2-23-2022 1:00 PM

Fate of Registered Studies From London, Ontario

Alex Bi, *The University of Western Ontario*

Supervisor: Martin, Janet, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Epidemiology and Biostatistics

Abstract

Introduction: Lack of study publication leads to bias in the scientific literature. It is important to better understand this phenomenon and find methods for mitigation.

Research Question: How many clinical trials registered on ClinicalTrials.gov in London, Ontario are started, completed, and published?

Methods: Data from all studies in the ClinicalTrials.gov registry associated with London, Ontario were collected, from registry conception until the end of 2017. We determined whether these registered studies were published by July 2020 and whether their first publication included their planned primary outcome at all. Main factors associated with non-publication were assessed using multivariable log-binomial regression. Multivariable modified Poisson regression was used to assess the association between enrollment size and publication. Time to publication was assessed using multiple linear regression.

Results: Of the registered studies (n = 2446), only 38% were published and 30% with their planned primary outcome. Median time to publication post-start was 53 months [IQR: 36, 75]. Factors associated with publication were randomized design, prospective registration, industry funding, drug study, and enrollment size ($p < 0.05$). Factors associated with shorter time to publication were positive results, prospective registration, and industry funding, while drug studies were associated with longer time to publication ($p < 0.05$). Surgical studies seemed to have decreased chances of publication and lengthened time to publication but was not statistically significant in either case.

Conclusions: A substantial proportion of clinical trials from London, Ontario remained unpublished. The factors predictive of non-publication and time to publication suggest potential avenues for increasing publication rates.

Keywords

Publication bias, selective outcome reporting, time to publication, clinical trial registration, ClinicalTrials.gov, research integrity

Summary for Lay Audience

An inherent limitation of scientific literature is that we only know the information that is publicly accessible through publications. Having as much information as possible is important when making decisions, especially when it comes to treating life-threatening illnesses. When studies are not published, the amount of information available is reduced.

Even worse is when certain factors make studies less likely to be published, such as unfavourable results or the type of treatments studied. *Publication bias* occurs when the amount of information available is skewed because some studies are less likely to be published than others. Additionally, there may be selective outcome reporting, where some studies are published with only a subset of their results. These issues are highly prevalent and hinder our ability to make reliable scientific decisions. Unfortunately, there is no way of knowing about every single study that ever existed. The closest thing we have are study registries like ClinicalTrials.gov, where studies are ideally registered before they are conducted or published.

The purpose of our study was to look at registered trials affiliated with London, Ontario, Canada to determine what proportion of these registered studies were ultimately published and to determine factors that may predict a study not being published. We found that fewer than 40% of these studies were published, which suggests that the published literature affiliated with London, Ontario represents fewer than half of all of its registered studies. We were able to quantify several factors associated with publication that could be addressed to increase publication rates in the future.

Acknowledgments

I would like to acknowledge my supervisor, Dr. Janet Martin, for her guidance through this whole project. I would also like to acknowledge Dr. Philip Jones for his help in the conception of our study, Dr. Riaz Qureshi for his insights in writing the manuscript, and Dr. Jessica Moodie for her technical expertise in our literature review.

Table of Contents

List of Tables

List of Figures

List of Appendices

Preface

The main topics of our project were publication and outcome reporting bias. Our goal was to better characterize and address these phenomena within the context of studies conducted in London, Ontario over the past 20 years. This paper aims to provide the justification for our project, the methodology used, and the results of our work. Our hope is that this project can serve as the foundation for further research and development in this area in order to improve research integrity and transparency.

Chapter 1

1 Introduction

Publication bias is an umbrella term for the impact of factors on publishing study results. However, the context and implications of this may not be well understood. This first chapter aims to provide an overview of the relevant concepts, the philosophical basis for preventing publication bias, and the current state of scientific research.

1.1 Key Concepts

Publication bias is defined as when "published literature is systematically unrepresentative of the population of completed studies" (Rothstein et al., 2006). In simpler terms, it is the failure to publish the results of a study due to internal or external factors (DeVito & Goldacre, 2019). This limits the readily available information in the scientific literature and can cripple the clinical decision-making of those that rely on that information (Chan et al., 2008). With as many as half of all clinical studies remaining unpublished (Song et al., 2010), it is imperative that we better understand publication bias to mitigate its adverse impacts. Although this issue is not limited to clinical interventions or specific study types, clinical studies will be the focus of this paper.

"Publication" refers to the public release of the results of a study in the form of a formal study report, which can include but is not limited to a manuscript, report, or conference abstract. They can be published in scientific journals (Blümle et al., 2014), conference proceedings, and other online platforms (Ganga & Edupuganti, 2017). Since it would be infeasible to identify every study that may have ever been conceptualized or conducted, metaresearchers rely on existing databases that record information on studies before they are conducted or published, such as public registries, ethics review submissions (Blümle et al., 2014), and clinical study repositories (Doshi et al., 2013). "Bias" implies a systematic skew in a certain direction due to an identifiable cause, which differentiates it from non-publication alone in that the latter looks at failure to publish without an inherent focus on why (Al-Durra et al., 2018; Jones et al., 2013). Studies of publication biases aim

to assess the factors associated with the non-publication of studies, which may include the direction and strength of the results (Dickersin & Min, 1993).

Outcome reporting bias is a subset of publication bias in that it is the non-publication of certain results within a given study (Dwan et al., 2008; Goldacre et al., 2019). That is, certain outcomes assessed in a study may be present in a publication while others are omitted or demoted to less important objectives. The potential causes of selective outcome reporting may be similar to those of publication bias as a whole (Song et al., 2010). However, it can be harder to detect because the selectively omitted results would not be obvious from the publication alone. One would have to obtain prior information on the planned outcomes for a given study, which may not be readily available, and compare them to the published outcomes to find such discrepancies (Dwan et al., 2008; Goldacre et al., 2019). To that end, registries can protect against selective reporting as studies must register their outcomes of interest and report any changes over the course of the study. This increases transparency, as any discrepancies would become apparent when comparing registry data to corresponding publications. Nevertheless, selective outcome reporting can lead to bias in the published literature and may be considered just as detrimental as the non-publication of entire studies, while being potentially more elusive.

1.2 Ethics of Publication Bias

Strech (2012) argues that the prevention of publication bias is an ethical obligation because it prevents potentially valuable information from being accessible. Not only does this skew the published literature, but the inclusion of unpublished results also has the potential to change the consensus on the effectiveness of treatments (Turner et al., 2008). The misrepresentation of treatment effects that results from keeping results unpublished can harm patients, waste valuable resources, and inadvertently exploit study participants (Strech, 2012).

The most obvious consequences of publication bias are the negative effects on patient health, especially if related to the nature of the study results. Published studies tend to report larger effect sizes and overestimate the benefits of a given treatment compared to unpublished studies (Moreno et al., 2009; Song et al., 2010). Published literature also

tends to underestimate harms (Dickersin & Chalmers, 2011) while unpublished literature provides much more information on adverse effects (Wieseler et al., 2013). Indeed, Turner et al. (2008) found that the inclusion of unpublished literature could dramatically shift or even reverse the original outlook. In their systematic review of various antidepressant trials approved by the Food and Drug Administration (FDA), they found that the overwhelming majority of resulting publications were positive. However, including unpublished FDA reports revealed almost as many negative results as positive, which seriously called into question the effectiveness of these antidepressants (Turner et al., 2008). These findings demonstrated that the published literature can fail to adequately encapsulate all the relevant data for a given treatment and can be heavily skewed.

Another issue is that publication bias contributes to the improper allocation of resources. A substantial amount of time and funding goes into conducting research so if the results of that work are not published, that money and effort goes to waste. Strand et al. (2017) estimates that roughly 50% of research funding goes into studies that remain unpublished. With an estimated 85% of research funding going to waste overall (Chalmers & Glasziou, 2009), non-publication is a waste of resources that should not be overlooked. Even poorly conducted or negative studies are beneficial when published because they can reduce redundant research and prevent other researchers from committing the same methodological errors (DeVito & Goldacre, 2019; Doshi et al., 2013; Song et al., 2010). Withholding data can also lead to ineffective or even dangerous medicines being purchased and funded, potentially wasting billions of dollars a year (Juni, 2002; Thaler et al., 2015). Therefore, failing to publish results is unethical and wastes limited resources that could have been used elsewhere.

Not publishing results is also an injustice to the people who agreed to participate in the trial. Research participants consent to give their time and energy towards a study, while often giving up their privacy, with the expectation that the results will benefit the rest of society if not themselves (Jones et al., 2013). Clinical interventions also pose some risk to the participants due to potential adverse effects and the forgoing of potentially better alternative treatments (Jones et al., 2013). Unfortunately, a large amount of patient data never sees the light of day (Jones et al., 2013; Kirkham et al., 2016). Failing to publish

this data is exploitative because it puts patients at risk and wastes their efforts while giving nothing back. Furthermore, researchers and participants alike may not know their study is redundant if past research is not published, which would needlessly put even more participants at risk. To put this into perspective, a meta-analysis of intravenous streptokinase trials prior to 1973 (inclusive) found that it significantly reduced mortality from acute infarctions. Even though adequate data already existed, 25 subsequent trials were conducted from 1973 to 1988, enrolling almost 35,000 additional patients (Lau et al., 2010). These subsequent trials were mostly redundant and likely deprived thousands of control patients from an already proven treatment.

1.3 Real-World Implications

An important example of the patient harm that can result from publication bias is the formerly FDA-approved rosiglitazone, a drug developed by GlaxoSmithKline (GSK) to reduce blood sugar in people with diabetes. It was speculated that the original approval in 1999 was based on limited publicly available data that was not adequately powered for cardiovascular events (Nissen & Wolski, 2007). After including both new and unpublished data from the FDA and GSK registries in their meta-analysis, Nissen & Wolski (2007) found that rosiglitazone significantly increased the odds of myocardial infarction by 43% (odds ratio [OR] 1.43 [95% confidence interval [CI]: 1.03, 1.98]) compared to alternative diabetic treatments. The increased risk of cardiovascular events eventually led the FDA to restrict this drug in September 2010 (FDA, 2018), but not before being used for over a decade. It is estimated to have caused 431 deaths for every 100,000 patients compared to safer alternatives (Loke et al., 2011), as well as 6000 to 8000 additional myocardial infarctions among US and UK patients in 2010 alone (Chan et al., 2014). Had the unpublished data been publicly available beforehand, the use of rosiglitazone in lieu of safer alternatives could have been prevented.

While outcome reporting bias may be more nuanced, its implications are no less important. One such case was celecoxib, an FDA-approved non-steroidal anti-inflammatory drug (NSAID) for treatment of arthritis. Both the drug trial and subsequent publication were funded by its manufacturer, Pharmacia (Juni, 2002; Silverstein et al., 2000). In their CLASS trial publication, Silverstein et al. (2000) originally concluded that

celecoxib significantly reduced annual incidence of gastrointestinal ulcer complications compared to other NSAIDs (0.44% vs. 1.27%, $p = 0.04$) during 6 months of treatment. What they neglected to mention was that the trial actually lasted 12 to 15 months. In fact, 15-month follow-up data from FDA reports showed it was barely better than the comparator ibuprofen (22.4% vs. 23%) (McCormack & Rangno, 2002). Furthermore, the majority of ulcer complications in the latter half of the study were from the celecoxib group (Juni, 2002). This supports that celecoxib does not actually cause fewer adverse effects compared to conventional NSAIDs, only delayed. Unfortunately, the omission of the 12- and 15-month data from the original publication led to a misleading perception of celecoxib as being safer. Despite its failure to demonstrate superiority long-term, it had over 14 million prescriptions in 2004 from the US alone, instead of cheaper alternatives (Chan et al., 2014). Even when studies are published, the selective reporting of results can overestimate treatment benefits and warp our perception of them.

1.4 Shortcomings of Research Infrastructure

For the purposes of our project, we used the ClinicalTrials.gov registry as our reference for planned clinical trials, which was created by the Food and Drug Administration Modernization Act (FDAMA) in 1997 and was made available to the public in 2000 (NLM, n.d.). The act required that all clinical trials in the United States of America be registered in the database at least before publication (Viteri-García et al., 2018), providing an effective record of virtually all planned clinical trials in the US. The later Food and Drug Administration Amendments Act (FDAAA) of 2007 required that clinical trials also report their results within one year of completion (Prayle et al., 2012). Additionally, the International Committee of Medical Journal Editors (ICMJE) began requiring prospective trial registration for journal publications in 2005 (De Angelis et al., 2004), which likely further encouraged researchers to register worldwide.

While Health Canada encourages trial sponsors to register in public registries like ClinicalTrials.gov (Canada.ca, 2016) and it remains a popular option among Canadian institutions, registration is not legally required as it is in the US (ClinRegs, 2021). This may limit the scope of the registry with regard to studies from Canadian research institutions. Indeed, many have called for Health Canada to match the FDA on this

particular legal requirement (Shuchman, 2013). Nevertheless, ClinicalTrials.gov remains a feasible and useful source of data on Canadian studies as Health Canada does not provide a comprehensive and comparable alternative (Canada.ca, 2016).

While one may assume that legally mandating registration and reporting results may increase accountability, some findings suggest that may not be the case. Although prospective registration should substantially decrease reporting bias in theory, Thaler et al. (2015) found it was not effective in preventing outcome discrepancies. Even more worrisome is that legal requirements themselves are not useful if they are not enforced. Despite the threat of fines by the FDA, Prayle et al. (2012) found that 78% of clinical trials still failed to report their results within one year of completion. To further shed light on this non-compliance, DeVito et al. (2018) developed the online FDAAA TrialsTracker, which tracks registered studies on ClinicalTrials.gov and records whether they have reported their results after study completion. To date, the tracker has revealed a substantial number of unreported results and unenforced fines. However, one of the drawbacks of such a generalized tool is that it is severely limited in depth. Taken at face value, a search of the prominent research institutions based in London, Ontario (i.e., Western University and Lawson Health Research Institute) would suggest that 80% of their clinical trials fail to report their results (as of July 31, 2020). While this may seem like a staggering proportion, the reality is that the tracker only found a total of five completed studies from those institutions (EBM DataLab, 2018), which seems implausible and makes the estimate largely unreliable. Although the tracker is useful for providing broad general statistics, its use in addressing specific regions is limited. This highlights the continued need for in-depth research on publication that is smaller in scope to allow for greater focus.

Even if study results are submitted to the FDA, they are not readily accessible in the scientific literature, let alone to the general public. Lack of time and interest remain commonly cited personal reasons among researchers for failing to publish (Timmer et al., 2002). That being the case, a potential remedy would be for study data to be published by other willing researchers. This is what the *Restoring Invisible and Abandoned Trials* (RIAT) initiative attempts to achieve, by publishing the results of previously unpublished

industry clinical study reports on behalf of the original researchers (Doshi et al., 2013). These reports are rarely freely accessible and may only be granted upon request to the companies or their investigators (Chan et al., 2014). Thus, the RIAT initiative aims to both increase the availability of such data and expose the vast registry of industry trials that would have otherwise remained hidden. Although the necessity is clear, it remains to be seen if such practises will become commonplace.

A potential misconception regarding failure to publish is that it is due to rejection by scientific journals. On the contrary, negative or non-significant results do not seem to prevent eventual journal acceptance (Song et al., 2010). In fact, Timmer et al. (2002) estimates that 76% of unpublished studies were never submitted to journals in the first place. Among the remaining that were rejected, half were due to improper methodology or uninteresting topics. Even the few studies that were initially rejected due to negative results eventually reached publication in other journals (Timmer et al., 2002). Furthermore, several journals exist that actively publish negative findings (Amsen, 2015; Horsley et al., 2015). Thus, it is clear that the main predictors of non-publication occur well before attempting to submit the write-up and cannot be solely attributed to the journals (Song et al., 2010; Thaler et al., 2015).

However, that is not to say that the current framework is without fault. While the Consolidated Standards of Reporting Trials (CONSORT) statement offers a standard for publication manuscripts to follow, journals are not forced to follow it (Goldacre et al., 2019) and many other areas lack such rigorous guidelines. Among systematic reviews, Potthast et al. (2014) found that almost 90% failed to include industry registries, suggesting inadequate methodology for including grey and unpublished literature. Sources that document the planning of a study, such as registries, protocols, and ethics review databases, often lack consistency and detail (Chan et al., 2014). In particular, public registry entries are often limited and contain far less information on study outcomes than clinical study reports from industry trials (Dwan et al., 2008; Wieseler et al., 2013). Even if one has access to these reports, they are notoriously long and difficult to navigate, with some spanning over 600 pages (Doshi & Jefferson, 2013). The lack of

standard guidelines in these crucial areas reduces the feasibility of these data sources, especially when it comes to detecting non-publication and outcome discrepancies.

1.5 Summary

Publication bias and selective outcome reporting skew the published literature and prevent an accurate representation of treatment outcomes. Non-publication is unethical because it wastes resources and can lead to patient harm, whether by impairing clinical decision-making or needlessly putting patients at risk. Regulations mandating publication are not sufficiently enforced and there lacks rigorous guidelines for grey literature. Fortunately, registries provide a useful tool for detecting non-publication and outcome discrepancies. However, that alone may not be enough to understand the nuances of publication bias and how to prevent it.

Chapter 2

2 Literature Review

Before beginning our study, we conducted a review of the relevant literature to assess what has already been done and inform our methodology. The following chapter will outline our review strategy and discuss our findings. Our focus was on publication bias, outcome reporting bias, and potential risk factors. A few studies had data on time to publication, which were included as well.

2.1 Review Methodology

The main topic of our review was the proportion of baseline studies that reached publication and what characteristics were associated with them. Any measures of discrepancies between planned and published outcomes, as well as time to publication analysis, were also reviewed if included in the publication. The purpose of our review was to explore the existing literature for relevant study data and methodology. This would give us a reference for our publication statistics as well as suggestions for potential covariates in our regression models. We also assessed the existing methodology for potential weaknesses that could be addressed in our study.

Following an exploratory search, a list of key terms was developed to represent our topic of interest: publication bias, outcome reporting bias, registries, ethics submissions, protocols, and RCTs. These terms were used to develop our systematic search strategy in the Ovid search engine (Wolters Kluwer, n.d.), which was then tailored to each of our intended databases of published medical literature: Medline and Embase. We completed the search on October 16, 2019, yielding a total of 1007 results after removing duplicates. (Appendix A)

We were interested in primary studies that compared pre-study documentation (i.e., registries, protocols, etc.) to corresponding publications of medical intervention RCTs. Studies that did not collect primary data or that did not analyze both components of this comparison together were not included. Published manuscripts, journal articles, and

abstracts were included. Review articles, presentations, and letters or editorials were excluded. Protocols were also excluded unless they included preliminary or final results.

A single reviewer (A. B.) conducted two stages of screening, using Mendeley to manage citations. In the first stage, the publication abstracts were reviewed for key terms, and clearly irrelevant results were excluded. Manual forward and backward citation searching, of included articles that strongly matched our key terms, was then conducted to find additional papers that may have been relevant. In the second stage, publication abstracts were reviewed for our inclusion criteria and those that met them were added to our final inclusion group. In cases where the abstract alone was inconclusive, the full text was reviewed. Screening was completed on December 14, 2019.

For our literature review, the full text of each of our final included studies was reviewed for relevant results. Data on publication rates, prevalence of selective outcome reporting, and time to publication intervals were extracted. Publication measures were converted to proportion published and time to publication measures were recorded as months. The statistical significance and direction of any comparative statistics for potential predictors of publication and time to publication were recorded. Acceptable comparative statistics included odds ratios (OR), relative risk (RR), hazard ratios (HR), and absolute differences. If these statistics were not present but the relevant data was available, they were manually calculated. Data sources and sample sizes were also recorded. Table 1 and Table 2 present an overview of our findings for publication and time to publication, respectively. Due to lack of consistency in outcome discrepancy analyses, they were not included in these tables and will instead be addressed in text.

Although the measures of proportion published and time to publication should theoretically have been comparable, we decided it was not reasonable to combine them via meta-analysis because the underlying studies differed greatly in inclusion criteria and baseline data sources. While the pre-existing results can serve as reference, they may not directly inform future studies in other contexts.

2.2 Types of Studies

A total of 15 primary studies were included in our literature review, which had a variety of retrospective data sources being used as their baseline reference for assessing publication bias (Table 1). Although their sources varied, each study followed an analogous structure; a pre-existing catalogue of planned or completed studies was used to form a retrospective cohort, followed by a search for corresponding publications.

Five of our included studies used public registries as their baseline, specifically ClinicalTrials.gov (Al-Durra et al., 2018; Chahal et al., 2012; Jones et al., 2013; Khan et al., 2014; Pica & Bourgeois, 2016). Seven of them looked at protocol submissions to various ethics review boards (Blümle et al., 2008; Blümle et al., 2014; Kasenda et al., 2014; Kirkham et al., 2016; Menzel et al., 2007; Rosenthal et al., 2015; von Elm et al., 2008). Two studies delved into industry registries, specifically that of GSK (Pang & Loke, 2011; Tfelt-Hansen, 2009). One analyzed government-funded studies, specifically those funded by the National Institutes of Health (NIH) (Dickersin & Min, 1993).

2.3 Proportion Published

None of the studies included in our literature review reported that 100% of their baseline studies were published, suggesting study loss is broadly prevalent in the scientific literature. Proportion published ranged from virtually none to almost all of their respective cohorts. These extremes likely represent vastly different methodology and timing. Dickersin & Min (1993) reported 93% of their cohort of studies was published. However, their analysis of NIH-funded studies had a publication window of over 9 years, allowing substantially more time for studies to reach publication than other similar studies, and excluded discontinued trials. In contrast, Tfelt-Hansen (2009) found that virtually none of their industry registry studies were published. While industry studies may indeed have questionable rates of publication (Doshi et al., 2013), a likely alternative explanation in this case would be the sample size of only six manually-found studies. Publication bias studies with large sample sizes and unrestrictive inclusion criteria (i.e., public registries and ethics submissions) tended to report fairly similar results to one another, ranging from 50% to 70% eventually being published (Table 1). While still a

wide range, we can roughly estimate from the existing literature that over half of all medical RCTs are published in some form. More restrictive data sources (i.e., industry registries) had comparatively lower rates of publication (Pang & Loke, 2011; Tfelt-Hansen, 2009).

Notably, Al-Durra et al. (2018) used a unique method of calculating proportion published by only including published studies that reported their registered primary outcomes. This created a hybrid measure that incorporated elements of both non-publication and selective outcome reporting, as opposed to the conventional overall proportion published. One could argue that such a measure is more effective at characterizing research loss, as published studies may not report all of their ascertained results.

The most common reason given for failure to publish was lack of time or interest, which made up 43% of responses (Dickersin & Min, 1993). Dickersin & Min (1993) found that an additional 38% of unpublished studies had general operational problems, which coincided with the 40% of discontinued studies citing poor recruitment reported by Kasenda et al. (2014). Incomplete analysis made up the remainder of the reasons (14.3%), with the rest being unknown (5.4%) (Dickersin & Min, 1993).

2.4 Selective Outcome Reporting

Among the publication studies that assessed outcome discrepancies, prevalence of selective outcome reporting was found to be widespread. Perhaps due to its nuanced nature, a variety of different classifications were used to characterize its prevalence. These classifications were generally not uniform across studies.

Overall, 29–35% of registered studies had a discrepancy in their primary outcomes (Chahal et al., 2012; Pang & Loke, 2011), while 64% had a discrepancy in their secondary outcomes (Pang & Loke, 2011) and as high as 80% had at least one discrepancy in either (Chahal et al., 2012). Al-Durra et al. (2018) estimated that of published trials, only 2% outright did not report their primary outcomes at all. This suggests that the vast majority of published studies report their primary outcomes in some form, even if not as the originally intended primary outcome. However, Blümle et al.

(2008) estimated that 43% of trial protocols fail to prespecify a primary outcome at all. This makes assessing discrepancies challenging in such cases and may reflect a lack of rigorous oversight in protocol submissions.

Some studies focussed on the publication of prespecified outcomes themselves. Kirkham et al. (2016) found that only 70% of all planned outcomes from published trials are published in full. When looking at patient outcomes, this dropped to 43%. Industry funding seemed to coincide with a decreased likelihood of fully reporting all outcomes, though not significantly (RR 0.64 [95% CI: 0.30, 1.38]) (Kirkham et al., 2016). Likewise, Pang & Loke (2011) found that industry trials tended to underreport outcomes related to treatment side effects, with 88% of all adverse effects going unreported. Furthermore, 93% of "serious" and 85% of "fatal" side effects were not reported in publications (Pang & Loke, 2011). This is especially worrisome when "harm" outcomes should actually increase odds of publication compared to efficacy outcomes (OR 3.57 [1.09, 11.11]), at least for surgical studies (Rosenthal et al., 2015). Whether the lack of full reporting reflects differing tendencies between types of interventions or a systemic problem among industry-conducted trials remains uncertain.

