AMERICAN UNIVERSITY OF BEIRUT

ASSESSING RISK OF BIAS ASSOCIATED WITH MISSING DICHOTOMOUS OUTCOME DATA IN META-ANALYSES: APPLICATION IN FIVE COCHRANE SYSTEMATIC REVIEWS

by
LARA ANDRE KAHALE

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science in Epidemiology
to the Department of Epidemiology and Population Health
of the Faculty of Health Sciences
at the American University of Beirut

Beirut, Lebanon
April 2015

AMERICAN UNIVERSITY OF BEIRUT

ASSESSING RISK OF BIAS ASSOCIATED WITH MISSING DICHOTOMOUS OUTCOME DATA IN META-ANALYSES: APPLICATION IN FIVE COCHRANE SYSTEMATIC REVIEWS

by
LARA ANDRE KAHALE

Approved by:

Dr. Elie Akl, Associate Professor Department of Internal Medicine Department of Epidemiology and Population Health	Advisor
Dr. Holger Schünemann, Professor and Chair Department of Clinical Epidemiology and Biostatistics McMaster University, Canada	Member of Committee
Dr. Monique Chaaya, Professor and Chair, Department of Epidemiology and Population Health	Member of Committee
Dr. Robert Habib, Professor Department of Internal Medicine	Member of Committee

Date of thesis defense: April 24, 2015

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Kahale (Last) Lara (First) Andre (Middle)

☑ Master's Thesis ○ Master's Project ○ Doctoral Dissertation

☐ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☑ I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project,** to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

Lara May 5, 2015

Signature Date

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and appreciation to my supervisor, mentor, and school Dr. Elie Akl, for his constant support and guide at both the academic and personal level. It has been an honor to work with him.

I would also like to thank Dr. Robert Habib for his bright ideas that always stimulated my critical thinking.

I would like to thank Dr. Monique Chaaya for her constant support and advice.

I would like to express my gratitude as well to Dr. Holger Schünemann for sharing his immense experience in health research methodology and providing critical feedback.

I am also grateful for my colleagues Maddalena Barbara, Irene Terrenato, Francesca Sperati, Elie Ramly and Ramy Ballout who helped with data abstraction.

A special thanks to my wonderful parents, brother, fiancée, and friends for their patience, support, and motivation throughout this challenging chapter of my life.

Above all, I thank God for everything**.**

AN ABSTRACT OF THE THESIS OF

Lara Andre Kahale for Master of Science
Major: Epidemiology

Title: Assessing risk of bias associated with missing dichotomous outcome data in meta-analyses: Application in five Cochrane systematic reviews

Background:
Missing participant data relates to trial participants for whom outcome data are not available for systematic review authors. There is no consensus on how systematic review authors should assess risk of bias associated with missing data for a given meta-analysis. One proposed approach is to evaluate the impact of different assumptions about missing data on the pooled effect estimate.

Objective:
To assess how different assumptions about the outcome of participants with missing data alters statistically significant pooled effect estimates of patient-important dichotomous outcomes in five Cochrane systematic reviews.

Methods:
We conducted this study using a series of five recently updated Cochrane systematic reviews addressing different clinical questions about anticoagulation in patients with cancer. We considered patients with missing data those described as having withdrawn consent, being lost to follow-up or having outcome not assessable. We focused on outcomes for which the primary meta-analysis, a complete case analysis, revealed statistically significant pooled effect estimates. We applied nine assumptions about the outcome of participants with missing data. Four of these assumptions are commonly used (e.g., best case scenario and worst case scenario). The remaining five assumptions are considered more plausible as they are based on incidences observed among participants followed-up in the trials. We assessed the number of assumptions under which each pooled effect estimate loses significance and changes direction.

Results:
We included 12 outcomes that had statistically significant results in the complete case analysis. The impact of the common assumptions varied significantly with no change for two of them (best case scenario and 'none had the event') and major change for the other two ('all had the event' and, particularly, worst case scenario). Under the plausible assumptions (based on the $RI_{LTFU/FU}$), an increasing number of outcomes, up to five, lost statistical significance, with one changing direction.

Discussion:
The impact of different assumptions about the outcome of participants with missing data on pooled effect estimates will help judge the associated risk of bias for a given meta-analysis. Our findings will inform recommendations on how to test the robustness of meta-analytical results in the presence of missing data.

CONTENTS

ILLUSTRATIONS

TABLES

ABBREVIATIONS

AC	AntiCoagulation
BCS	Best Case Scenario
CCA	Complete Case Analysis
CI	Confidence Interval
CVC	Central Venous Catheter
DVT	Deep Vein Thrombosis
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
IM	Informatively Missing
IPD	Individual Participant Data
IQR	InterQuartile Range
LMWH	Low Molecular Weight Heparin
MAR	Missing At Random
MCAR	Missing Completely At Random
MNAR	Missing Not At Random
MPD	Missing Participant Data
OR	Odds Ratio
PE	Pulmonary Embolism
QoE	Quality of Evidence

RCT	Randomized Controlled Trial
RI	Relative Incidence
RoB	Risk of Bias
RR	Relative Risk
SD	Standard Deviation
SR	Systematic Review
UFH	UnFractionated Heparin
VKA	Vitamin K Antagonist
VTE	Venous ThromboEmbolism
WCS	Worst Case Scenario

CHAPTER I

LITERATURE REVIEW

A. Systematic reviews and meta-analyses

Decision making by health care professionals and health policy makers has been challenged by the ever growing accumulation of evidence. In 1992, more than two million articles were published annually in over 20,000 journals; about 4,400 pages were devoted to approximately 1,100 articles in the British Medical Journal (BMJ) and New England Journal of Medicine, combined (Mulrow, 1994). A decade later, more than 8,200 articles were published in March 2002 in 341 journals of five primary care journal review services (ACP Journal Club, DynaMed, Evidence-Based Practice, Journal Watch, and QuickScan Reviews) (Alper et al., 2004). In 2010, the new journal PLoS ONE celebrated its 10,000th article since its inception in December 2006; and in 2013 only, it published 31,000 articles (Campbell, 2014; Van Gemert, 2010). Today, the MEDLINE database includes more than 21 million references to articles published in approximately 5,600 current biomedical journals from more than 80 countries (National Library of Medicine, 2014).

The need to summarize the unmanageable amounts of information accurately and reliably paved the way to the development and wide adoption of systematic review methodology. A systematic review is based on research methodology that aims to identify, select, appraise, and synthesize all research evidence relevant to a specific research question (Cook, Mulrow, &

Haynes, 1997) It also filters the biomedical/health literature from the insignificant, unreliable, or redundant information and sheds the light on the solid and critical studies that are worthy of reflection (Mulrow, 1994). Archie Cochrane, one of the pioneers of systematic reviews, stated in his influential book "Effectiveness and Efficiency. Random Reflections on Health Services": "It is surely a great criticism of our profession that we have not organized a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomized controlled trials"(Cochrane, 1973). His call for a collection of systematic reviews led to the foundation of evidence-based health care and consequently the creation of the Cochrane Collaboration in 1993 (Shah & Chung, 2009). The mission of the Cochrane Collaboration is to synthesize and disseminate up-to-date review of randomized controlled trials (RCTs) of healthcare interventions to guide:

- Healthcare professionals in clinical decision making (Cochrane Community, 2014);
- Researchers in refining hypotheses and avoiding pitfalls of previous work (Mulrow, 1994);
- Health policy makers in formulating guidelines and legislation concerning the use of certain diagnostic tests and treatment strategies (Dickersin & Berlin, 1992).

