January 2015

Reporting Of Results In Clinicaltrials.gov And High-Impact Journals: A Cross-Sectional Study

Jessica Elizabeth Becker
Yale School of Medicine, becker.jes@gmail.com

REPORTING OF RESULTS IN CLINICALTRIALS.GOV AND HIGH-IMPACT JOURNALS: A CROSS-SECTIONAL STUDY

A Thesis Submitted to the
Yale University School of Medicine
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Medicine

By

Jessica Elizabeth Mirabito Becker

Class of 2015

ABSTRACT

REPORTING OF RESULTS IN CLINICALTRIALS.GOV AND HIGH-IMPACT JOURNALS: A CROSS-SECTIONAL STUDY.

Jessica E. Becker, Harlan M. Krumholz, Gal Ben-Josef, and Joseph S. Ross. Section of General Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT.

In 2007, the FDA Amendments Act expanded requirements for ClinicalTrials.gov, a public clinical trial registry maintained by the U.S. National Library of Medicine, mandating results reporting within 12 months of trial completion for all FDA regulated drugs. We compared clinical trial results reported on ClinicalTrials.gov with corresponding published articles. We conducted a cross-sectional analysis of clinical trials published from July 1, 2010 through June 30, 2011 in high impact journals (impact factor ≥10) that were registered and reported results on ClinicalTrials.gov. We compared trial results reported on ClinicalTrials.gov and within published articles for the following: cohort characteristics, trial intervention, primary and secondary efficacy endpoint definition(s) and results, and adverse events. Of 95 included clinical trials registered and reporting results on ClinicalTrials.gov, there were 96 corresponding publications, among which 95 (99%) had at least one discrepancy in reporting of trial details, efficacy results, or adverse events between the two sources. When comparing reporting of primary efficacy endpoints, 132 (85%) were described in both sources, 14 (9%) were described only on ClinicalTrials.gov, and 10 (6%) only within articles. Results for 30 of 132 (23%) primary endpoints could not be compared because of reporting differences between the two sources (e.g., tabular versus graphics); among the

remaining 102, reported results were discordant for 21 (21%), altering interpretations for 6 (6%). When comparing reporting of secondary endpoints, 619 (30%) were described in both sources, 421 (20%) were described only on ClinicalTrials.gov, and 1049 (50%) only within articles. Results for 228 of 619 (37%) secondary endpoints could not be compared; among the remaining 391, reported results were discordant for 53 (14%). Among published clinical trials that were registered and reported results on ClinicalTrials.gov, nearly all had at least one discrepancy in reported results, including a fifth among primary endpoints. Our findings question the accuracy of both sources and raise concerns about the usefulness of results reporting to inform clinical practice and future research efforts.

ACKNOWLEDGEMENTS

I would, first and foremost, like to thank Dr. Joseph Ross, who has been the most supportive of mentors throughout my medical school career. He has guided me through all stages of the academic process, from brainstorming to data collection to presentation practices, and has served as an exemplary role model through his own work. Thank you for all of your mentorship! My perspective as a physician and researcher, and my career as a hopeful academic, will always reflect your influence!

I would also like to thank Dr. Harlan Krumholz for his insights, assistance, and support with the project and with my research trajectory during medical school. Additionally, I would like to thank Dr. Gal Ben-Josef, for her support and work on the project.

I would like to acknowledge the Yale University School of Medicine for its funding support for my research, both during my first-year summer and for my presentation at the Peer Review Congress in 2013.

Finally, I would like to thank my parents and sister for their unending support and helping me to get where I am today!

NOTE ABOUT THE TEXT

I am grateful to have had the opportunity to publish the work that follows in my thesis in several places. The main publication associated with my work was in *JAMA* in 2014. Additional publications include a Letter to the Editor at *Annals of Internal Medicine* and an online blog post publication on the *JAMA Internal Medicine Blog*. Please see a list of these citations below.

I was also fortunate to be able to present my work as a plenary podium presentation at the 7th International Congress on Peer Review and Biomedical Publication in September 2013 in Chicago, as well as at the 9th Annual Klingenstein Third Generation Foundation Child and Adolescent Psychiatry National Conference at the Mayo Medical School in February 2015.

Citations:

1. **Becker JE,** Krumholz HM, Ben-Josef G, and Ross JS. "Reporting of Results in ClinicalTrials.gov and High-Impact Journals." *JAMA* 311(10): 12 March 2014, 1063-5.

2. **Becker JE** and Ross JS. "Trial Registries, Journals, and Selective Endpoint Reporting: Still Biased After All These Years." *JAMA Internal Medicine Blog*: 21 April 2014. Internet: http://internalmedicineblog.jamainternalmed.com/2014/04/21/trial-registries-journals-and-selective-endpoint-reporting-still-biased-after-all-these-years/.

3. **Becker JE**, Ross JS. Reporting Discrepancies Between the ClinicalTrials.gov Results Database and Peer-Reviewed Publications. *Ann Intern Med* 161(10): 18 Nov 2014, 760.

TABLE OF CONTENTS

I. INTRODUCTION

A. Evidence-Based Medicine: A History

Evidence-based medicine has become the approach of modern physicians in delivering top-notch patient care. Defined as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients,"[1] evidence-based medicine emphasizes combining clinical judgment from experience with individual patients with data from studies that illuminate disease pathology and the safety and efficacy of medical interventions. It calls for "weighing the pros and cons of different treatments fairly."[2] While the term "evidence-based medicine" was first introduced to the modern medical literature in 1991 by Gordon Guyatt of McMaster Medical School in Ontario, Canada,[3,4] this approach to patient care has been cited as far back as the 9th century AD and in 19th century Europe.[1,5] Yet, once Guyatt brought the term into the literature, the academic importance of evidence-based medicine exploded. By 2004, less than fifteen years after the term entered the literature, there were over 20 textbooks, nine academic journals, and several online tutorials dedicated to promoting evidence-based medicine.[3]

B. Evidence-Based Medicine: The Importance of Clinical Trials

In the era of evidence-based medicine, there are several methodologies of obtaining clinical data that will be used to inform clinician decision-makers. These include retrospective observational studies, case studies, and reviews of basic and pathophysiologic research. The paragon of evidence informing practice, however, is the randomized, controlled clinical trial.[1,6-8] Indeed, since 1962, the United States Food and Drug Administration (FDA) has regulated the approval of drugs on the basis of efficacy and safety data from "adequate and well-controlled" clinical trials.[9] Moreover, approximately 20,000 new clinical trials of drugs and medical devices are registered each year in the United States.[10] These clinical trials represent a large investment in both the public and private sphere. A 2006 *Lancet* study of 28 Phase III randomized clinical trials funded by the NIH, for instance, found the cost of each trial ranged from $722,000 to $64 million,[11] or about $12 million per trial.[12]

C. Clinical Trial Results Dissemination

Clinical decision-making in a world of evidence-based medicine is complicated and depends not only on physicians' access to the most up-to-date data on drug safety and effectiveness, but also on habit, personal experience, colleague experience, patient preference, cost concerns, and other factors.[13,14] It may be argued, however, that whereas physicians' personal experience with a medication or technology may play heavily into the decision to stop using it in practice, the adoption of a new medication or technology relies heavily on evidence and knowledge transfer, as the physician has no

personal experience with it.[15] In a treatise on the diffusion, dissemination, and implementation of medical research into clinical decision-making, Lomas describes that the steps of first making clinical trial data available, the subsequent analysis of the data through meta-analysis and consensus statements, and ultimately the behavioral change of physician decision-makers in their prescribing decisions is a process in which each step relies on the previous step.[14] Thus, though not sufficient to alter physicians' medical decisions, the availability – or diffusion – of clinical data is a necessary step to altering physicians' medical decisions.[14] Furthermore, because "research information…is the core building block around which further efforts will be built," Lomas argues that "extensive attention" must be paid to the "validity and reliability of the research information."[14] Thus, without disregarding the importance of factors like physicians' willingness to change, patient preference, and clinical context of physician decision-making,[13] it is the availability of safety and efficacy information on drugs and medical innovations – which come from clinical trials – that is key to physician decision-making and that forms the basis of evidence-based medical practice.

It usually takes the results of more than one trial to truly influence medical practice.[16] Rather than read through each new trial and analyze data individually, however, most physicians rely on guidelines, consensus statements from professional organizations, and meta-analyses to summarize and recommend new treatment options.[13,14,17] In fact, while it has been demonstrated that a single clinical trial may not change clinical decision-making,[18] recommendations put forth in consensus statements have been estimated to have a conformation rate among physician decision-makers

averaging around 50 to 60 percent.[18] Moreover, policymakers generally assume that guidelines put out by professional organizations will serve to spread new knowledge about the effectiveness and safety of treatments.[15] For instance, in the case of a new recommendation from the U.S. Preventive Services Task Force (USPSTF) against the use of prostate-specific antigen (PSA) as a screen for prostate cancer, Sen demonstrated that PSA provision and ordering by physicians dropped substantially after the recommendation release.[15]

In addition to demonstrating the effect of this institutional recommendation in altering physicians' behavior, Sen went on to demonstrate that physicians rely on one another for practice decisions as well, by showing that the use of PSA by physicians in the same group or practice as a given physician was significantly and positively associated with PSA use by that physician.[15] Nair *et al.* also find that the prescribing behavior of expert "opinion leaders" has a significant impact on other physicians' behavior within their networks – but only after new clinical guidelines are introduced.[19] These studies indicate the importance of clinical evidence and its review by professional organizations in driving not only individual physician decision-making, but also driving decision-making among physician peers, which has safety and cost implications for an even greater swath of patients. Given the widespread dissemination and influence of clinical guidelines and their effect on medical decision-making, the full data picture must be used by experts in order to develop the guidelines. While efforts have been made to ensure the high quality of evidence that goes into consensus statements and

guidelines,[20] it is equally important to ensure that complete evidence is available to guideline-makers and individual physicians alike.

D. Publication Bias

Throughout recent history, the main venue for accessing clinical trial results has been in peer-reviewed medical journals.[21,22] Ideally, for full efficacy and safety information to reach the prescribing clinician decision-makers and professional organizations creating clinical guidelines, all clinical trial results would be published completely and accurately in the medical literature.[8] Unfortunately, however, a growing body of literature has indicated this is rarely the case. Frequently, the studies published are only those with significant or positive results, a phenomenon termed "publication bias."[23-32]

The earliest modern evidence for this phenomenon comes from a 1959 paper by Sterling, who demonstrated evidence of an unexpectedly large amount of studies with statistically significant results in four psychology journals.[30] More recently, many studies have continued to show that positive studies are more likely to be published than negative studies.[7,23,25-28,30-33] One way of demonstrating publication bias is through analysis of meeting abstracts, which has shown that clinical trial abstracts with positive results are more likely to be published in full after the meeting.[2,34,35] Another insight into publication bias comes from looking at the length of time to publication. Among mostly industry-funded research, it has been repeatedly found that 25% to 50% of clinical trials are not published after several years. [7,22,31,33,36-40] Even among publicly-funded trials,

Ross *et al.* demonstrated that only 46% of a cross-section of NIH-funded clinical trials were published within 30 months of trial completion, with 68% published overall after a median of 51 months after study completion.[22] Moreover, multiple studies have shown that time from trial completion to manuscript submission and to publication are significantly faster among studies with positive results as compared to those with negative results.[7,28,32,39] Thus, a large share of the results from clinical trials of drugs and devices, and particularly negative results, are delayed or never make it into the medical literature, where they have the potential to impact healthcare decision-making.

Reasons for not publishing may include, but are not limited to, selective publication, bias against reporting or publication of negative results, limited resources, and a desire to publish "attention-grabbing" results – all both on the part of investigators and of journal editors.[2,7,22,36,37,41] Though journal editors play a role in encouraging the publication of positive results, some evidence has demonstrated that studies more often remain unpublished because they are never submitted, rather than because they are rejected by reviewers.[7,26] Selective publication, however, may also be attributed to censorship by study sponsors, particularly in the case of pharmaceutical trials that are funded by the drug or device manufacturer as compared to an outside funding source.[2]

E. Outcome Reporting Bias

Even among clinical trials that are published, however, bias can exist in the data that are available in the literature. Among such trials, the phenomenon of "outcome

reporting bias" – when only specific outcomes are included in the publication – has become increasingly recognized and problematic to a transparent medical evidence base.[7,42,43]

As far back as 1986, Robert Simes used results from a clinical trial registry and from the published literature to demonstrate differences in the suggested efficacy of two chemotherapeutic agents.[29] He looked specifically at trials registered in the International Cancer Research Data Base (ICRDB), which at the time contained most cancer-related clinical trials funded by the NIH. He reviewed the ICRDB registry, along with the published literature, to examine the use of combination chemotherapy and single agent therapy in two kinds of cancer. In each case, he performed a pooled analysis and demonstrated that clinical trial results in the literature showed significant benefits with the use of each of the combination chemotherapeutic regimens as compared to the single alkylating agent. Pooling the results from the trial registry, however, demonstrated no effect, or less of a positive effect, of the combination chemotherapy as compared to the single agent regimen. Thus, he demonstrated a selection bias in the publication of the outcomes previously registered in ICRDB that led to a rosier picture of combination chemotherapy accessible to clinicians through the medical literature than the data truly suggested.

Outcome reporting bias is not limited to the oncologic literature, however. While the CONSORT Statement, first created in 1996 and updated as recently as 2010, lays out guidelines for effectively and transparently publishing clinical trial data, selective outcome reporting has continued to be demonstrated widely among published clinical

trial reports.[8,44] In more recent studies comparing data submitted to the FDA to the published literature, important safety and efficacy information of drugs across a variety of medical fields have repeatedly found to be lacking from the published literature.[37,41,45,46]

The first systematic, comprehensive look at the issue of outcome reporting bias in the literature came in 2004 from Chan *et al.*, who compared trial protocols submitted for Ethics Committee approval between 1994 and 1995 to the same trials published as journal articles between 1995 and 2003.[42] The group found that a striking 71% of trials had at least one efficacy or safety outcome listed in the trial protocol that went unreported in the corresponding publication; additionally, 62% had "major discrepancies" between the primary endpoints laid out in the protocol and those described in the corresponding publication.[42] Such discrepancies suggest that investigators or journal editors are shaping the medical literature to promote particular outcomes, rather than allowing for the complete results to be analyzed in clinical decision-making. Indeed, in this study, Chan *et al.* also surveyed trial authors, who cited lack of statistical significance, journal space constraints, and low clinical importance as the reasons that outcomes were selectively published.[42] These reports were consistent with the study's finding that statistically significant outcomes were more than twice as likely to be fully published than those that were not statistically significant, again directly impacting the available medical evidence base to imply more positive drug effects than the full data may really indicate.[42]

Chan and Altman went on to perform an additional study with quite similar results, examining primary and subsequent publications of clinical trials to identify outcomes reported in the Methods section of at least one publication whose results were never published in a Results section.[47] Chan *et al.* additionally demonstrated similar findings among a sample of exclusively federally-funded trials in Canada, with 88% of trials having at least one outcome going unreported in the publication and 40% with discordant primary outcome specification between the trial protocol and publication.[48] This study indicated that, though financial incentives may play into the desire to report only certain efficacy and safety outcomes in the published literature,[16] selective outcome reporting is not limited to industry-funded trials.

