for missing data; or timing for outcomes assessment). For example, a high-profile journal article concluded that celecoxib reduced major bleeding compared to ibuprofen and diclofenac based on 6-month results.(13) But the FDA reviews, which included results for the protocol-specified 1-year endpoints, indicated that celecoxib did not reduce major bleeding.(14)

We compared sponsor-submitted definitions and results posted on ClinicalTrials.gov with corresponding FDA-generated information posted on Drugs@FDA for trials used to support new drugs approvals: specifically, how often efficacy and adverse event outcomes could be compared and whether the posted data were consistent.

Methods

Sample collection

To identify 100 parallel-group, randomized trials that served as the basis of FDA new drug approvals (i.e., new molecular entities), we searched Drugs@FDA(15) beginning with approvals on January 1, 2013 (Figure). Each FDA approval package includes review documents authored by FDA staff (e.g., physicians, statisticians, pharmacologists). These documents, which summarize analyses of clinical and other data submitted in new drug applications (NDAs), are used by FDA to determine whether to approve marketing of new drugs or biological products for a particular use(s) (16). While FDA has made some drug approval packages and component review documents publicly available on Drugs@FDA since 1997 (12), recent federal law now requires FDA to post “action packages” systematically for original NDAs (1)

We manually searched FDA medical and statistical reviews to find all trials designated as “pivotal” and “supportive” by the FDA reviewer. We then sought to match these trials with corresponding ClinicalTrials.gov results database records, downloaded on March 15, 2015. Although the Clinicaltrials.gov and Drugs@FDA databases were created for different purposes, their content overlaps substantially (Table 1). Because FDA review documents do not currently list ClinicalTrials.gov Identifiers (NCT Numbers), we used the process described in the Figure to match trials between sources. We searched Drugs@FDA through July 2014 until reaching our target: 75 pivotal and 25 supportive parallel-group, randomized trials with some results data in both sources – comprising all trials that could be compared during this time frame. We hypothesized that documents available from Drugs@FDA would contain less results information for supportive trials than for pivotal trials, which provide the primary evidence for approval.

Data extraction

We created a structured data extraction form to capture detailed trial information from ClinicalTrials.gov and Drugs@FDA and revised it after a pilot test extracting information for 5 trials. The 6 major domains (Appendix 1) were:

1. **Trial characteristics** including drug indication, development phase, blinding, comparator, and basic trial data (number randomized, completed, age, and gender distributions).
2. **Primary outcome**, including **definition** using the following framework: domain, specific measurement, specific metric, method of aggregation (4) plus time frame, **analysis** (measurement population, methods to account for missing data), **result values** (consistency of results for each study arm, number analyzed, results), and **treatment effect** (differences in treatment effect and associated confidence interval and p-value between experimental and control arms).
3. **Secondary outcomes** (number, data availability, and outcome).
4. **Serious adverse events** (number analyzed and consistency of results).

5. **Deaths** (whether reported or mention that no deaths occurred and consistency of results).
6. Number of **other adverse events**.

In contrast to ClinicalTrials.gov records, which typically only present a single set of analyses per outcome, Drugs@FDA review documents often present multiple analyses, including those conducted by the sponsor and separately by the FDA statistician (e.g., sensitivity analyses with different measurement populations or different imputation methods for missing data). We extracted results from the FDA statistical reviewer's independent analyses (available for two-thirds of primary outcomes) or, when not available, the drug company's analyses, provided that the statistical reviewer explicitly indicated agreement with them (remaining primary outcomes).

Two assessors (EZ and Rachael Bornstein) systematically extracted -- and another (LS or SW) verified -- trial design, primary and secondary outcomes, adverse events, and death from both ClinicalTrials.gov and Drugs@FDA.

Data Comparison

Consistency of outcome definitions and analyses—The number and definitions of primary and secondary outcomes posted on ClinicalTrials.gov (and concordance in outcome “level”) were compared to those listed in Drugs@FDA. Primary outcomes definitions were considered discordant if a mismatch occurred at any level of the following framework: domain (e.g., anxiety), specific measurement (e.g. Beck Anxiety Inventory), specific metric (e.g. change from baseline), or method of aggregation (e.g. mean). We also used an alternative definition of discordance that excludes method of aggregation to account for researchers who feel it is unnecessary to prespecify statistical analysis plans prior to trial unblinding (11). In addition, we compared timing of the outcome assessment plus three key aspects of the primary outcome analyses: measurement population, crude or adjusted analysis, and method of handling missing data.

Consistency of results—We assessed the consistency of results reporting between ClinicalTrials.gov and Drugs@FDA using the approach adopted from Hartung et. al.(7) (detailed in Appendix 1). For example, results for the outcome measure, “% change from baseline of A1c,” were considered discordant if the reported values were not consistent to 1 decimal place (e.g., 0.094 is **not** consistent with 0.12 because it rounds to 0.09, but 0.115 would be consistent since it rounds to 0.12).