2.5 Time to Publication Interval

For studies registered on ClinicalTrials.gov, median time to publication following study completion ranged from 27 to 38 months (Jones et al., 2013; Khan et al., 2014). Jones et al. (2013) found the proportion published to be highest between 24 and 48 months, followed by a plateau thereafter. Al-Durra et al. (2018) also found that about half of their studies were published within 48–72 months. Pica & Bourgeois (2016) reported a mean of 29 months following primary completion. Overall, existing analyses suggest that about half of all studies are published within 4 years following conception and within 2–3 years following completion.

Unpublished trials that posted results to ClinicalTrials.gov did so after a median time of 22 months post-completion (Jones et al., 2013). For ethics submissions, Menzel et al. (2007) measured time to publication from the date of approval and reported a median time of 46 months. In addition, Khan et al. (2014) found that studies that published post-

2006 tended to do so sooner than those that published in 2005 and earlier, suggesting that time to publication may be decreasing over time.

2.6 Characteristics Associated With Publication

Registering or submitting a protocol prior to conducting a study is intended to increase transparency and encourage researchers to adhere to prespecified study plans. Having a prior record can hold researchers accountable for changes to the methodology made after the study has begun. Despite this reasoning, prospective registration was not significantly associated with increased publication in the existing literature. In fact, Al-Durra et al. (2018) and Khan et al. (2014) found that prospective registration may coincide with a decrease in odds of publication, though not statistically significant (OR 0.75 [0.44, 1.28] and 0.4 [0.1, 1.0], respectively). Prospective registration was also not significantly associated with time to publication (HR 1.2 [0.8, 1.8]) (Khan et al., 2014).

Funding source was one of the most commonly studied and controversial predictors. Industry sponsors may affect the conduct of a trial, even though the true extent of their influence is often omitted from the publication (Lundh et al., 2012). In fact, some sponsorship agreements have revealed that sponsors can review, terminate, and publish a study without needing the original authors' consent (Lundh et al., 2012). However, whether this translates to reducing publication rates is not clear. While some studies suggest industry funding is associated with decreased odds of publication, others did not find such an association. On the one hand, Jones et al. (2013), Kasenda et al. (2014), and Rosenthal et al. (2015) each reported a significant decrease in odds of publication associated with industry funding, relative to other funding sources (OR 0.48 [0.31, 0.74], 0.60 [0.43, 0.83], and 0.33 [0.12, 0.95], respectively). Data from von Elm et al. (2008) revealed comparable results (crude OR 0.55 [0.35, 0.86]) and Pica & Bourgeois (2016) reported an OR of 0.45 [0.27, 0.74] at 24 months after trial completion. On the other hand, Al-Durra et al. (2018), Blümle et al. (2008), Blümle et al. (2014), and Chahal et al. (2012) did not find a significant association (OR 0.75 [0.44, 1.28], 0.89 [0.42, 1.87], 0.92 [0.61, 1.37], and 0.28 [0.04, 1.79], respectively), along with Kirkham et al. (2016) (RR 1.20 [0.86, 1.67]). Reviewing the association between industry funding and time to publication revealed similar insights. Pica & Bourgeois (2016) found a significantly

lengthened time to publication associated with industry funding (33 vs. 24 months, $p < 0.001$), but Khan et al. (2014) did not find an association (HR 1.0 [0.6, 1.5]). Thus, there was no clear consensus on the relationship between funding source and likelihood of publication or time to publication. Heterogeneity and lack of power among existing studies may have prevented definitive conclusions.

Only one of the studies in our review assessed the type of intervention. Specifically, Rosenthal et al. (2015) did not find a significant association between surgical studies and odds of publication, compared to non-surgical studies (OR 0.76 [0.50, 1.13]).

Statistically significant results, irrespective of direction, was one of the most consistent predictors for publication, likely because researchers may be more motivated to publish such results that are deemed interesting. Both Dickersin & Min (1993) and Khan et al. (2014) reported that statistically significant results were associated with dramatically increased odds of publication (OR 12.3 [2.54, 60] and 4.3 [1.8, 10.2], respectively). Likewise, Tfelt-Hansen (2009) suggested that industry studies that failed to show adequate treatment benefits or superiority to alternatives were less likely to be published. Statistically significant results were also associated with significantly shorter time to publication (HR 1.9 [1.2, 2.8]) (Khan et al., 2014). Thus, statistical significance of results seems to be positively associated with both likelihood and speed of publication.

Study enrollment size was also positively associated with publication. Larger studies may have more statistical power, resources, and renown, increasing the probability of publication. Unfortunately, there was little replicable consistency in how its association was measured and some thresholds were seemingly arbitrary. Of the seven studies that analyzed the association between sample size and publication, three of them dichotomized size by the median (Blümle et al., 2008; Blümle et al., 2014; von Elm et al., 2008). Both Blümle et al. (2008) and Blümle et al. (2014) had a median of 120, with the latter reporting a significant association (Chi^2 test of independence, $p = 0.01$). With a median of 236, von Elm et al. (2008) also found that size was associated with an increase in odds of publication (OR 2.04 [1.23, 3.39]). One study dichotomized using a threshold of 100 participants (Dickersin & Min, 1993). The remaining three studies divided size

into three or more quartiles, increments, or categories, with all of them reporting a positive association between size and publication (Al-Durra et al., 2018; Kasenda et al., 2014; Pica & Bourgeois, 2016). Both Al-Durra et al. (2018) and Pica & Bourgeois (2016) used Chi^2 tests ($p = 0.001$ and $p < 0.001$, respectively). Kasenda et al., (2014) reported a small but significant increase with an OR of 1.05 [1.03, 1.09] per 100 increase in size. Only two of the seven studies did not find an association (Blümle et al., 2008; Dickersin & Min, 1993), likely due to a lack of statistical power in both cases. Nevertheless, none of the studies suggested a decrease in publication associated with enrollment size.

Early trial discontinuation was another factor evaluated in studies of publication bias. Intuitively, studies that failed to reach completion are less likely to reach publication as they may not have sufficient results to publish. Kasenda et al. (2014) and Rosenthal et al. (2015) estimated that discontinuation is associated with decreased odds of publication by as much as 70–75% (OR 0.31 [0.23, 0.44] and 0.24 [0.08, 0.69], respectively). With 16–25% of trials not reaching completion (Blümle et al., 2014; Kasenda et al., 2014; Pica & Bourgeois, 2016), trial discontinuation could be a major influence on failure to reach publication. Characteristics such as enrollment size (Pica & Bourgeois, 2016) and harm outcomes (Rosenthal et al., 2015) did not seem to be significantly associated with likelihood of discontinuation. While Rosenthal et al. (2015) did not find an association between industry funding and trial discontinuation (0.60 [0.20, 1.85]), Pica & Bourgeois (2016) found that industry funding was associated with a significant decrease in odds of discontinuation (0.46 [0.27, 0.77]), which somewhat contradicts its negative influence on publication. One interpretation could be that industry sponsorship and associated infrastructure may increase impetus and accountability for researchers to complete trials but not necessarily to publish them. Although discontinuation is not a focus of our study, reviewing its prevalence and predictors may still be relevant to our understanding of non-publication as a whole.

2.7 Knowledge Gap

Our literature review revealed some shortcomings in existing studies that we hoped to address. A prominent issue was the general lack of sufficient sample size to adequately explore associations between study characteristics and publication. Most studies did not

use multivariable regression analysis, which would have improved effect estimates and mitigated confounding. Another issue was the lack of consistency in analyzing enrollment size, particularly as a dichotomous or categorical variable. The difficulty in deciding on a consistent method of categorization suggests that size should be kept as a continuous variable instead, which could produce more valid results in analysis.

The majority of existing studies calculated a crude proportion published to quantify non-publication. Determining the rate of publications that include their prespecified primary outcome (Al-Durra et al., 2018) may be a more effective measure, as it considers selectively reported primary outcomes. Most studies also did not consider potential changes in publication measures and study characteristics over time, even though it would be unreasonable to assume them to be static. Time to publication may be another important factor in publication bias because studies may vary in how long it takes them to publish. However, analysis of potentially associated characteristics for publication time is scarce. Studies published sooner may also have greater impact than those published later, which makes shortening publication times beneficial. Plus, none of the existing studies measured time to publication from the registry start date or properly calculated effect estimates in terms of absolute difference in time. Furthermore, the quality of the protocols and registry entries themselves were largely ignored. Not only would poor outcome descriptions make it difficult to find discrepancies, but better documentation may also indicate better planning, which could increase likelihood of publication.

The heterogeneity in publication rates suggests that likelihood of publication and associated factors may vary widely, especially by region and data source. To ensure the applicability of our results, we determined we needed to quantify publication bias specific to research conducted in London, Ontario. To the best of our knowledge, no such study on publication bias has been conducted. While a broad tool like the FDAAA TrialsTracker may be useful as general reference, we have noted that it is lacking in sensitivity for this region. The fact that it suggests a mere 20% reporting rate (EBM DataLab, 2018) is all the more reason a proper in-depth study is necessary to affirm or correct those findings. To that end, data from a public registry would likely allow us to find a greater number and a broader range of relevant studies from London, Ontario.

Table 1

Summary of literature review results for publication rate

Study	Data Source	Sample Size	Proportion	Predictors
Al-Durra et al., 2018	ClinicalTrials.gov	556	0.73	size*, industry, prospective
Blümle et al., 2008	Ethics protocols	299	0.48	industry
Blümle et al., 2014	Ethics protocols	917	0.52	size*, industry
Chahal et al., 2012	ClinicalTrials.gov	34	0.59	industry
Dickersin & Min, 1993	NIH-funded trials	198	0.93	positive*, size, industry
Jones et al., 2013	ClinicalTrials.gov	585	0.71	industry†
Kasenda et al., 2014	Ethics protocols	1017	0.56	size*, industry†, discontinue†
Khan et al., 2014	ClinicalTrials.gov	143	0.64	positive*, prospective
Kirkham et al., 2016	Ethics protocols	308	0.54	industry
Menzel et al., 2007	Ethics protocols	99	0.7	
Pang & Loke, 2011	GSK registry	54	0.47	
Pica & Bourgeois, 2016	ClinicalTrials.gov	559	0.57	size*, industry†

Rosenthal et al., 2015	Ethics protocols	863	0.66	industry†, discontinue†, surgical
Tfelt-Hansen, 2009	GSK registry	6	0	negative
von Elm et al., 2008	Ethics protocols	451	0.52	size*, industry†

Note. This table lists the studies included in our literature review as well as their baseline data sources, sample sizes, and general findings. Parameter estimates and evaluated predictors for proportion published are shown.

*Significantly increases publication rate ($p < 0.05$).

†Significantly decreases publication rate ($p < 0.05$).

Table 2

Summary of literature review results for time to publication

Study	Time (Months)	Predictors
Al-Durra et al., 2018	post-start median 48–72	
Jones et al., 2013	post-completion median 27	
Khan et al., 2014	post-completion median 38	positive*, industry, prospective
Menzel et al., 2007	post-approval median 46	
Pica & Bourgeois, 2016	post-completion mean 29	industry†

Note. This table lists the parameter estimates and evaluated predictors for time to publication from the studies included in our literature review.

*Significantly shortens time to publication ($p < 0.05$).

†Significantly lengthens time to publication ($p < 0.05$).

Chapter 3

3 Visuals for Temporality of Publication

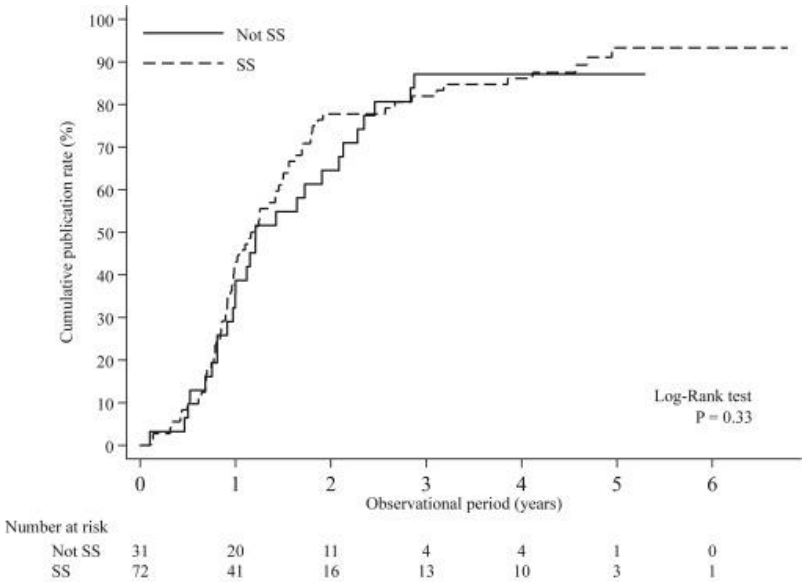
To the best of our knowledge, graphing publication temporally has not been commonplace when studying publication bias. Such graphs could show variations in publication rates over time and supplement conventional proportion statistics. In particular, a temporal graph would allow the observation of influences from past landmark events on publication rates. While some studies have touched on the concept, few have done so in the capacity we have planned. This chapter aims to summarize the existing work that inspired and helped develop our graphics.

3.1 Publication Rates Over Time

The concept of graphing the proportion of published studies across time as a variable has been done in other contexts. Tsujimoto et al. (2019) constructed a cumulative incidence curve for publication rate of systematic reviews over a period of roughly 6 years. This visual approach revealed an almost logarithmic curve with a steep initial incline around 1 year, which coincided with a median time to publication of 1.2 years (Figure 1) (Tsujimoto et al., 2019). Strand et al. (2017) employed a similar idea, but instead constructed a Kaplan-Meier survival curve and focussed on clinical trials (Figure 2). With time of funding as their starting point and publication of main results as their survival criteria, their graph was in stark contrast to that of Tsujimoto et al. (2019) and revealed an exponential curve with a steep incline much later, which corresponded with a longer median time to publication of 7.1 years (Strand et al., 2017). Comparing these findings demonstrates an observable difference between the publication trends of systematic reviews and clinical trials, with the bulk of systematic reviews being published much sooner.

Figure 1

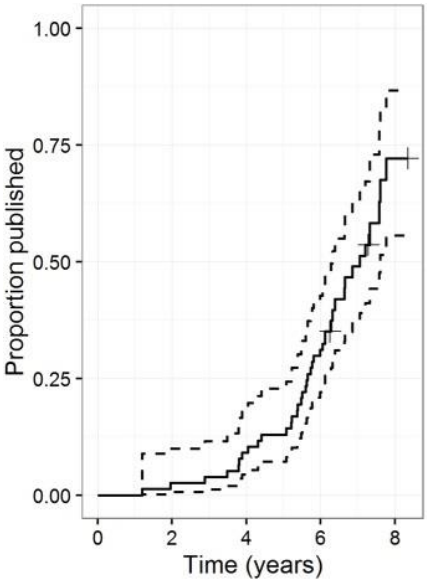
Cumulative publication rate of systematic reviews with statistically significant (SS) results (dashed line) and those without SS results (solid line)



Note. Cumulative incidence curves for publication of systematic reviews over 6 years. From "Statistical significance did not affect time to publication in non-Cochrane systematic reviews: A metaepidemiological study," by Y. Tsujimoto, Y. Tsutsumi, Y. Kataoka, H. Tsujimoto, Y. Yamamoto, D. Papola, G. H. Guyatt, S. Fukuhara, and T. A. Furukawa, 2019, *Journal of Clinical Epidemiology*, *115*, p. 25–34 (https://doi.org/10.1016/j.jclinepi.2019.06.015). Copyright © 2019 by Elsevier Inc. Reprinted with permission.

Figure 2

Kaplan-Meier survival curve (and 95% CIs as dashed lines) for time from funding to main paper for all papers combined



Note. Kaplan-Meier survival curves for publication of clinical trials over 8 years. From "Time to publication for publicly funded clinical trials in Australia: An observational study," by L. B. Strand, P. Clarke, N. Graves, and A. G. Barnett, 2017, *BMJ Open*, *7*(3), p. e012212 (https://doi.org/10.1136/bmjopen-2016-012212). Copyright © 2017 by BMJ Publishing Group Ltd. Reprinted with permission.

A plausible explanation for the difference in median time to publications would be the difference in amount of time and work necessary to carry out the two types of studies. However, if that were the sole reason, one would expect a simple horizontal translation going from one graph to the other. The clear differences in overall trends cannot be so easily explained and would not have been as apparent with the conventional statistics alone. Thus, the ability to document and compare publication trends shows the importance of visual representation and temporality in the analysis of publication rates.

Both of the temporal line graphs described above (Figure 1, Figure 2) used standardized timescales of relative years, showed the cumulative proportion of studies published instead of frequency, and had relatively narrow time ranges of 6 to 8 years. While these design decisions adequately served their intended purposes, both of these pre-existing

graphs were limited in their ability to display historical trends over longer periods of time. To that end, Blümle et al. (2014) provided an alternative method in constructing a bar graph of publication frequency by calendar year. Unfortunately, this graph was also limited and only showed the publications of ethics submissions made in 2000 alone. Thus, our goal was to expand on this concept and create a graph that would both use absolute time in years and span a much longer timeframe.

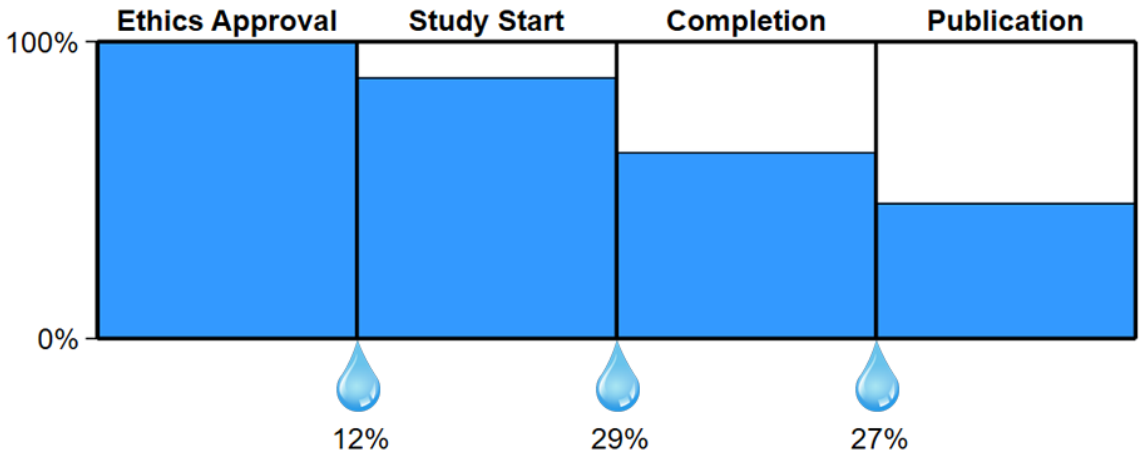
3.2 Leaky Buckets and Icon Arrays

A leaky bucket diagram is a graph that shows losses across stages of a sequential process, with each stage being represented by a rectangular "bucket" (Glasziou, 2005). Such a graph may be useful in portraying the losses in studies on their way to publication, where the "water" dripping from each bucket represents the proportion of studies that failed to reach the next stage. While a standard bar graph may accomplish the same task, the leaky bucket diagram may better resonate with readers as a relevant analogy due to the implied flow across the diagram and the ability to directly compare proportions between each stage. To the best of our knowledge, using registry data to construct such a diagram has not yet been done.

Blümle et al. (2014) was a key inspiration because they used metadata from a database of study proposals submitted for ethics review, which was comparable to information in public registries. Their results included the number of studies that were approved, started, completed, and published. While these statistics were not explicitly portrayed as stages in the research process, we believed they could be re-interpreted as such and have adapted them into a leaky bucket diagram as proof of concept (Figure 3).

Figure 3

Leaky bucket diagram

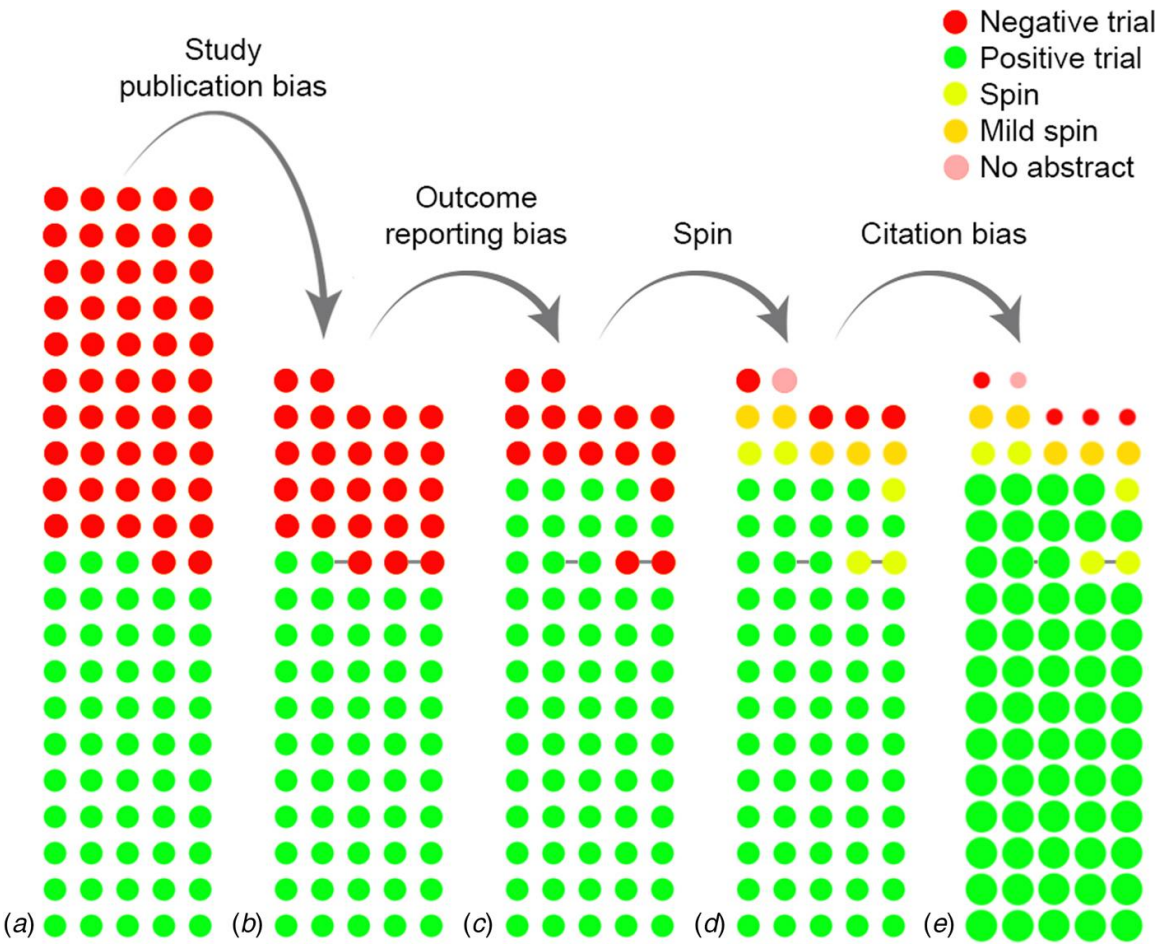


Note. Leaky bucket diagram showing the proportion of studies that progressed to each stage of the research process. Percentages along the bottom indicate relative proportion of studies lost from the previous stage. Data adapted from "Fate of clinical research studies after ethical approval – follow-up of study protocols until publication," by A. Blümle, J. J. Meerpohl, M. Schumacher, and E. von Elm, 2014, *PLoS ONE*, *9*(2), p. e87184 (https://doi.org/10.1371/journal.pone.0087184). CC BY 4.0.

Unfortunately, an inherent weakness of a leaky bucket diagram was its lack of ability to convey multiple values per stage without becoming too complicated, since a single value is used to summarize the differences between stages. Tsujimoto et al. (2019) showed the effectiveness of conveying multiple publication values in the same graph, specifically the proportion of significant and non-significant studies published. This suggested another type of graph may be necessary: the icon array (Figure 4) (de Vries et al., 2018). The icons in the array represent chunks of equal size and are grouped together to convey the number of entities per stage, similar to a leaky bucket diagram. The benefit of icon arrays is the ability to colour code these icons, which allows for a clear and convenient method of describing the distribution of any number of characteristics within a given stage. For the purposes of our study, we planned to create an icon array that showed the distribution of positive and negative studies across each stage of the progression from study registration to publication.

