Outcome Reporting in Surgical Randomized Controlled Trials

Peter Glen, MD, BSc

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the Master of Science degree in Epidemiology

Epidemiology and Community Medicine

Faculty of Medicine

University of Ottawa

Table of Contents

Figure Index

Table Index

Abstract

Background: In September 2005, scientific journals began requiring trial protocol registration to increase transparency and accountability.

Objective: My primary objectives were: develop a database of linked protocols and publications for surgical randomized control trials (RCTs); estimate the proportion published; and determine the proportion exhibiting selective outcome reporting.

Methods: A systematic search of the clinicaltrials.gov database was conducted identifying surgical RCTs, completed between 2006 and 2012. Protocols were linked with publications. Primary outcomes were compared.

Results: We identified a cohort of 743 surgical RCT protocols. The proportion of registered trials which published their primary results was 0.49 (n=364). The proportion of selective outcome reporting was estimated to be 0.244, significantly lower than the previous estimate ($p<0.001$).

Conclusion: More than half of the completed surgical RCTs were unpublished, and one quarter of those published selectively reported their primary outcome. This supports the notion that significant bias is present in the surgical literature.

1. Introduction

Properly conducted, randomized controlled trials (RCTs) are considered the gold standard for measuring the efficacy of an intervention[1]. In surgical disciplines, the underwhelming number of RCTs conducted and published has for many years been criticized [2,3]. As an example, in the surgical subspecialty of pediatric surgery, it was estimated that at the end of the 20th century, as few as 0.3% of all publications in the discipline were RCTs [4]. Thankfully, funding bodies, surgeons and epidemiologists have answered the call, and the rate at which RCTs are being conducted has been steadily increasing. The volume of surgical RCTs has increased by 50% over the past decade [5].

As the volume of RCTs has increased, systematic reviews and meta-analyses have become powerful tools capable of estimating with greater precision the effect of interventions across a number of studies[6]. However, the validity of meta-analyses depends heavily on the quality of the RCT data used and reported [7–10]. Conducting meaningful meta-analyses proves difficult if the original studies are heterogeneous in their study design and poor methodological reporting, and outcome reporting is present[11]. Consequently there have been efforts made to ensure that RCTs are conducted and reported with greater transparency through their design, conduct, and reporting phases.

An example of this effort is the Consolidated Standards of Reporting Trials (CONSORT) statement, which was first published in 1996 [12] and has since undergone two revisions [13,14]. The CONSORT statement consists of items which should be presented in publications so that readers can critically appraise the data

and be aware of all required trial elements. Since its inception, it has been adopted by most high impact journals which require investigators to demonstrate compliance with CONSORT when submitting articles [15,16]. By requiring authors to comply with the reporting standards, there has been a demonstrated improvement in the quality of the literature, and enabled the reader with the ability to critically assess the data [17–19]. Reporting trial design and conduct elements is indispensible in assessing the quality of the trial and robustness of results. It has been determined that ignoring even just one element, allocation concealment, is associated with an over estimation of treatment effect by as much as 30% [20,21].

For the past 10 years, interventional studies, including RCTs, have been required to register a trial protocol in a publicly accessible database prior to commencement of the trial. This initiative was adopted by the International Committee of Journal Editors (ICMJE) in 2004 [22], and became standard in September of 2005 [23]. This was meant to ensure trials establish a priori their intervention, comparator, target population, and outcome measures with the objective of reducing publication bias and selective outcome reporting. These efforts to improve trial transparency and complete reporting have been widely endorsed by researchers and top-tier journals [24–29].

The introduction of these requirements has improved the overall quality of reporting of clinical trials, however in medical and surgical trials many potential sources of bias remain [30–32]. This thesis will look at two important issues that affect the quality of surgical literature: selective outcome reporting bias and publication bias.

Problems with selective outcome reporting have been recognized as major sources of bias in both the surgical and medical literature [33,34]. A number of researchers have attempted to quantify the proportion of trials which show a discrepancy between registered primary outcomes and reported outcomes. The range of estimates is broad with estimates of reporting discrepancies ranging from 14%-100% [35,36].

Three sources of data have been used to evaluate trial protocol outcomes and their congruence with published outcomes. The first technique which has been employed by Dr. An-Wen Chan involved accessing trial protocols registered with funding agencies or regulatory agencies and comparing the protocol outcomes with the results presented in published trials. One study used 48 trial protocols accessed through the Canadian Institute of Health Research and determined that 40% showed evidence of selective outcome reporting[33]. Another study used 122 trials approved in Denmark between 1994-1995 and showed that at least one major reporting difference occurred in 62% of published trials[37]. In a study of 227 approved drug trials in Switzerland, Redmond estimated the rate of outcome reporting discrepancies to be 30%[38].

Another research method involves identifying published RCTs and then accessing their trial protocols to determine the proportion of trials which changed their primary outcome variable. In studies assessing outcomes published in medical trials the proportion has been estimated at between 14%-50%[34,35,39]. A similar technique was used in an assessment of surgical trials and the estimated reporting discrepancy was 29%[40].

Trial protocol registries have now enabled a third technique for assessing outcome reporting whereby the researcher starts with a cohort of trial protocols and then tracks forward to publication. The clinicaltrials.gov registry has been used as a starting point due to its accessibility and widespread adoption by investigators and sponsors. Evidence presented in two recent reviews of surgical trial reporting suggest that 45-50% of surgical RCTs show a discrepancy between the primary outcome measures listed in trial protocol, and the outcome reported in the published trial, a discrepancy known as selective outcome reporting [41,42]. In sports medicine trials, 35% changed their primary outcome[43], and in the anesthesia literature the rate is even more telling where one study estimated the proportion to be 100%[36].

The consequences of these discrepancies can be substantial. When trials either do not report, or change their primary outcome measures, they are more likely to produce spurious, chance results [37]. The advent of trial protocol registries (e.g. ClinicalTrials.gov) and stricter reporting guidelines required by journals trial publication allows researchers to compare the two and identify instances in which researchers have changed their reporting measures and thus identify instances of selective outcome reporting [32].

Publication bias is another area of concern. Previous estimates of the proportion of RCTs that go on to publication have linked conference abstracts with subsequent publication rates. Using this technique, the proportion published has been estimated at between 50-74%[44–46]. Now, trial protocol registries have proven useful in determining publication rates. Three recent reviews in which protocols registered on clinicaltrials.gov were used to identify subsequent published data

found that 2-4 years after study completion, trials were only published 45 to 65% of the time [47–49].

When trial data goes unpublished, the overall quality of the literature is diminished. In evaluating effectiveness of interventions, particularly when done with the technique of meta-analysis, negative trials, or trials showing no difference in effect are as important as those which show effect. The concern is that if null data goes unpublished, then only trials which have statistically significant results – even if they are spurious – will be published [32,33,50]. Of additional concern is that patients are being recruited to trials with the altruistic belief that the potential risks they are subject to will be offset by the gains made in health research. Not reporting trial results is as much a disservice to patients as it is to the literature. Additionally, the amount of money spent on research totals billions of dollars annually, and by extension there is tremendous waste of money when research is left unpublished. Quantifying this waste of resources is difficult to measure with precision, but it is reasonable to conclude that the loss of public tax dollars is significant [51].

Even as trial conduct and reporting has improved, there have been persistent issues with specific elements of surgical trials which have been criticized as potentially serious sources of bias [30,52,53]. These include the surgical learning curve, whereby mastery of the intervention can have a dramatic impact on outcomes; and standardization of the surgical protocol, an area vulnerable to bias if there are even slight deviations from one procedure to the next. In addition, there are a number of additional sources of potential bias which have been identified in clinical research [54] that have proven particularly difficult to control in surgical research. Blinding is an

area of particular concern, be it blinding of surgeons to the allocation process, blinding of patients to the procedure being performed, or blinding of the outcome assessors: all can introduce sources of significant bias [20,55,56]. Therefore, complete reporting of all study elements is crucial for the reader to interpret the validity of surgical trials.

Dr. Pierre-Alain Clavien has stated that surgeons must be aware of the biases in the literature, and that knowledge of publication and selective outcome reporting biases will enable practitioners to mitigate their effect [57]. Fully accessible trial information, from conception through publication, is heralded as the means by which biases can be identified and overcome. Such accessibility will ensure that high quality research is conducted and could lead to significant cost savings in research [58]. A commentary published in the *Journal of the American Medical Association* goes even further, calling for trial registration and reporting to be seen as a "public good" [59].

In 2005 Dr. John Ioannidis came to the conclusion that the published literature is not as good as consumers would like to believe. In fact he came to the conclusion that most of the published literature could in fact be "proven" to be false [60]. Although the significance of Dr. Ioannidis' claim has been challenged by Dr. Steve Goodman [61], both agree that published research remains imperfect in its interpretation and presentation of trial results.

If we accept the premise of this argument, then the detrimental impact of non-publication further distorts the validity of the literature [62]. Although the argument has been made that scientists should always publish their results [63] the message

has not yet resonated with researchers who are either unaware of the importance of publication, or who are unwilling to abide by this edict [64].

2. Thesis Objectives

There are three primary objectives to this thesis:

1. Identify a cohort of surgical randomized controlled trials

- Cohort of trials assessing surgical intervention

2. Link trial protocols to the primary trial results publication

- Determine proportion of trials published
- Identify factors associated with publication

3. Compare published trial outcome with the registered trial outcome

- Determine proportion of trials which selectively report primary outcome
- Identify factors associated with selective outcome reporting

3. Identifying Surgical Trial Protocols and Linked Primary Publications: Thesis Objectives 1 & 2

3.1. Introduction

This thesis was established on the principle that clinical trials should answer the primary research question specified in the trial protocol. We hypothesized that there is likely an important proportion of trials that would go unpublished upon their completion or termination and that there would be factors associated with non-publication.

A subgroup of members from the thesis advisory committee convened and defined three primary structural domains hypothesized to have bearing on whether a completed trial would be published. These domains were (1) Registration (2) Demographics and (3) Characteristics. A set of items (variables) was defined for each of these domains.

Objective 1 was to use registered trial protocol information available on the National Institute of Health (NIH) website “clinicaltrials.gov” as a basis for creating a database of surgical randomized controlled trials. Each of the eligible trials was systematically reviewed and information was abstracted for the defined variables identified a priori. Subsequently, a systematic literature search algorithm was performed to identify publications that arose from our database of trial protocols. This data was stored in a novel database named STARTOPP (*Surgical Trials and Reported Trial Outcomes, Protocols to Publications*).

Objective 2 was to determine the proportion of published trials and to determine if there were significant associations between the variables we had defined and the publication status of the trials.

3.2. Identification of Study Protocols (Objective 1): Materials and Methods

3.2.1. Initial Query of clinicaltrials.gov Database

The clinicaltrials.gov trial registry database was searched formally on January 8, 2014. An advanced search strategy was conducted to identify studies registered on the database on or before December 31, 2013 which pertained to interventional studies conducted in surgical disciplines. The "search terms" field of clinicaltrials.gov was used to identify trial protocols that contained the word "surgery" in any part of the protocol. Details of the search, and the definitions of the limiting fields, appear in Table 1.

Table 1. Search strategy used in the Advanced Search of clinicaltrials.gov registry.

Search Field	**Field Definition (clinicaltrials.gov)** [65]	**Input Terms**
Search Terms	Contains words related to the studies; general search of study fields including: study title, brief description, conditions, interventions, locations	**surgery** OR **surgical** OR **surgeon** OR **surg** OR **operation** **chirurgicale** OR **chirurgicale** OR **chirugien** OR **chirug**
Study Results	Limits search based on whether the study has results posted in the clinicaltrials.gov database	**All Studies**
Study Type	Limits the search by clinical study type (interventional study, observational study, expanded access)	**Interventional Study** (By definition: participants are assigned to receive one or more interventions)
First Received	The date that the summary clinical study protocol information was first submitted to the registry	On or Before **December 31, 2013**

Recruitment	
Conditions	
Conditions	
Interventions	
Title	
Outcome Measure	
Sponsor	*No Restrictions*
Location	
Gender	
Age Group	
Phase	
Funder Type	
Safety Issue	
Last Updated	

Content information for all studies identified using the search strategy was downloaded in the comma-separated values (.csv) format and used to populate an Excel database on a disk drive controlled by the investigators. Information for all 21 available protocol fields, the maximum field number made available by clinicaltrials.gov, was downloaded for each of the identified studies. The 21 protocol fields are filled both by researchers registering trials and automatically generated by the clinicaltrials.gov. The full list of registered fields appears in Appendix 1. This dataset was the foundation of the STARTOPP database.

3.2.2. Automated Protocol and Application of Eligibility Assessment

Trials were entered into the STARTOPP database with each row representing an individual registered trial and each column representing a “protocol field” obtained from clinicaltrials.gov. The sort function was used in Excel to identify studies which were not eligible for inclusion according to our study design (Table 2). The goal of this phase of protocol eligibility assessment was to restrict the dataset to

RCTs that started and finished collecting outcome data between January 2006 and December 2012.

Trial protocols were excluded using sequential searches; if a trial protocol was excluded, it was not assessed in the subsequent searches. Ineligible trial protocols were stored separately according to the criterion for which they were excluded. The sequence of protocol assessment and rationale for exclusion are shown in Table 2. In circumstances where the protocol field contained insufficient information in the protocol field (blank, unknown, etc.), the trial protocol remained eligible for inclusion.

Table 2. Assessment of eligibility criteria using protocol field information made available on clinicaltrials.gov by study investigators.

Sort Number	**Protocol Field**	**Exclusion Criterion**
1	Study Design: *Allocation*	Non-randomized
2	Study Design: *Interventional Model*	Single Group Assignment
3	Study Start Date	**Start < January 1, 2006**
4	Completion Date	**Completed > December 31, 2012**
5	Primary Completion Date	**Primary Complete > December 31, 2012**
6	Recruitment Status	Withdrawn (by definition: no participants recruited)
7	Recruitment Status	Not yet recruiting
8	Recruitment Status	Recruiting
9	Completion Date	Completed < January 1, 2006

3.2.3. Final Eligibility Assessment

Trial protocols which remained eligible following the automated assessment were then individually reviewed to define a cohort of trials which (1) employed a randomized control trial design and (2) assessed a surgical intervention.

Trials were excluded if they studied medical interventions solely (e.g. drugs or anesthetics) or if they studied non-surgical interventions (e.g. physiotherapy, decision-aids) in patients undergoing surgery. If the study did not enroll participants who were in need of a surgical intervention, they were also excluded (e.g. simulated surgery).

Eligibility was assessed by two independent reviewers (PG and RW) using a standardized abstraction form (Appendix 2). The item assessing eligibility on the form was considered valid if there was agreement of 90% or higher on whether the trial protocol was eligible for inclusion in the study. The reviewers independently assessed random samples of 5% of the trial protocols until the validation threshold was reached.

Once validated, the remaining trial protocols were assessed for eligibility by a single reviewer (PG). Eligible trials were stored in the STARTOPP database.