In 2008, Turner *et al.* built on Chan's work, examining phase 2 and 3 trials of 12 antidepressant medications that received FDA approval between 1987 and 2004.[37] The group examined data submitted to the FDA for marketing approval of these 12 medications, using the pre-specified trial protocols to evaluate the selective publication of outcomes of those trials in the published literature.[37] Turner uncovered wide-ranging bias – including significantly higher rates of publication of trials with outcomes deemed "positive" by the FDA as compared to those deemed "negative," as well as the spinning of results deemed "negative" by the FDA as positive in the literature.[37] In fact, the team found that results found to be positive by the FDA were 12 times as likely to be published in agreement with the FDA analysis as compared to those deemed negative or neutral by the FDA.[37] The group went on to analyze how these discrepancies could pose a threat to the validity of the studies in the published literature and found that tactics to

distort the results included using higher patient populations in the published sources; increasing the effect size, with a median increase of 32%; and omitting or altering the primary outcome from that pre-specified in the FDA protocol when publishing those with an initially negative finding.[37] Thus, this analysis further uncovered the threat to the sanctity of science and the validity of the published literature when outcomes are specifically selected to represent a story that may not be comprehensive and may even be false.

Beyond these studies, others have found extensive evidence of outcome reporting bias across medical specialties and countries. Ewart *et al*. compared trials published in high-impact journals from late 2006 through early 2007 to entries in clinical trial registries and found that nearly one-third had a change in the primary endpoint, with the change most often being deletion.[49] In a sample of trials published in high-impact journals in the fields of cardiology, rheumatology, and gastroenterology and registered in a trial registry, largely of which was ClincalTrials.gov, Mathieu *et al.* demonstrated that the primary trial outcomes registered varied from those that were published in approximately 30% of trials; this number was consistent across general medical and specialty journals and was similar to that found by Ewart *et al*.[49,50] Fifteen percent of the trials in this study had a primary outcome listed in the protocol that was omitted from the publication.[50] While only half of the discrepancies could be assessed for clinical impact, of those that could, over 80% favored the reporting of a statistically significant outcome in the publication, thus indicating the same bias that has been demonstrated elsewhere.[50] Dwan *et al.* found a higher rate of primary outcome

inconsistency, 47 to 74 percent, in their review of studies examining outcome reporting bias that compared trial protocols to publications.[7] Two additional studies have also demonstrated selective outcome reporting by showing significantly more complete reporting of efficacy and safety results in trial registry entries as compared to publications in a German sample and in a sample with registry entries on ClinicalTrials.gov, respectively.[51,52] Finally, Vedula *et al.* were able to go a step further beyond trial protocols or registry reports to access internal documents from Pfizer and Parke-Davis for trials of off-label use of gabapentin; they found that 8 of the 12 trials that were published had a discordant primary outcome between the internal document and the publication, thus demonstrating an even deeper-rooted issue in selective outcome reporting.[53]

These studies highlight that, among a range of sources and medical specialties, outcome reporting bias exists in the accessible medical literature and often exaggerates the efficacy or safety of drugs. The lack of reporting, or altered reporting, of primary endpoints in the published literature as compared to pre-specified protocols and registry entries is particularly troubling, as trials are designed with statistical power to address these outcomes as initially designated. Thus, the results of an altered primary outcome when published may not be as statistically reliable. Additionally, patients sign up for, and funders invest in, trials with particular clinical aims. Altering the primary endpoint for publication can prevent advancement and true understanding in clinical science to allow instead for the perceived advancement of academic or financial interests when publishing more positive results.

E. The Problem with Bias

Despite the mounting evidence that negative results are not published, at both the study- and outcome-level, physicians and the institutions crafting clinical guidelines need access to both positive and negative results in order to make fully informed, evidence-based clinical decisions and recommendations.[54] In fact, negative results are instrumental in helping physicians to evaluate the medications they currently have at their disposal and to "invalidat[e] previously accepted" medication usage.[1,31] Additionally, literature reviews and meta-analyses depend on fully informed data.[2,36,41] If trials or outcomes are selectively published, the numbers of patients can be difficult to glean when using the published literature for meta-analysis or literature review.[55] However, the complete story – positive and negative results – can only be revealed in the setting of publication bias if clinical trial data can be freely accessed by clinicians through mechanisms other than the published literature.

In addition to improving the quality and efficiency of evidence-based clinical decision-making, other factors necessitate the public dissemination of clinical trial data. Perhaps the most important of such factors is the ethical consideration to trial participants.[2,16,31,54,56] When patients participating in clinical trials give their time – and their bodies – to research investigators, they generally do so with the assumption that their donation of time and self will help to improve medical care for others. If the data obtained from these patients are not fully or accurately conveyed to clinical decision makers, the patient's donation is for naught. Further, public dissemination of trial

results may also improve the trust of the general public in medical science as an institution.[56]

On top of clinical and ethical concerns, unpublished clinical trial results represent a sunken investment. The cost of an individual phase III clinical trial has been estimated in the hundreds of millions of dollars.[57] The National Institutes of Health (NIH) spends over $3.5 billion of public funds on investments in clinical trials.[22] Thus, there is an additional obligation from investigators funded by public sources to disseminate their trial results in a freely-accessible, public setting.[2] Whether publicly- or privately-funded, unpublished research can also lead to the unbeknownst duplication of clinical trials, further increasing the cost burden to the health care system.[58]

F. Policies to Address Bias: Clinical Trial Registration and Results Reporting

To tackle the issues of selective publication and publication bias, evidence-based medicine experts have long supported the creation of broad-ranging public clinical trial registries, with the goal of prospectively registering basic information about trials at trial inception.[2,29,31,55,56,58,59] As such, trial investigators and sponsors can be more readily held accountable for not disseminating trial information and selectively reporting results. While clinical trial registries have existed in one form or another since at least the mid-1970s, the initial registries were generally specialty-specific, incomplete, and difficult to access.[60,61] The advent of the Internet age, however, has allowed for the creation of online clinical trial registries, with the potential to disseminate all clinical trial results in a complete and easily-accessible manner.[60]

In 1997, the United States Congress passed the FDA Modernization Act (FDAMA), mandating the Department of Health and Human Services, acting through the NIH, to develop a public, Internet-based clinical trial registry, or "data bank[,]…of information on clinical trials for drugs for serious or life- threatening diseases and conditions" submitted to the FDA as part of Investigational New Drug (IND) applications.[62,63] This law compelled the U.S. Department of Health and Human Services to create a public registry of ongoing, federally- and privately-funded clinical trials of drugs in the U.S. The law was intended to provide a publicly-available listing of ongoing trials for patients who wished to enroll, particularly those with rare illnesses who otherwise may not have knowledge or access to groundbreaking treatments for their conditions. In addition to providing information for patients, however, the registry was a first step in holding trial investigators and sponsors accountable for providing public access to trial information, an early step to improve transparency in the publicly-accessibly medical evidence base. The initial requirements for the data bank, as laid out by the law, included:

> ***"(A) A registry of clinical trials (whether federally or privately funded) of experimental treatments for serious or life-threatening diseases and conditions… Information provided shall consist of eligibility criteria for participation in the clinical trials, a description of the location of trial sites, and a point of contact for those wanting to enroll in the trial, and shall be in a form that can be readily understood by members of the public… (B) Information pertaining to experimental treatments for serious or life-threatening diseases and conditions that may be available-- (i) under a treatment investigational new drug application that has been submitted to the Secretary under section 561(c) of the Federal Food, Drug, and Cosmetic Act; or (ii) as a Group C cancer drug (as defined by the National Cancer Institute). The data bank may also include information pertaining to the results of clinical trials of such treatments… including information concerning potential toxicities or***

> ***adverse effects associated with the use or administration of such experimental treatments."***[63]

Thus, on top of requiring basic information about each trial, the law suggests but did not explicitly mandate the reporting of trial results and adverse events. Notably, device trials were excluded from these mandates, with the explicit statement in the law that including devices would be considered and possibly added.[63]

As a result of these requirements set out in FDAMA, the U.S. National Library of Medicine launched a public, Internet-based clinical trial registry, ClinicalTrials.gov, in 2000. As laid out in the law, the registry was created on behalf of the National Institutes of Health, under the Department of Health and Human Services. In addition to IND-related trials, the new ClinicalTrials.gov database allowed for the registration of any clinical trial, regardless of intervention kind, medical condition, or country of investigation; and within five years, the ClinicalTrials.gov database contained over 20,000 trial entries.[62] While other trial registries had been in place and subsequently developed, including the WHO's International Clinical Trial Registry Platform, ClinicalTrials.gov has remained the largest and most prominent such registry in the United States.[54,56,60,61] While ClinicalTrials.gov did not initially require trial results to be registered, the suggestion of including results laid out in FDAMA set the groundwork for a potential publicly-available register of clinical trial results, which would later be developed to improve access to trial data for clinical decision-makers and guideline-developing institutions.

In 2005, the International Committee of Medical Journal Editors (ICMJE), a group of general medical journal editors who work to set recommendations for the conduct and publication of medical research,[64] initiated a policy that all trials being considered for publication in one of its member journals must be registered in a public clinical trial registry at or before patient enrollment.[16] This move came explicitly in an attempt from journal editors to start addressing the issues of publication and outcome reporting bias and to improve clinical trial data transparency, with the policy announcement noting the "ideal" that eventually "if all trials are registered in a public repository at their inception, every trial's existence is part of the public record, and the many stakeholders in clinical research can explore the full range of clinical evidence."[16] The ICMJE defined clinical trial at the time as "any research project that prospectively assigns human subjects to intervention or comparison groups to study the cause-and-effect relationship between a medical intervention and a health outcome."[16] At the time, ClinicalTrials.gov was, to the knowledge of the ICMJE, the only established eligible registry fitting all requirements laid out in their new policy, such as being freely accessible, open to all registrants, run by a non-profit, and electronically searchable.[16] Despite concerns from trial investigators that this requirement would be taxing and eliminate competition,[16] the number of clinical trials registered in ClinicalTrials.gov jumped over 70% from 13,153 in the months preceding the announcement to 22,714 trials just one month after the policy went into effect in September 2005.[56,62]

In 2007, the ICMJE reevaluated this policy. At this time, the ICMJE broadened the definition of clinical trial to that being used by the World Health Organization (WHO) –

"any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes"[56] – in order to include preliminary trials in its mandate and thus in its goal of improving trial transparency in the world of evidence-based medicine. The group also endorsed the WHO ICTRP and its member registers as acceptable venues for clinical trial registration.[56] Recognizing a growing advocacy for reporting of clinical trial results to the databases in which registration was required, the ICMJE specifically laid out in their 2007 announcement that "results posted in the same clinical trials registry in which the primary registration resides…[would not be considered] to be previous publication if the results [we]re presented in the form of a brief (<500 words) structured abstract or table."[56] This move encouraged open access to trial results via trial registries while reassuring investigators who might participate that their work would still be competitive for publication.

In September of the same year, the U.S. Congress followed suit in utilizing the developed ClinicalTrials.gov registry to create a database of publicly-accessible clinical trial results, thus expanding the evidence base for clinical decision-making beyond publications, by passing the FDA Amendments Act (FDAAA), Section 801.[65] FDAAA updated the clinical trial registry requirements that had been laid out under the 1997 FDAMA law. Specifically, FDAAA expanded the requirements to mandate registration by trial sponsor or primary investigator at trial initiation for trials not only of drugs, as in FDAMA, but also of all biologic agents and medical devices regulated by the FDA and of pediatric postmarket surveillance studies required by the FDA.[54,63,65,66] Notably, unlike

the new ICMJE requirement, FDAAA continues to exclude phase 1 clinical trials and "small feasibility studies" from reporting requirements.[54,65,66] The law specifies that FDA regulation applies to interventional trials with at least one arm; with one or more U.S. sites; studying a drug, biologic, or device manufactured in the U.S. or its territories; or that are conducted under an investigational new drug application or investigational device exemption.[65,66]

More important, FDAAA mandates the reporting of the results of such trials ongoing on or after September 27, 2007. Specifically, trial results must be uploaded to ClinicalTrials.gov within 12 months of study completion, defined as completion of collection of the primary endpoint on the final study subject.[54,65] The specific data elements that are required by this mandate include reporting of: the study design, start date, target sample size, and the trial's primary and secondary outcomes; demographic and baseline characteristics of the sample, including enrollment and completion numbers; results of the primary and secondary outcomes, including results of tests of significance; and serious and frequent adverse events (>5% in any arm of the trial).[54,65] The reporting requirements apply to the trial's main results, not those of subgroup analyses. The law also established penalties for not complying with the reporting of trial information and results, including loss of NIH grant funding and fines of up to $10,000 per day.[54,65] Results of trials of FDA-approved drugs being studied for a new, unapproved indication have an additional year to report results; there is no requirement to post results for drugs that never received FDA approval or for drugs that received FDA approval before September 27, 2007 that no longer were the subject of clinical trials – a

large majority of prescription drugs currently in clinical practice.[54] Nonetheless, through this new regulation, the U.S. government took a firm stance in supporting evidence-based medicine by promoting clinical trial data transparency going forward. The law also attempted to increase data clarity by laying out requirements for the Department of Health and Human Services to link ClinicalTrials.gov entries to relevant results information already available on the Internet through the FDA, National Library of Medicine (NLM), or National Institutes of Health (NIH), such as FDA public health warnings, FDA drug approval packages, or peer-reviewed articles linked through Medline.[54,65]

G. Initial Evaluation of ClinicalTrials.gov

The mandated expansion of ClinicalTrials.gov under FDAAA to require results reporting for studies investigating medical products created a pool of clinical trial findings available to inform clinical practice and research. However, in a careful evaluation of the new legislation, Alastair Wood points out that "because the submitted data will not be peer-reviewed…interpreted, qualified, or explained, the results reported at ClinicalTrials.gov will complement rather than replace the thoughtful presentation and discussion of results characteristic of the best peer-reviewed publications."[54] Nonetheless, given the broad evidence base suggesting publication and outcome reporting bias in peer-reviewed publications, ClinicalTrials.gov may very well serve as a cache of data that may never make it into the peer-reviewed, published literature. As such, analysis of the completeness and accuracy of the contents of the ClinicalTrials.gov

registry after the implementation of FDAAA is important to ensure it can be a valuable resource for clinicians and researchers.