We analyzed the data at two levels: (1) numbers of trials and (2) numbers of primary outcomes or named serious adverse events, including death. While the latter approach explicitly shows the frequency of discrepancies for individual measures, it may overstate the distribution of discrepancies among trials because the number of potential discrepancies is the product of the outcomes times the number of study arms. Reporting numbers of trials mitigates this problem: a few discordant trials (even with many outcomes) would not overwhelm a majority of concordant trials. This approach, however, created an important challenge: how many discordant outcomes (or study arms) does it take to deem a trial “discordant” between the two sources? We used Hartung's approach and called a trial

“discordant” if data from ClinicalTrials.gov and Drugs@FDA were inconsistent for 1 or more result, “concordant” if all were consistent, and “cannot compare” if, in both sources, the outcomes did not match or the data were not posted.

Results

Overall, our sample of 100 parallel-group randomized trials (90% Phase 3) was used to support 32 recent new drug approvals from 24 pharmaceutical companies for a variety of indications, most commonly diabetes (42%) and pulmonary disease (19%) (Table 2). Most trials were double-blind (92%) and included a placebo control (73%). While trial results from both sources listed a median of 1 primary outcome measure (range: 1–8), the Drugs@FDA reviews listed a median of 5 (0–94) “key” secondary outcome measures compared to a median of 7 “prespecified” secondary outcome measures (0–227) posted on ClinicalTrials.gov.

The numbers of trial participants randomized and completing the trial were discordant in 24 and 31 trials, respectively, ranging from 0.1% discrepant as a proportion of the total number to 296%. Mean age was discordant in 4 trials, but could not be compared for 27 trials, largely because the data were presented in different formats (e.g., categories versus means).

Primary outcomes measures and outcomes

Primary outcomes were largely consistent across both sources (Table 3). Of 137 primary outcomes posted on ClinicalTrials.gov, Drugs@FDA mentioned 134 (98%)—among which only 4 had discordant outcome definitions – and concordantly specified 119 (87%) as “primary.”

Table 4 illustrates two types of discrepancies observed between ClinicalTrials.gov and Drugs@FDA primary outcomes. The first type of discrepancy occurred for outcomes specified as primary in both sources with discordant definitions. The 4 cases of discordance using the broader definition stemmed from differences in the use of measurement tools (1), methods of aggregation (2), and time frames (1). (Using the narrower discordance definition, we only observed 2 cases.)

The second type of discrepancy occurred for ClinicalTrials.gov primary outcomes that Drugs@FDA did not specify as “primary” (Table 4). Of the 14 cases, Drugs@FDA did not mention the outcome at all in 3, and mentioned, but did not specify as primary, secondary, or tertiary, in the remaining 11 cases. Ten of these discrepancies occurred in 8 trials where at least 1 other ClinicalTrials.gov primary outcome was also identified as primary in Drugs@FDA. For example, for a trial of Mipomersen (Kynamro) for hyperlipidemia (NCT00607373), ClinicalTrials.gov and Drugs@FDA matched on the primary outcome “% change in LDL,” but ClinicalTrials.gov also listed another primary outcome (“mean LDL at the end of the study”) which was listed without any specified level in Drugs@FDA.

In 3 of the 100 trials, none of the ClinicalTrials.gov primary outcomes was considered “primary” by FDA. For example, for a trial of Pomalidomide (Pomalyst) for multiple myeloma (NCT00833833), ClinicalTrials.gov listed 3 primary outcomes, none of which

matched a Drugs@FDA-specified primary outcome. The first ClinicalTrials.gov-listed primary outcome, dose-limiting toxicity from Phase 1 of the trial, was not mentioned in the available FDA documents. FDA deemed the second ClinicalTrials.gov-listed primary outcome, progression free survival (time-to-event analysis), to be “exploratory,” stating that objective response rate (not specified as either a primary or secondary outcome in ClinicalTrials.gov) should be “primary.” The third ClinicalTrials.gov-listed primary outcome, percent of people with disease progression or death, was mentioned in the FDA documents but without a specified outcome level].

Two analytic methods for the primary outcome, measurement population and whether results were crude or adjusted, were discordant for 9% (12/129 with data in Drugs@FDA) and 5% (7/129) of primary outcomes, respectively, and could not be compared for 10% (13/129) and 6% (8/129) (Table 3). One notable discordance occurred in the second pivotal trial testing Ospemifene (Osphena) for dyspareunia in postmenopausal women (NCT00276094): the ClinicalTrials.gov record reported an analysis of the full randomized population for 3 of the 5 primary outcomes, but the information posted on Drugs@FDA only presented the analysis for the subgroup for whom the drug had a benefit (i.e., women with dyspareunia as their most bothersome symptom). While 51% of primary outcomes had concordant descriptions of how missing data were handled, 47% could not be compared due to missing information in at least one source (i.e., not available for 56 trials in ClinicalTrials.gov and for 22 trials in Drugs@FDA).

Among the 100 trials, the number of people analyzed for the primary outcome was discordant in 23 trials and the results were discordant for 14 (additionally, 8 were not reported in either or both sources). (Table 5) Of 137 ClinicalTrials.gov-listed primary outcomes, 42 (31%) were discordant in the number of people analyzed and 22 (16%) had discordant results. Discordance was not evenly distributed across the 24 drug companies. (Appendix 2) For six companies listed as lead sponsors on ClinicalTrials.gov, half or more of primary outcomes across the trials sponsored by each were found to be discordant for either the number of people analyzed or the results: Biomarin (1 of 1 results discordant), Genzyme (4 of 8 number [of people] analyzed discordant), Iroko (2 of 3 results), Shinogi (13 of 13 number analyzed and 8 of 13 results), Takeda (6 of 12 number analyzed), and Trius (1 of 2 number analyzed). As hypothesized, fewer results were presented for supportive (4) than pivotal (24) trials in documents available on Drugs@FDA.