Figure 4

The cumulative impact of reporting and citation biases on the evidence base for antidepressants



Note. Icon array using circles to represent the distribution of studies across each category of bias. From "The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: The case of depression," by Y. A. de Vries, A. M. Roest, P. de Jonge, P. Cuijpers, M. R. Munafò, and J. A. Bastiaansen, 2018, *Psychological Medicine*, *48*(15), p. 2453–2455 (https://doi.org/10.1017/S0033291718001873). Copyright © 2018 by Cambridge University Press. Reprinted with permission.

Chapter 4

4 Methods

This chapter describes our retrospective cohort of all studies registered on ClinicalTrials.gov affiliated with London, Ontario. The primary objective of our study was to determine the proportion of registered studies that were ultimately published and what were the associated study characteristics. Secondary objectives were to estimate time to completion and time to publication for studies following study start. Lastly, we created visualizations for changes in publication over time.

4.1 Registered Studies

ClinicalTrials.gov is a global clinical trial registry hosted by the United States National Library of Medicine, within the National Institutes of Health (NLM, n.d.). We selected this registry as the source to identify our cohort of registered studies from London, Ontario since it is the largest clinical trials database and remains the most common avenue for registering studies in North America.

The registry outcome descriptions were assumed to be each study's preconceived and planned outcomes prior to being conducted. These served as reference for determining the quality of protocols and potential discrepancies in outcome reporting. When applicable, only the registry's most recent information at the time of our search was used.

Any interventional studies (i.e., clinical trials) that mentioned "London, Ontario" in their affiliation or location, registered from ClinicalTrials.gov inception to the end of 2017, were eligible for inclusion. Observational and expanded access (treatment of patients who cannot participate in a clinical trial) studies, as indicated in the registry, were excluded. Prior exploratory screening showed that limiting to studies as late as 2017 ensured a reasonable cohort size of over 2000 studies while also allowing more than 2 years for registered studies to reach publication.

4.2 Published Studies

The Ovid search engine (Wolters Kluwer, n.d.) was used to find corresponding publications for all registered studies (regardless of study status), via NCT number, in the following databases: Medline, Embase, PsycInfo, and Western Libraries. These databases were likely to reasonably capture any potential publication in journals or other publication outlets. Our search method assumed that publications would specifically mention their associated NCT numbers and that they would be indexed in each of our searched databases. However, exploratory searches found that NCT numbers were a practical and reliable method of matching any registered study on ClinicalTrials.gov to a corresponding publication if one existed. The cut-off date for primary publications was July 31, 2020, meaning a minimum time for publication window of about 2.6 years from time of registration until publication. Published manuscripts, journal articles, and abstracts were included. Review articles, presentations, and letters or editorials were excluded. Protocols were also excluded unless they included preliminary or final results.

The abstracts of search results were manually screened to determine whether the information in the title, objectives, and outcomes reasonably matched those in the registry entries. Due to the possibility of results being published under different authorship (Doshi, 2013), the names of the researchers were not used in matching. The screening was generally restricted to publication abstracts. However, the full-text was retrieved for screening if a standard abstract including the aforementioned elements was not found. If at least one result matched a corresponding registry entry, the study was considered published. In the case of multiple matching results, only the publication with the earliest date was used. This was to mitigate any advantages from multiple publications, such as having more time on later publications. Only studies that included results were considered as publications. Both descriptive and comparative statistics were accepted as results if indicated as outcomes by the publication's study objectives.

The earliest available publication date was recorded as the study publication date, including electronic publications and publications ahead of print. When the publication date was not available, the acceptance date was used instead. If neither was available,

then the submission date was used. For conference abstracts without a publication date, the earliest conference date was used.

4.3 Study Characteristics and Outcomes

Metadata for ClinicalTrials.gov registry entries were exported in .csv format and converted to a Microsoft Excel spreadsheet (.xlsx). The NCT number, year, title, intervention(s), and outcome(s) were used for study identification. Other noteworthy metadata included study status (completed, ongoing, terminated, etc.), study design (RCT), intervention type (surgery/procedure, drug, or other), funding source, and enrollment size. Study registration date (first posted on ClinicalTrials.gov), start date (first participant enrolled), primary completion date (last participant examined for primary outcome measure), and overall completion date (last participant examined for all outcome measures) were recorded as indicated in the registry (NLM, n.d.). Estimated values and dates were accepted when actual values were not available. When applicable, we calculated the time in months from the study start date to the study primary outcome completion date, overall completion date, and date of first publication.

For each study, we recorded whether randomized allocation was used to distribute participants between treatment groups, as opposed to non-randomized allocation (i.e., single group, manual assignment, etc.). A study was considered prospectively registered if the estimated or actual start date occurred later than the date of initial registration, and retrospectively otherwise. Funding sources were recorded as industry, academic/government/other, or both. Studies with no stated funding were included in "other". For the purposes of analysis, any industry funding at all, compared to no industry funding, was used as a study characteristic. Regarding intervention types, surgical studies were compared to non-surgical studies (drug or other) while drug studies were compared to non-drug studies (surgery or other).

In the registry, the inclusion of each study's planned outcome(s) was recorded. If primary and secondary outcomes were not clearly indicated in the registry entry, the first-listed outcome was considered primary. Partially inspired by Saldanha et al. (2014), the completeness of primary outcome descriptions was judged using three criteria: a specific

metric to assess the outcome (e.g., change in heart rate, not “heart health”); a unit or scale of measurement for the relevant data (e.g., beats per minute); and at least one timepoint of outcome ascertainment (e.g., after 1 hour).

Among publication abstracts, the inclusion of planned primary outcome at all was recorded. When multiple primary outcomes were mentioned in the registry, only the first-listed was considered. This was to prevent any potential advantages from studies that mentioned more than one planned primary outcome (i.e., more opportunities to obtain favourable results). A published outcome only had to fit the registry outcome metric to be considered a match with the planned outcome. We decided not to require additional outcome criteria, such as matching measures and timepoints, because not all registry entries provided these details. Although complete definitions would be ideal (Saldanha et al., 2014), we believed it would be improper to penalize those that simply put more effort into their outcome descriptions, since any discrepancies with the publication would only be apparent among registry entries that included such details in the first place.

Planned primary outcomes that were not reported as primary outcomes or were not reported at all in the publication were considered switched. If primary and secondary outcomes were not clearly indicated in the abstract, primacy was used as follows. Among the listed outcomes, the first half of the list was considered primary while the second half was considered secondary. If there was an odd number of outcomes, the middle outcome was considered primary. For example, if there were a total of five outcomes described with no clear differentiation, only the first three would be considered primary outcomes. Study registry entries that did not specify planned outcomes were not included. We also excluded descriptive and summary statistics unless otherwise stated as being the primary objective of the study. By definition, the difference between studies that published any results and those that switched their primary outcome would equal the number of studies that published their planned primary outcome as primary.

To assess favourability of study results, a study was considered positive if at least one published primary outcome was both statistically significant and directionally in favour of the study hypothesis, and negative otherwise. Since statistical significance and

direction are not usually applicable to descriptive outcomes, they were considered inherently positive when applicable.

4.4 Statistical Analysis

Data analysis was performed in Stata 16 (StataCorp LLC, College Station, TX). Summary statistics were reported for relevant metadata, characteristics, and outcome variables, in the form of medians, means, proportions (as percentages), or sums. A *paired t-test* and McNemar's Chi^2 test were conducted to determine whether overall proportion published and proportion published with primary outcome were significantly different from one another. The continuous variables of time to study completion for primary outcome, time to overall study completion, time to first publication, and enrollment size were plotted on histograms to detect skew. Heavy skew would indicate the median was more representative than the mean as a summary statistic and vice versa. Additionally, the prevalence of various dichotomous study characteristics each year was calculated. The number of registered studies that began enrollment each year and the annual average sample sizes were also recorded.

Log-binomial regression (log link function) was used to analyze the relationship between our dichotomous study characteristics and publication status (i.e., published or not published), as well as publication with planned primary outcome, in the form of relative risk. Relative risk was preferred over odds ratios as it tends to be more conservative. The main study characteristics used were randomized design (RCT), prospective registration, industry funding, surgical study, and drug study. An additional set of models were run to analyze the effects of registry outcome description quality, among entries that had descriptions, using each of our three criteria as individual covariates and adjusting for the main study characteristics listed above. For each model, multicollinearity between variables was checked by calculating the variance inflation factors (VIF) for each of the covariates. VIF values greater than 10 would indicate presence of multicollinearity (UCLA, n.d.). Due to a high degree of multicollinearity between the outcome description covariates (Table C5), they were regressed separately. Favourability of results was not included as a covariate for publication because we lacked data on the results of

unpublished studies. As our study design only assessed results from published studies, the inclusion of positive results in the regression model would heavily skew our regression.

Enrollment size was kept as a continuous variable, instead of dichotomization into a dummy variable, as we lacked justification for arbitrary divisions. Due to convergence issues with binomial regression, we used modified Poisson regression (robust variance) (Zou, 2004) to examine the relationship between enrollment size and publication, as well as publication with planned primary outcome, adjusting for the main study characteristics included in our log-binomial model. To check for overdispersion, likelihood ratio tests were conducted between Poisson regression and analogous negative binomial regression models. These tests did not show evidence of overdispersion (Table C9, Table C10), justifying the applicability of Poisson regression. Relative risk, in the form of incidence rate ratios (IRR), was reported for increases of 1000 participants in order to improve readability. Pearson and point-biserial correlation coefficients were also calculated, between enrollment size and publication, as supplementary statistics.

Among published studies, a multiple linear regression model was used to analyze the relationship between our dichotomous study characteristics and time to first publication. Linear regression was preferred for its ease of understanding and absolute output statistics (i.e., differences in time). The characteristics included were positive results, randomized design, prospective registration, industry funding, surgical study, and drug study. The assumptions of normality, homoskedasticity, and linearity were checked (Figure C1, Figure C2, Figure C3, Table C11, Figure C4). Due to substantial right skew, the outcome variable time to publication was natural log-transformed. A robust variance estimate was also used due to potential heteroskedasticity. Linearity was met because binary variables are inherently linearly related to the outcome.

In addition to linear regression, we were interested in exploring the feasibility of time-to-event analysis to further evaluate time to publication for positive versus negative studies. To that end, we interpreted time to publication as the length of time observed from baseline (first patient enrolled) to our event of interest, publication. A cumulative incidence graph was generated, and a log-rank test was performed, to assess the

association between positive results and time to publication. We also attempted to run a post hoc multivariable Cox regression model for time to publication using the same covariates as our multiple linear regression model. However, since both the Schoenfeld residuals test and our log-log plot for positive results demonstrated that the proportional hazards assumption was not met (Table C14, Figure C5), the parametric Weibull regression was used instead. We have included the regression output in our appendices as reference (Table C15).

4.5 Graphs and Visuals

One of our additional objectives was to convey our data visually using various graphs. When applicable, proportions were shown as percentages.

The proportion of studies published, by year started (first participant enrolled), was made into a bar graph to show potential trends over time. To complement this graph, a Chi^2 test and one-way ANOVA were used to examine whether there was any relationship between study start year and publication rates. Analysis was limited to 1997 and onwards due to sparsity of data before that year (Table D2), which roughly coincides with FDAMA in 1997 and the formalization of ClinicalTrials.gov (NLM, n.d.).

In addition to the proportion of studies published, we included a post hoc estimate of studies “Not Yet Published” for each year. This was added to mitigate the perceived effects of right censoring, as some studies may be less likely to publish due to having less time. To estimate this, we first assume the distribution of the time to publication variable (Figure B3) is representative of the general timeliness of publication of studies started in any given year within our dataset. It is worth noting that time to publication is measured in months, ends on our publication cut-off date (July 31, 2020), and has a range of 304 months (Table 3). We calculate the length of time from the beginning of a given year to our cut-off date, then find the equivalent percentile for this length of time in the distribution of the time to publication variable. This percentile represents the estimated proportion of studies started that year that have been published out of all studies that will be published, at least within the ~25 years of our time to publication variable. Dividing the actual proportion by the percentile allows us to calculate the estimated proportion of

all registered studies that will be published. The difference between the estimated and actual proportion is thus our estimate for proportion of studies that have yet to be published from that given year. This process was repeated for each year (Table D3) and was shown as part of the stacked bars on our graph.

The registry metadata we collected on study statuses of unpublished studies was simplified into three mutually exclusive study categories: completed, ongoing, or stopped. Studies that were “completed” maintained their original registry definition. “Ongoing” studies were those that were recruiting, enrolling, or active. “Stopped” studies included studies that were withdrawn, suspended, terminated, or had unknown status. The frequencies and proportions of studies in each category, along with published studies, were then graphed by year started in separate stacked bar graphs.

To supplement publication over time, a multiple line graph was created to show changes in the prevalence of our main dichotomous study characteristics over the years. This included randomized design, prospective registration, industry funding, surgical studies, and drug studies.

A leaky bucket diagram (Glasziou, 2005) was created to show the proportion of studies that progressed through the stages of the research process to publication. The stages were registration, study start, completion, publication, and publication with planned primary outcome at all.

An icon array was also generated to show the distribution of positive and negative studies that were registered, completed, published, and published with planned primary outcome at all. Since we did not have any data on the results of unpublished studies, they could not be accurately represented. Under the circumstances, unpublished studies may be considered negative in interpretation, as they inherently failed to produce and publish any results in the context of our study. However, this would likely overestimate the proportion of negative studies among unpublished registered or completed studies. While not a direct representation, we hoped it could serve as a proof of concept.

Chapter 5

5 Results

The following chapter will report the results of our statistical analysis and data synthesis, starting with a summary of the baseline characteristics for our cohort of registered studies from London, Ontario. Analyses for the outcome variables of publication and time to publication are reported separately. The last section will be a description of the visual graphics generated to represent our data. All data is as of July 31, 2020.

5.1 Descriptive Statistics of Registered Studies From London, Ontario

Table 3

Summary statistics for count and continuous variables

Statistic	N	Median	IQR (Q1, Q3)	Mean	SD	Min	Max
Counts							
Registered Studies	2446	-	-	-	-	-	-
Enrollment Size*	2410	250	610 (80, 690)	980	5093	0	164946
Time Intervals (Months)							
Primary Outcome Completion*	2435	35	36 (21, 57)	44	34	0	340
Overall Completion*	2298	41	48 (23, 71)	53	43	0	340
Publication*	925	53	39 (36, 75)	63	41	1	304

Note. This table shows the total number of registered studies and the summary statistics for our count/continuous variables: study enrollment size, time to complete examinations for primary outcome, time to complete examinations for all outcomes, and time to first publication. N = number of distinct observations, IQR = interquartile range, SD = standard deviation, Min = minimum observed value, Max = maximum observed value. *Right-skewed (Figure B1, Figure B2, Figure B3, Figure B5).

A total of 2446 study registry entries were included in our dataset, of which 2410 included data on actual or expected enrollment numbers. Median study enrollment size was 250 participants. In chronological order, the median time for a study to finish ascertaining its primary outcome was 35 months, to complete the entire study was 41 months, and to publish their results was 53 months. All time interval variables showed strong evidence of right-skew. (Table 3)

From 1997 to 2017, the annual average enrollment size of randomized studies decreased over the years from 4455 to 1676 participants, while the average enrollment size of non-randomized studies increased from 151 to 284. Overall study enrollment sizes decreased from an annual average of 2542 in 1997 to 1484 in 2017. (Table B3)

Table 4

Prevalence and frequencies of dichotomous variables

Statistic	Prevalence (%)	Frequency (N = 2446)
Published	38*	925
Primary Outcome Published	30*	743
Primary Outcome Switched	8	197
RCT	84	2050

Prospective Registration	49	1206
Industry Funding	65	1565
Surgical Study	11	259
Drug Study	72	1752
Outcome Mentioned	96	2358
Metric Defined	68	1664
Measurement Defined	88	2152
Timepoint Defined	92	2247

Note. This table shows the percentages and frequencies of our dichotomous parameters and study characteristics: publication, publication with primary outcome at all, omission/demotion of primary outcome, randomized design (RCT), prospective registration, any industry funding, surgical intervention, and drug intervention. Registry outcome characteristics are also included: mention of any outcome at all, a specific metric, a unit/scale of measurement, and a timepoint of ascertainment.
*Significantly different via *paired t-test* and McNemar's Chi^2 test, both $p < 0.001$ (Table B1, Table B2).

Overall, most studies were RCTs (84%) and received some form of industry funding (65%). Fewer than half (49%) were registered prior to beginning the study. For intervention types, 11% were surgical while 72% were drug studies. Nearly all registry entries mentioned an outcome of interest (96%). However, 32% of them did not provide their metric for assessing their outcome, 12% did not mention the measurement to be used, and 8% did not give at least one timepoint for outcome ascertainment. (Table 4)

In total, 38% of all studies were published. Only 30% were published with their planned primary outcome at all. About 8% of all studies switched their primary outcome in publication either by complete omission or demotion to a secondary outcome. Demotion

alone was rare and occurred in fewer than 1% of studies. Additionally, among published studies, almost 85% had positive results (Table D8). (Table 4)

5.2 Probability of Publication

Table 5

Adjusted RR for association between study characteristics and study publication as well as publication with primary outcome

Predictor	RR	95% CI	*p* value
Publication with Any Results			
RCT	1.46	1.22, 1.73	<0.001
Prospective Registration	1.15	1.04, 1.26	0.006
Industry Funding	1.67	1.45, 1.92	<0.001
Surgical Study	0.86	0.69, 1.09	0.22
Drug Study	1.26	1.09, 1.45	0.002
Publication with Primary Outcome			
RCT	1.50	1.22, 1.83	<0.001
Prospective Registration	1.20	1.06, 1.34	0.003
Industry Funding	1.85	1.56, 2.19	<0.001
Surgical Study	0.82	0.62, 1.08	0.16
Drug Study	1.16	0.98, 1.36	0.08

Note. This table shows the estimated relative risks (RR) for publication and publication with primary outcome at all, using multivariable log-binomial regression, associated with our dichotomous study characteristics: randomized design (RCT), prospective

registration, any industry funding, surgical intervention, and drug intervention. CI = confidence interval.

Relative risk for publication with any results was significantly increased for randomized versus non-randomized study design (RR 1.46 [95% CI: 1.22, 1.73]), prospective versus retrospective registration (1.15 [1.04, 1.26]), and industry versus other funding (1.67 [1.45, 1.92]). Similarly, relative risk for publication of planned primary outcome was also significantly increased for randomized design (1.50 [1.22, 1.83]), prospective registration (1.20 [1.06, 1.34]), and industry funding (1.85 [1.56, 2.19]). Surgical versus non-surgical studies were not significantly associated with publication or publication with primary outcome (0.86 [0.69, 1.09] and 0.82 [0.62, 1.08], respectively). Drug versus non-drug studies were significantly associated with increased publication with any results (1.26 [1.09, 1.45]) but not publication with primary outcome (1.16 [0.98, 1.36]). (Table 5)

Table 6

Adjusted RR for association between individual outcome description quality criteria and study publication as well as publication with primary outcome

Predictor	RR	95% CI	*p* value
Publication with Any Results			
Metric	1.61	1.39, 1.85	<0.001
Measurement	2.01	1.50, 2.68	<0.001
Timepoint	2.51	1.62, 3.88	<0.001
Publication with Primary Outcome			
Metric	1.65	1.40, 1.95	<0.001
Measurement	1.80	1.32, 2.45	<0.001
Timepoint	2.11	1.34, 3.33	0.001

Note. This table shows the estimated relative risks (RR) for publication and publication with primary outcome at all, using multivariable log-binomial regression, associated with our criteria for registry outcome description quality: mention of a specific metric, a unit/scale of measurement, and a timepoint of ascertainment. Regression models were run separately for each criterion due to high multicollinearity (Table C5, Table C6) and adjusted for randomized design, prospective registration, any industry funding, surgical intervention, and drug intervention. CI = confidence interval.

Relative risk for publication with any results was significantly and positively associated with defining the metric in the registry entry (1.61 [1.39, 1.85]), mentioning the measurement in the registry entry (2.01 [1.50, 2.68]), and describing at least one timepoint of outcome ascertainment in the registry entry (2.51 [1.62, 3.88]). Similarly, relative risk for publication with planned primary outcome was also significantly associated with including the metric (1.65 [1.40, 1.95]), measurement (1.80 [1.32, 2.45]), and timepoint (2.11 [1.34, 3.33]) in the registry entry. (Table 6)

Table 7

Adjusted IRR for association between enrollment size and study publication as well as publication with primary outcome, per 1000 increase in enrollment size

Predictor	IRR (per 1000)	95% CI	*p* value
Publication with Any Results			
Enrollment Size	1.014	1.007, 1.021	<0.001
Publication with Primary Outcome			
Enrollment Size	1.015	1.008, 1.021	<0.001

Note. This table shows the estimated incidence rate ratios (IRR) for publication and publication with primary outcome at all, using multivariable modified Poisson regression, associated with study enrollment size. The model was adjusted for randomized design,

prospective registration, any industry funding, surgical intervention, and drug intervention. IRRs are reported for increases of 1000 participants. CI = confidence interval.

Enrollment size was significantly and positively correlated with likelihood of publication (Table C1, Table C2). Relative risk for an increase in enrollment size was significantly associated with both publication and publication with primary outcome at all (IRR 1.014 [1.007, 1.021] and 1.015 [1.008, 1.021] per increase of 1000, respectively). (Table 7)

5.3 Time to Publication

Table 8

Association between study characteristics and time to publication

Predictor	Difference (Months)	95% CI	*p* value
Positive Results	-9.70	-15.75, -3.65	0.002
RCT	4.05	-3.84, 11.93	0.31
Prospective Registration	-16.12	-21.17, -11.07	<0.001
Industry Funding	-22.72	-30.43, -15.00	<0.001
Surgical Study	9.28	-2.01, 20.57	0.11
Drug Study	9.58	1.76, 17.40	0.02

Note. This table shows the estimated difference in months for time to publication, using multiple linear regression, associated with our dichotomous study characteristics: positive results, randomized design (RCT), prospective registration, any industry funding, surgical intervention, and drug intervention. CI = confidence interval.

Among published studies, positive results were significantly associated with decreased time to publication by almost 10 months relative to non-significant or unfavourable

results. Randomized design was not significantly associated with a change in time to publication (~4 months, $p = 0.31$). Prospective registration was significantly associated with decreased time to publication by over 16 months relative to retrospective registration. Receiving any form of industry funding was significantly associated with decreased time to publication by almost 23 months relative to only receiving funding from non-industry sources. Surgical studies were not significantly associated with a change in time to publication compared to non-surgical studies (~9 months, $p = 0.11$), while drug studies versus non-drug studies were significantly associated with an increase in time to publication of almost 10 months. (Table 8)

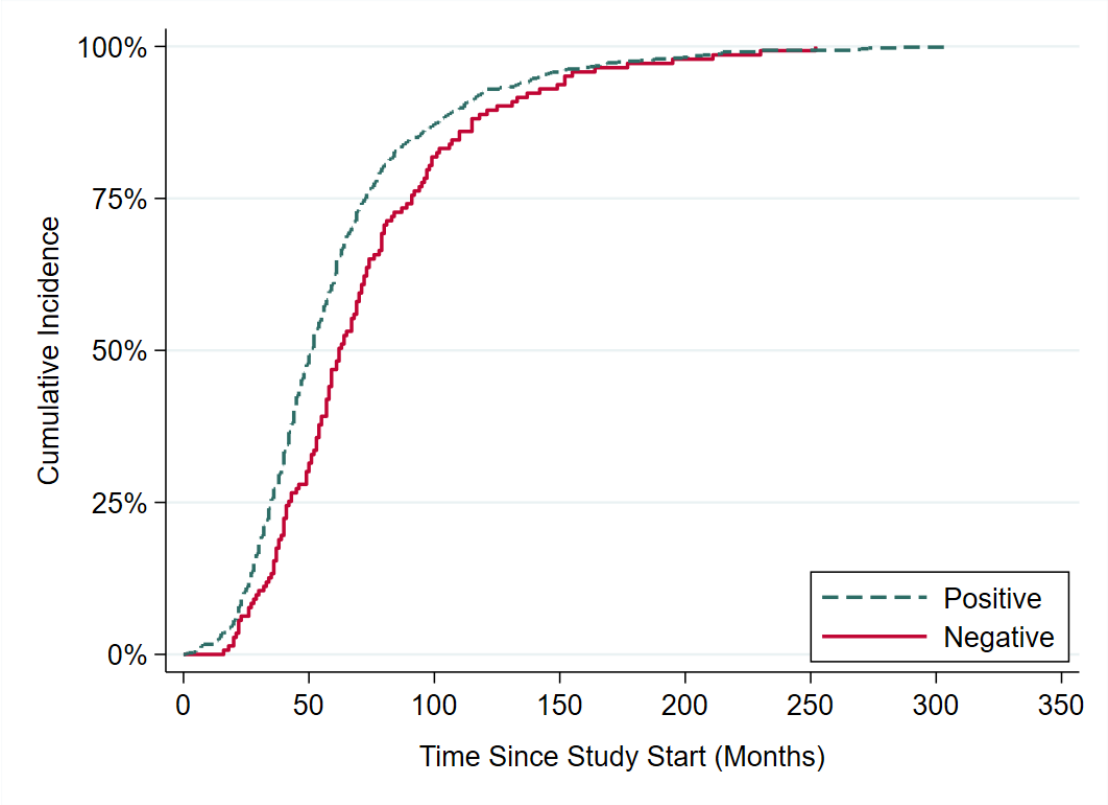


Figure 5. Cumulative incidence of publication for positive versus negative studies. This cumulative incidence graph shows the survival curves for study publication, after their first participant was enrolled, of published studies with positive (dashed line) and negative (solid line) results. Log-rank test showed a statistically significant difference between the survival curves ($p < 0.05$) (Table C13).