3.2.4. Abstraction of Trial Protocols

Information from eligible protocols was then abstracted for the categorical variables defined by the research team (Appendix 2. Summary of variables abstracted from trial protocols.).

This abstraction form was validated by two reviewers (PG, RW). The abstracted data elements were considered valid if there was at least 80% agreement for each of the items. The reviewers independently assessed random samples of 5% of the trial protocols until this validation threshold was reached.

Once validated, the abstraction form was then used to abstract data for all of the remaining protocols by one reviewer (PG).

The variables we identified for abstraction were intended to supplement the 21 standard categorical fields downloaded into STARTOPP from clinicaltrials.gov. Information for our additional variables required hand searching of the original trial protocols on the website. The following is a brief description of each of the reviewed variables listed on the abstraction form (Appendix 2. Summary of variables abstracted from trial protocols.).

1. **Region**. The continent in which the study recruited patients was identified from the protocol. Latin American and Caribbean countries were grouped with those from South America; Ukraine, Israel, and Turkey with those from Asia. If a study was conducted on multiple continents, then it was assigned to the 'Multiple' group.
2. **Trial Category**. Surgical studies were categorized according to the type of intervention understudy.
    i. *Technique* trials were designed to compare two or more surgical interventions (e.g. open surgery compared with laparoscopic).
    ii. *Device trials* were identified using the Health Canada definition medical devices which states that the primary effect is not

achieved by “pharmacological, immunological, or metabolic means”[66]. These devices were identified as being either implanted, meaning they were left in the patient, or external meaning they directly influenced the surgical field but were not left in the patient.

iii. *Dressing* trials were identified as those which studied the effect of a dressing placed on the surgical wound at the time of surgery.

iv. *Surgical versus non-surgical* trials were those in which either the intervention or control arm underwent a surgical procedure which was not conducted on the comparator arm.

3. **Age**. The age of participants was abstracted and categorized as *adult* (all over 18 years), *pediatric* (all under 18 years), *both*, or *unknown*.
4. **Funding**. The sources of trial funding were assessed. One trial funding source is required for registration on clinicaltrials.gov. At most two funding sources were abstracted from the trial protocols. *Industry* funded trials were defined as those which listed as a sponsor an identifiable private or public corporation, or a group of corporations. If an industry sponsor was identified anywhere in the trial sponsorship it was recorded in the abstraction form. The other funding sources defined were *government* (e.g. NIH or CIHR), *university*, *hospital*, and *private interest*.
5. **Registration Timing**. The date of trial registration date was compared with the trial start and reported completion dates of primary outcome follow up.

The study was then classified according to whether it was registered *before*, *during*, or *after* the trial. If the trial dates were missing, then the registration was categorized *unknown*.

6. **Surgical Specialty**. The surgical field in which the study was conducted was identified from the trial protocol. The surgical specialties recognized as training domains by the Royal College of Physicians and Surgeons of Canada were used: *Cardiac*, *General Surgery*, *Neurosurgery*, *Obstetrics and Gynecology*, *Ophthalmology*, *Orthopedic Surgery*, *Otolaryngology*, *Plastic Surgery*, *Urology*, *Vascular Surgery*. In addition, trials in the area of *Thoracic Surgery*, *Dentistry*, and those involving *multiple* specialties were captured.

3.2.5. Descriptive Analysis of Trial Protocols

The data abstracted from the trial protocols and the data downloaded from the clinicaltrials.gov website were combined in the dataset.

Using SAS 9.3 software (2011, SAS Institute Inc., Cary, NC, USA), descriptive statistics of the number, proportion, and 95% confidence intervals were generated for each of the variables described in Section 0.

3.3. Linked Publication Search (Objective 2): Materials and Methods

A standardized step-wise search strategy was used to identify publications associated with the registered study protocols. The search for published trials was performed by a single investigator (PG). This search used both the clinicaltrials.gov trial registration number (NCT Number) and if necessary, the name of the principal

investigator (Figure 1). If a publication was identified using the NCT Number then the name of the principle investigator was not used.

The search was conducted between February 10, 2014 and March 18, 2014. Any trials not published (electronically or in print) by March 18, 2014 were categorized as not published. This result was recorded as a binary variable in our database.

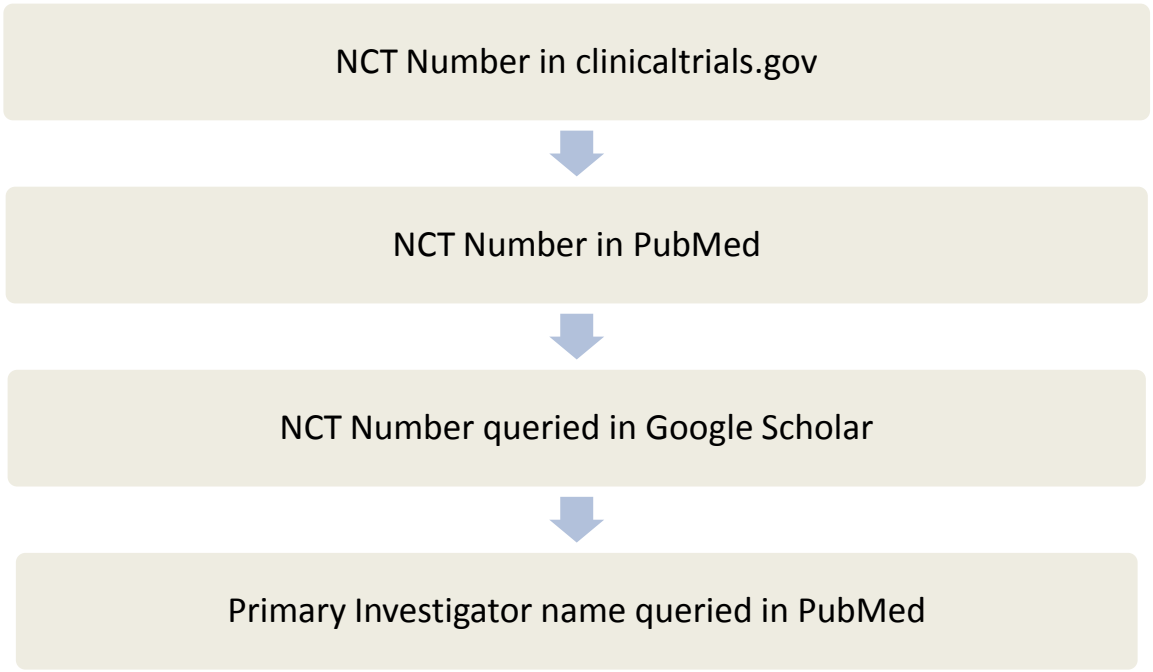

Figure 1. Search strategy used to identify publications arising from trial protocols.

In circumstances where multiple publications were found, then the first publication presenting descriptive results was selected for abstraction. Trials for which there was result data available on the clinicaltials.gov website but not found in

publication format were identified separately. In these circumstances, trial result data was not used for further analysis and the study was considered unpublished.

If there was no publication identified then the principle investigator's name was searched both on PubMed and in the clinicaltrials.gov database in an attempt to locate a contact email. This information was stored in our database for future use.

Published studies were retrieved in full text PDF format and archived on the investigator's computer. They were linked to the trial protocol by the unique NCT number. Demographic information for each published study was abstracted and recorded in the database:

1. Electronic publication date
2. Final publication date
3. Publishing journal
4. First author
5. Number of related published articles
6. Journal impact factor (As defined by the 2013 Thompson Reuters Journal Citation Report[67])

The earliest publication date, either the electronic publication date or the publication date, was used as the date of publication.

3.3.1. Statistical Analysis of Publications

Analysis was performed using the SAS 9.3 software (2011, SAS Institute Inc., Cary, NC, USA).

The primary outcome of interest was full journal publication of the registered protocol. The proportion of published trials was determined as the proportion of eligible protocols which resulted in a full journal publication; a 95% confidence interval was generated.

Publication was then analyzed as a binary variable to determine if there was an association with categorical variables of interest. The Kruskal-Wallis test was conducted to determine if there was a difference in the proportion of published trials with respect to each variable, and a univariable analysis was performed for each of the a predictor variables for the binary publication outcome. The reference category for each variable was set as the most common outcome variable. Odds ratios and 95% confidence intervals were generated for each of the univariable analyses.

Using the same binary outcome of publication, a multivariable analysis was performed using all categorical variables in a full model. Odds ratios and 95% confidence intervals were generated for each of the variables in the model.

Estimates for the time to publication were made by comparing the trial completion date with the date of publication. The association between mean time to publication and journal impact factor was analyzed using a linear regression model.

3.4. Results

3.4.1. Cohort of surgical trials retrieved from clinicaltrials.gov

The search of clinicaltrials.gov identified 17,498 trials registered on or before December 31, 2012 that fulfilled the initial search criteria (Section 3.2.1). The automated article eligibility assessment (Section 3.2.2) eliminated 14,009 protocols as outlined in Table 3.

Table 3. Trial protocols eliminated by automated article eligibility assessment.

Exclusion Criterion	**Number Eliminated**
Non-Randomized Trial	2428
Study Design: Single Group Assignment	2953
Study Start Date < January 1, 2006	2698
Completion Date > December 31, 2012	4090
Primary Outcome Completion > December 31, 2012	675
Recruitment Status 'Withdrawn'	143
Recruitment Status 'Not yet Recruiting'	176
Recruitment Status 'Recruiting'	830
Completion Date < January 1, 2006	16
Total Number of Trials Excluded	**14,009**

3.4.2. Validation of Protocol Abstraction Form

The trial protocol abstraction form (Appendix 2. Summary of variables abstracted from trial protocols.) was validated by comparing results from two independent abstractors according to the criteria outlined in Section 0. The form was validated after the first 5% (n=175) sample of trial protocols were reviewed, according to the pre-specified threshold of agreement.

Table 4. Validation of abstraction form with level of agreement for each item.

Trial Protocol Item	**Level of Agreement (Kappa)**
Inclusion	0.943
Region	0.932
Funding	0.864
Trial Category	0.886
Age	0.956
Surgical Specialty	1.000
Registration Status	0.909

3.4.3. Objective 1: Final Determination of Eligible Surgical RCT Protocols

The validated abstraction form was then applied to the cohort of 3,489 surgical trial protocols which remained eligible following the automated assessment. 2,746 trials were ineligible for inclusion (Table 5). A total of 743 surgical RCT protocols met full eligibility and comprised the final cohort for this study.

Table 5. Result of inclusion assessment of trial protocols eligible for hand search.

Exclude	**Number of Protocols**	**Percentage of Protocols**	**95% Confidence Interval**	
Medical Trial	1614	46.26	44.61	47.91
Not Surgical Subject	677	19.40	18.09	20.71
Not Surgical Intervention	417	11.95	10.87	13.03
Not an RCT	23	0.66	0.39	0.93
Ineligible Dates	15	0.43	0.21	0.65
Include	**743**	**21.30**	**19.94**	**22.66**

The 743 trial protocols determined to be eligible for the cohort of surgical randomized controlled trials for this study constituted 4.25% of the trial protocols accessed from the clinicaltrials.gov website.

The successful identification of this cohort trial protocols fulfilled Objective 1 of this thesis.

3.4.4. Descriptive Analysis of Eligible Trial Protocols in the STARTOPP Database

Trial variables within each domain of interest (Appendix 2. Summary of variables abstracted from trial protocols.) were assessed as categorical variables. Results are shown in Table 6.

Table 6. Number of trial protocols and total proportion for each trial variable for the 743 eligible trial protocols.

Variable	Category	Number	Proportion	95% Confidence Interval	
TRIAL DEMOGRAPHICS					
Region	North America	300	0.4038	0.3685	0.4391
	Europe	262	0.3526	0.3182	0.3870
	South America	29	0.0390	0.0251	0.0529
	Asia	101	0.1359	0.1113	0.1605
	Africa	31	0.0417	0.0273	0.0561
	Australasia	3	0.0040	-0.0005	0.0085
	Multiple	16	0.0215	0.0111	0.0319
	Not Specified	1	0.0013	-0.0013	0.0039
Funding Source **Maximum two abstracted per trial*	Government	48	0.0646	0.0469	0.0823
	University	259	0.3486	0.3143	0.3829
	Hospital	273	0.3674	0.3327	0.4021
	Industry	237	0.3190	0.2855	0.3525
	Private Interest	48	0.0646	0.0469	0.0823
	Not Declared	0	0.0000	-	-
Industry Funding	Yes	237	0.3190	0.2855	0.3525
	No	506	0.6810	0.6475	0.7145
TRIAL CHARACTERISTICS					
Surgical Specialty	Cardiac	51	0.0686	0.0504	0.0868
	Dentistry	21	0.0283	0.0164	0.0402
	ENT	40	0.0538	0.0376	0.0700
	General Surgery	197	0.2651	0.2334	0.2968
	Neurosurgery	44	0.0592	0.0422	0.0762
	Obstetrics and Gynecology	71	0.0956	0.0745	0.1167
	Ophthalmology	83	0.1117	0.0891	0.1343
	Orthopedic	110	0.1480	0.1225	0.1735
	Plastic	43	0.0579	0.0411	0.0747
	Thoracic	23	0.0310	0.0185	0.0435
	Urology	25	0.0336	0.0206	0.0466
	Vascular	23	0.0310	0.0185	0.0435
Type of Intervention	Technique	426	0.5734	0.5378	0.6090
	Implanted Device	218	0.2934	0.2607	0.3261
	External Device	26	0.0350	0.0218	0.0482
	Dressing	28	0.0377	0.0240	0.0514
	Surgical vs. Non-Surgical	45	0.0606	0.0434	0.0778
Subject Age	Adult	623	0.8385	0.8120	0.8650
	Pediatric	24	0.0323	0.0196	0.0450
	Both	34	0.0458	0.0308	0.0608

	Not Specified	62	0.0834	0.0635	0.1033
TRIAL REGISTRATION					
Recruitment Status	Completed	543	0.7308	0.6989	0.7627
	Terminated	118	0.1588	0.1325	0.1851
	Enrolling by invitation	14	0.0188	0.0090	0.0286

3.4.5. Objective 2: Proportion of Published Trials

Using the step-wise hierarchical search strategy shown in Figure 1 there were a total of 364 published trials identified, generating a trial publication proportion of 0.49 (95% CI [0.454, 0.526]). There were an additional 49 trial protocols (0.066) that were not published but made raw data results available through the clinicaltrials.gov website.

The three search engines used to identify published trials were each searched using the trial registration number. The proportion of published trials found with each of the search engines and the author search strategy are shown in Table 7.

Table 7. Results of search for published trials with number of published trials found for each reported by search engine.

Search Engine	**Published Trials**	**Proportion**	**95% Confidence Interval**	
NCT number clinicaltrials.gov	158	0.43	0.38	0.48
NCT Number Google Scholar	259	0.71	0.66	0.76
NCT Number PubMed	158	0.43	0.38	0.48
Author Search on PubMed * *Conducted if no publication found with NCT number*	92	0.25	0.21	0.29

3.4.6. Analysis of Variables Associated with Trial Publication

Trial publication was set as a binary outcome of interest. For each of the variables of interest a univariable analysis was performed. Reference variables were set to be the most common category for each variable. An odds ratio greater than 1 in a category indicates the category was more often associated with publication than the reference variable. Significance of association was assessed using 95% confidence intervals.