Comparisons of treatment effects, corresponding confidence intervals, and p values could not be made in about a third of the cases because the information was not available in one or both sources (e.g., treatment effects were not available in either or both sources for 26% of trials and 35% of primary outcomes). When treatment effect sizes were included in both sources, few trials (5) had discordant results, all with relative differences <10%.

Secondary outcomes

Drugs@FDA only lists secondary outcome measures considered “key” for regulatory decision-making by FDA reviewers, thus limiting our ability to validate ClinicalTrials.gov entries. Of the 1,927 secondary outcomes posted on ClinicalTrials.gov, Drugs@FDA

mentioned 1,061 (55%) and 981 (51%) were specified as “secondary;” results were included in Drugs@FDA for 367 (19%) (Table 3).

Harm reporting

Documents in Drugs@FDA frequently pooled serious adverse events across trials to support the overall assessment of harm for a drug. Our sample included 14 trials for which Drugs@FDA documents listed trial-specific serious adverse events that were also listed in ClinicalTrials.gov. (Table 5) Among the 14 trials with serious adverse events in both sources, the number of unique named events reported per trial was substantially higher in ClinicalTrials.gov (median 50, range 1 – 223) than in Drugs@FDA (median 9, range 1–37). Adverse event results were discordant in 7 (of 14) trials.

A total of 62 trials listed at least 1 death in either or both sources. Of the 25 trials that reported the number of deaths in both sources, 17 (68%) were discordant. For 15 of the 17 discordant results, fewer deaths were reported in ClinicalTrials.gov. Some, but not all discrepancies stem from time-frame differences. The largest discrepancy (104 deaths) was seen in a pivotal trial of Xofigo (radium 223 dichloride) for hormone-resistant prostate cancer (NCT00699751). The ClinicalTrials.gov record reported 6 deaths assessed “after the first injection of study treatment and within 12 weeks after the last injection of study treatment.” The Drugs@FDA documents, however, noted 110 deaths: 30 occurring during the treatment period (i.e., between first and last injection)+ 30 days, and an additional 80 deaths in the next 3 years. The time-frame difference cannot entirely account for the discrepancy because more deaths occurring over less time were reported in the Drugs@FDA documents than on ClinicalTrials.gov.

Discussion

We were able to match almost all of the primary outcome measures posted on ClinicalTrials.gov with publicly available data from Drugs@FDA in our sample. The primary outcome definitions and results listed in the ClinicalTrials.gov results database entries were largely consistent with those reported in corresponding review documents from Drugs@FDA. A minority, however, were nominally discordant. In 3 of the 100 trials examined, for example, no ClinicalTrials.gov-listed primary outcome matched any primary outcome identified on Drugs@FDA, raising questions about the validity of these ClinicalTrials.gov entries. Thus, our analysis suggests that Drugs@FDA may be a useful resource validating primary outcomes posted on the ClinicalTrials.gov results database, but required considerable effort for manual extraction from Drugs@FDA and matching of trials and primary outcomes between the two databases.

In contrast, Drugs@FDA was not as helpful in validating secondary outcome measures, adverse events information, and death. Only about half (51%) of secondary outcomes listed in ClinicalTrials.gov were explicitly identified as “secondary” in Drugs@FDA. As noted, FDA only reports “key” secondary outcome measures, whereas submissions to ClinicalTrials.gov are required to include all secondary outcome measures prespecified by the sponsor. However, the fact that some outcomes identified by the sponsor as “secondary” in ClinicalTrials.gov were presented but not labeled as “secondary” in Drugs@FDA raises

questions about the meaning and utility of the term “secondary outcome measure.” Serious adverse events and deaths could also not be validated for most trials because Drugs@FDA generally only reports such data aggregated across trials. These findings, consistent with other studies comparing ClinicalTrials.gov results entries and corresponding results publications (5–7), suggest that while the primary outcome measures in ClinicalTrials.gov generally appear to be valid, the information currently available from Drugs@FDA cannot be used to conduct routine validation of the full set of results submissions to ClinicalTrials.gov.

Our analysis also highlights mismatches in primary and secondary outcomes, which illustrate the different purposes of the two databases: sponsor-submitted results information with a focus on fidelity to the protocol as required by law for public posting on ClinicalTrials.gov and FDA-analyzed results for regulatory decision-making, some of which are subsequently available from Drugs@FDA. In some cases, the sponsor’s protocol may have been written without FDA input – or even in contradiction to FDA’s recommendations about design or analysis. (17, 18) Consequently, analyses by FDA staff may deviate from the sponsor’s pre-specified analysis plan if the FDA believes that there is a better way to analyze the data to inform its regulatory decisions. (11)

Our study has several limitations. We only analyzed parallel-group, randomized trials, so it is uncertain whether our findings generalize to other designs that may be the basis of FDA approval (e.g., single-arm or crossover studies). Further, while we cannot be certain that our sample is sufficiently representative of all trials used to support new drug approval, the analyzed trials are unselected, recent FDA new drug approval trials with results reported on ClinicalTrials.gov and cover a broad array of drug products and indications. Additionally, the study characteristics reported in Table 2 appear consistent with previously reported analyses of trials supporting FDA new drug approvals. (19, 20) Nevertheless, our sample only included 24 different drug companies, and discordance was mostly concentrated among a few (Appendix 2).

