Reporting Discrepancies between the ClinicalTrials.gov Results Database and Peer Reviewed Publications

Reporting Discrepancies between the ClinicalTrials.gov Results Database and Peer Reviewed Publications

Authors:

*Daniel Hartung, PharmD, MPH[1], Deborah A. Zarin, MD[2], Jeanne-Marie Guise, MD, MPH[3], Marian McDonagh, PharmD[3], Robin Paynter, MLS[3], Mark Helfand, MD, MS, MPH[3,4]

1. Oregon State University / Oregon Health & Science University College of Pharmacy; 3303 SW Bond Ave; CH12C; Portland, OR; 97239
   Phone: 503-706-9192
   Email: hartungd@ohsu.edu

2. Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Bethesda MD;

3. Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR, USA;

4. Portland VA Medical Center, Portland, OR.

*Corresponding Author

Acknowledgements: Luke Middleton, BS for developing abstraction tool and database. Kaylee Williamson, BS PharmD candidate, for assisting in sample validation

Funding: Daniel Hartung receives support from an AHRQ career development award (K12 HS019456).

Manuscript word count: 3512

Background: Result summaries are now required to be reported in ClinicalTrials.gov for many trials of drugs and devices.

Purpose: To evaluate the consistency of reporting in trials that are both registered in the ClinicalTrials.gov results database and published in the literature.

Data Sources: ClinicalTrials.gov results database, matched publications identified through both ClinicalTrials.gov and a manual search of two electronic databases.

Study Selection: 10% random sample of Phase III or IV trials with results in the ClinicalTrials.gov results database, completed before January 1, 2009, with two or more arms.

Data Extraction: One reviewer extracted data from ClinicalTrials.gov results database and matching publications. A subsample was independently verified. Basic design features and results were compared between reporting sources and discrepancies were summarized.

Data Synthesis: Of 110 reviewed trials with results, most were industry-sponsored, parallel design, drug studies. The most common inconsistency was the number of secondary outcome measures reported (80%). There were 16 trials (15%) that reported the primary outcome description inconsistently and 22 (20%) in which the primary outcome value was reported inconsistently. A total of 38 trials inconsistently reported the number of individuals with a serious adverse event (SAE), of which 33 (87%) reported more SAEs in ClinicalTrials.gov. Among the 84 trials that reported SAEs in ClinicalTrials.gov, 11 publications did not mention SAEs, 5 reported SAEs as zero or not occurring, and 21 reported a different number of SAEs. In 29 trials that reported deaths in ClinicalTrials.gov, 28% differed with the matched publication.

Limitations: Small sample that includes earliest results posted to the database and therefore may reflect inexperience with the submission process.

Conclusions: Reporting discrepancies between the ClinicalTrials.gov results database and matching publications are common. It is unclear which reporting source contains the most accurate account of trial results. ClinicalTrials.gov may provide a more comprehensive description of trial adverse events than the publication.

Primary funding source: AHRQ career development award (K12 HS019456)

Background

Medical decision makers who use clinical trial evidence most often rely on findings that are published in peer-reviewed journals. Selective reporting of clinical trial results is a well-documented problem that raises concerns about using journal publications.(1) Clinical trial registration is one mechanism aimed at reducing the impact of dissemination biases. Although numerous clinical trial registries exist, the single largest publicly accessible trial registry, and the only one with a results database, is ClinicalTrials.gov.(2) Administered through the National Library of Medicine (NLM), ClinicalTrials.gov was developed to provide the public with a web-based, searchable source of information about trials conducted within the US. In September of 2007 the Food and Drug Administration Amendments Act (FDAAA) was passed, greatly expanding the legal requirements for trial registration and mandating the creation of a publically accessible clinical trial results database (Section 801).(3) According to FDAAA Section 801 as of September 2008, basic summary results must be submitted for certain trials (called "applicable clinical trials" in the statute.) "Applicable clinical trials" include most phase II through IV trials of drugs, devices, or biologics regulated by the FDA having at least one site in the US or conducted under an investigational new drug application or investigational device exemption.(4) Several different elements are required to be reported including: number of participants entering and completing the study, number of participants analyzed, demographic data such as age and sex, summary results for all pre-specified primary and secondary outcome measures, and anticipated and unanticipated adverse events by organ system. Results are generally required to be reported within 1 year of study completion, although submission may be delayed if the drug or device is not yet approved or if an application for a new use is to be submitted.