Meta-analysis refers to the statistical methods applied to combine the results of the individual studies addressing the same research question (Higgins & Green, 2011). It is usually presented in a forest plot which summarizes results from individual trials included in the meta-analysis in a visual presentation (Juni & Egger, 2009). A landmark cumulative meta-analyses and its benefits is exemplified by the Oxford Database of Perinatal Trials, shown in Figure 1

(Antman, Lau, Kupelnick, Mosteller, & Chalmers, 1992). The figure shows the odds ratios and 95% confidence intervals for 17 trials comparing the effect of intravenous streptokinase and placebo or no therapy on mortality in patients who had been hospitalized for acute myocardial infarction. The left side of the figure shows that the streptokinase was protective against mortality in 14 out of the 17 trials, however only two achieved statistical significance. At the bottom, it is shown that the overall pooled effect estimate significantly favored treatment. The right side of the figure shows the cumulative meta-analysis of the same data, i.e., as if a new meta-analysis was performed each time the results of a new trial were reported (Mulrow, 1994). The first shade of statistical significance of treatment effect was found in year 1971 (with a two sided p-value of < 0.05). Then, the significance kept on refining up to year 1977 with a p-value of < 0.001 (Mulrow, 1994). The cumulative meta-analyses showed that intravenous streptokinase could have been proved to prevent mortality almost 20 years before the United States Food and Drug Administration approved it and consequently adopting it in clinical practice (Antman et al., 1992; Mulrow, 1994) Thus, meta-analyses can reduce the unnecessary time lost between new medical research discoveries and clinical implementation of effective diagnostic or treatment strategies

According to the Cochrane collaboration, certain study designs are more appropriate than others for answering particular questions (Higgins & Green, 2011). However, systematic reviews addressing questions about the effects of health care should focus primarily on RCTs. That is because randomization is the only way to eliminate the selection bias and balance

between the different interventions groups with respect to both known and unknown (or unmeasured) confounders or prognostic variables (Higgins & Green, 2011; Suresh, 2011).

B. MPD in randomized controlled trials

RCTs are expected to produce unbiased estimates of treatment effects compared to other study designs (Wood, White, & Thompson, 2004). Problems in the design and conduct of RCTs may present threats to the validity of their results. In practice, achieving this goal depends on the extent to which potential sources of bias have been avoided or minimized. For several reasons, RCTs might suffer from missing outcome data for a number of participants (Akl et al., 2012). Indeed, 60-89% of randomized controlled trials published in of the top five general medical journals have some missing participant data (MPD) (Akl et al., 2012). One study found that one quarter of 93 Health Technology Assessments trial suffered from MPD (Sylvestre, 2011). Similarly in mental health trials, it was found that MPD exceeded a rate of 50% for certain conditions (Mavridis, Chaimani, Efthimiou, Leucht, & Salanti, 2014). Despite persistent attempts by trial investigators to prevent incomplete outcome data, this phenomenon cannot be entirely eliminated (Altman, 2009).

The mechanism of MPD can be classified into three categories (Mavridis et al., 2014):

- *Missing completely at random:* the reason of missingness is related neither to participants' characteristics nor to the outcome, e.g., if a participant misses some appointments due to scheduling difficulties. This assumption means that the group of participants who provided data is a random sample of the total population.

- *Missing at random:* the reason of missingness is related to participants' characteristics but not the actual outcome, e.g., primary school children are randomized to different intervention groups to reduce school-related anxiety. Younger children are less likely to provide outcome data due to their age-related cognitive challenges. Thus, rates of MPD among younger children across groups are expected to be comparable, and consequently the outcomes for the younger children who dropped out are expected to be similar to outcomes for the younger children who completed the study.
- *Missing not at random (MNAR) or informatively missing (IM):* the reason of missingness is associated with the actual effect of the intervention, e.g., in mental health trials, placebo groups show larger dropout rate than patients treated with antipsychotics because of placebo's lack of efficacy. Thus, the effect estimate of the relative treatment would be biased when the analysis is based only on participants who completed the study.

In order to preserve the prognostic balance created by randomization, the intention-to-treat principle calls for trialists to include all randomized participants in the group to which they were allocated in superiority trials (Montori & Guyatt, 2001). The CONSORT (Consolidated Standards of Reporting Trials) statement is a set of guidelines that was established to improve the quality of reporting of clinical trials. It recommends intention-to-treat analysis as standard practice (Moher et al., 2012; Rennie, 2001). Though this principle is frequently applied, the intention-to-treat principle does not protect against bias associated with MPD (Montori & Guyatt, 2001). Indeed, MPD is still present in one quarter of RCT reports, and is more poorly reported than items specifically listed in CONSORT (Sylvestre, 2011). Moreover, one would

have to make assumptions about the outcomes of participants with MPD in order to include them in the analysis (Alshurafa et al., 2012). A common practice by most trials investigators is the inclusion participants with MPD in the denominators while calculating estimates of effect. This approach assumes that none of those participants with MPD experienced the outcome of interest. Consequently, reporting results of the effect of the intervention may be misleading given that this assumption is highly unlikely.

C. MPD in systematic reviews

By definition, systematic reviewers, similarly to trialists, do not know the actual outcomes of participants with MPD. The Cochrane handbook encourages systematic review authors to conduct intention-to-treat analysis by including all randomized participants in the analysis (Higgins & Green, 2011). The handbook, however, does not provide sufficient detailed guidance on how such analyses should be conducted. While proposals on how to address this issue exist, they are statistically sophisticated and may be challenging for common use (Ebrahim et al., 2013). An important challenge with abstracting data from RCT reports is that results of RCTs are usually presented together for fully observed and imputed outcomes (Mavridis et al., 2014). Consequently, systematic review authors conducting meta-analyses are not given choice but to synthesize outcome results as reported in RCTs, even when the imputation technique is inappropriate (Mavridis et al., 2014). In some cases, RCT reports present the outcomes for completers only as well as the results from the merged sample of observed and imputed outcomes (Mavridis et al., 2014). For example, an RCT report might indicate that a certain

number of participants withdrew consent, without indicating whether or not they were followed-up. Bias associated with MPD can be translated from RCTs into systematic reviews including these trials in the meta-analysis. Thus, MPD creates a serious problem in systematic reviews because MPD can bias estimates of the treatment effect and reduce statistical power.

D. Risk of bias associated with MPD

A crucial issue for all systematic reviews authors is the extent to which risk of bias associated with MPD reduces the confidence in results. The Cochrane Collaboration's Risk of Bias (RoB) tool was designed to help in assessing bias associated with a number of factors including random sequence generation, allocation concealment, blinding of personnel, participants, outcome adjudicators, and data analysts, selective reporting, and incomplete outcome data (Higgins & Green, 2011). The latter factor is likely the least developed component of the RoB tool. Indeed, a recently published study assessing stakeholders' experiences with and perceptions of the Cochrane RoB tool participants found that incomplete outcome data as one of the most difficult domains to assess (Savovic et al., 2014). They also requested more guidance on how to incorporate RoB assessments into meta-analyses and conclusion (Savovic et al., 2014).

Historically, arbitrary thresholds for acceptable rate of MPD have been suggested by different methodologists, e.g., less than 20%. However, the significance of particular rates of MPD is not associated with the magnitude of the rate; instead, it is highly dependent on the MPD mechanism which describes how propensity for MPD depends on the participant's characteristics and outcomes (Guyatt, Oxman, Vist, et al., 2011; Mavridis et al., 2014). Several approaches

have been proposed for dealing with MPD in systematic reviews. These approaches recommend conducting a complete case analysis for the primary analysis and some form of sensitivity analysis to evaluate robustness of results (Akl et al., 2013; Gamble & Hollis, 2005; J. P. Higgins, White, & Wood, 2008; Mavridis et al., 2014).