Table 8. Univariable analyses of study variables for the outcome of trial publication. Each variable was assessed for independence using the Kruskal-Wallis test of independence with result shown as group p-value (KWPV). Odds ratios and 95% Wald confidence limits are reported for each categorical variable.

Variable (KWPV)	**Categorical Variable**	**Odds Ratio** (published)	**95% Wald Confidence Limits**	
TRIAL DEMOGRAPHICS				
Region p = 0.0008	North America	1.00	Ref	Ref
	Africa	2.375	1.112	5.072
	Asia	1.306	0.829	2.056
	Australasia	3.00	0.269	33.453
	Europe	2.041	1.457	2.858
	South America	2.850	1.281	6.342
	Multiple	0.900	0.319	2.542
	Not Specified	<0.001	<0.001	>999.999
Primary Funder p = <0.0001	University	1.00	Ref	Ref
	Government	0.886	0.383	2.053
	Hospital	0.781	0.550	1.108
	Industry	0.356	0.239	0.532
	Private Interest	0.950	0.462	1.952
Industry Funding p = < 0.0001	No	1.00	Ref	Ref
	Yes	0.480	0.349	0.658
TRIAL CHARACTERISTICS				
Surgical Specialty	General Surgery	1.00	Ref	Ref

p = <0.0001	Cardiac	0.977	0.522	1.827
	Dentistry	0.622	0.252	1.532
	ENT	0.619	0.313	1.224
	Neurosurgery	0.176	0.080	0.386
	Ob-Gyn	0.934	0.539	1.620
	Ophthalmology	0.367	0.215	0.626
	Orthopedics	0.684	0.428	1.094
	Plastics	0.330	0.164	0.664
	Thoracic	0.365	0.148	0.901
	Urology	1.216	0.512	2.886
	Vascular	0.889	0.372	2.126
	Multiple	0.228	0.060	0.868
Trial Intervention Type p = 0.0688	Technique	1.00	Ref	Ref
	Dressing	0.676	0.312	1.464
	External Device	1.443	0.640	3.252
	Implanted Device	0.684	0.492	0.950
	Surgery vs. Non-surgical	0.601	0.321	1.124
Age of Participants p = 0.0715	Adult	1.00	Ref	Ref
	Pediatric	1.889	0.815	4.382
	Both	1.619	0.803	3.264
	Not Specified	1.678	0.986	2.854
TRIAL REGISTRATION				
Recruitment Status p = <0.0001	Completed	1.00	Ref	Ref
	Active, not recruiting	0.895	0.528	1.515
	Terminated	0.170	0.103	0.281
	Suspended	0.393	0.071	2.164
	Enrolling by Invitation	0.314	0.097	1.015
Registration Timing p = 0.0029	Before	1.00	Ref	Ref
	After	1.977	1.358	2.880
	During	1.453	1.026	2.059
	Unknown	2.317	0.869	6.183
Trial Completion Date p = 0.029	< 2007	1.00	Ref	Ref
	2008	0.908	0.418	1.969
	2009	0.879	0.429	1.801
	2010	0.699	0.349	1.402
	2011	0.790	0.395	1.579
	2012	0.371	0.184	0.747
	Unknown	0.649	0.232	1.814

A multivariable analysis was performed using the same set of predictor variables in a full model. Reference variables were set to be the most common category for each variable. An odds ratio greater than 1 in a category indicates the categorical variable was more often associated with publication than the reference category. Significance of association was assessed using 95% confidence intervals. This model found that trials whose recruitment had been terminated were less likely to be published than those which had been fully completed, and that trials completed in 2012 were less likely to have been published than those completed in 2006.

Table 9. Multiple logistic regression model relating predictor variables with the outcome of trial publication. Model likelihood ratio test (X^2 = 177.87, df = 44, p < 0.0001).

Variable Wald ChiSq p value	**Categorical Variable**	**Odds Ratio** (published)	**95% Wald Confidence Limits**	
TRIAL DEMOGRAPHICS				
Region p = 0.1308	North America	1.00	Ref	Ref
	Africa	0.773	0.314	1.904
	Asia	0.777	0.438	1.378
	Australasia	7.265	0.466	113.306
	Europe	1.493	0.971	2.294
	South America	1.914	0.747	4.903
	Multiple	0.837	0.262	2.678
	Not Specified	<0.001	<0.001	>999.999
Primary Funder p = 0.1621	University	1.00	Ref	Ref
	Government	0.988	0.368	2.656
	Hospital	0.631	0.414	0.961
	Industry	0.512	0.243	1.079
	Private Interest	1.023	0.444	2.356
Industry Funding p = 0.7649	No	1.00	Ref	Ref
	Yes	0.905	0.470	1.743
TRIAL CHARACTERISTICS				
Surgical Specialty	General Surgery	1.00	Ref	Ref

p = <0.0001	Cardiac	1.605	0.763	3.377
	Dentistry	0.613	0.220	1.708
	ENT	0.443	0.208	0.944
	Neurosurgery	0.219	0.091	0.529
	Ob-Gyn	1.247	0.656	2.369
	Ophthalmology	0.304	0.167	0.553
	Orthopedics	0.739	0.426	1.282
	Plastics	0.312	0.138	0.704
	Thoracic	0.385	0.141	1.050
	Urology	1.650	0.624	4.365
	Vascular	1.289	0.462	3.601
	Multiple	0.275	0.067	1.131
Trial Intervention Type p = 0.7054	Technique	1.00	Ref	Ref
	Dressing	1.231	0.477	3.179
	External Device	0.957	0.379	2.413
	Implanted Device	0.943	0.619	1.438
	Surgery vs. Non-surgical	0.596	0.283	1.255
Age of Participants p = 0.3450	Adult	1.00	Ref	Ref
	Pediatric	1.358	0.528	3.492
	Both	1.717	0.736	4.009
	Not Specified	1.520	0.826	2.799
TRIAL REGISTRATION				
Recruitment Status p = <0.0001	Completed	1.00	Ref	Ref
	Active, not recruiting	1.178	0.652	2.129
	Terminated	0.137	0.077	0.244
	Suspended	0.300	0.044	2.033
	Enrolling by Invitation	0.187	0.048	0.719
Registration Timing p = 0.5719	Before	1.00	Ref	Ref
	After	1.221	0.770	1.936
	During	1.312	0.871	1.977
	Unknown	2.027	0.265	15.520
Trial Completion Date p = 0.0060	< 2007	1.00	Ref	Ref
	2008	0.831	0.345	1.999
	2009	0.613	0.271	1.388
	2010	0.660	0.297	1.468
	2011	0.654	0.294	1.455
	2012	0.294	0.131	0.656
	Unknown	0.472	0.071	3.153

3.4.7. Time to Publish and Journal Impact Factor

The mean time to publication for each trial was assessed as the difference between the earliest trial publication date (first of online or print publication date) and the trial completion date. The mean time to publication was 21.7 months [1.93, 47.08] and median time to publish was 20 months.

The impact factor of the journals was assessed based on their established impact factor in 2013 [67]. The median journal impact was 3.3.

A regression analysis was used to determine if there was an association between the journal impact factor and the time to publication. There was poor correlation between the impact factor and the outcome of publication.

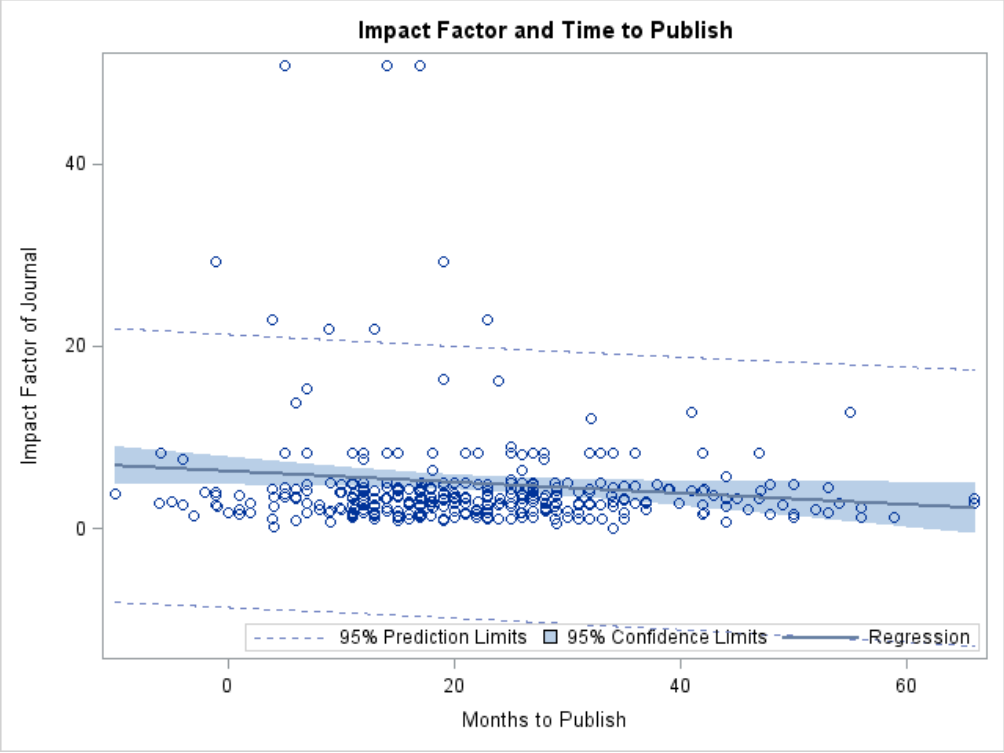

Figure 2. Regression analysis of time to publication (months) and journal impact factor for 264 publications (R^2 = 0.012).

3.5. Discussion

3.5.1. Protocol Identification and Database Creation

In Objective 1, we succeeded in creating a novel database of surgical RCTs which was built on the foundation of the clinicaltrials.gov database. Our final cohort of 743 surgical trials represented a small fraction, 4.25%, of all trial protocols downloaded from the site, only 0.5% of all trial protocols (139,063) registered with the site at the time we performed our search. This highlights two important results: (1) the number of registered surgical RCTs is very small and (2) a systematic search strategy is necessary to isolate relevant protocols.

When we accessed the clinicaltrials.gov dataset, we used a purposely broad search strategy to maintain maximum control over the inclusion criteria for our own dataset, and thus started with a set of 17,498 trial protocols. We eliminated a large number of studies (12,844, or 73%) based on study start dates and completion dates which fell outside of our range of interest, trial designs which were either non-randomized, or employed a single group assignment. Ultimately a hand search of 3,489 trials (20%) was required to constitute our final dataset of 743 trials.

By hand reviewing protocols, we were able to identify those which were truly evaluating a surgical intervention. This assessment could not have been performed efficiently using any of the advanced search strategies available through clinicaltrials.gov and highlights the importance of hand searching protocols, and one of the strengths of this study. The majority of the trials eliminated during this review were drug trials being performed using surgical subjects (46%). We did not consider these to be surgical studies as the surgical intervention itself was always held

constant. A smaller proportion of trials (12%) employed other non-surgical interventions, often physiotherapy and patient decision aids. Interestingly, there were nearly 20% of trials eliminated because the trial subjects were not undergoing any form of surgery; it was found that in many of these trials the key word “surgery” appeared as an exclusion criterion for the study. These exclusions suggest that our initial search strategy was sensitive enough to identify a near complete set of surgical RCTs registered with the website.

The descriptive statistics generated for the surgical trial protocols show trends in the data variables we sought to explore. Within the trial demographic domain we were interested in the region of trial conduct and the funding sources for the trials. Most of the studies were based in North America (40%) or Europe (35%). Only 2% were multi-continent studies. Interestingly, less than 0.5% of eligible protocols were based in Australasia. This should be interpreted as a cautionary note rather than an explanatory point. Requirements for protocol registration only require registration with one publically accessible database, the choice of database falls on the investigator. The low number of protocols registered from Australasia likely point to alternative registries, such as the *Australian New Zealand Clinical Trials Registry*.

Trial funding most often came from hospitals (37%), Universities (35%) and Industry (32%). A small number listed multiple funding sources. Our choice of categorical variables in this instance may not have been sensitive enough to identify all studies in which there was industry funding given the registry only requires the investigator to report the “primary funder”, not all funders. Furthermore, there was no way to determine whether hospital funding, or University funding, may have had

underlying associations with industry. It is conceivable that the number of trials receiving industry funding is in fact higher than the 32% we identified. The proportion of surgical trials sponsored by industry (32%) is approximately half of the 66% estimated for medical trials published in top journals[68].

Trial characteristics were another domain of interest. The majority of trials were conducted in general surgery (26.5%), Orthopedics (15%) and Ophthalmology (11%). Fewer trials were conducted in the sub specialized areas such as Urology, Vascular, and Thoracics (3% each). Nearly all trials enrolled subjects over 18 years of age (88%) with a much smaller number involving subjects under 18 years either as a part of the study population (5%) or exclusively (3%).

The type of surgical intervention was determined based on the description of the intervention provided by the study investigators. Many were technique trials (57%) or involved the comparison of a surgical intervention with a non-surgical intervention (6%). Medical devices either implanted (29%) or external (3.5%) and types of dressing (4%) made up the remainder of the trials.

The trial registration domain established in this thesis assessed the quality of the registered protocol. It was felt that an important quality indicator would be the timing of protocol registration relative to the start date of the trial. We hypothesized that this may be associated with trial publication in that those who registered after trial completion may have done so in order to satisfy journal criteria for publication. We found that a minority of trials (33%) were registered before the trial commenced, and that the remainder either registered during the conduct of the trial (37%) or after its completion (28%). This means that two-thirds of surgical trials did not follow the

clinicaltrials.gov policy of registering the protocol prior to randomizing the first participant. Given that we only considered trials eligible if they started after the ICMJE stated trial protocols must be registered, we expected that the majority would have registered in advance.

Only “closed” trials as defined by clinicaltrials.gov were eligible for this study. We had previously used dates to identify those studies whose outcome measures were registered to have been completed before December 31, 2012. Not surprisingly the majority of studies in our database were completed (73%) or terminated (16%). The group listed as “active, not recruiting” (8%) and “enrolling by invitation” (2%) were considered complete by our assessment and were considered publishable by our standard.

3.5.2. Proportion of Registered Trials Published (Objective 2)

Estimating the proportion of published trials was our second objective. Our search strategy identified a total of 364 published trials, a publication proportion of 0.49 (95% CI [0.454, 0.526]). This is in line with publication proportions estimated elsewhere in the literature[47,49]. Our method of searching three separate publication databases shows important distinctions and similarities. Interestingly, the Google Scholar search using the NCT number was the most sensitive way to find trials based on the registration number, with 71% of trials being found using this strategy. The additional search of the PubMed database using the author’s last name was a critical addition to the search strategy, adding a further 92 publications (25%) to the dataset.