ClinicalTrials.gov results database has the potential to be a great asset for clinicians, patients, and researchers, however the ultimate validity of posted results is unclear. In contrast to the scientific scrutiny trials undergo during peer review for journals, results posted to ClinicalTrials.gov go through a quality assurance process focusing on internal consistency and logic. Although a gold standard repository of clinical trial results does not exist, inconsistencies between the ClinicalTrials.gov results database and other sources of clinical trial data may be suggestive of validity problems in one or both sources. The goal of this study was to assess the consistency of results reporting in the ClinicalTrials.gov results database compared to those summarized in peer-reviewed journal publications.

Methods

Trial Selection

Studies were eligible for inclusion if they posted results to ClinicalTrials.gov, were interventional, and were phase III or IV. To allow sufficient time for publication, we limited our search to trials with a primary completion date before January 1, 2009 or start date before July 1, 2008 if the primary completion date field was not populated. Completed trials with results were sequenced in random order using Microsoft Excel™ (Redmond, WA) and screened for the presence of a matching publication until a 10% random sample of trials with results was obtained. Trials were excluded if they did not assign participants to 2 or more interventional groups.

Matching publications were identified in a sequential process first examining citations provided within ClinicalTrials.gov and then a manual search of two electronic bibliographic databases. PubMed citations embedded within ClinicalTrials.gov can be provided by the investigator or by NLM based on matching National Clinical Trial (NCT) identifiers.(5) A publication was considered a match if the intervention was the same and if one or more arms of the trial had an identical number of enrolled subjects. If relevant studies were not identified using citations provided within ClinicalTrials.gov, an electronic search of MEDLINE and the Cochrane Central Register of Controlled Trials was conducted using the study intervention(s), condition, principal investigator (if supplied) and date of trial completion as search criteria.

Data Abstraction and Comparisons

The following elements were abstracted and compared between the ClinicalTrials.gov results record and its corresponding publication(s): trial design, number of arms, primary outcome measure (POM) description(s), secondary outcome measure (SOM) descriptions, total enrollment, and primary outcome results. We also abstracted the number of individuals affected by at least 1 adverse event (AE) and the number of individuals at risk, as reported to ClinicalTrials.gov. Comparisons of counts (i.e. enrollment, participants analyzed for primary outcome, number with an AE) were considered discrepant if they were not an exact match. The primary outcome result was required to be consistent to one decimal place. In cases of multiple publications, inconsistencies between the ClinicalTrials.gov result record and descriptions in any of the associated publications were considered a discrepancy.

POM description inconsistencies were classified using an existing framework describing the specificity of outcome reporting in ClinicalTrials.gov.(6) The POM could deviate entirely in the domain measured or number of POMs reported, the measurement tool used (e.g. change in LDL versus change in total cholesterol), how the measure was used (e.g. % change from baseline versus absolute value), or the method of aggregation (e.g. HbA1c

<7% versus <8%). SOMs were considered consistent if they were mentioned in either the results or methods section of the publication and were listed in the ClinicalTrials.gov results record. For trials with multiple publications, we considered the aggregate number of SOM(s) across all associated publications. When evaluating POM reporting consistency, we first determined whether the description was consistent in both sources. When the POM(s) was consistent, we looked for discrepancies in either the reported value (e.g. mean response, count with outcome) or the number of individuals analyzed for the outcome (e.g. denominator, number analyzed). For trials where more than one POM was specified in both sources, any inconsistency in either result numerator or denominator was considered a discrepancy. If discrepancies in trial features resulted in downstream inconsistencies, only the highest order feature was compared to avoid double counting.

