

Electronic Thesis and Dissertation Repository

---

7-18-2017 12:00 AM

## Reducing Distortion – Identifying Areas to Improve the Quality of Randomized Clinical Trials Published in Anesthesiology Journals

Jeffrey T.Y. Chow, *The University of Western Ontario*

Supervisor: Dr. Philip M. Jones, *The University of Western Ontario*

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

© Jeffrey T.Y. Chow 2017

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Anesthesiology Commons](#), and the [Clinical Epidemiology Commons](#)

---

### Recommended Citation

Chow, Jeffrey T.Y., "Reducing Distortion – Identifying Areas to Improve the Quality of Randomized Clinical Trials Published in Anesthesiology Journals" (2017). *Electronic Thesis and Dissertation Repository*. 4681. <https://ir.lib.uwo.ca/etd/4681>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

## Abstract

Randomized clinical trials (RCTs) provide important evidence to inform clinical decision making; if these trials are of low quality, the resulting clinical decision will likely also be of low quality. The main purpose of this thesis was to conduct a series of methodological surveys that would identify potential areas of improvement in the quality of reporting for RCTs published in anesthesiology journals. Trial registration adequacy, adherence to CONSORT for Abstracts guidelines, and sample size calculation quality were all assessed, with a final chapter exploring the effect of industry funding on these methodological quality measures. While the results suggest improvement over time, the overall quality is still lacking. Industry sources funded a minority of the included RCTs, and did not appear to affect any of the measures of quality. More research is needed to confirm these findings and to identify tools for reducing the potential distortion emanating from low quality design and reporting.

## Keywords

Randomized clinical trial, Trial registration, CONSORT for abstracts, Sample size, Statistical power, Industry funding, Study design, Reporting quality.

## Co-Authorship Statement

**Chapter 3:** *Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals*

Co-authorship Jones PM, Chow JTY, Arango MF, Fridfinnson JA, Gai N, Lam K, Turkstra TP: JC and PJ designed the study, analyzed the data, and wrote the manuscript. JC, MA, JF, NG, KL, TT, and PJ conducted the study, collected data, and revised the manuscript. PJ conceived the study idea and supervised the study.

**Chapter 4:** *The degree of adherence to CONSORT reporting guidelines for the abstracts of randomized clinical trials published in anesthesiology journals*

Co-authorship Chow JTY, Turkstra TP, Yim E, Jones PM: JC and PJ designed the study, analyzed the data, and wrote the manuscript. JC, TP, EY, and PJ conducted the study, collected data, and revised the manuscript. PJ supervised the study.

**Chapter 5:** *Sample size calculations for randomized clinical trials published in anesthesiology journals: A comparison of 2010 versus 2016*

Co-authorship Chow JTY, Turkstra TP, Yim E, Jones PM: JC and PJ designed the study, analyzed the data, and wrote the manuscript. JC, TP, EY, and PJ conducted the study, collected data, and revised the manuscript. PJ supervised the study.

**Chapter 6:** *Impact of funding source on randomized clinical trials published in anesthesiology journals*

Co-authorship Chow JTY, Turkstra TP, Yim E, Jones PM: JC and PJ designed the study, analyzed the data, and wrote the manuscript. JC, TP, EY, and PJ conducted the study, collected data, and revised the manuscript. PJ supervised the study.

## Acknowledgments

First, I must express my sincere gratitude to my supervisor, Dr. Philip Jones. He has guided me both before and throughout my time in the MSc program. Without him, this thesis would never have been completed.

I would also like to thank my supervisory committee, Drs. Neil Klar and Janet Martin, for their valuable advice and feedback, and my family, Raymond, Mabel, and Jerry Chow for their continued support.

# Table of Contents

Abstract.....	i
Co-Authorship Statement.....	ii
Acknowledgments.....	iv
Table of Contents.....	v
List of Tables.....	viii
List of Figures.....	ix
List of Appendices.....	x
List of Abbreviations.....	xi
Chapter 1.....	1
1 Introduction.....	1
1.1 Background.....	2
1.2 Thesis Rationale.....	3
1.3 Thesis Objectives.....	7
1.4 Thesis Structure.....	8
1.5 Literature Cited.....	9
Chapter 2.....	13
2 Literature Review.....	13
2.1 Anesthesiology.....	14
2.2 Randomized Clinical Trials.....	15
2.3 Trial Registration.....	16
2.3.1 Reasons for Trial Registration.....	16
2.3.2 Mandatory Trial Registration.....	19
2.3.3 Quality of Trial Registration.....	21

2.4	Abstract Reporting and the CONSORT Statement.....	23
2.4.1	Abstracts .....	23
2.4.2	CONSORT Statement.....	24
2.4.3	CONSORT for Abstracts .....	25
2.5	Sample Size.....	27
2.5.1	Sample Size Calculation Assumptions .....	27
2.5.2	Quality of Sample Size Calculation .....	29
2.6	Impact of Funding Source on RCTs .....	30
2.7	Literature Cited .....	32
Chapter 3.....		45
3	Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals .....	45
3.1	Introduction.....	46
3.2	Methods.....	47
3.3	Results.....	52
3.4	Discussion.....	60
3.5	Literature Cited.....	66
Chapter 4.....		69
4	The degree of adherence to CONSORT reporting guidelines for the abstracts of randomized clinical trials published in anesthesiology journals.....	69
4.1	Introduction.....	70
4.2	Methods.....	72
4.3	Results.....	76
4.4	Discussion.....	80
4.5	Literature Cited .....	88
Chapter 5.....		92

5	Sample size calculations for randomized clinical trials published in anesthesiology journals: A comparison of 2010 versus 2016.....	92
5.1	Introduction.....	93
5.2	Methods.....	95
5.3	Results.....	99
5.4	Discussion.....	105
5.5	Literature Cited.....	110
	Chapter 6.....	114
6	Impact of funding source on randomized clinical trials published in anesthesiology journals.....	114
6.1	Introduction.....	115
6.2	Methods.....	116
6.3	Results.....	121
6.4	Discussion.....	127
6.5	Literature Cited.....	130
	Chapter 7.....	133
7	Integrated Discussion and General Conclusions.....	133
7.1	Overview.....	134
7.2	Integrated Discussion of Results.....	134
7.3	General Conclusions.....	136
7.4	Literature Cited.....	138
	Appendices.....	140
	Curriculum Vitae.....	150



## List of Tables

Table 3-1: Characteristics of trials .....	53
Table 3-2: Description of primary and secondary outcomes in adequately registered trials .....	56
Table 3-3: Differences between outcomes when comparing the published trial to the trial registry among adequately registered trials .....	58
Table 4-1: Characteristics of RCTs included for analysis .....	77
Table 4-2: Adherence to CONSORT-Abstract reporting items in RCTs .....	79
Table 4-3: Characteristics of secondary outcomes reported in the abstract.....	80
Table 5-1: Characteristics of included RCTs.....	101
Table 5-2: Reporting and accuracy of sample size assumptions in RCTs.....	102
Table 5-3: Replication of sample size calculations for included RCTs.....	104
Table 5-4: Comparison of estimated and reported effect sizes for binary outcomes.....	105
Table 6-1: Funding source of RCTs by year.....	123
Table 6-2: Results of simple regression models for the impact of industry funding in included RCTs from 2015/2016 only .....	125

## List of Figures

Figure 3-1: Flowchart for inclusion of trials defined as adequately registered .....	52
Figure 3-2: Number of randomized clinical trials published by journal over time .....	55
Figure 3-3: Percentage of adequately registered trials by journal over time .....	55
Figure 4-1: Distribution of overall CONSORT for Abstracts abstract-only score by journal over time .....	78
Figure 4-2: Distributions of overall CONSORT for Abstracts abstract-only score in 2010 and 2016.....	78
Figure 5-1: Flowchart for inclusion of trials.....	100
Figure 5-2: Distribution of primary post-hoc power stratified by statistical significance of the trial .....	104
Figure 6-1: Funding source of RCTs by journal for 2010 and 2016 .....	124

## List of Appendices

Appendix A: CONSORT for Abstracts checklist .....	140
Appendix B: Sample email to corresponding authors .....	141
Appendix C: Differences between outcomes when comparing the published trial to the trial registry among adequately registered trials (stratified by journal).....	142
Appendix D: Adapted version of CONSORT for Abstracts statement explanations .....	144
Appendix E: Adherence to CONSORT-Abstract reporting items in RCTs (breakdown by journal in 2016).....	148
Appendix F: Distribution of country origin for included RCTs .....	149

## List of Abbreviations

*A&A – Anesthesia & Analgesia*

*Acta – Acta Anaesthesiologica Scandinavica*

*Annals IM – Annals of Internal Medicine*

*BJA – British Journal of Anaesthesia*

CI – confidence interval

CIHR – Canadian Institutes of Health Research

*CJA – Canadian Journal of Anesthesia*

COMPare – Centre for Evidence-Based Medicine Outcome Monitoring Project

CONSORT – Consolidated Standards of Reporting Trials

CONSORT for Abstracts – Consolidated Standards of Reporting Trials for Abstracts

EBM – evidence-based medicine

*EJA – European Journal of Anaesthesiology*

EUCTR – European Union Clinical Trials Registry

FDAAA – Food and Drug Administration Amendments Act

ICMJE – International Committee of Medical Journal Editors

ICTRP – International Clinical Trials Registry Platform

IQR – interquartile range

IRR – incidence rate ratio

ISRCTN – International Standard Randomized Controlled Trial Number

*JAMA – Journal of the American Medical Association*

*NEJM – New England Journal of Medicine*

OR – odds ratio

*RAPM – Regional Anesthesia and Pain Medicine*

RCT – randomized clinical trial

RR – risk ratio

TCPS2 – Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans

WAME – World Association of Medical Editors

WHO – World Health Organization

WMA – World Medical Association

# Chapter 1

## 1 Introduction

---

---

## 1.1 Background

First coined by Gordon Guyatt in 1991,<sup>1</sup> evidence-based medicine (EBM) involves identifying the best available evidence, deciding the accuracy and reliability of that evidence, considering the trade-offs to alternative management strategies, and integrating patients' values and preferences so that the optimal clinical decision can be made.<sup>2</sup> The first of these principles is also one of the most important since, without relevant evidence, the process of EBM cannot be started. That is not to say that individual clinical expertise is unimportant, as the proponents of EBM have always emphasized the concept of integrating the best available evidence with clinical expertise.<sup>3</sup> In regards to therapy decisions, large randomized trials are at the top of the evidence hierarchy,<sup>2</sup> representing the best type of evidence when available. Randomized trials (and systematic reviews of multiple randomized trials) are the gold standard for determining whether a treatment is more beneficial than harmful.<sup>3</sup>

A randomized clinical trial (RCT) is a prospective cohort study where participants are allocated by chance to treatment and control groups.<sup>4</sup> Ideally, randomization eliminates selection bias since participants are randomly assigned to each study arm. Randomization also reduces confounding since both known and unknown potentially confounding patient factors are randomly distributed between study arms.<sup>4</sup> Even though RCTs are the best type of evidence, confidence in their results is decreased when there are design, execution, or reporting problems leading to a high risk of bias.<sup>2</sup> If low quality RCTs are used to make clinical decisions, those clinical decisions are likely to also be of low quality.

This thesis will focus on four aspects relating to the quality of reporting in RCTs: trial registration, abstract reporting guidelines, sample size calculations, and funding source. While some investigators have analyzed these factors in the top general medical journals, this thesis will focus specifically on the anesthesiology literature to provide a specialty-specific point of view; specifically, only RCTs published from the top six general (non-pain-centric) anesthesiology journals are included in the following chapters: *Anaesthesia*, *Anesthesia & Analgesia (A&A)*, *Anesthesiology*, *British Journal of Anaesthesia (BJA)*, *Canadian Journal of Anesthesia (CJA)*, and *European Journal of Anaesthesiology (EJA)*.<sup>5</sup>

## 1.2 Thesis Rationale

To ensure that the RCTs used for clinical decision making are useful, authors must ensure that their studies are appropriately designed, executed, and reported.<sup>2</sup> The key message in EBM is *best available evidence*: however, if the RCTs being used to make clinical decisions are distorted or biased, how can they provide reliable evidence? Distortion occurs when the design quality or reporting quality of a trial causes its true findings to become obscured, either through intentional or unintentional misrepresentation. This thesis reviews published RCTs in anesthesiology journals to evaluate and improve *design quality* by assessing trial registration, adequacy of sample size calculations, and the role (if any) of the trials' funding source. This thesis also examines *reporting quality* by assessing adherence to abstract reporting guidelines and discrepancies between reported and registered outcomes.



Adequate trial registration involves prospectively recording, in a publicly available trial registry, the precise intervention to be tested and outcomes to be measured, *prior* to the first patient being enrolled. The process of adequate trial registration aims to reduce publication and outcome reporting bias, thereby improving transparency.<sup>6</sup> Publication bias involves publication of research findings based on the nature and direction of the results while outcome reporting bias involves selectively reporting outcomes based on the nature and direction of the results.<sup>7</sup> If trials are registered before they are conducted, there would be a record regardless of whether or not they are published.<sup>6</sup> The issue of selective outcome reporting is also easier to recognize since readers and editors are able to compare what the investigators had stated in the registry with what they report in their manuscript. The 2005 statement from the International Committee of Medical Journal Editors (ICMJE) states that all ICMJE member journals will require registration in a public trials registry for a RCT to be considered for publication,<sup>8</sup> demonstrating the importance of adequate trial registration.

Recently, the Centre for Evidence-Based Medicine Outcome Monitoring Project (COMPare) team checked RCTs published in the top five medical journals for outcome switching and sent letters to the journal editors with their findings.<sup>9</sup> Of the 67 trials checked, they found that on average, each trial only reported 58.2% of their specified outcomes and added 5.3 new outcomes that were never registered.<sup>9</sup> Of the 201 RCTs published in top anesthesiology journals in 2013, only 36% in 2013 were adequately registered and 48% of those adequately registered RCTs had a major discrepancy between the trial registry and the published manuscript, suggesting an unacceptably high prevalence of outcome reporting bias.<sup>10</sup> Chapter 3 of this thesis will further explore this

issue by examining the trial registration rates and outcome discrepancy rates over time and between anesthesiology journals. By quantifying how often investigators are switching outcomes, authors may appreciate why they need to avoid this practice and journal editors will be able to identify these problems — potentially reducing publication bias and selective outcome reporting bias.

Another important aspect of trial reporting is the abstract of the published paper. Since abstracts are concise summaries of the entire RCT, clinicians and other readers often use the abstract to decide whether or not to read the full manuscript.<sup>11</sup> Due mainly to time constraints, many readers will never read any part of the paper except the abstract.<sup>12</sup> In other cases, the full-text may be irretrievable due to limited resources or language barriers.<sup>13</sup> Either way, it is clear that abstracts are extremely important in biomedical publishing, and it is likely that at least some of current clinical practice is being based solely on the evidence presented in abstracts.<sup>13</sup> In 2008, the Consolidated Standards of Reporting Trials for Abstracts (CONSORT for Abstracts) statement provided a list of essential items authors should include in abstracts when reporting trial results.<sup>14,15</sup> Comparing 2006 (pre-CONSORT for Abstracts) and 2009 (post-CONSORT for Abstracts) abstracts published in four anesthesiology journals, there were no clear improvements observed, with well-reported items continuing to be well-reported and poorly reported items remaining poorly reported.<sup>16</sup> Chapter 4 of this thesis updates the anesthesiology literature regarding the quality of RCT abstract reporting and determines whether abstracts are distorting the trials' findings by leaving items from the full-text manuscript out of the abstract.

Every RCT needs participants. However, determining the ideal number of participants that balances limited resources and the ability to detect an effect is difficult. The main objective of sample size estimation is to determine the minimum number of participants needed to detect a clinically relevant effect for the intervention.<sup>17</sup> The CONSORT statement recognizes the importance of sample size calculations, and recommends that authors should include how sample size was determined when publishing their findings.<sup>18</sup> Researchers found that even though sample size calculations were frequently reported (91.7%) in anesthesiology RCTs published in 2013, the required assumptions to replicate the sample size calculations were not consistently reported, with differences found between the estimated effect used for the sample size calculation and the actual values observed in the RCT.<sup>19</sup> Since these effect differences for each RCT were not quantified, chapter 5 of this thesis will help to show how many trials actually had different values than what was anticipated in the sample size calculations. This chapter will also look at the progress of anesthesiology RCTs from 2010 to 2016 in terms of sample size calculation quality, updating the literature using the most recently published trials.

In chapter 6, the role that the funding source has on RCTs will be explored using the outcomes measured in previous chapters. With the increased support from industry for biomedical research over the years, potential conflicts of interest stemming from financial interests have often been problematic for maintaining scientific integrity.<sup>20</sup> When the pharmaceutical industry and other related biotechnology firms fund clinical trials, achieving favourable results in those trials are critical to the company and there are financial pressures to achieve a certain outcome.<sup>21</sup> While many studies have shown that industry funded studies are more likely to have outcomes favouring the sponsor,<sup>22</sup> none

have focused on the effect in anesthesiology literature. This chapter fills that gap as well as examines the effect of industry funding using less commonly studied quality indicators such as trial registration adequacy, abstract reporting guidelines adherence, and sample size calculation quality to provide a different perspective on the discussion.

### 1.3 Thesis Objectives

As suggested by the title, the overall aim of this thesis is to identify areas where authors can reduce distortion and improve the quality of RCTs published in anesthesiology journals. To accomplish this objective, each of chapters 3 to 6 addresses one potential concept that authors may not have fully considered when designing, executing, and reporting their RCTs.

The first objective (Chapter 3) is to determine whether the reported outcomes of RCTs published in anesthesiology journals are the same as the outcomes originally registered by the investigators in publicly available clinical trial registries. Potential trends for this objective will be observed both across time and between journals.

The second objective (Chapter 4) involves determining the degree of adherence to CONSORT reporting guidelines for the abstracts of RCTs published in anesthesiology journals. The full-text manuscript will also be searched for items left out of the abstract to determine whether these abstracts are distorting the overall meaning of the trial (i.e. required items *are* reported in the full manuscript but *not* in the abstract).

The third objective (Chapter 5) is to determine whether the quality of conducting sample size calculations has improved from 2010 to 2016 for RCTs published in anesthesiology journals. The quality will be ascertained in two ways, by assessing whether authors report

the necessary elements to allow for replication of sample size calculations and whether the assumptions made by authors are similar to the values actually observed after conducting the trial. Using values reported in the manuscript, the post-hoc power for primary and secondary outcomes will also be calculated to determine whether there was enough statistical power to detect differences between study groups.

The fourth and final objective (Chapter 6) seeks to understand whether the funding source, namely industry funding, introduces biases to RCTs published in anesthesiology journals. This objective will show what types of funding sources support anesthesiology RCTs and whether industry funding affects the objectives outlined in earlier chapters such as registration rates, registered compared to reported outcomes, adherence to the CONSORT for Abstracts statement, quality of sample size calculations, etc.

## 1.4 Thesis Structure

In compliance with the standards set by The University of Western Ontario's School of Graduate and Postdoctoral Studies, this thesis is presented in the integrated-article format.<sup>23</sup> Chapter 2 provides a detailed and critical review of the literature pertaining to the quality of reporting for RCTs published in anesthesiology journals. This chapter comprehensively surveys and appraises the literature in regards to subject material presented in subsequent chapters such as anesthesiology, RCTs, the CONSORT statement, trial registration, sample size, and funding source impact. Each chapter from 3 to 6 is a manuscript that achieves one of the objectives described in the previous section.

Chapter 3, *Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals*, addresses the first objective, chapter 4, *The degree*

*of adherence to CONSORT reporting guidelines for the abstracts of randomized clinical trials published in anesthesiology journals*, addresses the second objective, chapter 5, *Sample size calculations for randomized clinical trials published in anesthesiology journals: A comparison of 2010 versus 2016*, addresses the third objective, and chapter 6, *Impact of funding source on randomized clinical trials published in anesthesiology journals*, addresses the fourth objective. Each individual chapter from 3 to 6 addresses discrete but related topics, with the intention of submitting each chapter as a separate manuscript for publication.

Chapter 3 is an extension of my undergraduate thesis<sup>24</sup> completed in partial fulfillment of the requirements for the degree of Bachelor of Medical Sciences, Honors Specialization in Epidemiology and Biostatistics. This undergraduate thesis (work performed before beginning the MSc degree) comprised the searching, screening, and extracting of data relating to study characteristics and registration status of eligible studies<sup>24</sup> while the rest of Chapter 3 is new work that determines the degree of primary and secondary outcome discrepancies in the included studies. Chapter 7, *Integrated Discussion and General Conclusions*, summarizes the main results of this thesis and relates each chapter's findings to each other.

## 1.5 Literature Cited

1. Guyatt GH. Evidence-based medicine. *ACP J Club*. 1991;114(2):A16. doi:10.7326/ACPJC-1991-114-2-A16.
2. Guyatt G, Rennie D, Meade MO, Cook DJ. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 3rd ed. United States of America: McGraw-Hill Education; 2015.

3. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71. doi:10.1136/bmj.312.7023.71.
4. Sessler DI, Imrey PB. Clinical Research Methodology 3: Randomized Controlled Trials. *Anesth Analg*. 2015;121(4):1052-1064. doi:10.1213/ANE.0000000000000862.
5. Thomson Reuters. Journal Citation Reports - Anesthesiology. ISI Web Knowl. 2016. <http://admin-apps.webofknowledge.com/JCR/JCR>. Accessed August 15, 2016.
6. Fredrickson MJ, Ilfeld BM. Prospective Trial Registration for Clinical Research. *Reg Anesth Pain Med*. 2011;36(6):619-624. doi:10.1097/AAP.0b013e318230fbc4.
7. Sterne J, Egger M, Moher D. Chapter 10: Addressing reporting biases. In: Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5. The Cochrane Collaboration; 2011.
8. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors. *N Engl J Med*. 2004;351(12):1250-1251. doi:10.1056/NEJMe048225.
9. Goldacre B, Drysdale H, Powell-Smith A, et al. The COMPare Trials Project. 2016. [www.COMPare-trials.org](http://www.COMPare-trials.org). Accessed December 1, 2016.
10. De Oliveira GS, Jung MJ, McCarthy RJ. Discrepancies Between Randomized Controlled Trial Registry Entries and Content of Corresponding Manuscripts Reported in Anesthesiology Journals. *Anesth Analg*. 2015;121(4):1030-1033. doi:10.1213/ANE.0000000000000824.
11. Sivendran S, Newport K, Horst M, Albert A, Galsky MD. Reporting quality of abstracts in phase III clinical trials of systemic therapy in metastatic solid malignancies. *Trials*. 2015;16(1):341. doi:10.1186/s13063-015-0885-9.
12. Saint S, Christakis DA, Saha S, et al. Journal reading habits of internists. *J Gen Intern Med*. 2000;15(12):881-884. doi:10.1046/j.1525-1497.2000.00202.x.

13. Wang L, Li Y, Li J, et al. Quality of reporting of trial abstracts needs to be improved: using the CONSORT for abstracts to assess the four leading Chinese medical journals of traditional Chinese medicine. *Trials*. 2010;11(1):75. doi:10.1186/1745-6215-11-75.
14. Hopewell S, Clarke M, Moher D, et al. CONSORT for Reporting Randomized Controlled Trials in Journal and Conference Abstracts: Explanation and Elaboration. von Elm E, ed. *PLoS Med*. 2008;5(1):e20. doi:10.1371/journal.pmed.0050020.
15. Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet*. 2008;371(9609):281-283. doi:10.1016/S0140-6736(07)61835-2.
16. Can OS, Yilmaz AA, Hasdogan M, et al. Has the quality of abstracts for randomised controlled trials improved since the release of Consolidated Standards of Reporting Trial guideline for abstract reporting? A survey of four high-profile anaesthesia journals. *Eur J Anaesthesiol*. 2011;28(7):485-492. doi:10.1097/EJA.0b013e32833fb96f.
17. Noordzij M, Tripepi G, Dekker FW, Zoccali C, Tanck MW, Jager KJ. Sample size calculations: Basic principles and common pitfalls. *Nephrol Dial Transplant*. 2010;25(5):1388-1393. doi:10.1093/ndt/gfp732.
18. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340(1):c869-c869. doi:10.1136/bmj.c869.
19. Abdulatif M, Mukhtar A, Obayah G. Pitfalls in reporting sample size calculation in randomized controlled trials published in leading anaesthesia journals: a systematic review. *Br J Anaesth*. 2015;115(5):699-707. doi:10.1093/bja/aev166.
20. Bekelman JE, Li Y, Gross CP. Scope and Impact of Financial Conflicts of Interest in Biomedical Research. *JAMA*. 2003;289(4):454-465. doi:10.1001/jama.289.4.454.
21. Lexchin J. Those Who Have the Gold Make the Evidence: How the Pharmaceutical Industry Biases the Outcomes of Clinical Trials of Medications. *Sci Eng Ethics*. 2012;18(2):247-261. doi:10.1007/s11948-011-9265-3.



22. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003;326(7400):1167-1170. doi:10.1136/bmj.326.7400.1167.
23. The University of Western Ontario: School of Graduate and Postdoctoral Studies. 8. Thesis. Grad Regul. 2016. [http://grad.uwo.ca/current\\_students/regulations/8.html](http://grad.uwo.ca/current_students/regulations/8.html). Accessed July 1, 2016.
24. Chow J, Jones P. Adequate Registration of Randomized Controlled Trials Published in Anesthesiology Journals. 2016.

## Chapter 2

### 2 Literature Review

---

---

## 2.1 Anesthesiology

One of the many specialties in medicine, anesthesiology is concerned with pain management, administering anesthesia, airway management, advanced life support, and perioperative care.<sup>1</sup> Because of the diverse nature of anesthesiology, anesthesiologists practice in a wide variety of settings such as in the operating room, where they are responsible for medical management and anesthetic care of patients; in the postanesthesia care unit, where they monitor and assess patients regaining consciousness; in pain management clinics, where they relieve pain for patients; and in critical care units, where they provide medical assessment, airway management, and respiratory support.<sup>2</sup> The field of anesthesiology is also referred to as *anaesthesiology*, *anaesthesia*, and *anesthesia*, with practitioners termed *an(a)esthesiologists* or *an(a)esthetists*.

In Canada, only physicians can have the responsibility for administering anesthesia, but may delegate certain tasks to anesthesia assistants or technicians.<sup>3</sup> The same situation occurs in the United States, with the exception of certified registered nurse anesthetists also able to administer anesthesia with or without physician supervision.<sup>3,4</sup> In Canada, the number of physicians in the anesthesiology specialty has been rising with 3,274 anesthesiologists in 2016, corresponding to 9.1 anesthesiologists per 100,000 people in the population.<sup>5</sup> Excluding on-call, Canadian anesthesiologists spent, on average, only 0.8 hours per week (1.6% of total hours worked per week excluding on-call) on research in 2014, with most of their time being spent on direct patient care and teaching.<sup>5</sup> It is therefore apparent that most anesthesiologists are not researchers *per se*.

## 2.2 Randomized Clinical Trials

A randomized clinical trial (RCT) is a prospective study that examines the effects of an intervention in human beings by randomly assigning participants to intervention and control groups, and following up on their health status at a later time.<sup>6</sup> The term randomized *clinical* trial is often used interchangeably with the term, randomized *controlled* trial, and both terms share the same acronym of RCT. In evidence-based medicine (EBM), RCTs (and systematic reviews or meta-analyses of RCTs) are considered to be at the top of the evidence hierarchy for therapeutic interventions since they present the lowest risk of bias.<sup>7</sup> The randomization aspect of RCTs reduces selection bias by having participants randomly allocated to study arms and reduces confounding by equally distributing subject characteristics between study arms.<sup>8</sup>

The criticisms of EBM can be categorized into five broad themes: overreliance on empiricism independent of physiological theory, narrow definition of high quality evidence disregarding other types of research, lack of evidence regarding the efficacy of EBM to improve the quality of healthcare, limited applicability when treating an individual patient, and potential reduction in a physician's autonomy to cut costs.<sup>9</sup> The evidence hierarchy in EBM should not imply that RCTs are the best type of study, since maximizing internal validity typically comes at the cost of reducing external validity, generalizability, and applicability.<sup>10</sup>

Looking at the number of publications from Canadian university anesthesiology departments in 2000 to 2004, RCTs were the most common type of study published at 18%, followed by case reports, reviews, and cohort studies.<sup>11</sup> However, the total number

of RCTs appeared to be decreasing from 2000 to 2004 even though the total number of anesthesia publications was constant with minor yearly fluctuations.<sup>11</sup> A follow-up study in the years 2005 to 2013 found that the declining trend did not continue beyond 2004, resulting in a slight overall increase in RCTs published.<sup>12</sup> A similar trend was observed in the United States with the overall number and percentage of clinical research studies increasing from 2001 to 2010.<sup>13</sup> In the top seven general anesthesiology journals from 1997 to 2006, the worldwide number of original research publications remained constant while the number of United Kingdom original research publications decreased by about 50%.<sup>14</sup> Unfortunately, the data was not stratified by type of publication, classifying original research publications as including all experimental research, RCTs, observational studies, and large case series with statistics applied.<sup>14</sup>

## 2.3 Trial Registration

### 2.3.1 Reasons for Trial Registration

The purpose of trial registration is to ensure that there is a public record of all trials conducted and that investigators commit to a certain trial protocol before conducting the study. Adequate trial registration involves prospectively (before the first patient is enrolled) registering an RCT in a publicly available trials registry,<sup>15</sup> ideally reducing publication bias and outcome reporting bias. While both biases fall under the general category of reporting bias, publication bias occurs at the study level while outcome reporting bias occurs at the outcome level.<sup>16</sup>

Publication bias involves deciding whether to publish research findings based on the nature and direction of the results.<sup>17</sup> This type of reporting bias has been extensively

studied, with a systematic review finding that positive trials were more likely to be published than negative or null trials (odds ratio (OR): 3.90; 95% confidence interval (CI): 2.68–5.65).<sup>18</sup> Potential reasons for this finding include authors failing to write manuscripts for trials with negative results, trials with negative results being peer reviewed less favourably, or journal editors being reluctant to publish negative results.<sup>17</sup> However, among submitted manuscripts to *Journal of the American Medical Association (JAMA)*, no statistically significant difference was found between the acceptance rates of submissions with positive results and negative results,<sup>19</sup> suggesting that publication bias may be due more to lack of manuscript submission than lack of manuscript acceptance.<sup>20</sup> Further studies have confirmed the investigator-based factors for nonpublication, such as lack of time to prepare manuscript, perceived low likelihood of being published, and difficulties with co-investigators.<sup>21–23</sup>

Trial registration aims to reduce publication bias by identifying all trials conducted, thereby encouraging investigators to publish their results after the trial has been completed.<sup>24</sup> By ensuring all RCTs are registered prior to patient enrollment, a public record of all RCTs being conducted will remain even if trials fail to be published. Though trial registration has certainly not eliminated publication bias, more trials are being registered and public trials registries allow for the extent of publication bias to be fairly easily examined and monitored.<sup>25</sup> In fact, a recent cross-sectional study of published randomized trials found the first inverse association between trial registration and significant positive outcomes in cardiovascular trials, suggesting that trial registration may already be starting to reduce publication bias.<sup>26</sup>

On the other hand, outcome reporting bias involves selectively reporting outcomes based on the nature and direction of individual outcomes' results.<sup>17</sup> Selective reporting can be divided into several subcategories: selective omission of outcomes, selective choice of data for an outcome such as reporting only one time point after many have been measured, selective reporting of analyses using the same data such as selecting from analyses using multiple cut-points, selective reporting of subsets of the data such as using only subscales of a full measurement scale that was measured, and selective under-reporting of the data that prevents future meta-analyses.<sup>27</sup> Selective outcome reporting affects the effect size and significance level estimates, resulting in less robust conclusions.<sup>28</sup> Many studies have examined outcome reporting bias, with a systematic review finding that statistically significant outcomes were more likely to be completely reported than non-statistically significant outcomes with ORs ranging from 2.2 to 4.7.<sup>29</sup> A survey of trial authors found that reasons for not reporting outcomes included lack of understanding about the importance of reporting negative results, the data perceived to be uninteresting, having too few events to be worth reporting, and perceived need for brevity from the journal.<sup>30</sup> Since most unreported outcomes were due to a lack of a significant difference between study arms, published results would likely be overoptimistic and biased.<sup>30</sup>

Trial registration requires investigators to publicly commit to outcomes prior to the study being conducted, reducing the chance of outcome reporting bias remaining undetected.