The univariable analysis we performed was based on our hypothesis that trial publication would be influenced by the categorical variables we had identified a priori. We found that there were signals of association between publication and a number of these variables.

In the trial demographics, region, funding source, and industry funding source were all found to be statistically significant based on the Kruskal-Wallis test of independence. Relative to those trials conducted in North America, trials conducted in Europe (OR 2.04 [1.46, 2.86]), Africa (OR 2.8 [1.1, 5.07]), and South America [1.28, 6.34]) were more likely to be published. This difference may be attributable to the use of the clinicaltrials.gov registry for my thesis which is based in North America. One explanation could be that researchers seeking to register their trial outside of a regional database is done with the intent to garner international attention to their trial. These researchers may be demonstrating a commitment to good reporting and proper trial conduct to ensure their trials are regarded positively.

When university funding was used as a reference variable our data show that Industry funded trials were less likely to be published (OR 0.36 [0.24, 0.53]). When taken as a simple binary variable, industry funded trials were less likely to be published than those trials which did not have industry funding (OR 0.48 [0.35, 0.66]). This is an important finding, and speaks to a need for further work to identify what factors influence non-publication amongst industry funded trials. It would stand to reason that industry sponsored publications would be in line with the profitability of the sponsor. Previous research has estimated that trials associated with industry funding are up to 3 times more likely to have statistically significant results when

compared to trials without industry funding[69,70]. Therefore it is possible that the larger proportion of unpublished industry sponsored trials would be less likely to have statistically significant findings and may be evidence of publication bias.

For the trial characteristic domain, trial intervention type and participant age did not reach significance on the Kruskal-Wallis test. Surgical specialty was found to have an association with trial publications. Relative to studies conducted in General Surgery, trials conducted in Neurosurgery (OR 0.18 [0.08, 0.39]), Ophthalmology (OR 0.37 [0.22, 0.63]), Plastic Surgery (OR 0.33 [OR 0.16, 0.66]), Thoracic Surgery (OR 0.37 [0.15, 0.90]) and multiple specialties (OR 0.23 [0.06, 0.87]) were less likely to have been published. An explanation may be the relative lack of sub-specialty journals with limited capacity for publications which could restrict the options for researchers looking to publish work aimed at the sub-specialist audience. Another explanation could be an association with other variables such as industry funding or trial quality, associations which could be explained in future work using the STARTOPP database.

In the domain of trial registration, all three categorical variables showed significant differences. When compared to trials which had been completed, those trials which had been terminated were less likely to have been published (OR 0.17 [0.10, 0.28]). This difference suggests that there may be biases that cause terminated trials to go unpublished, a significant finding which is supported by the literature[71]. It is our contention that the lessons learned from terminated trials are vital to the progress of science. This is an area which warrants further study to identify the rationale of both trialists and journals for non-publication of such trials.

As we had hypothesized, there was a difference in trial publication based on the timing of protocol registration. Relative to those trials registered in advance of commencing enrollment, those trials which registered their protocols during the study (OR 1.45 [1.03, 2.06]) or after study completion (OR 1.98 [1.36, 2.88]) were more likely to have been published. This may suggest that trials are registered post-hoc to satisfy the ICMJE rules for protocol registration.

Trial completion data was only found to vary significantly for trials completed in 2012 relative to those completed before 2007, with those completed in 2012 less likely to have been published (OR 0.37 [0.18, 0.75]). This is reinforced by our finding that the mean time to trial publication was 22 months, and median was 20 months. This exceeds the recommended time to publication of 12 months (reference). This also points to a potential limitation in our study as trials completed in the year 2012 may have gone on to be published and that in fact the publication proportion may be higher than we have estimated.

The multivariable analysis was performed using all of the variables we had identified within each of the domains to control for important hypothesized risk factors. Within the trial demographics domain, none of the variables were found to have statistical significance. Within the trial characteristics domain, there were significant differences identified for the surgical specialty variable. Trials conducted in ENT (OR 0.44 [0.21, 0.94]), neurosurgery (OR 0.22 [0.09, 0.53]), ophthalmology (OR 0.30 [0.17, 0.55]), and plastic surgery (OR 0.31 [0.14, 0.70]) were all less likely to be published when compared to those conducted in general surgery.

Within the trial registration domain there were two variables which were significant on the multivariable analysis. For trial recruitment, the reference category was a trial registered as completed. For this variable, studies which were terminated (OR 0.14 [0.08, 0.24]) or enrolling by invitation (OR 0.19 [0.05, 0.72]) were each less likely to be published. The low proportion of publications for terminated trials found on multivariable analysis is further evidence in support of a bias against publication of such trials. Non-publication of terminated trials is an issue that should be addressed by researchers and journal editorial boards – particularly as there are often good reasons for trial termination (e.g. patient safety).

Within the trial completion date variable we found that studies completed in 2012 were less likely to be published relative to those completed before 2007 (OR 0.29 [0.13, 0.66]). This reflects evidence found on the univariable analysis and further supports the notion that the true proportion of trials being published in our cohort may be higher than what we determined based solely on a lag in time to publication.

We hypothesized that the trials published soon after completion would be published in high impact factor journals. A linear regression was performed by impact factor and time to publication in months. There was no evidence of the hypothesized negative correlation with an R^2 value of 0.012. Average journal impact factor was 4.9 (1.12, 13.71), the clustering of journals with little variability in the impact factor made discriminating a difference difficult. Previous research has drawn similar conclusions; trials with positive results are more likely to be published

early, but the impact factor of the publishing journal is not associated with the time to publication or the statistical significance of the result [72].

3.6. Conclusions

Surgical RCT protocols form a small fraction of the trials registered on clinicaltrials.gov. We succeeded in developing a novel database of surgical RCTs by abstracting trial protocols registered on the site. Within at least 12 months of study completion, the proportion of trials that have been published is 0.49. A univariable analysis showed evidence of an association between non-publication of completed trials and industry funding (OR 0.48 [0.35, 0.66]) and trials which are terminated early (OR 0.17 [0.10, 0.28]). There is also evidence to suggest that trial protocols registered during and after study completion are more likely to be associated with publication.

4. Assessment of Reported Outcome Measures: Objective 3

4.1. Introduction

When there is a discrepancy between the primary outcome measure stated in the trial protocol at the start of the trial and the primary outcome measure reported in the publication, selective outcome reporting bias is present. Objective 3 was to assess the proportion of publications with selective outcome reporting in registered surgical RCTs. We hypothesized that the proportion of selective outcome reporting would be similar to previous literature estimates that suggests 45% of trials show evidence of biased reporting[41,42].

In this phase of the study, we used the STARTOPP database of surgical trial protocols and linked publications created in the first objective. A standardized review of the published articles was conducted to look for associations between study variables and the presence of selective outcome reporting. As was done during the protocol abstraction (0) three domains of trial demographics, trial characteristics, and trial registration were assessed. Additionally, a fourth domain of trial results was used. For this domain, new categorical variables were introduced and identified by systematic review of the publications.

4.2. Objective 3: Materials and Methods

4.2.1. STARTOPP Surgical RCT Database

The STARTOPP database of surgical trials developed in Chapter 2 was used for Objective 3. All trials which had been identified as published were eligible for inclusion.

4.2.2. Sample Size Calculation

The literature estimate of 45% selective outcome reporting was used to generate our sample size estimate; we hypothesized that this would be the estimated rate. To estimate whether the true proportion was 0.45 within a margin of error of 0.05 using a two sided test, a sample of 270 trials was required (Appendix 3. Sample Size Calculation).

4.2.3. Selection of Trials for Outcome Reporting Assessment

For efficiency, a random sample of 270 trials was selected from the cohort of 364 published surgical RCTs in the STARTOPP database. A random number generator was used in Excel to identify the random sample. The NCT number served as the unique identifier for each protocol and its linked publication.

4.2.4. Development of Abstraction Form

A reported outcome was to be considered congruent, that is not selectively reported, if the primary outcome measure stated in the protocol was the same as the

primary outcome reported in the publication. The research team determined that selective outcome reporting existed when one of the following conditions existed:

1. Primary outcome (Publication) was secondary outcome (Protocol)
2. Primary outcome (Publication) was not registered in the Protocol
3. Primary outcome (Protocol) was not reported in the Publication

Additionally, the abstraction of publications would identify variables within the publication that were hypothesized to have potential association with reporting of outcomes. A new domain of assessment, “Trial Reporting”, was developed by the research team with a series of six categorical variables of interest:

1. Statistical design
2. Statistical significance of outcome
3. Favoured outcome in trial
4. Presence of sample size calculation
5. Achievement of sample size
6. Author concluding statement (endorsement of results)

The abstraction form (**Error! Reference source not found.** was produced to apture these elements, and corresponding variables were added to the STARTOPP database.

4.2.5. Abstraction of Published Trials

The protocol outcomes and published outcomes for all 270 trials were independently assessed by two reviewers (PG and SB). The first two sections (abstraction of outcomes from protocol; comparison of reported and registered outcomes) of the abstraction form were completed in full by the two reviewers (**Error! Reference source not found.**). These sections of the form were focused n assessing the primary and secondary outcomes listed in the protocols and publications. Disagreements were resolved by consensus between the two reviewers. Irreconcilable differences were arbitrated by a third party (DF) when required.

Additional study variables within the trial reporting domain listed, in 4.2.4 and identified in sections three and four of the abstraction form, were abstracted by a single reviewer (PG).

4.2.6. Statistical Analysis of Published Trials

SAS 9.3 software (2011, SAS Institute Inc., Cary, NC, USA) was used for all statistical analyses. The database was exported to SAS from Excel.

The primary outcome of interest was selective outcome reporting and measured as the proportion and 95% confidence intervals of published trials where selective outcome reporting was present. This was then compared to the literature

estimate using a comparison of proportions with a Z statistic and 95% confidence interval.

Selective outcome reporting was then analyzed as a binary variable to determine if there was an association with the categorical variables of interest. The four domains of trial demographics, characteristics, registration, and reporting with their respective variables were measured. Variables in the trial reporting domain were analyzed using descriptive statistics. The number, proportion, and 95% confidence intervals were generated for each of the variables described in 4.2.4.The Kruskal-Wallis test was conducted to determine if there was a difference in selective outcome reporting with respect to each variable and a univariable analysis was performed for each variable. The reference category for each variable was set as the most commonly encountered category. Odds ratios and 95% confidence intervals were generated for each of the univariable analyses.

Using the same binary outcome of selective outcome reporting, a multivariable analysis was performed using all of the categorical variables in the four domains of assessment. Odds ratios and 95% confidence intervals were generated for each of the variables.

Trial sample sizes were analyzed as continuous variables. Two linear regression analyses were performed to compare initial study sample size with the recruited sample size, the second to compare the recruited sample size with the impact factor of the publishing journal.

4.3. Objective 3: Results

4.3.1. Reporting of Primary Outcome

The proportion of trials where the reported primary outcome was identical in the protocol and publication was 0.756 (Table 10); subsequently the proportion where there is evidence of selective outcome reporting is 0.244 (95% CI [0.193, 0.295]) of publications. This proportion is significantly lower ($p<0.001$) than the 0.45 proportion of selective outcome reporting estimated in the literature [41].

Table 10. Assessment of the reported primary outcome in the publication compared to the primary outcome registered in the trial protocol.

Published Primary Outcome	**Number**	**Proportion**	**95% Confidence Interval**	
Primary Outcome in Protocol	204	0.7556	0.7043	0.8069
Secondary Outcome in Protocol	18	0.0667	0.0369	0.0965
Not Reported in Protocol	48	0.1778	0.1322	0.2234
Total	**270**	**1.00**		

A second analysis was conducted to assess the reporting of the registered primary outcome in the protocol. This assessment adds information on what proportion of primary outcomes went unpublished. As presented in Table 11, the proportion of published trials which did not report the primary outcome registered in the protocol is 0.078.

Table 11. Assessment of the reported primary outcome compared to the primary outcome registered in the trial protocol.

Primary Outcome in Protocol	**Number**	**Proportion**	**95% Confidence Interval**	
Primary Outcome in Publication	204	0.7556	0.7043	0.8069
Secondary Outcome in Publication	35	0.1296	0.0895	0.1697
Not Reported in Publication	21	0.0778	0.0459	0.1097
No Primary in Protocol	10	0.0370	0.0145	0.0595
Total	**270**	**1.00**		

The number of primary outcomes and secondary outcomes registered in the trial protocols was captured as a continuous variable. The mean, median and mode number of primary outcomes in the protocol was 1 (IQR = 0). The minimum number of outcomes registered was 0 and the maximum 9.

For secondary outcomes, the median number of outcomes was registered was 2 (IQR = 3). The minimum number of secondary outcomes registered was 0 and the maximum 42.

4.3.2. Descriptive Statistical Analysis of Trial Reporting Variables

Categorical variables in the trial reporting domain were described (Table 12). Previously abstracted data in the other three domains was not re-evaluated. This assessment was conducted during the analysis of the entire cohort of trials presented in Table 6.

Table 12. Number of publications and total proportion for each trial reporting variable for the 270 publications assessed.

Trial Variable	Category	Number	Proportion	95% CI	
Statistical Design	Superiority	240	0.8889	0.8514	0.9264
	Equivalence/Non-Inferiority	18	0.0667	0.0369	0.0965
	Not Reported	12	0.0444	0.0198	0.0690
Statistical Significance in Primary Outcome	Yes	142	0.5259	0.4663	0.5855
	No	124	0.4593	0.3999	0.5187
	Not Comparative	4	0.0148	0.0004	0.0292
Outcome Favours	Intervention	171	0.6333	0.5758	0.6908
	Control	48	0.1778	0.1322	0.2234
	Equivalence/Non-Inferiority	45	0.1667	0.1222	0.2112
	Not Comparative	6	0.0222	0.0046	0.0398
Sample Size Calculation	Yes	182	0.6741	0.6182	0.7300
	No	88	0.3259	0.2700	0.3818
Sample Size Fulfillment	100% Fulfillment	62	0.2296	0.1794	0.2798
	>100% Fulfillment	88	0.3259	0.2700	0.3818
	<100% Fulfillment	40	0.1481	0.1057	0.1905
	Indeterminate	80	0.2963	0.2418	0.3508
Study Conclusion	Recommends Intervention	116	0.4296	0.3706	0.4886
	Recommends Intervention with Caution	111	0.4111	0.3524	0.4698
	Refutes Intervention	43	0.1593	0.1156	0.2030

4.3.3. Elements Associated with Proper Reporting of Primary Outcome

Categorical variables for trial demographics, characteristics, registration, and published articles were independently assessed for their association with the reported primary outcome. In this analysis, an odds ratio of greater than 1 indicates

that there is a higher likelihood of selective outcome reporting. Significance of association was assessed with 95% confidence intervals.

Table 13. Univariable analyses of study variables for the outcome of selective outcome reporting. Each variable assessed for independence using the Kruskal-Wallis test of independence with result shown as group p-value (KWPV). Odds ratios and 95% Wald confidence limits are reported for each categorical variable.