This study also highlights important limitations of both databases as public resources. ClinicalTrials.gov would be improved by added clarity to key definitions and by providing additional structure to certain data elements. For example, most pre-specified “secondary outcome measures” reported in ClinicalTrials.gov are identified as something other than “key” by FDA reviewers (e.g., “exploratory” or “safety”), raising questions about the meaning of the term, as generally understood. The recent notice of proposed rulemaking (NPRM) issued by the Department of Health and Human Services (HHS) proposed reserving the term “secondary” for pre-specified outcomes with a statistical analysis plan (outcomes not primary or secondary would be designated as exploratory) - and all submitted outcome definitions would need to specify the measurement, metric, method of aggregation, and time frame.(21) Doing so could help to prevent the cherry picking of result-driven false positive findings that pervades the published biomedical literature. (8–10) The NPRM also requested comments on whether to require the submission of full protocols to ClinicalTrials.gov, which some have suggested would reveal how secondary outcomes were prospectively defined and intended to be analyzed. (22, 23) Others have noted that greater structuring of ClinicalTrials.gov (24), such as separate data elements for measurement

population, handling of missing data, and all-cause mortality (6) would improve the utility of results reporting, though with added burden on data submitters.

The utility of Drugs@FDA for validating ClinicalTrials.gov entries specifically, and for third-party researchers in general, would be improved if NCT Numbers were included to provide unambiguous identification of trials, and if greater structure were provided.(12) For example, it was challenging to find unambiguously the trials of interest and to identify those sections of interest within the reviewed reports. FDA’s new Drug Trials Snapshots website (25) makes primary outcome results more accessible, but lacks a standard format for trial data and analytic methods.

Finally, ClinicalTrials.gov and Drugs@FDA, extraordinary resources that contribute to the transparent communication of trial results, could complement each other in other ways. ClinicalTrials.gov is an attempt to provide “fishbowl transparency,” a complete reporting of all summary data according to the pre-specified study protocol. (26) In contrast, Drugs@FDA, which reflects what the FDA thought was the proper analysis for regulatory action, provides a different kind of transparency - sometimes called “reasoned transparency” (26) or guidance and context for interpreting what lives in the fishbowl based on the independent opinions of their expert reviewers. Unfortunately, these opinions may be hard to find in review documents as currently formatted. Even if all the numbers reported in ClinicalTrials.gov were completely accurate, there will always be open questions about the trial design, conduct, or analysis which can affect conclusions about trial results, which are assessed in Drugs@FDA. Reviewers may have raised questions, for example, about the appropriateness of an active comparator used (or its dose) or about unvalidated outcome measures. Implementing better integration between ClinicalTrials.gov and Drugs@FDA for providing context (e.g., by including NCT Numbers in Drugs@FDA documents) would help combine the fishbowl and reasoned approaches to transparency. The result would be better health information and, perhaps ultimately, better health.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

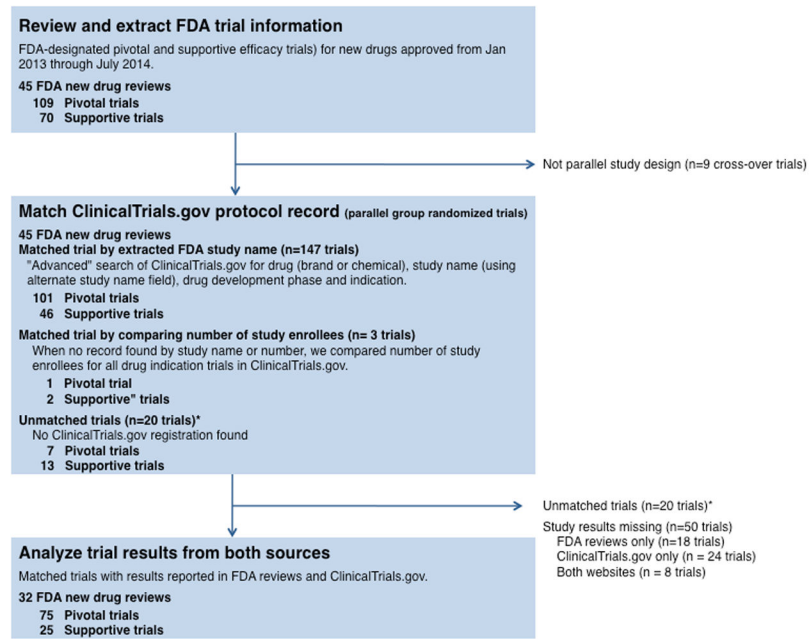
Primary Funding Source Supported in part by the Intramural Research Program of the National Library of Medicine (NLM), National Institutes of Health; LMS, SW were paid through a contract from NLM.