Initial studies examining the contents of ClinicalTrials.gov after FDAAA have indicated that despite the mandate and the purported penalties for non-compliance, fewer than one-quarter of trials required to report results to ClinicalTrials.gov had actually done so within a year of trial completion prior to 2011.[67,68] However, in subsequent years, the number of registered trials reporting results has increased.[69] Previous studies have examined trial information available on ClinicalTrials.gov and found fairly complete reporting of mandatory data elements, such as trial funder and intervention, but high rates of missing data for optional elements and uncertainty about information accuracy.[41,50,62] Because results reporting began relatively recently under FDAAA, no studies have yet thoroughly examined the completeness and accuracy of this information on ClinicalTrials.gov.

II. Statement of Purpose

The purpose of this project is to evaluate the trial data and results reported to ClinicalTrials.gov after the passing of the FDA Amendments Act and to examine how these data on ClinicalTrials.gov compare to data published in the high-impact medical literature for the same trials. While the results reported in either ClinicalTrials.gov or published articles have the potential to be incomplete and inaccurate, inconsistencies between these two sources of trial results will offer insights into whether and how results reported for all trials, published or not, can be used to inform clinical practice and future research efforts.

III. Specific Hypothesis

The results listed on ClinicalTrials.gov will be fairly complete for required data elements. The clinical trial endpoints and results defined and reported on ClinicalTrials.gov will vary from those in the published literature, in statistically significant ways, with a bias towards positive results in the literature.

IV. Specific Aims

Primary Aims:

1) To examine the completeness of trial results data reported on ClinicalTrials.gov as required by FDAAA within 12 months of trial completion.
    a. Specifically, to assess completeness of the following required data elements: cohort characteristics; intervention; primary and secondary endpoints; results of all primary and secondary endpoints; all serious adverse events; and all other adverse events with a frequency ≥ 5% in at least one study arm.
2) To better understand the accuracy of these trial data elements reported on ClinicalTrials.gov, as compared to corresponding trial data elements published in high-impact biomedical journals.

Secondary Aim:

1) To use secondary endpoint data reporting on ClinicalTrials.gov and in the published literature to evaluate the existence of positive outcome reporting bias in the published literature.
    a. Specifically, to evaluate the significance of secondary endpoints listed on ClinicalTrials.gov and to correlate secondary endpoint significance as listed on ClinicalTrials.gov with publication of the endpoint in the trial's corresponding journal article.

V. METHODS

A. Study Sample

Using information obtained from the NLM, members of the team (Dr. Ross and Gal Ben-Josef) identified all articles published in a Medline-indexed journal between July 1, 2010 and June 30, 2011 that were linked to a ClinicalTrials.gov identification number (n=4,586). These dates were chosen because we wanted to examine a one-year period, data collection was set to begin in January 2012, and we wanted to allow for a lag period after which investigators published their study to report results to ClinicalTrials.gov. The information obtained from the NLM included all reported data elements available on ClinicalTrials.gov, such as lead funder, study design, and condition studied. From this sample of published clinical trials, Dr. Ross and Ms. Ben-Josef identified a sub-sample of articles published in journals with an impact factor ≥ 10 (n=831), determined using Web of Knowledge (Thomson Reuters; New York, NY). Subsequently, we restricted our study to articles describing clinical trials for which results had been reported to ClinicalTrials.gov as of January 2012 (n=149). Finally, because FDAAA only requires reporting of main trial results, we excluded 53 articles that did not report main trial results or reported results of multiple trials (**Figure 1**).

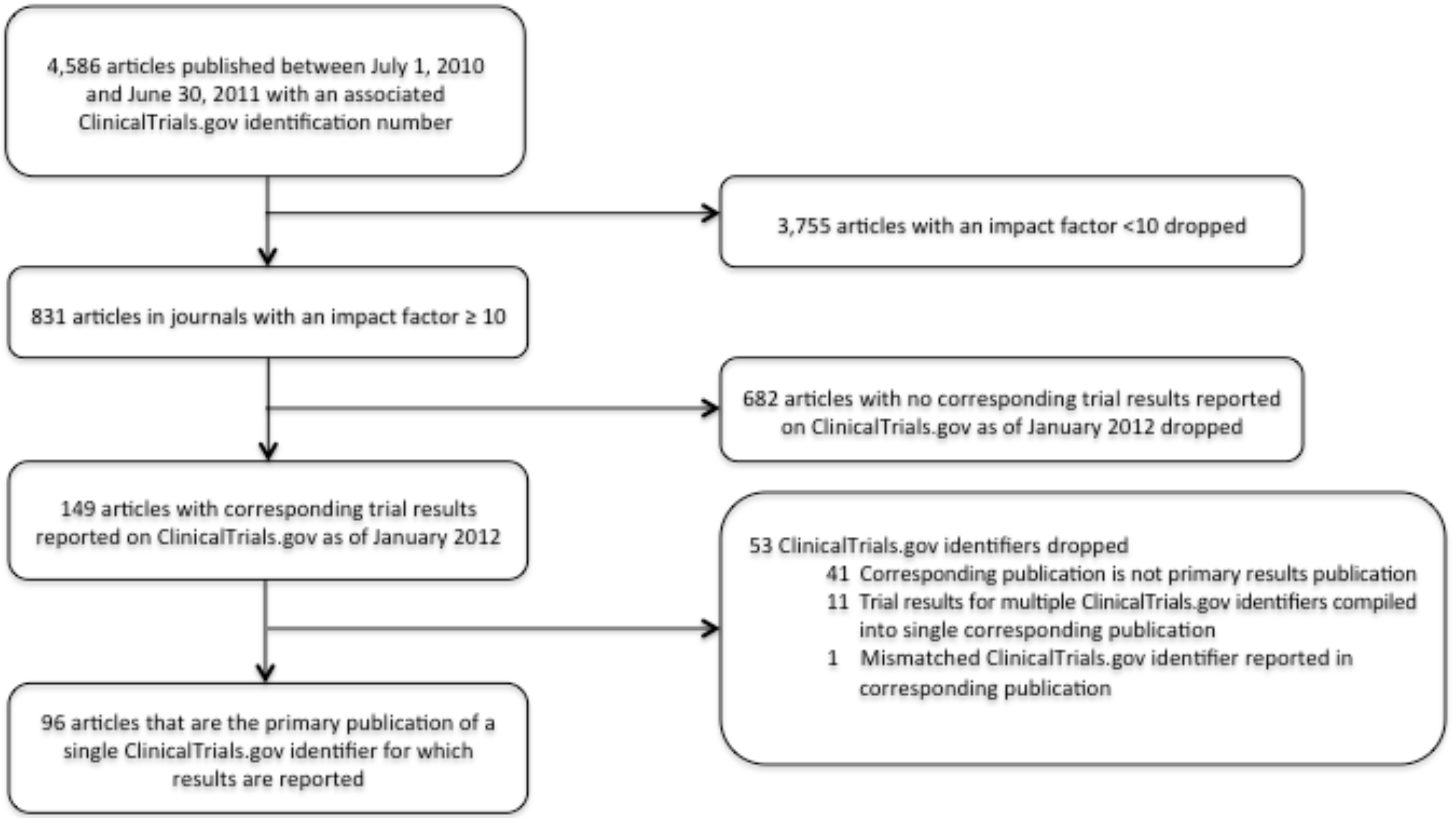

Figure 1: Flow diagram showing sample construction of clinical trials registered in and reporting main results on ClinicalTrials.gov that were published in a Medline-indexed journal with an impact factor ≥ 10 between July 1, 2010 and June 30, 2011.

B. Collecting Reported Information and Values

All data collection described below was performed by me (Jessica Becker). Gal Ben-Josef assisted in data collection from a fraction of the trials. A selection of data was also reviewed by Dr. Ross.

i. Cohort Characteristics and Trial Intervention

For each clinical trial, we collected information from both ClinicalTrials.gov and the corresponding publication on cohort characteristics, including enrollment numbers, completion numbers, and age and sex distributions, and the trial intervention. We considered study intervention information incomplete unless the definition (i.e. name), duration, frequency, and dosage of the intervention were each described.

ii. Primary Efficacy and Secondary Efficacy Endpoints

For each clinical trial, we collected the number of primary and secondary efficacy endpoints reported on ClinicalTrials.gov and in the publication along with the definition of each endpoint, including time point(s) at which the endpoint was measured and scale defined to measure the endpoint's results (e.g. a 54-point Attention Deficit Hyperactivity Disorder rating scale or the defined threshold value for the endpoint). If an outcome was reported at multiple time points, we considered each ascertainment point separately. Since FDAAA mandates that only endpoints evaluated for the entire study population be reported on ClinicalTrials.gov, we did not collect endpoints that were

defined only for a subgroup population or extension study endpoints. Next, for each endpoint, we collected results values reported on ClinicalTrials.gov and in the corresponding publication and noted when values were not reported in one source or the other.

iii. Adverse Events

For each trial, we collected the type and frequency of all serious adverse events affecting one or more trial participants and all other adverse events with at least 5% frequency in one trial arm from ClinicalTrials.gov and from the corresponding publication, in concordance with ClinicalTrials.gov definitions.[70]

C. Comparing Reported Information and Values

All data collection and comparisons described below were performed by me (Jessica Becker). Gal Ben-Josef assisted in data collection and comparison for a fraction of the trials. A selection of data, including all results discordant between sources, was also reviewed by Dr. Ross.

i. Cohort Characteristics and Trial Intervention

For each trial, we compared the reported information for cohort characteristics and intervention between ClinicalTrials.gov and corresponding publications, determining whether information was concordant, or in agreement. Each cohort

characteristic was considered concordant only if the values of the characteristic (e.g. number of participants enrolled or mean age) were numerically equal across sources for all trial arms. We could not compare cohort characteristics when sources differed in statistical analysis method, for instance, if one source reported medians and another means. We considered reported study interventions concordant if there was agreement in the definition, duration, frequency, and dosage of the intervention between both sources.

ii. Primary Efficacy and Secondary Efficacy Endpoints

For each trial, we compared the reported information for primary and secondary efficacy endpoints between ClinicalTrials.gov and corresponding publications. We considered endpoint definitions concordant if there was agreement in the described endpoint, time of endpoint ascertainment, and endpoint measurement scale between both sources.

For primary and secondary efficacy endpoints that were reported in both sources and defined concordantly, we determined whether results values were concordant (i.e., numerically equal), discordant (i.e., not numerically equal), or could not be verified. We could not verify results values when reporting was numerical in one source but graphical in the other, sources differed in statistical analysis method, or results were stratified differently between the sources.

For those primary efficacy endpoints with discordant results values, the medical student author of this thesis (myself) and Dr. Ross evaluated the results together to

determine whether the publication's values would lead to a different interpretation than the results reported on ClinicalTrials.gov would. We accomplished this by examining the magnitude of the difference, directionality of the results, and statistical significance testing from each source, when available. We did not examine the impact on study interpretation for secondary efficacy endpoint results.

iii. Adverse Events

For each trial, we compared the reported information for adverse events between ClinicalTrials.gov and corresponding publications, determining whether results were concordant, discordant, or could not be verified. We considered serious and other adverse event reporting concordant only if all serious events in one or more subjects or all other events with a frequency ≥ 5% reported within one source were also reported in the other, at equal numerical frequency for each event in all trial arms. We could not verify adverse event results values when sources used different adverse event reporting scales (e.g. serious and other categorization required by ClinicalTrials.gov versus the National Cancer Institute Common Terminology Criteria for Adverse Events [NCI CTCAE] 5-point graded scale in a publication) or if the publication combined serious and other adverse events into one category.

D. Statistical Analysis

All of the descriptive and statistical analyses described below, including the exploratory analysis of secondary endpoints, were performed by me (Jessica Becker), with assistance as needed and complete review both provided by Dr. Joseph Ross.

For sample description purposes, we collapsed the reported data elements available on ClinicalTrials.gov into clinically meaningful categories (**Table 1**). We then conducted a descriptive analysis, examining completeness of reported information on ClinicalTrials.gov and concordance between ClinicalTrials.gov and corresponding publications for trial cohort characteristics; intervention; primary and secondary endpoints; primary and secondary endpoint results; all serious adverse events; and all other adverse events with a frequency ≥ 5% in at least one study arm. For discordant primary endpoint results, we also characterized the frequency with which trial interpretation differed between the two sources.

As an exploratory analysis of secondary endpoints, to determine whether favorable secondary endpoints were more likely to be reported in corresponding publications, we classified the statistical significance of all secondary endpoints reported on ClinicalTrials.gov and examined the likelihood of reporting statistically significant endpoints in the corresponding publications using Chi-Square tests. In many cases, statistical testing was reported within ClinicalTrials.gov. However, when not reported, we performed statistical testing using reported results values, using QuickCalcs *t test calculator* (v3.5.4, GraphPad; La Jolla, CA) for comparisons of continuous endpoints and Epi Info™ (v7.1.1.0, Centers for Disease Control; Atlanta, GA) for comparisons of

categorical endpoints. A p-value ≤ 0.05 was used as a threshold for statistical significance for superiority studies, > 0.05 for non-inferiority studies.

Descriptive analyses were performed using Excel® (v14.3.1, Microsoft Corporation; Redmond, WA) and Chi-Square tests were performed using Epi Info™.

VI. RESULTS

A. Study Sample

We identified 96 trials reporting main results on ClinicalTrials.gov that were published in a Medline-indexed journal with an impact factor ≥ 10 between July 1, 2010 and June 30, 2011. One ClinicalTrials.gov identifier registered a trial conducted in two continents, with primary results for each continent published separately; for our analysis, these articles were considered to be two separate trials. Industry was the lead funder of nearly three-quarters of the trials (n=70; 73%). The most common conditions studied were cardiovascular disease, hyperlipidemia, or diabetes (n=21; 22%), cancer (n=20; 21%), and infectious diseases (n=19; 20%). Median trial enrollment size was 509 subjects (Inter-Quartile Range [IQR], 216–1281) and trials were most frequently published by *New England Journal of Medicine* (n=23; 24%), *Lancet* (n=18; 19%), and *Journal of the American Medical Association* (n=11; 12%). **Table 1** below describes the sample demographics in detail.