Predictor Variable Kruskal-Wallis p-value	**Category**	**Odds Ratio**	**95% Wald Confidence Limits**	
TRIAL DEMOGRAPHICS				
Region p = 0.1113	North America	1.00	Ref	Ref
	Africa	0.364	0.076	1.746
	Asia	0.880	0.382	2.025
	Australasia	2.000	0.121	33.095
	Europe	0.444	0.231	0.854
	South America	0.167	0.021	1.343
	Multiple	0.400	0.045	3.579
	Not Specified	-	-	-
Primary Funder p = 0.1811	University	1.00	Ref	Ref
	Government	0.01	<0.001	>999
	Hospital	1.429	0.767	2.662
	Industry	0.747	0.319	1.748
	Private Interest	0.485	0.103	2.292
Industry Funding p = 0.7743	No	1.00	Ref	Ref
	Yes	1.096	0.586	2.051
TRIAL CHARACTERISTICS				
Surgical Specialty p = 0.5176	General Surgery	1.00	Ref	Ref
	Cardiac	1.063	0.374	3.027
	Dentistry	0.354	0.042	2.963
	ENT	0.245	0.030	1.989
	Neurosurgery	1.276	0.230	7.067
	Ob-Gyn	1.241	0.456	3.377
	Ophthalmology	0.608	0.187	1.970
	Orthopedics	1.718	0.761	3.877
	Plastics	3.190	0.734	13.877
	Thoracic	0.532	0.061	4.671
	Urology	1.367	0.324	5.763
	Vascular	1.254	0.247	6.368
	Multiple	<0.001	<0.001	>999
Trial Intervention	Technique	1.00	Ref	Ref

Type p = 0.1520	Dressing	0.506	0.058	4.450
	External Device	1.013	0.303	3.388
	Implanted Device	0.622	0.317	1.222
	Surgery vs. Non-surgical	<0.001	<0.001	>999
Age of Participants p = 0.8285	Adult	1.00	Ref	Ref
	Pediatric	1.626	0.393	6.727
	Both	1.477	0.491	4.447
	Not Specified	1.083	0.409	2.872
TRIAL REGISTRATION				
Recruitment Status p = 0.6895	Completed	1.00	Ref	Ref
	Active, not recruiting	1.115	0.447	2.779
	Terminated	1.593	0.462	5.495
	Suspended	3.185	0.196	51.786
	Enrolling by Invitation	<0.001	<0.001	>999
Registration Timing p = 0.3089	Before	1.00	Ref	Ref
	After	1.140	0.539	2.411
	During	1.353	0.656	2.791
	Unknown	3.800	0.850	16.998
Trial Completion Date p = 0.0525	< 2007	1.00	Ref	Ref
	2008	1.008	0.282	3.604
	2009	0.615	0.176	2.146
	2010	1.322	0.409	4.272
	2011	0.491	0.145	1.668
	2012	1.159	0.341	3.936
	Unknown	5.597	0.772	40.571
TRIAL REPORTING				
Statistical Design p = 0.0099	Superiority	1.00	Ref	Ref
	Non-Inferiority	0.401	0.090	1.798
	Not Reported	4.495	1.373	14.708
Statistically Significant Outcome p = 0.0013	Yes	1.00	Ref	Ref
	No	0.770	0.436	1.361
	Not Comparative	<0.001	<0.001	>999
Outcome Favours p = 0.0038	Intervention	1.00	Ref	Ref
	Control	1.115	0.518	2.401
	Non-Inferiority	1.694	0.816	3.516
	Not Comparative	18.750	2.123	165.583
Sample Size	Yes	1.00	Ref	Ref

Calculation p = 0.0979	No	1.623	0.913	2.884
Sample Size Achievement p = 0.5425	100%	1.00	Ref	Ref
	< 100%	1.276	0.485	3.362
	> 100%	1.159	0.536	2.508
	Indeterminate	0.745	0.351	1.581
Study Conclusion p = 0.2108	Recommends Intervention	1.00	Ref	Ref
	Refutes Intervention	1.950	0.905	4.201
	Recommends Intervention with Caution	1.113	0.597	2.077

A multivariable analysis was performed using the same set of predictor variables. Reference variables were set by the most common categorical outcome for each variable. An odds ratio greater than 1 in a category indicates the categorical variable was more often associated with selective outcome reporting than the reference variable. Significance of association was assessed using 95% confidence intervals.

Table 14. Multiple logistic regression model relating predictor variables with the outcome of selective outcome reporting. Model likelihood ratio test (X^2 = 89.106, df = 56, p = 0.0032).

Predictor Variable Wald ChiSq p value	**Category**	**Odds Ratio**	**95% Wald Confidence Limits**	
TRIAL DEMOGRAPHICS				
Region p = 0.1533	North America	1.00	Ref	Ref
	Africa	0.331	0.026	4.159
	Asia	1.388	0.381	5.058
	Australasia	0.256	0.008	7.927
	Europe	0.296	0.102	0.861
	South America	0.458	0.030	6.989
	Multiple	1.285	0.042	39.206
	Not Specified	-	-	-
Primary Funder	University	1.00	Ref	Ref

p = 0.2185	Government	<0.001	<0.001	>999.999
	Hospital	1.114	0.449	2.761
	Industry	0.215	0.036	1.290
	Private Interest	0.257	0.034	1.950
Industry Funding p = 0.9333	No	1.00	Ref	Ref
	Yes	1.063	0.256	4.413
TRIAL CHARACTERISTICS				
Surgical Specialty p = 0.4968	General Surgery	1.00	Ref	Ref
	Cardiac	0.656	0.143	3.006
	Dentistry	0.207	0.015	2.804
	ENT	0.094	0.009	1.016
	Neurosurgery	1.149	0.126	10.486
	Ob-Gyn	1.608	0.377	6.862
	Ophthalmology	0.191	0.035	1.047
	Orthopedics	1.193	0.393	3.622
	Plastics	4.386	0.452	42.573
	Thoracic	0.769	0.060	9.926
	Urology	0.534	0.021	13.870
	Vascular	1.330	0.135	13.155
	Multiple	<0.001	<0.001	>999.999
Trial Intervention Type p = 0.5682	Technique	1.00	Ref	Ref
	Dressing	0.485	0.036	6.571
	External Device	0.699	0.132	3.714
	Implanted Device	0.424	0.156	1.154
	Surgery vs. Non-surgical	<0.001	<0.001	>999.999
Age of Participants p = 0.4394	Adult	1.00	Ref	Ref
	Pediatric	2.260	0.441	11.578
	Both	0.723	0.198	2.645
	Not Specified	0.723	0.198	2.645
TRIAL REGISTRATION				
Recruitment Status p = 0.8624	Completed	1.00	Ref	Ref
	Active, not recruiting	1.503	0.437	5.173
	Terminated	1.657	0.278	9.865
	Suspended	6.059	0.136	270.664
	Enrolling by Invitation	<0.001	<0.001	>999.999
Registration Timing p = 0.6824	Before	1.00	Ref	Ref
	After	0.661	0.212	2.061
	During	1.263	0.468	3.408
	Unknown	<0.001	<0.001	>999.999
Trial Completion Date	< 2007	1.00	Ref	Ref

p = 0.2600	2008	0.561	0.101	3.100
	2009	0.346	0.061	1.957
	2010	0.740	0.135	4.062
	2011	0.160	0.028	0.913
	2012	0.460	0.071	2.971
	Unknown	>999.999	<0.001	>999.999
TRIAL REPORTING				
Statistical Design	Superiority	1.00	Ref	Ref
	Non-Inferiority	0.481	0.052	4.411
p = 0.4715	Not Reported	3.115	0.346	28.087
Statistically Significant Outcome	Yes	1.00	Ref	Ref
	No	0.737	0.237	2.290
	Not Comparative	1.186	<0.001	>999.999
p = 0.8699				
Outcome Favours	Intervention	1.00	Ref	Ref
	Control	1.237	0.347	4.404
	Non-Inferiority	3.245	0.861	12.226
p = 0.3479	Not Comparative	>999.999	<0.001	>999.999
Sample Size Calculation	Yes	1.00	Ref	Ref
	No	2.085	0.293	14.840
p = 0.4631				
Sample Size Achievement	100%	1.00	Ref	Ref
	< 100%	0.272	0.066	1.123
	> 100%	0.882	0.311	2.501
p = 0.3101	Indeterminate	0.693	0.082	5.893
Study Conclusion	Recommends Intervention	1.00	Ref	Ref
	Refutes Intervention	2.348	0.570	9.677
p = 0.2163	Recommends Intervention with Caution	0.833	0.282	2.466

4.3.4. Assessing Study Conclusions

Within the trial reporting domain there were three variables which were considered related by the thesis advisory subcommittee. These variables were the study's concluding statement, statistically favoured outcome, and the overall statistical significance of the outcome variable. In Table 15 we report the number of

trials that fell within each category. For each trial both the favoured outcome, or the direction of effect, was evaluated as well as the statistical significance of the outcome. Results were considered statistically significant in a trial's reported outcome when the p-value was less than 0.05, or the reported 95% confidence interval did not cross 1.

Table 15. Assessment of author conclusion when compared to both the direction of effect (outcome favours) and statistical significance of the primary outcome measure; reported by number of trials and proportion of trials per conclusion group.

Study Conclusion	**Outcome Favours**	**Statistically Significant**	**Trials** (n)	**Group Proportion**	**95% CI**	
Recommends Intervention (N = 115)	Intervention	Yes	92	0.8000	0.727	0.873
		No	7	0.0609	0.017	0.105
	Control	Yes	1	0.0087	-0.008	0.030
		No	5	0.0435	0.006	0.081
	Equivalence	Yes	6	0.0522	0.012	0.093
		No	4	0.0348	0.001	0.068
Recommends Intervention with Caution (N = 106)	Intervention	Yes	28	0.2642	0.180	0.348
		No	32	0.3019	0.215	0.389
	Control	Yes	5	0.0472	0.007	0.088
		No	16	0.1509	0.083	0.219
	Equivalence	Yes	3	0.0283	-0.003	0.060
		No	22	0.2075	0.130	0.285
Refutes Intervention (N = 43)	Intervention	Yes	0	0.0000	-	-
		No	12	0.2791	0.145	0.413
	Control	Yes	6	0.1395	0.036	0.243
		No	15	0.1415	0.037	0.246
	Equivalence	Yes	1	0.0233	-0.022	0.068
		No	9	0.2093	0.088	0.331

4.3.5. Assessment of Sample Sizes

The sample size reported in the protocol and the sample recruited in the publication were compared. In 262 publications this comparison could be made, in

the 8 instances where the comparison could not be made there was no sample size reported in the protocol. The mean difference between the two was 6.8 with 95% CI [-3.4, 16.9]. The association is shown in the figure below.

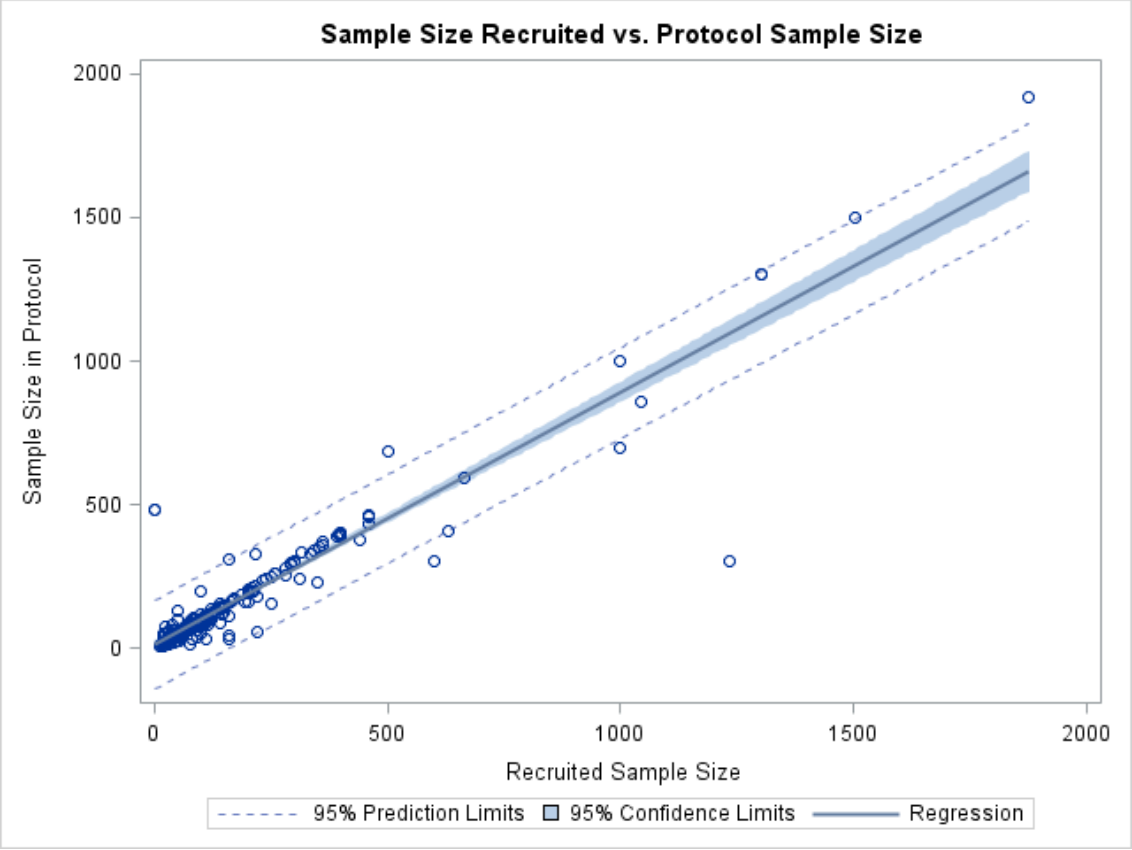

Figure 3. Linear regression of the sample size reported in the protocol and the sample size recruited as reported in the final publication (R^2 = 0.875).

Sample size estimates were assessed in the published trials. In 190 (70%) of the publications there was a reported sample size estimate, and in 182 of the trials (67.4%) there was a sample size calculation provided. In Figure 4 we report the comparison between these two sample sizes. The mean difference between the two was 7.2 with 95% CI [-6.3, 20.7]. The correlation is shown in the figure below.