Both the positive and negative study survival curves had a sharp initial incline until about 100 months, followed by a plateau. The overall 50% cumulative incidence roughly coincides with our median time to publication of 53 months (Table 3). While the initial incline was almost immediate for positive studies, negative studies did not have their first incidence of publication until around 15 months. Overall, positive studies were consistently published sooner than negative studies, which was confirmed to be statistically significant via log-rank test. (Figure 5)

5.4 Graphics

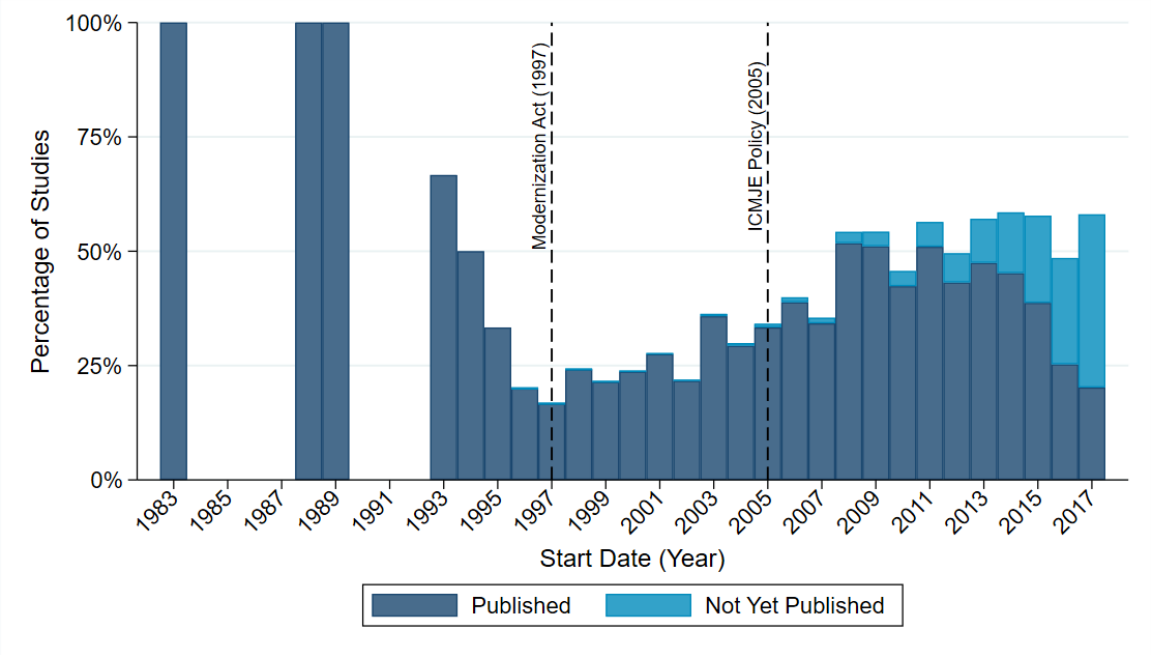


Figure 6. Study publication rates by start year. This stacked bar graph shows the annual percentage of all registered studies that are published or not yet published, by the year their first participant was enrolled, up to 2017. FDAMA (1997) and ICMJE (2005) are indicated by the vertical dashed lines.

Percentage values prior to 1997 were sporadic due to sparsity of data before formal adoption of the registry and can be considered outliers. From 1997 onwards, the actual percentage of studies published has steadily increased over time from under 20% to a high of around 50% in 2008. This was followed by a subsequent decrease in the mid 2010s, which could be expected due to right censoring and coincides with our median

time to publication of ~4 years (Table 3). Including not yet published studies seemed to mitigate this decrease and continues the reasonably linear upwards trend to an estimated percent published of nearly 60% in 2017. In addition, restricting to 1997 and onwards, overall publication was found to be significantly associated with the year the study was started (Table D4, Table D5). (Figure 6, Table D3)

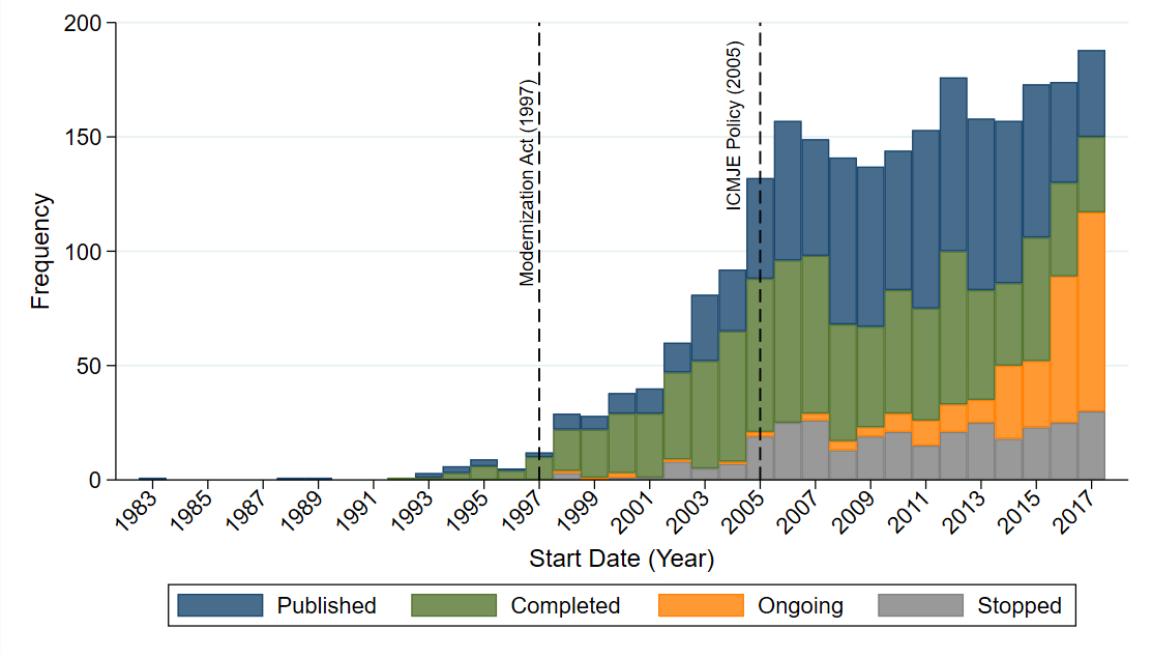


Figure 7. Fate of all registered studies by start year (frequencies). This stacked bar graph shows the frequency distribution of statuses of all registered studies, by the year their first participant was enrolled, up to 2017. FDAMA (1997) and ICMJE (2005) are indicated by the vertical dashed lines. Completed, ongoing, and stopped studies are unpublished. Studies cannot be in more than one status group at a time.

Frequency of all registered and started studies increased dramatically following formalization of the registry with a spike around 2005. Completed and published studies decreased in recent years, which coincided with a large increase in ongoing studies. As expected, the frequency of ongoing studies increased dramatically in more recent years. Frequency of terminated and stopped studies saw a rise in the mid 2000s but seems to have remained constant since. (Figure 7, Table D2)

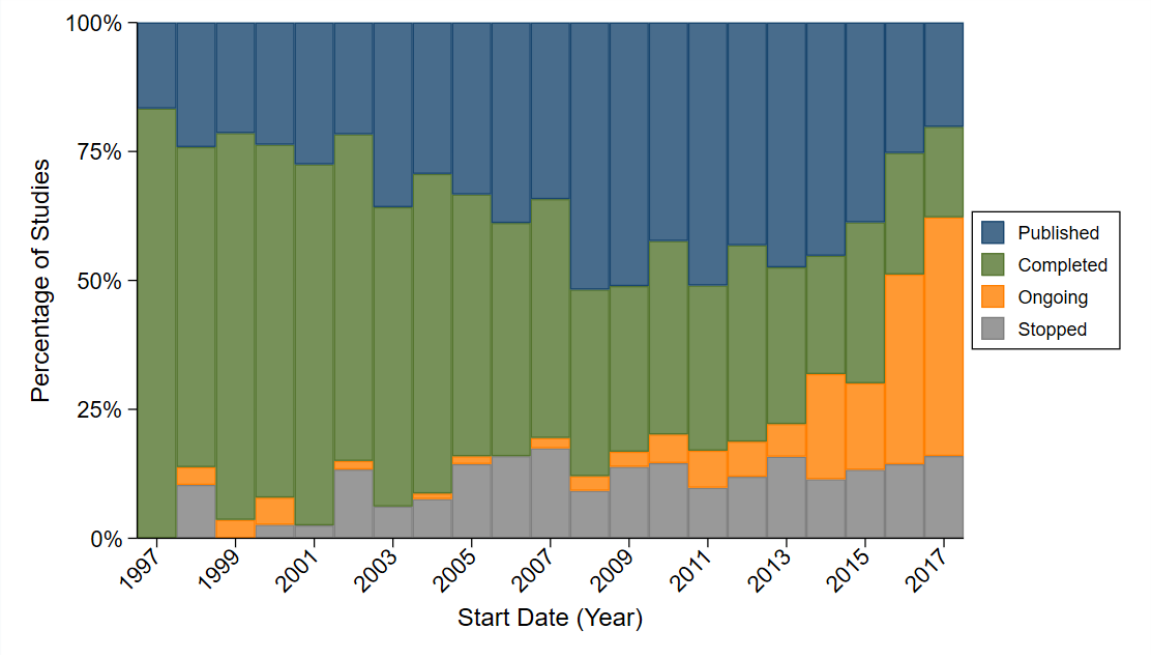


Figure 8. Fate of all registered studies by start year (proportions). This stacked bar graph shows the percentage distribution of statuses of all registered studies, by the year their first participant was enrolled, from 1997 to 2017.

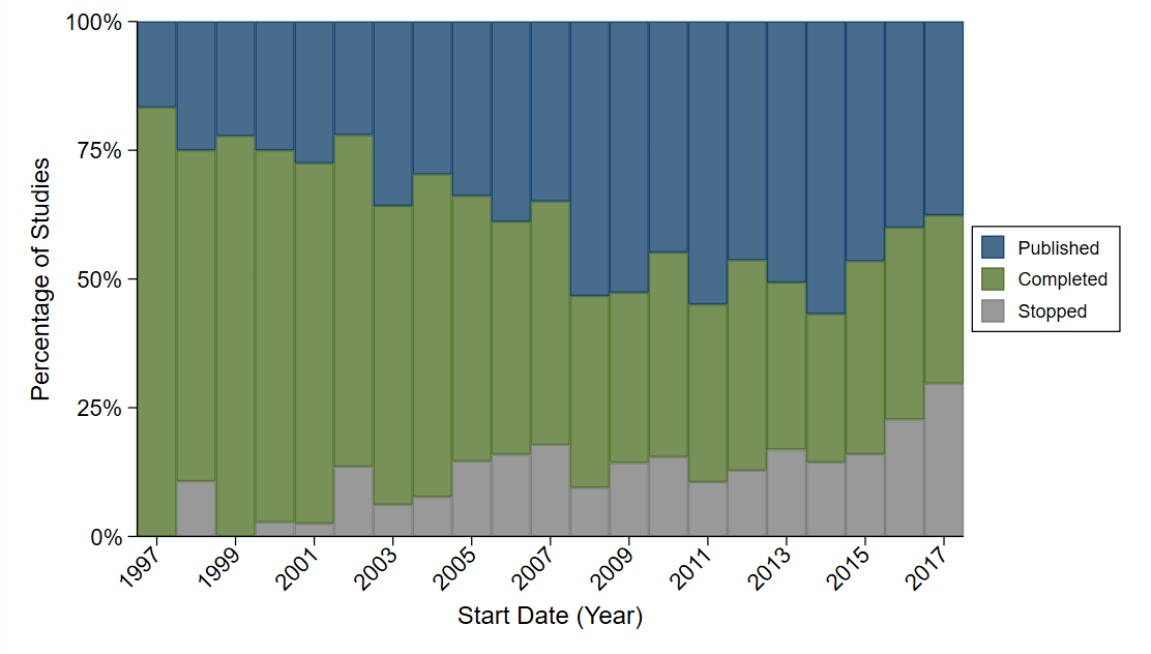


Figure 9. Fate of completed and stopped studies by start year (proportions). This stacked bar graph shows the percentage distribution of statuses of registered studies excluding ongoing studies, by the year their first participant was enrolled, from 1997 to 2017.

Figure 8 offers an alternative view to Figure 7 by displaying proportions of statuses (as percentages) instead of frequencies. Completed studies made up the largest percentage in the initial years, then gradually decreased over time. This decrease coincided with a gradual increase in published studies, as well as a dramatic increase in ongoing studies in the more recent years. Annual percentage of stopped or terminated studies seems to have remained fairly constant (Figure 8, Table D2). Excluding ongoing studies in Figure 9 did not substantially influence trends among published and completed studies. However, the percentage of stopped studies each year seems to have increased in more recent years relative to the other statuses (Figure 9).

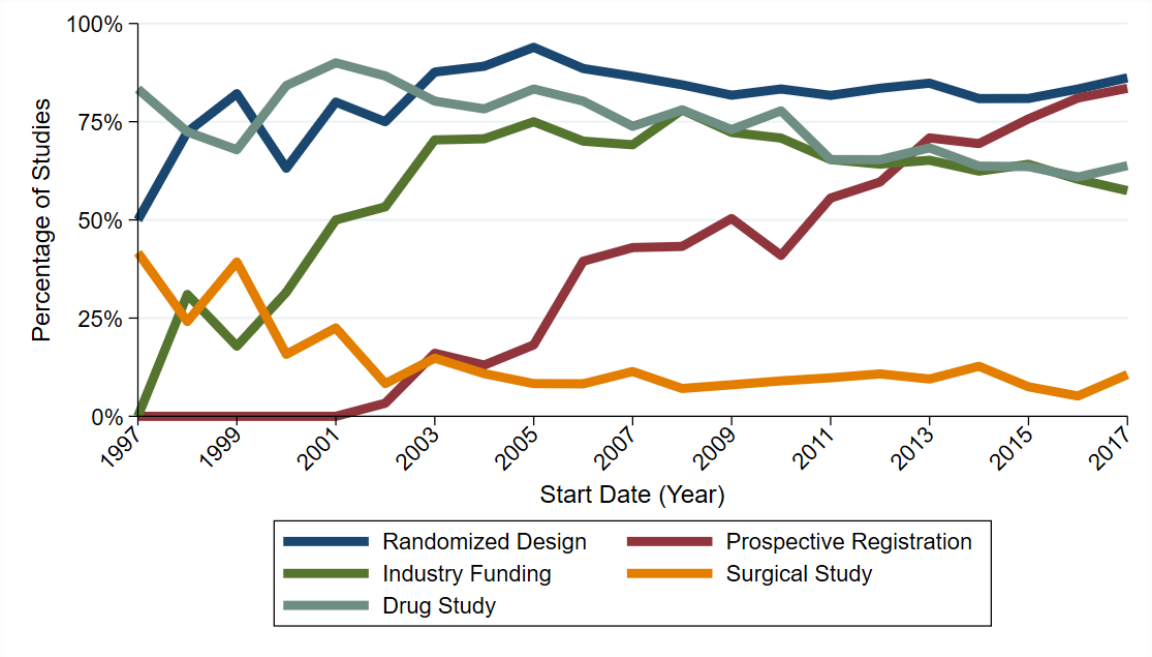


Figure 10. Prevalence of study characteristics by start year. This multiple line graph shows the prevalence over time of various dichotomous study characteristics among all registered studies, by the year their first participant was enrolled, from 1997 to 2017.

The percentage of studies started each year using randomized design initially increased from about 50% to a peak of over 90%, followed by a plateau in the mid-2000s, and has since remained steady at a little over 80%. The percentage of studies each year that were prospectively registered has increased almost linearly from virtually none prior to 2000 to over 80% in more recent years, which coincides with ClinicalTrials.gov going public in

the year 2000 (NLM, n.d.). The percentage of studies each year that were industry-funded increased dramatically from 1997 to the mid-2000s, reaching a peak of almost 80%, followed by a gradual decrease to about 60%. The percentage of surgical studies each year started at 40% in 1997 but declined rapidly until the early 2000s, remaining steady at around 10%. The percentage of drug studies each year has declined gradually over the years from over 80% to a little over 60%. (Figure 10, Table D6)

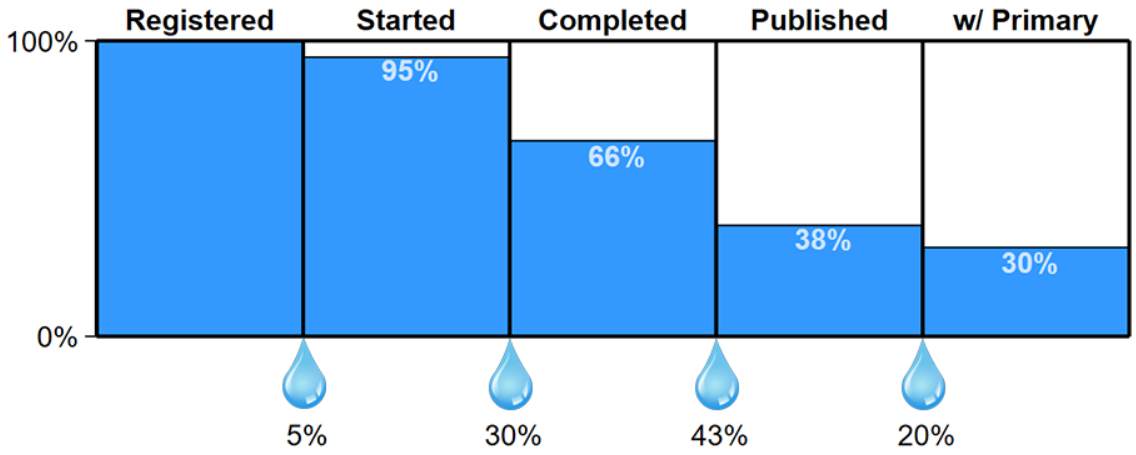


Figure 11. "Leaks" in the research process. This leaky bucket diagram shows the percentage of all registered studies that progress through each stage of the research process. Percentages at each bar level represent the overall proportion of all registered studies that reached that stage. Percentages along the bottom indicate the conditional proportion of studies lost relative to the previous stage.

The majority of studies that are registered seem to at least be able to start in some capacity, with only 5% failing to do so. Reaching publication seemed to be the largest point of failure, losing about 43% of the completed studies. Failing to reach completion after starting was another substantial leak, losing 30%. Among studies that reached publication, about 20% did not publish their planned primary outcome at all. (Figure 11, Table D7)

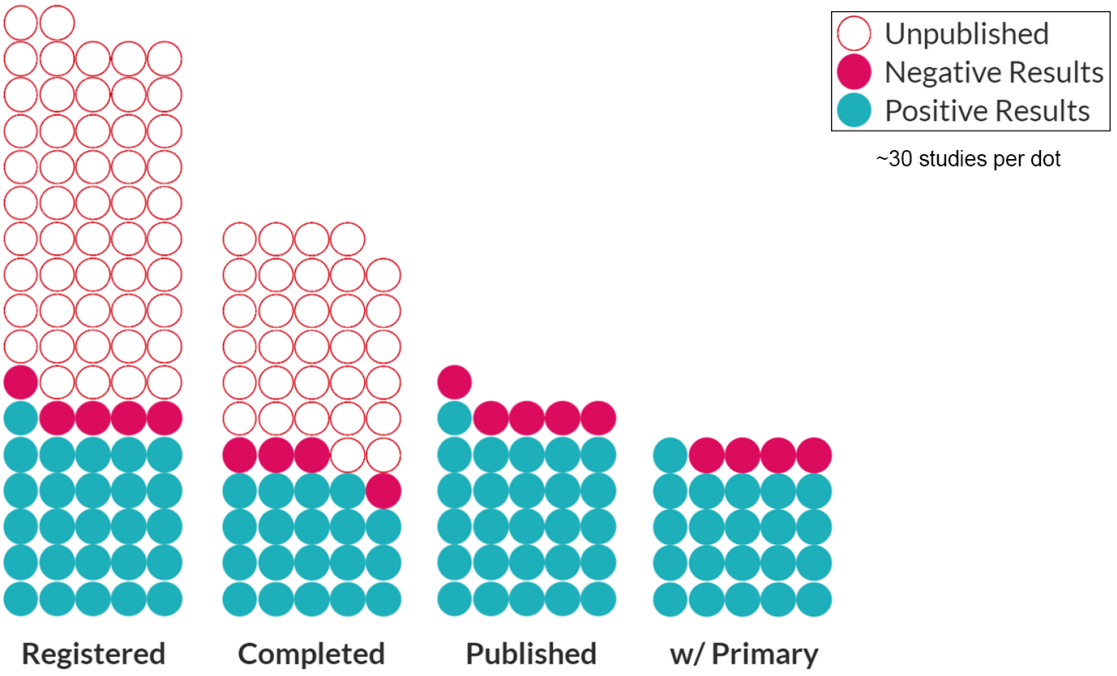


Figure 12. Distribution of positive and negative studies across stages of the research process. This icon array shows the proportion of studies within each stage with positive or negative results as represented by green and red dots, respectively. Each dot represents approximately 30 studies. Unfilled dots represent unpublished studies, which may be considered negative. The number of positive published studies is greater than that of completed studies due to publications from ongoing and stopped studies.

The overwhelming majority of studies that reach some form of publication had positive results, and this distribution did not differ greatly between those published and published with primary outcome. Studies published with any results and published with planned primary outcome had similar proportions of positive studies. The majority of studies that only reached registration or completion seemed to be negative or unpublished. Due to not having the results of unpublished registered studies, the number of positive studies in the registered and published stage are inherently the same. (Figure 12, Table D8)

Chapter 6

6 Discussion

This chapter will provide an overview of our results. We will reiterate our main findings, then discuss their implications and how they relate to existing literature. Potential shortcomings from our research methodology are also addressed.

6.1 Publication Bias

Over 60% of the registered studies in our cohort failed to reach publication. If we require that publications are published with their planned primary outcome at all, our estimate for failure to publish is significantly increased to 70% (Table 4). As expected, our findings were not as poor as the 80% failure rate in reporting of results for London, Ontario research shown by the FDAAA TrialsTracker (EBM DataLab, 2018). However, our study indicates a low rate of publication when compared to similarly conducted studies (Table 1). Furthermore, the significant decrease in proportion published when using more strict publication outcome criteria suggests outcome switching may be prevalent among the published studies in our cohort (Table B1, Table B2).