Moreover, the significance of MPD rates is highly associated with the rate of outcome event among the participants with observed available data (Guyatt, Oxman, Vist, et al., 2011). For example, MPD rate of 5% in both intervention and control groups would bring about little threat of bias if event rates were 20% and 40% in the two groups, respectively. However, if the event rates were 2% and 4%, MPD rate with 5% is much more problematic and creates more risk. This concludes that risk of bias is greater when the rate of MPD in relation to groups' event rates is high (Guyatt, Oxman, Vist, et al., 2011). Despite the relatively high rates of MPD, bias will result only if the rate of MPD missing data is imbalanced between groups or the association between MPD and the likelihood of events varies between the groups (Guyatt, Oxman, Vist, et al., 2011). However, because details about the latter assumption is not available from RCT reports (i.e. whether the relationship between MPD and the likelihood of events differs in intervention and control groups); high rates of MPD creates a serious threat of bias (Guyatt, Oxman, Vist, et al., 2011).

A recent study found that, almost one in every three RCTs with statistically significant results lost statistical significance when making plausible assumptions about the outcomes of participants with MPD (Akl et al., 2012). This reduces our confidence in the effect estimates resulting not only from these trials, but also from systematic reviews including these studies.

Thus, conducting sensitivity analyses based on different assumptions regarding the outcomes of participants with MPD may test the robustness of the results (i.e. the extent of risk of bias associated with MPD) (Akl et al., 2013; Ebrahim et al., 2013; Ebrahim et al., 2014).

E. Quality of evidence associated with MPD

Summarizing and assessing the quality of evidence generated by systematic reviews is crucial for decision makers including, clinicians, practice guidelines panelists, and policy makers. The ‘‘Grades of Recommendation, Assessment, Development, and Evaluation’’ (GRADE) approach provides guidance for rating quality of evidence, grading strength of recommendations in health care, and consequently moving from the evidence to a recommendation or decision (Guyatt, Oxman, Schunemann, Tugwell, & Knottnerus, 2011). In summary, the rating of the quality of evidence highly depends on the following criteria:

- *Design of studies included (specifically in relation to randomization)*: considering RCTs as high-quality evidence, whereas observational studies as low-quality evidence (Guyatt, Oxman, Schunemann, et al., 2011);
- *Risk of bias:* presenting an approach similar to the Cochrane RoB tool (sequence generation (selection bias), allocation sequence concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective outcome reporting (reporting bias) and other potential sources of bias (Guyatt, Oxman, Vist, et al., 2011; Higgins & Green, 2011);

- *Imprecision:* focusing on the consideration of confidence intervals around point estimates associated with each outcome (Guyatt, Oxman, Kunz, Brozek, et al., 2011);
- *Inconsistency:* addressing similarity of point estimates, the extent to which confidence intervals overlap, and the available statistical tests related to heterogeneity between study results (Guyatt, Oxman, Kunz, Woodcock, Brozek, Helfand, Alonso-Coello, Glasziou, et al., 2011);
- *Indirectness:* referring first to differences between the population, intervention, and outcome addressed in the included studies and those of interest to systematic review authors and guideline developers. Second, it refers to indirect comparisons in which one is interested in recommending between two agents that have each been tested against a third comparator, but not directly against each other (Guyatt, Oxman, Kunz, Woodcock, Brozek, Helfand, Alonso-Coello, Falck-Ytter, et al., 2011);
- *Publication bias:* referring to when studies without statistically significant results are less likely to be published than studies with statistically significant results (Guyatt, Oxman, Montori, et al., 2011);
- *Upgrading criteria:* is considered when methodologically rigorous observational studies show at least a two-fold reduction or increase in risk, and rating up two levels for at least a five-fold reduction or increase in risk. It is also considered in presence of a dose-response gradient, or a conclusion that plausible residual confounding would further support inferences regarding treatment effect (Guyatt, Oxman, Sultan, et al., 2011).

Through direct and indirect ways, MPD can contribute to the imprecision of results and amplifying the risk of bias, specifically attrition bias, leading ultimately to downgrading the quality of evidence.

F. Significance of topic

As discussed above, MPD can bias estimates of the treatment effect and reduce statistical power. A number of authors have proposed approaches for how systematic review authors should deal with trial participants with MPD when conducting meta-analyses. These approaches recommend conducting a complete case analysis for the main analysis and some form of sensitivity analysis to evaluate robustness of results. There is a lack of evidence about the impact of different approaches for dealing with MPD on pooled effect estimates. Consequently, the extent to which results of systematic reviews are vulnerable to MPD remains uncertain.

A protocol of a large methodological study has been planned to test the recommended approaches to dealing with MPD in systematic reviews (Akl, Kahale, Agarwal, et al., 2014). This study includes a representative sample of 100 Cochrane and non-Cochrane systematic reviews. Thus, this thesis project was designed as a pilot testing study for the larger methodological study. The findings of this pilot study would help inform the methodology for judging the risk of bias associated with MPD in systematic reviews.

G. Objective

The main objective is to assess how different assumptions about the outcome of participants with MPD alter statistically significant pooled effect estimates of patient-important dichotomous outcomes in five Cochrane systematic reviews. The ultimate aim is to develop the methodology for assessing the risk of bias associated with MPD.

CHAPTER II

METHODS

A. Definitions

MPD: outcome data of trial participants that are not available to the systematic reviewer authors (from the published RCT reports or personal contact with trial authors) for inclusion in their meta-analyses. MPD do not relate to any of the following:

- Missing studies (e.g., unpublished studies);
- Entire unreported outcomes (e.g., outcomes planned in trial protocols but not included in trial reports).

Cochrane systematic reviews: systematic reviews published in the Cochrane Database of Systematic Reviews.

Meta-analysis: the statistical synthesis of results from a series of studies collected systematically (Borenstein, Hedges, Higgins, & Rothstein, 2011).

A patient-important outcome: An outcome that matters to patients when they make decisions about interventions that affect their health, including pharmaceutical, behavioral or tests. We will use a previously developed hierarchy of outcomes for the selection of one outcome of interest as shown in Appendix 1 (Akl et al., 2009). Categories I, II, and III include patient-important outcomes. Category IV includes surrogate outcomes, which are not considered as patient-

important. For a composite outcome to be considered as patient-important, we will require all of its components to be patient-important (Akl et al., 2009).

Complete Case Analysis: only individuals whose outcome is known are included in the analysis; usually the point reference approach in many meta-analyses (Mavridis et al., 2014). Components of the complete case analysis consist of:

- Denominator: [number of participants randomized] – [number of participants most likely with MPD, both pre and post-intervention initiation] (see below section: "Challenge of identifying participants with missing participant data ");
- Numerator: number of participants with observed events (i.e., participants who suffered at least one event for the outcome of interest during their available follow-up time).

Sensitivity analysis: defined by the Cochrane Handbook as "a repeat of the primary analysis or meta-analysis, substituting alternative decisions or ranges of values for decisions that were arbitrary or unclear" (Higgins & Green, 2011).

B. Research design

We conducted an imputation study by carrying out a main meta-analysis, and several sensitivity analyses using different assumptions about outcome of participants with MPD. For feasibility issues, we conducted this study as we were updating a series of five Cochrane systematic reviews addressing different clinical questions about anticoagulation in patients with cancer. Topics covered by the series included:

- Parenteral anticoagulation for survival benefit (venous thromboembolism (VTE) thromboprophylaxis trials in ambulatory patients with cancer) (Akl, Kahale, Ballout, et al., 2014);
- Oral anticoagulation for survival benefit (Akl, Kahale, Terrenato, et al., 2014);
- Central venous catheters thromboprophylaxis (Akl, Ramly, et al., 2014);
- Perioperative thromboprophylaxis (Akl, Kahale, Sperati, et al., 2014);
- Long-term anticoagulation treatment of VTE (Akl, Kahale, Barba, et al., 2014)

The study sample consisted of all meta-analyses revealing statistically significant pooled effect estimates of patient-important outcomes of dichotomous measures.