Trial registration also makes the conduct and reporting of RCTs a more transparent process, with other researchers, reviewers, or editors able to compare the registered trial protocol with the final publication. One example of this occurring is the Centre for

Evidence-Based Medicine Outcome Monitoring Project (COMPare), where researchers monitor the RCTs published in the top five medical journals (*Annals of Internal Medicine* (*Annals IM*), the *BMJ*, *JAMA*, the *New England Journal of Medicine* (*NEJM*), and the *Lancet*) for outcome switching by comparing the trial protocol or trials registry entry with the trial report, writing letters to those journals and correcting the record for those trials.<sup>31</sup>

Even though trial registration improves study conduct transparency, many of its benefits can only be seen if a large majority of RCTs are adequately registered, suggesting the need for mandatory trial registration.<sup>32</sup> If unregistered or retrospectively registered trials can be published, some investigators may decide to only register a trial after a decision to publish is made, resulting in publication bias as there would be no record of unpublished studies.

### 2.3.2 Mandatory Trial Registration

In 2005, the International Committee of Medical Journal Editors (ICMJE) mandated that all member journals would require clinical trials to be registered in a public trials registry prior to patient enrollment in order to be considered for publication.<sup>33</sup> While the ICMJE does not advocate for a specific trials registry, they require that the registry must be publicly available at no charge, open to all prospective registrants, managed by a not-for-profit organization, possessing a mechanism to ensure registration data validity, and electronically searchable;<sup>33</sup> the following registries are currently deemed acceptable: [www.anzctr.org.au](http://www.anzctr.org.au), [www.clinicaltrials.gov](http://www.clinicaltrials.gov), [www.ISRCTN.org](http://www.ISRCTN.org), [www.umin.ac.jp/ctr/index/htm](http://www.umin.ac.jp/ctr/index/htm), [www.trialregister.nl](http://www.trialregister.nl), <https://eudract.ema.europa.eu/> (new registrations after 2011), and any registries that participate in the World Health Organization (WHO) International Clinical Trials Portal after June 2007



(<http://www.who.int/ictcp/network/primary/en/index.html>).<sup>34</sup> Another association of medical journal editors, the World Association of Medical Editors (WAME) also has a policy that clinical trials should be registered at their inception.<sup>35</sup>

The World Medical Association's (WMA) Declaration of Helsinki is a statement of ethical principles for medical research where human participants are used.<sup>36</sup> As part of this statement, the authors recommended that all research studies using human participants should be publicly registered prior to subject recruitment, suggesting that there is also an ethical obligation to prospectively register RCTs.<sup>36</sup> In fact, the WHO also believes that trial registration is a "scientific, ethical and moral responsibility," and has created an International Clinical Trials Registry Platform (ICTRP) to display entries from trial registries around the world.<sup>37</sup> The Consolidated Standards of Reporting Trials (CONSORT) statement specifies trial registration as one of the mandatory reporting items, with authors required to either provide the name of the trial registry along with the registration number or a reason for not registering their trial.<sup>38</sup>

In the United States, the first federal law to require trial registration for clinical trials conducted under investigational new drug applications was passed in 1997.<sup>39</sup> In 2007, the Food and Drug Administration Amendments Act (FDAAA) was passed, with section 801 expanding the requirements to include all clinical trials conducted in the United States or using interventions manufactured in the United States.<sup>40</sup> However, there is currently no legislation in Canada regarding mandatory clinical trial registration, with Health Canada only encouraging sponsors to register their trials in publicly accessible registries<sup>41</sup> even though the Tri-Council Policy Statement: Ethical Conduct for Research Involving

Humans (TCPS2) states that all clinical trials should be prospectively registered in a publicly accessible registry that is acceptable to the WHO or the ICMJE.<sup>42</sup>

The instructions to authors or similarly named sections in journals describe guidelines for authors to follow when submitting to specific journals and serve as the main source of information regarding journal policies. A study using the ICMJE member journal list and CONSORT adopting journal list in 2011 found that only 47% of journals on the ICMJE list and 69% on the CONSORT list had a statement in the instructions to authors section about trial registration, with 90% and 94% of these statements explicitly requiring trial registration, respectively.<sup>43</sup> These results are concerning as journals that are ICMJE members or CONSORT statement endorsers require trial registration, potentially misleading submitting authors on the journal's policy by not being explicit about trial registration. In terms of anesthesiology journals, *Anesthesiology*, *British Journal of Anaesthesia (BJA)*, and *European Journal of Anaesthesiology (EJA)* described their mandatory trial registration policy in their instructions to authors sections starting in 2013,<sup>44</sup> 2009,<sup>45</sup> and 2015,<sup>46</sup> respectively. *Anaesthesia* and *Anesthesia & Analgesia (A&A)* both confirmed a mandatory trial registration policy after emails to the editorial office,<sup>47,48</sup> while the *Canadian Journal of Anesthesia (CJA)* only strongly recommends trial registration in its instructions to authors.<sup>49</sup>

### 2.3.3 Quality of Trial Registration

Studies assessing study publication bias or outcome reporting bias found that the percentage of trial protocols resulting in a published manuscript ranged from 21–93%, but consistently showed that positive studies were more likely to be published compared to negative studies.<sup>50</sup> However, another study found that trial registration was only

marginally associated with a RCT having positive findings.<sup>51</sup> In terms of outcome reporting bias, 40–62% of trials had major discrepancies in the primary outcomes when comparing the protocol to the publication, with 13–31% of primary outcomes from the protocol not reported in the publication, 10–18% of primary outcomes reported in the publication but not described in the protocol, and statistically significant outcomes having higher odds of being fully reported.<sup>50</sup>

Focusing on studies examining trial registration, the median proportion of studies with an identified primary outcome discrepancy between the trials registry entry and the publication was 31% (interquartile range (IQR): 17 to 45%).<sup>52</sup> When able to be assessed, these outcome discrepancies often favoured statistically significant results with a median of 50% (IQR: 28 to 64%).<sup>52</sup> The proportion of primary outcome discrepancies was highly variable among included studies,<sup>52</sup> potentially because studies used different groups of RCTs from different specialties, journals, and time periods. This suggests that a specialty-specific perspective would provide the most accurate information about trial registration quality in a particular specialty.

For RCTs published in the top five high impact factor anesthesiology journals in 2013 (*Anaesthesia*, *A&A*, *Anesthesiology*, *BJA*, and *Regional Anesthesia and Pain Medicine (RAPM)*), only 35% were prospectively registered in a publicly accessible trials registry.<sup>53</sup> When focusing on all RCTs registered, regardless of before or after patient enrollment, 48% had a major discrepancy between the trials registry and the publication.<sup>53</sup> In *Acta Anaesthesiologica Scandinavica (Acta)*, the registration rates were found to significantly increase from 17.1% in trials starting subject enrollment before

2010 to 63.2% in trials starting subject enrollment after 2010, when *Acta* had implemented mandatory registration for RCTs.<sup>54</sup>

## 2.4 Abstract Reporting and the CONSORT Statement

### 2.4.1 Abstracts

An abstract is a concise summary of the entire published study and is often used by readers to determine whether or not to read the full-text.<sup>55</sup> In fact, the full-text manuscript does not exist beyond its abstract for most readers since the abstract is the only part that potential referees read when invited to review a manuscript, the only part available when searching through electronic databases, and the only part quickly accessible after readers see an interesting title in a journal.<sup>56</sup> For readers that decide to read the full-text, the abstract will have already left a first impression and set the tone for the rest of the text.<sup>56</sup> When reading medical journals, clinicians reported reading only the abstract about two thirds of the time, either using the abstract as a means for extracting information from the full-text or as a screening tool to determine which studies are actually worth reading.<sup>57</sup> Abstracts for RCTs can help clinicians quickly decide whether reading the full-text is worthwhile to ensure medical decisions are made with appropriate evidence.

In addition, there may be financial, information technology, time, or language barriers that prevent or reduce access to all full-text manuscripts, increasing the importance of abstracts including all essential elements of the trial. Clinicians practising in community-based settings or low income countries may not have the financial ability or appropriate technology to access all full-text manuscripts, potentially relying on publicly available abstracts as the only source for key RCT evidence.<sup>55</sup> Busy clinicians also may not have

enough time to access or read many full-text manuscripts, leaving abstracts as the primary source of staying updated with new medical developments.<sup>58</sup> Finally, language barriers may result in only abstracts being used to make clinical decisions since only translated versions of the abstract are easily available and accessible.<sup>59</sup> Thus, abstracts are important tools for clinicians to make decisions about medical interventions; it is critical that they are reported comprehensively and accurately.

#### 2.4.2 CONSORT Statement

There are numerous guidelines for RCTs, addressing both the trial protocol<sup>60</sup> and the resulting publication. The most commonly used set of RCT reporting guidelines is the Consolidated Standards of Reporting Trials (CONSORT) statement, endorsed by over 600 biomedical journals, the ICMJE, and the WAME.<sup>61</sup> Recently updated in 2010, the CONSORT statement was published as a 25-item checklist with a flow diagram<sup>62-70</sup> and as an explanation/elaboration document<sup>38,71</sup> to improve the reporting of RCTs. Though RCTs are the gold standard for evaluating healthcare interventions, the results may be biased if RCTs are not properly designed or conducted.<sup>72</sup> In order for readers to accurately assess the results of a RCT, complete and clear information needs to be adequately reported.<sup>62-70</sup> Though the 2010 CONSORT statement is applicable for all RCTs, its main focus is for individually randomized, two group, parallel trials so extensions to the CONSORT statement have been developed to provide further guidance.<sup>62-70</sup> One of the extensions to the CONSORT statement is the CONSORT for Abstracts statement, aimed at improving the reporting of RCT abstracts.<sup>73,74</sup> The CONSORT for Abstracts statement was also integrated into the 2010 CONSORT

statement with a reference to refer to the CONSORT for Abstracts statement for specific guidance on reporting a structured abstract.<sup>62-70</sup>

### 2.4.3 CONSORT for Abstracts

The CONSORT for Abstracts statement was published in 2008 as a 17-item checklist,<sup>73</sup> with an accompanying explanation and elaboration document.<sup>74</sup> The CONSORT for Abstracts statement specifies 16 essential items that need to be reported in journal abstracts for RCTs and an additional item for conference abstracts (Appendix A).<sup>73</sup> To develop the checklist, authors created a list of items using existing tools for quality assessment and empirical evidence, then used a modified Delphi consensus method to select possible items for further discussion and revisions at the CONSORT group meeting.<sup>73</sup> An important issue for improving abstract reporting is space limitation, but preliminary work showed that all checklist items could be included within a 250-300 word limit.<sup>73</sup> Contrary to popular belief, the current Medline character limit for abstracts is actually 10,000,<sup>75</sup> allowing for much longer abstracts to be published if desired by journals.

In 2006, most abstracts of RCTs published in four high impact, general medical journals specified the study type, study population, objectives, and trial interventions, but were deficient in reporting markers of methodological quality such as allocation concealment, blinding, use of intent-to-treat analysis, and extent of lost to follow-up.<sup>76</sup> After the publication of the CONSORT for Abstracts statement, a study found that in five high impact, general medical journals, two journals (*Annals IM* and the *Lancet*) had an active policy to enforce the CONSORT for Abstracts guidelines and had an immediate increase in the mean number of items reported, while three journals (*JAMA*, *NEJM*, and the *BMJ*)

did not have an active policy to enforce the CONSORT for Abstracts guidelines and did not have any change in the mean number of items reported.<sup>77</sup> However, for RCTs published in 2010 for four of those journals (the *BMJ*, *JAMA*, the *Lancet*, and *NEJM*), an analysis using an equal proportion of abstracts from each journal still showed inadequate abstract reporting, with fewer than 50% reporting the trial design, randomization, harms, and conclusions.<sup>78</sup> Many abstract reporting differences were found between journals, suggesting that individual journal guidelines and house style have an impact on the adequacy of abstract reporting.<sup>78</sup>

The most recent update using up to 100 RCTs published between 2011 and 2014 from five high impact, general medical journals (*Annals IM*, the *BMJ*, *JAMA*, the *Lancet*, and *NEJM*) still found an overall adherence of only 67%, ranging from 55% in the *NEJM* to 78% in the *Lancet*.<sup>79</sup> Though many journals have endorsed the CONSORT for Abstracts statement, full adherence to the guidelines is still lacking. Many other studies have focused on other medical areas to determine abstract adherence with similar conclusions (i.e. that abstract reporting needed to be improved).<sup>55,59,80-84</sup> Discordance was also observed in 10% of RCTs focusing on lung cancer when comparing the conclusions in the abstracts to the full-text manuscript, with the majority of discrepancies arising from the experimental arm being described more favourably in the abstract than in the full-text.<sup>58</sup>

Abstract reporting in anesthesiology journals have been similarly unimpressive. Comparing pre-CONSORT abstracts in 2006 to post-CONSORT abstracts in 2009 for RCTs in four high-profile anesthesiology journals (*Anaesthesia*, *A&A*, *Anesthesiology*, and *EJA*), there were improvements in abstract reporting for the checklist items of

blinding, harms, appropriate title, and primary outcome – methods.<sup>85</sup> However, overall, there were still on average, less than a third of the recommended CONSORT for Abstracts checklist items adequately reported in the abstract.<sup>85</sup> Interestingly, RCT abstract adherence in the *Korean Journal of Anesthesiology (KJA)* increased from 41.7% in 2006 (pre-CONSORT) to 53.0% in 2012 (post-CONSORT), mainly due to improvements in the checklist items of trial design, randomization, blinding, and number randomized.<sup>86</sup> No anesthesiology studies have reported on whether items not reported in the abstract were available in the full-text manuscript.

For the top six general (non-pain-centric) anesthesiology journals as determined by impact factor, the instructions to authors do not specifically address adherence to the CONSORT for Abstracts statement, though *Anaesthesia* requires an unstructured abstract no more than 250 words,<sup>87</sup> *A&A* requires a structured abstract no more than 400 words,<sup>88</sup> *Anesthesiology* requires a structured abstract no more than 250 words,<sup>44</sup> *BJA* requires a structured abstract no more than 250 words,<sup>45</sup> *CJA* requires a structured abstract no more than 250 words,<sup>49</sup> and *EJA* requires a structured abstract no more than 300 words,<sup>46</sup> However, all six journals expect authors to be in compliance with the general CONSORT statement, which includes adhering to the CONSORT for Abstracts statement.<sup>44–46,49,87,88</sup>

## 2.5 Sample Size

### 2.5.1 Sample Size Calculation Assumptions

The sample size is the number of participants included in the RCT and the sample size calculation is used to determine the minimum number of participants that need to be included in order to detect a clinically relevant treatment effect.<sup>89</sup> RCTs with



inappropriate sample size calculations may be unethical as having a low sample size may prevent detecting differences between groups while having a high sample size may unnecessarily put participants at risk.<sup>90</sup> While there is disagreement over whether underpowered trials are ethical, *a priori* sample size calculations are frequently used when planning RCTs.<sup>91</sup> The CONSORT statement also recommends authors to report how sample size was determined, ideally with a high power to detect a statistically significant clinically relevant difference if one exists.<sup>38,71</sup>

Sample size estimation methods vary based on the study design and outcome type so to stay consistent, sample size calculations in this thesis will be restricted to two arm, parallel, superiority RCTs with binary or continuous outcomes. To calculate the sample size, the expected outcomes for each group (which implies the clinically relevant effect size), the type I ( $\alpha$ ) error level, the type II ( $\beta$ ) error or power level, and the standard deviation (only for continuous outcomes) are needed.<sup>38,92</sup> Allowances made for attrition or non-compliance should also be included when appropriate.<sup>38,92</sup>

For continuous outcomes, the standard deviations of the outcomes are also needed to determine the estimated variability of the data.<sup>93</sup> Also known as the false-positive rate, the type I ( $\alpha$ ) error level is the probability of finding a difference between study arms when the difference does not actually exist.<sup>93</sup> On the other hand, the type II ( $\beta$ ) error level is also known as the false-negative rate and is the probability of failing to find a difference between study arms when a difference actually exists.<sup>93</sup> In practice, this is usually reported as the power ( $1 - \beta$ ) or the true-positive rate, which means the probability of finding a difference between study arms if a difference actually exists.<sup>93</sup> The type I ( $\alpha$ ) error level and the type II ( $\beta$ ) error or power level are normally specified by the

investigators prior to the trial being conducted while the expected outcomes and variability for each group are estimated from previous studies or clinical expertise. Inappropriately estimated values for assumptions may result in underpowered or overpowered studies.<sup>94</sup>

## 2.5.2 Quality of Sample Size Calculation

The proportion of RCTs with the sample size reported has generally increased over the years.<sup>95</sup> All two-arm, parallel group RCTs published in 2005 and 2006 for six high impact factor general medical journals (*Annals IM*, the *BMJ*, *JAMA*, the *Lancet*, *NEJM*, and *PLoS Medicine*) were assessed to determine the quality of their sample size calculations.<sup>96</sup> Of included RCTs, 95% reported an *a priori* sample size calculation but only 53% reported all the required sample size calculation parameters.<sup>96</sup> When comparing the sample size assumptions for the control group with the observed results, the absolute difference was greater than 30% for 31% of eligible RCTs and greater than 50% for 17% of eligible RCTs,<sup>96</sup> suggesting either inappropriate estimations of sample size assumptions or a great degree of difficulty for accurate estimations of sample size assumptions.<sup>96</sup> A study focusing on RCTs published in 2002 for four of those journals (*Annals IM*, *JAMA*, the *Lancet*, and *NEJM*) also found that 80% of the reported standard deviations were greater than the estimated standard deviations, resulting in underestimates of sample variability (and hence, sample size) for the majority of RCTs.<sup>97</sup>

From 2000 to 2006, the percentage of anesthesiology RCTs reporting sample size estimates significantly increased from 52% to 86%.<sup>98</sup> The quality of sample size calculations was also assessed for RCTs published in 2013 for ten anesthesiology journals (*Acta*, *Anaesthesia*, *A&A*, *Anesthesiology*, *BJA*, *CJA*, *EJA*, *Journal of*

*Neurosurgical Anesthesiology, Pediatric Anesthesia, and RAPM*).<sup>99</sup> Similar to general medical journals, 92% of RCTs reported a sample size calculation, but surprisingly, 80% of these RCTs reported enough assumptions for sample size replication.<sup>99</sup> Comparing replicated to reported sample sizes, 28.7% of RCTs exceeded a 10% difference.<sup>99</sup> While the study did not report the percentage of RCTs with a greater than 10% difference between expected and reported outcomes, the overall median differences between expected and observed outcomes was 9% (IQR: -5 to 30%) for dichotomous outcomes and 9% (IQR: -30 to 30%) for continuous outcomes.<sup>99</sup>

## 2.6 Impact of Funding Source on RCTs

The percentage of industry funding in the United States for medical research rose from 46% in 1994 to 58% in 2012, with the share of industry funding spent on phase three trials increasing by 36% from 2004 to 2011.<sup>100</sup> Since the economic wellbeing of companies rely on favourable outcomes in the RCTs they fund, this may lead to biased research.<sup>101</sup>

The impact that the funding source has on the outcomes of RCTs has been extensively studied, particularly the effect of industry funding. One systematic review identified all studies published up to 2002 that assessed the extent, impact, or management of financial relationships in trials.<sup>102</sup> RCTs funded by industry were found to yield pro-industry conclusions (OR: 4.14; 95% CI: 2.72–6.32) when compared to non-industry-funded RCTs, even though industry-funded RCTs and non-industry-funded RCTs were found to be of similar methodological quality.<sup>102</sup> A similar systematic review was performed the same year focusing on quantitative studies assessing the effect of pharmaceutical

company funding on methodological quality and outcomes.<sup>103</sup> Clinical trials funded by pharmaceutical companies were associated with favouring the funders' products and comparing the treatment to an inactive control.<sup>103</sup> Half of the included studies indicated that industry- and non-industry-funded studies were of similar quality while the other half found that industry-funded studies clinical trials were of better quality.<sup>103</sup> Another systematic review included studies published between 2003 and 2006, with 17 studies finding that industry-funded trials favoured the products of their funders and 2 studies finding no statistically significant association between funding source and outcome.<sup>104</sup> A recent Cochrane systematic review in 2011 found convincing and consistent evidence for industry bias in pharmaceutical RCTs but insufficient evidence for medical device RCTs.<sup>105</sup> Studies funded by industry had favourable efficacy results (risk ratio (RR): 1.32; 95% CI: 1.21–1.44) and favourable conclusions (RR: 1.31; 95% CI: 1.20–1.44) more often than studies not funded by industry.<sup>106</sup>

Four possible explanations for the positive outcomes of RCTs funded by industry are that industry only funds superior treatments, industry conducts poor quality trials, industry does not select an appropriate comparator, and industry contributes to publication bias by discontinuing or not publishing trials with non-favourable results (these factors are not mutually exclusive).<sup>103</sup> However, the underlying principle for RCTs being conducted is uncertainty over whether the treatment or control are more beneficial,<sup>107</sup> meaning that investigators would not be able know whether a treatment was superior until after the trial was performed. Studies have consistently shown that RCTs funded by industry are of similar or superior methodological quality to RCTs not funded by industry,<sup>104</sup> suggesting that poor quality trials are not an adequate explanation unless other measures of quality

such as proper selection of controls are used. Publication bias is another likely explanation, but the increasing adoption of mandatory trial registration policies is aimed at alleviating this problem.<sup>24</sup>

A study in 2015 found that, compared with cardiovascular RCTs funded by the federal government, those that were funded by industry were more likely to report positive or favourable outcomes and less likely to report unfavourable results, even when eliminating publication bias by only using data from a trials registry, [clinicaltrials.gov](http://clinicaltrials.gov). However, recent studies investigating the association between industry funding and trial quality in specific disciplines such as synbiotics/probiotics/prebiotics added to infant formula and rheumatoid arthritis have not found industry funding to be associated with a higher likelihood of positive outcomes,<sup>108,109</sup> suggesting that industry has varying effects on RCTs in different specialties. More research concerning the effect of industry funding in anesthesiology RCTs is needed, including an explanation for the process through which industry can bias results.

## 2.7 Literature Cited

1. Morra D, Torgerson C, Loblaw A. Canadian Medical Residency Guide: Taking Control of Your Future Medical Career and Financial Life. 10th ed. RBC Royal Bank and Medical Career Services; 2012.
2. American Society of Anesthesiologists. What is Anesthesiology. Lifeline to Mod Med. [http://www.asahq.org/lifeline/types of anesthesia/what is anesthesiology](http://www.asahq.org/lifeline/types%20of%20anesthesia/what%20is%20anesthesiology). Accessed August 15, 2016.
3. Canadian Coordinating Office for Health Technology Assessment. Pre-assessment: Surgical Anesthesia Delivered by Non-physicians. *Can Agency Drugs Technol Heal*. 2004;(37).

[https://www.cadth.ca/media/pdf/273\\_No37\\_surgicalanesthesia\\_preassess\\_e.pdf](https://www.cadth.ca/media/pdf/273_No37_surgicalanesthesia_preassess_e.pdf).

Accessed August 15, 2016.

4. MacIntyre P, Stevens B, Collins S, Hewer I. Cost of education and earning potential for non-physician anesthesia providers. *AANA J*. 2014;82(1):25-31.

5. Canadian Medical Association. Anesthesiology Profile. *Can Spec profiles*. 2016. <https://www.cma.ca/Assets/assets-library/document/en/advocacy/profiles/anesthesiology-e.pdf>. Accessed August 15, 2016.

6. Friedman LM, Furberg CD, DeMets DL, Reboussin DM, Granger CB. *Fundamentals of Clinical Trials*. 5th ed. Springer International Publishing; 2015. doi:10.1007/978-3-319-18539-2.

7. Guyatt G, Rennie D, Meade MO, Cook DJ. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 3rd ed. United States of America: McGraw-Hill Education; 2015.

8. Sessler DI, Imrey PB. Clinical Research Methodology 3: Randomized Controlled Trials. *Anesth Analg*. 2015;121(4):1052-1064. doi:10.1213/ANE.0000000000000862.

9. Cohen AM, Stavri PZ, Hersh WR. A categorization and analysis of the criticisms of Evidence-Based Medicine. *Int J Med Inform*. 2004;73(1):35-43. doi:10.1016/j.ijmedinf.2003.11.002.

10. Walach H, Loeff M. Using a matrix-analytical approach to synthesizing evidence solved incompatibility problem in the hierarchy of evidence. *J Clin Epidemiol*. 2015;68(11):1251-1260. doi:10.1016/j.jclinepi.2015.03.027.

11. Tsui BC., Li LX., Ma V, Wagner AM, Finucane BT. Declining randomized clinical trials from Canadian anesthesia departments? *Can J Anesth*. 2006;53(3):226-235. doi:10.1007/BF03022207.

12. Lam DKM, Tsui BCH. Analysis of the number of publications by Canadian anesthesiology departments from 2000-2013. *Can J Anesth*. 2015;62(9):1021-1022. doi:10.1007/s12630-015-0385-0.
13. Pagel PS, Hudetz J a. Recent trends in publication of basic science and clinical research by United States investigators in anesthesia journals. *BMC Anesthesiol*. 2012;12(1):5. doi:10.1186/1471-2253-12-5.
14. Feneck RO, Natarajan N, Sebastian R, Naughton C. Decline in research publications from the United Kingdom in anaesthesia journals from 1997 to 2006. *Anaesthesia*. 2008;63(3):270-275. doi:10.1111/j.1365-2044.2008.05475.x.
15. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors. *N Engl J Med*. 2004;351(12):1250-1251. doi:10.1056/NEJMe048225.
16. McGauran N, Wieseler B, Kreis J, Schüler Y-B, Kölsch H, Kaiser T. Reporting bias in medical research - a narrative review. *Trials*. 2010;11:37. doi:10.1186/1745-6215-11-37.
17. Sterne J, Egger M, Moher D. Chapter 10: Addressing reporting biases. In: Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1. The Cochrane Collaboration; 2011.
18. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Hopewell S, ed. *Cochrane Database Syst Rev*. 2009;(1):MR000006. doi:10.1002/14651858.MR000006.pub3.
19. Olson CM, Rennie D, Cook D, et al. Publication bias in editorial decision making. *JAMA*. 2002;287(21):2825-2828.
20. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H. Publication bias and clinical trials. *Control Clin Trials*. 1987;8(4):343-353.

21. Okike K, Kocher M, Mehlman C, Heckman J, Bhandari M. Publication Bias in Orthopaedic Research: An Analysis of Scientific Factors Associated with Publication in The Journal of Bone and Joint Surgery (American Volume). *J Bone Jt Surg*. 2008;90(3):595. doi:10.2106/JBJS.G.00279.
22. Weber EJ, Callahan ML, Wears RL, Barton C, Young G. Unpublished Research From a Medical Specialty Meeting. *JAMA*. 1998;280(3):257-259. doi:10.1001/jama.280.3.257.
23. Sprague S, Bhandari M, Devereaux PJ, et al. Barriers to full-text publication following presentation of abstracts at annual orthopaedic meetings. *J Bone Joint Surg Am*. 2003;85-A(1):158-163.
24. Abaid LN, Grimes DA, Schulz KF. Reducing Publication Bias Through Trial Registration. *Obstet Gynecol*. 2007;109(6):1434-1437. doi:10.1097/01.AOG.0000266557.11064.2a.
25. Dolgin E. Publication bias continues despite clinical-trial registration. *Nature*. 2009. doi:10.1038/news.2009.902.
26. Emdin C, Odutayo A, Hsiao A, et al. Association of Cardiovascular Trial Registration With Positive Study Findings. *JAMA Intern Med*. 2015;175(2):304-307. doi:10.1001/jamainternmed.2014.6924.
27. Higgins J, Altman D, Sterne J. Chapter 8: Assessing risk of bias in included studies. In: Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1. The Cochrane Collaboration; 2011.
28. Hutton JL, Williamson PR. Bias in meta-analysis due to outcome variable selection within studies. *J R Stat Soc Ser C (Applied Stat)*. 2000;49(3):359-370. doi:10.1111/1467-9876.00197.
29. Dwan K, Altman DG, Arnaiz JA, et al. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias. Siegfried N, ed. *PLoS One*. 2008;3(8):e3081. doi:10.1371/journal.pone.0003081.



30. 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. doi:10.1136/bmj.c7153.
31. Goldacre B, Drysdale H, Powell-Smith A, et al. The COMPare Trials Project. 2016. [www.COMParE-trials.org](http://www.COMParE-trials.org). Accessed December 1, 2016.
32. Fredrickson MJ, Ilfeld BM. Prospective Trial Registration for Clinical Research. *Reg Anesth Pain Med*. 2011;36(6):619-624. doi:10.1097/AAP.0b013e318230fbc4.
33. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors. *N Engl J Med*. 2004;351(12):1250-1251. doi:10.1056/NEJMe048225.
34. International Committee of Medical Journal Editors. Clinical Trials Registration. *Int Comm Med J Ed*. 2016. <http://www.icmje.org/about-icmje/faqs/clinical-trials-registration/>. Accessed August 15, 2016.
35. WAME Editorial Policy Committee. The Registration of Clinical Trials. WAME Editor Policy Comm Policy Statements. 2005. [http://www.wame.org/about/policy-statements#Trial Reg](http://www.wame.org/about/policy-statements#Trial%20Reg). Accessed August 15, 2016.
36. World Medical Association. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. 2013. <http://www.wma.net/en/30publications/10policies/b3/index.html>. Accessed July 1, 2016.
37. World Health Organization. International Clinical Trials Registry Platform (ICTRP). World Heal Organ. <http://www.who.int/ictrp/en/>. Accessed August 15, 2016.
38. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340(1):c869-c869. doi:10.1136/bmj.c869.
39. U.S. National Institutes of Health. History, Policies, and Laws. *ClinicalTrials.gov*. 2016. <https://clinicaltrials.gov/ct2/about-site/history>. Accessed August 15, 2016.

40. U.S. National Institutes of Health. FDAAA 801 Requirements. ClinicalTrials.gov. 2015. <https://clinicaltrials.gov/ct2/manage-recs/fdaaa#WhichTrialsMustBeRegistered>. Accessed August 15, 2016.
41. Health Canada. Health Canada's Clinical Trials Database. Heal Canada. 2016. <http://www.hc-sc.gc.ca/dhp-mps/prodpharma/databasdonclin/index-eng.php>. Accessed August 15, 2016.
42. Interagency Advisory Panel on Research Ethics. TCPS 2 (2014)— the latest edition of Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Panel Res Ethics. 2014. <http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/Default/>. Accessed August 15, 2016.
43. Hooft L, Korevaar DA, Molenaar N, Bossuyt PMM, Scholten RJPM. Endorsement of ICMJE's clinical trial registration policy: A survey among journal editors. *Neth J Med*. 2014;72(7):349-355.
44. Anesthesiology Editorial Office. Instructions for Authors. *Anesthesiology*. <http://anesthesiology.pubs.asahq.org/public/instructionsforauthors.aspx>. Accessed August 15, 2016.
45. Mahajan RP, British Journal of Anaesthesia Editorial Team. Instructions to Authors. *Br J Anaesth*. 2016. [http://www.oxfordjournals.org/our\\_journals/bjaint/for\\_authors/general.html](http://www.oxfordjournals.org/our_journals/bjaint/for_authors/general.html). Accessed August 15, 2016.
46. European Journal of Anaesthesiology Editorial Office. Guidance for Authors on the Preparation and Submission of Manuscripts to the European Journal of Anaesthesiology. *Eur J Anaesthesiol*. 2016. <http://edmgr.ovid.com/eja/accounts/ifaauth.htm>. Accessed August 15, 2016.
47. Jarvis-Atton S. Trial Registration - Anaesthesia. 2016.
48. Williams-Klamborowski C. Trial Registration - Anesthesia & Analgesia. 2016.

49. Backman S, Baker A, Beattie S, et al. 2011 Canadian Journal of Anesthesia Guide for Authors. *Can J Anesth Can d'anesthésie*. 2011;58(7):668-696. doi:10.1007/s12630-011-9516-4.
50. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias — An Updated Review. Boutron I, ed. *PLoS One*. 2013;8(7):e66844. doi:10.1371/journal.pone.0066844.
51. Odutayo A, Emdin CA, Hsiao AJ, et al. Association between trial registration and positive study findings: cross sectional study (Epidemiological Study of Randomized Trials—ESORT). *BMJ*. 2017;j917. doi:10.1136/bmj.j917.
52. Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Med*. 2015;13:282. doi:10.1186/s12916-015-0520-3.
53. De Oliveira GS, Jung MJ, McCarthy RJ. Discrepancies Between Randomized Controlled Trial Registry Entries and Content of Corresponding Manuscripts Reported in Anesthesiology Journals. *Anesth Analg*. 2015;121(4):1030-1033. doi:10.1213/ANE.0000000000000824.
54. Østervig RM, Sonne A, Rasmussen LS. Registration of randomized clinical trials - a challenge. *Acta Anaesthesiol Scand*. 2015;59(8):986-989. doi:10.1111/aas.12575.
55. Sivendran S, Newport K, Horst M, Albert A, Galsky MD. Reporting quality of abstracts in phase III clinical trials of systemic therapy in metastatic solid malignancies. *Trials*. 2015;16(1):341. doi:10.1186/s13063-015-0885-9.
56. Andrade C. How to write a good abstract for a scientific paper or conference presentation. *Indian J Psychiatry*. 2011;53(2):172. doi:10.4103/0019-5545.82558.
57. Saint S, Christakis DA, Saha S, et al. Journal reading habits of internists. *J Gen Intern Med*. 2000;15(12):881-884. doi:10.1046/j.1525-1497.2000.00202.x.