As one would expect, studies with randomized design were more likely to publish. This could be because RCTs may produce more reliable results than single-arm or non-randomized studies, which would perhaps be more readily accepted by journals. In contrast to much of the past literature (Jones et al., 2013; Kasenda et al., 2014; Pica & Bourgeois, 2016; Rosenthal et al., 2015; von Elm et al., 2008), industry funding was significantly associated with increased publication in our cohort. As registration has become the norm, the non-publication of trial results may have become subject to greater scrutiny. As a result, industry sponsors affiliated with London, Ontario may have started encouraging their constituents to publish instead of withholding research. This could include mandating better study planning and closer adherence to publication requirements from ICMJE or other organizations, ultimately increasing the likelihood of publication. Among our registered studies, drug studies made up the majority (Table 4) and seemed to

be associated with a significantly greater likelihood of overall publication than other types of interventional studies. However, this association was not significant for publication with primary outcome, suggesting that there may be some risk of outcome switching among drug studies. Similar to Rosenthal et al. (2015), we found signs that surgical studies could be associated with a decrease in publication compared to non-surgical studies, but our results were ultimately inconclusive. As with Rosenthal et al. (2015), this may have been due to a lack of statistical power for surgical studies. Surgical studies made up only 11% of our cohort (Table 4) and may also be decreasing in relative frequency over time (Figure 10). Nevertheless, the consistency of our findings suggests that surgical studies need to be studied further and with greater sample size. (Table 5)

Contrary to existing studies (Al-Durra et al., 2018; Khan et al., 2014), we found prospective registration to be significantly associated with an increase in both overall publication and publication with primary outcome, compared to retrospective registration (Table 5). Likewise, each of our separately modeled outcome description criteria also significantly increased publication, with some more than doubling the probability (Table 6). We found strong evidence of multicollinearity between these criteria (Table C5), suggesting that studies that meet one criterion are likely to meet the others and that there may be an underlying reason for this correlation. We speculate that prospective registration and better-defined outcomes reflect overall better planning and critical thinking prior to conducting the study. Not only do better outcome descriptions allow us to better diagnose publication biases, it seems they also increase chances of successful completion and eventual publication. This supports the need for more rigorous guidelines for registry entries and protocols.

In line with existing research (Al-Durra et al., 2018; Blümle et al., 2014; Kasenda et al., 2014; Pica & Bourgeois, 2016; von Elm et al., 2008), enrollment size was significantly associated with an increase in both publication and publication with primary outcome, with similar effect sizes for both (Table 7). For every 1000 additional study participants, the incidence rate of publication increased by 1.4–1.5%. These findings suggest that larger studies tend to have a greater likelihood of publication, which may reflect more

available resources and personnel. Smaller studies may also lack statistical power, leading to fewer significant results and lower rates of publication.

6.2 Time to Publication

Post-start time to publication in our cohort fell roughly in line with existing studies that measured post-start or post-approval (Table 2, Table 3). However, the difference between our median time to primary outcome completion and time to publication was fewer than 20 months (Table 3). This suggests a shorter post-completion time to publication than existing studies, which reported time intervals up to twice as long (Table 2). Also, the steep initial incline in cumulative incidence of publication (Figure 5) and strong evidence of right-skew (Figure B3) suggest that most studies that publish tend to do so earlier.

Positive studies were generally published sooner than negative studies (Figure 5, Table C13), reaffirming previous findings (Khan et al., 2014). In fact, positive results were associated with a reduction in time to publication by almost a year. Favourable results may motivate researchers to publish sooner, leading to further bias in the published literature. Prospective registration was associated with a reduction in time to publication of well over a year, which coincides with an increased likelihood of publication (Table 5). This may indicate better planning and foresight, leading to more efficient execution and faster completion. In contrast to existing literature (Pica & Bourgeois, 2016), industry funding was associated with a reduction in time to publication of over 2 years. Our findings coincided with a greater rate of overall publication from industry funding and an increase in industry-funded studies over time (Table 5, Figure 10), which could be attributed to greater pressure from industry sponsors to register and publish. Considering our median time to publication was roughly 4.5 years (Table 3), these study characteristics combined could contribute to a total reduction of up to 90%. (Table 8)

Although drug studies were associated with an increase in publication (Table 5), it was also associated with longer time to publication by almost a year. This suggests that drug studies overall tend to take longer to complete and publish. It is also possible that non-industry-funded and negative drug studies have a much harder time reaching publication than their counterparts. Additionally, both randomized design and surgical studies failed

to reach significance, suggesting that they do not have a substantial effect on time to publication. However, for surgical studies, the lack of significance may have been due to the lack of statistical power in our cohort (Table 4). (Table 8)

6.3 Publications Over Time

Both frequency and proportion of registered studies published seem to be steadily increasing over time, particularly following FDAMA in 1997 (NLM, n.d.) (Figure 6, Figure 7, Figure 8, Figure 9) and the ICMJE mandate in 2005 (De Angelis et al., 2004) (Figure 7). The proportion of prospectively registered studies each year have also been increasing since 2000 (Figure 10). This supports that the implementation of ClinicalTrials.gov and ICMJE registration requirements have been beneficial towards both increasing study registration and likelihood of publication. Although our graphs showed a sudden decrease in publications in the most recent few years, this can easily be explained by the increase in ongoing and completed studies that have simply not yet reached publication (Figure 7, Figure 8). Adding an estimate of such studies produced a more reasonable and consistent increasing trend (Figure 6). That being said, the observed decrease suggests that 2.5 years was simply not enough time for all studies in our cohort to reach publication. Nevertheless, these findings support a promising trend of increasing publication rates over time, consistent with the findings of Khan et al. (2014). Notably, frequency graphs seemed to be more effective at displaying sparse data than proportion graphs, particularly with data prior to 1997 (Figure 6, Figure 7).

The reasons for the observed increases in annual proportion published are uncertain. On the one hand, the increase has coincided with more prospective registration and a greater proportion of RCTs, which suggest better accountability and quality of research. While average sample sizes have somewhat decreased (Table B3), better methodology in recent studies could have improved efficiency and reduced the size needed to achieve adequate statistical power. On the other hand, the increased publication rate has also coincided with an increase in proportion of industry-funded studies and a relative decrease in surgical studies. This suggests that the increase in publication may instead be due to greater industry pressure and an increase in types of studies that are more likely to publish (Table 5). However, it should be noted that “industry-funded” includes studies

that were also funded by other sources, which likely underrepresents the proportion of academic and government-funded studies. It is also possible that some of the observed increase in registered industry-funded studies is a result of increasing adoption of public registries, in lieu of private industry registries, as opposed to a true increase in industry-funded research. (Figure 10)

Another noteworthy point is that the observed increase in total registered and prospectively registered studies (Figure 7, Figure 10), particularly following the ICMJE mandate in 2005 (De Angelis et al., 2004), suggests the possibility of some reverse causation. For example, a strong pre-existing desire from researchers or industry sponsors to publish may have pushed researchers to adhere to ICMJE publication policies, which requires them to register before starting their trial. Although prospective registration may be a predictive factor of publication (Table 5), the reverse relationship may be true where studies are prospectively registered specifically because they already have a high likelihood of publication success from other factors, such as high researcher motivation.

Our leaky bucket diagram (Figure 11) revealed that a study reaching completion and publication were the biggest bottlenecks in the research process. The vast majority of registered studies are started in some capacity but meet resistance either while conducting the study or submitting the results for publication. The last stage, while not the steepest relative drop, suggests that a troubling one in five published studies do not include their planned primary outcome at all.

The findings from our icon array diagram (Figure 12) were also troubling. While the vast majority of registered and completed studies were potentially negative, positive studies made up the majority of published studies. These findings support that positive studies are highly overrepresented in the published literature. Furthermore, there are noticeably fewer completed studies with positive results than published studies. This mismatch is possible because ongoing and stopped studies may still publish results. However, this could be indicative of selective outcome reporting because these studies would presumably be publishing before they finished ascertaining all their planned outcomes.

6.4 Implications for London, Ontario

It seems that research institutions affiliated with London, Ontario are headed in the right direction, as rates of prospective registration are increasing (Figure 10). Excluding the most recent years, annual publication rates also seem to be increasing at a steady pace (Figure 7). This could be due to increasing awareness of publication bias and better local research policies, including the introduction of registration requirements for ethics approval of human clinical trials at Western University (OHRE, 2016). Furthermore, studies that do publish generally do so sooner than contemporary estimates from other studies on time to publication (Table 2, Table 3). If these trends continue, research loss could be better mitigated in the future.

Unfortunately, the rate of failure to publish remains high for studies associated with London, Ontario over the past two decades. Compared to other studies on non-publication, our requirements for identifying acceptable corresponding publications were exceptionally lenient, essentially only requiring matching NCT numbers and some form of results. Even regarding whether publications included their planned primary outcome, we only required that it reasonably matched one of their published outcomes. We did not scrutinize changes in secondary outcomes, analysis plans, measurement timepoints, or other issues like interpretive bias (spin) (Chiu et al., 2017). Despite this leniency, our proportion published ranks as one of the lowest among existing studies (Table 1). In fact, our estimated 30–38% rate of publication is lower than any of the studies that focussed on select regions and institutions, which generally reported rates of 50% or more (Blümle et al., 2008, 2014; Kasenda et al., 2014; Kirkham et al., 2016; Menzel et al., 2007; Rosenthal et al., 2015; von Elm et al., 2008). While one could argue that our use of ClinicalTrials.gov did not adequately restrict our data to institutions based in London, Ontario, our findings are no less troubling for the region as a whole.

The increase in absolute number of registered studies and publications over time, while the overall rate of failure to publish remains high, could be due to increases in industry-funded studies (Figure 10), which may be more likely to publish and sooner (Table 5, Table 8). Further research is necessary to explore the barriers that academic and government-funded studies may face when reaching publication. Otherwise, the

published literature could become increasingly dominated by industry studies. Notably, the research affiliated with London, Ontario is already dominated by positive studies and drug studies (Figure 10, Figure 12). If researchers are reluctant to conduct or submit studies with anticipated low publication success, this would further bias the available data and be detrimental to evidence-based decision-making. In addition, an increase in overall publications does not necessarily mean a decrease in data loss. Our findings have shown that outcome switching is prevalent among studies affiliated with London, Ontario and can significantly impact measures of publication (Table 4). It is possible that the increase in publications may have coincided with an increase in selective outcome reporting.

Even if studies are legally required to submit their results to the FDA, many still fail to do so (Prayle et al., 2012) and this does not prevent discrepancies in outcome reporting (Thaler et al., 2015). Publication of results is largely voluntary and inconsistent at best, especially among studies lacking external pressure from industry sponsors. Unpublished data is not easily accessible to the public and will have limited impact. In the absence of regulatory pressure, the majority of London, Ontario research is likely still being wasted. Thus, it is clear that more improvements are necessary to the current research infrastructure in London, Ontario.

6.5 Limitations and Future Directions

A major limitation of our study was our data source. Some of the metadata on ClinicalTrials.gov may not have been reliable, especially since we accepted estimated values. For example, some studies had time to completion intervals of zero months (Table 3), which seems highly implausible and suggests the registry dates were incorrect. We did not include observational or qualitative studies in our cohort, which may have yielded different results from interventional studies. We also accepted studies that had any affiliation with London, Ontario instead of restricting to institutions based in the city. While this likely makes our data more representative of all research conducted in London, Ontario, it means that our results may include multicenter studies influenced by factors outside of the region. A potential solution could be to use a local registry instead, such as ethics review submissions. This could provide more accurate metadata, include non-clinical studies, and limit the regional scope more effectively. R scripts may also be

used to organize large sets of registry metadata further (Ramagopalan et al., 2014), instead of relying solely on the registry's existing filter functions.

Both our search for planned studies and corresponding publications were limited to publicly accessible databases, as we believed these to be representative of most clinical research. We did not reach out to the authors themselves for their manuscripts or publications, which would likely have increased our publication estimates. Additionally, more recently registered studies in our cohort may not have had adequate time to complete their studies and submit publications, leading to right censoring where publications from studies that are slower to publish may not have been detected within our search window. This could be mitigated by increasing the duration of follow-up and searching for publications at a later date. Our study also did not include grey literature, such as clinical study reports and registry results. Consequently, we were unable to compare the characteristics of published and unpublished literature, including whether favourability of study results could be associated with likelihood of publication. This also prevented us from accurately representing the distribution of positive and negative results among registered and completed studies (Figure 12). Future endeavours may include using grey literature sources and searching for trial results on ClinicalTrials.gov.

Another issue is that our search strategy for published literature relied heavily on publications including their trial NCT number, as authors may neglect to include it. A more comprehensive approach may entail searching other metadata entries, such as title and treatment, when the NCT number alone fails to find a corresponding publication. Furthermore, our screening of publications was mostly limited to study abstracts, which may not adequately represent all the published outcomes. For example, authors may choose to include some results in their appendices, which we would have missed.

We adopted a relatively simple approach in our analysis for the purposes of exploration. Most of our data interpretation assumed a linear and sequential progression through the research process. We assumed the registry and publication dates would represent common stages of the research process, but the amount of work done at any given stage may vary drastically from study to study. Our log-binomial and linear regression analysis

also did not account for right censoring, which could reduce the validity of our estimates. Future endeavours may include expanding on our time-to-event analysis, which may be better suited to the limitations of our data and allow us to better understand the influences of study characteristics on likelihood of publications over time. Potential models include the parametric Weibull regression, which would not require strong adherence to assumptions of temporality, censoring, or proportional hazards.

We did not assess outcome reporting bias in-depth. Further analysis of selective outcome reporting could give us a better understanding of data loss among published studies. For example, industry funding may increase likelihood of overall publication but have the opposite effect on publishing individual outcomes (Kirkham et al., 2016). Although harder to detect, selective outcome reporting and its causes are no less important and may warrant its own study. Additionally, although we assessed the associations of drug and surgical studies, the reasons for their differences in likelihood of publication are not clear. For example, different types of treatments may be subject to different thresholds for favourable results or adverse effects. To that end, a greater or targeted sample may be beneficial, as our cohort lacked statistical power for surgical studies. With regard to our analysis of registry outcome description quality, we did not consider the recent changes made to the ClinicalTrials.gov registration requirements in 2017, which now require much greater specificity in registered outcomes (i.e., specific metric, measurement, timepoint, etc.) (Collins & Burwell, 2016). Further research is necessary to assess the effectiveness of these updated requirements and their influence on publication.

While most of our graphs proved effective at visually conveying information, there were some areas that could be improved. Our proportion published over time graphs could include more landmark timepoints, such as local policy changes. Our leaky bucket and icon array diagrams could also be expanded to include more stages of the research process, such as obtaining funding, completing analysis, and journal submission.

Lastly, our study only covered a small portion of the research process: publication. We did not attempt to address issues in study design, data collection, analysis, or interpretation. Lack of bias prevention, poor quality data, and manipulation of statistics to

reach significance (p-hacking) may diminish the integrity and reproducibility of research (Munafò et al., 2017). With regard to our registry data collection, we only used the most recent information and did not consider potential changes over the course of the study. For example, researchers may have changed their hypotheses and outcomes after completing analysis to make their results appear more favourable (HARKing) (Munafò et al., 2017). Post hoc changes to registry data may bias results and threaten the reliability of registries as data sources. Thus, more in-depth research into other aspects of the research process is necessary to improve research transparency and inform better practises.

6.6 Conclusion

Publication bias is a widespread issue and London, Ontario is no exception. As many as two-thirds of registered studies fail to publish their results. While our cohort had lower publication rates than contemporary studies, those that did publish generally did so sooner. Prospective study registration and industry funding were associated with both a significantly increased likelihood of publication and shortened post-start time to publication. Randomized design, larger enrollment sizes, and better outcome planning were associated with increased likelihood of publication, while positive results were associated with shortened time to publication. Drug studies were more likely to publish but tended to take longer than non-drug studies.

Our incorporation of visual diagrams proved to enhance the understandability of our data and outlined details that may not have been obvious from the statistics alone. The observed increase in publications in London, Ontario seems to coincide with increased prevalence of randomized, prospectively registered, and industry-funded studies. These types of studies were also more likely to be published, suggesting that the increase in publication rate may be due to a shift in the distribution of study types as opposed to a general increase in all research. Furthermore, many bottlenecks remain in the research process and positive results are still disproportionately represented in the published literature. These findings bring us closer to understanding the factors that influence publication rates and are potential avenues for addressing publication bias.

Evidence suggests that preventing non-publication and encouraging timeliness of publication should start even before studies are conducted, as many potentially predictive factors exist as early as study registration. We believe that rigorous guidelines need to be developed and enforced for registry entries and study protocols. Requiring better planning and documentation, in addition to prospective registration, may increase the likelihood of publication and allow for greater transparency in research methodology. Pre-study sample size calculations could also help to prevent underpowered studies. Greater effort is still necessary to explore and address barriers to different types of interventional studies that may have lower publication rates. New policies could be introduced in London, Ontario that require certain types of studies (i.e., non-industry, surgical, etc.) to submit or publish their results within a set timeframe, introducing greater regulatory pressure. This would ensure that clinical studies of all types may reach publication in a timely manner. Another approach to decreasing research waste may instead be increasing the accessibility of unpublished research. Opening up industry registries and simplifying clinical study reports can greatly increase their usage as sources of information. In addition, more consistent reporting in public registries like ClinicalTrials.gov, coupled with greater pressure to submit study results, may pave the way for innovations in automated tools. For example, a system could be developed to automatically produce a meta-analysis of relevant registry outcome data based on a string of parameters, such as intervention and outcome. This would remove several barriers to accessing unpublished data and greatly reduce the need for formal publications.

Although awareness of publication bias is increasing, it is clearly still a prevalent issue. Our findings have produced new insights on the influences of publication, as well as time to publication, that may inform additional measures for increasing publication rates. Not only would addressing publication bias improve the overall quality of clinical research, but it would also reduce bias in the published literature and increase access to important data for clinical decision-making. As a result, we could ultimately prevent research loss, wasted resources, and further patient harm. We hope that our work here can serve as a foundation for future research and a template for other regions or institutions to follow.

Bibliography

Al-Durra, M., Nolan, R. P., Seto, E., Cafazzo, J. A., & Eysenbach, G. (2018). Nonpublication rates and characteristics of registered randomized clinical trials in digital health: Cross-sectional analysis. *Journal of Medical Internet Research*, *20*(12), e11924. https://doi.org/10.2196/11924

Amsen, E. (2015). Publishing negative results and underlying research data [PowerPoint slides]. *F1000Posters*, *6*, 180. https://f1000research.com/slides/1097637

Argyra, E., & Moka, E. (2013). Methodology, publication and retrospective bias. *Regional Anesthesia and Pain Medicine*, *38*(5 SUPPL. 1), E1–E121. https://doi.org/10.1097/AAP.0b013e3182a6a39d

Blümle, A., Antes, G., Schumacher, M., Just, H., & von Elm, E. (2008). Clinical research projects at a German medical faculty: Follow-up from ethical approval to publication and citation by others. *Journal of Medical Ethics*, *34*(9), e20–e20. https://doi.org/10.1136/jme.2008.024521

Blümle, A., Haag, T., Balmford, J., Rücker, G., Schumacher, M., & Binder, N. (2020). A multi-state model analysis of the time from ethical approval to publication of clinical research studies. *PloS One*, *15*(3). https://doi.org/10.1371/JOURNAL.PONE.0230797

Blümle, A., Meerpohl, J. J., Schumacher, M., & von Elm, E. (2014). Fate of clinical research studies after ethical approval – follow-up of study protocols until publication. *PLoS ONE*, *9*(2), e87184. https://doi.org/10.1371/journal.pone.0087184

Blümle, A., Wollmann, K., Bischoff, K., Kapp, P., Lohner, S., Nury, E., Nitschke, K., Zähringer, J., Rücker, G., & Schumacher, M. (2021). Investigator initiated trials versus industry sponsored trials - translation of randomized controlled trials into clinical practice (IMPACT). *BMC Medical Research Methodology*, *21*(1). https://doi.org/10.1186/S12874-021-01359-X

Canada.ca. (2016, June 7). *Health Canada's Clinical Trials Database*. https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/health-canada-clinical-trials-database.html

Chahal, J., Tomescu, S. S., Ravi, B., Bach, B. R., Ogilvie-Harris, D., Mohamed, N. N., & Gandhi, R. (2012). Publication of sports medicine–related randomized controlled trials registered in ClinicalTrials.gov. *The American Journal of Sports Medicine*, *40*(9), 1970–1977. https://doi.org/10.1177/0363546512448363

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *The Lancet*, *374*(9683), 86–89. https://doi.org/10.1016/S0140-6736(09)60329-9

Chan, A.-W. (2008). Bias, spin, and misreporting: Time for full access to trial protocols and results. *PLoS Medicine*, *5*(11), e230. https://doi.org/10.1371/journal.pmed.0050230

Chan, A.-W., Song, F., Vickers, A., Jefferson, T., Dickersin, K., Gøtzsche, P. C., Krumholz, H. M., Ghersi, D., & van der Worp, H. B. (2014). Increasing value and reducing waste: Addressing inaccessible research. *The Lancet*, *383*(9913), 257–266. https://doi.org/10.1016/S0140-6736(13)62296-5

Chan, A.-W., Tetzlaff, J. M., Gotzsche, P. C., Altman, D. G., Mann, H., Berlin, J. A., Dickersin, K., Hrobjartsson, A., Schulz, K. F., Parulekar, W. R., Krleza-Jeric, K., Laupacis, A., & Moher, D. (2013). SPIRIT 2013 explanation and elaboration: Guidance for protocols of clinical trials. *BMJ*, *346*, e7586–e7586. https://doi.org/10.1136/bmj.e7586

Chen, R., Desai, N., Ross, J., Zhang, W., Chau, K., Wayda, B., Murugiah, K., Lu, D., Mittal, A., & Krumholz, H. (2016). Publication and reporting of clinical trial results: Cross sectional analysis across academic medical centers. *BMJ (Clinical Research Ed.)*, *352*. https://doi.org/10.1136/BMJ.I637

Chiu, K., Grundy, Q., & Bero, L. (2017). "Spin" in published biomedical literature: A methodological systematic review. *PLoS Biology*, *15*(9). https://doi.org/10.1371/JOURNAL.PBIO.2002173

ClinRegs. (2021, August 6). *Clinical research regulation for Canada*. https://clinregs.niaid.nih.gov/country/canada#_top

Collins, F. S., & Burwell, S. M. (2016). Clinical trials registration and results information submission. *Federal Register*, *81*(183), 64982–65157. https://www.federalregister.gov/documents/2016/09/21/2016-22129/clinical-trials-registration-and-results-information-submission

Coronado-Montoya, S., Levis, A. W., Kwakkenbos, L., Steele, R. J., Turner, E. H., & Thombs, B. D. (2016). Reporting of positive results in randomized controlled trials of mindfulness-based mental health interventions. *PLOS ONE*, *11*(4), e0153220. https://doi.org/10.1371/journal.pone.0153220

De Angelis, C., Drazen, J. M., Frizelle, F. A., Haug, C., Hoe, J., Horton, R., Kotzin, S., Laine, C., Marusic, A., Overbeke, A. J. P. M., Schroeder, T. V., Sox, H. C., & Van Der Weyden, M. B. (2004, September). *Clinical trial registration: A statement from the International Committee of Medical Journal Editors*. International Committee of

Medical Journal Editors. http://www.icmje.org/news-and-editorials/clin_trial_sep2004.pdf

de Vries, Y. A., Roest, A. M., de Jonge, P., Cuijpers, P., Munafò, M. R., & Bastiaansen, J. A. (2018). The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: The case of depression. *Psychological Medicine*, *48*(15), 2453–2455. https://doi.org/10.1017/S0033291718001873

DeVito, N. J., Bacon, S., & Goldacre, B. (2018). FDAAA TrialsTracker: A live informatics tool to monitor compliance with FDA requirements to report clinical trial results. *BioRxiv*, 266452. https://doi.org/10.1101/266452

DeVito, N. J., & Goldacre, B. (2019). Catalogue of bias: Publication bias. *BMJ Evidence-Based Medicine*, *24*(2), 53–54. https://doi.org/10.1136/bmjebm-2018-111107

Dickersin, K., & Chalmers, I. (2011). Recognizing, investigating and dealing with incomplete and biased reporting of clinical research: From Francis Bacon to the WHO. *Journal of the Royal Society of Medicine*, *104*(12), 532–538. https://doi.org/10.1258/jrsm.2011.11k042

Dickersin, K., & Min, Y. I. (1993). NIH clinical trials and publication bias. *The Online Journal of Current Clinical Trials*, *Doc No 50*. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med3&NEWS=N&AN=8306005

Doshi, P., Dickersin, K., Healy, D., Vedula, S. S., & Jefferson, T. (2013). Restoring invisible and abandoned trials: A call for people to publish the findings. *BMJ*, *346*, f2865–f2865. https://doi.org/10.1136/bmj.f2865

Doshi, P., & Jefferson, T. (2013). Clinical study reports of randomised controlled trials: An exploratory review of previously confidential industry reports. *BMJ Open*, *3*(2), e002496. https://doi.org/10.1136/bmjopen-2012-002496

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A.-W., Cronin, E., Decullier, E., Easterbrook, P. J., Von Elm, E., Gamble, C., Ghersi, D., Ioannidis, J. P. A., Simes, J., & Williamson, P. R. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE*, *3*(8), e3081. https://doi.org/10.1371/journal.pone.0003081

EBM DataLab. (2018). *FDAAA TrialsTracker*. https://fdaaa.trialstracker.net/

Evers, J. L. H. (2015). Towards prospective trial registration in Human Reproduction: Step 2 of 3. *Human Reproduction*, *30*(7), 1517–1518. https://doi.org/10.1093/humrep/dev092