C. Eligibility criteria

We restricted to quantitative data, thus semi-quantitative or qualitative synthesis was excluded. Inclusion criteria for a meta-analysis that meets all the following criteria:

- Reported an effect estimate expressed as a dichotomous measure (including relative risk (RR) or odds ratio (OR); excluding those produced by generic inverse variance method);
- Reported a patient-important outcome;
- Reported a statistically significant pooled effect estimate under complete case analysis from at least two trials; statistical significance refers to p-value < 0.05 or confidence interval (CI) not including 1.0.

We focused on systematic reviews with a statistically significant pooled effect estimate for a patient-important efficacy outcome because they are the most likely to influence clinical practice.

D. Data abstraction

In reference to the outcome addressed in the eligible meta-analysis, we collected the following information:

- Number of trials included;
- Numerator and denominator used in the meta-analysis for each group for each trial;
- Categories of participants that could potentially be counted as having MPD:
  - "mistakenly randomized or inappropriately excluded or ineligible participants";
  - "did not receive any treatment";
  - "withdrew consent";
  - "outcome not assessable";
  - "dead";
  - "experienced adverse events";
  - "non-compliant";
  - "discontinued prematurely";
  - "cross-over";
  - "moved out of country";

  - "lost to follow-up" (lost to follow-up for reasons not considered in our other categories);
  - Other similarly described categories.
- Number of participants within each of the above category;
- Pooled relative effect measure (RR or OR) under complete case analysis and its associated CI, p-value, and measure of heterogeneity (I^2);
- Analysis model used (i.e., random effect or fixed effect);
- Statistical method used (e.g., Mantel-Haenszel, or Peto);

Table 1 presents the numerical details that are used to deal with MPD in the sensitivity analyses.

E. Data analysis

1. *Challenge of identifying participants with MPD*

During the process of data abstraction, we noticed that different RCTs report participant categories differently in terms of MPD (e.g., those described as "withdrew consent" or "experienced adverse events"). Also, we found it challenging to ascertain which categories the trialists actually followed-up (i.e., did not suffer MPD) and which categories they did not (i.e., suffer MPD).

For example, figures 2 and 3 are taken from one of the included RCTs published in a high impact journal and correspond respectively to the study flow diagram and the outcomes frequency table (Agnelli et al., 2009). Figure 2 shows that in the intervention group 12

participants were “lost to follow-up “and 57 “withdrew consent”. Figure 3 shows that all these participants were included in the analyses (See ‘N’ in the top row). So in fact, the trialists managed to include participants who definitely had MPD (“lost to follow-up” in this case) in the analysis. In the absence of any further clarification (whether in the table or in the text), it is possible that participants who “withdrew consent” were actually followed up and subsequently included in the analysis. However, it is also possible that those who withdrew consent had premature end of follow-up (consequently MPD) but the trialists still managed to include them in the analysis, in the same way they included “lost to follow-up participants”. Thus, while examining that trial report, we could not know whether that group of participants who “withdrew consent” did or did not have MPD. The above scenario would introduce bias when analyzing data. Indeed, there is a risk of double counting of events under the following scenario: in case those participants actually had (and were counted by trialists as having had) the outcome of interest, but the systematic reviewers count them as having MPD, and make an assumption that participants with MPD had the outcome of interest. To deal with this issue, we made the following considerations:

- “mistakenly randomized or inappropriately excluded or ineligible participants” and “did not receive any treatment” participant categories, which are defined prior to the initiation of the study intervention (pre-intervention), most likely have MPD;
- “withdrew consent”, “outcome not assessable”, “moved out of country”, and “lost to follow-up” participant categories, which are defined after the initiation of the study intervention (post-intervention), most likely have MPD;

- “dead”, “experienced adverse events”, “non-compliant”, “discontinued prematurely” and “cross-over” (and similarly described) participant categories, less likely have MPD.

2. *Size of categories of participants with potential MPD*

We conducted a descriptive analysis of the different reporting aspects of MPD. To estimate the extent of MPD in the RCTs, we calculated for each RCT included in the eligible meta-analyses, the percentage of participants belonging to each of categories of participants that could potentially be counted as having MPD. We calculated the mean, standard deviation (SD), and interquartile range (IQR) of these percentages of participants belonging to each of the categories across all the RCTs included.

3. *Impact of MPD on effect estimates*

We evaluated the effect of several assumptions about the outcomes of participants with MPD data on the effect estimates of the eligible meta-analyses. For each eligible meta-analysis, we conducted several sensitivity analyses to assess the risk of bias associated with MPD. Those sensitivity analyses used nine assumptions about the outcomes of participants with MPD and evaluated whether these assumptions altered the effect estimates. Four of these assumptions are commonly used:

- None of the participants with MPD in both intervention and control groups had the outcome of interest;

- All of the participants with MPD in both intervention and control groups had the outcome of interest;
- Best case scenario: all participants with MPD in the intervention group had a favorable event of the outcome of interest, and all participants with MPD in the control group had an unfavorable event of the outcome of interest;
- Worst case scenario: all participants with MPD in the intervention group had an unfavorable event of the outcome of interest, and all participants with MPD in the control group had a favorable event of the outcome of interest.

Although the above assumptions are commonly used in the literature, they are not plausible (White, Horton, Carpenter, & Pocock, 2011). Thus, we consider the remaining five and increasingly stringent assumptions more plausible as they are based on incidences observed among participants with complete follow-up data. The increasingly stringent order is chosen in order to progressively challenge the statistical significance of the results of the primary analysis (Akl et al., 2013; Ebrahim et al., 2013).

For relative risk (RR) showing a reduction in effect (RR<1), we will use the following five increasingly stringent but plausible assumptions (Akl et al., 2013; Ebrahim et al., 2013) about the relative incidence (RI) among those with MPD (lost to follow-up (LTFU)) compared to those with available data (followed-up (FU)) in the same group:

- For the control group, $RI_{LTFU/FU}$ =1 for the intervention group, $RI_{LTFU/FU}$ = 1;
- For the control group, $RI_{LTFU/FU}$ = 1; for the intervention group, $RI_{LTFU/FU}$ = 1.5;
- For the control group, $RI_{LTFU/FU}$ = 1; for the intervention group, $RI_{LTFU/FU}$ = 2;

- For the control group, $RI_{LTFU/FU} = 1$; for the intervention group $RI_{LTFU/FU} = 3$;
- For the control group $RI_{LTFU/FU} = 1$; for the intervention group, $RI_{LTFU/FU} = 5$.

Alternatively, for RR showing an increase in effect (RR>1), we switched the above assumptions between the control and interventions groups (i.e., used $RI_{LTFU/FU} = 1$ for the intervention group) as follows:

- For the control group, $RI_{LTFU/FU} = 1$ for the intervention group, $RI_{LTFU/FU} = 1$;
- For the control group, $RI_{LTFU/FU} = 1.5$; for the intervention group, $RI_{LTFU/FU} = 1$;
- For the control group, $RI_{LTFU/FU} = 2$; for the intervention group, $RI_{LTFU/FU} = 1$;
- For the control group, $RI_{LTFU/FU} = 3$; for the intervention group $RI_{LTFU/FU} = 1$;
- For the control group $RI_{LTFU/FU} = 5$; for the intervention group, $RI_{LTFU/FU} = 1$.