58. Altwaairgi AK, Booth CM, Hopman WM, Baetz TD. Discordance Between Conclusions Stated in the Abstract and Conclusions in the Article: Analysis of Published Randomized Controlled Trials of Systemic Therapy in Lung Cancer. *J Clin Oncol*. 2012;30(28):3552-3557. doi:10.1200/JCO.2012.41.8319.
59. Wang L, Li Y, Li J, et al. Quality of reporting of trial abstracts needs to be improved: using the CONSORT for abstracts to assess the four leading Chinese medical journals of traditional Chinese medicine. *Trials*. 2010;11(1):75. doi:10.1186/1745-6215-11-75.
60. Tetzlaff JM, Chan A-W, Kitchen J, Sampson M, Tricco AC, Moher D. Guidelines for randomized clinical trial protocol content: a systematic review. *Syst Rev*. 2012;1(1):43. doi:10.1186/2046-4053-1-43.
61. The CONSORT Group. Impact of CONSORT. CONSORT. <http://www.consort-statement.org/about-consort/impact-of-consort>. Accessed August 15, 2016.
62. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med*. 2010;152(11):726-732. doi:10.7326/0003-4819-152-11-201006010-00232.
63. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med*. 2010;8:18. doi:10.1186/1741-7015-8-18.
64. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.
65. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol*. 2010;63(8):834-840. doi:10.1016/j.jclinepi.2010.02.005.
66. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Obstet Gynecol*. 2010;115(5):1063-1070. doi:10.1097/AOG.0b013e3181d9d421.

67. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomized trials. *Open Med.* 2010;4(1):e60-8.
68. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *PLoS Med.* 2010;7(3):e1000251. doi:10.1371/journal.pmed.1000251.
69. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Trials.* 2010;11:32. doi:10.1186/1745-6215-11-32.
70. Schulz KF, Altman DG, Moher D, Fergusson D. CONSORT 2010 changes and testing blindness in RCTs. *Lancet.* 2010;375(9721):1144-1146. doi:10.1016/S0140-6736(10)60413-8.
71. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol.* 2010;63(8):e1-e37. doi:10.1016/j.jclinepi.2010.03.004.
72. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ.* 2001;323(7303):42-46. doi:10.1136/bmj.323.7303.42.
73. Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet.* 2008;371(9609):281-283. doi:10.1016/S0140-6736(07)61835-2.
74. Hopewell S, Clarke M, Moher D, et al. CONSORT for Reporting Randomized Controlled Trials in Journal and Conference Abstracts: Explanation and Elaboration. von Elm E, ed. *PLoS Med.* 2008;5(1):e20. doi:10.1371/journal.pmed.0050020.
75. U.S. National Library of Medicine. MEDLINE/PubMed Data Element (Field) Descriptions. *US Natl Libr Med.* 2016. <https://www.nlm.nih.gov/bsd/mms/medlineelements.html>. Accessed August 15, 2016.

76. Berwanger O, Ribeiro RA, Finkelsztein A, et al. The quality of reporting of trial abstracts is suboptimal: Survey of major general medical journals. *J Clin Epidemiol*. 2009;62(4):387-392. doi:10.1016/j.jclinepi.2008.05.013.
77. Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. *BMJ*. 2012;344:e4178. doi:10.1136/bmj.e4178.
78. Ghimire S, Kyung E, Kang W, Kim E. Assessment of adherence to the CONSORT statement for quality of reports on randomized controlled trial abstracts from four high-impact general medical journals. *Trials*. 2012;13:77. doi:10.1186/1745-6215-13-77.
79. Hays M, Andrews M, Wilson R, Callender D, O'Malley PG, Douglas K. Reporting quality of randomised controlled trial abstracts among high-impact general medical journals: a review and analysis. *BMJ Open*. 2016;6(7):e011082. doi:10.1136/bmjopen-2016-011082.
80. Fleming PS, Buckley N, Seehra J, Polychronopoulou A, Pandis N. Reporting quality of abstracts of randomized controlled trials published in leading orthodontic journals from 2006 to 2011. *Am J Orthod Dentofac Orthop*. 2012;142(4):451-458. doi:10.1016/j.ajodo.2012.05.013.
81. Seehra J, Wright NS, Polychronopoulou A, Cobourne MT, Pandis N. Reporting Quality of Abstracts of Randomized Controlled Trials Published in Dental Specialty Journals. *J Evid Based Dent Pract*. 2013;13(1):1-8. doi:10.1016/j.jebdp.2012.11.001.
82. Cui Q, Tian J, Song X, Yang K. Does the CONSORT checklist for abstracts improve the quality of reports of randomized controlled trials on clinical pathways? *J Eval Clin Pract*. 2014;20(6):827-833. doi:10.1111/jep.12200.
83. Chen Y, Li J, Ai C, et al. Assessment of the Quality of Reporting in Abstracts of Randomized Controlled Trials Published in Five Leading Chinese Medical Journals. Earnest CP, ed. *PLoS One*. 2010;5(8):e11926. doi:10.1371/journal.pone.0011926.

84. Chhapola V, Tiwari S, Brar R, Kanwal SK. An interrupted time series analysis showed suboptimal improvement in reporting quality of trial abstract. *J Clin Epidemiol.* 2016;71:11-17. doi:10.1016/j.jclinepi.2015.10.013.
85. Can OS, Yilmaz AA, Hasdogan M, et al. Has the quality of abstracts for randomised controlled trials improved since the release of Consolidated Standards of Reporting Trial guideline for abstract reporting? A survey of four high-profile anaesthesia journals. *Eur J Anaesthesiol.* 2011;28(7):485-492. doi:10.1097/EJA.0b013e32833fb96f.
86. Shin WJ, Kim YO, Oh JH, Chung JS, Kim KH. Is there any quality improvement in the randomized controlled trial abstracts in the Korean Journal of Anesthesiology after the publication of the CONSORT abstract guidelines in 2008? *Korean J Anesthesiol.* 2015;68(4):420-422. doi:10.4097/kjae.2015.68.4.420.
87. Klein AA, Anaesthesia Editorial Team. Author Guidelines. *Anaesthesia.* [http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1365-2044/homepage/ForAuthors.html](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1365-2044/homepage/ForAuthors.html). Accessed August 15, 2016.
88. Anesthesia & Analgesia Editorial Office. Instructions for Authors. *Anesth Analg.* 2016. <http://edmgr.ovid.com/aa/accounts/ifaauth.htm>. Accessed August 15, 2016.
89. Noordzij M, Tripepi G, Dekker FW, Zoccali C, Tanck MW, Jager KJ. Sample size calculations: Basic principles and common pitfalls. *Nephrol Dial Transplant.* 2010;25(5):1388-1393. doi:10.1093/ndt/gfp732.
90. Gupta K, Attri J, Singh A, Kaur H, Kaur G. Basic concepts for sample size calculation: Critical step for any clinical trials! *Saudi J Anaesth.* 2016;10(3):328. doi:10.4103/1658-354X.174918.
91. Schulz KF, Grimes D a. Sample size calculations in randomised trials: mandatory and mystical. *Lancet.* 2005;365(9467):1348-1353. doi:10.1016/S0140-6736(05)61034-3.
92. Campbell M, Julious S, Altman D. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in paired comparisons. *BMJ.* 1995;311(7013):1145-1148.

93. Lerman J. Study design in clinical research: sample size estimation and power analysis. *Can J Anesth*. 1996;43(2):184-191.
94. Tavernier E, Giraudeau B. Sample Size Calculation: Inaccurate A Priori Assumptions for Nuisance Parameters Can Greatly Affect the Power of a Randomized Controlled Trial. *PLoS One*. 2015;10(7):e0132578. doi:10.1371/journal.pone.0132578.
95. Chen H, Zhang N, Lu X, Chen S. Caution regarding the choice of standard deviations to guide sample size calculations in clinical trials. *Clin Trials*. 2013;10(4):522-529. doi:10.1177/1740774513490250.
96. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials: review. *BMJ*. 2009;338(1):b1732-b1732. doi:10.1136/bmj.b1732.
97. Vickers AJ. Underpowering in randomized trials reporting a sample size calculation. *J Clin Epidemiol*. 2003;56(8):717-720. doi:10.1016/S0895-4356(03)00141-0.
98. Greenfield MLVH, Mhyre JM, Mashour GA, Blum JM, Yen EC, Rosenberg AL. Improvement in the Quality of Randomized Controlled Trials Among General Anesthesiology Journals 2000 to 2006: A 6-Year Follow-Up. *Anesth Analg*. 2009;108(6):1916-1921. doi:10.1213/ane.0b013e31819fe6d7.
99. Abdulatif M, Mukhtar A, Obayah G. Pitfalls in reporting sample size calculation in randomized controlled trials published in leading anaesthesia journals: a systematic review. *Br J Anaesth*. 2015;115(5):699-707. doi:10.1093/bja/aev166.
100. Moses H, Matheson DHM, Cairns-Smith S, George BP, Palisch C, Dorsey ER. The Anatomy of Medical Research. *JAMA*. 2015;313(2):174. doi:10.1001/jama.2014.15939.
101. Lexchin J. Sponsorship bias in clinical research. *Int J Risk Saf Med*. 2012;24(4):233-242. doi:10.3233/JRS-2012-0574.
102. Bekelman JE, Li Y, Gross CP. Scope and Impact of Financial Conflicts of Interest in Biomedical Research. *JAMA*. 2003;289(4):454-465. doi:10.1001/jama.289.4.454.



103. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003;326(7400):1167-1170. doi:10.1136/bmj.326.7400.1167.
104. Sismondo S. Pharmaceutical company funding and its consequences: A qualitative systematic review. *Contemp Clin Trials*. 2008;29(2):109-113. doi:10.1016/j.cct.2007.08.001.
105. Bero L. Industry Sponsorship and Research Outcome. *JAMA Intern Med*. 2013;173(7):580. doi:10.1001/jamainternmed.2013.4190.
106. Lundh A, Sismondo S, Lexchin J, Busuioc O a, Bero L. Industry sponsorship and research outcome. In: Lundh A, ed. *Cochrane Database of Systematic Reviews*. Vol 12. Chichester, UK: John Wiley & Sons, Ltd; 2012:MR000033. doi:10.1002/14651858.MR000033.pub2.
107. Djulbegovic B. Acknowledgment of uncertainty: a fundamental means to ensure scientific and ethical validity in clinical research. *Curr Oncol Rep*. 2001;3(5):389-395.
108. Mugambi MN, Musekiwa A, Lombard M, Young T, Blaauw R. Association between funding source, methodological quality and research outcomes in randomized controlled trials of synbiotics, probiotics and prebiotics added to infant formula: A Systematic Review. *BMC Med Res Methodol*. 2013;13(1):137. doi:10.1186/1471-2288-13-137.
109. Khan NA, Lombeida JI, Singh M, Spencer HJ, Torralba KD. Association of industry funding with the outcome and quality of randomized controlled trials of drug therapy for rheumatoid arthritis. *Arthritis Rheum*. 2012;64(7):2059-2067. doi:10.1002/art.34393.

## Chapter 3

### 3 Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals

Philip M. Jones<sup>1,2</sup>, Jeffrey T. Y. Chow<sup>1</sup>, Miguel F. Arango<sup>2</sup>, Jason A. Fridfinnson<sup>2</sup>, Nan Gai<sup>2</sup>, Kevin Lam<sup>1</sup>, Timothy P. Turkstra<sup>2</sup>

<sup>1</sup>Department of Epidemiology & Biostatistics, Schulich School of Medicine & Dentistry, The University of Western Ontario

<sup>2</sup>Department of Anesthesia & Perioperative Medicine, Schulich School of Medicine & Dentistry, The University of Western Ontario

---

---

**Note:** A version of this chapter has been submitted and accepted for publication:

Jones PM, Chow JTY, Arango MF, Fridfinnson JA, Gai N, Lam K, Turkstra TP. Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals. *Anesth Analg*. 2017; *in-press*

### 3.1 Introduction

Generally considered to be at the top of the evidence hierarchy for evaluating therapeutic interventions, randomized clinical trials (RCTs) and systematic reviews/meta-analyses of RCTs provide important evidence for clinical decision-making.<sup>1</sup> While RCTs provide high quality evidence by minimizing selection bias through random allocation to groups, there are many other factors that may distort or bias the results. One tool used to improve RCT quality is trial registration, whose purpose is to reduce publication bias (i.e., not publishing studies deemed to be negative, uninteresting, or potentially damaging to the study's sponsor) and selective outcome reporting (i.e., publishing only a subset of all outcomes measured, often favouring outcomes that are statistically significant).<sup>2</sup>

Recognizing the importance of trial registration, the International Committee of Medical Journal Editors (ICMJE) mandates that all clinical trials commencing after July 1, 2005 need to be registered in a public trials registry in order to be published in any of the ICMJE member journals.<sup>3</sup>

Pre-specifying outcomes helps protect against data dredging and/or selective reporting, where only favourable outcomes are reported and unfavourable outcomes are obscured.<sup>4</sup> If authors elect to report different outcomes than the outcomes they registered, trial registries can also serve as an audit trail for the authors' original intentions, allowing researchers to check published RCTs for potential outcome switching.<sup>5</sup> Other researchers have reviewed the overall prevalence of outcome switching in the medical literature, finding a median discrepancy rate of 41% (interquartile range [IQR]: 33 – 48%) when comparing prospectively registered outcomes to reported outcomes.<sup>6</sup> Though selective reporting has been studied in general and in specialty journals, only one study has

focused on the anesthesiology specialty, finding that 48% of registered trials had a major discrepancy when comparing the registry entry and the published manuscript.<sup>7</sup> While De Oliveira et al. included RCTs published in one year (2013) from the top five general anesthesiology journals by impact factor,<sup>7</sup> three of the included journals, *Anesthesiology*, *Anaesthesia*, and *Anesthesia & Analgesia* only mandate trial registration for clinical trials that begin patient enrollment after 2013,<sup>8</sup> 2014,<sup>9</sup> and 2014,<sup>10</sup> respectively. No study has yet performed a longitudinal analysis of outcome discrepancies in the anesthesia literature, nor has any study performed a detailed analysis of secondary outcome discrepancies.

Our objective was to expand upon previous work by longitudinally examining the rate of adequate trial registration in the anesthesiology literature, focusing on RCTs published in 2007, 2010, 2013, and 2015 for the top six general anesthesiology journals. We also planned to thoroughly investigate outcome discrepancies (i.e., differences between the outcomes registered and the outcomes reported) in both primary and secondary outcomes.

## 3.2 Methods

Publicly available, our study protocol was finalized in September 2016 and all analyses (except for the raw incidences of adequate trial registration) were conducted after this time.<sup>11</sup>

Using the journals' official websites, Tables of Contents were independently hand-searched by two reviewers to identify eligible RCTs, with disagreements resolved by consensus. RCTs were only included if published in 2007, 2010, 2013, or 2015 in one of the top six general anesthesiology journals as determined by impact factor: *Anaesthesia*,

*Anesthesia & Analgesia, Anesthesiology, British Journal of Anaesthesia, Canadian Journal of Anesthesia, and European Journal of Anaesthesiology.*<sup>12</sup> An electronic database was used to systematically screen and extract data from RCTs regarding general study characteristics and trial registration status. An RCT was defined as a prospective study assessing randomly allocated health care interventions in human participants and identified by authors clearly reporting random allocation of participants to the study arms through words like “random”, “randomized”, and “randomised” in the publication. Observational studies (e.g., case-control, or cohort studies), learning curve studies, cadaver studies, cost-effectiveness studies, dose-finding or dose-response studies which were not designed to test a clinical intervention, and diagnostic test accuracy studies were excluded. We also excluded meta-analyses, editorials, letters to the editor, narrative reviews, animal studies, manikin studies, simulation studies, duplicate reports, re-analyses of previously published RCTs, and studies published in the correspondence section or in supplemental issues. If studies included two trials within one article, only data from the larger trial or phase was used.

RCTs were considered adequately registered if the trial was registered in a publicly available trial registry prior to the first patient being enrolled and if the registry entry had a clearly defined primary outcome. The trial registration status was systematically ascertained through the use of full-text searching, trial registry searching, and emails to the corresponding author. Trial registry searching used the following registries: Clinicaltrials.gov, the International Standard Randomized Controlled Trial Number Register (ISRCTN), and the World Health Organization Clinical Trials Search Portal. If we could not positively determine a trial’s registration status by examining the full text of

the article or *via* registry searches, we sent up to two standardized emails to the corresponding author with a one-week gap between emails to determine the registration status. If the registration status of the trial still could not be determined after the two emails, the trial was considered to be unregistered. Trials registered using only the EudraCT database (<https://www.clinicaltrialsregister.eu>) were considered not registered since trials contained within the European Union Clinical Trials Registry (EUCTR) were not made publicly available until recently.<sup>13</sup>

For each registered trial, the date of trial registration and first patient enrollment were collected to determine whether the trial was adequately registered. Trial registration date was defined as the date submitted for trial registration in the trials registry. If the date of first patient enrollment was not specified, the study start date was used instead. If the study start date specified only the month and year, trials were considered to be inadequately registered if the month occurred after the trial registration date. Trials registered after the first patient was enrolled or labelled as retrospectively registered in the trials registry were considered to be inadequately registered. An outcome was considered to be clearly defined if the outcome was clearly and unambiguously identified with a specific time frame in the trials registry. Only RCTs that were assessed to be adequately registered trials were further analyzed to determine the extent of outcome discrepancies.

The journal, year, funding source, number of authors, sample size, and adequate registration status were also extracted from each included RCT to describe the general characteristics of the included RCTs. To address the primary objective, the number of outcomes reported in the trial registry, the number of outcomes reported in the published

article, and the presence of outcome discrepancies for included RCTs were also assessed. If the primary outcome was not explicitly described in the publication, the outcome used in the sample size calculation was considered to be the primary outcome. The number of participants per secondary outcome reported was calculated by dividing the sample size (number of participants randomized) by the number of secondary outcomes reported. Due to the number of RCTs included, RCTs were assigned to a single reviewer for data extraction and a second reviewer was consulted if a variable was unclear, with resolution by consensus.

To provide more granularity, the type of outcome discrepancy was extracted with as much detail as possible. RCTs were considered to have a primary outcome discrepancy if:

- A) a registered primary outcome was not reported as a primary outcome (e.g., *24-hr morphine consumption* registered but not reported as a primary outcome)
- B) a reported primary outcome was not registered as a primary outcome (e.g., *6-hr morphine consumption* reported but not registered)
- C) the timing of a reported primary outcome was different from the timing of a registered primary outcome (e.g., the registered outcome was *30-day mortality* but the reported outcome was *7-day mortality*).

A registered primary outcome that was not reported as a primary outcome might have been not reported at all or reported as a secondary outcome, while a reported primary outcome that was not registered as a primary outcome might have been not registered at all or registered as a secondary outcome.

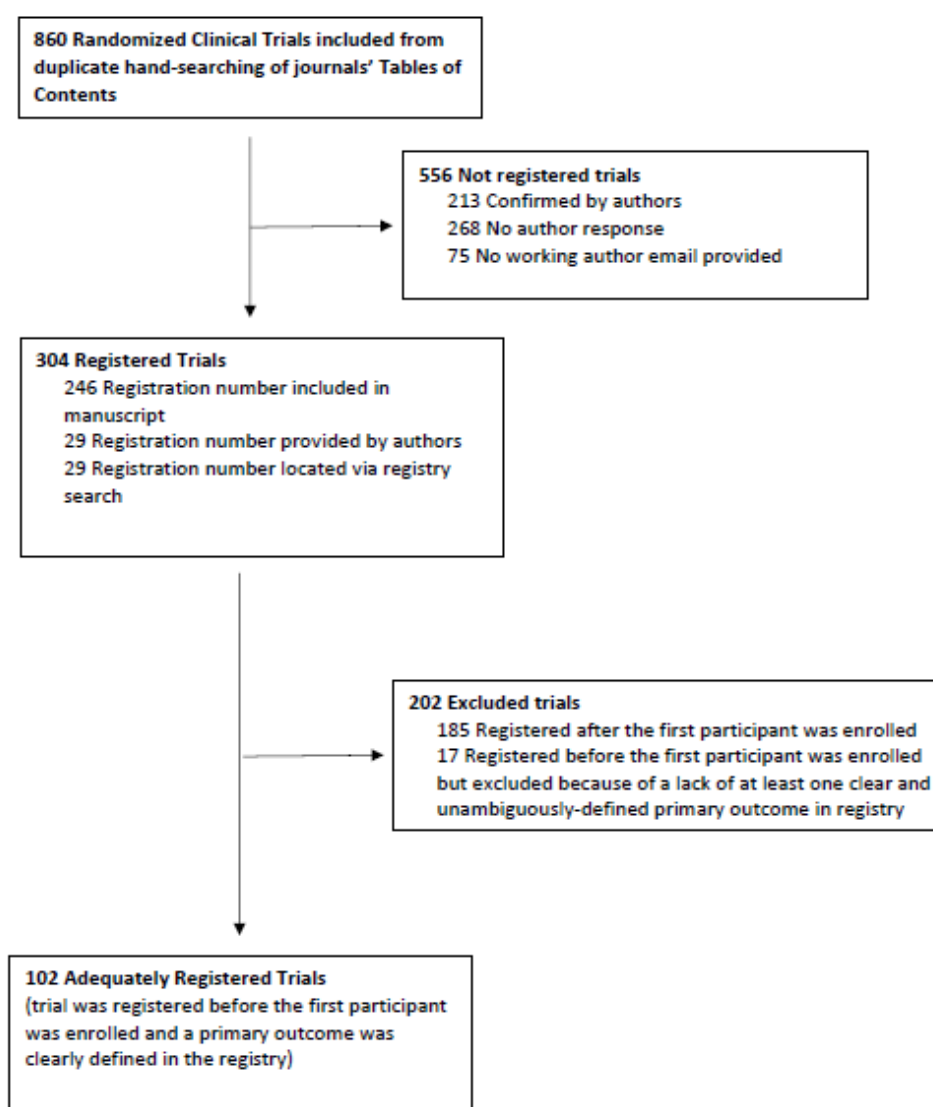
Secondary outcomes were defined using a similar structure. A registered secondary outcome that was not reported as a secondary outcome might have been not reported or reported as a primary outcome while a reported secondary outcome that was not registered might have been not registered or registered as a primary outcome. Because of this hierarchical detail, some outcomes are identical but phrased differently (i.e., a *registered* primary outcome *reported* as a secondary outcome is the same as a *reported* secondary outcome *registered* as a primary outcome; a *reported* primary outcome *registered* as a secondary outcome is the same as a *registered* secondary outcome *reported* as a primary outcome).

When an outcome discrepancy was identified, the reviewer determined the discrepancy to favour statistical significance if the reported outcome in question was statistically significant, as described by the authors. However, when a *registered* primary outcome was *reported* as a secondary outcome, the discrepancy was said to favour statistical significance if the registered primary outcome in question was *not* statistically significant. In discrepancies where a registered primary or secondary outcome was not reported at all, statistical significance favouring could not be determined. Outcomes were considered statistically significant if reported thus by the authors. Stata 14 statistical software (College Park, Texas) was used for all calculations and figures. For each of the four years measured, the number and percentage of RCTs with outcome discrepancies were summarized along with other descriptive statistics such as trial characteristics and number of outcomes. No inferential statistical tests were performed.



### 3.3 Results

Of the 860 RCTs that fulfilled the inclusion/exclusion criteria, 556 were not registered and 202 were either registered after the first participant was enrolled or did not have a clear unambiguously-defined primary outcome in the registry entry, resulting in 102 adequately registered trials assessed for the outcome discrepancy part of the study (Figure 3-1). Table 3-1 describes the general characteristics for all of the included RCTs.



**Figure 3-1: Flowchart for inclusion of trials defined as adequately registered**

**Table 3-1: Characteristics of trials**

Data presented are the number of trials / eligible trials (%) unless otherwise stated. Percentages may not sum to 100% due to rounding.

<sup>a</sup> Adequate registration means that the trial was registered before the first participant was enrolled and that a primary outcome was clearly defined in the registry.

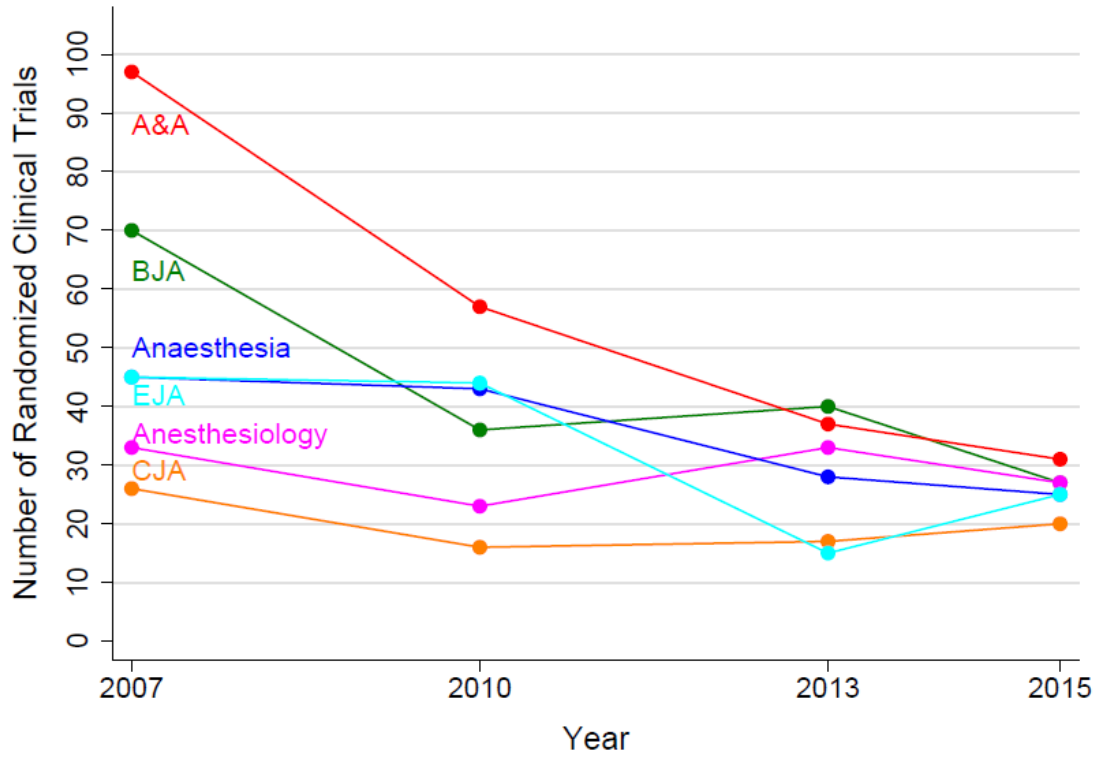
<sup>b</sup> No trial registration was located during the systematic search using the full-text article, clinical trial registries, and corresponding author emails as described in the Methods of this manuscript

Abbreviations: IQR, inter-quartile range.

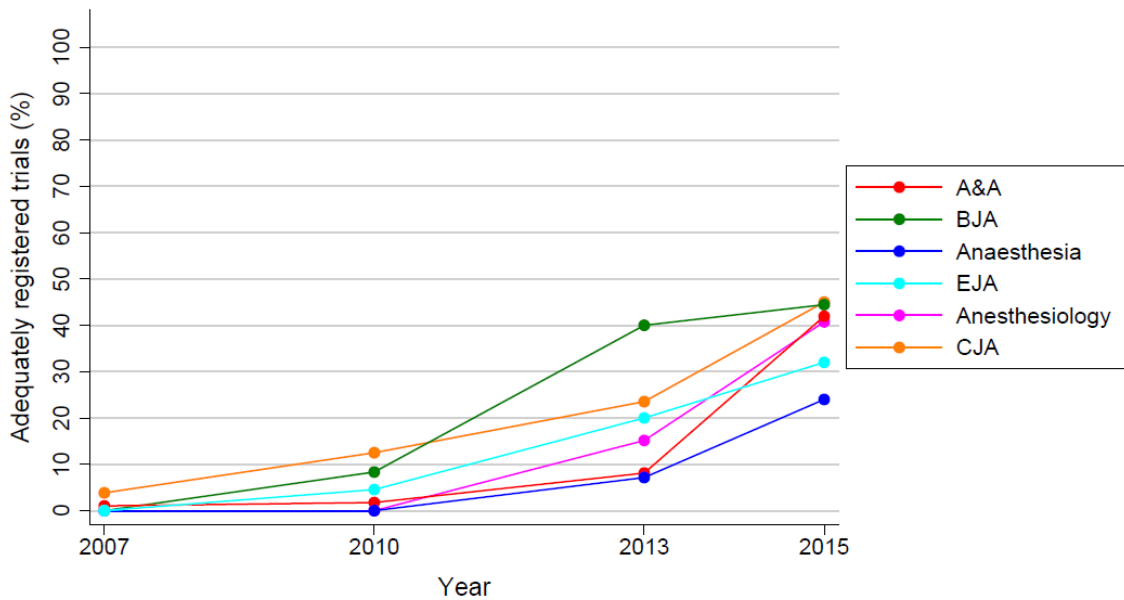
	2007 (n=316)	2010 (n=219)	2013 (n=170)	2015 (n=155)
<b>Journal</b>				
Anaesthesia	45/316 (14%)	43/219 (20%)	28/170 (16%)	25/155 (16%)
Anesthesia & Analgesia	97/316 (31%)	57/219 (26%)	37/170 (22%)	31/155 (20%)
Anesthesiology	33/316 (10%)	23/219 (11%)	33/170 (19%)	27/155 (17%)
British Journal of Anaesthesia	70/316 (22%)	36/219 (16%)	40/170 (24%)	27/155 (17%)
Canadian Journal of Anesthesia	26/316 (8%)	16/219 (7%)	17/170 (10%)	20/155 (13%)
European Journal of Anaesthesiology	45/316 (14%)	44/219 (20%)	15/170 (9%)	25/155 (16%)
Number of authors, median (IQR)	6 (4 – 7)	6 (5 – 7)	6 (5 – 8)	6 (5 – 8)
Trial sample size, median (IQR)	66 (40 – 110)	70 (44 – 120)	89 (56 – 150)	92 (57 – 150)
<b>Trials which were adequately registered <sup>a</sup></b>	2/316 (0.6%)	8/219 (4%)	33/170 (19%)	59/155 (38%)
Registration identified in article	1/2 (50%)	8/8 (100%)	29/33 (88%)	54/59 (92%)
Registration not identified in article	1/2 (50%)	0/8 (0%)	4/33 (12%)	5/59 (8%)
<b>Inadequately registered trials <sup>a</sup></b>	314/316 (99%)	211/219 (96%)	137/170 (81%)	96/155 (62%)
No trial registration located <sup>b</sup>	285/314 (91%)	167/211 (79%)	73/137 (53%)	31/96 (32%)
Registration occurred after the first participant was enrolled	22/314 (7%)	34/211 (16%)	45/137 (33%)	56/96 (58%)
No or unclear primary outcome was specified in registry	0/314 (0%)	1/211 (0.5%)	8/137 (6%)	8/96 (8%)
Registration occurred after the first participant was enrolled and no/unclear primary outcome was specified in the registry	7/314 (2%)	9/211 (4%)	11/137 (8%)	1/96 (1%)

Figure 3-2 demonstrates a decreasing number of RCTs published in each journal over time, with only 155 RCTs identified in 2015 (compared to 316 in 2007). However, over time, an increasing proportion of these RCTs were adequately registered (Figure 3-3). This trend was consistent across all included journals, but overall, only 38% of RCTs were adequately registered in 2015. A detailed breakdown of why RCTs were considered inadequately registered is available in Table 3-1, with the reason “no trial registration located” decreasing over time and the reason “registration occurred after first patient being enrolled” correspondingly increasing over time.

The majority of adequately registered trials had only one primary outcome, but almost one-third of RCTs had more than one (Table 3-2). There was a large discrepancy between the median [IQR] number of secondary outcomes registered (4 [1 - 8]) and the number of secondary outcomes reported (18 [10 – 29]) (Table 3-2). The median [IQR] number of study participants per secondary outcome reported declined from 2007 (23 [13 – 33]) to 2015 (6 [2 – 11]) (Table 3-2).



**Figure 3-2: Number of randomized clinical trials published by journal over time**



**Figure 3-3: Percentage of adequately registered trials by journal over time**

An adequately registered trial was defined as being registered before the first participant was enrolled with a primary outcome clearly defined in the registry.