FDA. (2018, August 2). *FDA drug safety communication: Updated Risk Evaluation and Mitigation Strategy (REMS) to restrict access to rosiglitazone-containing medicines including Avandia, Avandamet, and Avandaryl*. https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-updated-risk-evaluation-and-mitigation-strategy-rems-restrict-access

Fuentes Camps, I., Rodríguez, A., & Agustí, A. (2018). Non-commercial vs. commercial clinical trials: A retrospective study of the applications submitted to a research ethics committee. *British Journal of Clinical Pharmacology*, *84*(6), 1384–1388. https://doi.org/10.1111/bcp.13555

Ganga, V., & Edupuganti, M. (2017). Publication timeliness of the randomized controlled trials of drug therapy for inflammatory arthritis. *Arthritis and Rheumatology*, *69*(Supplement 10). http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed18&NEWS=N&AN=618912440

Gilbody, S. M., & Song, F. (2000). Publication bias and the integrity of psychiatry research. *Psychological Medicine*, *30*(2), 253–258. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medc&NEWS=N&AN=10824646

Glasziou, P. (2005). *The epidemiology of ignorance*. Evidence-Based Health Care International Conference. https://www.ebhcconference.org/previous_editions/2005/presentations/4-11/1.Glasziou Paul.pdf

Goldacre, B., Drysdale, H., Dale, A., Milosevic, I., Slade, E., Hartley, P., Marston, C., Powell-Smith, A., Heneghan, C., & Mahtani, K. R. (2019). COMPare: A prospective cohort study correcting and monitoring 58 misreported trials in real time. *Trials*, *20*(1), 118. https://doi.org/10.1186/s13063-019-3173-2

Hawe, P. (2012). The truth, but not the whole truth? Call for an amnesty on unreported results of public health interventions. *Journal of Epidemiology and Community Health*, *66*(4), 285–285. https://doi.org/10.1136/jech.2011.140350

Hodgson, R., Allen, R., Broderick, E., Bland, J., Dumville, J. C., Ashby, R., Bell-Syer, S., Foxlee, R., Hall, J., Lamb, K., Madden, M., O'Meara, S., Stubbs, N., & Cullum, N. (2014). Funding source and the quality of reports of chronic wounds trials: 2004 to 2011. *Trials*, *15*(1), 19. https://doi.org/10.1186/1745-6215-15-19

Hopewell, S., Clarke, M. J., Stewart, L., & Tierney, J. (2007). Time to publication for results of clinical trials. *Cochrane Database of Systematic Reviews*, *2010*(1). https://doi.org/10.1002/14651858.MR000011.pub2

Horsley, J., Wade, S., & Byrne, M. (2015, February 25). *Positively negative: A new PLOS ONE collection focusing on negative, null and inconclusive results*. PLOS Collections. https://collectionsblog.plos.org/positively-negative-new-plos-one-collection-focusing-negative-null-inconclusive-results-everyone/

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Jia, Y., Huang, D., Wen, J., Liang, J., Qureshi, R., Wang, Y., Rosman, L., Chen, Q., Robinson, K. A., Gagnier, J. J., Celentano, D. D., & Ehrhardt, S. (2021). Association between switching of primary outcomes and reported trial findings among randomized drug trials from China. *Journal of Clinical Epidemiology*, *132*, 10–17. https://doi.org/10.1016/J.JCLINEPI.2020.11.023

Jones, C. W., Handler, L., Crowell, K. E., Keil, L. G., Weaver, M. A., & Platts-Mills, T. F. (2013). Non-publication of large randomized clinical trials: Cross sectional analysis. *BMJ*, *347*, f6104–f6104. https://doi.org/10.1136/bmj.f6104

Jones, P. M. (2016). Publication bias in the anesthesiology literature: Shifting the focus from the "positive" to the "truth." *Canadian Journal of Anesthesia/Journal Canadien d'anesthésie*, *63*(6), 658–663. https://doi.org/10.1007/s12630-016-0632-z

Jones, P. M., Chow, J., Arango, M., Fridfinnson, J., Gai, N., Lam, K., & Turkstra, T. (2017). Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals. *Anesthesia and Analgesia*, *125*(4), 1292–1300. https://doi.org/10.1213/ANE.0000000000002272

Jørgensen, L., Gøtzsche, P. C., & Jefferson, T. (2018). The Cochrane HPV vaccine review was incomplete and ignored important evidence of bias. *BMJ Evidence-Based Medicine*, *23*(5), 165–168. https://doi.org/10.1136/bmjebm-2018-111012

Juni, P. (2002). Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ*, *324*(7349), 1287–1288. https://doi.org/10.1136/bmj.324.7349.1287

Kaiser, K. A., Cofield, S. S., Fontaine, K. R., Glasser, S. P., Thabane, L., Chu, R., Ambrale, S., Dwary, A. D., Kumar, A., Nayyar, G., Affuso, O., Beasley, M., & Allison, D. B. (2012). Is funding source related to study reporting quality in obesity or nutrition randomized control trials in top-tier medical journals? *International Journal of Obesity*, *36*(7), 977–981. https://doi.org/10.1038/ijo.2011.207

Kasenda, B., von Elm, E., You, J., Blümle, A., Tomonaga, Y., Saccilotto, R., Amstutz, A., Bengough, T., Meerpohl, J. J., Stegert, M., Tikkinen, K. A. O., Neumann, I., Carrasco-Labra, A., Faulhaber, M., Mulla, S. M., Mertz, D., Akl, E. A., Bassler, D., Busse, J. W., … Briel, M. (2014). Prevalence, characteristics, and publication of

discontinued randomized trials. *JAMA*, *311*(10), 1045. https://doi.org/10.1001/jama.2014.1361

Khan, N. A., Singh, M., Spencer, H. J., & Torralba, K. D. (2014). Randomized controlled trials of rheumatoid arthritis registered at ClinicalTrials.gov: What gets published and when. *Arthritis & Rheumatology*, *66*(10), 2664–2674. https://doi.org/10.1002/art.38784

Kirkham, J. J., Dwan, K. M., Blümle, A., von Elm, E., & Williamson, P. R. (2016). How much participant outcome data is missing from sight: Findings from a cohort of trials submitted to a German research ethics committee. *PLOS ONE*, *11*(6), e0157883. https://doi.org/10.1371/journal.pone.0157883

Klassen, T. P., Wiebe, N., Russell, K., Stevens, K., Hartling, L., Craig, W. R., & Moher, D. (2002). Abstracts of randomized controlled trials presented at the Society for Pediatric Research meeting. *Archives of Pediatrics & Adolescent Medicine*, *156*(5), 474. https://doi.org/10.1001/archpedi.156.5.474

Lau, J., Antman, E. M., Jimenez-Silva, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (2010). Cumulative meta-analysis of therapeutic trials for myocardial infarction. *Http://Dx.Doi.Org/10.1056/NEJM199207233270406*, *327*(4), 248–254. https://doi.org/10.1056/NEJM199207233270406

Lensen, S., Jordon, V., Showell, M., Showell, E., Shen, V., Venetis, C., & Farquhar, C. (2015). Publication bias in the field of subfertility. *Human Reproduction*, *30*(SUPPL. 1). https://doi.org/10.1093/humrep/30.Supplement-1.1

Li, T., Boutron, I., Salman, R. A.-S., Cobo, E., Flemyng, E., Grimshaw, J. M., & Altman, D. G. (2017). Review and publication of protocol submissions to Trials – what have we learned in 10 years? *Trials*, *18*(1), 34. https://doi.org/10.1186/s13063-016-1743-0

Loke, Y. K., Kwok, C. S., & Singh, S. (2011). Comparative cardiovascular effects of thiazolidinediones: Systematic review and meta-analysis of observational studies. *BMJ*, *342*, d1309–d1309. https://doi.org/10.1136/bmj.d1309

Lundh, A., Krogsbøll, L. T., & Gøtzsche, P. C. (2012). Sponsors' participation in conduct and reporting of industry trials: A descriptive study. *Trials*, *13*(1), 146. https://doi.org/10.1186/1745-6215-13-146

McCormack, J. P., & Rangno, R. (2002). Digging for data from the COX-2 trials. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *166*(13), 1649–1650. http://www.ncbi.nlm.nih.gov/pubmed/12126315

Menzel, S., Uebing, B., Hucklenbroich, P., & Schober, O. (2007). Evaluation von Forschungsvorhaben nach einem positiven Votum einer Ethikkommission. *DMW -*

Deutsche Medizinische Wochenschrift, *132*(44), 2313–2317. https://doi.org/10.1055/s-2007-991648

Moreno, S. G., Sutton, A. J., Turner, E. H., Abrams, K. R., Cooper, N. J., Palmer, T. M., & Ades, A. E. (2009). Novel methods to deal with publication biases: Secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications. *BMJ*, *339*, b2981–b2981. https://doi.org/10.1136/bmj.b2981

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E. J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour 2017 1:1*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Niebling, W. (2011). Evidenztransfer in die Praxis – Hindernisse und Barrieren. *Zeitschrift Für Evidenz, Fortbildung Und Qualität Im Gesundheitswesen*, *105*(9), 646–651. https://doi.org/10.1016/j.zefq.2011.10.027

Nissen, S. E., & Wolski, K. (2007). Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *New England Journal of Medicine*, *356*(24), 2457–2471. https://doi.org/10.1056/NEJMoa072761

NLM. (n.d.). *ClinicalTrials.gov*. Retrieved July 13, 2020, from https://clinicaltrials.gov/

OHRE. (2016, May 26). *Clinical trials registration requirements*. Western University. https://applywesternrem.uwo.ca/Personalisation/DownloadTemplate/47

Olson, C. M. (2002). Publication bias in editorial decision making. *JAMA*, *287*(21), 2825. https://doi.org/10.1001/jama.287.21.2825

Pang, C., & Loke, Y. (2011). Publication bias: Extent of unreported outcomes in conference abstracts and journal articles when compared with the full report in the GSK trials register. *British Journal of Clinical Pharmacology*, *71*(6), 971–994. https://doi.org/10.1111/j.1365-2125.2011.03930.x

Pica, N., & Bourgeois, F. (2016). Discontinuation and nonpublication of randomized clinical trials conducted in children. *PEDIATRICS*, *138*(3), e20160223–e20160223. https://doi.org/10.1542/peds.2016-0223

Potthast, R., Vervolgyi, V., McGauran, N., Kerekes, M. F., Wieseler, B., Kaiser, T., R., P., V., V., N., M., M.F., K., & B., W. (2014). Impact of inclusion of industry trial results registries as an information source for systematic reviews. *PloS One*, *9*(4), e92067–e92067. https://doi.org/10.1371/journal.pone.0092067

Prayle, A. P., Hurley, M. N., & Smyth, A. R. (2012). Compliance with mandatory reporting of clinical trial results on ClinicalTrials.gov: Cross sectional study. *BMJ*, *344*, d7373–d7373. https://doi.org/10.1136/bmj.d7373

Ramagopalan, S., Skingsley, A. P., Handunnetthi, L., Klingel, M., Magnus, D., Pakpoor, J., & Goldacre, B. (2014). Prevalence of primary outcome changes in clinical trials registered on ClinicalTrials.gov: A cross-sectional study. *F1000Research, 3*. https://doi.org/10.12688/F1000RESEARCH.3784.1

Rezaeian, M. (2016). Reducing publication bias in biomedical research: Reviewing and registering protocol with a suitable journal. *Journal of Clinical Epidemiology*, *69*, 248–249. https://doi.org/10.1016/j.jclinepi.2015.05.007

Richards, S. M., & Burrett, J. A. (2013). A proposal for reducing the effect of one of many causes of publication bias. *Trials*, *14*(1), 41. https://doi.org/10.1186/1745-6215-14-41

Rivara, F. P., & Cummings, P. (2002). Publication bias: The problem and some suggestions. *Archives of Pediatrics & Adolescent Medicine*, *156*(5), 424–425. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=11980544

Rosenthal, R., Kasenda, B., Dell-Kuster, S., von Elm, E., You, J., Blümle, A., Tomonaga, Y., Saccilotto, R., Amstutz, A., Bengough, T., Meerpohl, J. J., Stegert, M., Tikkinen, K. A. O., Neumann, I., Carrasco-Labra, A., Faulhaber, M., Mulla, S., Mertz, D., Akl, E. A., … Briel, M. (2015). Completion and publication rates of randomized controlled trials in surgery. *Annals of Surgery*, *262*(1), 68–73. https://doi.org/10.1097/SLA.0000000000000810

Ross, J. S., Tse, T., Zarin, D. A., Xu, H., Zhou, L., & Krumholz, H. M. (2012). Publication of NIH funded trials registered in ClinicalTrials.gov: Cross sectional analysis. *BMJ*, *344*(7838). https://doi.org/10.1136/BMJ.D7292

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). Publication bias in meta-analysis: Prevention, assessment and adjustments. In *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. John Wiley and Sons. https://doi.org/10.1002/0470870168

Saldanha, I. J., Dickersin, K., Wang, X., & Li, T. (2014). Outcomes in Cochrane systematic reviews addressing four common eye conditions: An evaluation of completeness and comparability. *PLoS One*, *9*(10). https://doi.org/10.1371/JOURNAL.PONE.0109400

Schmucker, C., Schell, L. K., Portalupi, S., Oeller, P., Cabrera, L., Bassler, D., Schwarzer, G., Scherer, R. W., Antes, G., von Elm, E., & Meerpohl, J. J. (2014). Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. *PLoS ONE*, *9*(12), e114023. https://doi.org/10.1371/journal.pone.0114023

Schuit, E., & Ioannidis, J. P. (2016). Network meta-analyses performed by contracting companies and commissioned by industry. *Systematic Reviews*, *5*(1), 198. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med12&NEWS=N&AN=27884175

Shuchman, M. (2013). Health Canada's new clinical trials database should be mandatory, says expert. *Canadian Medical Association Journal*, *185*(11), 946–946. https://doi.org/10.1503/cmaj.109-4526

Silverstein, F. E., Faich, G., Goldstein, J. L., Simon, L. S., Pincus, T., Whelton, A., Makuch, R., Eisen, G., Agrawal, N. M., Stenson, W. F., Burr, A. M., Zhao, W. W., Kent, J. D., Lefkowith, J. B., Verburg, K. M., & Geis, G. S. (2000). Gastrointestinal toxicity with Celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis. *JAMA*, *284*(10), 1247. https://doi.org/10.1001/jama.284.10.1247

Skeldon, S., & Goldenberg, L. (2015). The methodological quality and publication of industry versus non-industry trials for testosterone treatment. *Journal of Urology*, *193*(4 SUPPL. 1), e104–e105. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed16&NEWS=N&AN=71858037

Smith, E. R. (2004). The importance of negative and neutral clinical trials. *The Canadian Journal of Cardiology*, *20*(12), 1267–1268. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=15494781

Song, F., Parekh, S., Hooper, L., Loke, Y., Ryder, J., Sutton, A., Hing, C., Kwok, C., Pang, C., & Harvey, I. (2010). Dissemination and publication of research findings: An updated review of related biases. *Health Technology Assessment*, *14*(8). https://doi.org/10.3310/hta14080

Strand, L. B., Clarke, P., Graves, N., & Barnett, A. G. (2017). Time to publication for publicly funded clinical trials in Australia: An observational study. *BMJ Open*, *7*(3), e012212. https://doi.org/10.1136/bmjopen-2016-012212

Strech, D. (2012). Normative arguments and new solutions for the unbiased registration and publication of clinical trials. *Journal of Clinical Epidemiology*, *65*(3), 276–281. https://doi.org/10.1016/j.jclinepi.2011.07.002

Takeda, A., Loveman, E., Harris, P., Hartwell, D., & Welch, K. (2008). Time to full publication of studies of anti-cancer medicines for breast cancer and the potential for publication bias: A short systematic review. *Health Technology Assessment (Winchester, England)*, *12*(32), iii–46. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med6&NEWS=N&AN=18831948

Texier, G., Meynard, J.-B., Michel, R., Migliani, R., & Boutin, J.-P. (2007). Should we try to publish a study that concludes nothing? *The Medical Press*, *36*(3), 541–550. https://doi.org/10.1016/j.lpm.2007.01.015

Tfelt-Hansen, P. C. (2009). Unpublished clinical trials with sumatriptan. *The Lancet*, *374*(9700), 1501–1502. https://doi.org/10.1016/S0140-6736(09)61906-1

Thaler, K., Kien, C., Nussbaumer, B., Van Noord, M. G., Griebler, U., Klerings, I., & Gartlehner, G. (2015). Inadequate use and regulation of interventions against publication bias decreases their effectiveness: A systematic review. *Journal of Clinical Epidemiology*, *68*(7), 792–802. https://doi.org/10.1016/j.jclinepi.2015.01.008

Timmer, A., Hilsden, R. J., Cole, J., Hailey, D., & Sutherland, L. R. (2002). Publication bias in gastroenterological research – a retrospective cohort study based on abstracts submitted to a scientific meeting. *BMC Medical Research Methodology*, *2*(1), 7. https://doi.org/10.1186/1471-2288-2-7

Tsujimoto, Y., Tsutsumi, Y., Kataoka, Y., Tsujimoto, H., Yamamoto, Y., Papola, D., Guyatt, G. H., Fukuhara, S., & Furukawa, T. A. (2019). Statistical significance did not affect time to publication in non-Cochrane systematic reviews: A metaepidemiological study. *Journal of Clinical Epidemiology*, *115*, 25–34. https://doi.org/10.1016/j.jclinepi.2019.06.015

Tuma, R. S. (2010). New law may be having some effect on publication bias. *JNCI Journal of the National Cancer Institute*, *102*(5), 290–292. https://doi.org/10.1093/jnci/djq051

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*, *358*(3), 252–260. https://doi.org/10.1056/NEJMsa065779

UCLA. (n.d.). *Regression with Stata chapter 2 – regression diagnostics*. Retrieved August 19, 2021, from https://stats.idre.ucla.edu/stata/webbooks/reg/chapter2/stata-webbooksregressionwith-statachapter-2-regression-diagnostics/

Vines, G. (1997). Publish or be damned. *New Scientist (1971)*, *155*(2099), 53. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=16127834

Viteri-García, A., DeVito, N. J., & Goldacre, B. (2018). Too few trials or too few reported trials? *Evidence-Based Dentistry*, *19*(3), 67–68. https://doi.org/10.1038/sj.ebd.6401317

von Elm, E., Rollin, A., Blümle, A., Huwiler, K., Witschi, M., & Egger, M. (2008). Publication and non-publication of clinical trials: Longitudinal study of applications submitted to a research ethics committee. *Swiss Medical Weekly*, *138*(13–14), 197–203. https://doi.org/2008/13/smw-12027

Weiser, M., Haim, A., Ross, J., Laughren, T., Kalali, A., DeLisi, L., & Krystal, J. (2016). Improving transparency of clinical trial data. *Neuropsychopharmacology*, *41*(S1), S1–S115. https://doi.org/10.1038/npp.2016.239

Wieseler, B., Wolfram, N., McGauran, N., Kerekes, M. F., Vervölgyi, V., Kohlepp, P., Kamphuis, M., & Grouven, U. (2013). Completeness of reporting of patient-relevant clinical trial outcomes: Comparison of unpublished clinical study reports with publicly available data. *PLoS Medicine*, *10*(10), e1001526. https://doi.org/10.1371/journal.pmed.1001526

Wolters Kluwer. (n.d.). *Ovid*. Retrieved July 31, 2020, from https://ovidsp.ovid.com/

Yordanov, Y., Dechartres, A., Atal, I., Tran, V.-T., Boutron, I., Crequit, P., & Ravaud, P. (2018). Avoidable waste of research related to outcome planning and reporting in clinical trials. *BMC Medicine*, *16*(1), 87. https://doi.org/10.1186/s12916-018-1083-x

Zanchetti, A., & Mancia, G. (1998). Searching for information from unreported trials--amnesty for the past and prospective meta-analyses for the future. *Journal of Hypertension*, *16*(2), 125. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medc&NEWS=N&AN=9535137

Zou, G. (2004). A modified poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology*, *159*(7), 702–706. https://doi.org/10.1093/AJE/KWH090

Zucker, K., Pengel, L. H., & Liu, L. Q. (2012). The rate of publication of abstracts in solid organ transplantation presented at scientific meetings. *Transplantation*, *94*(SUPPL. 10S), 246. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed13&NEWS=N&AN=71250262

Appendices

Appendix A: Literature Review Systematic Search Strategy and Results

Key search terms: publication bias, outcome reporting bias, registries, ethics submissions, protocols, RCTs

MEDLINE search strategy:

1. Exp Publication Bias/
2. (publication bias or unreport* or incomplete report* or “not reported” or outcome omission or partial report* or outcome suppress* or non*publication or report* bias).ti,ab,tw.
3. 1 or 2
4. Exp Registries/
5. (register* protocol or register* trial or register* study or register* method or trial register* or ethics submission* or ethics approval or ethics application*).ti,ab,tw.
6. 4 or 5
7. Exp Clinical Protocols/
8. Protocol$.ti,ab,tw.
9. 7 or 8
10. Exp Randomized Controlled Trials as Topic/
11. (RCT or randomi#ed controlled trial*).ti,ab,tw.
12. 10 or 11
13. 6 or 9
14. 3 and 12 and 13

EMBASE search strategy:

1. Exp publishing/
2. (publication bias or unreport* or incomplete report* or “not reported” or outcome omission or partial report* or outcome suppress* or non*publication or report* bias).ti,ab,tw.
3. 1 or 2

4. Exp register/

5. (register* protocol or register* trial or register* study or register* method or trial register* or ethics submission* or ethics approval or ethics application*).ti,ab,tw.

6. 4 or 5

7. Exp clinical protocol/

8. Protocol$.ti,ab,tw.

9. 7 or 8

10. Exp "randomized controlled trial (topic)"/

11. (RCT or randomi#ed controlled trial*).ti,ab,tw.

12. 10 or 11

13. 6 or 9

14. 3 and 12 and 13

Ovid search of MEDLINE (n = 475) and EMBASE (n = 714) – Oct. 16, 2019

Total results after removing duplicates: 1007

Level 1 Exclusion: 842

Level 1 Inclusion: 165 + 20 additional papers found

Level 2 Exclusion: 170

Level 2 Inclusion: 15 primary studies on RCTs – Dec. 14, 2019

165 of 1007 results from the systematic search were included for final screening, with 20 additional papers found from manual and forward/backward citation searching. A total of 15 relevant primary studies were included.