Grant support: None

References

1. Food and Drug Administration Amendments Act of 2007: Public Law 110–85.
2. ClinicalTrials.gov. [Accessed February 24, 2016] “Basic Results” Data Element Definitions (DRAFT). Apr. 2015 Available at: https://prsinfo.clinicaltrials.gov/results_definitions.html
3. §402(j)(a)(5)(C) of the Public Health Service Act, as amended by FDAAA.
4. Zarin D, Tse T, Williams R, Califf R, Ide N. The ClinicalTrials.gov results database - update and key issues. *N Engl J Med.* 2011; 364:852–60. [PubMed: 21366476]

5. Becker J, Krumholz H, Ben-Josef G, Ross J. Reporting of results in ClinicalTrials.gov and high-impact journals. *JAMA*. 2014; 311:1063–65. [PubMed: 24618969]
6. Earley A, Lau J, Uhlig K. Haphazard reporting of deaths in clinical trials: a review of cases on ClinicalTrials.gov records and matched publications - a cross sectional study. *BMJ Open*. 2013; 3(1)
7. Hartung D, Zarin D, Gulse J, McDonagh M, Paynter R, Helfand M. Reporting discrepancies between ClinicalTrials.gov results database and peer-reviewed publications. *Ann Intern Med*. 2014; 160:477–83. [PubMed: 24687070]
8. Chan A, Hrobjartsson A, Haahr M, Gotzsche P, Altman D. Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles. *JAMA*. 2004; 291:2457–65. [PubMed: 15161896]
9. Turner E, Matthews A, Lindartos E, Tell R, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008; 358:252–60. [PubMed: 18199864]
10. Vedula S, Bero L, Scherer R, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. *N Engl J Med*. 2009; 361:1963–71. [PubMed: 19907043]
11. Food and Drug Administration. [Accessed February 24, 2016] CDER Manual of Policies and Procedures. Office of Translational Sciences. Good Review Practice: Statistical Review Template. 2012. MAPP 6010.4. Available at: <http://www.fda.gov/downloads/aboutfda/centersoffices/officeofmedicalproductsandtobacco/cder/manualofpoliciesprocedures/ucm313814.pdf>
12. Turner E. A taxpayer-funded clinical trials registry and results database. *PLOS Med*. 2004; 1(3):e60. [PubMed: 15562322]
13. Silverstein F, Faich G, Goldstein J, et al. Gastrointestinal Toxicity With Celecoxib vs Nonsteroidal Anti-inflammatory Drugs for Osteoarthritis and Rheumatoid Arthritis: The CLASS Study: A Randomized Controlled Trial. *JAMA*. 2000; 284:1247–55. [PubMed: 10979111]
14. Juni, P.; RA; Dieppe, P. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs?; *BMJ*. 2002. p. 324<http://dx.doi.org/10.1136/bmj.324.7349>
15. Food and Drug Administration. Drugs@FDA. Available at: <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/>
16. Turner E. How to access and process FDA drug approval packages for use in research. *BMJ*. 2013
17. Woloshin S, Schwartz L. U.S. Food and Drug Administration approval of flibanserin: Even the score does not add up. *JAMA Intern Med*. 2016
18. Woloshin S, Schwartz L, Frankel B, Faerber A. US Food and Drug Administration and design of drug approval studies. *JAMA*. 2014; 312:2163–65. [PubMed: 25423223]
19. Lee K, Bacchetti P, Sim I. Publication of clinical trials supporting successful new drug applications: A literature analysis. *PLoS Med*. 2008; 5(9):e191. [PubMed: 18816163]
20. Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: Review of publication and presentation. *PLoS Med*. 2008; 5(11):e217. [PubMed: 19067477]
21. Zarin D, Tse T, Sheehan J. The Proposed Rule for U.S. Clinical Trial Registration and Results Submission. *N Engl J Med*. 2015; 372:174–80. [PubMed: 25539444]
22. Chan A. Access to clinical trial data. *BMJ*. 2011; 342:d80. [PubMed: 21228020]
23. Miller J. Registering clinical trial results: the next step. *JAMA*. 2010; 303(8):773–4. [PubMed: 20179288]
24. Gopal A, Desai N, Tse T, Ross J. Reporting of noninferiority trials in Clinicaltrials.gov and corresponding publications. *JAMA*. 2015; 313(11):1163–5. [PubMed: 25781447]
25. Food and Drug Administration. [Accessed February 25, 2016] Drug Trials Snapshots. Available at: <http://www.fda.gov/Drugs/InformationOnDrugs/ucm412998.htm>
26. Boudier F, Way D, Lofstedt R, Evensen D. Transparency in Europe: A Quantitative Study. *Risk Analysis*. 2015; doi: 10.1111/risa.12386

**Figure.**

Trial search and selection.

*The 20 unmatched trials were from 8 new drug reviews: 4 had other matched trials, 4 did not. The 50 trials without results in both sources were from 21 new drug reviews: 12 had other matched trials, 9 did not.