AEs did not become a mandatory reporting element until September 2009, and are summarized in the ClinicalTrials.gov results record in two tables: serious AEs (SAE) and other, non-serious, AEs (OAE). The FDA defines SAEs as any event that results in death, is life-threatening, requires or extends hospitalization, results in significant incapacity or interferes with normal life functions, or causes a congenital anomaly or birth defect.(7) We compared the total number of SAEs reported in ClinicalTrials.gov to the total reported in the corresponding publication(s). In cases where the SAE counts differed, we compared the risk difference (experimental arm risk – control arm risk) reported in ClinicalTrials.gov to the published estimate. For trials with multiple experimental arms, we selected the arm of primary interest stated in the paper or if multiple FDA approved dosing arms were assessed, we combined these to compare against the control. For OAEs, we restricted our comparison to specific AEs that could be matched to the publication without ambiguity and which were not also reported as an SAE in order to eliminate the possibility of double counting participants who may have had both a serious and non-serious AE. We distinguished publications reporting only treatment-related (attributable) AEs because ClincalTrials.gov requires reporting AEs regardless of attribution. Finally, we compared the number of deaths reported between sources. In ClinicalTrials.gov deaths can be reported as an outcome, in the participant flow section, or as an SAE. If death was not a primary our secondary outcome, we compared deaths reported in the Participant Flow or SAE Section of ClinicalTrials.gov to the number reported in the publication. We classified the sources as discrepant only if counts of death reported in both sources differed.

A second reviewer independently assessed reporting discrepancies between the ClinicalTrials.gov results record and the matched publication in a 20% random sample (22 trials) for all comparisons. Agreement between the primary (DH) and secondary abstractor (KW) was very high with a Kappa averaged across categories of 0.98 and no single category with a Kappa below 0.91.

Results

Figure 1 describes the flow of trials from the initial ClinicalTrials.gov candidate pool to final study sample. A total of 1,669 phase III and IV trials with posted results were initially identified through a query of ClinicalTrials.gov on February 15, 2011. After excluding trials with a primary completion date after January 1, 2009 and those not completed or terminated, 1,120 trials remained. We randomly screened 357 potentially includable trials until a 10% sample (n=110) was achieved. Three trials reported results in multiple publications. Table 1 describes the characteristics of the 110 matched trials and the 195 unmatched trials. A majority of studies were industry funded, parallel design, trials of drugs. Unmatched trials were more likely to investigate something other than a drug or device and less likely to be a cardiovascular trial. Twenty-nine trials (26%) described more than one POM in ClinicalTrials.gov.

Table 2 summarizes reporting discrepancies between the ClinicalTrials.gov results database and the matching publication(s). Sixteen trials (15%) had discrepant POM descriptions. In nine (56% of 16) of these, POM descriptions reported in ClinicalTrials.gov were not reported as POMs in the publication. POM descriptions in all but one of these studies were reported as SOMs in the publication. The only publication not reporting POM descriptions reported in ClinicalTrials.gov omitted three POMs related to pharmacokinetic outcomes (NCT00158600). On average, the publication listed 2.4 more SOMs than the ClinicalTrials.gov results database. Three trials (3%) reported enrollment results inconsistently. The inconsistencies in enrollment reflected differences in up to 14% of the total enrollment.

There were 22 (20%) trials that inconsistently reported the primary outcome result (Supplemental Table 1). Seven trials (32% of 22) reported larger treatment effects in the publication relative to ClinicalTrials.gov and two trials (9% of 22) reported larger treatment effects in ClinicalTrials.gov. For the seven trials with larger treatment effects reported in the publication, the median relative difference in treatment effect was 10% (min - max1% to 270%). On an absolute scale most discrepancies were small and did not affect the statistical significance of the reported results.

Of the 104 ClinicalTrials.gov entries reporting information about SAEs, 84 trials reported at least one SAE. Among these 84 entries, 11 publications did not mention SAEs and 5 reported SAEs as zero or not occurring. In total, 38 trials had SAE reporting discrepancies (Supplemental Table 2). For 33 of these trials (87% of 38 trials) more SAEs were reported in the ClinicalTrials.gov registry than in the publications. Four trial publications reported only treatment-related SAEs. When the risk of SAEs was higher for the experimental arm in ClinicalTrials.gov, 17 of 20 (85%) publications reported SAE risks more favorable to the experimental arm. For three of these trials, the

publication reported that the risk of an SAE was lower in the experimental arm compared control. Two publications reported attenuated SAE risk differences that would imply 10 and 500 more patients needing to be treated with the intervention in order to cause one SAE. For the remainder of publications (n=12), SAEs were either reported as zero (n=3) or not reported (n=9). For these twelve trials the number needed to harm for the intervention ranged from 5 to 125 (median = 37). When the SAE rate was higher for the control group in ClinicalTrials.gov, 7 of 15 (47%) publications reported differences that were even more favorable towards the intervention. In 3 trials the SAE risks could not be compared clearly between groups.