Appendix B: Supplementary Data and Calculations for Descriptive Statistics

Table B1

Paired t-test *between study publication with any results and publication with primary outcome (*p < *0.001)*

```
. ttest publish == publishprimary

Paired t test
```

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
publish	2,446	.3781684	.0098071	.4850291	.3589374	.3973995
publis~y	2,446	.3037612	.0093005	.4599748	.2855236	.3219989
diff	2,446	.0744072	.0053074	.2624861	.0639998	.0848146

```
    mean(diff) = mean(publish - publishprimary)                   t =  14.0196
Ho: mean(diff) = 0                                     degrees of freedom =     2445

Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

Table B2

*Exact McNemar's Chi^2 test between study publication with any results and publication with primary outcome (*p *< 0.001)*

```
. mcc publish publishprimary
```

Cases	Controls Exposed	Unexposed	Total
Exposed	743	182	925
Unexposed	0	1521	1521
Total	743	1703	2446

```
McNemar's chi2(1) =    182.00    Prob > chi2 = 0.0000
Exact McNemar significance probability       = 0.0000

Proportion with factor
        Cases       .3781684
        Controls    .3037612     [95% Conf. Interval]
                    ---------    --------------------
        difference  .0744072      .0635983   .0852161
        ratio       1.244953      1.205872     1.2853
        rel. diff.  .1068702       .092197   .1215435

        odds ratio          .     48.83916          .   (exact)
```

Both *paired t-test* (Table B1) and McNemar's Chi^2 test (Table B2) showed a statistically significant difference between the publication outcome measures (both with $p < 0.001$). The McNemar's Chi^2 test should be robust to structural zeros due to its paired nature, our use of the exact form, and our large sample size.

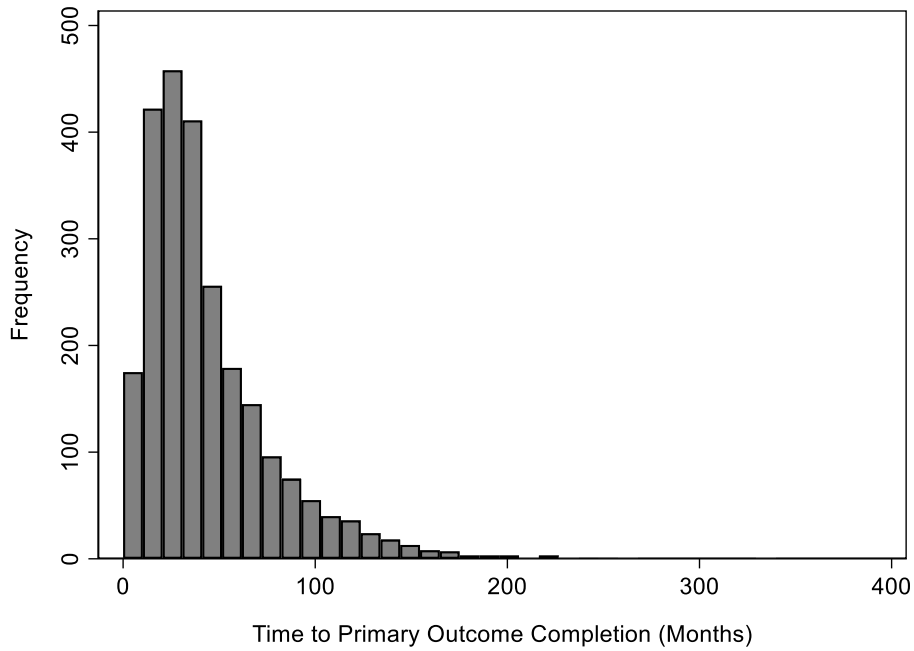


Figure B1. Histogram of time to primary outcome variable.

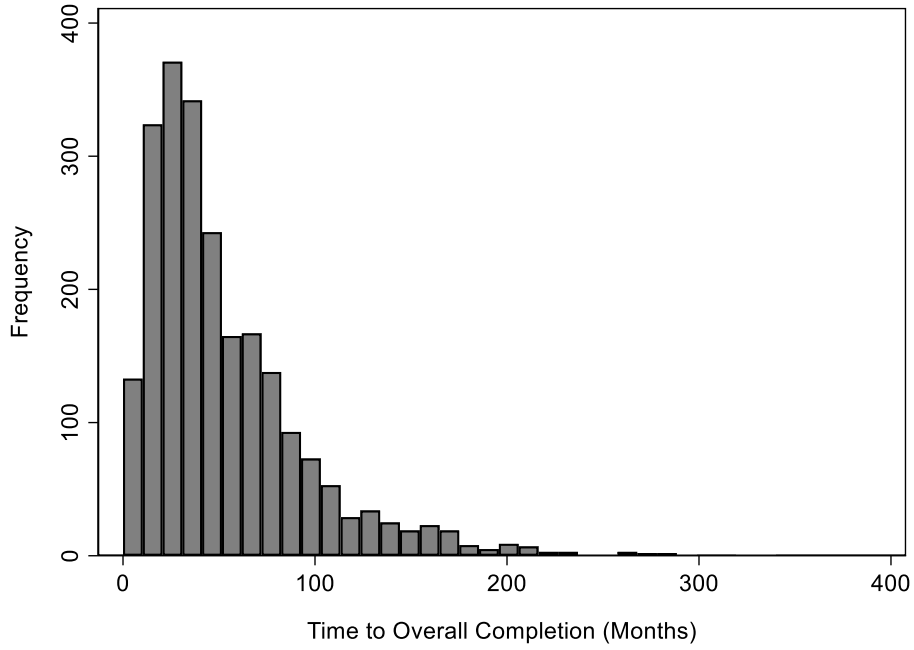


Figure B2. Histogram of time to overall completion variable.

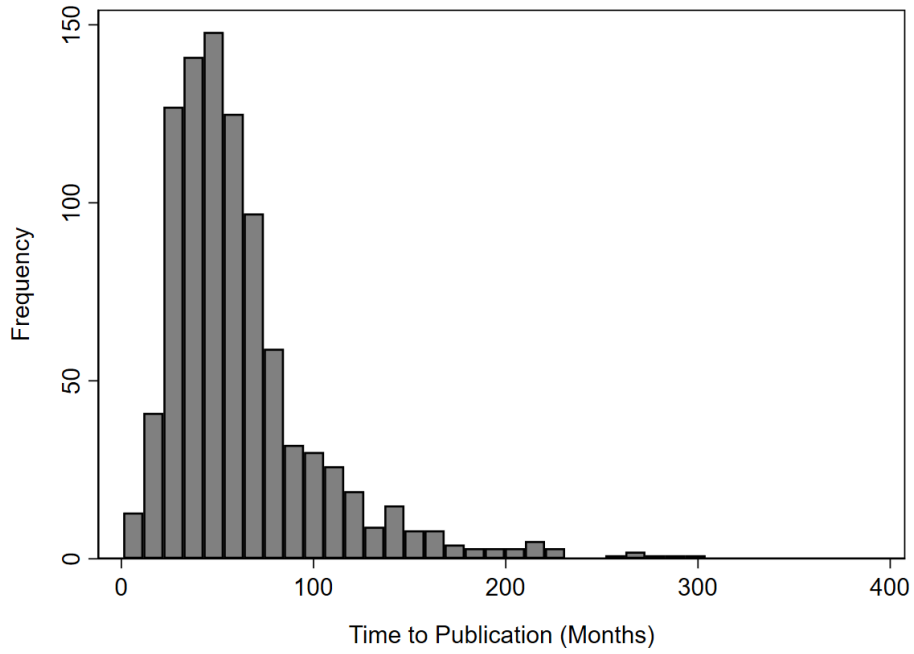


Figure B3. Histogram of time to publication variable.

Histograms for time to primary outcome (Figure B1), time to overall completion (Figure B2), and time to publication (Figure B3) demonstrated clear right skew.

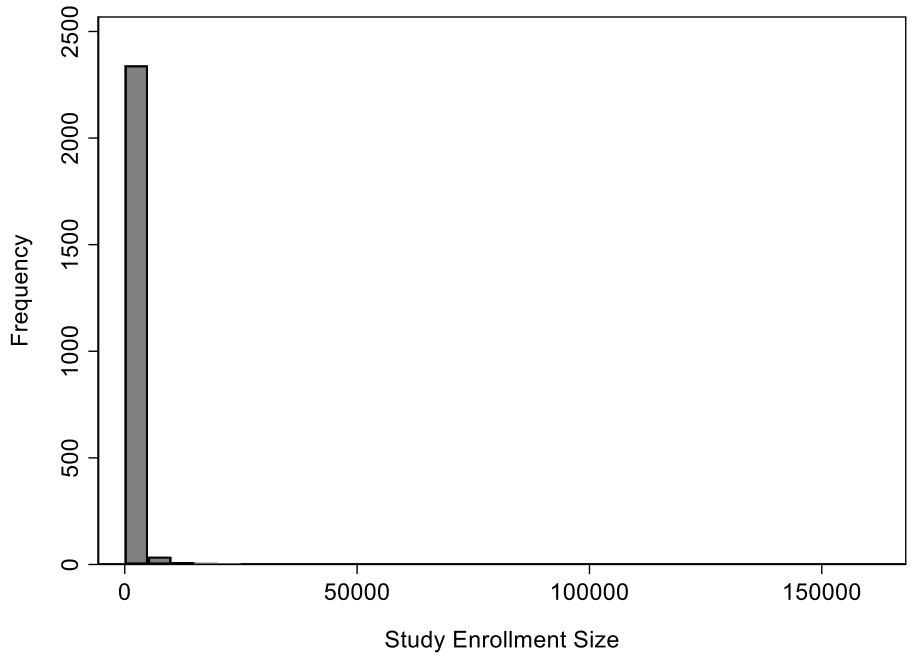


Figure B4. Histogram of entire enrollment size variable.

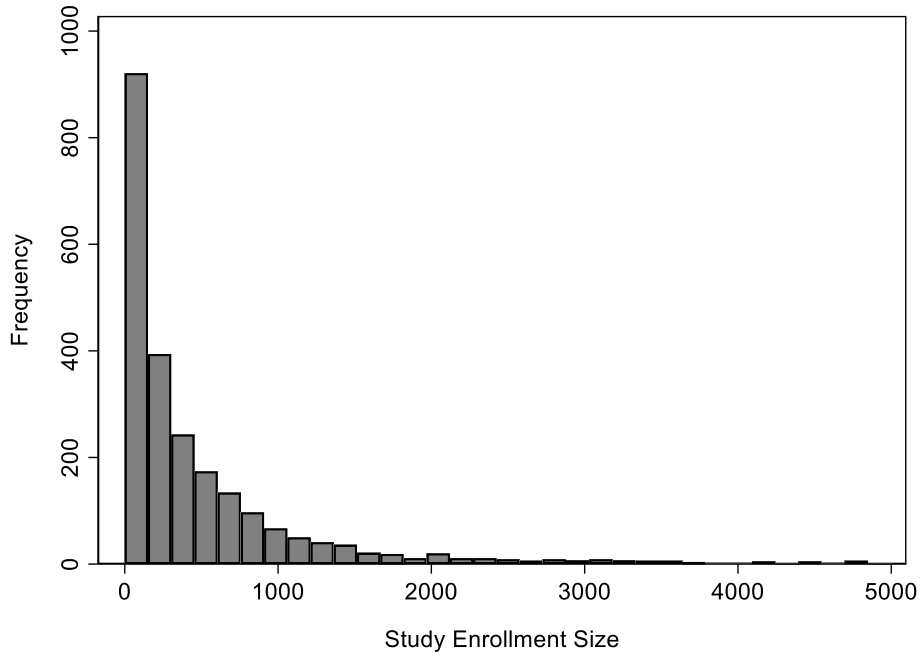


Figure B5. Histogram of enrollment size variable up to 5000.

Due to the extreme range of enrollment sizes (Figure B4), a truncated version was generated to provide a clearer view (Figure B5). Both demonstrated a clear right skew.

Table B3

Annual average enrollment sizes of registered studies with randomized versus non-randomized design or either, by year of first participant enrolled

Year	Mean Enrollment Size		
	Randomized Design	**Non-Randomized Design**	**Either**
1983	1441	-	1441
1988	-	573	573
1989	154	-	154
1992	9000	-	9000
1993	953	10	639
1994	2000	-	2000
1995	1302	1534	1354
1996	314	-	314
1997	4455	151	2542
1998	1426	201	1108
1999	496	324	460
2000	1234	124	852
2001	842	6356	1629

2002	602	116	505
2003	1441	912	1399
2004	542	627	550
2005	627	364	613
2006	1103	668	1053
2007	822	255	745
2008	962	254	851
2009	857	195	736
2010	1372	441	1217
2011	1921	257	1617
2012	530	211	477
2013	1361	246	1191
2014	1142	515	1022
2015	769	229	666
2016	1111	187	957
2017	1676	284	1484
Overall	**1088**	**394**	**980**

Appendix C: Supplementary Data and Calculations for Comparative Analysis

Table C1

*Pairwise Pearson's coefficient between study publication and enrollment size (*p < *0.05)*

```
. pwcorr publish enrollment, sig star(.05)

             | publish enroll~t
-------------+------------------
     publish |   1.0000
             |
  enrollment |   0.1072*  1.0000
             |   0.0000
```

Table C2

*Point-biserial correlation for binary publication variable and enrollment size (*p < *0.001)*

```
. pbis publish enrollment

(obs= 2410)
Np= 922  p= 0.38
Nq= 1488  q= 0.62
------------------+------------------+------------------+------------------+
Coef.= 0.1072          t= 5.2905        P>|t| = 0.0001        df=   2408
```

Both Pearson coefficient (Table C1) and point-biserial correlation (Table C2) showed a statistically significant positive correlation ($p < 0.05$ and $p < 0.001$, respectively) between publication and enrollment size.

Table C3

Multivariable log-binomial regression model for association between study characteristics (randomized design, prospective registration, industry funding, surgical study, and drug study) and study publication

```
. glm publish rct prospective industry1 surgical drug, fam(bin) link(log) nolog
> eform

Generalized linear models                         Number of obs   =      2,446
Optimization     : ML                             Residual df     =      2,440
                                                  Scale parameter =          1
Deviance         =  3081.709711                   (1/df) Deviance =   1.262996
Pearson          =  2452.293273                   (1/df) Pearson  =   1.005038

Variance function: V(u) = u*(1-u)                 [Bernoulli]
Link function    : g(u) = ln(u)                   [Log]

                                                  AIC             =   1.264804
Log likelihood   = -1540.854856                   BIC             =  -15955.68
```

publish	Risk Ratio	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
rct	1.456379	.1290497	4.24	0.000	1.224192	1.732604
prospective	1.14669	.0570195	2.75	0.006	1.040207	1.264073
industry1	1.668885	.1191111	7.18	0.000	1.451025	1.919456
surgical	.864594	.102476	-1.23	0.220	.6853676	1.090689
drug	1.255393	.0916781	3.11	0.002	1.087974	1.448574
_cons	.150844	.0166006	-17.19	0.000	.1215772	.187156

Note: _cons estimates baseline risk.

```
. vif, uncentered
```

Variable	VIF	1/VIF
rct	3.68	0.271419
drug	3.48	0.287111
industry1	3.25	0.307768
prospective	1.81	0.551338
surgical	1.20	0.834972
Mean VIF	2.69	

Table C4

Multivariable log-binomial regression model for association between study characteristics (randomized design, prospective registration, industry funding, surgical study, and drug study) and study publication with primary outcome

```
. glm publishprimary rct prospective industry1 surgical drug, fam(bin) link(log)
>  nolog eform

Generalized linear models                         Number of obs   =      2,446
Optimization     : ML                             Residual df     =      2,440
                                                  Scale parameter =          1
Deviance         =  2863.582781                   (1/df) Deviance =     1.1736
Pearson          =  2453.288745                   (1/df) Pearson  =   1.005446

Variance function: V(u) = u*(1-u)                 [Bernoulli]
Link function    : g(u) = ln(u)                   [Log]

                                                  AIC             =   1.175627
Log likelihood   = -1431.791391                   BIC             =  -16173.81
```

publishprim~y	OIM Risk Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
rct	1.495059	.155692	3.86	0.000	1.219035	1.833583
prospective	1.196195	.0711832	3.01	0.003	1.064507	1.344173
industry1	1.847975	.1586211	7.15	0.000	1.561828	2.186548
surgical	.8159375	.1174736	-1.41	0.158	.6153273	1.081951
drug	1.155279	.0961768	1.73	0.083	.9813508	1.360032
_cons	.1145942	.0148594	-16.71	0.000	.0888766	.1477536

Note: _cons estimates baseline risk.

```
. vif, uncentered
```

Variable	VIF	1/VIF
rct	3.68	0.271419
drug	3.48	0.287111
industry1	3.25	0.307768
prospective	1.81	0.551338
surgical	1.20	0.834972
Mean VIF	2.69	

Log-binomial regression models for publication (Table C3) and publication with primary outcome (Table C4) did not show evidence of multicollinearity (VIF < 10).

Table C5

Variance inflation factors of log-binomial regression covariates (randomized design, prospective registration, industry funding, surgical study, and drug study), for association with study publication, including all outcome description criteria (metric, measurement, and timepoint)

```
. qui glm publish metric measurement timepoint rct prospective industry1 surgical
>  drug if outcomementioned == 1, fam(bin) link(log) nolog eform

. vif, uncentered
```

Variable	VIF	1/VIF
timepoint	13.51	0.074007
measurement	12.48	0.080140
rct	5.75	0.173909
metric	4.53	0.220975
drug	4.00	0.249763
industry1	3.62	0.275984
prospective	2.14	0.467262
surgical	1.28	0.783508
Mean VIF	5.91	

Table C6

Variance inflation factors of log-binomial regression covariates (randomized design, prospective registration, industry funding, surgical study, and drug study), for association with study publication, including one outcome description criterion (metric)

```
. glm publish metric rct prospective industry1 surgical drug if outcomementioned
>  == 1, fam(bin) link(log) nolog eform

Generalized linear models                         Number of obs   =      2,358
Optimization     : ML                             Residual df     =      2,351
                                                  Scale parameter =          1
Deviance         =  2949.341796                   (1/df) Deviance =   1.254505
Pearson          =  2388.729313                   (1/df) Pearson  =   1.016048

Variance function: V(u) = u*(1-u)                 [Bernoulli]
Link function    : g(u) = ln(u)                   [Log]

                                                  AIC             =   1.256718
Log likelihood   = -1474.670898                   BIC             =  -15307.51
```

publish	Risk Ratio	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
metric	1.605119	.1151422	6.60	0.000	1.394591	1.847428
rct	1.375973	.1200412	3.66	0.000	1.159712	1.632561
prospective	1.057097	.0524897	1.12	0.263	.9590669	1.165148
industry1	1.549251	.110801	6.12	0.000	1.346618	1.782374
surgical	.8873885	.1057291	-1.00	0.316	.7025808	1.120808
drug	1.235307	.0909432	2.87	0.004	1.069324	1.427054
_cons	.124654	.0147398	-17.61	0.000	.098868	.1571655

Note: _cons estimates baseline risk.

```
. vif, uncentered
```

Variable	VIF	1/VIF
rct	4.10	0.244115
drug	3.76	0.265964
industry1	3.48	0.287065
metric	3.39	0.295199
prospective	2.00	0.499023
surgical	1.20	0.832822
Mean VIF	2.99	

Including all outcome description criteria as covariates in the same regression model showed strong evidence of multicollinearity (Table C5). Therefore, each criterion was modeled individually for publication and publication with primary outcome, which reduced multicollinearity to acceptable levels in every case (VIF < 10). Output and variance inflation factors for an example model, including only one of the criteria variables (mention of metric), are given above (Table C6).

Table C7

Multivariable modified Poisson regression model for association between study characteristics (enrollment size, randomized design, prospective registration, industry funding, surgical study, and drug study) and study publication

```
. gen enroll1000 = enrollment/1000
(36 missing values generated)

. poisson publish enroll1000 rct prospective industry1 surgical drug, irr vce(
> robust) nolog

Poisson regression                              Number of obs     =      2,410
                                                Wald chi2(6)      =     136.90
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -1754.5205               Pseudo R2         =     0.0295
```

publish	IRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
enroll1000	1.014067	.0035416	4.00	0.000	1.00715	1.021033
rct	1.41943	.1246182	3.99	0.000	1.195041	1.68595
prospective	1.115876	.0560786	2.18	0.029	1.011203	1.231383
industry1	1.65028	.1180411	7.00	0.000	1.434409	1.898639
surgical	.874039	.1085815	-1.08	0.278	.6851512	1.115001
drug	1.27985	.0941576	3.35	0.001	1.107993	1.478364
_cons	.1537276	.0174915	-16.46	0.000	.1229986	.1921337

Note: _cons estimates baseline incidence rate.

Table C8

Multivariable modified Poisson regression model for association between study characteristics (enrollment size, randomized design, prospective registration, industry funding, surgical study, and drug study) and study publication with primary outcome

```
. poisson publishprimary enroll1000 rct prospective industry1 surgical drug, irr
>  vce(robust) nolog

Poisson regression                              Number of obs     =      2,410
                                                Wald chi2(6)      =     118.45
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -1565.8953               Pseudo R2         =     0.0311

------------------------------------------------------------------------------
             |               Robust
publishprim~y |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  enroll1000 |    1.01471   .0034438     4.30   0.000     1.007983    1.021483
         rct |   1.444343   .1485751     3.57   0.000     1.180619    1.766977
  prospective |   1.159702   .0693164     2.48   0.013       1.0315    1.303837
    industry1 |   1.815224   .1561268     6.93   0.000     1.533623    2.148532
     surgical |   .8265732   .1220181    -1.29   0.197     .6189101    1.103914
         drug |   1.183309   .0996283     2.00   0.046     1.003301    1.395612
        _cons |   .1182062   .0157312   -16.05   0.000     .0910668    .1534338
------------------------------------------------------------------------------
Note: _cons estimates baseline incidence rate.
```

Table C9

Multivariable negative binomial regression, including likelihood ratio test with Poisson regression (p > 0.05), for association between study characteristics (enrollment size, randomized design, prospective registration, industry funding, surgical study, and drug study) and study publication

```
. nbreg publish enroll1000 rct prospective industry1 surgical drug, irr nolog

Negative binomial regression                    Number of obs     =      2,410
                                                LR chi2(6)        =     106.74
Dispersion     = mean                           Prob > chi2       =     0.0000
Log likelihood = -1754.5205                     Pseudo R2         =     0.0295
```

publish	IRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
enroll1000	1.014067	.0032939	4.30	0.000	1.007632	1.020544
rct	1.41943	.1509771	3.29	0.001	1.152328	1.748443
prospective	1.115876	.0737372	1.66	0.097	.9803209	1.270174
industry1	1.65028	.1434097	5.76	0.000	1.391836	1.956714
surgical	.8740389	.1253044	-0.94	0.348	.6599338	1.157607
drug	1.27985	.1150776	2.74	0.006	1.073059	1.526493
_cons	.1537276	.0203793	-14.13	0.000	.1185523	.1993397
/lnalpha	-20.9446	.			.	.
alpha	8.01e-10	.			.	.

```
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline incidence rate.
LR test of alpha=0: chibar2(01) = 0.00                 Prob >= chibar2 = 1.000
```

Table C10

Multivariable negative binomial regression, including likelihood ratio test with Poisson regression (p > 0.05), for association between study characteristics (enrollment size, randomized design, prospective registration, industry funding, surgical study, and drug study) and study publication with primary outcome

```
. nbreg publishprimary enroll1000 rct prospective industry1 surgical drug, irr nolog

Negative binomial regression                    Number of obs     =      2,410
                                                LR chi2(6)        =     100.41
Dispersion     = mean                           Prob > chi2       =     0.0000
Log likelihood = -1565.8953                     Pseudo R2         =     0.0311
```

publishprimary	IRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
enroll1000	1.01471	.003583	4.14	0.000	1.007712	1.021757
rct	1.444343	.1726447	3.08	0.002	1.142681	1.825643
prospective	1.159702	.0855519	2.01	0.045	1.003581	1.340108
industry1	1.815224	.1796738	6.02	0.000	1.495123	2.203858
surgical	.826573	.1357878	-1.16	0.246	.5990285	1.140552
drug	1.183309	.1172615	1.70	0.089	.9744221	1.436974
_cons	.1182062	.0176099	-14.33	0.000	.0882737	.1582885
/lnalpha	-20.18261	.			.	.
alpha	1.72e-09	.			.	.

```
Note: Estimates are transformed only in the first equation.
Note: _cons estimates baseline incidence rate.
LR test of alpha=0: chibar2(01) = 0.00                  Prob >= chibar2 = 1.000
```

Likelihood ratio tests were statistically non-significant and did not suggest negative binomial regression better fit the data (both $p > 0.05$), demonstrating lack of evidence for overdispersion (Table C9, Table C10). Therefore, Poisson regression was adequate (Table C7, Table C8).

```
. qui reg timepublication significant rct prospective industry1 surgical drug

. capture program drop eda

. program define eda
  1. set graphics off
  2. set scheme s1mono
  3. qui hist `1', name(eda1, replace)
  4. qui graph box `1', name(eda2, replace)
  5. qui kdensity `1', ep normal name(eda3, replace)
  6. qui qnorm `1', name(eda4, replace)
  7. set graphics on
  8. graph combine eda1 eda2 eda3 eda4, name(eda, replace)
  9. end

. predict resid, resid
(1,521 missing values generated)

. eda resid
```

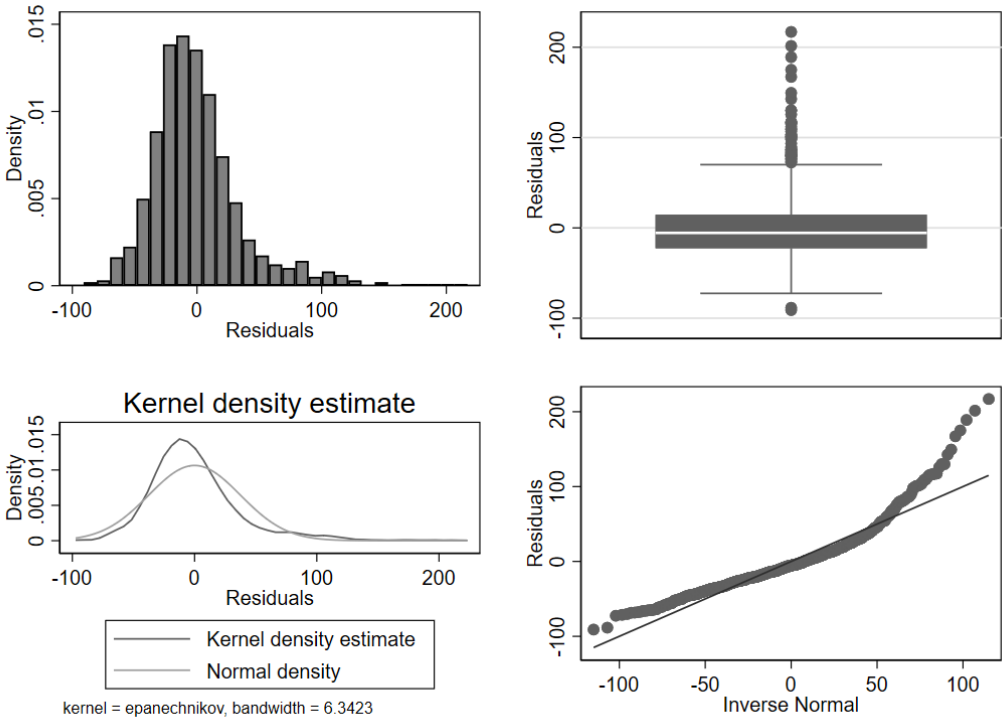


Figure C1. Residual and density plots of multiple linear regression for association between study characteristics and time to publication.