Table 1

Comparison of ClinicalTrials.gov and FDA reviews on Drugs@FDA

Item	ClinicalTrials.gov	FDA reviews on Drugs@FDA
Description	Trial registry and results database – summary protocol information and aggregate results	Publically available FDA drug approval packages including FDA medical, statistical and other reviews
Purpose	Document clinical trials conducted and improve access to clinical trial results	Summarize basis for FDA drug approval including key efficacy and safety results
Party responsible for information	Trial sponsor or principal investigator	FDA clinical and statistical experts review sponsor-submitted data
Target audience	Registry – potential trial participants, public	FDA Advisory Committees, FDA decision makers; and researchers
	Results database – readers of the medical literature	
	Website accessible to anybody	Website accessible to anybody
Data format	Structured, tabular format with minimal narrative.	Embedded in pdf of narrative memo,
Data reported	Summary protocol information	Summaries of study designs including primary and key secondary outcomes
	Aggregate results for all primary and secondary outcomes specified in the protocol, all serious adverse events, other adverse events	Summaries of FDA analyses and/or sponsor aggregate analyses of primary and key secondary endpoints, serious adverse events, deaths
Selected limitations (relevant to this analysis)	Protocol and aggregate results information not externally validated	Only key secondary endpoints mentioned
	Analytic details reported varies across records (and sometimes missing); usually only one analytic method reported per outcome measure	Adverse events and deaths often combined across trials

Table 2

Description of randomized parallel group pivotal and supportive efficacy trials for new drugs approved by the FDA from January 2013 through July 2014 with results reported in ClinicalTrials.gov and in Drugs@FDA.

Characteristic reported in ClinicalTrials.gov (unless otherwise specified)	# Trials (n=100) unless otherwise specified
Unique drugs	32 drug products
Pharmaceutical company	24 companies as CTgov-designed lead sponsors
Drug indication	
Diabetes	42 (42%)
Pulmonary	19 (19%)
Oncology	9 (9%)
Infectious disease	8 (8%)
Endocrine (other)	7 (7%)
Pain	6 (6%)
Inflammatory arthritis	3 (3%)
Inflammatory bowel disease	3 (3%)
Female sexual dysfunction	2 (2%)
Osteoporosis	1 (1%)
Drugs@FDA designation for approval	
Pivotal	75 (75%)
Supportive	25 (25%)
Drug development phase	
Phase 1/2	1 (1%)
Phase 2	9 (9%)
Phase 3	90 (90%)
Blinding	
Double-blind	92 (92%)
Open label	8 (8%)
Study groups (median (range))	3 groups (2–12)
2	43 (43%)
3	25 (25%)
>3	32 (32%)
Comparator	
Placebo only	65 (65%)
Active only	26 (26%)
Placebo and active	8 (8%)
Other*	1 (1%)
>1 primary outcome	18 (18%)
Median no. of primary outcomes (range) per trial	
ClinicalTrials.gov	1 primary outcome (1–8)
Drugs@FDA	1 primary outcome (1–8)
Median no. of secondary outcomes (range) per trial	
ClinicalTrials.gov	7 secondary outcomes (0–227)

Characteristic reported in ClinicalTrials.gov (unless otherwise specified)	# Trials (n=100) unless otherwise specified
Drugs@FDA	5 secondary outcomes (0–94)
Median number randomized (25ile 75ile), n	562 people (124, 1554)
Comparison of basic trial data	
Total number of people randomized	
Concordant	76
Discordant	24
Median (range) in difference in total number randomized	11 people (1–331)
Median (range) of discrepancy as percent of total randomized	2% (0.1%–54%)
Cannot compare	0
Total number of people completing trial	
Concordant	69
Discordant	31
Median (range) in difference in total number completed	62 people (1–441)
Median (range) of discrepancy as percent of total completed	13% (0.2%–296%)
Cannot compare	12 (12 not included in Drugs@FDA)
Mean age	
Concordant	69
Discordant	4
Cannot compare	27 (10 not reported in CTgov, 17 not included in Drugs@FDA)
Sex distribution (n or %)	
Concordant	80
Discordant	3
Cannot compare	17 (17 not included in Drugs@FDA)

* No true control group since trial compared drug alone to drug plus steroids

Abbreviations: CTgov = ClinicalTrials.gov; Drugs@FDA = publicly accessible FDA review documents

Table 3

Primary and secondary outcome mentions, discordance, and level shifts between ClinicalTrials.gov and Drugs@FDA. “Cannot compare” means outcomes were not matched in both sources or data were not reported in at least one source.

Primary outcomes	# Measures
ClinicalTrials.gov primary outcomes	n=137
Not mentioned in Drugs@FDA *	3 (2%)
Mentioned in Drugs@FDA	134 (98%)
Discordant outcome definition **	4 (3%)
Different method of aggregation	2
Different measurement tool	1
Different time frame	1
Concordant outcome definition	130 (95%)
No result data included in Drugs@FDA	1 (1%)
Results concordant	107 (78%)
Results discordant	22 (16%)
Specified in Drugs@FDA as a....	n=137
Primary outcome	119 (87%)
Secondary outcome	0 (0%)
Tertiary or exploratory outcome	0 (0%)
Outcome level unspecified	18 (13%)
Not mentioned in Drugs@FDA *	3
Mentioned but outcome level unspecified	11
Discordant outcome definition **	4
Primary outcome analyses with data included in Drugs@FDA	n=129
Measurement population	
Concordant	104 (81%)
Discordant	12 (9%)
Cannot compare	13 (10%)
	((11 not in CTgov, 5 not in Drugs@FDA)
Crude or adjusted analysis	
Concordant	114 (88%)
Discordant	7 (5%)
Cannot compare	8 (6%)
	8 not in CTgov, 2 not in Drugs@FDA
Method of handling missing data	
Concordant	66 (51%)
Discordant	2 (2%)
Cannot compare	61 (47%)
	56 not in CTgov, 22 not in Drugs@FDA
Secondary outcomes	
ClinicalTrials.gov secondary outcomes	n=1,927