Table 2 presents the numerical details of the different assumptions that are used to deal with MPD. We used the following calculations for each study group:

- Denominator: (number of participants randomized) - (number of participants most likely with MPD, pre-intervention initiation);
- Numerator: (number of participants with observed events) + (number of participants most likely with MPD, post-intervention initiation, with assumed events).

Assumed events are calculated by applying the a priori assumptions to the participants considered most likely with MPD post-intervention initiation.

Each of these methods generated a set of values for the numerator and denominator in each group of each RCT included in the meta-analysis. We pooled the generated data of all RCTs in the Review Manager software to conduct the sensitivity meta-analyses (The Nordic

Cochrane Centre, 2014). We used the same pooled relative effect measure, the same analysis model, and the same statistical method used in the main (complete case) meta-analysis to generate the effect estimate of the sensitivity analyses. The term "assumption pooled effect estimates" is used to refer to the results of those sensitivity analyses.

We presented the results of the sensitivity analyses graphically in a color-coded table. The table showed how each of the nine previously described assumptions affected the statistical significance of the pooled effect estimates. It also showed the relative risks and their corresponding confidence intervals. The color coding is as follows:

- Green color refers to the pooled effect estimates that maintained direction and significance;
- Yellow color refers to the pooled effect estimates that maintained direction however lost significance;
- Purple color refers to the pooled effect estimates that changed direction and lost significance;
- Red color refers to the pooled effect estimates that changed direction and became significant.

We then calculated the percentage of effect estimates that were no longer significant for each of the assumptions. We started with the worst-case scenario assumption. If the assumption pooled effect estimates are robust to that assumption, then we would be assured and interpret that the risk of bias associated with MPD is low. This approach is recommended especially when there are relatively few participants with MPD (Akl et al., 2013). Afterwards, we examined the

impact of the other plausible assumptions. We judged the risk of bias associated with MPD to be low in case the estimates remain similar under different assumptions. To the extent that estimates remain similar when making the assumptions, the lower the risk of bias associated with MPD.

We were also interested to find what variables could predict the probability of effect estimates losing their significance. Thus, we conducted a multiple logistic regression using STATA version 10 software with the "losing statistical significance of effect estimates under an assumption with high variability of results" as the dependent variable and the following general and methodological characteristics as the independent variables:

- *Rate of MPD:* defined as the number of participants most likely to have MPD divided by the total number of randomized participants for each eligible meta-analysis;
- *Rate of participants with observed events:* defined as the number of participants with observed events divided by the total number of randomized participants for each eligible meta-analysis;
- *Number of trials included in each eligible meta-analysis.*

These covariates were selected since we think they might be related to the possibility of losing significance.

F. Ethical consideration

Since this study involves no human subjects, no Institutional Board Review approval was sought.

CHAPTER III

RESULTS

A. Description of eligible meta-analyses

The sample of the five systematic reviews included 58 trials and 75 meta-analyses for patient-important dichotomous outcomes. Of those 75 meta-analyses, 12 had pooled effects that remained significant under complete case analysis. Table 3 displays the general characteristics of the systematic reviews.

- The first systematic review entitled "*Oral* anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation" (*oral* systematic review) compared vitamin K antagonists to no intervention or placebo and included seven RCTs with 1,770 participants who completed follow-up (Akl, Kahale, Terrenato, et al., 2014). This systematic review conducted nine meta-analyses under the complete case analysis where only two safety outcomes were significantly associated with the intervention: major bleeding (RR 4.24; 95%CI 1.86, 9.65) and minor bleeding (RR 3.19; 95% CI 1.83, 5.55).
- The second systematic review entitled "Low molecular weight heparin versus unfractionated heparin for *perioperative* thromboprophylaxis in patients with cancer" (*perioperative* systematic review) compared low-molecular weight heparin to unfractionated heparin and included 16 RCTS with the 12,890 participants who

completed follow-up (Akl, Kahale, Sperati, et al., 2014). Of 13 different outcomes evaluated, only one safety outcome yielded a significant effect estimate under complete case analysis: wound hematoma (RR0.68, 95%CI 0.43, 0.84).

- The third systematic review entitled "Anticoagulation for the *long-term* treatment of venous thromboembolism in patients with cancer" (*long-term* treatment systematic review) compared low-molecular weight heparin to vitamin K antagonists included 10 RCTs with 1,981 participants with complete follow-up (Akl, Kahale, Barba, et al., 2014). This systematic review evaluated 16 outcomes of which one was found significant under the complete case analysis: recurrent venous thromboembolism (RR 0.50; 95%CI 0.35, 0.71).
- The fourth systematic review entitled "*Parenteral* anticoagulation in ambulatory patients with cancer" (*parenteral* systematic review) comparing low-molecular weight heparin to no intervention or placebo included 15 RCTs with 7,662 participants with complete follow-up (Akl, Kahale, Ballout, et al., 2014). Of the 18 different meta-analyses conducted, three efficacy outcomes and one safety outcome yielded significant effect estimates under the complete case analysis: symptomatic venous thromboembolism (RR 0.56; 95%CI 0.42, 0.74), symptomatic deep vein thrombosis (RR 0.49; 95%CI 0.28, 0.86) , symptomatic pulmonary embolism (RR 0.59; 95%CI 0.37, 0.96) , and minor bleeding (RR 1.32; 95%CI 1.02, 1.71).
- The fifth systematic review entitled "Anticoagulation for people with cancer and *central venous catheters*" (*central venous catheter* systematic review) included a total of 10

RCTs which reported on three different comparisons: (1) low-molecular weight heparin versus no intervention or placebo, (2) vitamin K antagonists versus no intervention or placebo, and (3) low-molecular weight heparin versus vitamin K antagonists (Akl, Ramly, et al., 2014). Six RCTs reported on the first comparison with 1,448 participants with complete follow-up; only one efficacy outcome out of seven outcomes was significantly associated with the intervention under the complete case analysis: symptomatic deep vein thrombosis (RR 0.48; 95%CI 0.27,0.86). Regarding the second comparison, five RCTs were included 666 participants with complete follow-up; one efficacy outcome out of seven outcomes significantly associated with the intervention under the complete case analysis: asymptomatic deep vein thrombosis (0.43; 95%CI 0.30, 0.62). The third comparison included three RCTs with 620 participants with complete follow-up; one efficacy and one safety outcomes out of seven outcomes yielded significant results: asymptomatic deep vein thrombosis (RR 1.74;95%CI 1.20, 2.52) and thrombocytopenia (RR 3.73;95%CI 2.62, 6.16).

B. Size of categories of participants with potential MPD

Table 4 shows the mean, standard deviation (SD), and interquartile ranges (IQR) for the percentage of participants belonging to each category (described above) across the included trials. On average and across both groups these percentages were:

- 1.15% (SD 2.72; IQR 0.00-1.30%) for participants most likely had MPD due to reasons prior to treatment initiation;

- 4.92% (SD 8.89; IQR 0.00-7.22%) most likely MPD due to reasons post treatment initiation;
- 7.37% (SD 12.66; IQR 0.00-10.74%) less likely had MPD.

Thus, the overall percentage was 13.60% (SD 17.14; IQR 0.00-24.78%). The mean percentage of participants who less likely had MPD was significantly higher in the intervention group (5.39%) compared to the control group (4.40%) with a p-value of 0.003.