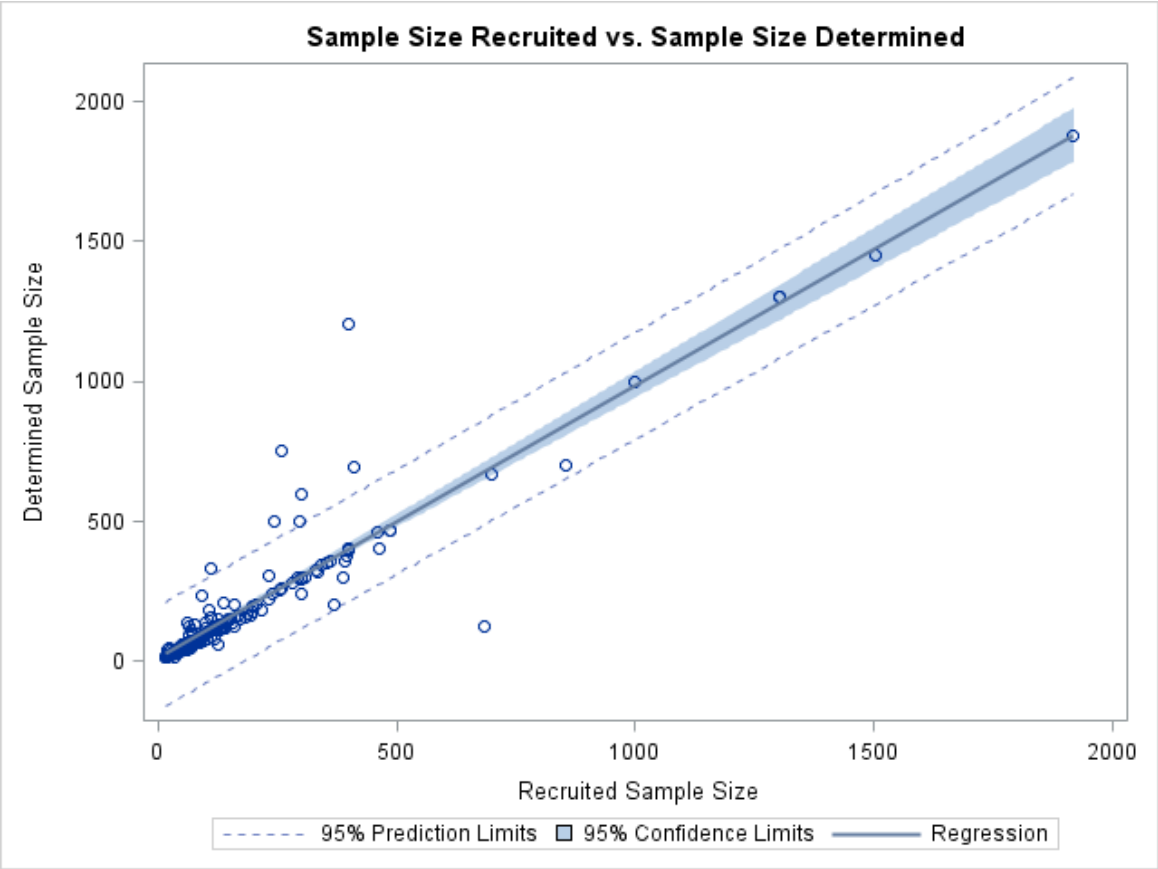

Figure 4. Linear regression of the estimated sample size reported in the publication and the sample size recruited (R^2 = 0.869).

An analysis was conducted comparing the recruited sample size with the impact factor of the publishing journal. Impact factors were available for 263 of the publications at the time of writing[67].

Figure 5. Linear regression of the sample size recruited in the study and the impact factor of the publishing journal (R^2 = 0.041).

4.4. Discussion

Detecting evidence of selective outcome reporting requires insight on the researcher's primary research objective. The cohort of trial protocols generated in this research project functioned in this capacity. Previous studies have used evaluated samples of published trials and then retrospectively review registered protocols associated with those trials. Our study was unique in that it captured a larger sample of protocols, protocols which were required to establish an a priori research objective.

We hypothesized that this prospective technique would be sensitive in detecting evidence of selective outcome reporting. In the 270 trials we evaluated we established a selective outcome reporting proportion of 0.244 which was significantly lower ($p<0.0001$) than the estimated proportion of 0.45 previously reported. Importantly, this proportion is reflective of the published outcomes. As shown earlier, over half of the studies which were completed went unpublished. This low proportion of selective reporting may reflect more stringent journal editing which is eliminating the selective reporting phenomenon.

The explanation for this difference may be due to differences between the manner in which selective outcome reporting was detected in this trial compared to the trial conducted by Hannink[41]. In the Hannink trial, they did not include trials which had registered their protocols during or after trial conduct, a criterion we did not apply. Our study included all trials regardless of when the trial was registered, though in doing so we allowed for trials which perhaps were registering outcomes they had already deemed suitable for publication irrespective of whether they were a

priori outcomes. This could lower our overall rate of selective outcome reporting. In addition, they only used so called “high power” journals, limiting the overall pool of trials and generalizability. Again, by capturing publications from all journals, our study offers a more robust estimate of the selective outcome reporting rate in the published literature.

Among the trials where there was evidence of selective outcome reporting, 48 (0.73) reported a primary outcome which was not registered in the protocol. This is a break from the scientific method whereby the research study should be designed to answer the research question. In these circumstances, it is impossible for the reader to interpret whether the result is spurious, and to determine whether there is any clinical relevance to the reported outcome. In the circumstances where a secondary outcome is promoted to the primary study outcome, as existed in 18 (27%) of the trials, the reader is able to interpret the relative clinical significance of the new outcome, however whether the statistical significance is spurious would be difficult to ascertain given sample sizes would not have been estimated based on the secondary outcome of interest.

When evaluating selective outcome reporting based on the protocol objectives, the overall proportion of inconsistent outcome measures remains the same. The subtle difference of looking forward allowed us to detect primary outcome measures which were dropped altogether from the published paper. In 21 (7.8%) of the publications, the registered primary outcome was eliminated entirely. This should be flagged as a major violation of protocol when one considers that

studies are granted ethical approval based mainly on the clinical importance of the primary research objective.

The protocols were fairly homogenous with respect to the number of registered primary outcomes. One outcome was registered for the majority of trials (IQR was 0). There was a range however, with up to 9 primary outcomes registered for a single trial. In circumstances where there was more than one primary outcome, the one on which the sample size calculation was based was first used to determine which was the main outcome. The registration of secondary outcome measures was more varied. Between 0 and 42 secondary outcome measures were registered.

The research team identified six categorical variables of interest which were abstracted from the published articles. The statistical design of this cohort of surgical RCTs was mainly superiority trials (89%), with a minority being equivalence or non-inferiority trials (7%). Sample size calculations were provided in 67% of publications. This number is surprisingly low, and confirms that at least 33% of trials are deficient in reporting this crucial element identified in the CONSORT guidelines[13]. Such a deviation indicates that not only are investigators falling short of the guidelines, but the journals and editors are too. This brings into question whether journals should be responsible for adjudicating trial conduct and reporting, or whether there is a need for oversight of Journals themselves. Such oversight might be warranted if the situation of poor compliance with the CONSORT statement continues to be an issue.

Statistically significant outcomes were reported in half of the published trials (53%). This is an interesting finding which supports one of the hypotheses that there would be a large proportion of trials published because they presented a statistically significant primary outcome measure. However, given that we are unable to assess the statistical significance of outcome measures in trials which were unpublished, we cannot make any definitive conclusions. Other authors have investigated this association and estimated that 41% of negative (statistically insignificant) trials are published and 73% of positive trials are published [73].

Further to this, the concluding statement and author recommendation was assessed for each published article. In 43% the authors definitively endorsed the intervention under study. A further 41% at least recommended the intervention, often with the caveat that further investigation was required to determine what the statistical or clinical significance was. These numbers were further evaluated in Table 15 where the statistical significance of the outcomes was compared to the study conclusions. Here we found that in studies where the authors recommended their intervention, 80% had found statistical significance in favour of their intervention. These 92 trials represent only 34% of the total published articles. By this standard, we concluded that there were situations in which authors made definitive conclusions endorsing their interventions when they have not demonstrated statistically significant evidence in favour of the intervention. In one instance, the authors not only failed to demonstrate statistical significance they had in fact showed evidence that was conclusively against the intervention under study. This is in keeping with the “spin” that is added to publications to enhance impact [74].

Selective outcome reporting was evaluated as a binary variable, and a univariable analysis was performed in an effort to identify factors which may be associated with poor outcome reporting. Interestingly, the domains of trial demographics, trial characteristics, and trial registration which had shown evidence of association during the assessment of publication status were no longer significant when reporting outcomes were assessed. None of the variables within these domains showed evidence of difference when evaluated with a Kruskal-Wallis test.

Within the trial reporting domain there were statistically significant differences within three groups. Trials which did not report a statistical design had greater instances of selective outcome reporting than trials which used a superiority design (OR 4.495 [1.37, 14.71]). The other two variables which showed group differences were ‘statistically significant outcome’ and ‘outcome favours’. However interpretation of within group differences did not detect any meaningful result.

Perhaps most important was that there was no association found between the study’s concluding statement and selective outcome reporting. This suggests that the conclusions drawn from this cohort of surgical RCTs is based primarily on properly reported outcomes.

The multivariable analysis used all of the variables within all four of the domains. There were no variables identified as significantly associated with the outcome.

Study sample sizes were assessed to determine how the protocol sample sizes correlated with the recruited sample sizes. We identified strong correlation (R^2 = 0.875) between the two, and strong correlation between the recruited sample size

and determined sample size in the published article (R^2 = 0.869). This result confirms that studies adhere closely to their sample size estimates. Finally we compared the recruited sample size with the journal impact factor. We had hypothesized that large studies may be more likely to be published in high impact factor journals. We did not find any evidence to support this, with poor correlation (R^2 = 0.041).

4.5. Conclusions

The proportion of selective outcome reporting in surgical RCTs is lower than previous estimates published for surgical trials. In our cohort of surgical RCTs the proportion of selective outcome reporting was 24.4%. We did not detect any variables which were conclusively associated with selective outcome reporting.

5. Summary and Future Directions

The three main objectives of this thesis were to:

1. Develop a database of surgical randomized controlled trials
2. Determine the publication proportion of registered and completed surgical RCTs
3. Determine the proportion of publications with selective outcome reporting in surgical RCTs

We successfully completed all three objectives by utilizing the publically accessible clinicaltrials.gov trial registry database. By systematically reviewing trial protocols and by systematically searching electronic databases of publications we were able to create a set of linked surgical RCT protocols and publications to complete our first objective. In answering our second objective we determined that the proportion of publication for completed surgical RCTs is 0.49, and that the average time to publication was 22 months from trial completion. To answer our third objective we compared the registered primary outcome with the reported primary outcome and determined that the proportion with changed primary outcomes is 0.244, significantly less than expected ($p<0.001$).

In constructing our database we built off the variables available on clinicaltrials.gov using a set of new categorical variables which were abstracted from the protocols and published articles. There were associations found that suggested trials with industry funding were more likely to be unpublished when compared to those without industry funding (OR 0.48 [0.349, 0.658]), or when compared to those

that were funded by universities (OR 0.36 [0.24, 0.53]). Another association suggests that when compared to protocols registered before the trial start date those registered during (OR 1.45 [1.03, 2.06]) and after (OR 1.98 [1.36, 2.88]) the study were more likely to be published. These findings suggest that there may be underlying publication biases, and consequently unethical research methodology as outlined by the Committee on Publication Ethics endorsed by journals such as *Annals of Surgery* [75]. When selective outcome reporting was used as the outcome of interest there were no clear associations that could be drawn from the variables we had evaluated.

The requirement for trial protocol registration made our study possible; however, as evidenced by recent reviews of adherence to reporting guidelines such as the CONSORT criteria[10]**,** there is imperfect adherence to this guideline. This discrepancy may be due to researchers not registering all details because they believe them unimportant, or because there is no oversight on elements included in the protocols. In addition to imperfect trial registration, the quality of trial protocols is highly variable due to the fact that there are few required elements in trial protocol registration and due to the fact that protocols must be added and maintained by the principal investigators with little oversight on the accuracy of the data. The accessibility of the clinicaltrials.gov database has allowed us to generate a database of nearly 750 trial protocols. In the future we will be able to add to our database by systematically reviewing protocols which were registered after the dates we had restricted. This will allow us to conduct research to look at how trial protocol registry changes over time, and to determine whether protocol quality is changing and if

there is suggestion of changing rates of publication or outcome reporting. The design of our database also gives the opportunity to add trials registered with different registries, such as Europe's clinicaltrialsregister.eu, and Australia's anzctr.org.au registry. By adding from these registries we will strengthen our estimation of global reporting of surgical RCTs.

While we succeeded in answering our questions on proportion of publication and selective outcome reporting a more pressing question remains: what are the biases that limit full publication and drive investigators to change their outcome measures? Getting at these motivations will be a focus of future work.

While searching for published trials, I also attempted to collect as many email addresses as I could for authors whose registered protocols had gone unpublished. This was not always possible, as more often than not contact information for the principle investigator was left off protocols. I was left to then search for authors by name on PubMed and Google Scholar, a difficult and imperfect technique. Regardless I was able to populate a list of over 200 email addresses for unpublished trial protocols. Previous work by Chapman found that frequently investigators of unpublished trials are not contactable[76]. More recently Smyth et al. have conducted qualitative research suggesting that non-publication may be impacted by investigators trying to publish results in prestigious journals, thus impeding the timing of publication [77].

Developing a survey questionnaire to address questions of publication bias would be a delicate task; however, candid responses would be invaluable. Was your work rejected for publication, and why? Did you choose not to publish your

work, and why? Was there pressure from industry sponsors to leave work unpublished? Did your results go against your hypothesis? Did your outcome measures reach statistical significance?

As the focus on systematic reviews[78] continues to grow in epidemiology, the importance of complete and proper outcome reporting will progress in tandem. A simplistic tenet of research is that if it isn't published then it never happened. Now with protocol databases we should be consider altering this view. If the research question is important enough to design a trial and subject patients to inherent risks and benefits then it should be unquestionably important enough for publication. We should encourage a new paradigm whereby if a trial protocol is registered then it should result in trial publication.

6. Lessons Learned and New Perspectives

My first exposure to the scientific method was in grade 8. My then science teacher, Ms. Myles, tasked us with designing a question that we were subsequently tasked with answering. I chose to study erosion. My materials were five rocks of similar size, four types of citrus fruit (of similar size), one scale, and my parent's whirlpool bathtub. Needless to say, this erosion study did not make waves in the literature; but that was not the point. Education is a process, the discoveries made on the journey can be inconsequential or they can be life changing. Completing this thesis was an extension of this learning process, as important personally as it was professionally.

A research question does not have to change the world; however a study which enrolls human participants can have profound implications on one person's world. It is why my personal philosophy that a researcher holds a fiduciary duty to the research subject to follow through and publish results. In performing this work I at times felt like a cynical auditor, looking for the errors and deficiencies in the labour of others. I read over 3400 research protocols and over 350 published trials. This process confirmed that research is hard work, the global body of evidence is increasing exponentially, and no study is exactly the same as another. I was not surprised to find incomplete protocols not linking to publications: there is stiff competition. However I would posit that this difficulty to publish should either be ameliorated by the journals or studies should be focused on answering "publishable" questions.

In presenting this work at various research days and meetings I have learned that not everyone shares this perspective. One memorable question from a colleague was to the effect of 'So what? If the research wasn't published, why would I care?' I understood then what isolation felt like. Had I wandered into an ivory tower of academic obscurity where I was to reside alone fighting, for justice in the literature? This is, fortunately, not the case. I have found solace in the fact those furthest from scientific research are often the most intrigued by the questions I was asking. With a rudimentary understanding of the statistical implications they understood the gravity of neglecting an unpublished body of literature as big, or possibly even bigger than the published works.