A total of 35 (34 %) of 95 trials that reported one or more OAEs in their ClinicalTrials.gov entry had at least one reporting discrepancy. In 19 trials (54% of 35) the publication reported fewer OAEs than the ClinicalTrials.gov record. Eleven (31% of 35) trials had more individuals with an OAE in the publication compared to ClinicalTrials.gov, and five trials (14% of 35) had reporting differences in both directions. There were nine trials that reported zero OAEs in ClinicalTrials.gov, of which five (56% of 9) reported one or more OAE in the publication.

There were 81 trials that did not report on deaths in ClinicalTrials.gov. Of these, 14 (17%) had deaths reported in the matched publication. In 16 of 29 trials that reported deaths in ClinicalTrials.gov, the publication reported the same number of deaths (Supplemental Table 3). In five others, deaths were reported as counts in one source and as survival analysis (without counts) in the other source. Counts of death were discrepant in the remaining eight trials. Death was a primary outcome or part of a composite primary outcome in two of these trials. One of these had a relatively small discrepancy (NCT00379769); in the other, a large trial of irbesartan (NCT00095238), the discrepancy was large (1003 deaths in ClinicalTrials.gov vs. 881 in the publication) but may have been due to different lengths of follow-up. There were six discrepancies among trials that reported deaths in the Participant Flow or SAE Section of ClinicalTrials.gov. In three of five trials that reported deaths in the SAE section, more deaths were reported in the publication than in the results database.

Discussion

Reporting discrepancies between the ClinicalTrials.gov results database and matching publications were common for several key trial attributes and results. Overall 20% of trials inconsistently reported the primary outcome result, although only a few could be considered potentially meaningful discrepancies. Descriptions of POMs were different between the ClinicalTrials.gov results database and publication 15% of the time, most often when one or more primary outcomes was dropped from primacy. This estimate is lower than other studies that have explored inconsistencies between clinical trial protocols and published results (62%)(8) or trial registry entries and journal publications (31%).(9, 10) The lower proportion of discrepant POMs found in our sample may reflect improved reporting when summary results as opposed to just a description of outcomes (as a requirement for registration) are recorded. Like previous studies, we found that 80% of trials contained a SOM reporting discrepancies.(10, 11) Huic et al compared 9 World Health Organization (WHO) Minimum Data Set trial registration elements from ClinicalTrial.gov to corresponding publications and found 65% differed on how SOMs were reported.(11) Similar to our finding, the most common SOM differences they noted were outcomes listed in the publication but missing from ClinicalTrials.gov. While this may reflect incomplete reporting in the ClinicalTrials.gov database, it could also indicate the misrepresentation of post hoc analyses as pre-specified SOMs in the publication.

Adverse events were reported inconsistently in over one-third of trials. Omission or under-reporting in the publication was the predominant inconsistency, with the vast majority of discrepant trials reporting fewer SAEs in the publication than in ClinicalTrials.gov. Under-reporting of AEs, even when not differential between groups, is of great concern because it can minimize impressions of the overall safety of an intervention.(12) Most inconsistencies were either complete SAE non-reporting or differences in 10 or fewer individuals that did not alter the direction of risk. However, when discrepant trials reported an increased SAE risk with the intervention relative to the control group in ClinicalTrials.gov the published account of this risk was almost universally less pronounced (i.e. more favorable to the intervention). In three trials the discrepancies reversed the direction of risk. For example, in one trial (NCT00323492) the publication reported a 4.5% decrease in the risk for an SAE with the intervention (emtricitabine/tenofovir) in contrast to ClinicalTrials.gov which reported an 8.2% increase in SAE risk. Publications infrequently provided detailed SAE descriptions and it was unclear if adverse events classified as serious in ClinicalTrials.gov were reclassified in the publication. Although there is some inherent subjectivity in the FDA's standard criteria for an SAE, determination of adverse events as serious should not change depending on reporting sources.(7) We identified two trial publications in type 2 diabetes with ambiguous and potentially misleading