**Table 3-2: Description of primary and secondary outcomes in adequately registered trials**

Data presented are the number of trials (%) unless otherwise stated. An outcome was considered to be clearly defined if the outcome was clearly and unambiguously identified with a specific time frame in the trials registry. By definition, all primary outcomes for the adequately registered trials in this Table were clearly defined.

Abbreviations: IQR, interquartile range.

	2007 (n=2)	2010 (n=8)	2013 (n=33)	2015 (n=59)
<b>Registered outcomes in trial registry</b>				
Only one primary outcome registered in trial registry	1 (50%)	4 (50%)	25 (76%)	42 (71%)
More than one primary outcome registered in trial registry	1 (50%)	4 (50%)	8 (24%)	17 (29%)
Number of secondary outcomes registered in trial registry, median (IQR)	2.5 (0 – 5)	7.5 (1.5 – 12.5)	3 (1 – 7)	4 (2 – 8)
Number of trials with at least one unclear secondary outcome registered in trial registry	0 (0%)	3 (38%)	10 (30%)	16 (27%)
<b>Reported outcomes in published article</b>				
One primary outcome reported in published article	2 (100%)	4 (50%)	27 (82%)	42 (71%)
More than one primary outcome reported in published article	0 (0%)	4 (50%)	6 (18%)	17 (29%)
Number of secondary outcomes reported in published article, median (IQR)	5.5 (5 – 6)	19.5 (7.5 – 29.5)	12 (8 – 26)	19 (11 – 36)
Number of trials with at least one unclear secondary outcome reported in published article	0 (0%)	1 (13%)	4 (12%)	2 (3%)
Number of participants per secondary outcome reported, median (IQR)	23 (13 – 33)	4 (3 – 16)	6 (3 – 14)	6 (2 – 11)

Table 3-3 presents the data on outcome discrepancies, showing that 92% of adequately registered trials in 2015 had at least one primary or secondary outcome discrepancy, with 59% of them favouring statistical significance. Results for RCTs published in 2007 are difficult to interpret due to the small number of adequately registered trials, but RCTs published in 2013 and 2015 had a slightly smaller percentage of RCTs with at least one primary or secondary outcome discrepancy compared to RCTs published in 2007 and 2010. When stratifying by type of outcome discrepancy, adequately registered trials published in 2007 and 2010 had a higher percentage of having at least one primary outcome discrepancy but the trend is reversed for secondary discrepancies, with adequately registered trials published in 2013 and 2015 having a higher percentage of having at least one secondary outcome discrepancy. Correspondingly, in 2015, many of the outcome discrepancies were related to issues with the secondary outcome, with 90% of adequately registered trials having at least one secondary outcome discrepancy. In 2015, 86% of adequately registered trials had at least one reported secondary outcome not registered, accounting for most of the secondary outcome discrepancies observed.

Overall, the outcome discrepancies with the lowest percentages were having at least one registered secondary outcome reported as a primary outcome (6%) and having at least one registered primary outcome not reported at all (11%). The number and percentage of outcome discrepancies stratified by journal can be found in Appendix C.

**Table 3-3: Differences between outcomes when comparing the published trial to the trial registry among adequately registered trials**

A: Data presented are the number of trials (%). All percentages are column percentages with the denominator set to the total number of trials published.

B: Data presented are the *proportion* of trials where the identified discrepancy favoured statistical significance (see article text for details).

All uses of ‘registered’ pertain to the outcome as registered in the trial registry and all uses of ‘reported’ pertain to the outcome as reported in the published manuscript. If there are one or more occurrences of the discrepancy in a trial, the trial will be counted as having the described discrepancy. Subcategories are not mutually exclusive so if individual studies have more than one discrepancy, the sum of the subcategories will be larger than the parent category.

“N/A” represents values where it could not be determined whether the discrepancy favoured statistical significance: either when the discrepancy involved an outcome not being reported or when there were no studies with the discrepancy.

<sup>a</sup>Trials with at least one discrepancy between the primary or secondary outcomes reported in the published article and those registered in the trial registry.

<sup>b</sup>Trials with at least one discrepancy between the primary outcome reported in the published article and that registered in the trial registry.

<sup>c</sup>Trials with at least one discrepancy among any of the secondary outcomes reported in the published article and those registered in the trial registry.

\* While the wording is slightly different, both rows are equivalent, as both show a registered primary outcome being reported as a secondary outcome.

\*\* While the wording is slightly different, both rows are equivalent, as both show a registered secondary outcome being reported as a primary outcome.

	2007 (n=2)		2010 (n=8)		2013 (n=33)		2015 (n=59)	
	A	B	A	B	A	B	A	B
Trials with any primary or secondary outcome discrepancy <sup>a</sup>	2 (100%)	1.00	8 (100%)	0.50	30 (91%)	0.63	54 (92%)	0.59
<b>Trials with primary outcome discrepancies <sup>b</sup></b>	1 (50%)	1.00	6 (75%)	0.67	14 (42%)	0.79	25 (42%)	0.56
Registered primary outcome not reported as primary outcome	1 (50%)	0.00	3 (38%)	0.33	7 (21%)	0.43	14 (24%)	0.00
Registered primary outcome not reported at all	1 (50%)	N/A	1 (13%)	N/A	3 (9%)	N/A	6 (10%)	N/A
Registered primary outcome reported as secondary outcome *	0 (0%)	N/A	2 (25%)	0.50	5 (15%)	0.60	9 (15%)	0.00
Reported primary outcome not registered as primary outcome	1 (50%)	1.00	4 (50%)	0.75	7 (21%)	0.86	17 (29%)	0.71
Reported primary outcome not registered	1 (50%)	1.00	4 (50%)	0.75	5 (15%)	0.80	13 (22%)	0.62
Reported primary outcome registered as secondary outcome **	0 (0%)	N/A	0 (0%)	N/A	2 (6%)	1.00	4 (7%)	1.00
Timing of reported primary outcome different from primary outcome registered	0 (0%)	N/A	0 (0%)	N/A	5 (15%)	0.60	8 (14%)	0.38
<b>Trials with secondary outcome discrepancies <sup>c</sup></b>	1 (50%)	1.00	7 (88%)	0.43	30 (91%)	0.50	53 (90%)	0.53
Registered secondary outcome not reported as secondary outcome	0 (0%)	N/A	2 (25%)	0	9 (27%)	0.22	17 (29%)	0.24
Registered secondary outcome not reported at all	0 (0%)	N/A	2 (25%)	N/A	7 (21%)	N/A	13 (22%)	N/A
Registered secondary outcome reported as primary outcome **	0 (0%)	N/A	0 (0%)	N/A	2 (6%)	1.00	4 (7%)	1.00
Reported secondary outcome not registered as secondary outcome	1 (50%)	1.00	7 (88%)	0.43	26 (79%)	0.50	51 (86%)	0.45
Reported secondary outcome not registered	1 (50%)	1.00	7 (88%)	0.29	25 (76%)	0.44	51 (86%)	0.45
Reported secondary outcome registered as primary outcome *	0 (0%)	N/A	2 (25%)	0.50	5 (15%)	0.60	9 (15%)	0.00
Timing of reported secondary outcome different from secondary outcome registered	0 (0%)	N/A	1 (13%)	0.00	10 (30%)	0.30	9 (15%)	0.33



### 3.4 Discussion

We found at least two major problems in the anesthesia literature — low rates of adequate RCT registration and high rates of discrepancies between the outcomes *registered* at a trials registry and the outcomes actually *reported* in a journal. While the proportion of adequately registered trials trended upward for all six journals over time, the overall percentage of 38% in 2015 is still inadequate. However, not all journals had the same proportion of adequately registered trials.

Although transparency of research conduct is the primary reason for an investigator to register their RCT, for some researchers, journal-specific policies may have impacted their registration decision. *Anesthesiology*, *British Journal of Anaesthesia (BJA)*, and *European Journal of Anaesthesiology (EJA)* described their mandatory trial registration policy in their Instructions to Authors sections starting in 2013,<sup>14</sup> 2009,<sup>15</sup> and 2015,<sup>16</sup> respectively. *Anaesthesia* and *Anesthesia & Analgesia (A&A)* both confirmed a mandatory trial registration policy starting in 2014 after emails to the editorial office,<sup>9,10</sup> while the *Canadian Journal of Anesthesia (CJA)* still only strongly recommends trial registration in its Instructions to Authors section.<sup>17</sup> Interestingly, despite the lack of a policy mandating RCT registration, the *CJA* did not have the lowest proportions of adequately registered trials — indicating that written journal policy is not necessarily the primary determinant of acceptance of inadequately registered RCTs.

Journals requiring RCTs to be registered in a public trials registry should, by definition, have a near 100% proportion of adequately registered trials but the results of this study demonstrate otherwise. We suspect this is because not all of the details of trial

registration are systematically being checked by journals — particularly whether trials are registered before the first study participant is enrolled. Also known as retrospective registration, registering trials after patient enrollment may not help to reduce publication and selective reporting bias.<sup>18</sup> The largest problem with retrospective registration is that it makes it appear as though there is complete alignment between registration and reporting, giving the impression of complying with a journal's registration policy without actually doing so. We therefore recommend heightened vigilance at all journals to recognize this problem.

We compared the results of our study with the De Oliveira et al. study that also examined the proportion of adequately registered trials in *Anaesthesia*, *Anesthesia & Analgesia*, *Anesthesiology*, and *British Journal of Anaesthesia* in 2013.<sup>7</sup> Our study expanded on this work in two major ways: first, by examining only adequately registered RCTs (defined as a trial which was registered in a publicly available trial registry prior to the first patient being enrolled and a registry entry containing a clearly defined primary outcome), and second, by performing a longitudinal examination over four different years rather than just a single year. Overall, the De Oliveira et al. study found that 35% of published clinical trials in anesthesia literature were prospectively registered,<sup>7</sup> about two times larger than the 19% we found (for those journals in 2013) in our study. For each individual journal, we found fewer RCTs published and fewer adequately registered RCTs than did De Oliveira et al.<sup>7</sup> This difference is likely due to different inclusion/exclusion criteria since De Oliveira et al. identified trials through a PubMed search<sup>7</sup> while we used hand-searching of each journal's Table of Contents and inclusion criteria that were more stringent (e.g., exclusion of dose-finding or dose-response studies

which were not designed to test a clinical intervention). As noted above, our stricter definition of what constituted an adequately registered RCT may have also contributed to the discrepancy since the De Oliveira et al. study only required adequately registered trials to be prospectively registered.<sup>7</sup>

Even within adequately registered RCTs, significant discrepancies between registered and published outcomes can affect the quality of published trials.<sup>2</sup> In some cases, this outcome switching is warranted and authors should explain the reasons for the changes, but outcome switching may also be used to spin and distort the data to have more favourable conclusions.<sup>6,19</sup> While the number of adequately registered trials with at least one primary or secondary outcome discrepancy in 2015 showed a decrease from 2007 and 2010, 92% of adequately registered trials published in 2015 still had at least one primary or secondary outcome discrepancy. For adequately registered trials published in 2015, 90% of trials had at least one secondary outcome discrepancy, with most of these discrepancies due to the reported secondary outcome not being registered. Historically, serendipitous findings (which are often secondary outcomes) have been responsible for many important scientific discoveries, so researchers should feel free to report these unexpected findings. However, to maintain the important scientific tradition of being open and transparent about a study's methods, these unregistered outcomes should be labelled as "post-hoc" or "exploratory".

The Centre for Evidence-Based Medicine Outcome Monitoring (COMPare) project also found that in the top five medical journals, each trial silently added, on average, 5.3 new outcomes.<sup>5</sup> In the current study, 42% of RCTs published in 2015 had at least one primary outcome discrepancy, which is concerning since the primary outcome is the objective of

the trial and distortion of the primary outcome will affect every aspect of the trial. No specific type of primary outcome discrepancy accounted for most of the discrepancies, but when the reported primary outcome was not registered as the primary outcome, there was a higher chance that the discrepancy favoured statistical significance. While a systematic review of all studies assessing outcome discrepancies found a median percentage of primary outcome discrepancies of 32% (IQR: 25% – 45%) and a median percentage of secondary outcome discrepancies of 54% (IQR: 33% – 68%),<sup>6</sup> the current study found 45% and 89%, respectively, suggesting either that RCTs in anesthesiology journals may have more outcome discrepancies than general and other specialty journals or that different studies may use varying methodologies. Comparing again to the De Oliveira et al. study which focused on anesthesiology journals, the authors found that 43% of trials had a primary outcome discrepancy and 79% of trials had a secondary outcome discrepancy,<sup>7</sup> similar to the results of our study.

With each successive year, more adequately registered RCTs have registered and reported more than one primary outcome. The median number of registered and reported secondary outcomes fluctuates between years but the number of reported secondary outcomes is always higher than the number of registered secondary outcomes. This corresponds to the high percentage of secondary outcome discrepancies observed, especially with reported secondary outcomes that were not registered. To maintain the effectiveness of trial registration, investigators should ensure that all outcomes are included in the registry entry, even secondary outcomes. The number of participants divided by the number of secondary outcomes reported is consistent around 6 except in 2007, though this may be influenced by the low number of adequately registered trials

published in 2007. When more variables are measured without increasing the sample size, there is a higher probability of getting a statistically significant result by chance alone.<sup>20</sup> There is considerable variation in the number of secondary outcomes reported but investigators should avoid measuring too many secondary outcomes if the sample size is small. It is concerning that, in 2015, 25% of adequately registered RCTs had two or fewer study participants per secondary outcome analyzed.

An unexpected finding of our study (an unplanned, *post-hoc* secondary outcome) that is deeply concerning is the large decline in the absolute number of RCTs being reported in the anesthesia literature. Potential reasons for this include a lack of grant funding to support research, the ever-increasing and frustrating bureaucratic “red tape” involved in conducting research, a shift in RCT publication from the highest impact factor journals to lower impact factor journals, or researchers eschewing randomized study designs in favour of observational designs (which may also involve less regulation and be less onerous to conduct).

Limitations of our study include the arbitrary selection of years examined, the number of reviewers, the exclusion of EUCTR registered RCTs, and the presence of non-responding corresponding authors. In this study, only the years of 2007, 2010, 2013, and 2015 were included with no data pertaining to RCTs published in unmeasured years. While two reviewers were used to screen for eligible RCTs, only one reviewer extracted data from each RCT, potentially increasing the amount of error. To reduce this error, reviewers were trained prior to data extraction with the same guidelines and a second reviewer was consulted if any variable was unclear.

Another limitation was that all RCTs registered in the EUCTR were considered to be unregistered, and subsequently excluded from analyses concerning adequately registered RCTs. Since the EUCTR was only available for public access starting in 2011,<sup>13</sup> RCTs in 2007 and 2010 using that registry would not have been publicly available. To maintain consistency, RCTs in subsequent years were also considered to be unregistered as many studies published in 2013 and 2015 would likely have commenced their trial planning and/or recruitment stages before 2011. For this reason, approximately six and four RCTs were excluded in 2013 and 2015, respectively. Since articles with authors that were unable to be contacted after two attempts were considered unregistered, the proportion of adequately registered trials may have been underestimated if nonresponding authors had actually registered their trials. There is potential for bias as authors with recently published studies may be more easily contacted. However, there is only a small chance of this misclassification as publications and clinical trial registries were also searched for relevant registration information.

In conclusion, an alarming proportion of published randomized clinical trials in the anesthesia literature are still inadequately registered despite long-standing international guidelines recommending trial registration. We suggest that journal editors need to be more vigilant when enforcing their registration policies. Peer reviewers should also consider it a responsibility to reconcile the registered and published outcomes when reviewing manuscripts. Unfortunately, even amongst adequately registered trials, important discrepancies still sometimes exist between the registered outcomes and the reported outcomes. Primary outcome switching can systematically distort the anesthesia literature and potentially negatively affect patient care if clinicians base their practice on

distorted evidence. Continued longitudinal research of the incidence and severity of these problems is warranted.

### 3.5 Literature Cited

1. Guyatt G, Rennie D, Meade MO, Cook DJ. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd ed. United States of America: McGraw-Hill Education; 2015.
2. Jones PM. Publication bias in the anesthesiology literature: shifting the focus from the 'positive' to the 'truth'. *Can J Anesth Can d'anesthésie*. 2016;63(6):658-663. doi:10.1007/s12630-016-0632-z.
3. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors. *N Engl J Med*. 2004;351(12):1250-1251. doi:10.1056/NEJMe048225.
4. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials. *JAMA*. 2004;291(20):2457-2465. doi:10.1001/jama.291.20.2457.
5. Goldacre B, Drysdale H, Powell-Smith A, et al. The COMPare Trials Project. 2016. [www.COMParE-trials.org](http://www.COMParE-trials.org). Accessed December 1, 2016.
6. Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Med*. 2015;13:282. doi:10.1186/s12916-015-0520-3.
7. De Oliveira GS, Jung MJ, McCarthy RJ. Discrepancies Between Randomized Controlled Trial Registry Entries and Content of Corresponding Manuscripts Reported in Anesthesiology Journals. *Anesth Analg*. 2015;121(4):1030-1033. doi:10.1213/ANE.0000000000000824.

8. Editorial Office. Instructions for Authors. *Anesthesiology*. 2015.  
<http://anesthesiology.pubs.asahq.org/public/instructionsforauthors.aspx>. Accessed August 1, 2016.
9. Jarvis-Atton S. Trial Registration - Anaesthesia. 2016.
10. Williams-Klamborowski C. Trial Registration - Anesthesia & Analgesia. 2016.
11. Chow J, Jones P. Identifying areas to improve the quality of reporting in randomized clinical trials published in anesthesiology journals: A study protocol for a series of literature surveys assessing quality of trial registration, adherence to abstract reporting guidelines. 2016. doi:10.6084/m9.figshare.4490582.v1.
12. Thomson Reuters. Journal Citation Reports - Anesthesiology. ISI Web Knowl. 2016.  
<http://admin-apps.webofknowledge.com/JCR/JCR>. Accessed August 15, 2016.
13. Benstetter M, Haubenreisser S. EU Clinical Trials Register goes live. Eur Med Agency. 2011.  
[http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Press\\_release/2011/03/WC500104032.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2011/03/WC500104032.pdf). Accessed August 1, 2016.
14. Anesthesiology Editorial Office. Instructions for Authors. *Anesthesiology*.  
<http://anesthesiology.pubs.asahq.org/public/instructionsforauthors.aspx>. Accessed August 15, 2016.
15. Mahajan RP, British Journal of Anaesthesia Editorial Team. Instructions to Authors. *Br J Anaesth*. 2016.  
[http://www.oxfordjournals.org/our\\_journals/bjaint/for\\_authors/general.html](http://www.oxfordjournals.org/our_journals/bjaint/for_authors/general.html). Accessed August 15, 2016.
16. European Journal of Anaesthesiology Editorial Office. Guidance for Authors on the Preparation and Submission of Manuscripts to the European Journal of Anaesthesiology. *Eur J Anaesthesiol*. 2016. <http://edmgr.ovid.com/eja/accounts/ifauth.htm>. Accessed August 15, 2016.



17. Backman S, Baker A, Beattie S, et al. 2011 Canadian Journal of Anesthesia Guide for Authors. *Can J Anesth Can d'anesthésie*. 2011;58(7):668-696. doi:10.1007/s12630-011-9516-4.
18. Harriman SL, Patel J. When are clinical trials registered? An analysis of prospective versus retrospective registration. *Trials*. 2016;17(1):187. doi:10.1186/s13063-016-1310-8.
19. 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(20):2058-2064. doi:10.1001/jama.2010.651.
20. Friedman LM, Furberg CD, DeMets DL, Reboussin DM, Granger CB. *Fundamentals of Clinical Trials*. 5th ed. Springer International Publishing; 2015. doi:10.1007/978-3-319-18539-2.

## Chapter 4

### 4 The degree of adherence to CONSORT reporting guidelines for the abstracts of randomized clinical trials published in anesthesiology journals

Jeffrey T. Y. Chow<sup>1,3</sup>, Timothy P. Turkstra<sup>2,3</sup>, Edmund Yim<sup>3</sup>, Philip M. Jones<sup>1,2,3</sup>

<sup>1</sup>Department of Epidemiology & Biostatistics

<sup>2</sup>Department of Anesthesia & Perioperative Medicine

<sup>3</sup>Schulich School of Medicine & Dentistry, The University of Western Ontario

---

---

## 4.1 Introduction

Randomized clinical trials (RCTs) are generally considered to be at the top of the evidence hierarchy when determining the benefits and harms of therapeutic interventions.<sup>1</sup> Abstracts are concise summaries of the entire RCT and are often used by readers to determine whether reading the full-text manuscript is warranted.<sup>2</sup> Interestingly, the *BMJ* found that it was valid for editors to reject a manuscript submission based only on reading the abstract, with an estimated 15-25% of papers rejected on that initial abstract read-through.<sup>3</sup> In the study, editors would read the abstract only, record their initial decision (or inability to make a decision), and then read the entire manuscript, processing it as usual.<sup>3</sup> For manuscripts that the *BMJ* editors determined could be rejected after only reading the abstract, the final decision after the entire process was still rejection.<sup>3</sup> Abstracts are also important tools for clinical decision-making because there may be financial, information technology, time, or language barriers that prevent or reduce access to all full-text manuscripts.<sup>2,4,5</sup> Indeed, just over 50% of studies published in 2009 and indexed on PubMed were found to be open access, leaving almost half of the studies in the biomedical field with the full-text only accessible by subscription.<sup>6</sup> Together, these factors imply that many clinicians may use the abstracts of clinical trials to guide clinical decision-making without ever having read the full text of a particular article. Therefore, it is clear that the quality of abstract reporting is of paramount importance.

To improve the quality of reporting RCT abstracts, an extension of the Consolidated Standards of Reporting Trials (CONSORT) statement was developed and published in 2008.<sup>7,8</sup> This *CONSORT for Abstracts* statement specified a minimum set of criteria that

authors should include in the abstract of a RCT, ensuring that readers have the necessary detail and clarity to assess a RCT's validity and applicability.<sup>7</sup>

A study comparing 2006 (pre-CONSORT) and 2009 (post-CONSORT) abstracts published in four high profile anesthesiology journals found some improvements in areas such as blinding and harms, with an overall 2.5% increase in the percentage of recommended criteria included.<sup>9</sup> However, on average, there were still fewer than a third of the checklist items reported in each RCT published in 2009.<sup>9</sup> For RCTs published in four high-impact general medical journals in 2010, some items were adequately reported while some items such as trial design, randomization, blinding, harms, conclusions, and funding were reported in fewer than 50% of RCTs.<sup>10</sup>

No studies in the anesthesiology literature have examined whether the items omitted from the abstract are present in the body of the manuscript. If the items are not reported in both the abstract and the full-text, the underlying problem may be inadequate trial reporting and not inadequate abstract reporting. If items are not reported in the abstract but are present in the full-text, the CONSORT for Abstracts guidelines are not being adhered to even though the information is available in the full-text. This lack of adherence may intentionally or unintentionally cause abstract readers to have a distorted impression of the RCT's validity or applicability. "Spin" and misrepresentation of the salient trial details in the abstract have an impact on clinicians' interpretations of an RCT, where a small modification in an abstract's focus can make clinicians more likely to consider a treatment beneficial despite the primary outcome being statistically nonsignificant.<sup>11</sup>

The objective of this study was to determine the degree of adherence to CONSORT's abstract reporting guidelines for RCTs published in anesthesiology journals. For checklist items not reported in the abstract, this study also identified whether they were reported in the full-text manuscript to determine the potential for inadequate abstract reporting distorting or "spinning" a RCT's findings. The top six general (non-pain-centric) anesthesiology journals as determined by impact factor were included to further examine between-journal differences. Our hypothesis was that newer trials would adhere to the CONSORT for Abstracts guidelines more completely compared to older trials; since journals specify different formats and restrictions for reporting abstracts, they would also be an important predictor of adherence. This study updated the literature on anesthesiology RCT abstract reporting by using RCTs published in 2016 and comparing them with RCTs published in 2010 to determine whether there was an improvement over time.

## 4.2 Methods

An electronic database was set up to identify and retrieve all RCTs published in the top six general anesthesiology journals as determined by impact factor in 2016: *Anaesthesia*, *Anesthesia & Analgesia (A&A)*, *Anesthesiology*, *British Journal of Anaesthesia (BJA)*, *Canadian Journal of Anesthesia (CJA)*, and *European Journal of Anaesthesiology (EJA)*.<sup>12</sup> Since the CONSORT for Abstracts statement was published in 2008,<sup>7,8</sup> 2010 was selected as the first year for analysis. To show a longitudinal trend and provide the most current results, RCTs were only included if published in 2010 or 2016.

Using the websites for each included journal, the Table of Contents' for each month, excluding supplemental issues, were independently hand-searched by two reviewers to identify eligible RCTs. Disagreements were assessed by a third reviewer with a decision made through consensus by all three reviewers. A RCT was defined as a prospective study assessing randomly allocated health care interventions in human participants, with random allocation of participants to study arms clearly reported (i.e. the use of the words "random", "randomized", or "randomised"). As a result, observational studies (e.g. case-control or cohort studies), learning curve studies, cadaver studies, cost-effectiveness studies, dose-finding or dose-response studies which were not designed to test a clinical intervention, diagnostic test accuracy studies, meta-analyses, editorials, narrative reviews, animal studies, manikin studies, and re-analyses of previously published studies were all excluded. RCTs published in the correspondence or letters to the editor sections were also excluded. If two phases were reported in one study, the phase with the highest number of participants was used for data extraction. RCTs identified as meeting the inclusion/exclusion criteria were retrieved and stored in a prospectively created electronic database. More details regarding methodology can be found in our publicly available study protocol.<sup>13</sup>

Data extraction was performed for all RCTs using the created electronic database. Data extraction for included RCTs were performed by one of four reviewers. All reviewers were trained together to ensure consistency of data extraction. Ambiguous data points were assessed by a second reviewer with a decision made through consensus. General study characteristics were extracted including journal, year, multicentricity, study design, sample size, and trial registration status. The number randomized was used for the sample

size. Adequate trial registration was defined as the trial being registered in a publicly available trial registry before the first participant was enrolled with a clearly defined primary outcome. To identify a RCT's registration status, the full-text was first searched for trial registration information. If none was located, searches were performed at the following websites: [www.clinicaltrials.gov](http://www.clinicaltrials.gov), the International Standard Randomized Controlled Trial Number Register (ISRCTN), and the World Health Organization (WHO) Clinical Trials Search Portal (which simultaneously searches 16 clinical trial registries). If no information was found during the trial registry search, two attempts (separated by a week) were made to contact the corresponding author using a standardized email asking whether or not the trial was registered, what the registration number was, and which registry was used (Appendix B). The RCT was assumed to be not registered if no registration information was identified using the steps above. Trials that were only registered at EudraCT (<https://www.clinicaltrialsregister.eu>) were considered not adequately registered since the information contained in that database was not publicly available until the release of the European Union Clinical Trials Register (EUCTR) in 2011.<sup>14</sup>

Adequate trial registration was determined using the date of trial registration and date of first patient enrollment. However, if the trial registry indicated that a RCT was retrospectively registered, this was accepted at face value. If the trial registration date occurred before the first patient enrollment date, the RCT was considered to be adequately registered. The date of trial registration was determined by the date submitted for trial registration as specified in the trial registry. The date of first patient enrollment was determined by the date specified in the full-text publication. If no date was provided

in the publication, the study start date as specified in the clinical trial registry was used instead. If no date was provided in the publication or trial registry, the RCT was considered inadequately registered. For study start dates that only included the month and year, RCTs were considered to be adequately registered if the month was during or before the trial registration date.

Since no validated scoring system for abstract quality exists, the 16 checklist items from the CONSORT for Abstracts statement<sup>7,8</sup> were used to create a convenience score as a proxy for RCT abstract reporting quality. A brief description of each item was adapted from the CONSORT for Abstracts statement explanations and presented alongside each checklist item during data extraction to improve quality (Appendix D). Each criterion was measured as being reported in abstract, not reported in abstract but reported in full-text manuscript, or not reported in abstract or full-text manuscript, then used to calculate summary measures such as number of trials with the checklist item reported and the number of trials with the checklist item reported in the full-text only. An overall score was calculated by counting the number of criteria reported in each RCT's abstract. Since the CONSORT for Abstracts statement only focuses on primary outcomes, the number of secondary outcomes reported in the abstract and the number of statistically significant secondary outcomes reported in the abstract were also extracted to calculate a percentage of secondary outcomes reported in the abstract that were statistically significant.

Descriptive statistics were used to present the raw number and percentage of checklist items reported for included RCTs. To assess differences between years, the Pearson chi-square test was used for categorical variables and the Wilcoxon rank-sum (Mann-Whitney) test was used for continuous variables. The Kruskal-Wallis test was used to



assess differences between journals stratified for each year of publication. To improve clarity in the results, the overall CONSORT for Abstracts score is reported as the *abstract-only score* and the number of full-text only (reported in the manuscript but not in the abstract) items is reported as the *full-text-only score*. Stata 13 statistical software (StataCorp LLC, College Park, Texas) was used for all analyses.  $P < 0.05$  was considered statistically significant.

### 4.3 Results

Table 4-1 describes the general characteristics for the 395 included RCTs, with 219 published in 2010 and 176 published in 2016. While the total number of RCTs decreased from 2010 to 2016, *Anesthesiology* and *CJA* had an absolute increase in the number of included RCTs. Most RCTs were conducted at a single-centre and had a parallel group superiority design. From 2010 to 2016, the number of single-centre RCTs decreased from 94% to 86% (Pearson chi-square:  $p=0.02$ ), the number of adequately registered RCTs increased from 4% to 39% (Pearson chi-square:  $p < 0.001$ ), and the median sample size increased from 68 to 86 (Wilcoxon rank-sum:  $p=0.03$ ).

**Table 4-1: Characteristics of RCTs included for analysis**

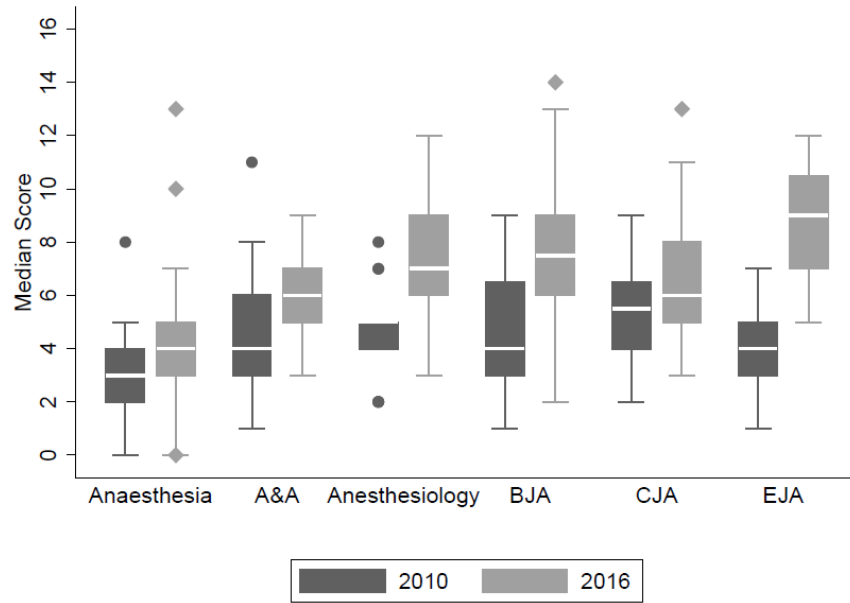
Data presented are the number of trials (%) unless otherwise stated. Percentages may not add up to 100% due to rounding. IQR means interquartile range.

<sup>1</sup> Multi-centre is defined as participants recruited from more than one institution or clinic

<sup>2</sup> Adequate registration is defined as the trial being registered before the first participant was enrolled and that a primary outcome was clearly defined in the registry.