We present in table 5 the general characteristics of the 12 outcomes which remained significant under complete case analysis. For each outcome, we show the rate of participants with observed events and the rate of participants with MPD. The rate of participants with observed events ranged from 1.16% to 22.65%. The lowest rate occurred for symptomatic pulmonary embolism outcome in the *parenteral* systematic review which enrolled a large number of participants (6,493) in nine RCTs. On the other hand, the highest rate of observed events occurred in thrombocytopenia outcome in *central venous catheter* systematic review which included the lowest number of participants (340) in only two RCTs. The rate of participants most likely had MPD ranged from 0.8% to 13.7%. For some outcomes, the rate of MPD was equal to or exceeded the rate of participants with observed events. Some of these outcomes include major bleeding outcome in the *oral* systematic review (7.64% and 6.71% respectively), symptomatic deep vein thrombosis (DVT) outcome in the *parenteral* systematic review (3.64% and 2.32% respectively), and symptomatic DVT in the *central venous catheter* systematic review comparing low-molecular weight heparin versus placebo (13.73% and 3.55% respectively).

All meta-analyses reported relative risks using random effect analysis model and Mantel-Haenszel statistical method. We presented the quality of evidence of each outcome as per the GRADE approach. The outcomes that were graded by a high rating were symptomatic venous thromboembolism, symptomatic deep vein thrombosis, and symptomatic pulmonary embolism outcomes in the *parenteral* systematic reviews. All three outcomes had similar low rates of MPD (3.4%, 3.64/%, and 3.65% respectively). Six outcomes were graded as having moderate quality of evidence: major and minor bleeding in *oral* systematic reviews, wound hematoma in *perioperative* systematic review, recurrent venous thromboembolism in *long-term* systematic reviews, minor bleeding in *parenteral* systematic review, and symptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to placebo.

C. Impact of MPD on pooled effect estimates

As an example, we present in appendix 2 the numerical data and results of the main analysis and sensitivity analysis of one RCT included in the wound hematoma in the *perioperative* systematic review (Bergqvist et al., 1990). The results of the 108 sensitivity analyses are graphically presented in a color-coded table (see Table 6).

For the four common assumptions, the percentage of effect estimates that lost significance varied from 0% to 92%:

- None of the 12 effect estimates lost significance under “best case scenario” assumption.

- One estimate lost significance under “none of participants with MPD had the event” assumption (minor bleeding outcome in *parenteral* systematic review).
- Three estimates lost significance under the assumption “all participants with MPD had the event” (symptomatic deep vein thrombosis in *parenteral* systematic review, symptomatic pulmonary embolism in *parenteral* systematic review, and symptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to placebo).
- All estimates lost significance under “worst case scenario assumption” except for the estimate of the outcome with the lowest rate of MPD (minor bleeding in *oral* systematic review). Almost half of these estimates changed direction and became insignificant (wound hematoma in *perioperative* systematic review, minor bleeding in *parenteral* systematic review, symptomatic pulmonary embolism in *parenteral* systematic review, symptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to placebo, and asymptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to vitamin K antagonist.

For the increasing range of plausible assumptions, the percentage of effect estimates that lost significance varied from 8% to 50%:

- Under $RI_{1/1}$ assumption, one estimate lost significant (minor bleeding in *parenteral* systematic review);

- Under $RI_{1.5/1}$ assumption, two estimates lost significant (minor bleeding in *parenteral* systematic review, asymptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to vitamin K antagonists);
- Under $RI_{2/1}$ assumption, one more estimate lost significance in addition to the two previous estimates mentioned in $RI_{1.5/1}$ assumption (thrombocytopenia in *central venous catheter* systematic review comparing low-molecular weight heparin to vitamin K antagonists);
- Under $RI_{3/1}$ assumption, in addition to the three estimates mentioned in the previous assumption, one further estimate lost significance (symptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to placebo);
- Under the most stringent assumption $RI_{5/1}$, almost half of the estimates lost significance (wound hematoma in *perioperative* systematic review, minor bleeding in *parenteral* systematic review, symptomatic deep vein thrombosis in *parenteral* systematic review, symptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to placebo, and thrombocytopenia in *central venous catheter* systematic review comparing low-molecular weight heparin to vitamin K antagonists) of which one changed direction as well (asymptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to vitamin K antagonists).

We used multiple logistic regression to explore the association of "percentage of effect estimates losing significance under the $RI_{5/1}$ assumption" and other general and methodological characteristics including number of RCTs included, rate of participants with observed events, and rate of participants most likely had MPD. The latter assumption was selected since it displayed the highest variability of impact on the 12 effect estimates (almost 50%). In our small sample, the model was not found to be statistically significant at α of 0.05 with a p-value of 0.428. Table 7 shows the adjusted OR of each covariate with the corresponding p-values and 95%CI.

CHAPTER IV

DISCUSSION AND RECOMMENDATION

A. Summary of findings

For the four common assumptions, the percentage of effect estimates that lost significance varied from 0% (best case scenario) to 8% (none of the participants with MPD had the event) to 25% (all participants with MPD had the event) to 92% (worst case scenario). In the worst case scenario, 42% of the effect estimates changed direction, i.e. a protective effect became harmful, and a harmful effect became protective. Whereas for the increasing range of plausible assumptions, the percentage of effect estimates that lost significance varied from 8% ($RI_{1/1}$) to 17 % ($RI_{1.5/1}$) to 25% ($RI_{2/1}$) to 33% ($RI_{3/1}$) to % 50% ($RI_{5/1}$) .

B. Interpretation of results

Up to a third of significant effect estimates of dichotomous outcomes that are patient-important lost significance when different assumptions about participants with MPD were applied. Though very extreme and unrealistic, the worst case scenario assumption is used to verify the robustness of the effect estimates (Unnebrink & Windeler, 2001). Our findings showed that only one effect estimate retained statistical significance under a worst case scenario (minor bleeding in the *oral* systematic review); this might be explained by the negligible rate of MPD (0.8%). On the other extreme, all effect estimates with the best case scenario assumptions

remained significant. The latter finding is not surprising since this assumption works towards the statistical significance and amplifies the effect. Under the other less plausible assumptions (none of the participants with MPD had the event and all participants with MPD had the event), one and three effect estimates lost significance, respectively.

The impact of the more plausible assumptions on the effect estimate varied among outcomes with no particular trend of association with the rate of missingness. Both outcomes, minor bleeding in the *parenteral* systematic review (low rate of MPD 3.5%) and asymptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to vitamin K antagonist (high rate or MPD 10.5%), lost significance under most assumptions. We notice that outcomes with CI of borderline significance are more likely to lose significance across different assumptions. For instance, minor bleeding in *parenteral* systematic review had a borderline significance CI with a low limit of 1.02. Similarly asymptomatic deep vein thrombosis in *central venous catheter* systematic review comparing low-molecular weight heparin to vitamin K antagonists had a borderline significance CI with low limit of 1.20.On the other hand, the outcome symptomatic pulmonary embolism in *parenteral* systematic review, despite its borderline significance of CI with a higher limit of 0.96, was robust to all assumptions except for two most common implausible assumptions “all participants with MPD had the event” and “worst case scenario”. Thus, the findings show that effect estimates are dependent on characteristics beyond the extent of MPD of trial participants, including but not limited to rate of observed events.

Our findings showed a significant difference between the size of participants who less likely had MPD across the intervention group and control group. This might have occurred due to chance especially due to the small sample size, thus should be explored with a larger representative sample.

As previously noted, the rating of the quality of evidence might be affected bias associated with MPD. In our sample, the evidence of an outcome which lost significance under most assumptions (minor bleeding in the *parenteral* systematic review) was rated moderate whereas that of an outcome which robust all assumptions except for worst case scenario (asymptomatic DVT in the *parenteral* systematic review) was rated high. This is explained by the fact that a large number of other factors are considered when rating the quality of evidence (Guyatt, Oxman, Schunemann, et al., 2011).