Primary outcomes	# Measures
Not mentioned in Drugs@FDA ***	866 (45%)
Mentioned in Drugs@FDA	1,061 (55%)
No result data included in Drugs@FDA	694 (36%)
Result data included in Drugs@FDA	367 (19%)
Specified in Drugs@FDA as a....	n=1,927
Primary outcome	2 (0.1%)
Secondary outcome	981 (51%)
Tertiary or exploratory outcome	20 (1%)
Outcome level unspecified	924 (48%)
Not mentioned in Drugs@FDA ***	866
Mentioned but outcome level unspecified	58

* The same "Not mentioned in Drugs@FDA " data are presented twice for clarity but are not double counted.

** The same "Discordant outcome definition" data are presented twice for clarity but are not double counted

*** The same "Not mentioned in Drugs@FDA " data are presented twice for clarity but are not double counted

Abbreviations: CTgov = ClinicalTrials.gov; Drugs@FDA = publicly accessible FDA review documents

Table 4

Discrepancies between ClinicalTrials.gov and Drugs@FDA primary outcomes: outcomes primary in both sources with discordant definitions (n=4 primary outcomes) and ClinicalTrials.gov primary outcomes not identified as “primary” in Drugs@FDA (n=14 primary outcomes). For trials with at least 1 concordant primary outcome, the concordant outcomes are in grey.

Trial ID (Drug)	ClinicalTrials.gov	Drugs@FDA	Problem
Outcomes primary in both sources with discordant definitions			
NCT00729469 Ospemifene (Osphena) for dyspareunia in postmenopausal women	1. Change from baseline in vaginal dyspareunia on a 4-point ordinal scale from none to severe - Ordinal results	Change from baseline in vaginal dyspareunia on a 4-point ordinal scale from none to severe - Means	Different methods of aggregation (2 primary outcomes): ClinicalTrials.gov provides the full ordinal distribution (which better illustrates clinical effect), but the primary outcome in Drugs@FDA approval tested means.
	2. Change from baseline in vaginal dryness on a 4-point ordinal scale from none to severe - Ordinal results • Change from baseline in the vaginal dyspareunia and dryness stratum in the percent of parabasal cells, superficial cells, and vaginal pH (6 outcomes)	Change from baseline in vaginal dryness on a 4-point ordinal scale from none to severe - Means • Change from baseline in the vaginal dyspareunia and dryness stratum in the percent of parabasal cells, superficial cells, and vaginal pH (6 outcomes)	• Same
NCT01010061 Obinutuzumab (Gazyza) for chronic lymphocytic leukemia	3. Progression-free survival Investigator-judged events	Progression-free survival Independent Review Committee-judged events	Different measurement tool: Drugs@FDA measured PFS “based on Independent Review Committee (IRC) assessments will be the basis of regulatory decisions in the United States. PFS is defined as the time from randomization to the first occurrence of progression, relapse, or death from any cause”. PFS in ClinicalTrials.gov was based on investigator-judged events.
NCT01167881 Empagliflozin (Jardiance) for type 2 diabetes	4. Mean change from baseline in HbA1C Final results of 2 year trial	Mean change from baseline in HbA1C Interim 1-year results of 2 year trial	Different time frame: Drugs@FDA-reported results incomplete (only present 1-year interim results)
Outcomes primary in ClinicalTrials.gov but not considered primary in Drugs@FDA			
NCT00833833 Pomalidomide (Pomalyst) for multiple myeloma	1. Proportion with dose-limiting toxicity 2. Progression free survival 3. Proportion with progression free survival events		Drugs@FDA did not mention Drugs@FDA specified as exploratory Note: Drugs@FDA specified objective response rate as the primary
NCT01010061 Obinutuzumab (Gazyza) for chronic lymphocytic leukemia	4. Proportion with progression-free events – Investigator-judged events		Drugs@FDA did not specify as outcome
NCT00607373 Mipomersen (Kynamro) for hyperlipidemia	5. Mean LDL (end of trial) • % Change in LDL	• % Change in LDL	Drugs@FDA did not specify as outcome • Matched primary outcome