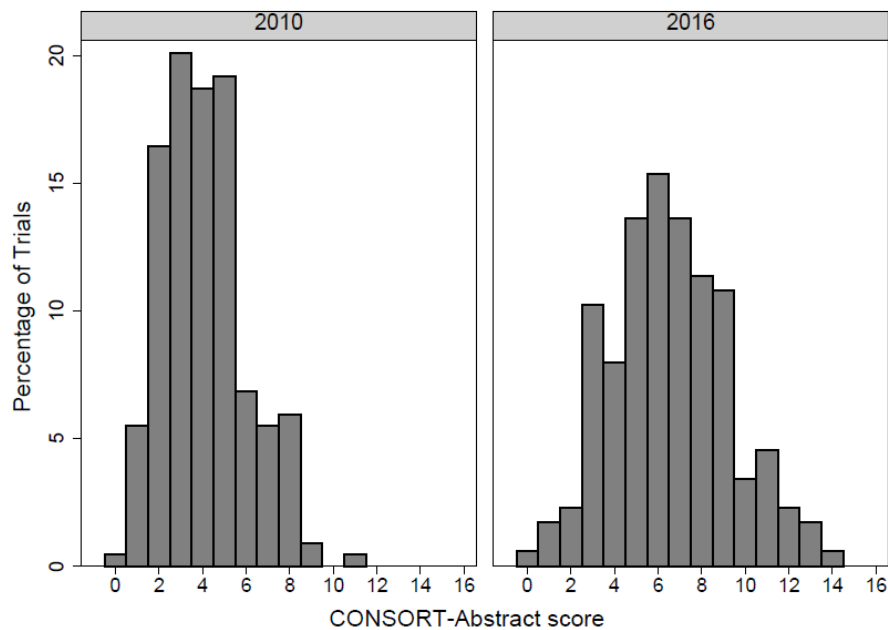
	<b>2010 (n=219)</b>	<b>2016 (n=176)</b>
<b>Journal</b>		
Anaesthesia	43 (20%)	38 (22%)
Anesthesia & Analgesia	57 (26%)	33 (19%)
Anesthesiology	23 (11%)	26 (15%)
British Journal of Anaesthesia	36 (16%)	34 (19%)
Canadian Journal of Anesthesia	16 (7%)	21 (12%)
European Journal of Anaesthesiology	44 (20%)	24 (14%)
<b>Multicentricity<sup>1</sup></b>		
Single-centre	205 (94%)	152 (86%)
Multi-centre	14 (6%)	24 (14%)
<b>Study Design</b>		
Parallel (superiority)	188 (86%)	138 (78%)
Parallel (non-inferiority/equivalence)	8 (4%)	16 (9%)
Crossover	15 (7%)	19 (11%)
Other	8 (4%)	3 (2%)
<b>Trial Registration<sup>2</sup></b>		
Adequately Registered	8 (4%)	68 (39%)
Inadequately Registered	211 (96%)	108 (61%)
<b>Median Sample Size (IQR)</b>	68 (44–120)	86 (50–139)

While the overall CONSORT for Abstracts abstract-only score increased from 2010 to 2016 for each journal, the median score was below 10 (out of a possible 16) for every journal (Figure 4-1) in both years analyzed. The distributions of the abstract-only scores were statistically significantly different between years; the median abstract-only score increased from 4 points [interquartile range (IQR): 3 to 5] in 2010 to 6 points [IQR: 5 to 8] in 2016 (Wilcoxon rank-sum:  $p < 0.0001$ ) (Figure 4-2; Table 4-2). When analyses were stratified by year, there was a statistically significant difference for the abstract-only score between the different journals of publication (Kruskal-Wallis:  $p < 0.0001$  for 2010 and  $p < 0.0001$  for 2016) (Figure 4-1).



**Figure 4-1: Distribution of overall CONSORT for Abstracts abstract-only score by journal over time**

White line represents median; box represents 25th and 75th percentiles; each whisker represents 1.5 times interquartile range; dots represent outside values



**Figure 4-2: Distributions of overall CONSORT for Abstracts abstract-only score in 2010 and 2016**

Wilcoxon Rank-Sum test for difference in distribution:  $p < 0.0001$

**Table 4-2: Adherence to CONSORT-Abstract reporting items in RCTs**

(A) Number of trials (%) with the checklist item reported

(B) Of trials with the checklist item not reported in the abstract, the number (%) of these trials with the checklist item reported in the manuscript

<sup>1</sup> Abstract-only score: Number of CONSORT for Abstracts criteria reported in the abstract(Wilcoxon rank-sum test for difference in distribution between years:  $p < 0.0001$ )<sup>2</sup> Full-text-only score: Number of CONSORT for Abstracts criteria reported only in the full-text manuscript(Wilcoxon rank-sum test for difference in distribution between years:  $p = 0.02$ )

	2010		2016	
	A (n=219)	B	A (n=176)	B
<b>Title</b>				
"Randomized" in title	56 (26%)	N/A	131 (74%)	N/A
"Randomized" in abstract	145 (66%)	18/18 (100%)	40 (23%)	5/5 (100%)
<b>Trial Design</b>	45 (21%)	128/174 (74%)	57 (32%)	81/119 (68%)
<b>Methods</b>				
Participants	10 (5%)	91/209 (44%)	74 (42%)	99/102 (97%)
Interventions	156 (71%)	62/63 (98%)	106 (60%)	70/70 (100%)
Objective	54 (25%)	70/165 (42%)	63 (36%)	86/113 (76%)
Outcome	38 (17%)	74/181 (41%)	101 (57%)	49/75 (65%)
Randomization	3 (1%)	150/216 (69%)	3 (2%)	152/173 (88%)
Blinding (masking)	82 (37%)	107/137 (78%)	66 (38%)	93/110 (85%)
<b>Results</b>				
Numbers randomized	81 (37%)	101/138 (73%)	74 (42%)	102/102 (100%)
Recruitment	5 (2%)	182/214 (85%)	33 (19%)	140/143 (98%)
Numbers analyzed	30 (14%)	169/189 (89%)	37 (21%)	139/139 (100%)
Outcome	13 (6%)	30/206 (15%)	71 (40%)	25/105 (24%)
Harms	108 (49%)	75/111 (68%)	104 (59%)	46/72 (64%)
<b>Conclusions</b>	203 (93%)	11/16 (69%)	161 (91%)	14/15 (93%)
<b>Trial registration</b>	13 (6%)	43/206 (21%)	70 (40%)	75/106 (71%)
<b>Funding</b>	0 (0%)	153/219 (70%)	0 (0%)	168/176 (95%)
<b>Median Score (IQR)</b>	4 (3–5) <sup>1</sup>	7 (6–9) <sup>2</sup>	6 (5–8) <sup>1</sup>	8 (7–9) <sup>2</sup>

The median full-text-only score increased from 7 points [IQR: 6 to 9] in 2010 to 8 points [IQR: 7 to 9] in 2016 (Wilcoxon rank-sum:  $p = 0.02$ ) (Table 4-2). There was also a statistically significant difference for the full-text-only score between the different journals of publication (Kruskal-Wallis:  $p < 0.0001$  for 2010 and  $p < 0.0001$  for 2016). A breakdown of the number of trials with the checklist item reported in the abstract, not reported in abstract but reported in manuscript, and not reported in abstract or manuscript is presented in Table 4-2. Appendix E presents the same information for RCTs published in 2016 stratified by each journal. Without adjusting for multiple comparisons, Pearson's

chi-squared tests for differences between journals suggest that the journal had a statistically significant effect on adherence to the CONSORT for Abstracts guidelines for every criterion except for randomization, numbers randomized, and numbers analyzed (Appendix E). The mean percentage of statistically significant secondary outcomes in the abstract decreased slightly from 60% in 2010 to 55% in 2016 though the standard deviations in both years were large (Table 4-3).

**Table 4-3: Characteristics of secondary outcomes reported in the abstract**

IQR means interquartile range and SD means standard deviation

<sup>1</sup> Using the secondary outcomes reported in the abstract for each trial, the number of statistically significant secondary outcomes was divided by the number of secondary outcomes to get the percentage. The mean and standard deviation were calculated using the percentages calculated for each trial.

	<b>2010</b>	<b>2016</b>
Median number (IQR) of secondary outcomes reported in abstract	3 (1–4)	3 (2–4)
Median number (IQR) of statistically significant secondary outcomes reported in abstract	1 (0–3)	1 (0–2)
Mean percentage (SD) of secondary outcomes reported in the abstract that were statistically significant <sup>1</sup>	60% (37%)	55% (39%)

## 4.4 Discussion

The results of this study suggest that many anesthesiology RCTs are not following the CONSORT for Abstracts guidelines when publishing their final manuscripts. A previous study examining adherence to the CONSORT for Abstracts guidelines for four general anesthesiology journals found the mean proportion of items reported to be 26.6% [95% CI: 25.5 to 28.0] in 2006 and 29.0% [95% CI: 27.2 to 30.2] in 2009.<sup>9</sup> The median number of items reported in this study was four in 2010 and six in 2016, which corresponds to 25% and 38%, respectively. Since the CONSORT for Abstracts guidelines were published in 2008, this increase may be due to more authors and editors being aware of them. However, there are still many criteria that are consistently not being reported by

authors. Since the results suggest that there are differences between journals, journal-specific policies may have some effect. Indeed, they may be the principal driver of the differences observed.

For each of the six included anesthesiology journals, the instructions to authors do not specifically address adherence to the CONSORT for Abstracts statement. However, all six journals require or recommend compliance with the general CONSORT guidelines, which includes adhering to the CONSORT for Abstracts guidelines.<sup>15-20</sup> In terms of abstract-specific instructions, *Anaesthesia* requires an unstructured abstract no more than 250 words,<sup>15</sup> *Anesthesia & Analgesia* requires a structured abstract no more than 400 words,<sup>19</sup> *Anesthesiology* requires a structured abstract no more than 250 words,<sup>16</sup> *British Journal of Anaesthesia* requires a structured abstract no more than 250 words,<sup>17</sup> *Canadian Journal of Anesthesia* requires a structured abstract no more than 250 words,<sup>18</sup> and *European Journal of Anaesthesiology* requires a structured abstract no more than 300 words.<sup>20</sup> Though there are strict word limits that authors cannot exceed in their abstract, 250 to 300 words is sufficient to address all items in the CONSORT for Abstracts guidelines.<sup>7,8</sup> In addition, MEDLINE no longer truncates abstracts at a specific word count and instead allows up to 10,000 characters in an abstract,<sup>21</sup> questioning whether journals should be imposing a specific word limit at all.

The proportions of RCTs that met each criterion in 2010 were similar to a previous study, with many RCTs not meeting the CONSORT for Abstracts guidelines.<sup>9</sup> Since that study found that more criteria were met after the CONSORT for Abstracts guidelines were published,<sup>9</sup> it would be expected that this study would see continued increases in adherence. Other than the interventions and objectives criteria, the percentage of RCTs in

2016 that reported each criterion in the abstract increased from a previous study in anesthesiology journals.<sup>9</sup>

Since the first thing readers see is the title, it is important to identify that the trial is randomized in the title. In Table 4-2, this CONSORT for Abstracts criterion was separated into three sections (reported in the title, reported in the abstract, or reported in the full-text manuscript) to show where authors were first describing that their trial was randomized. From 2010 to 2016, the percentage of RCTs with the word *randomized* in the title almost tripled, but almost a quarter of RCTs in 2016 still did not have the word *randomized* in the title. This is a very simple item that authors can easily achieve and editors can easily check for.

As expected, most RCTs included had a parallel groups and superiority design. However, many RCTs did not state whether they were conducting a superiority study in the abstract. Usually, the only way to determine whether the study was a superiority trial was from the objectives statement at the end of the Introduction or from the sample size calculation. In 2016, 68% of RCTs did not describe the trial design in the abstract, and of those, 68% had the trial design described in the full-text. The percentage of RCTs reporting the trial design in the abstract increased but among RCTs that did not report the trial design in the abstract, the percentage of RCTs reporting the trial design in the full-text decreased. A potential reason for this finding is that most trials conducted were superiority trials with parallel arms, with authors only specifying when they were using an alternative study design. With the increase in alternative study designs being used, authors should ensure that they specify the trial design in the abstract even when a parallel group, superiority trial is being conducted.

This lack of specificity in the abstract also applies to the objective criterion where many RCTs only stated that they compared one intervention to another intervention in the abstract, not to determine which one was better. A previous study in anesthesiology journals found that around 90% of RCTs met this criterion,<sup>9</sup> but this is likely due to a less stringent standard for having a clear objective. Another example of a lack of specificity is the Methods outcome criterion, where the primary reason for RCTs not meeting this criterion was not identifying which outcome was the primary outcome. The percentage of RCTs that describe their primary outcome increased from 2010 to 2016, but more authors need to ensure that they specify a primary outcome.

Many criteria (participants, interventions, randomization, blinding, numbers randomized, recruitment, and numbers analyzed) show similar trends in terms of reporting in the abstract compared to reporting in the full-text only, with a high percentage of RCTs reporting the information in the full-text only. The participants criterion in the CONSORT for Abstracts guidelines consists of two dimensions, the eligibility criteria for participants and the setting where the data were collected.<sup>7</sup> In 2016, only 42% of RCTs met this criterion, mostly due to the setting not being reported in the abstract. However, almost all RCTs that did not meet the criterion in the abstract had sufficient information in the full-text. As a result, abstract-only readers would not necessarily be able to know which patient population the results applied to.

For the interventions criterion, even though a majority of the RCTs described the interventions in the abstract, all RCTs that did not do so described the intervention in the full-text. Thus, for some RCTs (29% in 2010 and 40% in 2016), the abstract would be omitting important details even though the information was available in the full-text. The



randomization criterion shows a more extreme trend, where the way randomization was conducted was rarely described in the abstract, with less than 2% in both years, consistent with a previous study.<sup>9</sup> However, 88% of the trials in 2016 described this information in the full-text, suggesting that authors were choosing not to report this information in the abstract, not that the information was unavailable. Most authors just reported that the RCT was randomized in the abstract and not how participants were allocated to each group. Most RCTs in general medical journals also failed to meet this criterion, with fewer than 8% of RCTs in the *New England Journal of Medicine*, *Journal of the American Medical Association*, and *BMJ* reporting the method of random sequence generation in the abstract.<sup>10</sup> However, in the *Lancet*, 89% of RCTs reported the method of random sequence generation and 39% reported the method of allocation concealment in the abstract.<sup>10</sup>

For the blinding criterion, a similar trend was seen, but with a smaller gap since we considered RCTs as meeting the criterion even if they only used the terms *single* or *double blinding*. Even when using this relaxed criterion, only about 37% of RCTs reported the blinding in the abstract. This may be due to RCTs that did not use blinding in the RCT rarely mentioning a lack of blinding in the abstract. The recruitment criterion also followed this trend since RCTs that did not terminate early rarely reported the trial status. The proportion reporting in the full-text was almost 100% since we considered RCTs to meet this criterion if the number recruited met or exceeded the target sample size. Either RCTs should always report the trial status or this item should only be relevant when a trial is terminated early.

For the numbers randomized and numbers analyzed criteria, the percentage of RCTs adhering to the CONSORT for Abstracts guidelines increased by about 5% from 2010 to 2016. However, many RCTs did not have sufficient information in the abstract, often due to reporting a total sample size and not the number per group. The proportion of RCTs meeting the numbers analyzed criterion was lower than the numbers randomized criterion since many RCTs did not mention in the abstract if all patients were analyzed with no loss to follow-up. However, likely due to the increasingly widespread use of the CONSORT flow diagram,<sup>22,23</sup> all RCTs that did not meet these two criteria in the abstract reported the information in the full-text.

The percentage of RCTs that met the Results outcome criterion in the abstract increased from 6% in 2010 to 40% in 2016. While this is a large increase, more than half of the RCTs were still not adequately reporting this item. Most RCTs typically only reported a result for each group and a p-value, with few studies reporting an effect size and precision for that effect size. By comparison, in 2006, 62% of RCTs in major general medical journals reported the effect size and confidence interval in the abstract.<sup>24</sup> Of the RCTs in 2016 that did not meet the criterion in the abstract, only 24% met the criterion in the full-text. Since most anesthesiology RCTs were incorrectly reporting their results, the CONSORT for Abstracts guidelines was not able to be followed because the general CONSORT guidelines were not being followed. The harms criterion has a similar trend, though only a minority of RCTs were not reporting the harms. The percentage of RCTs reporting the harms in the abstract increased from 49% in 2010 to 59% in 2016 while the percentage of RCTs reporting the harms in the full-text decreased from 68% to 64%. This

suggests that the lack of harms reporting in the abstract is partially due to harms not being assessed in the RCT.

The conclusions criterion was the best reported, with over 90% of RCTs having a conclusion in the abstract. Corresponding with results from Chapter 3 suggesting an increase in trials being registered over time, the proportion of RCTs reporting trial registration in the abstract also increased. However, many RCTs still only reported the trial registration in the full-text and not the abstract. A distinct journal difference was observed in 2016, with almost no RCTs from *Anaesthesia*, *Anesthesia & Analgesia*, and *Anesthesiology* reporting the trial registration in the abstract even though 25%, 51%, and 65% of these RCTs had the trial registration reported in the full-text, respectively. In contrast, *British Journal of Anaesthesia*, *Canadian Journal of Anesthesia*, *European Journal of Anaesthesiology* reported the trial registration in the abstract for 53%, 54%, and 35% of the RCTs, respectively.

Just like in a previous study,<sup>9</sup> the funding source was never reported in the abstract. However, the percentage of RCTs that reported the funding source in the full-text increased from 70% in 2010 to 95% in 2016. Despite CONSORT for Abstracts guidelines, the funding source is always reported in the full-text only and not the abstract. Perhaps the authors of the CONSORT for Abstracts guidelines should consider removing this criterion or making it only applicable for conference abstracts, similar to the contact details for the corresponding author.

Potential limitations must be considered when interpreting the conclusions of this study. Since there are many factors contributing to whether an item is reported in the abstract,

such as space limitation, journal policy, author preference, and lack of available information, the exact reason for each RCT failing to meet a criterion is hard to determine even though there is an overall trend. Though two reviewers screened for eligible RCTs, each RCT had data extraction performed by one reviewer, potentially increasing the amount of error if the reviewers were inconsistent. To reduce the subjectivity of criteria and potential error, reviewers were trained prior to data extraction with the same guidelines and a second reviewer was consulted if any variable was unclear.

Another limitation is that cross-over trials are more likely to meet certain criteria since the same group is used for both treatment and control. For example, in the trial design criterion, cross-over trials do not need to specify that it is a parallel group design or in the numbers randomized criterion, the total number of participants is the same as the participants per arm. Finally, the overall adherence and full-text-only scores considered each CONSORT for Abstracts criterion as equally important, preventing an assessment of how much distortion the failure to report would cause a typical reader.

In conclusion, RCTs in the top six general anesthesiology journals have improved their adherence to the CONSORT for Abstracts guidelines from 2010 to 2016, but the level of adherence is still lacking. Even in 2016, around 75% of RCTs meet fewer than half of the 16 criteria with no RCTs reporting all 16 items in the abstract. Other than the Results outcome criterion, a majority of the RCTs have the information present in the full-text, but not reported in the abstract. This suggests that per the CONSORT for Abstracts guidelines, the abstracts for many anesthesiology RCTs are an incomplete summary of the entire manuscript. An alternative explanation is that some items are unrealistic for real-world application, with more investigation needed to determine which items would

provide the most impactful information. More research needs to be conducted to determine what tools can be used to improve the abstract reporting for anesthesiology RCTs and whether the CONSORT for Abstracts statement is appropriate.

## 4.5 Literature Cited

1. Guyatt G, Rennie D, Meade MO, Cook DJ. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd ed. United States of America: McGraw-Hill Education; 2015.
2. Sivendran S, Newport K, Horst M, Albert A, Galsky MD. Reporting quality of abstracts in phase III clinical trials of systemic therapy in metastatic solid malignancies. *Trials*. 2015;16(1):341. doi:10.1186/s13063-015-0885-9.
3. Groves T, Abbasi K. Screening research papers by reading abstracts. *BMJ*. 2004;329:470-471. doi:10.1136/bmj.329.7464.470.
4. Altwaairgi AK, Booth CM, Hopman WM, Baetz TD. Discordance Between Conclusions Stated in the Abstract and Conclusions in the Article: Analysis of Published Randomized Controlled Trials of Systemic Therapy in Lung Cancer. *J Clin Oncol*. 2012;30(28):3552-3557. doi:10.1200/JCO.2012.41.8319.
5. Wang L, Li Y, Li J, et al. Quality of reporting of trial abstracts needs to be improved: using the CONSORT for abstracts to assess the four leading Chinese medical journals of traditional Chinese medicine. *Trials*. 2010;11(1):75. doi:10.1186/1745-6215-11-75.
6. Kurata K, Morioka T, Yokoi K, Matsubayashi M. Remarkable Growth of Open Access in the Biomedical Field: Analysis of PubMed Articles from 2006 to 2010. Scherer RW, ed. *PLoS One*. 2013;8(5):e60925. doi:10.1371/journal.pone.0060925.
7. Hopewell S, Clarke M, Moher D, et al. CONSORT for Reporting Randomized Controlled Trials in Journal and Conference Abstracts: Explanation and Elaboration. von Elm E, ed. *PLoS Med*. 2008;5(1):e20. doi:10.1371/journal.pmed.0050020.

8. Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet*. 2008;371(9609):281-283. doi:10.1016/S0140-6736(07)61835-2.
9. Can OS, Yilmaz AA, Hasdogan M, et al. Has the quality of abstracts for randomised controlled trials improved since the release of Consolidated Standards of Reporting Trial guideline for abstract reporting? A survey of four high-profile anaesthesia journals. *Eur J Anaesthesiol*. 2011;28(7):485-492. doi:10.1097/EJA.0b013e32833fb96f.
10. Ghimire S, Kyung E, Kang W, Kim E. Assessment of adherence to the CONSORT statement for quality of reports on randomized controlled trial abstracts from four high-impact general medical journals. *Trials*. 2012;13:77. doi:10.1186/1745-6215-13-77.
11. Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of Spin in the Abstracts of Articles Reporting Results of Randomized Controlled Trials in the Field of Cancer: The SPIIN Randomized Controlled Trial. *J Clin Oncol*. 2014;32(36):4120-4126. doi:10.1200/JCO.2014.56.7503.
12. Thomson Reuters. Journal Citation Reports - Anesthesiology. ISI Web Knowl. 2016. <http://admin-apps.webofknowledge.com/JCR/JCR>. Accessed August 15, 2016.
13. Chow J, Jones P. Identifying areas to improve the quality of reporting in randomized clinical trials published in anesthesiology journals: A study protocol for a series of literature surveys assessing quality of trial registration, adherence to abstract reporting guidelin. 2016. doi:10.6084/m9.figshare.4490582.v1.
14. Benstetter M, Haubenreisser S. EU Clinical Trials Register goes live. Eur Med Agency. 2011. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Press\\_release/2011/03/WC500104032.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2011/03/WC500104032.pdf). Accessed August 1, 2016.
15. Klein AA, Anaesthesia Editorial Team. Author Guidelines. *Anaesthesia*. [http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1365-2044/homepage/ForAuthors.html](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1365-2044/homepage/ForAuthors.html). Accessed August 15, 2016.

16. Anesthesiology Editorial Office. Instructions for Authors. *Anesthesiology*. <http://anesthesiology.pubs.asahq.org/public/instructionsforauthors.aspx>. Accessed August 15, 2016.
17. Mahajan RP, British Journal of Anaesthesia Editorial Team. Instructions to Authors. *Br J Anaesth*. 2016. [http://www.oxfordjournals.org/our\\_journals/bjaint/for\\_authors/general.html](http://www.oxfordjournals.org/our_journals/bjaint/for_authors/general.html). Accessed August 15, 2016.
18. Backman S, Baker A, Beattie S, et al. 2011 Canadian Journal of Anesthesia Guide for Authors. *Can J Anesth Can d'anesthésie*. 2011;58(7):668-696. doi:10.1007/s12630-011-9516-4.
19. Anesthesia & Analgesia Editorial Office. Instructions for Authors. *Anesth Analg*. 2016. <http://edmgr.ovid.com/aa/accounts/ifaauth.htm>. Accessed August 15, 2016.
20. European Journal of Anaesthesiology Editorial Office. Guidance for Authors on the Preparation and Submission of Manuscripts to the European Journal of Anaesthesiology. *Eur J Anaesthesiol*. 2016. <http://edmgr.ovid.com/eja/accounts/ifaauth.htm>. Accessed August 15, 2016.
21. U.S. National Library of Medicine. MEDLINE/PubMed Data Element (Field) Descriptions. *US Natl Libr Med*. 2016. <https://www.nlm.nih.gov/bsd/mms/medlineelements.html>. Accessed August 15, 2016.
22. Toerien M, Brookes ST, Metcalfe C, et al. A review of reporting of participant recruitment and retention in RCTs in six major journals. *Trials*. 2009;10(52). doi:10.1186/1745-6215-10-52.
23. Hopewell S, Hirst A, Collins GS, Mallett S, Yu L-M, Altman DG. Reporting of participant flow diagrams in published reports of randomized trials. *Trials*. 2011;12(253). doi:10.1186/1745-6215-12-253.

24. Berwanger O, Ribeiro RA, Finkelsztejn A, et al. The quality of reporting of trial abstracts is suboptimal: Survey of major general medical journals. *J Clin Epidemiol.* 2009;62(4):387-392. doi:10.1016/j.jclinepi.2008.05.013.



## Chapter 5

### 5 Sample size calculations for randomized clinical trials published in anesthesiology journals: A comparison of 2010 versus 2016

Jeffrey T. Y. Chow<sup>1,3</sup>, Timothy P. Turkstra<sup>2,3</sup>, Edmund Yim<sup>3</sup>, Philip M. Jones<sup>1,2,3</sup>

<sup>1</sup>Department of Epidemiology & Biostatistics

<sup>2</sup>Department of Anesthesia & Perioperative Medicine

<sup>3</sup>Schulich School of Medicine & Dentistry, The University of Western Ontario

---

---

## 5.1 Introduction

In evidence-based medicine, randomized clinical trials (RCTs) are considered to provide the highest quality of evidence with the least amount of bias,<sup>1</sup> representing the gold standard for assessing therapeutic interventions.<sup>2</sup> In every study, participants are needed to provide data. However, the ideal number of participants cannot be determined without appropriate calculations. A sample size calculation involves determining the minimum number of participants needed to detect a treatment effect that is clinically relevant.<sup>3</sup> An inadequately small sample size may result in the inability to detect a precise effect, if present, while a needlessly large sample size may result in extra participants being exposed to the potential risks of the trial.<sup>4</sup> While there are ethical debates over whether underpowered trials should be conducted, an *a priori* sample size calculation should still be conducted and reported to ensure methodological quality.<sup>5</sup>

The Consolidated Standards of Reporting Trials (CONSORT) statement is a set of guidelines aimed at improving the quality of RCT reporting.<sup>6</sup> Item seven in the most recent checklist focuses on sample size, with recommendations to explain how sample size was determined and, ideally, to possess a high probability of detecting a statistically significant, clinically relevant difference if one exists.<sup>6</sup> The components needed for a sample size calculation are: the expected outcomes for each group (which implies the clinically relevant effect size), the type I ( $\alpha$ ) error level, the type II ( $\beta$ ) error or power ( $1-\beta$ ) level, the standard deviation for continuous outcomes, and any allowances made for attrition or non-compliance, if appropriate.<sup>6,7</sup>

In six high impact factor general medical journals, 95% of two-arm, parallel group, superiority RCTs reported a sample size calculation but only 53% reported all parameters required for sample size calculation.<sup>8</sup> For the RCTs that reported all parameters for sample size calculation, investigators replicated the sample size calculations and found that 18% of the replicated sample sizes were more than 10% lower than the reported sample size and 12% were more than 10% higher than the reported sample size.<sup>8</sup> In ten high impact factor anesthesiology journals, similar results were found for RCTs published in 2013, with 8.3% not reporting a sample size calculation and 28.7% of sample size recalculations more than 10% different from the reported sample size.<sup>9</sup> While the investigators found differences between the expected and actual values for each sample calculation parameter,<sup>9</sup> only the medians and interquartile ranges of the differences were presented and not the number of RCTs with over 10% difference between the expected and actual values.

The current study focused on anesthesiology RCTs, specifically two-arm, parallel group, superiority RCTs published in the top six general anesthesiology journals as determined by impact factor. The first objective of this study was to compare RCTs published in 2010 and 2016 in terms of the proportion that performed an *a priori* sample size calculation and the proportion that reported the necessary elements to allow for replication. This would update the literature using the most recent RCTs and examine potential trends over time in sample size calculation reporting. The second objective was to compare the expected sample size calculation parameters with the actual values observed in the results. The third objective was to perform post-hoc power calculations for the primary and first two reported secondary outcomes of each RCT. While post-hoc

power does not help with interpreting the results of a study since, by definition, a nonsignificant difference will always have low post-hoc power,<sup>10</sup> this exercise was not intended to interpret the results, but solely to examine the suitability of the sample size calculation.

## 5.2 Methods

An electronic database had been created for a previous project to collect all RCTs published in 2010 and 2016 from the top six general (non-pain-centric) anesthesiology journals as determined by impact factor: *Anaesthesia*, *Anesthesia & Analgesia (A&A)*, *Anesthesiology*, *British Journal of Anaesthesia (BJA)*, *Canadian Journal of Anesthesia (CJA)*, and *European Journal of Anaesthesiology (EJA)*.<sup>11</sup> Detailed screening and data extraction have been described previously in Chapters 3 and 4, with the study protocol publicly available as well.<sup>12</sup>

A RCT was defined as a prospective study that assessed randomly allocated health care interventions in human participants where authors had clearly reported that participants had been randomly allocated to study arms (i.e. the use of the words “random”, “randomized”, or “randomised”). The Table of Contents located on each journal’s website was independently searched by two reviewers for eligible RCTs. Any disagreements were discussed with a third reviewer with a decision made by consensus. Because only RCTs were included, all observational studies (e.g. case-control or cohort studies), learning curve studies, cadaver studies, cost-effectiveness studies, dose-finding or dose-response studies which were not designed to test a clinical intervention, diagnostic test accuracy studies, meta-analyses, editorials, narrative reviews, animal

studies, manikin studies, and re-analyses of previously published studies were excluded. RCTs published in the correspondence section, the letters to the editor section, or supplemental issues were also excluded. If a study reported more than one phase, data was only extracted from the phase with the highest number of participants. RCTs from this database were automatically screened for inclusion in the current study, with RCTs only eligible if there were two study arms and a parallel group, superiority design.

General characteristics such as journal, year, multicentricity, trial registration status, and sample size had been previously extracted for each RCT and those values were used to calculate summary measures for RCTs included in this study. RCTs were deemed adequately registered if the trial was registered in a publicly available trials registry prior to the first participant being enrolled with a clearly defined primary outcome. To determine trial registration status, the assigned reviewer searched the full-text for trial registration information. If no information was found, the reviewer searched [www.clinicaltrials.gov](http://www.clinicaltrials.gov), the International Standard Randomized Controlled Trial Number Register (ISRCTN), and the World Health Organization (WHO) Clinical Trials Search Portal. If no information was found during the registry search, two attempts separated by one week were made to contact the corresponding author via email for registration information. If none of these steps produced registration information, the study was considered not registered. Trials registered at the European Union Clinical Trials Register (EUCTR) were considered to be not registered to ensure consistency since the registry was only made publicly available in 2011.<sup>13</sup> Sample size was defined as the number randomized in the trial. The statistical software used for each RCT was also extracted, with the primary statistical package recorded if more than one was reported.

Data extraction for the current study was performed by one of four reviewers for data extraction. A second reviewer provided input if variables were unclear and a decision was made by consensus. During data extraction, the full-text of each included RCT was searched for the presence of a sample size calculation. Analyses were only performed for RCTs that had a sample size calculation that was amenable to replication, defined as using a clearly identified outcome that was continuous or binary in a standard sample size calculation procedure. For RCTs with sample size calculations amenable to replication, the assumptions made by the investigators were recorded such as the expected outcome for the treatment group, the expected outcome for the control group, the type I ( $\alpha$ ) error level, the type II ( $\beta$ ) error level, and whether any allowances were made for attrition or non-compliance. The standard deviation was also recorded for RCTs that used a continuous outcome in their sample size calculation. The actual values for the treatment group outcome, control group outcome, and standard deviation were also recorded to determine if the difference between actual and expected values was within 10% of the expected value. The justifications for these assumptions were recorded using a categorical outcome with options for published trial, meta-analysis, pilot study, observational data, unspecified, and other.

For studies that provided all assumptions needed to replicate the sample size (i.e. expected outcome for treatment group, expected outcome for control group, type I ( $\alpha$ ) error level, Type II ( $\beta$ ) error level, and standard deviation for continuous outcomes), the sample size was replicated using both the user-written add-on command, *ssi*,<sup>14</sup> and the *power* command in Stata. When provided in the sample size calculation, the allowances made for attrition or non-compliance were included in the recalculation to ensure

accuracy. The two replicated sample sizes were compared with the reported target sample size, with the sample sizes deemed to be equivalent for the purposes of this study if either replicated sample size had a difference of 10% or less on either side of the target sample size. When the difference (positive or negative) was greater than 10%, a discrepancy was noted and values checked again to ensure accuracy. For one-sided sample size calculations, the type I ( $\alpha$ ) error level was doubled and the same procedures followed. During data extraction, some trials with binary outcomes were found to not specify whether relative or absolute differences were expected. In those cases, both alternatives were tried with the replicated sample size closest to the reported target sample size used.