reporting of serious hypoglycemic episodes. In one trial (NCT00313313), the publication described no cases of hypoglycemia judged to be an SAE, however the ClinicalTrials.gov entry recorded 2 patients having an SAE of hypoglycemia (one in each dosing arm of the study drug). Another publication (NCT00494013) mentioned "two of seven patient-reported severe hypoglycemia episodes as serious AEs" but failed to attribute to a specific study group. The ClinicalTrials.gov record for this trial indicated these events occurred in the active treatment arm (long-acting insulin). A similar pattern was also observed in the reporting of OAEs, although a more focused examination of specific adverse events by disease state is needed.

In our sample, only a quarter of trials reported on deaths in ClinicalTrials.gov. It is likely that, in most cases, omission of death data from ClinicalTrials.gov occurred because there were no deaths in trial. However, in 17% of trials that failed to report deaths in ClinicalTrials.gov, deaths were documented in the publication. When ClinicalTrials.gov reported deaths, the number was inconsistent with the publication in about one quarter of trials. Reporting of deaths was more consistent when they were included in the Outcomes section of ClinicalTrials.gov. Among 14 trials where death was a prespecified outcome, we found only one discrepancy that could be considered meaningful. While our sample was small, our results suggest that reporting of deaths in the SAE section of ClinicalTrials.gov was often inconsistent. Earley and colleagues note that ClinicalTrials.gov does not a have a uniform template for how deaths are reported and internal consistency is sometimes problematic.(13)

Under-reporting of adverse events is a major concern because it can distort how decision-makers balance the benefits and harms of medical interventions. Even when the inconsistencies are minor in individual studies, as is the case for several of the trials analyzed, these distortions can be amplified when results are combined within systematic reviews.(12, 14) Suboptimal AE reporting may relate to space restriction imposed by journals, the use of study designs that poorly measure harms, or purposeful concealing of unfavorable data.(15-17) It is unclear from our study why some trials reported AE data more consistently than others. In general, however, ClinicalTrials.gov appears to provide a more comprehensive summary of AEs.

This study has several limitations. First, the study sample consisted of trials that were completed by January 1, 2009. These trials were likely among the first posted to the ClinicalTrials.gov results database and may contain inconsistencies that reflect investigators' inexperience with the system for entering results. Reporting consistency may be improving as investigators become more familiar with the data submission process. The ClinicalTrials.gov registry allows investigators to change their registered protocol using a track change function that is archived in a companion website (http://clinicaltrials.gov/archive/). We did not evaluate changes in POMs or SOMs archived over time relative to what was reported in the matched publication, only what was reported in the

results record. Modification of registered clinical trial protocols, specifically POM and SOM additions or deletions, is common prior to publication and may partially explain why POM discrepancies were lower in our study relative to other estimates.(11) Although we attempted to find matching publication through both citations within ClinicalTrials.gov or through search of two electronic bibliographic databases, some matches may have been overlooked. Finally, many of discrepancies were only observed in a small number of trials and estimates should be regarded as preliminary.

This analysis has documented that different instances of reporting results from a given trial frequently lead to discrepant data. Although there are many possible explanations for such discrepancies, these findings contribute to the growing sense that the process of taking the initially collected "raw" participant level data and deriving the ultimately reported aggregate or "summary" data involves a series of decisions that are not entirely prespecified or objective; different iterations of the process thus produce somewhat different results. It is uncertain if discrepancies we observed represent deliberate misrepresentation, reporting carelessness, the influence of journal editors, or simply an evolution of investigators' thinking or analytic approach over time. For example, the process of peer-review may introduce modifications in how results are analyzed or reported that may contrast with data submitted to the ClinicalTrials.gov results registry. Many of the primary outcome result discrepancies appear to be small inconsistencies that may be errors in data entry or be the result of additional or modified analyses requested by the specific journal. However, it is important to note that our analysis examined the summary metric (e.g., mean response rate) and not the associated statistical analysis. We believe that it is uncommon for peer review to lead to actual changes in the data, as opposed to changes in the types of statistical analyses and resulting inferences that are considered appropriate, although future research might examine this issue further. If investigators do not (or cannot) provide consistent quantitative summaries of the fundamental features of their trials, then one must question how accurate either reporting source could be. Because there is no gold standard clinical trial reporting source, for now the only possible path to resolving discrepancies is to seek clarifications from the investigator. Although FDA commonly use independent analysis of participant-level data for product reviews, clinicians, patients, and other decision-makers generally rely on summary data from journal articles and other sources to inform their decisions. While there is great interest in making participant-level clinical trial data publically available for independent analysis and dissemination, models that balance public and private data use concerns are just now beginning to emerge.(18, 19) It remains unclear if greater reporting transparency, up to and including access to participant-level data, will improve the reliability, and ultimately the validity, of clinical trial research for decision makers.