Table 1: Characteristics of clinical trials registered on ClinicalTrials.gov with primary results published in a biomedical journal with impact factor ≥ 10 between July 1, 2010 and June 30, 2011.

Trial Characteristics	
Trial Sponsor, No. (%)	
Industry	70 (73)
NIH or other government agency	9 (9)
Non-profit organization or other funding agency	17 (18)
Phase 3 & 4 trials, No. (%)	68 (71)
Conditions examined, No. (%)	
Diabetes, hyperlipidemia, or cardiovascular disease	21 (22)
Cancer	20 (21)
Infectious diseases	19 (20)
Neurologic	9 (9)
Hematologic	7 (7)
Other conditions	20 (21)
Trial enrollment size, median (range)	509 (216-1281)
Corresponding publication journal, No. (%)	
New England Journal of Medicine	23 (24)
The Lancet	18 (19)
Journal of the American Medical Association	11 (12)
Other journal	44 (46)

B. Data Completion

All trials reported cohort enrollment, age, and sex distributions in both sources. Trial completion rate was explicitly described in both sources, or could be calculated from number enrolled and withdrawn, for 90 trials (94%). All trials provided some intervention information in both sources; 16 (17%) were missing data on dosage amount or timing in at least one source. For primary outcomes, 91 trials (95%) described at least one primary efficacy outcome in both sources, whereas 5 (5%) had no primary efficacy outcomes in either source, instead defining only primary safety outcomes in both sources. Similarly, 94 publications (98%) and 89 ClinicalTrials.gov entries (93%) described at least one secondary efficacy outcome, whereas 95 trials (99%) reported adverse events in both sources.

C. Overall Registry-Article Comparison

Overall, 95 of 96 trials were found to have at least one discrepancy between results reported to ClinicalTrials.gov and corresponding published articles: 29 (30%) discordant trial cohort descriptions; 15 (16%) discordant intervention definitions; 27 (28%) discordant primary efficacy outcome definitions or results values; 91 (95%) discordant secondary efficacy outcome definitions or results values; and 50 (52%) discordant adverse events reporting. **Table 2** below summarizes the completion and comparison findings, while the sections below enumerate the comparison findings in detail.

Table 2: Reporting and comparison of results information on ClinicalTrials.gov and in publications among trials published in a biomedical journal with impact factor ≥ 10 between July 1, 2010 and June 30, 2011 that were registered and reported results on ClinicalTrials.gov (n=96).

Results Information	**Trials Reporting, No. (%)**		**Comparison of Reported Information among Trials Reporting in Both Sources, No. (%)**		
	ClinicalTrials.gov	**Publication**	**Concordant**	**Discordant**	**Could Not Be Compared**
Cohort Characteristics					
Enrollment No.	96 (100)	96 (100)	94 (98)	2 (2)	0 (0)
Completion Rate	90 (94)	90 (94)	70 (78)	20 (22)	0 (0)
Sample Age Distribution	96 (100)	96 (100)	56 (58)	6 (6)	34 (35)
Sample Sex Distribution	96 (100)	96 (100)	85 (89)	9 (9)	2 (2)
Trial Intervention	96 (100)	96 (100)	65 (68)	15 (16)	16 (17)
Efficacy Endpoints					
Primary*	91 (95)	91 (95)	81 (61)	21 (16)	30 (23)
Secondary†	89 (93)	94 (98)	338 (55)	53 (9)	228 (37)

* For primary efficacy endpoints, same 91 trials defined a total of 156 primary efficacy endpoints in either ClinicalTrials.gov or the corresponding publication, 132 (85%) of which were described in both sources.

† For secondary efficacy endpoints, 96 trials defined a total of 2089 secondary efficacy endpoints in either ClinicalTrials.gov or the corresponding publication, 619 (30%) of which were described in both sources.

D. Cohort and Intervention Registry-Article Comparison

Enrollment size could be compared for all 96 trials and was discordant for 2 (2%), varying by 1 to 2 individuals per trial arm. We were able to compare trial completion rates for 90 trials with complete information; of these, trial completion was discordant for 20 (22%), with a median difference of 10.5 percentage points (IQR: 2.3-31.1). We could not compare sex distributions for 2 trials (2%); among the remaining 94, sex distribution was discordant for 9 (10%), with a median difference of 0.70 percentage points (IQR: 0.17-4.7). We could not compare age distributions for 34 trials (35%); among the remaining 62, age distributions were discordant for 6 (10%), varying by a median of 0.60 years (IQR: 0.13-1.3). Among 80 trials with complete trial intervention descriptions in both sources, 15 descriptions (19%) were discordant, most often because of different dosages, different frequencies or duration of intervention, or description of an additional intervention or placebo. See **Appendix Table A1** for complete enumeration of intervention discordances.

E. Primary Outcome Registry-Article Comparison

Among 91 trials defining primary efficacy endpoints, there were 156 endpoints designated within either ClinicalTrials.gov or published articles, 132 (85%) of which were

described in both sources, 14 (9%) of which were described only on ClinicalTrials.gov, and 10 (6%) of which were described only in articles. The median number of primary efficacy endpoints was 1 (IQR 1.0-1.0) per ClinicalTrials.gov entry and 1 (IQR 1.0-1.3) per publication.

Thirty (23%) of 132 concordantly defined endpoints reported results values that could not be compared, often due to graphical versus numerical reporting. Among 102 endpoints that could be compared, results values for 21 (21%) were discordant. Therefore, in total, only 81 of 156 endpoints (52%) designated as primary efficacy outcomes within either ClinicalTrials.gov or the published article were described in both sources and reported verifiable and concordant results. **Figure 2** below summarizes the findings for primary endpoint reporting.

Among 21 trials with discordant primary efficacy outcome results values, discrepancies led to differences in trial interpretation between sources for 6 (29%; 7% of trials with primary efficacy endpoints). An example was reporting of statistically different primary outcomes between trial arms on ClinicalTrials.gov but not significantly different in the publication.[71,72] **Table 3** below highlights the primary endpoint results discordances that altered trial interpretation.

For the remaining 15 trials (71%) with discordant primary outcome results values, discrepancies did not lead to differences in trial interpretation. For instance, in one case, primary endpoint results reported in the paper for the treatment group were reported as results for the placebo group on ClinicalTrials.gov, and vice versa;[73,74] however, since the results were not statistically different, trial interpretation was

unchanged. In another instance, the discordant primary outcome results were close in absolute value, in the same relative directionality between trial arms, and statistically significant between trial arms in both sources.[75,76] For a complete enumeration of the primary endpoint results discordances that did and did not alter trial interpretation, please refer to **Appendix Table A2**.

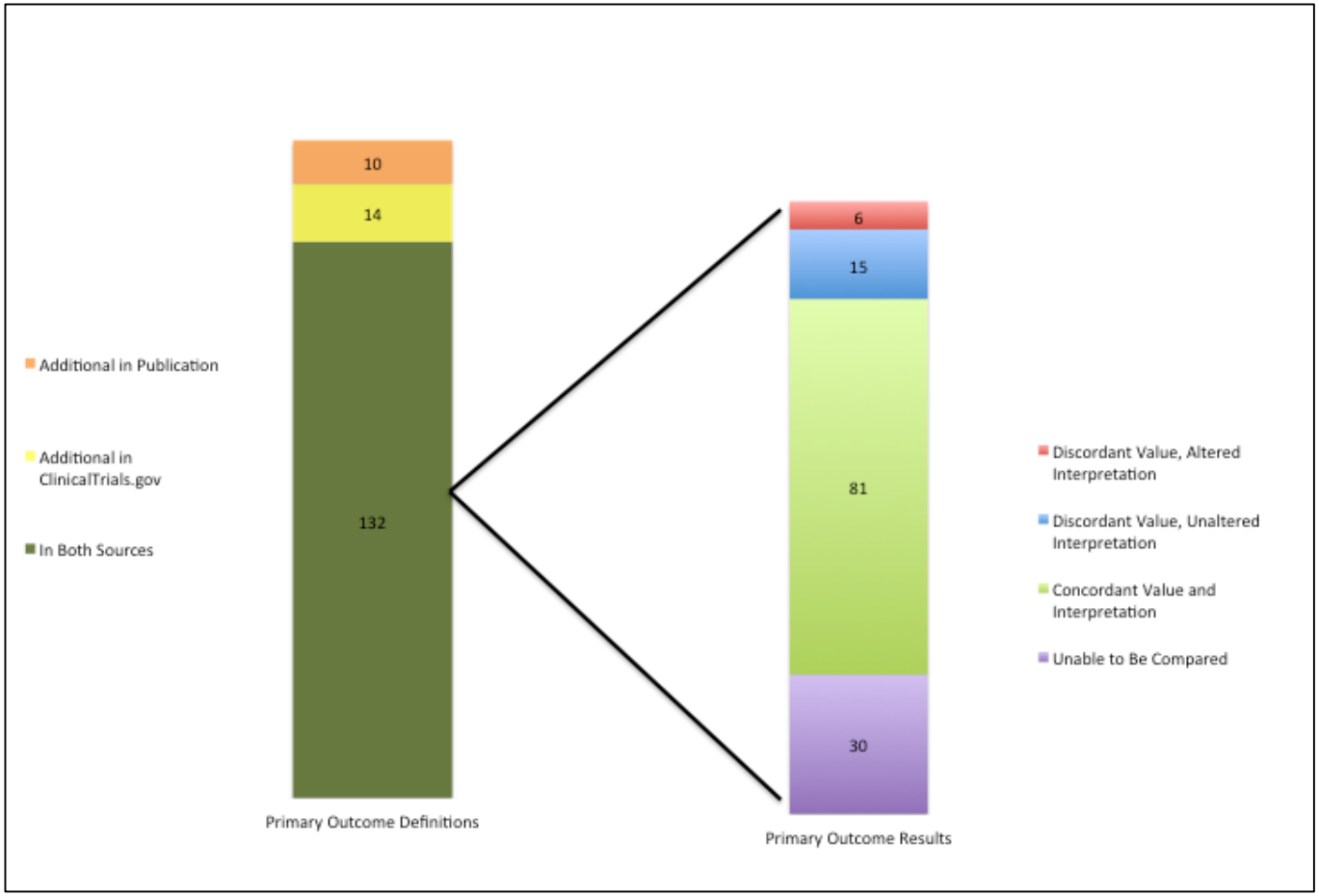

Figure 2: Primary efficacy outcome endpoint definitions and results values reported on ClinicalTrials.gov and in corresponding journal articles for clinical trials registered in and reporting main results on ClinicalTrials.gov that were published in a Medline-indexed journal with an impact factor ≥ 10 between July 1, 2010 and June 30, 2011 (n=91).

Table 3: Discordant primary efficacy endpoint results reported on ClinicalTrials.gov and in corresponding publication that altered trial interpretation (n=6).

Trial ID	Primary Efficacy Outcome		Explanation of Altered Trial Interpretation
	ClinicalTrials.gov Reported Results	**Publication Reported Results**	
NCT00094887	**Median Hours to Resolution of Vaso-occlusive Pain Crisis (95% CI)**		Time to resolution in both groups is substantially lower on ClinicalTrials.gov than in the article, altering clinical interpretation.
	Inhaled Nitric Oxide: 61.83, 95% CI: (41.75, 78.00); Placebo: 55.16, 95% CI: (46.00, 72.00); No statistical analysis provided.	Inhaled Nitric Oxide: 73.0, 95% CI: (46.0, 91.0); Placebo: 65.5, 95% CI: (48.1, 84.0); P = 0.87	
NCT00108953	**Median Time to Progression (95% CI)**		Median time to progression in both groups is substantially higher on ClinicalTrials.gov, altering clinical interpretation.
	Sorafenib + Doxorubicin: 263 days, 95% CI: (146, 384); Placebo + Doxorubicin: 147 days, 95% CI: (66, 244); P = 0.016	Sorafenib + Doxorubicin: 6.4 months, 95% CI: (4.8, 9.2); Placebo + Doxorubicin: 2.8 months, 95% CI: (1.6, 5); P = 0.02	
NCT00177671	**Number of Participants With Recurrence of Major Depression**		Percentage of participants with major depression recurrence is lower on ClinicalTrials.gov and hazard ratio on ClinicalTrials.gov is 2-fold greater.
	Donepezil: 19/67, 95% CI: (16, 31); Placebo: 11/63, 95% CI: (6, 18); HR=3.97, SD=2.09, 95% CI: (1.00, 4.41); P=0.05	Donepezil: 35%; 95% CI: (24%, 46%); Placebo: 19%, 95% CI: (9%, 29%); HR=2.09, 95% CI: (1.00, 4.41), λ^2=3.97; P=0.05	

NCT00281918	**Progression-free Survival (PFS), median**		PFS is substantially lower in the rituximab arm reported on ClinicalTrials.gov than in the article, altering clinical interpretation.
	Fludarabine/Cyclophosphamide: 981.0 days, Range: (1, 1343); Fludarabine/ Cyclophosphamide/Rituximab: 1212.0 days, Range: (1, 1372); P<0.0001	Fludarabine/Cyclophosphamide: 32.8 months, 95% CI: (29.6, 36.0); Fludarabine/ Cyclophosphamide/ Rituximab: 51.8 months, 95% CI: (46.2, 57.6); P<0.0001	
NCT00404079	**Roland Morris Disability Questionnaire, 1 year**		ClinicalTrials.gov score is higher for both trial arms and statistical testing results are different in the two sources, leading to a difference in trial interpretation.
	Glucosamine Sulphate: 9, SD: 4; Placebo: 9, SD: 4; Odds Ratio: 4.5 ± 4; P=0.05	Glucosamine Sulphate: 4.8, 95% CI: (3.9, 5.6); Placebo: 5.5, 95% CI: (4.7, 6.4); P=0.50	
NCT00426751	**Number of Participants With Complete Sum ST Resolution 60 Min After Percutaneous Coronary Intervention (Intent-to-Treat Population)**		Confidence interval of the adjusted difference between arms crosses zero on ClinicalTrials.gov and does not in the article, suggesting a difference in statistical testing of results between the two sources, leading to a difference in trial
	Eptifibatide: 124/214; Abciximab: 103/196; Adjusted Difference: 6.8%, 95% CI: (-3.0%, 16.6%)	Eptifibatide: 62.6%; Abciximab: 56.3%; Adjusted Difference: 7.1%, 95% CI: (2.7%, 17.0%)	

			interpretation.