The outcome type, actual outcome value for the treatment group, actual outcome value for the control group, standard deviation for the treatment group, standard deviation for the control group, and the specified type I ( $\alpha$ ) error level were extracted for the primary and first two unique secondary outcomes (secondary outcomes were considered to be not unique if they assessed the same outcome but at different time points). If there was more than one primary outcome, the one used in the sample size calculation was selected. For sample size calculations using a one-sided test, the type I ( $\alpha$ ) error level was doubled for the primary outcome only. Only binary and continuous outcomes (reporting a mean and standard deviation) were included to ensure post-hoc power could be calculated. If a secondary outcome was not binary or continuous, the next secondary outcome was used instead. Using the reported sample size and assuming equal allocation to groups, the user-written add-on command, *ssi*, in Stata<sup>14</sup> was used to calculate post-hoc power. The sample size was defined as the number of participants randomized. The mean for the

post-hoc powers of each RCT's two secondary outcomes was recorded, with an overall mean and standard deviation calculated for all RCTs.

The post-hoc powers for each RCT were categorized as meeting the minimum thresholds (70% and 80%) or not. An *a priori* decision was made to use an 80% threshold value since most RCTs target a power of 80%, and a 70% threshold value to assess the number of RCTs close to meeting the target threshold. Due to the relationship between post-hoc power and statistical significance, the post-hoc power results were then stratified by whether the RCT was statistically significant. A RCT was defined as statistically significant if the manuscript indicated statistical significance for the primary outcome used in the sample size calculation.

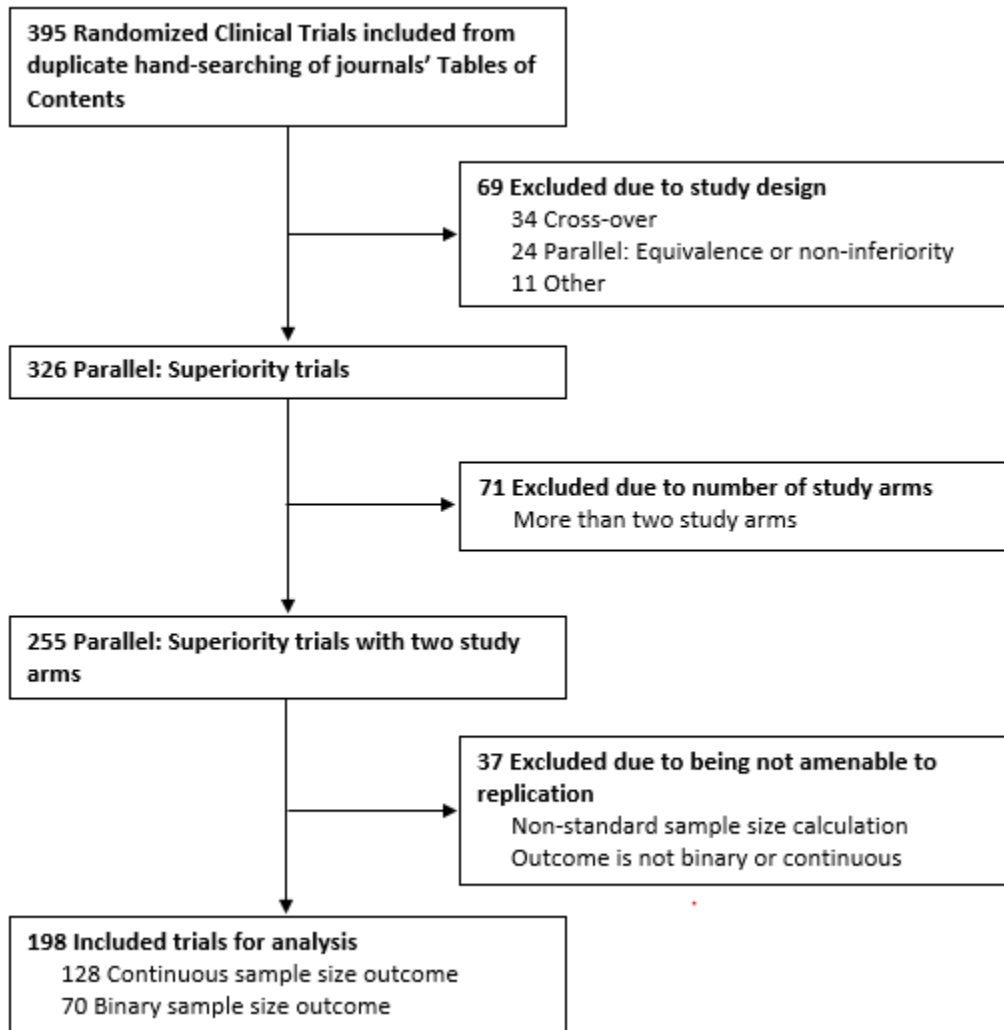
Descriptive statistics were used to present the raw numbers and percentages for all measured outcomes. All analyses were conducted using Stata 13 statistical software (StataCorp LLC, College Park, Texas).

### 5.3 Results

Of the 395 RCTs identified through duplicate hand-searching of the Table of Contents for the top six general anesthesiology journals (*Anaesthesia*, *A&A*, *Anesthesiology*, *BJA*, *CJA*, and *EJA*), 255 RCTs met the inclusion criteria of being parallel, superiority RCTs with two study arms (Figure 5-1). There were 143 RCTs from 2010 and 112 RCTs from 2016, with decreases in the number of RCTs from each journal other than Anesthesiology and CJA (Table 5-1). Most RCTs occurred at a single centre. Though SPSS was the most commonly used statistical software, the number of RCTs using SPSS decreased from 2010 to 2016; this coincided with increases in the use of Stata, SAS, and R. The



percentage of adequately registered RCTs increased almost seven-fold from 2010 to 2016 even though more than half of RCTs in 2016 were still inadequately registered (Table 5-1).



**Figure 5-1: Flowchart for inclusion of trials**

**Table 5-1: Characteristics of included RCTs**

Data presented are the number of trials (%) unless otherwise stated. Percentages may not add up to 100% due to rounding. IQR means interquartile range.

<sup>1</sup> Multi-centre is defined as participants recruited from more than one institution or clinic

<sup>2</sup> Adequate registration is defined as the trial being registered before the first participant was enrolled with a primary outcome clearly defined in the registry.

	<b>2010 (n=143)</b>	<b>2016 (n=112)</b>
<b>Journal</b>		
Anaesthesia	28 (20%)	22 (20%)
Anesthesia & Analgesia	37 (26%)	21 (19%)
Anesthesiology	11 (8%)	13 (12%)
British Journal of Anaesthesia	24 (17%)	22 (20%)
Canadian Journal of Anesthesia	12 (8%)	17 (15%)
European Journal of Anaesthesiology	31 (22%)	17 (15%)
<b>Multicentricity<sup>1</sup></b>		
Single-centre	137 (96%)	97 (87%)
Multi-centre	6 (4%)	15 (13%)
<b>Statistical Software</b>		
SPSS	68 (48%)	46 (41%)
Stata	10 (7%)	14 (13%)
SAS	8 (6%)	18 (16%)
R	1 (1%)	8 (7%)
JMP	2 (1%)	3 (3%)
Other	27 (19%)	14 (13%)
GraphPad (Prism/InStat)	9 (6%)	7 (6%)
SigmaPlot/SigmaStat	2 (1%)	3 (3%)
Statistica	2 (1%)	1 (1%)
StatView	4 (3%)	0 (0%)
Other	10 (7%)	3 (3%)
Not Mentioned	27 (19%)	9 (8%)
<b>Trial Registration<sup>2</sup></b>		
Adequately Registered	8 (6%)	46 (41%)
Inadequately Registered	135 (94%)	66 (59%)
<b>Median Sample Size (IQR)</b>	<b>63 (41–101)</b>	<b>80 (52–135.5)</b>

Only 110 RCTs in 2010 and 88 RCTs in 2016 had a sample size calculation amenable to replication, with a standard sample size calculation that used a clearly identified binary or continuous outcome (Table 5-2). The percentage of RCTs reporting all assumptions (expected outcome for treatment group, expected outcome for control group, type I error level, type II error level, and standard deviation for continuous outcomes) increased from 51% in 2010 to 84% in 2016. However, the difference between the actual and expected values for most RCTs was usually greater than 10% of the expected value, with negligible improvement from 2010 to 2016. Most RCTs based their sample size calculations on published trials, with the percentage continuing to increase in 2016 (Table 5-2).

**Table 5-2: Reporting and accuracy of sample size assumptions in RCTs**

Data presented are the number of trials with the specific criterion reported / number of eligible trials (%) unless otherwise specified. Percentages may not sum to 100% due to rounding.

<sup>1</sup> Number of trials where the actual value is within 10% of the assumption / number of trials with both the specific assumption and the actual value is reported (%).

<sup>2</sup> For trials using a continuous outcome in their sample size calculation

<sup>3</sup> Subgroups are not mutually exclusive so percentages will sum to over 100%

	2010	2016
<b>Sample size calculation amenable to replication</b>	110/143 (77%)	88/112 (79%)
No sample size calculation	16/33 (48%)	4/24 (17%)
Not amenable to replication	17/33 (52%)	20/24 (83%)
<b>Assumptions</b>		
Estimated outcome for treatment group	70/110 (64%)	76/88 (86%)
Actual values within 10% <sup>1</sup>	14/62 (23%)	17/68 (25%)
Estimated outcome for control group	70/110 (64%)	76/88 (86%)
Actual values within 10% <sup>1</sup>	26/62 (42%)	24/68 (35%)
Type I ( $\alpha$ ) error level	108/110 (98%)	88/88 (100%)
Type II ( $\beta$ ) error or power level	108/110 (98%)	88/88 (100%)
Standard deviation <sup>2</sup>	36/74 (49%)	50/54 (93%)
Actual values within 10% <sup>1,2</sup>	3/21 (14%)	4/21 (19%)
Allowances made for attrition or non-compliance	30/110 (27%)	47/88 (53%)
<b>All assumptions reported</b>	<b>56/110 (51%)</b>	<b>74/88 (84%)</b>
<b>Justification for assumptions<sup>3</sup></b>		
Published trial (RCT)	36/110 (33%)	45/88 (51%)
Meta-analysis	3/110 (3%)	1/88 (1%)
Pilot study	30/110 (27%)	21/88 (24%)
Observational data	14/110 (13%)	14/88 (16%)
Unspecified	18/110 (16%)	8/88 (9%)
Other	18/110 (16%)	7/88 (8%)

When replicating the initial sample size calculation for RCTs with all assumptions reported, 70% of RCTs estimated a sample size similar to the recalculated sample size, after including a 10% margin of error (Table 5-3). The average post-hoc power for primary and secondary outcomes decreased from 2010 to 2016, with large standard deviations. When categorizing the post-hoc power as meeting thresholds of 70% or 80%, most statistically significant RCTs had a primary post-hoc power greater than or equal to the threshold; as expected, RCTs that were not statistically significant had a low post-hoc power. For statistically significant RCTs, the percentage of RCTs having a post-hoc secondary outcome power  $\geq 70\%$  or  $\geq 80\%$  was about half of the percentage of RCTs having a post-hoc primary outcome power  $\geq 70\%$  (Table 5-3). Figure 5-2 shows the primary post-hoc power distributions stratified by whether the RCT is statistically significant, with RCTs that are not statistically significant having a lower post-hoc primary outcome power than RCTs that are statistically significant. Though the interquartile ranges are large, the median post-hoc power for the primary outcome appears to decrease from 2010 to 2016 irrespective of statistical significance (Figure 5-2).

**Table 5-3: Replication of sample size calculations for included RCTs**

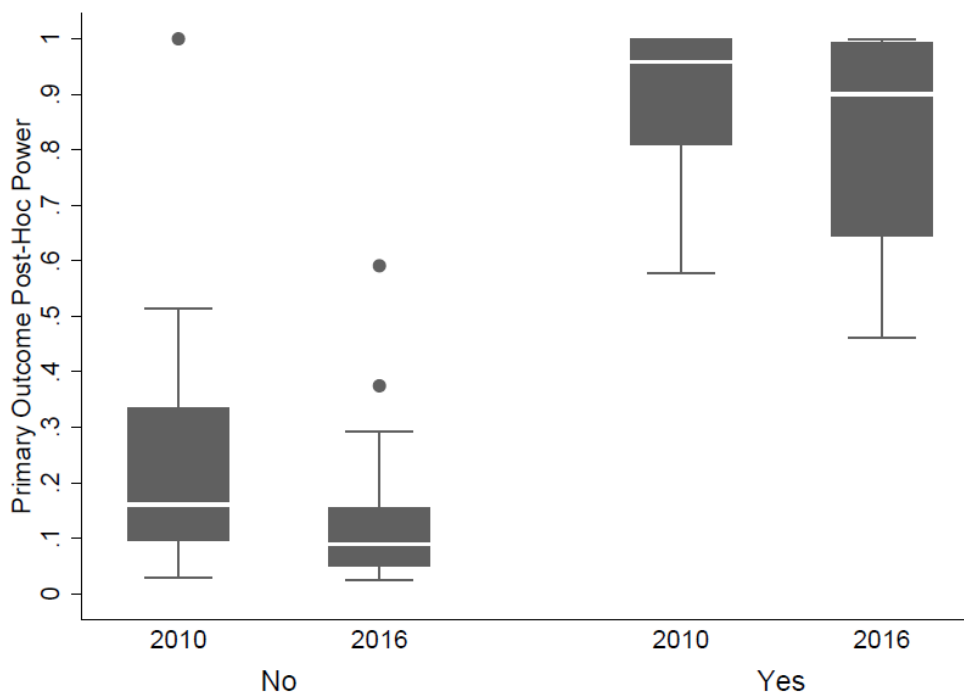
Data presented are the number of trials / number of eligible trials (%) unless otherwise stated.

<sup>1</sup> Number of trials where the replicated initial sample size is within 10% of the reported target sample size / number of trials with all sample size assumptions reported (%)

<sup>2</sup> Data presented as mean (standard deviation) post-hoc power; N = number of eligible trials

<sup>3</sup> Statistical significance as defined by what was reported in the manuscript for the primary outcome

	2010	2016
<b>Replication of initial sample size calculation<sup>1</sup></b>	39/56 (70%)	52/74 (70%)
<b>Average post-hoc power</b>		
Primary outcome <sup>2</sup>	0.65 (0.36)	0.49 (0.38)
	N=69	N=57
Average of first two reported secondary outcomes <sup>2</sup>	0.50 (0.31)	0.40 (0.31)
	N=101	N=80
<b>Number of statistically significant RCTs<sup>3</sup></b>	76/110 (69%)	50/88 (57%)
Primary outcome $\geq$ 80% post-hoc power	33/43 (77%)	19/29 (66%)
Primary outcome $\geq$ 70% post-hoc power	37/43 (86%)	20/29 (69%)
Secondary outcomes $\geq$ 80% post-hoc power	19/70 (27%)	11/43 (26%)
Secondary outcomes $\geq$ 70% post-hoc power	25/70 (36%)	16/43 (37%)
<b>Number of non-statistically significant RCTs<sup>3</sup></b>	34/110 (31%)	38/88 (43%)
Primary outcome $\geq$ 80% post-hoc power	1/25 (4%)	0/28 (0%)
Primary outcome $\geq$ 70% post-hoc power	1/25 (4%)	0/28 (0%)
Secondary outcome $\geq$ 80% post-hoc power	1/31 (3%)	1/37 (3%)
Secondary outcome $\geq$ 70% post-hoc power	5/31 (16%)	1/37 (3%)



**Figure 5-2: Distribution of primary post-hoc power stratified by statistical significance of the trial**

White line represents median; box represents 25th and 75th percentiles; each whisker represents 1.5 times interquartile range; dots represent outside values

For RCTs where the primary outcome used in the sample size calculation was binary (and when expected and actual values were reported for both the treatment and control groups), the majority of RCTs expected a larger effect size than the actual effect size observed (Table 5-4). From 2010 to 2016, there was an increase in the number and percentage of RCTs overestimating their effect sizes, but the mean absolute difference between the actual and expected effect sizes decreased (Table 5-4).

**Table 5-4: Comparison of estimated and reported effect sizes for binary outcomes**

Only trials where the primary outcome used in the sample size calculation was binary with estimated and actual values reported for both the treatment and control groups were included.

Data presented are the number of trials (%) unless otherwise stated.

<sup>1</sup> Data presented as mean (standard deviation) absolute difference between the actual effect size and the estimated effect size (only trials that had the effect sizes overestimated were included)

	<b>2010 (n=27)</b>	<b>2016 (n=30)</b>
Actual effect size < Estimated effect size	16 (59%)	23 (77%)
Actual effect size ≥ Estimated effect size	11 (41%)	7 (23%)
Difference between actual and estimated <sup>1</sup>	8.85% (6.75%)	6.31% (6.48%)

## 5.4 Discussion

In the top six general anesthesiology journals, the percentage of two-arm parallel group, superiority RCTs with a sample size calculation increased from 89% in 2010 to 96% in 2016. This is consistent with a previous study that found that 92% of anesthesiology RCTs in 2013 reported a sample size calculation.<sup>9</sup> While this is encouraging, these sample size calculations may have been based on assumptions different from reality. Consistent with previous studies in top general medical journals and anesthesiology journals,<sup>8,9</sup> about 30% of the initial size calculation replications were more than 10% different than the value estimated by the authors, suggesting that either researchers were calculating their sample sizes incorrectly or using non-standard sample size calculation formulas without specifying as such.

The assumptions needed for conducting sample size calculation were well reported, with increases seen in all sample size calculation parameters when comparing RCTs in 2016 to RCTs in 2010. The percentages in 2016 were also higher than a similar study in 2013,<sup>9</sup> suggesting an increase in sample size calculation reporting quality over time. Though Type I and Type II error levels were reported for all RCTs, the expected values for the treatment and control group were still below 90% in 2016, suggesting that the reason over 10% of RCTs were not explaining how the sample size was calculated was that the expected outcomes for the treatment and control group were not specified.

More concerningly, the postulated effect estimates themselves were often different from what was observed. Most RCTs that reported both expected and reported values had discrepancies larger than 10%. Of the RCTs from 2016 with a continuous outcome used in the sample size calculation, only 19% found a pooled standard deviation that was within 10% of the expected pooled standard deviation. Large differences between estimates and actual sample size calculation parameters were also found in general medical journals.<sup>8</sup> Around 80% of RCTs in general medical journals underestimated the sample size, with the actual standard deviation greater than the expected standard deviation.<sup>15</sup> This may be due to different techniques being available to estimate the standard deviation using previously published literature or pilot studies, with some methods such as using the standard deviation from a small pilot study having a 50% chance of underestimating the standard deviation and resulting in an underpowered study.<sup>16</sup> Investigators should be cautious when estimating sample size calculation parameters as the benefits of performing a sample size calculation are diminished if the estimates are inaccurate. However, the difficulty with estimating assumptions must also

be recognized; if the sample size calculation parameters are known with certainty, there may not be a need for a trial to be conducted. Since estimates are usually conducted in one of two ways, by using the treatment effect that is considered clinically meaningful or by using the treatment effect that is expected,<sup>17</sup> the inaccuracy of estimates may be due to the intervention being ineffective or the potential effect being poorly estimated. For RCTs where the clinically important difference is used as the anticipated effect size in the sample size calculation, discrepancies would not imply any incorrect estimation by investigators. Due to the potential for meta-analyses to aggregate the results of RCTs, under-powered trials may still contribute valuable evidence if there is high methodological rigour, clear reporting, and a lack of publication bias.<sup>5</sup>

With the sample size calculation assumptions different from what is observed, it would be helpful to also assess the power of the RCTs being conducted. While the importance of ensuring studies are adequately powered using *a priori* sample size calculations is widely accepted, calculating post-hoc power is inappropriate for interpreting the results of a study.<sup>10,18</sup> However, this study performed post-hoc power to examine the effects of inaccurately estimating sample size calculation assumptions. Another study performed similar calculations for anesthesiology RCTs that had a negative primary outcome, finding a mean post-hoc power (95% confidence interval (CI)) of 0.20 (95% CI: 0.13 to 0.27).<sup>9</sup> Most RCTs use a 0.80 power level in their sample size calculation, but in 2016, the average post-hoc power using the actual values for the parameters used in the sample size calculation was only 0.49, with only 66% of statistically significant RCTs having a post-hoc primary outcome power of at least 80%. If the sample size calculation



assumptions are inaccurately estimated such that RCTs are inadequately powered, the value in performing sample size calculations is diminished.

Since there is a direct relationship between p-values and post-hoc power, outcomes that are not statistically significant correspond to low post-hoc power.<sup>19</sup> The distribution of the post-hoc power for primary outcomes, as shown in Figure 5-1, is consistent with this concept. On the other hand, the post-hoc power of secondary outcomes is unlikely to correspond to the statistical significance of a RCT's primary outcome. A common concern when a trial has multiple secondary outcomes is an increase in the Type I error due to multiple testing.<sup>20</sup> These secondary outcomes are often considered exploratory or hypothesis-generating, with most *a priori* sample size calculations only planning for adequate power for the primary outcome. By calculating the average post-hoc power for secondary outcomes, the results provide further support that secondary outcomes should be interpreted cautiously due to their low power. Interestingly, almost all RCTs that were not statistically significant had mean secondary outcome post-hoc powers below 70% as well. One potential explanation is that if a RCT is inadequately powered for the primary outcome, the secondary outcomes will likely be even more underpowered.

Most sample size calculation parameters are poorly estimated and result in inadequately powered RCTs. Two potential solutions include improving estimates or using alternative methods to determine the sample size. Termed “sample size samba” or “delta inflation,” investigators commonly start with the number of available (or fundable) participants and adjust their estimates of the sample size calculation assumptions to justify their sample size.<sup>5,21</sup> Increasing the expected effect size will increase the power,<sup>22</sup> allowing for a smaller sample size to be calculated. For included RCTs with a binary outcome, the

majority find a smaller effect size than what was expected, showing one parameter that may be manipulated to produce a feasible sample size. If estimates are being manipulated in this way, it should be no surprise when the actual values are different from what is expected. Among the flaws of the traditional sample size calculation is the difficulty in accurately estimating all parameters and the design-use mismatch resulting from using a p-value cut-off to design a study that should not be interpreted based on a single p-value.<sup>23</sup> While performing sample size calculations can provide valuable information, less importance should be placed on these calculations due to the inherent subjectivity of the estimates.<sup>5</sup> In the reporting of RCTs, confidence intervals may provide a better measure of uncertainty and better help readers understand the strength of the evidence.<sup>24</sup> Some other alternatives to sample size calculations include using hybrid approaches, value of information methods, pragmatic methods based on cost or feasibility, sensitivity analyses, or using the same sample size as previous analogous studies.<sup>23,25,26</sup>

The results of this study must be interpreted considering its limitations. Since this study only included two-arm, parallel group, superiority RCTs published in the top anesthesiology journals as determined by impact factor, the generalizability of the findings may be limited. While this study focused on the most commonly used trial design, different sample size calculations are used for different trial designs. Since only information from the published manuscript was used, definitive conclusions about sample size calculations being performed *a priori* cannot be made. Previous research has identified discrepancies in sample size calculations between published manuscripts and their original protocols<sup>27</sup> so there is a chance of authors creating a sample size calculation based on the number of participants recruited. Another limitation was that analyses could

only be performed for RCTs that reported the relevant values. For example, RCTs that did not report all necessary sample size calculation assumptions could not be replicated to check for accuracy. For continuous outcomes, only those with a mean and standard deviation were used to compute post-hoc power, potentially making the results not applicable to RCTs with non-normal data reporting median and interquartile ranges.

While two reviewers were used to screen for eligible RCTs, data extraction for each RCT was only performed by a single reviewer. To reduce potential inconsistency, all reviewers were trained using the same guidelines, with data quality checks throughout the process.

In conclusion, almost all RCTs published in the top six general anesthesiology journals reported a sample size calculation. In 2016, most two-arm, parallel group, superiority RCTs included enough information in the methods section to allow for replication of the sample size calculation. While this represented a large increase from 2010, the outcome values and variability used in the sample size calculation were often different from what was observed, with most actual values having a difference from the expected value greater than 10%. The majority of RCTs using binary outcomes had a smaller reported effect size than was expected in the sample size calculation. More research needs to be conducted into how sample size calculation parameter estimates can be improved or whether alternative methods should be used to determine a trial's ideal sample size.

## 5.5 Literature Cited

1. Guyatt G, Rennie D, Meade MO, Cook DJ. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd ed. United States of America: McGraw-Hill Education; 2015.

2. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the Gold Standard — Lessons from the History of RCTs. Malina D, ed. *N Engl J Med*. 2016;374(22):2175-2181. doi:10.1056/NEJMms1604593.
3. Noordzij M, Tripepi G, Dekker FW, Zoccali C, Tanck MW, Jager KJ. Sample size calculations: Basic principles and common pitfalls. *Nephrol Dial Transplant*. 2010;25(5):1388-1393. doi:10.1093/ndt/gfp732.
4. Gupta K, Attri J, Singh A, Kaur H, Kaur G. Basic concepts for sample size calculation: Critical step for any clinical trials! *Saudi J Anaesth*. 2016;10(3):328. doi:10.4103/1658-354X.174918.
5. Schulz KF, Grimes D a. Sample size calculations in randomised trials: mandatory and mystical. *Lancet*. 2005;365(9467):1348-1353. doi:10.1016/S0140-6736(05)61034-3.
6. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340(1):c869-c869. doi:10.1136/bmj.c869.
7. Campbell M, Julious S, Altman D. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in paired comparisons. *BMJ*. 1995;311(7013):1145-1148.
8. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials: review. *BMJ*. 2009;338(1):b1732-b1732. doi:10.1136/bmj.b1732.
9. Abdulatif M, Mukhtar A, Obayah G. Pitfalls in reporting sample size calculation in randomized controlled trials published in leading anaesthesia journals: a systematic review. *Br J Anaesth*. 2015;115(5):699-707. doi:10.1093/bja/aev166.
10. Goodman SN, Berlin JA. The Use of Predicted Confidence Intervals When Planning Experiments and the Misuse of Power When Interpreting Results. *Ann Intern Med*. 1994;121(3):200-206. doi:10.7326/0003-4819-121-3-199408010-00008.

11. Thomson Reuters. Journal Citation Reports - Anesthesiology. ISI Web Knowl. 2016. <http://admin-apps.webofknowledge.com/JCR/JCR>. Accessed August 15, 2016.
12. Chow J, Jones P. Identifying areas to improve the quality of reporting in randomized clinical trials published in anesthesiology journals: A study protocol for a series of literature surveys assessing quality of trial registration, adherence to abstract reporting guidelin. 2016. doi:10.6084/m9.figshare.4490582.v1.
13. Benstetter M, Haubenreisser S. EU Clinical Trials Register goes live. Eur Med Agency. 2011. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Press\\_release/2011/03/WC500104032.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2011/03/WC500104032.pdf). Accessed August 1, 2016.
14. Jones PM. SSI: Stata module to estimate sample size for randomized controlled trials. IDEAS. 2010. <https://ideas.repec.org/c/boc/bocode/s457150.html>. Accessed September 1, 2016.
15. Vickers AJ. Underpowering in randomized trials reporting a sample size calculation. *J Clin Epidemiol*. 2003;56(8):717-720. doi:10.1016/S0895-4356(03)00141-0.
16. Chen H, Zhang N, Lu X, Chen S. Caution regarding the choice of standard deviations to guide sample size calculations in clinical trials. *Clin Trials*. 2013;10(4):522-529. doi:10.1177/1740774513490250.
17. Wittes J. Sample size calculations for randomized controlled trials. *Epidemiol Rev*. 2002;24(1):39-53.
18. Levine M, Ensom MHH. Post Hoc Power Analysis: An Idea Whose Time Has Passed? *Pharmacother Off J Am Coll Clin Pharm*. 2001;21(4):405-409. doi:10.1592/phco.21.5.405.34503.
19. Hoenig JM, Heisey DM. The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis. *Am Stat*. 2001;55(1):19-24. doi:10.1198/000313001300339897.

20. Feise RJ. Do multiple outcome measures require p-value adjustment? *BMC Med Res Methodol.* 2002;2(1):8. doi:10.1186/1471-2288-2-8.
21. Aberegg SK, Richards DR, O'Brien JM. Delta inflation: a bias in the design of randomized controlled trials in critical care medicine. *Crit Care.* 2010;14(2):R77. doi:10.1186/cc8990.
22. Wilson Van Voorhis CR, Morgan BL. Understanding Power and Rules of Thumb for Determining Sample Sizes. *Tutor Quant Methods Psychol.* 2007;3(2):43-50. doi:10.20982/tqmp.03.2.p043.
23. Bacchetti P. Current sample size conventions: Flaws, harms, and alternatives. *BMC Med.* 2010;8(1):17. doi:10.1186/1741-7015-8-17.
24. Bacchetti P, Leung JM. Sample size calculations in clinical research. *Anesthesiology.* 2002;97(4):1028-1029.
25. Ciarleglio MM, Arendt CD. Sample size determination for a binary response in a superiority clinical trial using a hybrid classical and Bayesian procedure. *Trials.* 2017;18(1):83. doi:10.1186/s13063-017-1791-0.
26. Bacchetti P, McCulloch CE, Segal MR. Simple, Defensible Sample Sizes Based on Cost Efficiency. *Biometrics.* 2008;64(2):577-585. doi:10.1111/j.1541-0420.2008.01004\_1.x.
27. Chan A-W, Hrobjartsson A, Jorgensen KJ, Gotzsche PC, Altman DG. Discrepancies in sample size calculations and data analyses reported in randomised trials: comparison of publications with protocols. *BMJ.* 2008;337:a2299. doi:10.1136/bmj.a2299.

## Chapter 6

### 6 Impact of funding source on randomized clinical trials published in anesthesiology journals

Jeffrey T. Y. Chow<sup>1,3</sup>, Timothy P. Turkstra<sup>2,3</sup>, Edmund Yim<sup>3</sup>, Philip M. Jones<sup>1,2,3</sup>

<sup>1</sup>Department of Epidemiology & Biostatistics

<sup>2</sup>Department of Anesthesia & Perioperative Medicine

<sup>3</sup>Schulich School of Medicine & Dentistry, The University of Western Ontario

---

---

## 6.1 Introduction

Randomized clinical trials (RCTs) are prospective cohort studies where interventions are randomly allocated to participants.<sup>1</sup> The methodological rigour of RCTs reduces the amount of bias inherent in clinical investigations, and this reduction in bias results in RCTs occupying the top of the evidence hierarchy for therapeutic interventions.<sup>2</sup> However, other factors may potentially negatively affect RCT quality, diminishing the credibility of some RCTs. One of these factors is the RCT's funding source.

The past few decades have seen a large increase in the number of industry-sponsored trials,<sup>3</sup> with industry funding surpassing government or public funding in 2001.<sup>4</sup> Systematic reviews have often concluded that industry-sponsored studies are more likely to reach conclusions favourable to the sponsor.<sup>5-7</sup> Reasons suggested for these findings include methodological quality, trial designs, publication biases, and interpretations of results.<sup>5,6,8</sup> However, there is still no clear explanation for the association between industry funding and positive results.<sup>9</sup>

Another aspect of funding source not studied is its effect, if any, in the anesthesiology literature. In the United States, anesthesiology continues to be underrepresented in National Institutes of Health (NIH) funding, with only 40% of current academic anesthesiology departments having at least one NIH grant credited to a faculty member or trainee and only 1% of the NIH budget going to anesthesiology departments, despite anesthesiologists representing 6% of the medical workforce.<sup>10</sup> In Canada, the Canadian Institutes of Health Research (CIHR), the main source of government funding for medical research, has reduced funding by 7.5% adjusted for inflation from 2007 to



2013.<sup>11</sup> With low government funding, other support sources such as industry funding may be employed at a more frequent rate compared to other specialties. Though more than half of original research in anesthesiology are clinical trials,<sup>12</sup> results from Chapter 3 show that the absolute number of anesthesiology RCTs have decreased over time.

The effect of funding source on RCT quality has not yet been studied in the anesthesiology literature so we sought to elucidate the role of funding source, if any, in the anesthesiology literature. Our first objective was to determine what types of financial support fund RCTs in anesthesiology and whether there have been differences in funding source over time. Our second objective addressed whether the source of funding would bias the design and reporting quality of RCTs (as discussed in previous chapters). Trial registration, outcome switching, adherence to CONSORT for Abstracts, sample size calculations, and post-hoc power were used as outcomes to provide a new perspective for the potential distorting effect of funding on trial reporting. To align our study with the previous literature,<sup>5-7</sup> our hypothesis was that most RCTs in the top six general anesthesiology journals would not be funded by industry, but those that were may potentially have had their quality negatively affected.