"Insanity: doing the same thing over and over again and expecting different results" - Albert Einstein

I believe there is much to learn from those trials which are not published. It is almost certain that the question will be posited by someone else in the world, someone who would benefit from basic answers to questions of feasibility, safety, and effect. With the hundreds of billions of dollars being invested in research we cannot afford to reinvent ourselves.

My work with the University of Ottawa Department of Epidemiology and Community Medicine taught me how to set a goal and achieve it; with the real lessons learned from the bumps in the road. Designing a database is easy; designing an effective database is the challenge. I was fortunate to work with two mentors, Dean Fergusson and Husein Moloo, who ensured the variables I recorded were both comprehensive and important. Despite this there were a few instances

where I found myself going back and reviewing protocols and publications again to tease out new variables of interest. When completing a project of this magnitude it is a task which is best done from the outset.

Finding publications is an art. For others it is a science. I benefited from the instruction of librarians both at The Ottawa Hospital and the University of Ottawa who taught me how to search through the volume of published trials on PubMed and Google Scholar using rigorous techniques. By far the hardest task was locating trials when there was no linked NCT number. This required searching through volumes of work tied to the principle investigator to identify the study under question. This would then require accessing abstracts and full publications to determine if it was indeed the same study. This took a substantial amount of time to perform, particularly in the cases where no study could be located. This search yielded 92 additional publications, but could not locate 379. If I were to conduct this study again I am not sure I would use the same approach.

In this study the use of categorical variables was essential. Designing categorical variables for this study was a big challenge, and in performing the analysis I realized that the reader could be left to question their meaning and significance. I feel that the choice of variables reflects the exploratory nature of this study. The measure of association found between industry and publication for instance, was meant to generate hypotheses for future work, not reach definitive conclusions. The finding that most researchers endorse the use of their intervention is interesting, but little more. I hope that this work will help guide future

investigations into these areas with an eye to teasing out the variables that warrant closer study.

I have returned to my life as a general surgery resident with a fresh perspective on evidence based medicine. Having sorted through thousands of trials I understand the vastness of the literature. My work is now focused on translating what is known into practice, with the benefit of knowing “best practice” is a distilled fraction of the evidence gained from research globally. For some questions the answers have been answered conclusively. But for most the answer is unclear, and I am quick to remind myself of the fog that is added from the unpublished and selectively published work.

7. Conclusion

The work presented in this thesis demonstrates that surgical randomized controlled trials are often unpublished, and in those that are unpublished a significant number show evidence of selective outcome reporting. The reasons behind these discrepancies remain uncertain; however we have found evidence that funding sources, surgical specialties, and year of publication are variables which are associated with these differences. Future work is warranted to determine what the motivations are that lead to unpublished research, and research that varies from its stated objectives.

Acknowledgements

I entered the epidemiology program at the University of Ottawa with an understanding that I would learn a set of skills that would transform me into a "surgeon-scientist". I knew that it would be a challenging transformation and one which would be shaped by the teachers and colleagues I would spend my two years with. I was fortunate to study under the tutelage of a superb research team: Dean Fergusson, Husein Moloo, and Rodney Breau. Thank you for your deft guidance throughout this process, you have instilled in me a passion and capacity that will be with me always.

To my office mates Michaël and Shane: thank you both for sharing your experience, and mentoring me through this thesis process. Both of you are accomplished physician-scientists who set a benchmark locally for professionalism and accomplishment in your fields. I am honoured to call you my friends.

I am grateful to the funding support to complete this project which I received from the Division of General Surgery, Clinician Investigator Program and the University of Ottawa. Research is a valuable enterprise, and I hope you will see in me the rewards of your investment.

For my wife Ashley, an accomplished individual who knows what it takes to reach her goals, thank you for pushing me to achieve my own. Your unwavering, loving support has brought this to fruition. Now the hard work begins!

And finally to Jack. You will be among those who inherit this world. We cannot be perfect, so we can always be better.

References

1 Spitzer W, Ronald J, Bayne D, Charron K, Fletcher S, Frappier-Davignon L, *et al.* The periodic health examination. *CMAJ*. 1979; **121**: 1193–1254.

2 Solomon M, McLeod R. Should we be performing more randomized controlled trials evaluating surgical operations? *Surgery*. 1995; **118**: 459–467.

3 Horton R. Surgical research or comic opera: questions, but few answers. *Lancet*. 1996; **347**: 984–985.

4 Hardin WD, Stylianos S, Lally KP. Evidence-based practice in pediatric surgery. *J Pediatr Surg*. 1999 May; **34**: 908–912.

5 Ahmed Ali U, van der Sluis PC, Issa Y, Habaga IA, Gooszen HG, Flum DR, *et al.* Trends in worldwide volume and methodological quality of surgical randomized controlled trials. *Ann Surg*. 2013 Aug; **258**: 199–207.

6 Ho PM, Peterson PN, Masoudi F a. Evaluating the evidence: is there a rigid hierarchy? *Circulation*. 2008 Oct 14; **118**: 1675–1684.

7 Fleiss J, Gross A. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique. *J Clin Epidemiol*. 1991; **44**: 127–139.

8 Moher D, Pham B, Jones A, Cook DJ, Jadad a R, Moher M, *et al.* Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998 Aug 22; **352**: 609–613.

9 Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ*. 2001; **323**: 42–46.

10 Falagas ME, Grigori T, Ioannidou E. A systematic review of trends in the methodological quality of randomized controlled trials in various research fields. *J Clin Epidemiol*. Elsevier Inc.; 2009; **62**: 227–231.

11 Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999 Sep 15; **282**: 1054–1060.

12 Begg C, Cho M, Eastwood S, Els D, Horton R, Moher D, *et al.* Improving the Quality of Reporting of Randomized Controlled Trials The CONSORT Statement. *JAMA*. 1996; **276**: 637–639.

13 Moher D, Schulz K, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *Ann Intern Med*. 2001; **134**: 657–662.

14 Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, *et al.* CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010 Mar 23; **340**: c869.

15 Altman DG. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions to authors. *BMJ*. 2005; **330**: 1056–1057.

16 Hopewell S, Altman DG, Moher D, Schulz KF. Endorsement of the CONSORT Statement by high impact factor medical journals: a survey of journal editors and journal 'Instructions to Authors'. *Trials*. 2008 Jan; **9**: 20.

17 Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials. *JAMA*. 2001; **285**: 1992–1995.

18 Hopewell S, Dutton S, Yu L. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ*. 2010; **340**: c723.

19 Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Cochrane Database Syst Rev*. Systematic Reviews; 2012; **1**: 60.

20 Schulz KF, Chalmers I, Grimes D a, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA*. 1994 Jul 13; **272**: 125–128.

21 Moher D, Altman DG, Schulz KF, Elbourne DR. Opportunities and challenges for improving the quality of reporting clinical research: CONSORT and beyond. *CMAJ*. 2004 Aug 17; **171**: 349–350.

22 De Angelis CD, Drazen JM, Frizelle FA, Hoey J, Horton R. Clinical trial registration. *JAMA*. 2004; **292**: 1363–1364.

23 De Angelis CD, Drazen JM, Frizelle FA, Hoey J, Horton R. Is this clinical trial fully registered? *JAMA*. 2005 Jun 15; **293**: 2927–2929.

24 Altman DG. Poor-Quality Medical Research. *JAMA*. 2002; **287**: 2765–2767.

25 Moher D, Bernstein A. Registering CIHR-funded randomized controlled trials: a global public good. *CMAJ*. 2004; **171**: 750–751.

26 Steinbrook R. Public registration of clinical trials. *N Engl J Med*. 2004; **351**: 315–317.

27 Strech D. Normative arguments and new solutions for the unbiased registration and publication of clinical trials. *J Clin Epidemiol*. Elsevier Inc; 2012; **65**: 276–281.

28 Dickersin K, Rennie D. Registering clinical trials. *JAMA*. 2003; **290**: 516–523.

29 Chalmers I, Glasziou P, Godlee F. All trials must be registered and the results published. *BMJ*. 2013; **346**: f105.

30 McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *BMJ*. 2002 Jun 15; **324**: 1448–1451.

31 Balasubramanian SP, Wiener M, Alshameeri Z, Tiruvoipati R, Elbourne D, Reed MW. Standards of reporting of randomized controlled trials in general surgery: can we do better? *Ann Surg*. 2006 Nov; **244**: 663–667.

32 Gurusamy KS, Gluud C, Nikolova D, Davidson BR. Assessment of risk of bias in randomized clinical trials in surgery. *Br J Surg*. 2009 Apr; **96**: 342–349.

33 Chan A-W, Krleza-Jerić K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ*. 2004 Sep 28; **171**: 735–740.

34 Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of Registered and Published Primary Outcomes in Randomized Controlled Trials. *JAMA*. 2009; **302**: 977–984.

35 Li X-Q, Yang G-L, Tao K-M, Zhang H-Q, Zhou Q-H, Ling C-Q. Comparison of registered and published primary outcomes in randomized controlled trials of gastroenterology and hepatology. *Scandanavian J Gastroenterol*. 2013; **48**: 1474–1483.

36 Smith SM, Wang AT, Pereira A, Chang RD, McKeown A, Greene K, *et al.* Discrepancies between registered and published primary outcome specifications in analgesic trials: ACTTION systematic review and recommendations. *Pain*. International Association for the Study of Pain; 2013; **154**: 2769–2774.

37 Chan A-W, Hrobjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials. *JAMA*. 2004; **291**: 2457–2465.

38 Redmond S, Von Elm E, Blümle A, Gengler M, Gsponer T, Egger M. Cohort study of trials submitted to ethics committee identified discrepant reporting of outcomes in publications. *J Clin Epidemiol*. 2013; **66**: 1367–1375.

39 Dwan K, Altman DG, Cresswell L, Blunderll M, Gamble C, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Libr*. 2011; **4**: 1–50.

40 Killeen S, Sourallous P, Hunter I a, Hartley JE, Grady HLO. Registration rates, adequacy of registration, and a comparison of registered and published primary outcomes in randomized controlled trials published in surgery journals. *Ann Surg*. 2014; **259**: 193–196.

41 Hannink G, Gooszen HG, Rovers MM. Comparison of registered and published primary outcomes in randomized clinical trials of surgical interventions. *Ann Surg*. 2013 May; **257**: 818–823.

42 Rosenthal R, Dwan K. Comparison of randomized controlled trial registry entries and content of reports in surgery journals. *Ann Surg*. 2013 Jun; **257**: 1007–1015.

43 Chahal J, Tomescu SS, Ravi B, Bach BR, Ogilvie-Harris D, Mohamed NN, *et al.* Publication of Sports Medicine-Related Randomized Controlled Trials Registered in ClinicalTrials.gov. *Am J Sports Med*. 2012; **40**: 1970–1977.

44 Klassen TP, Wiebe N, Russell K, Stevens K, Hartling L, Craig WR, *et al.* Abstracts of randomized controlled trials presented at the society for pediatric research meeting. *Arch Pediatr Adolesc Med*. 2002; **156**: 474–479.

45 Krzyzanowska MK. Large Randomized Trials Presented at an Oncology Meeting. *JAMA*. 2003; **290**: 495–501.

46 Hopewell S, McDonald S. Full publication of trials initially reported as abstracts in the Australian and New Zealand Journal of Medicine 1980-2000. *Intern Med J*. 2003; **33**: 192–194.

47 Ross JS, Mulvey GK, Hines EM, Nissen SE, Krumholz HM. Trial publication after registration in ClinicalTrials.Gov: a cross-sectional analysis. *PLoS Med*. 2009 Sep; **6**: e1000144.

48 Bourgeois F, Murthy S, Mandl K. Outcome reporting among drug trials registered in ClinicalTrials.gov. *Ann Intern Med*. 2010; **153**: 158–166.

49 Huser V, Cimino JJ. Linking ClinicalTrials.gov and PubMed to track results of interventional human clinical trials. Sampson M, editor. *PLoS One*. 2013 Jul 9; **8**: e68409.

50 Mahid SS, Qadan M, Hornung C a., Galandiuk S. Assessment of publication bias for the surgeon scientist. *Br J Surg*. 2008; **95**: 943–949.

51 Kaiser J. Study claims $28 billion a year spent on irreproducible biomedical research. *Science (80-)* [Internet]. : Online. Available from: http://news.sciencemag.org/biology/2015/06/study-claims-28-billion-year-spent-irreproducible-biomedical-research

52 Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien P-A, Reeves BC, *et al.* Challenges in evaluating surgical innovation. *Lancet*. Elsevier Ltd; 2009 Sep 26; **374**: 1097–1104.

53 Cook JA. The challenges faced in the design, conduct and analysis of surgical randomised controlled trials. *Trials*. 2009 Jan; **10**: 9.

54 Schulz K, Chalmers I, Hayes R, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995; **273**: 408–412.

55 Bhandari M, Richards R, Sprague S, Schemitsch EH. The quality of reporting of randomized trials in the Journal of Bone and Joint Surgery from 1988 through 2000. *J Bone Jt Surg*. 2002; **84**: 388–396.

56 Karanicolas P, Bhandari M, Taromi B, Akle E, Bassler D, Alonso-Coelleo P, *et al.* Blinding of outcomes in trials of orthopaedic trauma: an opportunity to enhance the validity of clinical trials. *J Bone Jt Surg*. 2008; **90**: 1026–1033.

57 Clavien P a., Puhan M a. Biased reporting in surgery. *Br J Surg*. 2014; **101**: 591–592.

58 Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, *et al.* Increasing value and reducing waste: Addressing inaccessible research. *Lancet*. 2014; **383**: 257–266.

59 Rodwin MA, Abramson JD. Clinical Trial Data as a Public Good. *JAMA*. 2012; **308**: 871.

60 Ioannidis JP a. Why most published research findings are false. *PLoS Med*. 2005; **2**: 0696–0701.

61 Goodman S, Greenland S. Assessing the Unreliability of the Medical Literature: a Response To‘ Why Most Published Research Findings Are False’. *PLoS Med* [Internet]. 2007; **4**: 135. Available from: http://www.bepress.com/jhubiostat/paper135

62 Knottnerus JA, Tugwell P. The potential impact of unpublished results. *J Clin Epidemiol*. 2013; **66**: 1061–1063.

63 Kheifets L, Olsen J. Should epidemiologists always publish their results? Yes, almost always. *Epidemiology*. 2008; **19**: 532–533.

64 Smyth RMD, Kirkham JJ, Jacoby a, Altman DG, Gamble C, Williamson PR. Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. *BMJ*. 2011; **342**: c7153.

65 ClinicalTrials.gov Web site [Internet]. Available from: http://www.clinicaltrials.gov

66 Drug and Medical Device Combination Product Decisions. Ottawa; 1998.

67 JCR Impact Factors List [Internet]. Thompson Reuters. 2013. Available from: http://thomsonreuters.com/journal-citation-reports/

68 Clifford TJ, Barrowman NJ, Moher D. Funding source, trial outcome and reporting quality: are they related? Results of a pilot study. *BMC Health Serv Res*. 2002; **2**: 18.