Note: NCT is term used by ClinicalTrials.gov when assigning a unique clinical trial identifier; CI=Confidence Interval

F. Secondary Outcome Registry-Article Comparison

Among the 96 trials, there were 2089 endpoints designated as secondary efficacy endpoints within either ClinicalTrials.gov or a published article, 619 (30%) of which were described in both sources, 421 (20%) of which were described only on ClinicalTrials.gov, and 1049 (50%) of which were described only in the article. There was a median of 5 (IQR 2-12) secondary efficacy endpoints per ClinicalTrials.gov entry and 11 (IQR 7-21) per publication.

Among the 619 secondary endpoints defined in both sources, results of 228 (37%) could not be compared, often because results values were not reported within ClinicalTrials.gov. Among 391 comparable secondary endpoints, 53 (14%) had discordant results. Therefore, in total, only 338 of 2089 endpoints (16%) designated as a secondary outcome within either ClinicalTrials.gov or the published article were described in both sources and reported verifiable and concordant results. **Figure 3** below summarizes the secondary endpoint completion and comparison findings.

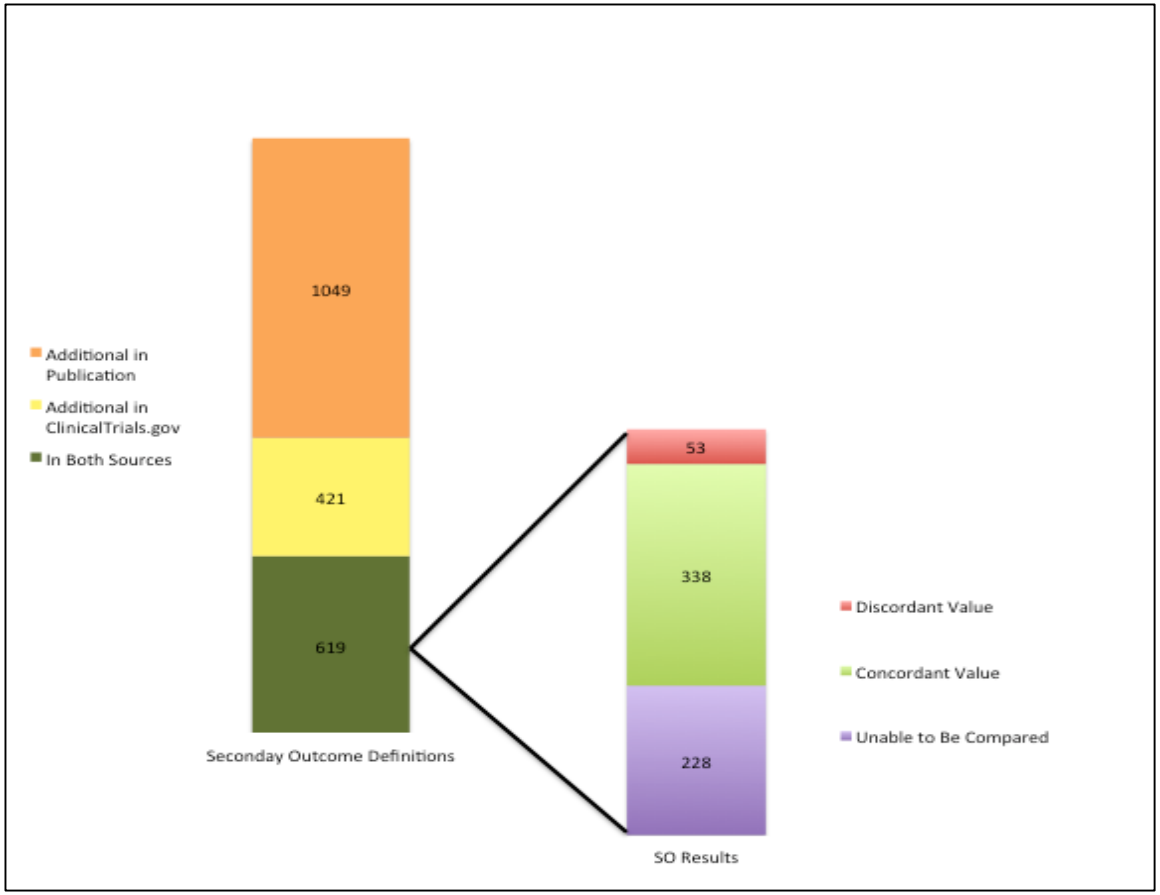

Figure 3: Secondary efficacy outcome endpoint definitions and results values reported on ClinicalTrials.gov and in corresponding journal articles for clinical trials registered in and reporting main results on ClinicalTrials.gov that were published in a Medline-indexed journal with an impact factor ≥ 10 between July 1, 2010 and June 30, 2011 (n=96).

G. Secondary Endpoint Sub-Analysis

As a secondary analysis, we examined the likelihood of reporting statistically significant endpoints in corresponding publications for the 1040 secondary efficacy endpoints listed on ClinicalTrials.gov. Statistical significance could not be determined for 184 (18%) endpoints because results were not reported on ClinicalTrials.gov, the trial only contained one arm, or results were in the form of a median or other summary statistic that required the underlying distribution to determine statistical significance, and no statistical analysis was provided on ClinicalTrials.gov. Among the remaining 856, 384 (45%) were statistically significant. However, only 559 (65%) of these 856 secondary endpoints were reported in the publication, and secondary endpoints with statistically significant results were more likely to be published when compared with endpoints whose results were not statistically significant (71% versus 61%; Odds Ratio=1.60, 95% CI=1.20-2.13, P=0.001).

H. Safety Outcome Registry-Article Comparison

Among 95 trials with adverse events results in both sources, we could not compare serious adverse events for 33 (36%) and other adverse events for 31 (33%), most often because of differences in severity stratification between the two sources. Among 62 trials with comparable serious adverse events, 39 (63%) were discordant, because of differing event criteria or frequency of reported serious adverse events. Among 64 trials with comparable other adverse events, 46 (72%) were discordant,

because of differing event criteria, frequency, or frequency reporting threshold of reported other adverse events. Overall, only 14 of 96 trials (15%) had concordant serious and other adverse event reporting between the two sources.

VII. DISCUSSION

A. Summary

In a sample of 96 trials published during a recent one-year period in high-impact journals that were registered within and reported results on ClinicalTrials.gov, we found that results reporting information available on ClinicalTrials.gov, including descriptions of the study cohort and intervention, as well as primary and secondary outcomes and safety outcomes, was complete for more than 90% of registered trials. However, upon comparing this information to the information provided in corresponding published articles, nearly every trial had at least one discrepancy in the results descriptions or values reported on ClinicalTrials.gov. Moreover, in many instances, results for both sources could not be compared because of differences in the presentation or analysis of the results. Our findings raise concerns about whether and which of the results reported on ClinicalTrials.gov or in published articles were accurate and about the usefulness of current clinical trial result reporting efforts to inform research and practice, as publicly reported results at times disagree with, or even contradict, findings reported in the peer-reviewed literature.

Many of the discrepancies between results reported on ClinicalTrials.gov and in the published articles had important implications for clinical practice and interpretation of trial findings. For instance, for some trials, the study intervention differed between what was reported on ClinicalTrials.gov and the published article, leading to uncertainty over how to interpret the results for practice. Similarly, a third of discordantly reported primary outcome results led to different trial interpretation between the two sources

and no discrepancy of this nature should ever occur. Discrepant primary outcome definitions and results have obvious potential to confuse or potentially mislead investigators, clinicians, and patients about the effectiveness of an intervention. However, it is also important to note that many of the discrepancies observed were not clinically meaningful. For instance, we observed minor discrepancies in the number of subjects studied as well as in the results values reported for primary and secondary outcomes and safety outcomes, many of which did not lead to different trial interpretation between the two sources.

B. Setting in the Developing Literature

As compared to studies done prior to ours, our findings were similar in terms of demonstrating outcome reporting bias in the published literature. Ewart *et al.,* for instance, found just over 30% of primary outcomes differed between trial registry – including, but not limited to, ClinicalTrials.gov – and publications in high-impact journals in a cross-section examined in 2007, before the passing of FDAAA.[49] While we found 15% of primary outcomes were described discordantly between ClinicalTrials.gov and the high-impact publications in our sample, it is promising that our rate was approximately half of that found by Ewart *et al.* now that the FDAAA mandates are in place.

Since the completion of our study, similar studies have come out in the literature. In one, by Hartung *et al.,* clinical trial results reported on ClinicalTrials.gov were compared to those reported in corresponding journal publications for a random

sample of phase 3 and 4 clinical trials.[77] Unlike our study, which examined studies in the highest-impact journals that likely represented a "best-case" scenario, this trial examined a random sample with trials published in a variety of journals. Strikingly, and reassuringly, our results were highly consistent with those uncovered by Hartung and colleagues.

Indeed, Hartung and colleagues found that 15% of trials reported discrepant primary outcome descriptions between ClinicalTrials.gov and corresponding publications, whereas 20% inconsistently reported primary outcome values.[77] We found that 15% of primary outcomes were described inconsistently between ClinicalTrials.gov and corresponding publications and 16% of primary results values were discordant between the two sources. Hartung and colleagues also found that 35% of trials had reporting discrepancies for serious adverse events.[77] We, too, found that 41% of trials had discrepant serious adverse event reporting. As in our study, this group similarly found that inconsistencies were frequently due to underreporting or omission of adverse events from publications. Whereas Hartung and colleagues found that 37% of trials with at least one frequently reported adverse event in ClinicalTrials.gov had discrepant adverse event reporting, we found slightly higher rates of these discrepancies.[77] Although the results reported on ClinicalTrials.gov and in corresponding publications are not yet consistent, it is reassuring that the studies examining this issue are.

Similarly, Killeen *et al.* published a study in *Annals of Surgery* to compare trial registration and primary outcome reporting in trials published specifically in surgery

journals.[78] Like ours, their sample came from studies published in the ten surgical journals with the highest impact factors, and all were either members of the ICMJE or specifically required trial preregistration as a prerequisite to publication. However, the trials in this sample were registered in several different clinical trial registries, including, but not limited to, ClinicalTrials.gov. They found that that fewer than half of the 246 trials in their sample were adequately registered on ClinicalTrials.gov, which included those that were registered after the study completion as well as those that did not have a primary endpoint, had an unclear primary endpoint, or had a primary endpoint without a defined timepoint in the registry entry.[78] Among the 108 trials deemed to be "adequately registered," the group found that nearly 30% had a discrepancy between the registered and published primary outcome, a number about twice as high as that found in our sample of trials published across specialties.[78] Among those results that could be compared, the group found the discrepancy favored a positive result in over 90% of cases – a high, but perhaps not surprising, number.[78] This finding was in line with our secondary endpoint sub-analysis, which also suggested positive outcome reporting bias among publications, as well as with previous findings in the literature.[7] Another study by Chahal *et al.* examining 34 clinical trials of operative and non-operative interventions in the field of orthopedic surgery registered in ClinicalTrials.gov and reported in publications showed that 80% had at least one discrepant piece of information between the two sources, with 35% of trials having a discordance between the registered and published primary endpoint.[79] As shown by these two recent studies, the findings in the specifically surgical world are not far off from, if not wrought by more

discrepancy than, those in the high-impact general medical literature as demonstrated by our work.

C. Implications

There are several possible explanations for the discrepancies we observed between the results information and values reported on ClinicalTrials.gov and in published articles. The most likely cause of these differences is reporting and typographical errors. Results reporting is a relatively new phenomena and is the responsibility of trial investigators and funders, who are less experienced with public dissemination of their findings and may in fact have become accustomed to a system where journal editors provide critical peer review and data checks to ensure accuracy. Similarly, the space available to report results in published articles is limited, potentially leading to incomplete reporting in articles while complete reporting is available on ClinicalTrials.gov. While this explanation may justify some discrepancies in adverse event reporting, particularly less serious adverse events, which have been shown to be less likely to be published,[80] it cannot account for the sizable discrepancy in secondary endpoint reporting. Other possibilities include differing reporting requirements, such that the format required by ClinicalTrials.gov may have differed from the journals, possibly leading to re-analysis and error; or that results might have been posted to ClinicalTrials.gov before the trial was completed, or alternatively have been published before the final data set was locked.

Lastly, and most worrisomely, discrepancies may also be due to investigators and funders disseminating more favorable findings in published articles, which are more likely to be read by clinicians and influence practice. We determined that statistically significant secondary endpoint results were more likely to be included in published articles, suggesting that investigators or publishers may be biased in the selection of endpoints to include in publications. Thus, our work, along with the other recent findings in the literature, raises important concerns that even with clinical trial registration and results reporting, selective result reporting continues to distort the medical evidence published in biomedical journals. Other examples of selectively reporting favorable outcomes, even despite public availability of trial findings through alternative sources, have been described.[7,37,42,47-50,52]

ClinicalTrials.gov has great potential to ensure the availability of complete and comprehensive results of all studies investigating medical products to inform and improve the practice of evidence-based clinical medicine, addressing known problems with the slow and incomplete dissemination of research findings via peer-reviewed published articles.[22,38,39,41,81] However, our study raises questions about the accuracy of the reported information, suggesting that further efforts are needed to improve the information being made available to investigators, clinicians, and patients. One solution may be to provide training to investigators and funders in the complete and accurate reporting of trial results, or simply to increase the resources available to ClinicalTrials.gov so that the organization could hire staff to input results for reporting. Another may be to offer external peer review of the reported results, providing

independent review of study reports before public results reporting. Alternatively, review of the results on ClinicalTrials.gov could be routinely undertaken by journal editors and reviewers upon receiving a trial for submission for publication. Finally, perhaps investigators could make their clinical trial data available for public use, allowing independent verification of the reported results,[82] or at least Clinical Study Reports or trial protocols, which can be similarly scrutinized and provide substantially more information than published articles.[31,52]

D. Limitations

There are important limitations that must be considered when interpreting our analysis. First, our study was limited to clinical trials that were registered and reported results on ClinicalTrials.gov and had main results published in high-impact journals. Despite studying a one year period, our sample for analysis only included 96 trials. Moreover, since half of registered studies are never published[41] and high-impact journals are members of the International Committee of Medical Journal Editors, which has been requiring clinical trial registration for nearly 10 years,[16] our study may represent the best case scenario in terms of the completeness and accuracy of results reporting. Nevertheless, in this early period after FDAAA enactment when few trials are reporting results,[68] our results can be used to inform and improve the results reporting system being used by ClinicalTrials.gov, ensuring its impact on clinical research and practice. Additionally, we were conservative in our determination of results concordance and discordance. We did not compare results when reporting was

numerical in one source but graphical in the other, sources differed in statistical analysis method, or results were stratified differently between the sources. Moreover, even when differences were observed, we were cautious when determining whether these differences would lead to different trial interpretation.