Table 1: Description of matched and unmatched trial characteristics as reported in ClinicalTrials.gov

Characteristic*	Matched Sample (n=110)	Percent	Unmatched Sample (n=195)	Percent
Median enrollment (IQR)	352	(537)	263	(486)
Study Condition[A]				
Cancer	8	7%	5	3%
Cardiovascular	28	26%	17	9%
Endocrine & Metabolic	13	12%	19	10%
Mental Health	15	14%	14	7%
Respiratory	6	6%	29	15%
Other	40	36%	111	57%
Sponsor[B]				
Government	1	1%	0	0%
Industry	98	89%	183	94%
Other	11	10%	12	6%
Intervention Type[C]				
Device	3	3%	10	5%
Drug	104	95%	158	81%
Other	3	3%	27	14%
Design				
Parallel	104	95%	179	92%
Cross-over	6	6%	13	7%
Factorial	0	0%	3	2%
Allocation Arms				
2	73	66%	138	71%
3	22	20%	34	17%
>3	15	14%	23	12%
More than 1 primary outcome	29	26%	52	27%

*All characteristics except for applicable clinical trial, ICMJE status, and study condition were abstracted from the ClinicalTrials.gov registry.

[A]Trial conditions were classified into one of 6 disease categories listed.

[B]If multiple Sponsors were listed, the information provider was considered the primary sponsor. 'Other' sponsors included foundations, health system, universities, etc.

[C]Intervention categorization was based on hierarchy of interventions: device, drug, or other (e.g. any trial with an arm with a device was categorized as a device trial, "other" trials had no devices or drugs).