```
. qui reg lntime significant rct prospective industry1 surgical drug

. predict lnresid, resid
(1,521 missing values generated)

. eda lnresid
```

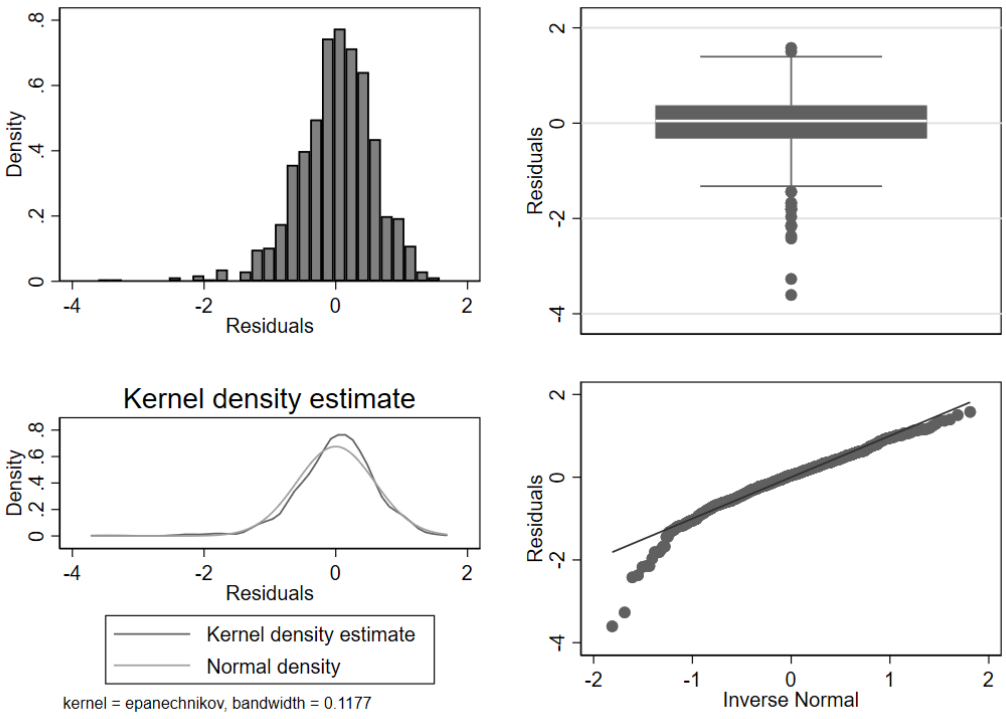


Figure C2. Residual and density plots of multiple linear regression for association between study characteristics and log-transformed time to publication.

Multiple linear regression model for effects of covariates on time to publication failed to meet assumptions of normality and homoskedasticity. The residual plot, boxplot, kernel density curve, and Q-Q plot for time to publication showed strong evidence of right skew (Figure C1). Therefore, log transformation of the time to publication variable was necessary, which improved its normality (Figure C2).

```
rvfplot, yline(0)
```

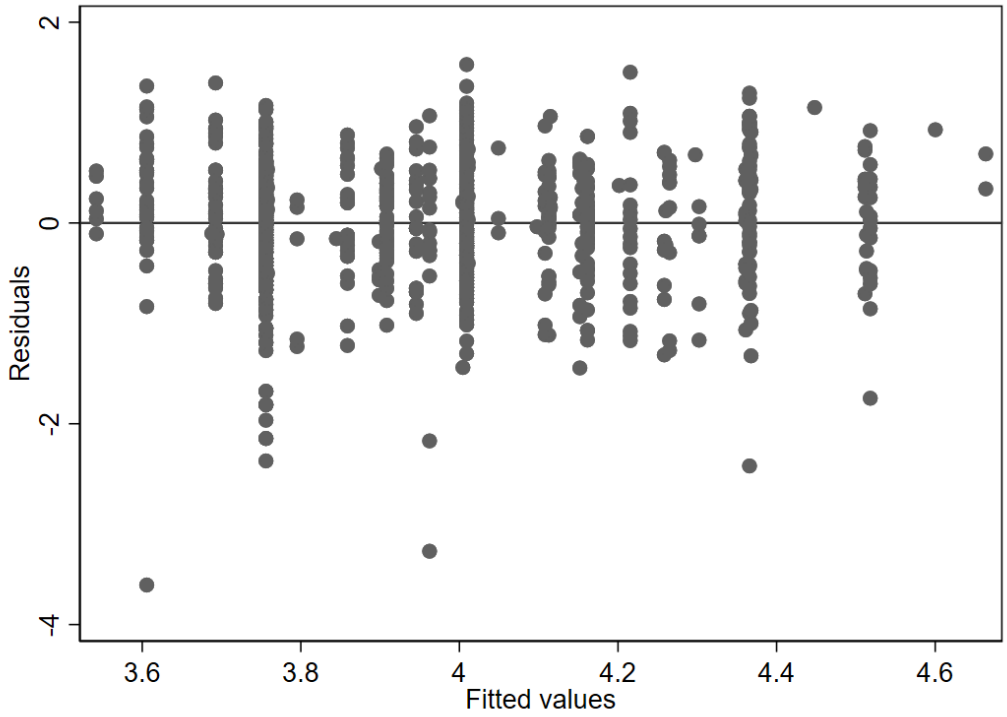


Figure C3. Residual versus fitted outcome plot for log-transformed time to publication in multiple linear regression model.

Table C11

Additional heteroskedasticity tests for log-transformed time to publication multiple linear regression model (both p *> 0.05)*

```
. estat imtest

Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	31.72	21	0.0625
Skewness	23.33	6	0.0007
Kurtosis	4.05	1	0.0443
Total	59.09	28	0.0005

```
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of lntime

         chi2(1)      =     0.48
         Prob > chi2  =   0.4880
```

Although the tests for constant variance were inconclusive (Table C11), the residual versus fitted outcome plot (Figure C3) showed some evidence of heteroskedasticity. Therefore, a robust variance parameter was used in the linear regression model (Table C12).

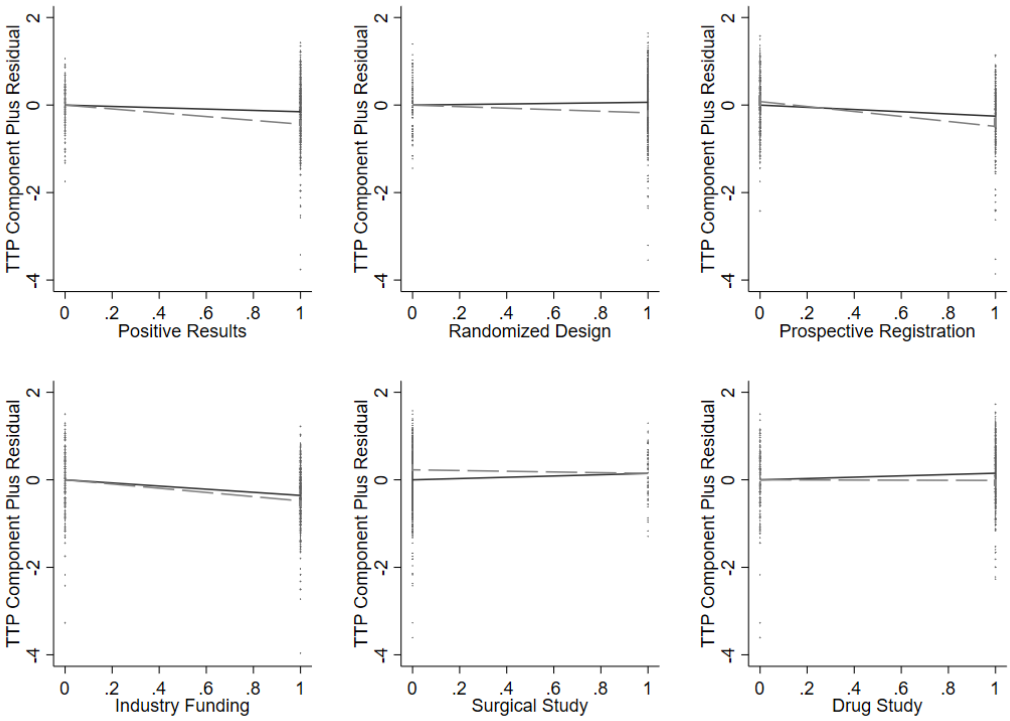


Figure C4. Component plus residual plots for predictors of log-transformed time to publication multiple linear regression model.

Due to the dichotomous nature of the covariates in the linear regression model, the predictor-outcome relationships were linear (Figure C4) and inherently met the assumptions of linearity.

Table C12

Analogous structural equation function of multiple linear regression model for association between study characteristics (positive results, randomized design, prospective registration, industry funding, surgical study, and drug study) and log-transformed time to publication, using robust variance

```
. qui gsem lntime <- significant rct prospective industry1 surgical drug, vce(rob
> ust)

. margins, expression(exp(predict(eta))*(exp((_b[/var(e.lntime)])/2))) dydx(signi
> ficant rct prospective industry1 surgical drug)

Average marginal effects                        Number of obs     =        925
Model VCE    : Robust

Expression   : exp(predict(eta))*(exp((_b[/var(e.lntime)])/2))
dy/dx w.r.t. : significant rct prospective industry1 surgical drug
```

	Delta-method dy/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
significant	-9.700985	3.086775	-3.14	0.002	-15.75095	-3.651017
rct	4.047654	4.022	1.01	0.314	-3.835322	11.93063
prospective	-16.12075	2.576861	-6.26	0.000	-21.1713	-11.07019
industry1	-22.71838	3.936931	-5.77	0.000	-30.43462	-15.00213
surgical	9.2766	5.760386	1.61	0.107	-2.013549	20.56675
drug	9.578251	3.990913	2.40	0.016	1.756205	17.4003

Table C13

*Log-rank test to compare survival to publication of positive versus negative studies (*p < *0.05)*

```
. qui stset timepublication, fail(publish)

. sts test significant

        failure _d:  publish
  analysis time _t:  timepublication
```

Log-rank test for equality of survivor functions

significant	Events observed	Events expected
0	143	176.99
1	782	748.01
Total	925	925.00

chi2(1) = 8.30
Pr>chi2 = 0.0040

Log-rank test showed statistically significant difference between survival distributions of positive and negative studies for reaching publication ($p < 0.05$) (Table C13).

Table C14

Schoenfeld residuals test for proportional hazards assumption of multivariable Cox regression for association between study characteristics (positive results, randomized design, prospective registration, industry funding, surgical study, and drug study) and survival to publication ($p < 0.05$*)*

```
. qui stcox significant rct prospective industry1 surgical drug, nolog

. estat phtest, detail

      Test of proportional-hazards assumption

      Time:  Time
```

	rho	chi2	df	Prob>chi2
significant	-0.09261	7.91	1	0.0049
rct	0.02661	0.66	1	0.4170
prospective	0.03452	1.15	1	0.2837
industry1	0.01978	0.35	1	0.5524
surgical	0.05666	2.94	1	0.0865
drug	0.01676	0.25	1	0.6204
global test		12.86	6	0.0454

```
. stphplot, by(significant) adjust(rct prospective industry1 surgical drug) legend(
> on order(2 "Positive" 1 "Negative")) plot1(lcolor(cranberry) lwidth(medthick) mco
> lor(cranberry)) plot2(lcolor(emerald) lwidth(medthick) lpattern(dash) mcolor(emer
> ald) msymbol(diamond)) graphregion(color(white))

        failure _d:  publish
  analysis time _t:  timepublication
```

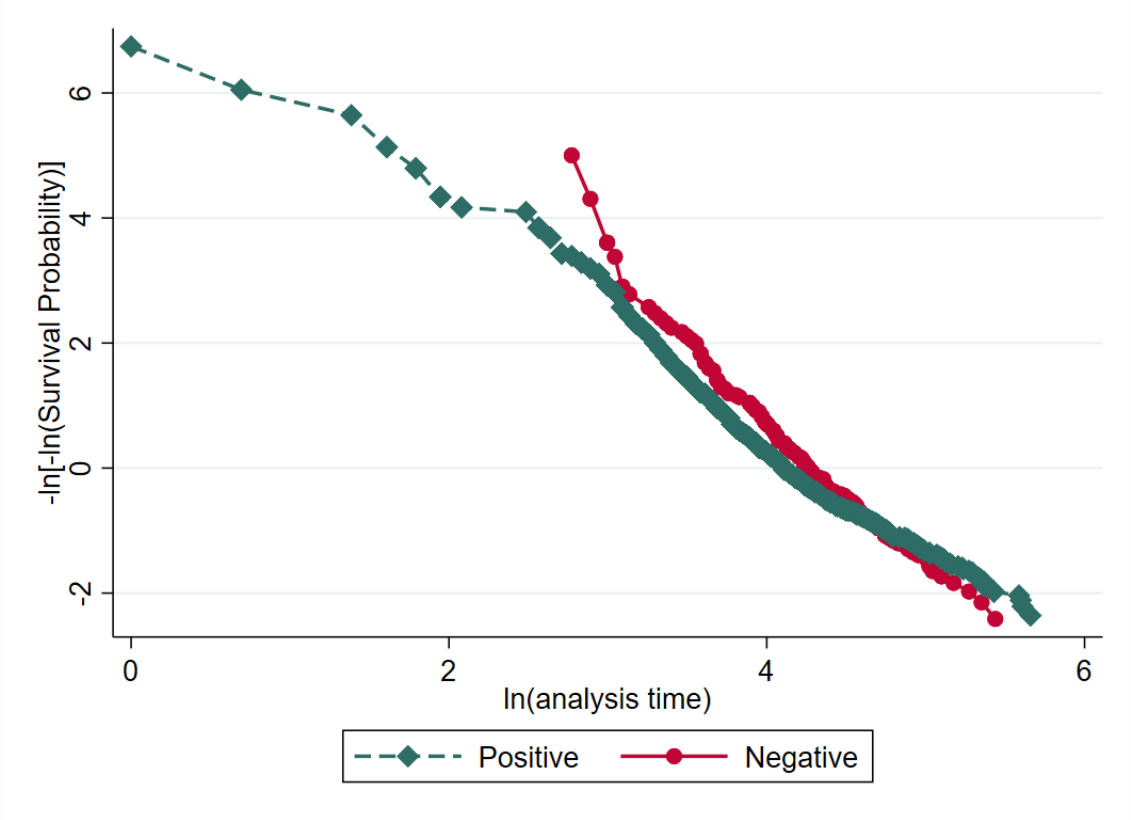


Figure C5. Log-log plot for "survival" to publication of positive (green dashed line) versus negative (red solid line) studies, adjusting for randomized design, prospective registration, industry funding, surgical study, and drug study.

Table C15

Multivariable Weibull regression for association between study characteristics (positive results, randomized design, prospective registration, industry funding, surgical study, and drug study) and survival to publication

```
. streg significant rct prospective industry1 surgical drug, distribution(weibull
> ) nolog

        failure _d:  publish
  analysis time _t:  timepublication

Weibull PH regression

No. of subjects =          925                 Number of obs    =         925
No. of failures =          925
Time at risk    =        58068
                                               LR chi2(6)       =      230.88
Log likelihood  =   -810.87321                 Prob > chi2      =      0.0000
```

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
significant	1.112103	.1022851	1.16	0.248	.9286584	1.331784
rct	.9331299	.0992145	-0.65	0.515	.7575977	1.149332
prospective	1.748559	.1224987	7.98	0.000	1.52422	2.005917
industry1	2.421869	.2230656	9.60	0.000	2.021859	2.901017
surgical	.9105783	.1309652	-0.65	0.515	.6868976	1.207098
drug	.8573642	.0730822	-1.81	0.071	.7254515	1.013263
_cons	.0001189	.0000338	-31.76	0.000	.000068	.0002076
/ln_p	.6571822	.024278	27.07	0.000	.6095982	.7047661
p	1.929348	.0468407			1.839692	2.023373
1/p	.5183098	.0125835			.4942241	.5435692

Note: _cons estimates baseline hazard.

As the Schoenfeld residuals test was statistically significant ($p < 0.05$) (Table C14) and the log-log plot showed the positive and negative study survival functions were not parallel (lines crossed) (Figure C5), there is evidence that our data does not meet the proportional hazards assumption. Therefore, a parametric model such as Weibull regression should be used instead (Table C15).

Appendix D: Supplementary Data and Calculations for Graphics

Table D1

Percentage and frequency of registered studies in each status category

Study Status	Percentage (%)	Frequency (N = 2446)
Completed	66.43	1625
Active, not recruiting	10.47	256
Enrolling by invitation	0.25	6
Recruiting	6.58	161
Terminated	10.83	265
Suspended	0.16	4
Unknown status	4.50	110
Withdrawn	0.78	19
Cumulative Total	**100**	**2446**

Table D2

Annual frequencies of study publication and simplified status, by year of first participant enrolled

Year	Published	Unpublished	Completed	Ongoing	Stopped
1983	1	0	0	0	0
1988	1	0	0	0	0
1989	1	0	0	0	0

1992	0	1	1	0	0
1993	2	1	1	0	0
1994	3	3	3	0	0
1995	3	6	6	0	0
1996	1	4	4	0	0
1997	2	10	10	0	0
1998	7	22	18	1	3
1999	6	22	21	1	0
2000	9	29	26	2	1
2001	11	29	28	0	1
2002	13	47	38	1	8
2003	29	52	47	0	5
2004	27	65	57	1	7
2005	44	88	67	2	19
2006	61	96	71	0	25
2007	51	98	69	3	26
2008	73	68	51	4	13
2009	70	67	44	4	19
2010	61	83	54	8	21

2011	787	75	49	11	15
2012	76	100	67	12	21
2013	75	83	48	10	25
2014	71	86	36	32	18
2015	67	106	54	29	23
2016	44	130	41	64	25
2017	38	150	33	87	30

Note. "Unpublished" is the sum of completed, ongoing, and stopped studies for each year.

Table D3

Annual actual and estimated percentages of studies that are or will be published, by year of first participant enrolled

Year	Actual Published (%)	Estimated Total Published (%)	Not Yet Published (%)
1983	100	100	0
1988	100	100	0
1989	100	100	0
1992	0	0	0
1993	66.67	66.67	0
1994	50	50	0
1995	33.33	33.33	0

1996	20	20.02	0.02
1997	16.67	16.70	0.04
1998	24.14	24.22	0.08
1999	21.43	21.55	0.12
2000	23.68	23.84	0.15
2001	27.50	27.68	0.18
2002	21.67	21.88	0.21
2003	35.80	36.31	0.51
2004	29.35	29.93	0.58
2005	33.33	34.15	0.81
2006	38.85	39.93	1.08
2007	34.23	35.45	1.23
2008	51.77	54.24	2.46
2009	51.09	54.26	3.17
2010	42.36	45.67	3.31
2011	50.98	56.41	5.43
2012	43.18	49.56	6.38
2013	47.47	57.10	9.63
2014	45.22	58.51	13.28

2015	38.73	57.78	19.05
2016	25.29	48.53	23.24
2017	20.21	58.06	37.85

Sample estimation for number of studies not yet published from 2016 (Table D3):

From Table D2, we see that 44 out of 174 studies (25.29%) started in year 2016 have already been published. Length of time from beginning of 2016 to end of July 2020 is then calculated in months as follows:
(2017 – 2016) × 12 + 43 = 55 months

```
. egen perc = mean((timepublication < 55) / (timepublication < .))

. di perc
.52108109
```

Using Stata, we determine 55 months to be roughly the 52.11th percentile of the time to publication variable distribution. This means the proportion of studies published since 2016 is estimated to represent only 52.11% of all studies that will be published from that year. Thus, we divide the actual proportion published by the percentile number (as a proportion) to get the estimated total proportion of studies that have and will be published given time:
0.2529/0.5211 = 0.4853

We then find the difference between the estimated and actual proportion to get the estimated proportion of not yet published studies:
0.4853 – 0.2529 = 0.2324 = 23.24%

Table D4

Chi^2 test of independence between study publication and year of first participant enrolled, from 1997 onwards (p < *0.001)*

```
. tab publish year if year >= 1997, chi2

 Pearson chi2(20) = 107.5368   Pr = 0.000
```

Table D5

*One-way ANOVA between study publication and year of first participant enrolled, from 1997 onwards (*p *< 0.001)*

```
. oneway publish year if year >= 1997

                        Analysis of Variance
    Source              SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between groups      25.2685893     20   1.26342946      5.58     0.0000
 Within groups      543.139017   2398   .226496671
------------------------------------------------------------------------
    Total           568.407606   2418   .235073452

Bartlett's test for equal variances:  chi2(20) =  22.6524  Prob>chi2 = 0.306
```

Both the Chi^2 test of independence (Table D4) and one-way ANOVA (Table D5) showed a statistically significant relationship between study publication and start year (both $p < 0.001$). Both tests were restricted to 1997 and onwards due to sparsity of preceding data (Table D2), which also coincides with the year ClinicalTrials.gov was formalized.

Table D6

Characteristics of registered studies, by year of first participant enrolled (annual)

Year	Registered Studies	Randomized Design (%)	Industry Funding (%)	Surgical Study (%)	Drug Study (%)
1983	1	100	0	0	0
1988	1	0	0	100	100
1989	1	100	100	0	100
1992	1	100	0	100	0
1993	3	66.67	0	0	66.67
1994	6	100	16.67	0	83.33

1995	9	77.78	22.22	44.44	77.78
1996	5	100	0	40	40
1997	12	50	0	41.67	83.33
1998	29	72.41	31.03	24.14	72.41
1999	28	82.14	17.86	39.29	67.86
2000	38	63.16	31.58	15.79	84.21
2001	40	80	50	22.50	90
2002	60	75	53.33	8.33	86.67
2003	81	87.65	70.37	14.81	80.25
2004	92	89.13	70.65	10.87	78.26
2005	132	93.94	75	8.33	83.33
2006	157	88.54	70.06	8.28	80.25
2007	149	86.58	69.13	11.41	73.83
2008	141	84.40	78.01	7.09	78.01
2009	137	81.75	72.26	8.03	72.99
2010	144	83.33	70.83	9.03	77.78
2011	153	81.70	65.36	9.80	65.36
2012	176	83.52	64.20	10.80	65.34
2013	158	84.81	65.19	9.49	68.35

2014	157	80.89	62.42	12.74	63.69
2015	173	80.92	64.16	7.51	63.58
2016	174	83.33	60.34	5.17	60.92
2017	188	86.17	57.45	10.64	63.83
Overall	**2446**	**83.81**	**63.98**	**10.59**	**71.63**

Table D7

Percentage of studies that progressed to each stage of the research process

Stage	Percentage of Total Registered Studies (%)	Percentage of Previous Stage (%)
Registered	100	-
Started	94.72	94.72
Completed	66.43	70.13
Published	37.82	56.93
Published with Primary	30.38	80.32

Table D8

Frequency and percentage of positive and negative studies across relevant stages of the research process

Stage	Positive Results	Negative Results	Unpublished	Total	Percentage Positive (%)
Registered	782	143	1521	2446	31.97
Completed	574	107	944	1625	35.32
Published	782	143	-	925	84.54
Published with Primary	616	127	-	743	82.91

Note. The “started” stage was not included due to redundancy with “registered”.

Curriculum Vitae

Name:	Alex Bi
Post-secondary Education and Degrees:	University of Western Ontario London, Ontario, Canada 2014-2018 B.M.Sc. University of Western Ontario London, Ontario, Canada 2018-Present M.Sc.
Honours and Awards:	Continuing Admission Scholarship (UWO) 2014-2015, 2017-2018 Dean's Honour List (UWO) 2015-2018 Western Graduate Research Scholarship (UWO) 2018-2020
Related Work Experience	Teaching Assistant – Multivariable Biostatistics 9521B University of Western Ontario 2020