## 6.2 Methods

To facilitate previous projects, an electronic database was set up including all RCTs published in 2007, 2010, 2013, 2015, and 2016 for the top six general (non-pain-centric) anesthesia journals as determined by impact factor: *Anaesthesia*, *Anesthesia & Analgesia* (A&A), *Anesthesiology*, *British Journal of Anaesthesia* (BJA), *Canadian Journal of Anesthesia* (CJA), and *European Journal of Anaesthesiology* (EJA).<sup>13</sup> Other than funding

source, all data had previously been collected (see chapters 3, 4, and 5). The following five paragraphs outline the screening and data extraction methods that were conducted in previous chapters, with the study protocol also being publicly available.<sup>14</sup>

Two reviewers independently searched the Table of Contents on each journal's website to screen for eligible RCTs, with disagreements resolved by consensus with a third reviewer. RCTs were eligible if they were prospective studies that assessed randomly allocated health care interventions in human participants and identified by authors clearly reporting that patients had been randomly allocated to study arms (i.e. the use of the words "random", "randomized", or "randomised"). Studies were not included if they were observational studies (e.g. case-control, or cohort studies), learning curve studies, cadaver studies, cost-effectiveness studies, dose-finding or dose-response studies which were not designed to test a clinical intervention, diagnostic test accuracy studies, meta-analyses, editorials, narrative reviews, animal studies, manikin studies, simulation studies, duplicate reports, re-analyses of previously published RCTs, studies published in the correspondence section, letters to the editor, or studies part of supplemental issues.

RCTs identified as eligible were prospectively included in an electronic database for data extraction. Data extraction was conducted by one reviewer with any uncertainty discussed with a second reviewer and a decision made by consensus. After extracting general characteristic information such as year and journal, the trial registration status was assessed. An RCT was only included if the RCT had been registered in a publicly available trials registry prior to the first patient being enrolled and had a clearly defined primary outcome. If registration information was not found in the manuscript, trial registries ([www.clinicaltrials.gov](http://www.clinicaltrials.gov), the International Standard Randomized Controlled

Trial Number Register, and the WHO Clinical Trials Search Portal) were searched. If no results were found after the search, the corresponding author was contacted using two standardized emails with one week in between. If the date of trial registration occurred after the date of first patient enrollment, the RCT was considered to be inadequately registered. If the first patient enrollment date was not provided, the study start date was used instead. A RCT was considered to have a clearly defined primary outcome if the primary outcome was clearly and unambiguously defined with a specific time frame in both the trials registry and the manuscript.

The presence of outcome discrepancies was assessed for all included RCTs in 2007, 2010, 2013, and 2015. A RCT was considered to have a primary outcome discrepancy if a registered primary outcome was not reported as a primary outcome, a reported primary outcome was not registered as a primary outcome, or the timing of a reported primary outcome differed from a registered primary outcome. A RCT was considered to have a secondary outcome discrepancy if a registered secondary outcome was not reported as a secondary outcome, a reported secondary outcome was not registered as a secondary outcome, or the timing of a reported secondary outcome differed from a registered secondary outcome. Any outcome discrepancy was defined by a RCT having either a primary or secondary outcome discrepancy.

In Chapters 4 and 5, data extraction was conducted by one of four assigned reviewers. Only RCTs published in 2010 and 2016 were included. Uncertain variables were discussed with a third reviewer, with a decision made by consensus. The general characteristics for each RCT were extracted such as journal, year, multicentricity, study design, sample size, and trial registration status. Sample size was defined as the number

randomized and adequately registered was defined as a RCT registered in a publicly available trials registry prior to the first patient being enrolled with a clearly defined primary outcome. Reviewers assessed the abstract and full-text for each item in the CONSORT for Abstracts statement, with each item recorded as reported in the abstract, not reported in abstract but reported in manuscript, or not reported in abstract or manuscript. While there are 16 items in the CONSORT for Abstracts statement, each individual item was not analyzed separately for the current study. The outcomes assessed as being reported in the abstract were title, trial design, methods (at least five of six criteria met), results/conclusions (at least five of six criteria met), and all checklist items (number of criteria met).

For sample size outcomes, only two-arm, parallel group, superiority RCTs were included for analysis. In Chapter 5, each RCT was examined for the presence of a sample size calculation. If there was a sample size calculation, the text would be examined to determine if all assumptions necessary for replication were included such as expected outcome for treatment group, expected outcome for control group, type I ( $\alpha$ ) error level, type II ( $\beta$ ) error/power level, and standard deviation for continuous outcomes. The RCT was also classified as statistically significant if the manuscript indicated statistical significance for the primary outcome used in the sample size calculation. For RCTs that included all assumptions necessary for replication, the sample size calculation was repeated using the user-written add-on command, *ssi*,<sup>15</sup> and the power command in Stata. RCTs were considered to be accurately replicated if either replicated sample size had a difference from the reported target sample size that was 10% or less. For the primary and first two unique secondary outcomes (i.e. not the same time point) reported, the outcome

type, actual outcome value for the treatment group, the actual outcome value for the control group, the standard deviation for the treatment group, the standard deviation for the control group, and the specified type I ( $\alpha$ ) error level were extracted. Only binary and continuous outcomes were used. Along with the sample size extracted previously, these values were used with the user-written add-on command, *ssi*, in Stata<sup>15</sup> to calculate the post-hoc power for each outcome assuming equal allocation of participants. For each RCT, the mean of the two secondary outcomes was taken to produce a single value for the secondary outcome post-hoc power.

To assess the potential impact of industry funding, RCTs published in 2016 were used for all outcomes, except for the outcome discrepancy outcomes which used RCTs published in 2015 due to data limitations. During data extraction for the current study, the funding source for each RCT was recorded using a categorical outcome with options: *no funding*; *only industry*; *university, hospital, or government*; *private non-profit*; *multiple sources of funding (including industry involvement)*; *multiple sources of funding (no industry involvement)*; *other*; or *not reported*. Using this variable, two groupings were created for analysis. Grouping One compared *industry-only* funding to *non-industry-only* funding, with non-industry-only funding comprising of RCTs with either no funding; university, hospital or government; private non-profit; other; multiple sources of funding (including industry); or multiple sources of funding (no industry). Grouping Two compared *any industry funding* to *no industry funding*, with any industry funding comprising of RCTs with either only industry or multiple sources of funding (including industry) and no industry funding comprising of RCTs with either no funding; university, hospital or government; private non-profit; other; or multiple sources of funding (no industry).

Descriptive statistics were used to present the raw numbers and percentages for all measured outcomes. A chi-squared test for linear trend with the year published as a continuous outcome was used to assess the change in funding between the measured years: 2007, 2010, 2013, 2015, 2016. To determine whether the type of funding had an effect on trial reporting, simple logistic regression was used for binary outcomes and simple linear regression was used for continuous outcomes. All outcomes were binary (i.e., the RCT either met the criteria or not) except for the overall CONSORT for Abstracts checklist item score, primary outcome post-hoc power, and secondary outcome post-hoc power. Stata 13 statistical software (College Park, Texas) was used to conduct all analyses. The analysis was conducted twice for each outcome, once with Grouping One as the explanatory variable where industry-only funding was compared to non-industry-only funding, and once with Grouping Two as the explanatory variable where any industry funding was compared to no industry funding. RCTs where the funding source was not reported were excluded from these analyses.

### 6.3 Results

A total of 1,036 RCTs were included, with 316 RCTs from 2007, 219 RCTs from 2010, 170 RCTs from 2013, 155 RCTs from 2015, and 176 RCTs from 2016 (Table 6-1). A decreasing number of RCTs were identified with each successive year, except in 2016, which had more RCTs than 2013 and 2015. The number and percentage of RCTs where the funding source was *not reported* consistently decreased with each successive year examined from 2007 to 2016. In every year examined, the *university, hospital, or government* category was the most prevalent funding source for RCTs. Industry funding represented a minority of anesthesiology RCTs, when measured using Grouping One or

Grouping Two (Table 6-1). Figure 6-1 illustrates the funding source for each of the included journals in 2010 and 2016. Notably, in 2016, *Anesthesiology* had the highest percentage of RCTs funded by industry and *CJA* had no RCTs funded by industry (Figure 6-1).

The results of simple regression models assessing the impact of industry funding on included RCTs from 2015/2016 only were inconclusive (Table 6-2). For all outcomes where linear and logistic regression was able to be performed, only the *trial design* element of the CONSORT for Abstracts statement for Grouping Two was statistically significant. All RCTs in 2015 funded fully or in part by industry had at least one outcome discrepancy and no RCTs in 2016 funded fully or in part by industry reported all sample size assumptions. For post-hoc power outcomes, regression models were also performed stratified by whether the RCT was statistically significant, though some could not be performed since the primary outcome post-hoc power was always low for RCTs that were not statistically significant (Table 6-2).

**Table 6-1: Funding source of RCTs by year**

Data presented are the number of trials (%). Percentages may not add up to 100% due to rounding.

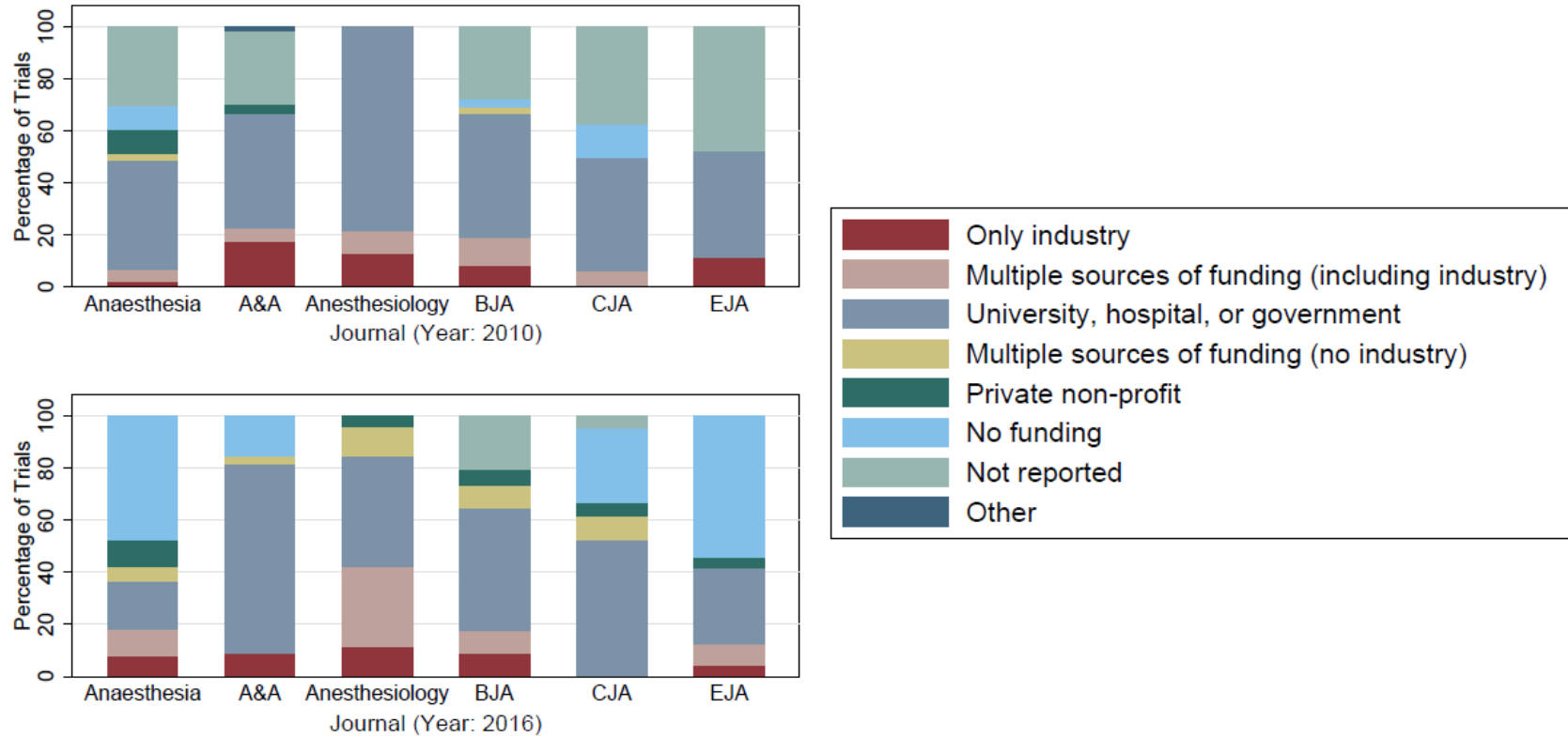
<sup>1</sup> Non-industry-only funding is the sum of trials with: no funding; university, hospital or government; private non-profit; other; or multiple sources of funding

<sup>2</sup> Grouping 2 consists of trials from 2010, 2015, and 2016 where the funding source is reported. *Any industry* refers to the funding source being only industry or having multiple sources of funding with at least one being industry; *no industry* refers to the funding source being non-industry or having multiple sources of funding with none being industry.

Statistical Significance: Denotes the p-value for the chi-squared test for linear trend

<b>Funding Source</b>	<b>2007 (n=316)</b>	<b>2010 (n=219)</b>	<b>2013 (n=170)</b>	<b>2015 (n=155)</b>	<b>2016 (n=176)</b>	<b>Statistical Significance</b>
Not reported	156 (49%)	66 (30%)	15 (9%)	12 (8%)	8 (5%)	p<0.0001
<b>Grouping One</b>						
No funding	12 (4%)	7 (3%)	37 (22%)	37 (24%)	42 (24%)	p<0.0001
Only industry	29 (9%)	22 (10%)	13 (8%)	9 (6%)	13 (7%)	p=0.0001
University, hospital, or government	82 (26%)	103 (47%)	82 (48%)	76 (49%)	76 (43%)	p=0.08
Private non-profit	8 (3%)	6 (3%)	5 (3%)	8 (5%)	9 (5%)	p=0.76
Other	0 (0%)	1 (0%)	0 (0%)	0 (0%)	0 (0%)	p=0.51
Multiple sources of funding	29 (9%)	14 (6%)	18 (11%)	13 (8%)	28 (16%)	p=0.45
Non-industry only funding <sup>1</sup>	131 (41%)	131 (60%)	142 (84%)	134 (86%)	155 (88%)	p=0.0001
<b>Grouping Two<sup>2</sup></b>	N/A	(n=153)	N/A	(n=143)	(n=168)	
Any industry	N/A	34 (22%)	N/A	19 (13%)	30 (18%)	N/A
No industry	N/A	119 (78%)	N/A	124 (87%)	138 (82%)	N/A





**Figure 6-1: Funding source of RCTs by journal for 2010 and 2016**

**Table 6-2: Results of simple regression models for the impact of industry funding in included RCTs from 2015/2016 only**

The Trial Registration section uses 2015 data while the CONSORT and Sample Size sections use 2016 data. Data presented are the odds ratio (95% CI; p-value) unless otherwise stated. Grouping One compares industry-only funding to non-industry-only funding (reference group). Grouping Two compares any industry funding to no industry funding (reference group).

n = number of trials included in the regression model

<sup>1</sup> Outcome was met if at least five of six CONSORT for Abstracts criteria were reported in the abstract

<sup>2</sup> Data presented as average difference in score (95% CI; p-value)

<sup>3</sup> Statistical significance as defined by what was reported in the manuscript for the primary outcome

<sup>4</sup> Number of trials (%) where the replicated initial sample size is within 10% of the reported target sample size

<sup>5</sup> Data presented as average difference in power (95% CI; p-value). If more than one primary outcome was specified in the manuscript, the primary outcome refers to the primary outcome used in the sample size calculation. Secondary outcomes refer to the average of the first two reported secondary outcomes.

\*1 Logistic regression could not be performed since all RCTs in the industry group (8 trials) had at least one discrepancy

\*2 Logistic regression could not be performed since no RCT in the industry group (13 trials) reported at least five of six items

\*3 Logistic regression could not be performed since all RCTs in the industry group (16 trials) reported all sample size assumptions

\*4 Linear regression could not be performed since all RCTs in the industry group (2 trials) could not have the primary post-hoc power calculated

\*5 Logistic regression could not be performed since all RCTs had a primary post-hoc power below the threshold

	<b>Grouping One</b> <b>(Industry-only funding vs non-industry- only funding)</b>	<b>Grouping Two</b> <b>(any industry funding vs no industry funding)</b>
<b>Trial Registration</b>		
Adequately registered (n=143)	1.26 (0.32 to 4.91; 0.74)	1.15 (0.43 to 3.07; 0.78)
Primary outcome discrepancy (n=56)	1.26 (0.16 to 9.65; 0.82)	1.29 (0.29 to 5.75; 0.74)
Secondary outcome discrepancy (n=56)	0.32 (0.03 to 3.68; 0.40)	0.81 (0.08 to 8.04; 0.86)
Any outcome discrepancy (n=56)	*1	*1
<b>CONSORT - Abstract Adherence</b>		
Title (n=168)	0.52 (0.16 to 1.68; 0.29)	1.16 (0.46 to 2.93; 0.75)
Trial design (n=168)	2.45 (0.78 to 7.67; 0.12)	2.29 (1.02 to 5.10; 0.04)
Methods (5/6 items) <sup>1</sup> (n=168)	*2	0.36 (0.05 to 2.90; 0.28)
Results/Conclusions (5/6 items) <sup>1</sup> (n=168)	0.50 (0.06 to 4.07; 0.48)	0.96 (0.30 to 3.07; 0.95)
Trial Registration (n=168)	0.74 (0.22 to 2.52; 0.63)	0.57 (0.23 to 1.36; 0.19)
Overall score <sup>2</sup> (n=168)	-0.58 (-2.16 to 1.00; 0.47)	0.20 (-0.90 to 1.30; 0.72)
<b>Sample Size</b>		
All sample size assumptions reported (n=83)	*3	*3
Statistically significant primary outcome <sup>3</sup> (n=83)	1.92 (0.35 to 10.52; 0.44)	0.92 (0.31 to 2.77; 0.89)
Replication of initial sample size calculation <sup>4</sup> (n=70)	2.58 (0.47 to 14.31; 0.26)	2.55 (0.78 to 8.34; 0.11)
Primary outcome post-hoc power <sup>5</sup> (n=55)	0.40 (-0.04 to 0.85; 0.07)	0.04 (-0.27 to 0.35; 0.81)
For trials that are statistically significant <sup>3,5</sup> (n=28)	0.07 (-0.22 to 0.35; 0.64)	0.00 (-0.26 to 0.26; 0.99)
For trials that are not statistically significant <sup>3,5</sup> (n=27)	*4	-0.02 (-0.18 to 0.14; 0.77)
Secondary outcomes post-hoc power <sup>5</sup> (n=71)	0.02 (-0.24 to 0.29; 0.86)	-0.14 (-0.33 to 0.06; 0.16)
For trials that are statistically significant <sup>3,5</sup> (n=38)	-0.04 (-0.40 to 0.32; 0.83)	-0.20 (-0.49 to 0.09; 0.18)
For trials that are not statistically significant <sup>3,5</sup> (n=33)	0.04 (-0.30 to 0.37; 0.83)	-0.05 (-0.26 to 0.15; 0.60)
Primary outcome ≥ 80% post-hoc power (n=55)	4.93 (0.42 to 58.57; 0.19)	0.88 (0.15 to 5.06; 0.89)
For trials that are statistically significant <sup>3</sup> (n=28)	1.33 (0.11 to 16.74; 0.82)	0.60 (0.07 to 5.03; 0.64)
For trials that are not statistically significant <sup>3</sup> (n=27)	*5	*5
Primary outcome ≥ 70% post-hoc power (n=55)	4.50 (0.38 to 53.29; 0.22)	0.80 (0.14 to 4.59; 0.80)
For trials that are statistically significant <sup>3</sup> (n=28)	1.12 (0.09 to 14.20; 0.93)	0.50 (0.06 to 4.23; 0.53)
For trials that are not statistically significant <sup>3</sup> (n=27)	*5	*5
Secondary outcomes ≥ 80% post-hoc power (n=71)	0.01 (-0.30 to 0.33; 0.94)	-0.09 (-0.32 to 0.14; 0.46)
For trials that are statistically significant <sup>3</sup> (n=38)	-0.01 (-0.50 to 0.47; 0.95)	-0.11 (-0.52 to 0.29; 0.57)
For trials that are not statistically significant <sup>3</sup> (n=33)	-0.03 (-0.30 to 0.23; 0.80)	-0.04 (-0.20 to 0.13; 0.64)
Secondary outcomes ≥ 70% post-hoc power (n=71)	-0.05 (-0.40 to 0.30; 0.78)	-0.15 (-0.41 to 0.11; 0.24)
For trials that are statistically significant <sup>3</sup> (n=38)	-0.13 (-0.66 to 0.40; 0.62)	-0.24 (-0.68 to 0.20; 0.28)
For trials that are not statistically significant <sup>3</sup> (n=33)	-0.03 (-0.30 to 0.23; 0.80)	-0.04 (-0.20 to 0.13; 0.64)

## 6.4 Discussion

While there have been absolute increases to medical research funding in the United States, the overall proportions from each funding source have remained relatively constant.<sup>16</sup> More trials registered in ClinicalTrials.gov from 2006 to 2014 have been funded by industry, but the proportion of industry-funded trials has not increased.<sup>17</sup> In anesthesiology, the results are similar with the proportion of RCTs funded by industry showing a slight decrease from 2007 to 2016. In 2016, only 7% of RCTs were funded by solely industry sources and 18% by at least one industry source. By comparison, 37% of interventional RCTs registered in ClinicalTrials.gov,<sup>18</sup> 44% of pulmonary, critical care, and sleep medicine RCTs,<sup>19</sup> 60% of psychiatry RCTs,<sup>20</sup> and 62% of rheumatoid arthritis RCTs<sup>21</sup> were found to be funded by industry sources. This suggests that industry funding does not have as large an impact in anesthesiology, where most RCTs are funded by university, hospital, or government (Figure 6-1).

There have been many studies comparing industry-sponsored studies to non-industry-sponsored studies, with industry-sponsored studies historically being more likely to report favourable efficacy results, less likely to report evidence of harm, and more likely to report favourable conclusions.<sup>7,22,23</sup> However, no difference was found when assessing whether the risk of bias (methodological quality scores, sequence generation, concealment of allocation, loss to follow-up, etc.) for trials differed by funding source, other than industry-sponsored trials being more likely to report adequate blinding.<sup>23</sup> The current study explored facets of methodological quality that had not been commonly researched in terms of the funding source impact. Our analysis found no significant differences between industry funded RCTs and non-industry funded RCTs in terms of

trial registration, outcome switching, adherence to CONSORT for Abstracts, sample size calculations, or post-hoc power. With the wide 95% confidence intervals around our point estimates, either there was no industry funding effect or there were not enough RCTs present in the analysis to be able to detect a real effect (i.e., we lacked statistical power).

For the outcomes concerning trial registration, no differences were found for whether a RCT was adequately registered, whether a RCT had at least one primary outcome discrepancy, or whether a RCT had at least one secondary outcome discrepancy. Though not significant, there was a trend that RCTs funded by industry were more likely to be adequately registered and have at least one primary outcome discrepancy, but less likely to have at least one secondary outcome discrepancy. This corresponds to a study using ClinicalTrials.gov records finding that compared to academic, non-profit, and government organizations, industry funded trials were more likely to be registered before or within three months of the trial start date.<sup>24</sup> However, when the primary and secondary outcome discrepancy outcomes were combined into one composite outcome assessing whether a RCT had at least one discrepancy regardless of it being primary or secondary, logistic regression could not be performed since all RCTs funded by industry had at least one discrepancy.

Even when using two grouping methods, one classifying RCTs into industry-only funding and one classifying RCTs into any industry funding, no differences were found when comparing to RCTs not funded by industry. From Table 6-2, the trial design criterion when comparing RCTs with any industry funding and no industry funding was the only regression model that was statistically significant. However, with wide

confidence intervals approaching an odds ratio of one and the high number of statistical tests performed, this marginally statistically significant result was possibly due to chance. Due to the low number of included RCTs, all journals were pooled together in the analyses. However, different journals had different proportions of RCTs being funded by industry sources, potentially influencing the results.

These findings highlighted one of the major limitations of this study — adequate sample size. Since the number of RCTs was dependent on the number found in the journals from previous chapters, no *a priori* sample size calculations were performed. As a result, this study may have been underpowered to identify the effect industry funding may have had on RCT methodological quality, if such an effect truly existed. Having only a minority of RCTs funded by industry sources would exaggerate the problem of low sample size and contribute to the limitation of inadequate power. Another limitation was that data extraction for each RCT was performed by a single reviewer. To reduce potential error, all reviewers were trained prior to data extraction using the same guidelines and unclear variables were discussed during the process. RCTs where the funding source was not reported present another limitation, as it would not be apparent whether they were funded by industry or not. The proportion of RCTs with the funding source not reported decreased with each successive year measured. In 2016, only two journals had RCTs where the funding source was not reported, comprising 5% of eligible RCTs.

In conclusion, a minority of RCTs published in the top six general anesthesiology journals were funded by industry sources. Contrary to our hypothesis, no methodological quality differences were found between industry and non-industry funded sources in terms of trial registration, outcome switching, adherence to CONSORT for Abstracts,

sample size calculations, or post-hoc power. The low number of industry-funded RCTs and lack of an industry funding effect suggest that, in contrast to other areas of medical research, the funding source may not be a significant problem in anesthesiology RCTs. However, more studies are needed to confirm the lack of an industry funding effect and more research needs to be conducted to determine whether other measures of quality of trial reporting are impacted by the funding source.

## 6.5 Literature Cited

1. Sessler DI, Imrey PB. Clinical Research Methodology 3: Randomized Controlled Trials. *Anesth Analg*. 2015;121(4):1052-1064. doi:10.1213/ANE.0000000000000862.
2. Guyatt G, Rennie D, Meade MO, Cook DJ. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 3rd ed. United States of America: McGraw-Hill Education; 2015.
3. Buchkowsky SS, Jewesson PJ. Industry Sponsorship and Authorship of Clinical Trials Over 20 Years. *Ann Pharmacother*. 2004;38(4):579-585. doi:10.1345/aph.1D267.
4. Patsopoulos NA. Origin and funding of the most frequently cited papers in medicine: database analysis. *BMJ*. 2006;332(7549):1061-1064. doi:10.1136/bmj.38768.420139.80.
5. Bekelman JE, Li Y, Gross CP. Scope and Impact of Financial Conflicts of Interest in Biomedical Research. *JAMA*. 2003;289(4):454-465. doi:10.1001/jama.289.4.454.
6. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003;326(7400):1167-1170. doi:10.1136/bmj.326.7400.1167.
7. Sismondo S. Pharmaceutical company funding and its consequences: A qualitative systematic review. *Contemp Clin Trials*. 2008;29(2):109-113. doi:10.1016/j.cct.2007.08.001.

8. Lexchin J. Those Who Have the Gold Make the Evidence: How the Pharmaceutical Industry Biases the Outcomes of Clinical Trials of Medications. *Sci Eng Ethics*. 2012;18(2):247-261. doi:10.1007/s11948-011-9265-3.
9. Gluud LL. Bias in Clinical Intervention Research. *Am J Epidemiol*. 2006;163(6):493-501. doi:10.1093/aje/kwj069.
10. Shwinn DA, Balsler JR. Anesthesiology Physician Scientists in Academic Medicine. *Anesthesiology*. 2006;104(1):170-178.
11. Canadian Association of University Teachers. Federal Funding of Basic Research. *CAUT Educ Rev*. 2013;13(1):1-6.
12. Chilkoti G, Sharma CS, Kochhar A, Agrawal D, Sethi AK. An overview of clinical research for anesthesiologists. *J Anaesthesiol Clin Pharmacol*. 2010;26(4):446-450.
13. Thomson Reuters. Journal Citation Reports - Anesthesiology. ISI Web Knowl. 2016. <http://admin-apps.webofknowledge.com/JCR/JCR>. Accessed August 15, 2016.
14. Chow J, Jones P. Identifying areas to improve the quality of reporting in randomized clinical trials published in anesthesiology journals: A study protocol for a series of literature surveys assessing quality of trial registration, adherence to abstract reporting guidelin. 2016. doi:10.6084/m9.figshare.4490582.v1.
15. Jones PM. SSI: Stata module to estimate sample size for randomized controlled trials. IDEAS. 2010. <https://ideas.repec.org/c/boc/bocode/s457150.html>. Accessed September 1, 2016.
16. Moses H, Matheson DHM, Cairns-Smith S, George BP, Palisch C, Dorsey ER. The Anatomy of Medical Research. *JAMA*. 2015;313(2):174. doi:10.1001/jama.2014.15939.
17. Ehrhardt S, Appel LJ, Meinert CL. Trends in National Institutes of Health Funding for Clinical Trials Registered in ClinicalTrials.gov. *JAMA*. 2015;314(23):2566. doi:10.1001/jama.2015.12206.



18. Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A. Characteristics of Clinical Trials Registered in ClinicalTrials.gov, 2007-2010. *JAMA*. 2012;307(17):1838. doi:10.1001/jama.2012.3424.
19. Todd JL, White KR, Chiswell K, Tasneem A, Palmer SM. Using ClinicalTrials.gov to Understand the State of Clinical Research in Pulmonary, Critical Care, and Sleep Medicine. *Ann Am Thorac Soc*. 2013;10(5):411-417. doi:10.1513/AnnalsATS.201305-111OC.
20. Perlis RH, Perlis CS, Wu Y, Hwang C, Joseph M, Nierenberg AA. Industry Sponsorship and Financial Conflict of Interest in the Reporting of Clinical Trials in Psychiatry. *Am J Psychiatry*. 2005;162(10):1957-1960. doi:10.1176/appi.ajp.162.10.1957.
21. Khan NA, Lombeida JI, Singh M, Spencer HJ, Torralba KD. Association of industry funding with the outcome and quality of randomized controlled trials of drug therapy for rheumatoid arthritis. *Arthritis Rheum*. 2012;64(7):2059-2067. doi:10.1002/art.34393.
22. Bero L. Industry Sponsorship and Research Outcome. *JAMA Intern Med*. 2013;173(7):580. doi:10.1001/jamainternmed.2013.4190.
23. Lundh A, Sismondo S, Lexchin J, Busuioc O a, Bero L. Industry sponsorship and research outcome. In: Lundh A, ed. *Cochrane Database of Systematic Reviews*. Vol 12. Chichester, UK: John Wiley & Sons, Ltd; 2012:MR000033. doi:10.1002/14651858.MR000033.pub2.
24. Zarin DA, Tse T, Williams RJ, Rajakannan T. Update on Trial Registration 11 Years after the ICMJE Policy Was Established. *N Engl J Med*. 2017;376(4):383-391. doi:10.1056/NEJMSr1601330.

## Chapter 7

### 7 Integrated Discussion and General Conclusions

---

---

## 7.1 Overview

This thesis is comprised of a series of four articles aimed at identifying potential areas where authors can reduce distortion by improving the quality of their research reporting. Randomized clinical trials (RCTs) from across the world were included, and a breakdown by country was created using Tableau Public 10.2 (Appendix E). Focusing on RCTs published in the top six general anesthesiology journals, the four overall objectives were:

- 1) To determine the rate of adequate trial registration and whether the reported outcomes were the same as the outcomes originally registered in publicly available clinical trial registries
- 2) To determine the degree of adherence to CONSORT for Abstracts reporting guidelines and the availability of that information in the full-text manuscript
- 3) To determine the quality of sample size calculations and assess the post-hoc power for primary and secondary outcomes
- 4) To determine the rate of industry funding and whether it has an effect on the methodological quality, as measured by the previous objectives

## 7.2 Integrated Discussion of Results

Methodological surveys are similar to systematic reviews since they both aggregate data from published studies. However, methodological surveys use information about how studies are conducted instead of the results of those studies. Chapters 3, 4, and 5 are methodological surveys while Chapter 6 explores the impact of funding on those measures of methodological quality. In each chapter, positives and negatives about the quality of anesthesiology RCTs were found.

Chapter 3 showed an increasing number of RCTs being adequately registered from 2007 to 2015. However, only 38% of RCTs in 2015 were adequately registered. Of the adequately registered RCTs, 92% had at least one primary or secondary discrepancy between the registry entry and the published article, with 42% having at least one primary outcome discrepancy and 90% having at least one secondary outcome discrepancy. These results were similar to a previous study in 2013<sup>1</sup> and suggest that more work needs to be done to improve trial registration, both in terms of increasing prospective trial registration and ensuring that registry entries are followed.

Chapter 4 assessed the adherence to the CONSORT for Abstracts guidelines in RCTs from 2010 and 2016. Consistent with previous research,<sup>2</sup> many RCTs did not meet the CONSORT for Abstracts guidelines. Though improvement was seen from 2010 to 2016, the overall level of adherence was still lacking with around 75% of RCTs meeting fewer than half of the 16 criteria and no RCTs reporting all 16 items in the abstract. For most criteria, the information not reported in the abstract was available in the full-text, suggesting that abstracts were not adequately summarizing the entirety of the reported RCT.