IQR=interquartile range

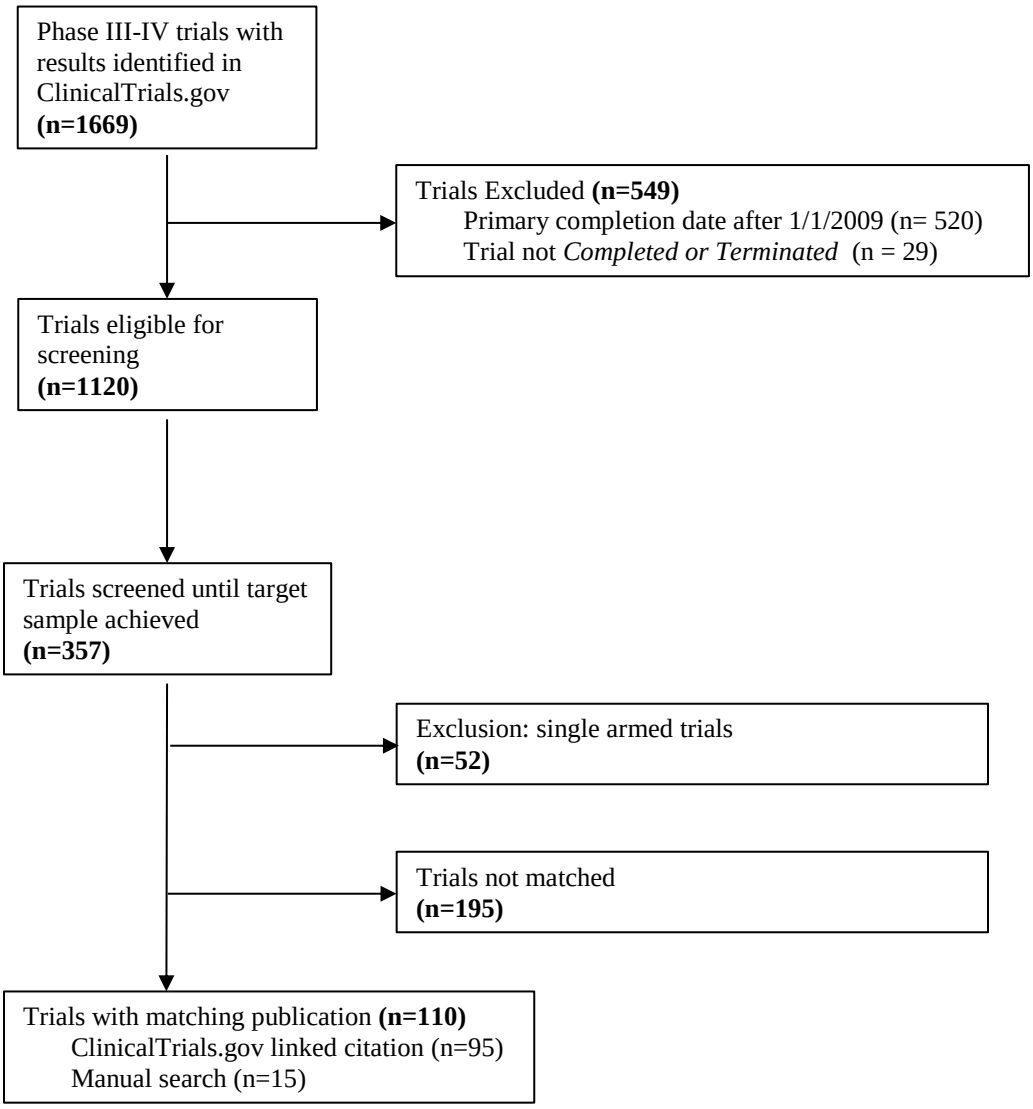
Table 2: Summary of reporting discrepancies between ClinicalTrials.gov records (n=110) and matched publications (n=113)

	Count	Percent
Design (n=110)		
Trial design	0	0
Number of Arms	1	1%
POM description		
Number of POMs not consistent	9	8%
Different measurement tools used	1	1%
Measurement tool used differently	6	6%
SOM descriptions		
Trials with discrepant number of SOMs	88	80%
Mean (SD) number of secondary outcomes in publication	9.6 (7.9)	
Mean (SD) number of secondary outcomes in ClinicalTrials.gov results record	7.2 (8.1)	
Difference (95% confidence interval)	2.4 (1.1 - 3.8)	
Results (n=110)		
Total enrollment	3	3%
min, max discrepancy percent of maximum enrollment	1% - 14%	
Primary outcome results		
Larger treatment effect in publication	7	32% [a]
Larger treatment effect in ClinicalTrials.gov	2	9% [a]
Other discrepancies*	13	59% [a]
Adverse Events (n=104); 6 trials posted prior to AE reporting requirement		
Serious AEs		
Trials with >=1 SAE reported in ClinicalTrials.gov (n=84)		
Discrepant: not reported in publication	11	11% [b]
Discrepant: reported as zero or not occurring in publication	5	5% [b]
Discrepant: different number reported in publication	21	20% [b]
Trials with 0 SAE reported in ClinicalTrials.gov (n=20)		
Discrepant: *>=1* SAE reported in publication	1	1% [b]
Other AEs		
Trials with >=1 OAE reported in ClinicalTrials.gov (n=95)		
Discrepant: comparable categories with differential reporting	35	34% [b]
Discrepant: AE not reported in publication	1	1% [b]
Trials with 0 OAE reported in ClinicalTrials.gov (n=9)		
Discrepant: *>=1* OAE reported in publication	5	5% [b]
Deaths		
Trials reporting deaths in ClinicalTrials.gov (n=29)		
Discrepant: different than number reported in publication	7	24% [d]
Discrepant: not reported in publication	1	3% [d]
Trials not reporting deaths in ClinicalTrials.gov (n=81)		
One or more deaths reported in publication	14	17% [c]
Zero deaths reported in publication	28	35% [c]
Deaths not reported in publication	39	48% [c]