E. Conclusion

In conclusion, among trials published during a recent one year period in high-impact journals that were registered within and reported results on ClinicalTrials.gov, we found that results reporting information available on ClinicalTrials.gov was predominantly complete. However, upon comparing this information to corresponding published articles, nearly every trial had at least one discrepancy in the results descriptions or values reported on ClinicalTrials.gov, questioning the accuracy of both sources and raising concerns about the usefulness of results reporting to inform clinical practice and future research efforts.

REFERENCES

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *Bmj.* 1996;312(7023):71-72.
2. Johnson RT, Dickersin K. Publication bias against negative results from clinical trials: three of the seven deadly sins. *Nature clinical practice. Neurology.* 2007;3(11):590-591.
3. Zimerman AL. Evidence-based medicine: a short history of a modern medical movement. *The virtual mentor : VM.* 2013;15(1):71-76.
4. GH G. Evidence-based medicine. *ACP journal club.* 1991;114(suppl 2):A-16.
5. Zarrintan S, Aslanabadi S, Rikhtegar R. Early contributions of Abu Bakr Muhammad Ibn Zakariya Razi (865-925) to evidence-based medicine. *International journal of cardiology.* 2013;168(1):604-605.
6. Laurence J. No more boring science, no more waste in clinical trials. *Translational research : the journal of laboratory and clinical medicine.* 2009;153(1):1-3.
7. Dwan K, Gamble C, Williamson PR, Kirkham JJ, Reporting Bias G. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PloS one.* 2013;8(7):e66844.
8. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Annals of internal medicine.* 2010;152(11):726-732.
9. Junod S. FDA and Clinical Drug Trials: A Short History. *About FDA* http://www.fda.gov/AboutFDA/WhatWeDo/History/Overviews/ucm304485.htm. Accessed 27 Dec, 2014.
10. Trends, Charts, and Maps. *ClinicalTrials.gov* 2014; https://clinicaltrials.gov/ct2/resources/trends. Accessed Dec 27, 2014.

11. Johnston SC, Rootenberg JD, Katrak S, Smith WS, Elkins JS. Effect of a US National Institutes of Health programme of clinical trials on public health and costs. *Lancet.* 2006;367(9519):1319-1327.

12. Glasziou P, Djulbegovic B, Burls A. Are systematic reviews more cost-effective than randomised trials? *Lancet.* 2006;367(9528):2057-2058.

13. Kanouse DE, Jacoby I. When does information change practitioners' behavior? *International journal of technology assessment in health care.* 1988;4(1):27-33.

14. Lomas J. Diffusion, dissemination, and implementation: who should do what? *Annals of the New York Academy of Sciences.* 1993;703:226-235; discussion 235-227.

15. Sen A. Physician Peer Effects, Organization, and the "Disadoption" of Low-Value Services. The Wharton School of the University of Pennsylvania; 2014.

16. DeAngelis CD, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *Jama.* 2004;292(11):1363-1364.

17. Lomas J, Anderson GM, Domnick-Pierre K, Vayda E, Enkin MW, Hannah WJ. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. *The New England journal of medicine.* 1989;321(19):1306-1311.

18. Kanouse DE, Kallich JD, Kahan JP. Dissemination of effectiveness and outcomes research. *Health policy.* 1995;34(3):167-192.

19. Nair HS, Manchanda P, Bhatia T. Asymmetric social interactions in physician prescription behavior: The role of opinion leaders. *Journal of Marketing Research.* 2010;47(5):883-895.

20. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj.* 2008;336(7650):924-926.

21. Manning PR, Denson TA. How internists learned about cimetidine. *Annals of internal medicine.* 1980;92(5):690-692.

22. Ross JS, Tse T, Zarin DA, Xu H, Zhou L, Krumholz HM. Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. *Bmj.* 2012;344:d7292.

23. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet.* 1991;337(8746):867-872.

24. Dickersin K, Min YI. Publication bias: the problem that won't go away. *Annals of the New York Academy of Sciences.* 1993;703:135-146; discussion 146-138.

25. Dickersin K, Min YI. NIH clinical trials and publication bias. *The Online journal of current clinical trials.* 1993;Doc No 50:[4967 words; 4953 paragraphs].

26. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *Jama.* 1992;267(3):374-378.

27. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H, Jr. Publication bias and clinical trials. *Controlled clinical trials.* 1987;8(4):343-353.

28. Song F, Parekh S, Hooper L, et al. Dissemination and publication of research findings: an updated review of related biases. *Health technology assessment.* 2010;14(8):1-+.

29. Simes RJ. Publication bias: the case for an international registry of clinical trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 1986;4(10):1529-1541.

30. Dickersin K. How important is publication bias? A synthesis of available data. *AIDS education and prevention : official publication of the International Society for AIDS Education.* 1997;9(1 Suppl):15-21.

31. Chalmers I. Underreporting research is scientific misconduct. *Jama.* 1990;263(10):1405-1408.

32. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *The Cochrane database of systematic reviews.* 2009(1):MR000006.

33. Lee K, Bacchetti P, Sim I. Publication of clinical trials supporting successful new drug applications: a literature analysis. *PLoS medicine.* 2008;5(9):e191.

34. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *The Cochrane database of systematic reviews.* 2007(2):MR000005.

35. Liebeskind DS, Kidwell CS, Sayre JW, Saver JL. Evidence of publication bias in reporting acute stroke clinical trials. *Neurology.* 2006;67(6):973-979.

36. Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. *PLoS medicine.* 2008;5(11):e217; discussion e217.

37. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *The New England journal of medicine.* 2008;358(3):252-260.

38. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *Bmj.* 1997;315(7109):640-645.

39. Dwan K, Altman DG, Arnaiz JA, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PloS one.* 2008;3(8):e3081.

40. Hall R, de Antueno C, Webber A. Publication bias in the medical literature: a review by a Canadian Research Ethics Board. *Can J Anaesth.* 2007;54(5):380-388.

41. Ross JS, Mulvey GK, Hines EM, Nissen SE, Krumholz HM. Trial publication after registration in ClinicalTrials.Gov: a cross-sectional analysis. *PLoS medicine.* 2009;6(9):e1000144.

42. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *Jama.* 2004;291(20):2457-2465.

43. Menzel HaK, E. Social Relations and Innovation in the Medical Profession: The Epidemiology of A New Drug. *Public Opinion Quarterly.* 1955;19(4):337-352.

44. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *Jama.* 1996;276(8):637-639.

45. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *The New England journal of medicine.* 2007;356(24):2457-2471.

46. Bennett CL, Silver SM, Djulbegovic B, et al. Venous thromboembolism and mortality associated with recombinant erythropoietin and darbepoetin administration for the treatment of cancer-associated anemia. *Jama.* 2008;299(8):914-924.

47. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *Bmj.* 2005;330(7494):753.

48. Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* 2004;171(7):735-740.

49. Ewart R, Lausen H, Millian N. Undisclosed changes in outcomes in randomized controlled trials: an observational study. *Annals of family medicine.* 2009;7(6):542-546.

50. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *Jama.* 2009;302(9):977-984.

51. Riveros C, Dechartres A, Perrodeau E, Haneef R, Boutron I, Ravaud P. Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. *PLoS medicine.* 2013;10(12):e1001566; discussion e1001566.

52. Wieseler B, Kerekes MF, Vervoelgyi V, McGauran N, Kaiser T. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. *Bmj.* 2012;344:d8141.

53. Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. *The New England journal of medicine.* 2009;361(20):1963-1971.

54. Wood AJ. Progress and deficiencies in the registration of clinical trials. *The New England journal of medicine.* 2009;360(8):824-830.

55. Rennie D. Fair conduct and fair reporting of clinical trials. *Jama.* 1999;282(18):1766-1768.

56. Laine C, Horton R, DeAngelis CD, et al. Clinical trial registration: looking back and moving ahead. *Lancet.* 2007;369(9577):1909-1911.

57. Roy ASA. *Stifling New Cures: The True Cost of Lengthy Clinical Trial Drugs.* Manhattan Institute for Policy Research;2012 March.

58. Dickersin K. Why register clinical trials?--Revisited. *Controlled clinical trials.* 1992;13(2):170-177.

59. Chalmers TC. Randomize the first patient. *The New England journal of medicine.* 1977;296(2):107.

60. Easterbrook PJ. Directory of registries of clinical trials. *Statistics in medicine.* 1992;11(3):345-359.

61. Easterbrook PJ. Directory of registries of clinical trials. *Statistics in medicine.* 1992;11(3):363-423.

62. Zarin DA, Tse T, Ide NC. Trial Registration at ClinicalTrials.gov between May and October 2005. *The New England journal of medicine.* 2005;353(26):2779-2787.

63. Food and Drug Modernization Act of 1997. In: Congress t, ed. *Public Law 105-115*1997.

64. Editors ICoMJ. About ICMJE. 2014; http://www.icmje.org/about-icmje/. Accessed December 21, 2014.

65. Food and Drug Administration Amendments Act of 2007. In: Congress t, ed. Vol Public Law 110-852007.

66. FDAAA 801 Requirements. *Submit Studies* https://clinicaltrials.gov/ct2/manage-recs/fdaaa. Accessed 21 Dec 2014.

67. Law MR, Kawasumi Y, Morgan SG. Despite law, fewer than one in eight completed studies of drugs and biologics are reported on time on ClinicalTrials.gov. *Health affairs.* 2011;30(12):2338-2345.

68. Prayle AP, Hurley MN, Smyth AR. Compliance with mandatory reporting of clinical trial results on ClinicalTrials.gov: cross sectional study. *Bmj.* 2012;344:d7373.

69. Number of Registered Studies Over Time. *Trends, Charts, and Maps* 2013; http://www.clinicaltrials.gov/ct2/resources/trends - RegisteredStudiesOverTime. Accessed March 6, 2013.

70. ClinicalTrials.gov.

ClinicalTrials.gov "Basic Results" Data Element Definitions (DRAFT). 2013; http://prsinfo.clinicaltrials.gov/results_definitions.html. Accessed March 3, 2013.

71. Wilkens P, Scheel IB, Grundnes O, Hellum C, Storheim K. Effect of glucosamine on pain-related disability in patients with chronic low back pain and degenerative lumbar osteoarthritis: a randomized controlled trial. *JAMA : the journal of the American Medical Association.* 2010;304(1):45-52.

72. ClinicalTrials.gov. The Glucosamine-study: The Effect of Glucosamine in Treatment of Chronic Low Back Pain. 2011; http://www.clinicaltrials.gov/ct2/show/NCT00404079?term=nct00404079&rank=1. Accessed 31 March, 2013.

73. Quinn JF, Raman R, Thomas RG, et al. Docosahexaenoic acid supplementation and cognitive decline in Alzheimer disease: a randomized trial. *JAMA : the journal of the American Medical Association.* 2010;304(17):1903-1911.

74. ClinicalTrials.gov. DHA (Docosahexaenoic Acid), an Omega 3 Fatty Acid, in Slowing the Progression of Alzheimer's Disease. 2013; http://www.clinicaltrials.gov/ct2/show/NCT00440050?term=docosahexanoic+acid&rank=19. Accessed 31 March, 2013.

75. Cicardi M, Levy RJ, McNeil DL, et al. Ecallantide for the treatment of acute attacks in hereditary angioedema. *The New England journal of medicine.* 2010;363(6):523-531.

76. ClinicalTrials.gov. Efficacy and Safety Study of DX-88 to Treat Acute Attacks of Hereditary Angioedema (HAE). 2010; http://www.clinicaltrials.gov/ct2/show/NCT00262080?term=NCT00262080&rank=1. Accessed 31 March, 2013.

77. Hartung DM, Zarin DA, Guise JM, McDonagh M, Paynter R, Helfand M. Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications. *Annals of internal medicine.* 2014;160(7):477-483.

78. Killeen S, Sourallous P, Hunter IA, Hartley JE, Grady HL. Registration rates, adequacy of registration, and a comparison of registered and published primary outcomes in randomized controlled trials published in surgery journals. *Annals of surgery.* 2014;259(1):193-196.

79. Chahal J, Tomescu SS, Ravi B, et al. Publication of sports medicine-related randomized controlled trials registered in ClinicalTrials.gov. *The American journal of sports medicine.* 2012;40(9):1970-1977.

80. Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. *Archives of internal medicine.* 2009;169(19):1756-1761.

81. Ross JS, Mocanu M, Lampropulos JF, Tse T, Krumholz HM. Time to Publication Among Completed Clinical Trials. *JAMA internal medicine.* 2013:1-3.

82. Ross JS, Krumholz HM. Ushering in a new era of open science through data sharing: the wall must come down. *JAMA : the journal of the American Medical Association.* 2013;309(13):1355-1356.