A systematic review conducted by Akl et al. applied similar methods for the same objective however their sample composed of 235 reports of RCTs recently published in five top general medical journals instead of systematic reviews (Akl et al., 2009). Similar to our findings, they deduced that the interpretation of results of RCTs published in top medical journals is highly influenced by the plausible assumptions regarding outcomes of participants with MPD, however slightly different percentages. Results of 19% of RCTs lost significance when none of the participants with MPD were assumed to have had the event of interest, 17% when all participants with MPD were assumed to have had the event, and 58% under worst case scenario assumption. Under the same more plausible assumptions, Akl et al found that, results of 0% to 33% RCTs were no longer significant (Akl et al., 2012).

C. Strengths and limitations

Our study has several strengths. It focuses on systematic reviews with a statistically significant pooled effect estimate for a patient-important efficacy outcome because they are the most likely to influence clinical practice. For that reason, we did not explore how making assumptions about MPD would affect non-statistically significant effect estimates. It appears that this might be the first study to apply a recommended approach to dealing with MPD in systematic reviews by conducting a complete case analysis in the main analysis and some form of sensitivity analysis to assess the robustness of results of the main analysis (Akl et al., 2013; Gamble & Hollis, 2005; J. P. Higgins et al., 2008; Mavridis et al., 2014). Three out of 100 clinical systematic reviews published in 2012 conducted sensitivity analyses using worst-case scenario or best case scenario to evaluate the robustness of results (unpublished data).

Using more plausible assumptions with the commonly used less plausible assumptions informed the evaluation of the robustness of the results. Moreover, the graphical color-coded presentation of the sensitivity analyses facilitated judging the risk of bias associated with MPD.

A main limitation of our study is its generalizability because of its small sample size. However, external validity was not the purpose of this study; it has been designed as feasibility pilot study of a larger study which includes a representative sample of Cochrane and non-Cochrane systematic reviews. Another limitation might be the focus on the change in significance and not on the change in effect estimates which might be of more importance and relevance in some situations.

This study focused on dichotomous outcome data and not on continuous measures. Given the substantial difference between continuous and dichotomous data in terms of methodological and statistical issues, the impact of MPD on continuous outcomes must be studies independently.

Making assumptions implies imputing the outcomes of participants with MPD. That would increase the total number of events and result in inappropriate narrowing of the CI of the effect estimate (Akl et al., 2009). In order to avoid such inappropriate narrowing of the CI, a group of methodologists recommended certain statistical approaches to take into account the uncertainty associated with imputing outcomes (White, Higgins, & Wood, 2008; White, Welton, Wood, Ades, & Higgins, 2008). One recommended approach consists of using a "variance inflation factor" to inflate the standard errors or variances of effect estimates. Though very crucial, taking uncertainty into account does not create a critical issue to our context, because the aim of conducting assumptions here is to assess the risk of bias associated with MPD and not to estimate the best estimate of a treatment effect.

D. Implications for practice

Our findings will inform systematic review authors regarding what assumptions for MPD should be used to test the robustness of the meta-analytical results. This depends on the research question itself and not on the findings of the meta-analysis. Akl et al recommend using a reasonable assumption with respect to the reason of missingness (Akl et al., 2012). For example, an assumption that all patients with MPD experienced an unfavorable outcome event

might be more plausible when patients are expected to adhere with the trial protocol and follow-up. This might be the case of a research question evaluating a drug to prevent rejection after cardiac transplantation. Whereas systematic reviews assessing the efficacy of drug to treat headaches may be more inclined to assume that none of the participants with MPD had the outcome of interest (persistent headache).

Thus, our study has important implications for trialists, editors of medical journals, systematic reviewers, and users of medical/health literature. Trial investigators should aim to reduce the extent of MPD by thorough and extensive follow-up, clear and transparent reporting of participants with MPD in terms of baseline characteristics, number per group, number per outcome, reasons for missingness, and potential implications for their primary analysis. This could be enforced by editors of medical/health journals by emphasizing the use of the CONSORT statement, specifically the reporting the participant flow diagram , the number of participants with MPD per group, the reasons for missingness, and the number of participants included in the analysis per group. Moreover, assessing risk of bias associated with MPD at the systematic review level will provide insight about the extent of which MPD compromises trust in statistically significant results. Ultimately guideline panels grading the evidence and stating recommendations are better guided.

Usually meta-analyses based on aggregate data (which is usually extracted from study reports) are challenged by the limited availability of participant data. Individual participant data (IPD) meta-analyses are commonly described as the gold standard of systematic reviews because they include both published and unpublished data of greater numbers of participants (Vale et al.,

2015). Thus more reliable summaries of trial results are presented and more detailed results could be generated (Vale et al., 2015). For instance, we would have tested other imputation methods (such as multiple imputations and regression models) if IPD had been available (Akl et al., 2009).

Besides being time and effort consuming, data extraction from eligible RCTs, is subject to bias especially when the data relates to MPD. Unfortunately, reporting MPD in most RCTs is not explicit and transparent enough for systematic reviewers to account for in their meta-analyses. Thus the judgment of what is MPD and what is followed- up is left up to the systematic reviewers. Hence, different systematic reviewers studying the same research question might end up with different results of meta-analysis based on the way they decided to identify and handle MPD. This problem is prevalent since there does not exist for systematic reviewers, uniform standards to report MPD or deal with MPD. None can deny that the reporting of RCTs has improved since the CONSORT statement publication in 1996 (Begg et al., 1996). The statement dwelled well upon reporting randomization and blinding; however nothing explicit was stated regarding reporting MPD. For that reason, we suggest that trialists start to report RCT data in a standardized file that is compatible with meta-analysis software. Such transparency would not only facilitate the work of the systematic review and limit any room for subjectivity and selection and analysis bias, it is also as crucial to identify bias and assess the validity. Improving the reporting of RCTs would provide the solutions for challenges related to identifying participants with MPD. Suggested improvements in RCT reporting include: reporting of MPD by outcome and not by participants; specifying for each outcome and by study group the number

with MPD and reasons; explicitly reporting how the RCT analysis dealt with MPD; and making IPD from RCT publicly available.

E. Implications for research

This thesis is a pilot study of a larger definitive project assessing the risk of bias associated with MPD in 100 Cochrane and non-Cochrane systematic reviews (Akl, Kahale, Agarwal, et al., 2014). This project includes a representative sample of systematic reviews, and thus the results can be generalized with much confidence. It will include the adoption of systematic and transparent methods, including specific and explicit eligibility criteria, sensitive search strategies, and duplicate and independent processes for study selection, data abstraction, and data interpretation. The team conducting this project has extensive experience in completing methodological studies.

We will consider in the definitive project the challenges faced and integrate the lessons learned from the pilot study:

- Identifying participants with MPD;
- Exploring the categories of trial participants the systematic review authors considered as having MPD in parallel to how trialists reported on these participants in trials;
- Dealing with MPD in meta-analyses if trialists already applied assumptions about outcomes of participants with MPD in the main analysis;
- Analyzing efficacy outcomes separately from safety ones since each type is reported and analyzed differently in RCTs especially with respect to participants with MPD;

- Exploring RoB associated with MPD from the lens of quality of evidence;
- Exploring whether the probability of effect estimates losing their significance is associated with other methodological characteristics, including (in addition to the rate of missingness and rate of observed events):
    - *Differential missingness:* defined as the difference in MPD rate between the intervention and control group;
    - *Borderline significance:* defined as the difference between the border of the confidence interval closer to 1 and the null effect 1;
    - *Magnitude of RR: defined as the difference between the RR measure and the null effect 1*(for RR less than 1, we will invert);
    - *Duration of follow-up*.

Future studies might also want to explore the above associations by conducting simulations.