SD=standard deviation, POM=primary outcome measure, SOM=secondary outcome measure, AE= adverse event; SAE= serious adverse event, OAE=Other adverse event, a=denominator is 22 trials with primary outcome result discrepancy; b=denominator is 104 trials reporting SAE; c=denominator is 81 trials; d=denominator is 29 trials

*7 trials (NCT00886600, NCT00308711, NCT01218958, NCT00432237, NCT00313820, NCT00452426, NCT00337727) had inconsistent analysis denominators that did not impact reported outcomes, 2 trials (NCT00852917, NCT00422734) had multiple primary outcomes where direction of discrepancy differed between outcomes, 2 trials (NCT00287053, NCT00806403) had transposition errors where values or denominators were reversed between arms, 1 trial (NCT00029172) did not report the outcome by treatment group in ClinicalTrials.gov, 1 trial (NCT00494013) had a discrepancy in reported outcome values for each arm but the differences between arms were consistent

Figure 1: Summary of trial selection



References:

1. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, et al. Dissemination and publication of research findings: an updated review of related biases. Health technology assessment (Winchester, England). 2010;14(8):iii, ix-xi, 1-193.

2. Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DAB. Issues in the Registration of Clinical Trials. JAMA: The Journal of the American Medical Association. 2007;297(19):2112-20.

3. Zarin DA, Tse T. Moving Toward Transparency of Clinical Trials. Science. 2008;319(5868):1340-2.

4. Tse T, Williams RJ, Zarin DA. Reporting "Basic Results" in ClinicalTrials.gov. Chest. 2009;136(1):295-303.

5. Clinical Trial Registry Numbers in MEDLINE®/PubMed® Records. US National Library of Medicine2011; Available from: http://www.nlm.nih.gov/bsd/policy/clin_trials.html.

6. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov Results Database — Update and Key Issues. New England Journal of Medicine. 2011;364(9):852-60.

7. US Department of Health and Human Services. Food and Drug Administration. 21 CFR 312.32(a). http://www.gpo.gov/fdsys/pkg/CFR-2013-title21-vol5/pdf/CFR-2013-title21-vol5-part312.pdf (accessed 10/28/2013).

8. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials. JAMA: The Journal of the American Medical Association. 2004;291(20):2457-65.

9. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of Registered and Published Primary Outcomes in Randomized Controlled Trials. JAMA: The Journal of the American Medical Association. 2009;302(9):977-84.

10. Ewart R, Lausen H, Millian N. Undisclosed Changes in Outcomes in Randomized Controlled Trials: An Observational Study. The Annals of Family Medicine. 2009;7(6):542-6.

11. Huić M, Marušić M, Marušić A. Completeness and Changes in Registered Data and Reporting Bias of Randomized Controlled Trials in ICMJE Journals after Trial Registration Policy. PLoS ONE. 2011;6(9):e25258.

12. Fu R, Selph S, McDonagh M, Peterson K, Tiwari A, Chou R, et al. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. Ann Intern Med. 2013;158(12):890-902.

13. Earley A, Lau J, Uhlig K. Haphazard reporting of deaths in clinical trials: a review of cases of ClinicalTrials.gov records and matched publications-a cross-sectional study. BMJ open. 2013;3(1).

14. Carragee EJ, Hurwitz EL, Weiner BK. A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. The spine journal : official journal of the North American Spine Society. 2011;11(6):471-91.

15. Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. Arch Intern Med. 2009;169(19):1756-61.

16. Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. JAMA. 2001;285(4):437-43.

17. Ioannidis JP. Adverse events in randomized trials: neglected, restricted, distorted, and silenced. Arch Intern Med. 2009;169(19):1737-9.

18. Krumholz HM, Ross JS. A Model for Dissemination and Independent Analysis of Industry Data. JAMA: The Journal of the American Medical Association. 2011;306(14):1593-4.

19. Ross JS, Krumholz HM. Ushering in a new era of open science through data sharing: the wall must come down. JAMA. 2013;309(13):1355-6.