APPENDIX

Appendix Table A1: **Trial intervention descriptions reported on ClinicalTrials.gov and in the corresponding journal article that were determined to be discordant (n=15).**

ClinicalTrials.gov Identifier	**Trial Intervention Description[a]**		**Trial Intervention Discrepancy Explanation**
	ClinicalTrials.gov	**Published Article**	
NCT00099268	Carbidopa/Levodopa/Entacapone: Patients received Carbidopa/ levodopa/entacapone tablets. The study was designed as **a flexible dose trial (200-1000 mg/day levodopa). The target dose was 400 mg/day levodopa** administered orally as 4 equal doses 4 times a day with 3.5-hour dosing intervals for a treatment period of 134 to 208 weeks. Carbidopa/Levodopa: Patients	L-dopa/carbidopa in both groups was initiated at a dose of **50/12.5mg twice daily, and titrated to 100/25 (target dose) or 150/37.5mg 4x daily** administered at 3.5-hour intervals. For patients in the LCE group, **entacapone 200mg** was	Dosage of intervention ClinicalTrials.gov states the levodopa dose is flexible, up to 1000 mg/day, while the article lists three distinct doses with a maximum of 600 mg/day levodopa.

	received Immediate release carbidopa/levodopa tablets. The study was designed as **a flexible dose trial (200-1000 mg/day levodopa). The target dose was 400 mg/day levodopa** administered orally as 4 equal doses 4 times a day with 3.5-hour dosing intervals for a treatment period of 134 to 208 weeks.	administered with each LC dose.	
NCT00108953	Sorafenib + Doxorubicin: "Sorafenib + Doxorubicin" -- combination therapy: **Sorafenib (Nexavar, BAY43-9006) 200 mg tablets by mouth (orally) twice daily** + doxorubicin 60 mg/m^2 intravenous infusion every 21 days for 6 cycles (18 weeks). Placebo + Doxorubicin: "Placebo + Doxorubicin" -- monotherapy: Sorafenib (Nexavar, BAY43-9006) **matching placebo tablets by mouth (orally) twice daily** + doxorubicin 60 mg/m^2 intravenous infusion every 21 days for 6 cycles (18 weeks).	Patients received 60 mg/m2 of doxorubicin intravenously every 21 days for a maximum of 360 mg/m^2 plus either **400 mg of sorafenib or placebo orally twice daily**.	Dosage of intervention ClinicalTrials.gov does not state number of 200 mg tablets and thus suggests Sorafenib and placebo dosing at 400 mg/day, while article states dosing is 800 mg/day for each.
NCT00141102	Celecoxib: 200 milligrams (mg) twice daily (BID) plus omeprazole	Patients were randomly assigned	Description of placebo

	placebo and diclofenac slow release (SR) placebo Oral Diclofenac Plus Omeprazole: Oral diclofenac SR (75 mg BID) plus omeprazole (20 mg once daily [QD]) and **celecoxib placebo.**	in a 1:1 ratio to receive either celecoxib 200 mg twice a day (Pfizer Inc, New York, NY, USA) or diclofenac slow release 75 mg twice a day (Novartis Pharmaceuticals UK Ltd, Camberley, UK) plus omeprazole 20 mg once a day (AstraZeneca LP, Westborough, MA, USA) for 6 months.	Additional celecoxib placebo is described on ClinicalTrials.gov and not in the article.
NCT00154310	Everolimus + Mycophenolate Sodium: **Everolimus tablets orally twice a day to maintain a level of 6- 10 ng/mL** and enteric-coated mycophenolate sodium orally twice a day to achieve a target dose of 1440 mg/day. Corticosteroids were added to the immunosuppressive regimen with a minimum dose of 5 mg prednisolone or equivalent and had to be continued throughout the first year. Cyclosporine	**Everolimus (0·75 mg twice a day, orally)** was started the day after the 4·5-month assessment. Ciclosporin replacement was done in a stepwise manner (step 1: 50%, step 2: 25%, step 3: 0) in 4 weeks or less. **Target trough**	Dosage of intervention While target trough concentrations of everolimus are concordant between ClinicalTrials.gov and the article, the everolimus dosing (0.75 mg twice daily) is enumerated only in the article and not

	withdrawal started from Month 4.5 post-transplant. Cyclosporine + Mycophenolate Sodium: **Cyclosporine tablets orally twice a day to achieve protocol specific target levels** and enteric-coated mycophenolate sodium orally twice a day to achieve a target dose of 1440 mg/day. Corticosteroids were added to the immunosuppressive regimen with a minimum dose of 5mg prednisolone or equivalent and had to be continued throughout the first year.	**concentrations of everolimus were 3–8 ng/mL in step 1 and 6–10 ng/mL thereafter. Target ciclosporin concentrations were based on either C-0h (trough concentration) or C-2h (drug concentration 2 h after dose) in whole blood, according to local practice. Up to month 4·5 (time of randomisation), C-0h target was 150–220 ng/mL, and C-2h were 1100–1400 ng/mL for month 1, 950–1300 ng/mL for month 2, and 800–1200 ng/mL for month 3 onwards. From months 4·5–6 after transplantation, C-0h (and C-2h)**	on ClinicalTrials.gov. Cyclosporine numerical dosage is not specified on ClinicalTrials.gov, while the target levels are detailed in full in the article.

		targets were 120–180 ng/mL (700–1000 ng/mL), and after month 6, 100–150 ng/mL (500–800 ng/mL). Corticosteroids were mandatory (≥5 mg/day prednisolone or equivalent) and were administered according to local practice. Target dose of enteric-coated mycophenolate sodium was 1440 mg/day, orally, for both groups throughout the first year; dose adjustments were permitted.	
NCT00171210	Crossover: Participants treated with Deferoxamine (DFO) during the core study and crossed over to receive Deferasirox (ICL670) orally once a day during the	**Deferasirox dose was initially assigned according to LIC [liver iron concentration]** at	Dosage of intervention Basis of dosage is described as body

	extension study. **Dosage based on body weight.** ICL670: Participants treated with Deferasirox (ICL670) orally once a day during the core study and continued this treatment in the extension study. **Dosage based on body weight.**	the start of deferasirox treatment, whereby patients with LIC values of 2-3, > 3-7, > 7-14, and > 14 mg Fe/g dry weight (dw) were assigned deferasirox doses of 5, 10, 20, or 30 mg/kg/ day, respectively.	weight on ClinicalTrials.gov and LIC in the article.
NCT00177671	Donepezil: Treatment with antidepressants (escitalopram (10mg to 20mg daily), **venlafaxine** (150mg to 300mg daily), duloxetine (20mg to 120mg daily) plus donepezil (5mg to 10mg daily). Placebo: Treatment with antidepressants (escitalopram (10mg to 20mg daily), **venlafaxine** (150mg to 300mg daily), duloxetine (20mg to 120mg daily) plus placebo.	Patients initially received open antidepressant pharmacotherapy with escitalopram oxalate (≤20 mg/d). Those not responding fully were switched to a serotonin noradrenergic reuptake inhibitor (duloxetine hydrochloride, ≤120 mg/d), followed as needed by **aripiprazole**	Description of intervention Additional antidepressants include venlafaxine on ClinicalTrials.gov only and aripiprazole in the article only.

		augmentation (≤15 mg/d) to achieve full response.	
NCT00197106	Salmeterol/FP 50/100 Mcg Plus Placebo: One puff Salmeterol/FP 50/100 mcg plus **one puff placebo** (matching one puff of FP in the 200 mcg group) BID via DISKUS inhaler.	Symptomatic children were randomized to either FP, 200 mg twice a day Diskus (FP group), or salmeterol/FP, 50/100 twice a day Diskus (SFP group), for 26 weeks.	Description of placebo Additional placebo described on ClinicalTrials.gov.
NCT00243919	Early Locomotor Training Program: Stepping on a treadmill with partial body weight support and manual assistance as needed for 20-30 minutes at 2.0 mph, followed by a progressive overground walking program for **20 minutes** delivered at 2 months post-stroke. Late Locomotor Training Program: Stepping on a treadmill with partial body weight support and manual assistance as needed for 20-30 minutes at 2.0 mph, followed by a progressive	Locomotor training included stepping on a treadmill with partial body-weight support and manual assistance as needed for 20 to 30 minutes at 3.2 km per hour (0.89 m per second [2.0 mi per hour]), followed by a progressive program of walking over ground for **15 minutes.**	Duration of administration Overground walking program is described as lasting 20 minutes on ClinicalTrials.gov and 15 minutes in the article.

	overground walking program for **20 minutes** delivered at 6 months post-stroke.		
NCT00255840	First line antiretroviral regimen monitored by a HIV-trained medical doctor: 1) Stavudine **(>60 kg: 40 mg twice daily and <60 kg: 30 mg twice daily)** 2) Lamivudine (150mg twice daily) and 3) Efavirenz (600mg daily). For women of child bearing potential with a CD4+ count <250 cells/mm3, Nevirapine (200 mg daily x 14 days, then 200 mg twice daily) and for women with a CD4+ count > 250 cells/mm3, Lopinavir/ritonavir (400/100mg twice daily).	Regimens initially prescribed by the clinical safety team included a nucleoside backbone of stavudine and lamivudine, with a choice of efavirenz, nevirapine, or lopinavir plus ritonavir. The initial dose of stavudine was **40 mg daily** for individuals weighing more than 60 kg, which was reduced to **30 mg for all patients** from mid-2007 in line with WHO recommendations. Efavirenz was the preferred non-nucleoside for men and women not	Frequency of administration ClinicalTrials.gov states stavudine dosage is 80 mg or 60 mg per day, while article states dosage is 40 mg or 30 mg per day.

		wishing to become pregnant and willing to maintain both barrier and hormonal contraception throughout the study. Women of childbearing potential were prescribed nevirapine if their CD4+ lymphocyte count was less than 250 cells per µL, or lopinavir plus ritonavir if their count was 250 cells per µL or greater. Pregnant women, who were allowed to enrol after their first trimester, were prescribed either nelfinavir or lopinavir plus ritonavir.	
NCT00298766	Single Agent VELCADE: Bortezomib **0.7, 1.0, 1.3 and 1.6**	Patients were enrolled to receive	Dosage of intervention

	mg/m^2 once weekly (QW) 4 doses in a 5 week cycle, and **0.7, 1.0, 1.3 mg/m^2** twice weekly (BIW) 4 doses in a 3 week cycle	bortezomib at **doses up to 1.6 mg/m^2** on days 1, 8, 15, and 22 of 35-day cycles (once-weekly regimen) and then at **doses up to 1.3 mg/m^2** on days 1, 4, 8, and 11 of 21-day cycles (twice-weekly regimen); details of the phase 1 dose-escalation component have been reported previously.	The article lists that there are dosages up to the maximum dose, but only enumerates the maximum dose for each cycle, while ClinicalTrials.gov enumerates all dose amounts.
NCT00423670	PegIntron **(1.5 µg/kg, once weekly [QW])** plus ribavirin (800 to 1400 mg/day) for 48 weeks.	Part 1 consisted of five treatment groups: one was a control group in which patients **received peginterferon alfa-2b 1·5 µg/kg plus ribavirin 800–1400 mg per day for 48 weeks (PR48**). Two groups, the lead-in groups, received	Frequency of administration PegIntron 1.5 µg/kg dosage is described as once weekly on ClinicalTrials.gov but appears to be daily in the article.

		peginterferon alfa-2b 1·5 µg/kg and ribavirin 800–1400 mg daily for 4 weeks (PR4)…	
NCT00463567	Indacaterol 150 µg (Continued Into Stage 2): In the morning, Indacaterol 150 µg once daily orally inhaled via a single dose dry powder inhaler (SDDPI) + **Placebo to Indacaterol delivered via SDDPI** + Placebo to Formoterol delivered via Aerolizer. In the evening, Placebo to Formoterol delivered via Aerolizer. Participated in the 2 week Stage 1 and continued treatment up to 26 weeks in Stage 2. Placebo to Formoterol inhalation in the morning and in the evening was discontinued after Stage 1. Indacaterol 300 µg (Continued Into Stage 2): In the morning, Indacaterol 300 µg once daily orally inhaled via a SDDPI + **Placebo to Indacaterol delivered via SDDPI** + Placebo to Formoterol delivered via	Patients received double-blind indacaterol 150 or 300 mg or placebo via single-dose dry powder inhaler, or open-label tiotropium 18 mg via HandiHaler (Boehringer Ingelheim, Ridgefield, CT). (Blinded tiotropium was not available.) Treatments were taken once daily at 08:00 to 10:00.	Description of placebo Additional placebo to Indacaterol is described on ClinicalTrials.gov but not in the article.

	Aerolizer. In evening, Placebo to Formoterol delivered via Aerolizer. Participated in the 2 week Stage 1 and continued treatment up to 26 weeks in Stage 2. Placebo to Formoterol inhalation in the morning and in the evening was discontinued after Stage 1. Tiotropium 18 µg (Continued Into Stage 2): Tiotropium 18 µg dry powder capsules delivered (open label) via manufacturer's proprietary SDDPI, (Handihaler®). Participated in the 2 week Stage 1 and continued treatment up to 26 weeks in Stage 2. Placebo (Continued Into Stage 2): In the morning, Placebo to Indacaterol delivered via two SDDPI devices + Placebo to Formoterol delivered via Aerolizer. In the evening, Placebo to Formoterol delivered via Aerolizer. Participated in the 2 week Stage 1 and continued treatment up to 26 weeks in		

	Stage 2. Placebo to Formoterol inhalation in the morning and in the evening was discontinued after Stage 1.		
NCT00551642	Inhaled Nitric Oxide (NO): Inhaled NO administered by nasal continuous positive airway pressure, nasal cannula or face mask at 5 parts per million (ppm) **for a maximum of 21 days** Placebo (Nitrogen): Placebo Nitrogen gas administered by nasal continuous positive airway pressure, nasal cannula or face mask at 5 parts per million **for a maximum of 21 days**	Infants were given inhaled nitric oxide (5 parts per million [ppm]) or placebo-equivalent nitrogen gas…Therapy was given **for at least 7 days**, **up to a maximum of 21 days**. If patients needed mechanical ventilation for less than 7 days, therapy was completed through a nasal continuous positive airway pressure or nasal cannula.	Duration of administration The article lists a minimum and maximum duration of inhaled NO, but ClinicalTrials.gov reports only a maximum, implying there is no minimum duration.
NCT00642174	Prasugrel Then Clopidogrel: Prasugrel: Oral prasugrel 60-mg loading dose, followed **by 6 to 9 days of prasugrel** 10-mg/day tablet maintenance dose. Clopidogrel: Oral clopidogrel 600-mg loading dose, followed **by 6 to**	Patients who met all criteria for enrolment were randomized at the first visit to double-blind treatment of either prasugrel 60	Description of placebo & Duration of administration Placebo tablets are described in the article but not on

	9 days of clopidogrel 150-mg/day tablet maintenance dose. Clopidogrel Then Prasugrel: Clopidogrel: Oral clopidogrel 600-mg loading dose, followed **by 6 to 9 days** of clopidogrel 150-mg/day tablet maintenance dose. Prasugrel: Oral prasugrel 60-mg loading dose, followed by **6 to 9 days** of prasugrel 10-mg/day tablet maintenance dose.	mg LD orally followed by 10 mg/day MD for **7 days** or clopidogrel 600 mg LD orally [75 mg tablets Plavixw (clopidogrel bisulfate; Bristol–Myers Squibb /sanofi-aventis, Bridgewater, NJ, USA)] followed by 150 mg/day MD for **7 days** (Figure 1A)…**Patients were administered an equal number of identical tablets for either the LDs (six prasugrel 10 mg tablets and eight placebo tablets or eight clopidogrel 75 mg tablets and six placebo tablets) or MDs (one prasugrel 10 mg tablet and two placebo tablets or**	ClinicalTrials.gov. ClinicalTrials.gov states that dose duration is 6 to 9 days, while the article specifies a 7 day duration explicitly.