69 Bhandari M, Busse JW, Jackowski D, Montori VM, Schünemann H, Sprague S, *et al.* Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. *CMAJ*. 2004 Feb 17; **170**: 477–480.

70 Shah R V, Albert TJ, Bruegel-Sanchez V, Vaccaro AR, Hilibrand AS, Grauer JN. Industry support and correlation to study outcome for papers published in Spine. *Spine (Phila Pa 1976)*. 2005; **30**: 1099–1104.

71 Kasenda B, von Elm E, You J, Blümle A, Tomonaga Y, Saccilotto R, *et al.* Prevalence, characteristics, and publication of discontinued randomized trials. *JAMA*. 2014; **311**: 1045–1051.

72 Suñé P, Suñé JM, Montoro JB. Positive Outcomes Influence the Rate and Time to Publication, but Not the Impact Factor of Publications of Clinical Trial Results. *PLoS One*. 2013; **8**: 1–8.

73 Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev*. 2009; 1–26.

74 Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and Interpretation of Randomized Controlled Trials With Statistically Nonsignificant Results for Primary Outcomes. *JAMA*. 2010; **303**: 2058–2064.

75 Consensus statement on the adoption of the COPE guidelines. *Ann Surg*. 2010; **252**: 2.

76 Chapman S, Shelton B, Fitzgerald J, Harrison E. Discontinuation and non-publication of surgical randomised controlled trials : observational study. *BMJ*. 2014; 1–15.

77 Smyth RM, Jacoby A, Altman DG, Gamble C, Williamson PR. The natural history of conducting and reporting clinical trials: interviews with trialists. *Trials*. 2015; **16**: 1–12.

78 Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Med*. 2010; **7**.

Appendix 1. Protocol fields downloaded from clinicaltrials.gov

Fields defined according to the glossary of terms available on clinicaltrials.gov[65].

Field	Registered Field	Field Definition
1	**Rank**	The order in which studies appear, sorted by relevance
2	**NCT Number**	Unique clinicaltrials.gov identifier; NCT and 8 digit number
3	**Title**	Name given to study by investigators
4	**Recruitment**	**Categorical field**. Identifies stage of clinical trial, and whether or not it is open for enrollment
5	**Study Results**	**Categorical field**. Identifies studies which have results available on the clinicaltrials.gov site
6	**Conditions**	The disorder or disease which is understudy
7	**Intervention**	**Categorical field** with free text. Describes the process understudy
8	**Sponsor/Collaborators**	The organization or person overseeing the conduct of the study, and the organization which provides support for the study
9	**Gender**	**Categorical field**. Identifies sex of the persons understudy
10	**Age**	**Categorical field**. Groups age of persons understudy
11	**Phases**	**Categorical field**. Employs standard definition of study phase according to the study objective and number of patients
12	**Enrollment**	The number of participants understudy
13	**Funded By**	**Categorical field**. Identifies those studies with

		monetary support from industry or government
14	**Study Type**	**Categorical field**. Identifies the method used in the study as interventional, observational, or expanded access
15	**Study Designs**	Hierarchical listing or investigative methods: • Primary Purpose • Intervention model • Masking • Allocation
16	**Other Identifiers**	Unique identification numbers assigned to the study by other agencies (such as funders or other trial registries)
17	**First Received**	Date on which summary protocol information was first submitted to clinicaltrials.gov
18	**Start Date**	The date that enrollment of participants for a clinical study begins
19	**Completion Date**	Participants are no longer being examined or treated
20	**Last Updated**	The last date on which changes to the study information were submitted
21	**Last Verified**	The last date on which the study information was confirmed to be accurate
22	**Results First Received**	The date on which results of the study were first submitted to the clinicaltrials.gov results database
23	**Acronym**	*Not defined*
24	**Primary Completion**	The date on which the last participant in a clinical study was examined for data that was used to assess the primary outcome
25	**Outcome Measure**	The planned measurement used to determine the outcome; divided according to primary and

		secondary measures
26	**URL**	Link to study protocol on clinicaltrials.gov

Appendix 2. Summary of variables abstracted from trial protocols.

Variable	Classification	Definition
Assessment of Eligibility		
Inclusion	Include	*Meets inclusion criteria*
	Exclude: Medical Trial	*Medical intervention in a surgical patient*
	Exclude: Not Surgical Intervention	*Non-medical intervention in surgical patient (i.e. physiotherapy)*
	Exclude: Not Surgical Subject	*Study does not use a surgical intervention or use surgical patients as subjects*
	Exclude: Not RCT	
	Exclude: Dates	*Study starts < 2006; Completed > 2012*
Contact Author	Name of Principle Investigator or Trial Coordinator	
Domain: Trial Demographics		
Region	North America	
	South America	*Includes Latin America and Caribbean*
	Europe	*Includes Ukraine*
	Asia	*Includes Turkey, Israel, Middle East*
	Africa	
	Australasia	*Includes Australia and the Pacific Islands*
	Multiple	*More than one Region in study*
	Not Specified	*No Study Location Specified*
Funding Source	Government	
	Hospital	*Funding from hospital agency - Hospital funding (private or public) not assessed*
	Industry	*Funding from provided by a publically traded or private corporation*
	University	*Funding from university agency - University funding (private or public) not assessed*
	Private Interest	*Private funding agency (e.g. NGO)*
Industry Funding	Yes	*At least one funding source listed is a publically traded or private corporation*

	No	
Domain: Trial Characteristics		
Trial Intervention Type	Technique	
	Implanted Device	
	External Device	
	Dressing	*Intervention is surgical wound dressing*
	Non-Surgical	*No surgery one group*
Age	Adult	*All subjects > 18*
	Pediatric	*All subjects < 18*
	Both	
	Not Specified	*No information on age of subjects available*
Surgical Specialty	Cardiac	
	Dentistry	
	Otolaryngology	
	General Surgery	
	Neurosurgery	
	Obstetrics and Gynecology	
	Ophthalmology	
	Orthopedics	
	Plastics	
	Thoracics	
	Urology	
	Vascular	
	Multiple	
Domain: Trial Registration		
Timing	Before	*Registered before trial start date (inclusive of trial registration month)*
	During	*Registered after trial start date, before trial completion (inclusive of completion month)*

	After	*Registered after trial completion*
	Unknown	

Appendix 3. Sample Size Calculation

The following sample size estimate was used to determine the number of published trials that would be needed to estimate the validity of the literature estimate of selective outcome reporting.

p = 0.45 estimated proportion of selective outcome reporting [41,42]

α = 0.10

$Z_{1-(\alpha/2)}$ = 1.645

E = 0.05

$$n = p(1-p)\left(\frac{Z_{1-\left(\frac{\propto}{2}\right)}}{E}\right)^2$$

$$n = 0.45(1-0.45)\left(\frac{1.645}{0.05}\right)^2$$

$$n = 267.89$$

Therefore 270 trials are required to estimate the true proportion of trials demonstrating selective outcome reporting.

Appendix 4. Published Trial Abstraction Form

Date of Abstraction	yyyy – mm – dd
Trial NCT #	
Reviewer Initials	

*NOTE: Reviewer must simultaneously abstract from clinicaltrials.gov protocol and publication.

1) Abstraction of outcomes from protocol. The primary outcome of interest is the original primary outcome recorded by the author.

Protocol Outcomes		
Primary Outcome Stated	Yes	No
****** For remainder of abstraction use "Original Primary Outcome". Exception: "Current Primary Outcome" registered before trial start date*****		
"Current Primary Outcome" matches "Original Primary Outcome"	Yes	No
Number of Primary Outcomes	1	>1
State Primary Outcome:		
Number of Secondary Outcomes Listed	#	
State Secondary Outcomes:		

2) Comparison of outcomes reported in publication with outcomes registered in protocol.

Published Outcomes	
Primary Outcome from Protocol is:	☐ Primary Outcome in Publication ☐ Secondary Outcome in Publication ☐ Not Reported in Publication
Primary Outcome in Publication is:	☐ Primary Outcome in Protocol ☐ Secondary Outcome in Protocol ☐ Not Reported in Protocol (*new outcome)
State Primary outcome in Publication:	
Proportion of Secondary Outcomes from Protocol Reported in Publication (%):	(x / y) %

3) Statistical measures and significance of primary outcome measure in publication; assessment of sample size, and sample size calculation.

Published Outcome and Sample Size		
Trial Design given by Publication:	□ Superiority □ Non-Inferiority □ Equivalence □ Not reported	
State measure of effect for primary Published outcome:		
State value of the primary Published outcome:		
State p value of primary outcome (if given):		
State 95% CI of primary outcome (if given):		
Is primary outcome statistically significant:	Yes	No
Outcome favours (irrespective of significance):	□ Intervention □ Control □ Non-Inferiority/Equivalence	
Is Outcome Measure Appropriate for Primary Outcome	Yes	No
Is Sample Size Calculation Given in Publication	Yes	No
Is Sample Size Based on Primary Protocol Outcome	Yes	No
Determined Sample Size Required given by Publication	#	
Determined Sample Size Required given by Protocol	#	
Recruited Sample Size	#	

4) Assessment of Concluding Statement given by authors in the publication.

Concluding Statement	
Concluding Statement given in abstract/conclusion of the Publication:	□ Recommends Intervention □ Refutes Intervention □ Recommends Intervention with Caution

Appendix 5. List of Variables Stored in Database

FIELD	FORMAT	VALUES	MISSING	SOURCE
RANK	#	**UNIQUE**	N/A	Abst
NCT_NUM	Text	**UNIQUE**	N/A	CT
INCLUSION	Text	Include Exclude Start < 2006 Exclude Complete > 2012 Exclude non-RCT Exclude Type 1 Exclude not surgical subject Exclude Surgeon non-surg	N/A	Abst
AUTHOR	Text	*FF*	.	CT
TITLE	Text	*FF*	N/A	CT
REGION	Text	NA SA Europe Asia Africa Australasia Multi Not Specified	.	Abst
SURG_SPEC	Text	Cardiac Dentistry ENT General Surgery Neurosurgery Ob-Gyn Ophthalmology Ortho Plastics Thoracic Urology Vascular Multiple	.	Abst
CENTRES	Text	Single Multiple Unknown	.	Abst
RECRUITMENT	Text	Active, not recruiting Completed Enrolling by invitation Suspended Terminated	.	CT
TRIAL_TYPE	Text	Type 2 - technique Type 2 - implanted device	.	Abst

		Type 2 - External device Type 2 - Dressing Type 3 - Surg v non-surg		
INTERVENTION	Text	*FF*	MAND	CT
SPONSOR	Text	*FF*	MAND	CT
GENDER	Text	Both Female Male	.	CT
AGE_CT	Text	*FF*		CT
AGE	Text	Adult Pediatric Both Not Specified	.	Abst
PHASES	Text	*FF*	.	CT
ENROLLMENT	#	*Continuous*	0	CT
FUNDING_CT	Text	*FF*		CT
FUND_PRIM	Text	Government University Hospital Industry Private Interest Not Declared	.	Abst
FUND_SEC	Text	Government University Hospital Industry Private Interest Not Declared	.	
STUDY_TYPE	Text	*FF*		CT
STUDY_DESIGN	Text	*FF*	.	CT
RECEIVED	Date	*Continuous*	N/A	CT
START	Date	*Continuous*	.	CT
COMPLETE	Date	*Continuous*	.	CT
REG_TIME	Text	Before During After Unknown	.	Abst
UPDATE	Date	*Continuous*	N/A	CT
VERIFIED	Date	*Continuous*	N/A	CT
COMPLETE_PRIM	Date	*Continuous*	.	CT
NCT_CT	#	Binary 1: Yes 0: No	9	Abst
CT_RESULT	#	Binary 1: Yes 0: No	9	Abst
NCT_PUBMED	#	Binary 1: Yes	9	Abst

		0: No		
NCT_GOOG	#	Binary 1: Yes 0: No	9	Abst
AUTH_PUBMED	#	Binary 1: Yes 0: No	9	Abst
PUBLISHED	#	Binary 1: Yes 0: No		Abst
E_PUB	Date	Continuous		Abst
J_PUB	Date	Continuous		Abst
JOURNAL	Text	*FF*		Abst
J_AUTHOR	Text	*FF*		Abst
PUB_NUM	#	Continuous	.	Abst
J_IMPACT	#	Continuous	.	Abst
MIN_PUB_DATE	Date	Continuous	.	Abst
MIN_COMP_DATE	Date	Continuous	.	Calc (CT)
DESIGN_PROT	Text	RCT Unknown	N/A	Abst
PRIM_STATED	Text	Yes No	N/A	Abst
PRIM_CONGRUENT	Text	Yes No No Primary Listed	N/A	Abst
SAMPLE_SIZE	#	*Continuous*	.	Abst
NUM_PRIM_PROT	#	*Continuous*	N/A	Abst
P_OUT_1	Text	*FF*	.	Abst
P_OUT_2	Text	*FF*	.	Abst
P_OUT_3	Text	*FF*	.	Abst
P_OUT_4	Text	*FF*	.	Abst
P_OUT_5	Text	*FF*	.	Abst
NUM_SEC_PROT	#	*Continuous*	N/A	Abst
S_OUT_1	Text	*FF*	.	Abst
S_OUT_2	Text	*FF*	.	Abst
S_OUT_3	Text	*FF*	.	Abst
S_OUT_4	Text	*FF*	.	Abst
S_OUT_5	Text	*FF*	.	Abst
S_OUT_6	Text	*FF*	.	Abst
S_OUT_7	Text	*FF*	.	Abst
S_OUT_8	Text	*FF*	.	Abst
S_OUT_9	Text	*FF*	.	Abst
S_OUT_10	Text	*FF*	.	Abst
DESIGN_PUB	Text	RCT Prospective Cohort	.	Abst
PRIM_PROT	Text	Primary Outcome in Publication Secondary Outcome in Publication Not Reported in	.	Abst

		Publication No Primary in Protocol		
PRIM_PUB	Text	Primary Outcome in Protocol Secondary Outcome in Protocol Not Reported in Protocol	.	Abst
PRIM_OUT_PUB	Text	*FF*	.	Abst
NUM_SECOND	#	*Continuous*	.	Abst
STAT_DESIGN	Text	Superiority Non-Inferiority Equivalence Not Reported	N/A	Abst
EFFECT_MEASURE	Text	*FF*		Abst
P_OUT_VAL	Text	*FF*	.	Abst
PVALUE	Text	*FF*	.	Abst
95CI	Text	*FF*	.	Abst
STAT_SIG	Text	Yes No Not Comparative		Abst
OUT_FAVOUR	Text	Intervention Control Non-Inferiority/Equivalence Not Comparative		Abst
APPROPRIATE	Text	Yes No		Abst
SS_CALC	Text	Yes No		Abst
SS_CALC_PRIM	Text	Yes No		Abst
SS_DET	#	*Continuous*	.	Abst
SS_REC	#	*Continuous*	.	Abst
STUDY_CONC	Text	Recommends Intervention Refutes Intervention Recommends Intervention with Caution		Abst