Trial ID (Drug)	ClinicalTrials.gov	Drugs@FDA	Problem
NCT00794664 Mipomersen (Kynamro) for hyperlipidemia	6. Mean LDL (end of trial) • % Change in LDL	• % Change in LDL	Drugs@FDA did not specify as outcome • Matched primary outcome
NCT00706849 Mipomersen (Kynamro) for hyperlipidemia	7. Mean LDL (end of trial) • % Change in LDL	• % Change in LDL	Drugs@FDA did not specify as outcome • Matched primary outcome
NCT00770146 Mipomersen (Kynamro) for hyperlipidemia	8. Mean LDL (end of trial) • % Change in LDL	• % Change in LDL	Drugs@FDA did not specify as outcome • Matched primary outcome
NCT01164501 Empagliflozin (Jardiance) for type 2 diabetes	9. Mean change from baseline in HbA1c for mild renal impairment 10. Mean change from baseline in HbA1c for moderate renal impairment • Mean change from baseline in HbA1C for mild or moderate renal impairment	• Mean change from baseline in HbA1C for mild or moderate renal impairment	Drugs@FDA did not specify subgroup results as outcomes • Matched primary outcome
NCT01164501 Alogliptin/pioglitazone (Oseni) for type 2 diabetes	11. Mean change from baseline in HbA1c –Dosing groups combined • Mean change from baseline in HbA1C	• Mean change from baseline in HbA1C	Drugs@FDA did not specify combined dosing group results as outcome • Matched primary outcome
NCT00829166 Alogliptin/pioglitazone (Kadcyla) for type 2 diabetes	12. Proportion alive at 1 year 13. Proportion alive at 2 years • Progression free survival • Overall survival	• Progression free survival • Overall survival	Drugs@FDA did not mention Drugs@FDA did not mention • Matched primary outcome • Matched primary outcome
NCT00808132 Estrogen/bazedoxifene (Duavee) for osteoporosis	14. Endometrial hyperplasia (using an alternate definition) • Endometrial hyperplasia	• Endometrial hyperplasia	Drugs@FDA recommended a sensitivity analysis with alternate definition (i.e. using the most severe rather than the majority pathology reading for disagreement) • Matched primary outcome

Abbreviations: CTgov = ClinicalTrials.gov; Drugs@FDA = publicly accessible FDA review documents

Table 5

Consistency of results reported in ClinicalTrials.gov primary outcome, serious adverse events, and death with Drugs@FDA reviews. “Cannot compare” means outcomes were not matched in both sources or data were not reported in at least one source. “Not in Drugs@FDA” means either ClinicalTrials.gov outcome did not match Drugs@FDA (not mentioned by Drugs@FDA or discordant definition) or data not reported. “Not in CTgov” only means data not reported.

Result	Trials		Outcome results
Primary outcome (n)	n=100 trials		n=137 outcomes
Number of people analyzed			
Concordant	71 (71%)		87 (64%)
Discordant	23 (23%)		42 (31%)
Cannot compare	6 (6%)		8 (6%) [8 not included in Drugs@FDA]
Result data in each group			
Concordant	78 (78%)		107 (78%)
Discordant	14 (14%)		22 (16%)
Cannot compare	8 (8%)		8 (6%) [8 not included in Drugs@FDA]
Treatment effect size			
Concordant	69 (69%)		83 (61%)
Discordant*	5 (5%)		6 (4%)
Cannot compare	26 (26%)	48 (35%) [40 not included in Drugs@FDA, 38 not reported in CTgov]	
95% Confidence interval for treatment effect			
Concordant	57 (57%)		69 (50%)
Discordant	5 (5%)		6 (4%)
Cannot compare	38 (38%)	62 (45%) [51 not included in Drugs@FDA, 46 not reported in CTgov]	
p value for treatment effect			
Concordant	67 (67%)		87 (64%)
Discordant	3 (3%)		3 (2%)
Cannot compare	30 (30%)	47 (34%) [27 not included in Drugs@FDA missing, 37 not reported in CTgov]	
Serious adverse events			
1 serious adverse event in either source	96 (96%)		
Results reported in;			
ClinicalTrials.gov only	82 (82%)		
Drugs@FDA only	0 (0%)		
Both sources	14 (14%)		
No serious adverse event reported	4 (4%)		
ClinicalTrials.gov adverse events	96 trials		4,983 adverse events named
Results mentioned in Drugs@FDA	14 trials		145 adverse events named (3% of CTgov total)
Serious adverse events in both sources	n=14 trials		n=145 adverse events
Number of people analyzed for event			
Concordant	14 (100%)		145 (100%)

Result	Trials	Outcome results
Adverse event results		
Concordant	7 (50%)	106 (73%)
Discordant	7 (50%)	39 (27%) Median difference (range)
More events in Drugs@FDA		29 (20%) 2 (1–58) events
More events in ClinicalTrials.gov		10 (7%) 2 (1–12) events
Death		
1 death in either or both sources	62 (62%)	
Result reported in:		
ClinicalTrials.gov only	16 (16%)	
Drugs@FDA only	21 (21%)	
Both sources	25 (25%)	
No deaths reported	38 (38%)	
Death results in both sources	n=25 trials	
Concordant	8 (32%)	
Discordant	17 (68%)	Median difference (range)
More deaths reported in Drugs@FDA	15 (60%)	4 (1–104)
More deaths in ClinicalTrials.gov	2 (8%)	2 (1–3)

* Nominal difference < 10% relative difference (absolute mean difference of 0.007 to 1.2)

Abbreviations: CTgov = ClinicalTrials.gov; Drugs@FDA = FDA publicly accessible review documents, Percentages may not add to 100% because of rounding