		two clopidogrel 75 mg tablets and one placebo tablet).	
NCT00703118	T12/PR48: 12 weeks of 750 mg telaprevir q8h followed by 4 weeks of Placebo in combination with 48 weeks of **Peg-IFN-alfa-2a and RBV at standard doses** T12(DS)/PR48: 4 weeks of Placebo followed by 12 weeks of 750 mg telaprevir q8h in combination with 48 weeks of **Peg-IFN-alfa-2a and RBV at standard doses** Pbo/PR48: 48 weeks of **Peg-IFN-alfa-2a and RBV at standard doses**	Telaprevir (Tibotec) was administered orally at a dose of 750 mg every 8 hours; **peginterferon alfa-2a (Pegasys, Roche) was administered subcutaneously at a dose of 180 µg per week; and ribavirin (Copegus, Roche) was administered orally at a dose of 1000 to 1200 mg per day…** In the T12PR48 group, 266 patients were assigned to receive telaprevir, peginterferon, and ribavirin for 12 weeks, followed by placebo plus peginterferon and ribavirin for 4	Description of placebo & Dosage of intervention Additional placebo is described for the PR48 (control) group in the article, but not on ClinicalTrials.gov. ClinicalTrials.gov reports only that doses are standard, while the article specifies numeric values for dosing.

		weeks, and then peginterferon plus ribavirin alone for 32 weeks. In the lead-in T12PR48 group, 264 patients were assigned to receive placebo, peginterferon, and ribavirin for 4 weeks, followed by telaprevir plus peginterferon plus ribavirin for 12 weeks, and then peginterferon plus ribavirin alone for 32 weeks. In the PR48 (control) group, 132 patients were assigned to receive **placebo**, peginterferon, and ribavirin for 16 weeks, followed by peginterferon plus ribavirin for 32 weeks. In all the groups, study drugs	

		were administered for 48 weeks**.**	

[a] Entries in the Trial Intervention Description columns are taken verbatim from ClinicalTrials.gov and the corresponding article, respectively, for each ClinicalTrials.gov identifier provided (emphasis added by this author).

Appendix Table A2: Primary efficacy outcome results values reported on ClinicalTrials.gov and in the corresponding journal article that were determined to be discordant (n=21).

ClinicalTrials.gov Identifier and Primary Efficacy Outcome	**Primary Efficacy Outcome Results**		**Difference in Trial Interpretation due to Discrepancy?**
	ClinicalTrials.gov	**Published Article**	
NCT00090285 Incidence of Human Papillomavirus (HPV) Related External Genital Warts, Perineal Intraepithelial Neoplasia (PIN), Penile, Perianal or Perineal Cancer, incidence per 100 person-years	qHPV Vaccine: 0.1 Placebo: 1.0 Percent relative risk reduction: 90.6 95% CI: (70.1, 98.2)	qHPV Vaccine: 0.11 Placebo: 1.10 Observed efficacy: 90.4% 95% CI: (69.2%, 98.1%) P<0.001	No The incidence values and observed efficacy are very similar between sources, likely with a similar test of statistical significance.

NCT00094887 Time to Resolution of Vaso-occlusive Pain Crisis (VOC), median hours (95% CI)	Inhaled Nitric Oxide: 61.83 95% CI: (41.75, 78.00) Placebo: 55.16 95% CI: (46.00, 72.00) No statistical analysis provided.	Inhaled Nitric Oxide: 73.0 95% CI: (46.0, 91.0) Placebo: 65.5 95% CI: (48.1, 84.0) P = 0.87	Yes The time to resolution in both groups is substantially lower on ClinicalTrials.gov than in the article, with the treatment time on ClinicalTrials.gov lower than the placebo time in the article, even if the test of statistical significance was likely the same in both sources.
NCT00108953 Time to Progression, median	Sorafenib + Doxorubicin: 263 days 95% CI: (146, 384) Placebo + Doxorubicin: 147 days 95% CI: (66, 244)	Sorafenib + Doxorubicin: 6.4 months 95% CI: (4.8, 9.2) Placebo + Doxorubicin: 2.8 months 95% CI: (1.6, 5)	Yes The median time to progression in both groups is substantially higher on ClinicalTrials.gov (9.6 and 4.8 months, respectively), even if the test of statistical significance between the groups was the same in both sources.

	P = 0.016	P = 0.02	
NCT00177671 Number of Participants With Recurrence of Major Depression	Donepezil: 19/67 95% CI: (16, 31) Placebo: 11/63 95% CI: (6, 18) HR = 3.97 SD = 2.09 95% CI: (1.00, 4.41) P = 0.05	Donepezil: 35% 95% CI: (24%, 46%) Placebo: 19% 95% CI: (9%, 29%) HR = 2.09 95% CI: (1.00, 4.41) λ^2=3.97 P = 0.05	Yes The percentage of participants with major depression recurrences is lower (28.4% and 17.5%, respectively) on ClinicalTrials.gov than in the article, and the hazard ratio listed on ClinicalTrials.gov is 2-fold higher than that listed in the article (3.97 vs. 2.09).
NCT00262080 Treatment Outcome Score at 4 Hours Post-Dose, mean units	Ecallantide: 49.5 SD: 59.43 Placebo: 18.5 SD: 67.78	Ecallantide: 46.8 SD: 59.3 Placebo: 21.3 SD: 69.0	No The absolute treatment outcome score and difference between arms is similar in both sources, and the test of statistical

	P = 0.037	P = 0.004	significance shows a significant difference in both sources.
NCT00281918 Progression-free Survival (PFS), median	Fludarabine/Cyclophospha mide: 981.0 days Range: (1, 1343) Fludarabine/Cyclophospha mide/ Rituximab: 1212.0 days Range: (1, 1372) P<0.0001	Fludarabine/Cyclophosphami de: 32.8 months 95% CI: (29.6, 36.0) Fludarabine/Cyclophosphami de/ Rituximab: 51.8 months 95% CI: (46.2, 57.6) P<0.0001	Yes PFS is substantially lower in the rituximab arm reported on ClinicalTrials.gov (39.8 months) than in the article, even if the test of statistical significance is the same in both sources.
NCT00335452 First Occurrence of CV Death / MI / Stroke - ASA	Clopidogrel + ASA Low Dose:	Clopidogrel + ASA Low Dose: 549/12,579	No The composite primary outcome results are quite close,

Dose Comparison	546/12,563 Clopidogrel + ASA High Dose: 527/12,498 P = 0.6047	Clopidogrel + ASA High Dose: 530/12,507 P = 0.61	with nearly identical tests of significance, in the two sources.
NCT00377260 The Time to Resolution of Symptoms, Defined as Acute Otitis Media-Severity of Symptoms (AOM-SOS) Score of 0 or 1, According to Treatment Assignment, Day 4, participants	Amoxicillin-Clavulanate: 87/144 Placebo: 78/147 No statistical analysis provided for this timepoint.	Amoxicillin-Clavulanate: 61% Placebo: 54% No statistical analysis provided for this timepoint.	No ClinicalTrials.gov percentages (60.4% and 53.1%, respectively) are very similar to those in the article, likely with a similar test of statistical significance.
NCT00377260 The Time to Resolution of	Amoxicillin-Clavulanate: 114/144	Amoxicillin-Clavulanate: 80% Placebo: 74%	No ClinicalTrials.gov percentages (79.1%

Symptoms, Defined as Acute Otitis Media-Severity of Symptoms (AOM-SOS) Score of 0 or 1, According to Treatment Assignment, Day 7, participants	Placebo: 106/147 No statistical analysis provided for this timepoint.	No statistical analysis provided for this timepoint.	and 72.1%, respectively) are very similar to those in the article, likely with a similar test of statistical significance.
NCT00377260 The Time to Resolution of Symptoms, Defined as Acute Otitis Media-Severity of Symptoms (AOM-SOS) Score of 0 or 1 on Two Consecutive Occasions, According to Treatment	Amoxicillin-Clavulanate: 58/144 Placebo: 51/147 No statistical analysis provided for this timepoint.	Amoxicillin-Clavulanate: 41% Placebo: 36% No statistical analysis provided for this timepoint.	No ClinicalTrials.gov percentages (40.2% and 34.7%, respectively) are very similar to those in the article, likely with a similar test of statistical significance.

Assignment, Day 4, participants			
NCT00377260 The Time to Resolution of Symptoms, Defined as Acute Otitis Media-Severity of Symptoms (AOM-SOS) Score of 0 or 1 on Two Consecutive Occasions, According to Treatment Assignment, Day 7, participants	Amoxicillin-Clavulanate: 96/144 Placebo: 76/147 No statistical analysis provided for this timepoint.	Amoxicillin-Clavulanate: 67% Placebo: 53% No statistical analysis provided for this timepoint.	No ClinicalTrials.gov percentages (66.7% and 51.7%, respectively) are very similar to those in the article, likely with a similar test of statistical significance.
NCT00404079 Roland Morris Disability Questionnaire, 1 year	Glucosamine Sulphate: 9, SD: 4; Placebo: 9, SD: 4 Odds Ratio: 4.5 ± 4; P=0.05	Glucosamine Sulphate: 4.8, 95% CI: (3.9, 5.6) Placebo: 5.5, 95% CI: (4.7, 6.4)	Yes ClinicalTrials.gov score is higher for both trial arms and statistical testing results are different in the two sources, leading to a difference in trial

		P=0.50	interpretation.
NCT00421733 Change From Baseline to the Last On-treatment Measurement in Urine Albumin to Creatinine Ratio (UACR) Levels Determined From the First Morning Void (FMV) Urine Collections Comparing Placebo to the Combined Paricalcitol Treatment Groups (1 Mcg and 2 Mcg), mean log mg/g creatinine	Placebo: -0.03 SD: 0.61 Combined Paricalcitol 1 Mcg and 2 Mcg: -0.18 SD: 0.70 Paricalcitol 1 Mcg: -0.15 SD: 0.72 Paricalcitol 2 Mcg: -0.22 SD: 0.69 Placebo vs. Combined Paricalcitol 1 Mcg and 2 Mcg: P = 0.071	Placebo: -3% 95% CI: (-16, 13) Combined Paricalcitol 1 Mcg and 2 Mcg: -16% 95% CI: (-24, -9) Paricalcitol 1 Mcg: -14% 95% CI: (-24, -1) Paricalcitol 2 Mcg: -0.20% 95% CI: (-30, -8) Placebo vs. Combined Paricalcitol 1 Mcg and 2 Mcg: P = 0.071 Placebo vs. Paricalcitol 1 Mcg:	No Change from baseline UACR levels are very similar with identical tests of significance.

	Placebo vs. Paricalcitol 1 Mcg: P = 0.229 Placebo vs. Paricalcitol 2 Mcg: P = 0.053	P = 0.23 Placebo vs. Paricalcitol 2 Mcg: P = 0.053	
NCT00426751 Number of Participants With Complete Sum ST Resolution (STR) 60 Min After Percutaneous Coronary Intervention (PCI) (Intent-to-Treat Population)	Eptifibatide: 124/214 Abciximab: 103/196 Adjusted Difference: 6.8% 95% CI: (-3.0%, 16.6%)	Eptifibatide: 62.6% Abciximab: 56.3% Adjusted Difference: 7.1% 95% CI: (2.7%, 17.0%)	Yes Confidence interval of the adjusted difference between arms crosses zero on ClinicalTrials.gov and does not in the article, suggesting a difference in the statistical significance of the results between the sources.

NCT00440050 Rate of Change on the ADAS-Cog 11, mean	Placebo: 7.98 SD: 9.84 Docosahexaenoic Acid (DHA): 8.27 SD: 8.9 No statistical analysis provided.	Placebo: 8.27 95% CI: (6.72, 9.82) Docosahexaenoic Acid (DHA): 7.98 95% CI: (6.51, 9.45) P = 0.41	No While values appear to be switched for the DHA and placebo arms between the two sources, interpretation is unchanged since the difference between values is not statistically significant.
NCT00440050 Rate of Change on CDR-SOB	Placebo: 2.87 SD: 2.93 Docosahexaenoic Acid (DHA): 2.93 SD: 2.83 No statistical analysis provided.	Placebo: 2.93 95% CI: (2.44, 3.42) Docosahexaenoic Acid (DHA): 2.87 95% CI: (2.44, 3.30) P = 0.68	No While values appear to be switched for the DHA and placebo arms between the two sources, interpretation is unchanged since the difference between values is not statistically significant.

NCT00484315 Target Lesion Failure (TLF) at 12 Months Post-index Procedure, number of participants.	TAXUS Element: 51/915 TAXUS Express: 19/309 P = 0.9996	TAXUS Element: 52/922 TAXUS Express: 19/313 P = 0.78	No Number of participants with TLF is similar between sources, and both sources report no significant difference between arms.
NCT00547534 Number of Participants With Progression Free Survival (PFS) at 2 Years	Lymphoma subjects: 13/29	Lymphoma subjects: 47% (of 29 evaluable patients)	No The number of participants with PFS at 2 years is approximately the same in both sources (45% and 47%, respectively), even if the reported values were different.
NCT00765817 Change in Glycosylated Hemoglobin (HbA1c), least squares mean	Exenatide: -1.71 SE: 0.09 Placebo: -1.00	Exenatide: -1.74 95% CI: (-1.91, -1.56) Placebo: -1.04	No The percentage change in HbA1c is very similar between sources, with an identical test of

percentage	SE: 0.09 P<0.001	95% CI: (-1.22, -0.86) P<0.001	statistical significance.
NCT00808236 Achieve Return of Spontaneous Circulation (ROSC)	RhinoChill: 33/83 Control: 43/99 No Statistical Analysis Provided.	RhinoChill: 35/93 Control: 43/101 P = 0.48	No The number who achieved ROSC is similar between sources, likely with a similar test of statistical significance.
NCT00894543 Change in Daily Severity of Hot Flashes Between Baseline and Week 8 as Assessed by Prospective Daily Diaries, mean	Escitalopram: -0.53 95% CI: (-0.64, -0.40) Placebo: -0.30 95% CI: (-0.42, -0.17) P<0.001	Escitalopram: -0.52 95% CI: (-0.64, -0.40) Placebo: -0.30 95% CI: (-0.42, -0.17) P<0.001	No The change in daily severity of hot flashes in the escitalopram arm is very similar between sources, and the test of statistical significance is identical.