Chapter 5 showed similar trends to the previous two chapters with the sample size calculation quality increasing over time but still remaining inadequate in 2016. With analyses restricted to two-arm, parallel group, superiority RCTs, almost all RCTs from 2016 performed a sample size calculation and 84% reported all necessary assumptions to allow for replication of the sample size calculation. However, similar to previous studies in top general medical journals and anesthesiology journals,<sup>3,4</sup> around 30% of sample size calculation replications were more than 10% different from the target sample size

reported. More concerningly, the expected outcome values and variability were different from what was observed, with most RCTs having discrepancies between the expected and reported assumptions larger than 10%. In terms of post-hoc power, statistically significant RCTs had a higher primary outcome post-hoc power than non-statistically significant RCTs, as expected from theory.<sup>5</sup> However, only 66% of statistically significant RCTs had a post-hoc primary outcome power of 80% or more, and the average post-hoc power for secondary outcomes was, as expected, even lower than that of the primary outcome.

Chapter 6 further explored the measures of methodological quality described in previous chapters by comparing RCTs funded by industry to those not funded by industry. A minority of included RCTs were funded by industry sources, across all years. Two groupings were used to classify RCTs as industry-funded: only industry and any industry. However, neither showed a difference when compared to no industry funded in terms of trial registration, outcome discrepancies, adherence to CONSORT for Abstracts guidelines, quality of sample size calculations, or post-hoc power.

### 7.3 General Conclusions

In evidence-based medicine, RCTs (and systematic reviews of their results) are typically considered to be at the top of the evidence hierarchy due to their lower risk of bias relative to non-randomized trials.<sup>6</sup> However, their findings may still be distorted if careful attention is not paid to the design and reporting quality. Trial registration helps to reduce both publication and outcome reporting bias,<sup>7</sup> reducing the chance that the published literature is distorted in favour of statistically significant RCTs. Abstracts are intended to

be concise and accurate summaries of the entire study, and they are frequently assessed by readers who never read the full trial report. Therefore, abstract reporting guidelines should be followed more closely to ensure that all relevant information is included.<sup>8,9</sup> Since most of the information being left out of the abstract is later presented in the full-text, adherence to the abstract reporting guidelines can be easily improved. Sample size calculations are intended as a guide to determine the ideal number of participants to be recruited, balancing the need to be able to detect a precise effect with the need to avoid exposing extra participants to potential risks.<sup>10</sup> However, inaccuracies in sample size calculations may result in readers having a distorted understanding of the effects of an intervention, if not properly reported. Any discussion of potential literature distortion should also consider the possible impacts of funding source since there may be financial pressures to achieve certain results.<sup>11</sup>

While the quality of reporting in anesthesiology RCTs has improved over time, this thesis identified several areas that can still be improved. In terms of *design quality*, trial registration and quality of sample size calculations need to be improved, while in terms of *reporting quality*, there needs to be improved adherence to abstract reporting guidelines and reduced discrepancies between registered and reported outcomes. Industry funding does not appear to be a major concern for anesthesiology RCTs with only a minority being industry-funded and no differences observed for the above measures of quality. When interpreting the findings from this thesis, careful consideration for the generalizability is needed as only RCTs from the top six general anesthesiology journals were included for analyses. As well, only a limited number of years were evaluated, reducing the ability to detect trend and differences over time due to the potential for

conflicting findings in unmeasured years. More research is needed to confirm the results of these studies as well as to identify potential tools that can be used to improve the quality of RCTs. More research is also needed to ensure that current trial registration policies, CONSORT for Abstracts reporting guidelines, and sample size calculation methods are suitable for reducing distortion in anesthesiology RCTs.

## 7.4 Literature Cited

1. De Oliveira GS, Jung MJ, McCarthy RJ. Discrepancies Between Randomized Controlled Trial Registry Entries and Content of Corresponding Manuscripts Reported in Anesthesiology Journals. *Anesth Analg*. 2015;121(4):1030-1033. doi:10.1213/ANE.0000000000000824.
2. Can OS, Yilmaz AA, Hasdogan M, et al. Has the quality of abstracts for randomised controlled trials improved since the release of Consolidated Standards of Reporting Trial guideline for abstract reporting? A survey of four high-profile anaesthesia journals. *Eur J Anaesthesiol*. 2011;28(7):485-492. doi:10.1097/EJA.0b013e32833fb96f.
3. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials: review. *BMJ*. 2009;338(1):b1732-b1732. doi:10.1136/bmj.b1732.
4. Abdulatif M, Mukhtar A, Obayah G. Pitfalls in reporting sample size calculation in randomized controlled trials published in leading anaesthesia journals: a systematic review. *Br J Anaesth*. 2015;115(5):699-707. doi:10.1093/bja/aev166.
5. Goodman SN, Berlin JA. The Use of Predicted Confidence Intervals When Planning Experiments and the Misuse of Power When Interpreting Results. *Ann Intern Med*. 1994;121(3):200-206. doi:10.7326/0003-4819-121-3-199408010-00008.

6. Guyatt G, Rennie D, Meade MO, Cook DJ. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd ed. United States of America: McGraw-Hill Education; 2015.
7. Fredrickson MJ, Ilfeld BM. Prospective Trial Registration for Clinical Research. *Reg Anesth Pain Med*. 2011;36(6):619-624. doi:10.1097/AAP.0b013e318230fbc4.
8. Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet*. 2008;371(9609):281-283. doi:10.1016/S0140-6736(07)61835-2.
9. Hopewell S, Clarke M, Moher D, et al. CONSORT for Reporting Randomized Controlled Trials in Journal and Conference Abstracts: Explanation and Elaboration. von Elm E, ed. *PLoS Med*. 2008;5(1):e20. doi:10.1371/journal.pmed.0050020.
10. Gupta K, Attri J, Singh A, Kaur H, Kaur G. Basic concepts for sample size calculation: Critical step for any clinical trials! *Saudi J Anaesth*. 2016;10(3):328. doi:10.4103/1658-354X.174918.
11. Lexchin J. Those Who Have the Gold Make the Evidence: How the Pharmaceutical Industry Biases the Outcomes of Clinical Trials of Medications. *Sci Eng Ethics*. 2012;18(2):247-261. doi:10.1007/s11948-011-9265-3.



## Appendices

### Appendix A: CONSORT for Abstracts checklist

Items to include when reporting a randomized trial in a journal or conference abstract

<b>Item</b>	<b>Description</b>
Title	Identification of the study as randomized
Authors*	Contact details for the corresponding author
Trial design	Description of the trial design (e.g. parallel, cluster, non-inferiority)
<b>Methods</b>	
Participants	Eligibility criteria for participants and the settings where the data were collected
Interventions	Interventions intended for each group
Objective	Specific objective or hypothesis
Outcome	Clearly defined primary outcome for this report
Randomization	How participants were allocated to interventions
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment
<b>Results</b>	
Numbers randomized	Number of participants randomized to each group
Recruitment	Trial status
Numbers analysed	Number of participants analysed in each group
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision
Harms	Important adverse events or side effects
Conclusions	General interpretation of the results
Trial registration	Registration number and name of trial register
Funding	Source of funding

\*this item is specific to conference abstracts

Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet*. 2008;371(9609):281-283. doi:10.1016/S0140-6736(07)61835-2.

**Appendix B: Sample email to corresponding authors**

Dear Dr *Corresponding Author*,

We are researchers from the University of Western Ontario in London, Ontario, Canada. We are conducting a study of published literature in Anesthesiology journals. We are contacting you because you were the corresponding author on the paper "*Name of published article*" published in "*Journal*" in *Month, Year*.

One of the items we are studying is clinical trial registration (such as at a site like [www.clinicaltrials.gov](http://www.clinicaltrials.gov)). We could not find any evidence that your trial was registered, and hence we are writing to see if your trial was indeed registered at any clinical trial registry.

If your trial was registered, would you please reply to this email to tell us where it was registered?

Even if your paper was not registered, we would appreciate it if you could reply to this email to let us know that it wasn't.

Your assistance is greatly appreciated.

Kind regards,

*Researcher*

on behalf of the Research Team.

-----  
Study Contact: Philip M Jones, MD MSc (Clinical Trials) FRCPC (Principal Investigator)

Associate Professor

Department of Anesthesia & Perioperative Medicine

Program in Critical Care, Department of Medicine

Department of Epidemiology & Biostatistics

University of Western Ontario / London Health Sciences Centre

(Office use only - Article ID: *Article ID*)

### **Appendix C: Differences between outcomes when comparing the published trial to the trial registry among adequately registered trials (stratified by journal)**

A: Data presented are the number of trials (%). All percentages are column percentages with the denominator set to the total number of trials published.

B: Data presented are the proportion of trials where the identified discrepancy favoured statistical significance (see article text for details).

All uses of 'registered' pertain to the outcome as registered in the trial registry and all uses of 'reported' pertain to the outcome as reported in the published manuscript. If there are one or more occurrences of the discrepancy in a trial, the trial will be counted as having the described discrepancy. Subcategories are not mutually exclusive so if individual studies have more than one discrepancy, the sum of the subcategories will be larger than the parent category.

"N/A" represents values where it could not be determined whether the discrepancy favoured statistical significance: either when the discrepancy involved an outcome not being reported or when there were no studies with the discrepancy.

<sup>a</sup>Trials with at least one discrepancy between the primary or secondary outcomes reported in the published article and those registered in the trial registry.

<sup>b</sup>Trials with at least one discrepancy between the primary outcome reported in the published article and that registered in the trial registry.

<sup>c</sup>Trials with at least one discrepancy among any of the secondary outcomes reported in the published article and those registered in the trial registry.

\* While the wording is slightly different, both rows are equivalent, as both show a registered primary outcome being reported as a secondary outcome.

\*\* While the wording is slightly different, both rows are equivalent, as both show a registered secondary outcome being reported as a primary outcome.

	Anaesthesia (n=8)		Anesthesia & Analgesia (n=18)		Anesthesiology (n=16)		British Journal of Anaesthesia (n=31)		Canadian Journal of Anesthesia (n=16)		European Journal of Anaesthesiology (n=13)	
	A	B	A	B	A	B	A	B	A	B	A	B
Trials with any primary or secondary outcome discrepancy <sup>a</sup>	4 (50%)	1.00	17 (94%)	0.53	14 (88%)	0.71	30 (97%)	0.67	16 (100%)	0.75	13 (100%)	0.15
<b>Trials with primary outcome discrepancies<sup>b</sup></b>	1 (13%)	1.00	9 (50%)	0.44	7 (44%)	0.86	20 (65%)	0.65	5 (31%)	1.00	4 (31%)	0.25
Registered primary outcome not reported as primary outcome	0 (0%)	N/A	6 (33%)	0.17	3 (19%)	0.00	12 (39%)	0.17	2 (13%)	0.50	2 (15%)	0.00
Registered primary outcome not reported	0 (0%)	N/A	3 (17%)	N/A	2 (13%)	N/A	4 (13%)	N/A	1 (6%)	N/A	1 (8%)	N/A
Registered primary outcome reported as secondary outcome*	0 (0%)	N/A	3 (17%)	0.33	1 (6%)	0.00	10 (32%)	0.20	1 (6%)	1.00	1 (8%)	0.00
Reported primary outcome not registered as primary outcome	0 (0%)	N/A	5 (28%)	0.60	4 (25%)	1.00	12 (39%)	0.83	4 (25%)	1.00	4 (31%)	0.25
Reported primary outcome not registered	0 (0%)	N/A	4 (22%)	0.50	3 (19%)	1.00	11 (35%)	0.82	2 (13%)	1.00	3 (23%)	0.00
Reported primary outcome registered as secondary outcome**	0 (0%)	N/A	1 (6%)	1.00	1 (6%)	1.00	1 (3%)	1.00	2 (13%)	1.00	1 (8%)	1.00
Timing of reported primary outcome different from primary outcome registered	1 (13%)	1.00	4 (22%)	0.00	3 (19%)	1.00	4 (13%)	0.50	0 (0%)	N/A	1 (8%)	0.00
<b>Trials with secondary outcome discrepancies<sup>c</sup></b>	4 (50%)	0.75	16 (89%)	0.44	13 (81%)	0.62	30 (97%)	0.53	16 (100%)	0.69	12 (92%)	0.17
Registered secondary outcome not reported as secondary outcome	0 (0%)	N/A	5 (28%)	0.20	5 (31%)	0.20	7 (23%)	0.14	7 (44%)	0.29	4 (31%)	0.25
Registered secondary outcome not reported	0 (0%)	N/A	4 (22%)	N/A	4 (25%)	N/A	6 (19%)	N/A	5 (31%)	N/A	3 (23%)	N/A
Registered secondary outcome reported as primary outcome**	0 (0%)	N/A	1 (6%)	1.00	1 (6%)	1.00	1 (3%)	1.00	2 (13%)	1	1 (8%)	1
Reported secondary outcome not registered as secondary outcome	3 (38%)	1.00	16 (89%)	0.38	12 (75%)	0.50	30 (97%)	0.50	14 (88%)	0.71	10 (77%)	0.00
Reported secondary outcome not registered	3 (38%)	1.00	16 (89%)	0.38	12 (75%)	0.50	29 (94%)	0.45	14 (88%)	0.64	10 (77%)	0.00
Reported secondary outcome registered as primary outcome*	0 (0%)	N/A	3 (17%)	0.33	1 (6%)	0.00	10 (32%)	0.20	1 (6%)	1.00	1 (8%)	0.00
Timing of reported secondary outcome different from secondary outcome registered	1 (13%)	0.00	5 (28%)	0.40	1 (6%)	1.00	6 (19%)	0.17	3 (19%)	0.33	4 (31%)	0.25

## Appendix D: Adapted version of CONSORT for Abstracts statement explanations

1. Abstract (if criteria reported in abstract)
2. FulltextOnly (if criteria not reported in abstract but reported in fulltext)
3. NotReported (if criteria not reported in abstract and not reported in fulltext)

Criteria	Explanation
<b>Title: Identification of the study as randomized</b>	<p>Authors should state explicitly in the title that the participants were randomly assigned to their comparison groups.</p> <ul style="list-style-type: none"> <li>- Title - Criteria reported in title (counts if randomized is in the subtitle)</li> <li>- Abstract - Criteria not reported in title but reported in abstract</li> <li>- FulltextOnly - Criteria not reported in title or abstract but reported in fulltext</li> <li>- NotReported - Criteria not reported in title, abstract, or fulltext"</li> </ul>
<b>Trial Design: Description of the trial design</b>	<p>The design of the trial should be described: E.G. parallel group, cluster randomized, crossover, factorial, superiority, equivalence/noninferiority, etc.</p> <p>Note: Trial design needs to be clear. I.E. If it is a superiority trial, should state that design is to test whether one is better. Parallel group does not need to be explicitly stated if it is mentioned that subjects are randomized to different groups.</p> <ul style="list-style-type: none"> <li>- If authors only state that they are comparing two interventions (not to see whether one is better or to see whether they are equivalent), no credit should be given since they are not clear enough. However, if the sample size calculation shows that superiority is hypothesized, "trial design" would be credited with "full text only."</li> </ul>
<b>Participants: Eligibility criteria for participants and the settings</b>	<p>A clear description of BOTH the <b>trial participants</b> and <b>setting in which they were studied</b> is needed. Participant eligibility criteria may relate to demographics, clinical diagnosis, and comorbid conditions while the trial may be performed in a particular setting (e.g. primary, secondary, or tertiary care).</p>
<b>Interventions: Interventions intended for each group</b>	<p>The essential features of the experimental and comparison interventions should be described. Authors should report details about the interventions (e.g. dose, route of administration, duration of administration, frequency, surgical procedure, or manufacturer of inserted device).</p>

	<ul style="list-style-type: none"> <li>- As long as clear and specific what was used</li> </ul>
<b>Objective: Specific objective or hypothesis</b>	<p>A clear statement of the specific objective or hypothesis addressed in the trial. If more than one objective is addressed, the main objective (i.e. based on the prespecified primary outcome) should be indicated and only key secondary objectives stated.</p> <ul style="list-style-type: none"> <li>- If authors only state that they are comparing two interventions (not to see whether one is better or to see whether they are equivalent), no credit should be given since they are not clear enough.</li> </ul>
<b>Outcome: Clearly defined primary outcome for this report</b>	<p>Authors should explicitly state the primary outcome for the trial and when it was assessed (e.g., the time frame over which it was measured).</p> <ul style="list-style-type: none"> <li>- Need to specify “primary outcome/endpoint”</li> </ul>
<b>Randomization: How participants were allocated to interventions</b>	<p>The method for assigning participants to interventions is clearly described.</p> <p>Note: Need to specify how participants were allocated (just a computer-generated randomization list is not enough)</p> <p>Examples of approaches used to ensure adequate concealment include: centralised (e.g. allocation by a central office) or pharmacy-controlled randomization; sequentially numbered identical containers that are administered serially to participants; on-site computer system combined with allocations kept in a locked, unreadable computer file that investigators can access only after the characteristics of an enrolled participant are entered; and sequentially numbered, opaque sealed envelopes.</p>
<b>Blinding (masking): Participants, caregivers, or outcome assessors blinded to group assignment</b>	<p>It is important that authors describe whether or not participants, those administering the intervention (usually health-care providers), and those assessing the outcome (the data collectors and analysts) were blinded to the group allocation. Authors should report if any form of blinding (such as blinding of data analysts) was used.</p> <ul style="list-style-type: none"> <li>- If there is no blinding, authors should state “no blinding, no masking, etc.” in order to count</li> <li>- Using a placebo counts as participant blinding</li> <li>- Using terms such as “single” or “double” blinding fulfills this criteria</li> </ul> <p>Blinding refers to the practice of keeping the trial participants, care providers, data collectors, and sometimes those analysing the data, unaware of which intervention is being administered to which</p>

	<p>participant, so that they will not be influenced by that knowledge. The term masking is sometimes used instead of blinding.</p>
<p><b>Numbers randomized: Subjects randomized to each group</b></p>	<p>The number of participants <b>randomized</b> to each intervention group is an essential element of the results of a trial.</p> <ul style="list-style-type: none"> <li>- Only counts if a number is provided for each group</li> </ul>
<p><b>Recruitment: Trial status</b></p>	<p>Authors should describe the status of the trial and whether it is still ongoing, closed to recruitment, or closed to follow-up. If the trial has stopped earlier than planned it is important to say why.</p> <ul style="list-style-type: none"> <li>- Note: If they reach their target sample size, we can assume that the trial is completed (full-text only if target sample size and number studied is only in full-text)</li> <li>- Okay if they provided a date of when data collection ended</li> </ul> <p>Possible reasons for early termination include: slow accrual rates, poor data quality, poor adherence, resource deficiencies, unacceptable harms or large benefits, or emerging information that makes the trial irrelevant, unnecessary, or unethical.</p>
<p><b>Numbers analysed: Participants analysed in each group</b></p>	<p>Authors should report the number of participants included <b>in the analysis</b> for each intervention group.</p> <ul style="list-style-type: none"> <li>- State numbers analyzed for entire trial or for primary outcome (stating for a secondary outcome only does not count I.E. 10/20 people achieved the secondary outcome is not enough but 10/20 people achieved the primary outcome is enough)</li> <li>- Only counts if a number is provided for each group</li> </ul>
<p><b>Outcome: A result for each group, overall effect size, and precision</b></p>	<p>For the primary outcome, authors should report trial results as a summary of the outcome in each group (e.g. the number of participants with or without the event, or the mean and standard deviation of measurements), together with the contrast between groups known as the effect size (with uncertainty such as 95% CI). For binary outcomes, the effect size could be the relative risk, relative risk reduction, odds ratio, or risk difference. For survival time data, the measurement could be the hazard ratio or difference in median survival time. For continuous data, the effect measure is usually the difference in means.</p> <ul style="list-style-type: none"> <li>- For primary outcome only (need all of the following elements)</li> <li>- Continuous: Mean of each group, effect size (difference between groups), 95% CI (some measure of uncertainty for these values)</li> <li>- Binary (# with/without event in each group, effect size, 95%</li> </ul>

	<p>CI</p> <ul style="list-style-type: none"> <li>- Effect size (comparison between groups) can be difference, ratio, odds, etc.</li> </ul> <p>Authors should present confidence intervals for the contrast between groups and as a measure of the precision (uncertainty) of the estimate of the effect.</p> <p>For abstracts not reporting the “primary” outcome of the trial (e.g. abstracts focusing on safety data or economic impacts), the secondary nature of the outcomes should be indicated, and, where possible, sufficient details of the primary outcome should be included to allow other findings to be taken in the proper context.</p>
<b>Harms: Important adverse events or side effects</b>	Authors should describe any important adverse (or unexpected) effects of an intervention in the abstract. If no important adverse events have occurred, the authors should state this explicitly.
<b>Conclusions: General interpretation of the results</b>	The conclusions of the trial, consistent with the results reported in the abstract, should be clearly stated along with their clinical application (avoiding over-generalisation). Authors should balance the benefits and harms in their conclusions. Where applicable, authors should also note whether additional studies are required before the results are used in clinical settings.
<b>Trial registration: Registration number and name of trial register</b>	<p>In an abstract reporting a trial, authors should provide details of the trial registration number and name of trial register.</p> <ul style="list-style-type: none"> <li>- Automatically completed as “notreported” if trial marked as not registered on registration tab OR registration info provided after email to author.</li> </ul>
<b>Funding: Source of funding</b>	<p>Authors should report the source of funding for the trial as this is important information for readers assessing a trial. Similarly, authors should report any other sources of support, such as in the preparation of the abstract, presentation, or manuscript.</p> <ul style="list-style-type: none"> <li>- Funding source needs to be in abstract in order to input “abstract”</li> </ul>

1. Hopewell S, Clarke M, Moher D, et al. (2008) CONSORT for Reporting Randomized Controlled Trials in Journal and Conference Abstracts: Explanation and Elaboration. PLoS Med 5:e20. doi: 10.1371/journal.pmed.0050020

2. Hopewell S, Clarke M, Moher D, et al. (2008) CONSORT for reporting randomised trials in journal and conference abstracts. Lancet 371:281–3. doi: 10.1016/S0140-6736(07)61835-2



### Appendix E: Adherence to CONSORT-Abstract reporting items in RCTs (breakdown by journal in 2016)

Abstract: Proportion of trials with the checklist item reported

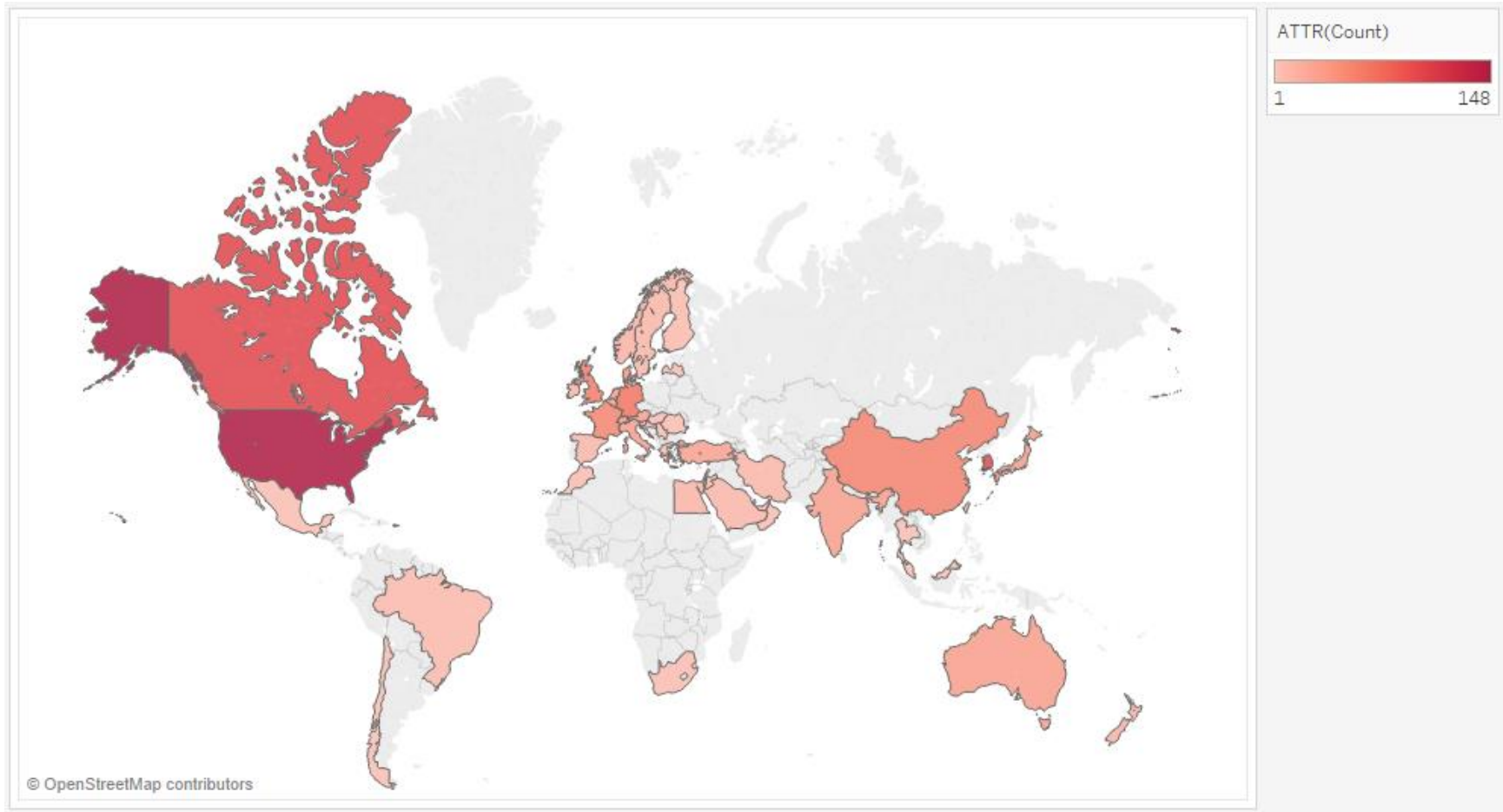
Full-text only: Of trials with the checklist item not reported in the abstract, the proportion of trials with the checklist item reported in the manuscript

A=Anaesthesia (n=38); B=Anesthesia & Analgesia (n=33); C=Anesthesiology (n=26); D=British Journal of Anaesthesia (n=34); E=Canadian Journal of Anesthesia (n=21); F= European Journal of Anaesthesiology (n=24)

*p-value*: Pearson's chi-squared test for differences between journals

	Abstract							Full-text only						
	A	B	C	D	E	F	<i>p</i>	A	B	C	D	E	F	<i>p</i>
Title	0.35	0.49	0.41	0.53	0.76	0.44	0.005	1	1	1	1	1	1	N/A
Trial Design	0.09	0.20	0.57	0.29	0.24	0.29	<0.001	0.56	0.50	0.39	0.67	0.32	0.60	<0.001
Methods														
Participants	0.20	0.04	0.33	0.41	0.05	0.25	<0.001	0.79	0.50	0.39	0.26	0.73	0.25	<0.001
Interventions	0.42	0.71	0.78	0.71	0.62	0.78	0.001	0.57	0.29	0.22	0.29	0.38	0.22	0.001
Objective	0.07	0.39	0.43	0.30	0.32	0.32	0.001	0.54	0.26	0.47	0.41	0.46	0.29	0.015
Outcome	0.19	0.38	0.49	0.41	0.46	0.29	<0.001	0.58	0.22	0.39	0.21	0.38	0.12	<0.001
Randomization	0.01	0.02	0.02	0.01	0	0.01	0.434	0.81	0.73	0.80	0.77	0.89	0.65	0.011
Blinding (masking)	0.19	0.44	0.49	0.44	0.30	0.40	0.001	0.70	0.46	0.39	0.44	0.65	0.41	<0.001
Results														
Numbers randomized	0.53	0.31	0.29	0.39	0.49	0.37	0.946	0.47	0.47	0.63	0.57	0.46	0.51	0.946
Recruitment	0.05	0.07	0.06	0.09	0.19	0.18	0.005	0.90	0.78	0.90	0.90	0.78	0.63	0.027
Numbers analyzed	0.15	0.14	0.27	0.17	0.24	0.12	0.053	0.85	0.76	0.71	0.83	0.73	0.75	0.053
Outcome	0.10	0.27	0.33	0.16	0.51	0.09	<0.001	0.17	0.16	0.27	0.09	0.14	0.04	0.088
Harms	0.37	0.57	0.51	0.64	0.54	0.60	0.038	0.41	0.30	0.31	0.24	0.16	0.34	0.003
Conclusions	0.72	0.98	1	0.97	0.95	0.97	<0.001	0.23	0.02	0	0.01	0.05	0.01	<0.001
Trial registration	0.01	0.01	0	0.53	0.54	0.35	<0.001	0.25	0.51	0.65	0.14	0.16	0.06	<0.001
Funding	0	0	0	0	0	0	N/A	0.84	0.82	1	0.76	0.81	0.69	<0.001

**Appendix F: Distribution of country origin for included RCTs**



## Curriculum Vitae

**Name:** Jeffrey Tin-Yu Chow

**Post-secondary Education and Degrees:** The University of Western Ontario  
London, Ontario, Canada  
2012–2016 B.MSc.

The University of Western Ontario  
London, Ontario, Canada  
2016–2017 M.Sc.

**Honours and Awards:** Ontario Graduate Scholarship (OGS)  
Ontario Ministry of Training, Colleges and Universities  
2016–2017

Western Graduate Research Scholarship (WGRS)  
The University of Western Ontario  
2016–2017

### Publications:

Gan XT, Rajapurohitam V, Xue J, Huang C, Bairwa S, Tang X, **Chow JT**, Liu MF, Chiu F, Sakamoto K, Wagner KU, Karmazyn M. Myocardial Hypertrophic Remodeling and Impaired Left Ventricular Function in Mice with a Cardiac-Specific Deletion of Janus Kinase 2. *Am J Pathol.* 2015;185(12):3202-3210. doi:10.1016/j.ajpath.2015.08.007. **Co-author**

**Chow JTY**, Lam K, Naeem A, Akanda ZZ, Si FF, Hodge W. The pathway to RCTs: how many roads are there? Examining the homogeneity of RCT justification. *Trials.* 2017;18:51. doi:10.1186/s13063-017-1804-z. **First Author**

Chow JTY, Hutnik CML, Solo K, Malvankar-Mehta MS. When is Evidence Enough Evidence? A Systematic Review and Meta-Analysis of the Trabectome as a Solo Procedure in Patients with Primary Open Angle Glaucoma. *J Ophthalmol.* 2017;2017:1-8. doi: 10.1155/2017/2965725. **First Author**

Jones PM, **Chow JTY**, Arango MF, Fridfinnson JA, Gai N, Lam K, Turkstra TP. Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals. *Anesth Analg*. 2017; *in-press*. **Co-author**

**Presentations:**

Gan XT, Rajapurohitam V, Huang C, Xue J, Chow J, Karmazyn M. Cardiac-Specific Deletion of Janus Kinase 2 (JAK2) Reveals Marked Cardiac Pathology Indicative of Hypertrophy, Heart Failure and Sudden Cardiac Death. Poster presentation at Frontiers in CardioVascular Biology 2014, Barcelona, Spain from July 4–6, 2014. **Co-author**

**Chow J**, Lam K, Akanda Z, Naeem A, Hodge W. The Pathway to RCTs, How Many Roads are There? Oral presentation at Western University Ophthalmology Research Day 2015, London, Ontario on November 6, 2015. **Presenter**

Rakibuz-Zaman M, **Chow J**, Si F, Hutnik C, Nicolela M, Crichton A, Birt C, Damji K, Sogbesan E, Gozdik D, Shuba L, Dorey M, Ford B, Hodge W. Repeat SLT v. ALT Interim Analysis of Demographics and Safety. Oral presentation at Western University Ophthalmology Research Day 2015, London, Ontario on November 6, 2015. **Co-author**

**Chow JT**, Jones PM, Arango MF, Fridfinnson JA, Gai N, Lam K, Turkstra TP. Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals. Oral presentation at Canadian Society of Epidemiology and Biostatistics 2017 Biennial Conference, Banff, Ontario from May 30–June 2, 2017. **Presenter**

Hutnik C, **Chow J**, Malvankar M. When is Evidence Enough Evidence? Oral presentation at Canadian Ophthalmology Society Annual Meeting 2017, Montreal, Quebec from June 15-18, 2017 **Co-author**

**Chow JT**, Jones PM, Arango MF, Fridfinnson JA, Gai N, Lam K, Turkstra TP. Comparison of registered and reported outcomes in randomized clinical trials published in anesthesiology journals. Oral presentation at Canadian Anesthesiologists' Society Annual Meeting 2017, Niagara Falls, Ontario from June 23–26, 2017. **Presenter**