Table 1: Numerical information from each trial to be used in the sensitivity analyses

	Intervention group			Control group		
	# Randomized	# Participant with MPD	# Observed events	# Randomized	# Participant with MPD	# Observed events
Number of trial participants	A	B	c	E	F	g

The number of participants with MPD is the sum of the number of participants in each of the categories of participants known or assumed to have MPD (see section “Challenge of identifying participants with missing participant data”)

Table 2: Numerical details of different methods to be used in the sensitivity analyses

Analytic method	Intervention group		Control group	
	Numerator	Denominator	Numerator	Denominator
Complete case analysis	C	A-B	g	E-F
None of participants with MPD had the outcome of interest	C	A	g	E
All participants with MPD had the outcome of interest	B + c	A	F + g	E
Best case scenario	C	A	F + g	E
Worst case scenario	B + c	A	g	E
Using the concept of $RI_{LTFU/FU}$ ¶	[B . ***y*** . c/(A-B)] + c	A	[F . ***z*** . g/(E-F)] + g	E

Refer to table 1 for what values letters a-g.

¶ *y* and *z* refer to $RI_{LTFU/FU}$ in the intervention and control group respectively (Akl et al., 2009)

Table 3: General characteristics of the five systematic reviews

Systematic Review	Comparison group	# Included studies	# Participants who completed follow-up	# Meta-analyses	# Outcomes with significant effect estimate
Oral AC (Reference #6)	VKA vs placebo	7	1770	9	2
Perioperative AC (Reference #5)	LMWH vs UFH	16	12890	13	1
Long-term AC (Reference #4)	LMWH vs VKA	10	1981	16	1
Parenteral AC (Reference #8)	LMWH vs placebo	15	7662	18	4
AC for CVC (Reference #9)	LMWH vs placebo	6	1448	7	1
	VKA vs placebo	5	666	5	1
	LMWH vs VKA	3	620	7	2

AC= AntiCoagulation; *CVC*= Central Venous Catheter; *LMWH*= Low Molecular Weight Heparin; *PE*= Pulmonary Embolism; *UFH*= UnFractionated Heparin; *VKA*= Vitamin K Antagonist

Table 4: Summary of MPD categories for all 12 eligible meta-analyses

Participant category	Mean± SD (IQR)			p-Value*
	Both groups (% of randomized)	Intervention (% of randomized)	Control (% of randomized)	
Most likely MPD pre-treatment [1]	1.15± 2.72 (0 - 1.30)	1.28± 3.01 (0 - 1.39)	0.94± 2.70 (0 - 0.43)	0.269
Most likely MPD post-treatment [2]	4.92± 8.89 (0 - 7.22)	5.39± 9.50 (0 - 7.66)	4.40± 8.41 (0 - 5.95)	0.003
Less likely MPD[3]	7.37± 13.66 (0 - 10.74)	7.34± 13.43 (0 - 11.30)	7.54± 14.37 (0 - 13.90)	0.684
Total	13.61± 17.14 (0 - 24.78)	14.27± 17.61 (0 - 24.91)	12.94± 16.93 (0 - 22.37)	0.079

[1] This category includes the participants: "mistakenly randomized or inappropriately excluded or ineligible participants" and "did not receive any treatment";
[2] This category includes the participants: "withdrew consent", "outcome not assessable", "moved out of country", and "lost to follow-up";
[3] This category includes the participants: "dead", "experienced adverse events", "non-compliant", "discontinued prematurely", "cross-over", and other similarly described participants.

*p-Value reflects the comparison of intervention group versus control group

Table 5: General characteristics of the 12 eligible meta-analyses

Systematic Review (comparison)	Outcome RR; [95%CI] by CCA	# Included trials	# Participants randomized	# Participants with observed events	Rate of observed events (%)	Number of participants most likely to have MPD	Rate of participant with MPD (%)	Quality of Evidence (GRADE)
Oral AC (VKA vs placebo)	Major bleeding 4.24; [1.86, 9.65]	4	1282	86	6.71	98	7.64	moderate
	Minor Bleeding 3.19; [1.83, 5.55]	4	875	143	16.34	7	0.80	moderate
Perioperative AC (LMWH vs UFH)	Wound Hematoma 0.68, [0.52, 0.88]	6	2550	216	8.47	89	3.49	moderate
Long-term AC (LMWH vs VKA)	Recurrent VTE 0.50; [0.35, 0.71]	5	1814	124	6.84	26	1.43	moderate
Parenteral AC (LMWH vs placebo)	Minor Bleeding 1.32; [1.02, 1.71]	13	7148	263	3.68	248	3.47	moderate
	Symptomatic VTE 0.56; [0.42, 0.74]	13	7067	278	3.93	240	3.40	high

	Symptomatic DVT 0.49; [0.28, 0.86]	9	6456	150	2.32	235	3.64	high
	Symptomatic PE 0.59; [0.37, 0.96]	9	6493	75	1.16	237	3.65	high
AC for CVC (LMWH vs placebo)	Symptomatic DVT 0.48; [0.27,0.86]	6	1551	55	3.55	213	13.73	moderate
AC for CVC (VKA vs placebo)	Asymptomatic DVT 0.43; [0.30, 0.62]	3	699	103	14.74	60	8.58	low
AC for CVC (LMWH vs VKA)	Asymptomatic DVT 1.74; [1.20, 2.52]	3	640	95	14.84	67	10.47	low
	Thrombocytopenia 3.73; [2.26, 6.16]	2	340	77	22.65	22	6.47	low

AC= AntiCoagulation; *CCA* = Complete Case Analysis; *CVC*= Central Venous Catheter; *DVT*= Deep Vein Thrombosis; *LMWH*= Low Molecular Weight Heparin; *MPD*= Missing Participant Data; *PE*= Pulmonary Embolism; *UFH*= UnFractionated Heparin; *VKA*= Vitamin K Antagonist; *VTE*= Venous ThromboEmbolism.

Table 6: Results of sensitivity analyses on significant pooled effect estimates under CCA

SR (comparison)	Outcome RR; [95%CI] By CCA	BCS RR; [95%CI]	None had the event RR; [95%CI]	$RI_{1/1}$ RR; [95%CI]	$RI_{1.5/1}$ RR; [95%CI]	$RI_{2/1}$ RR; [95%CI]	$RI_{3/1}$ RR; [95%CI]	$RI_{5/1}$ RR; [95%CI]	All had the event RR; [95%CI]	WCS RR; [95%CI]	QoE (GRADE)
Oral AC (VKA vs placebo)	Major bleeding 4.24; [1.86, 9.65]	4.88; [1.59, 14.94]	4.23; [1.86, 9.62]	3.95; [1.95, 7.98]	3.79; [1.97, 7.31]	3.64; [1.96, 6.73]	3.32; [1.90, 5.81]	2.79; [1.62, 4.81]	2.25; [1.94, 4.40]	1.87; [0.65, 5.36]	Moderate
	Minor Bleeding 3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	3.19; [1.83, 5.55]	Moderate
Perioperative AC (LMWH vs UFH)	Wound Hematoma 0.68, [0.52, 0.84]	0.52; [0.40, 0.66]	0.67; [0.52, 0.87]	0.67; [0.52, 0.86]	0.69; [0.54, 0.89]	0.71; [0.55, 0.91]	0.74; [0.58, 0.95]	0.81; [0.61, 1.08]	0.81; [0.66, 0.99]	1.22; [0.70, 2.12]	Moderate
Long-term AC (LMWH vs VKA)	Recurrent VTE 0.50; [0.35, 0.71]	0.45; [0.32, 0.63]	0.50; [0.35; 0.71]	0.49; [0.34, 0.70]	0.50; [0.35, 0.71]	0.50; [0.35, 0.71]	0.50; [0.35, 0.71]	0.53; [0.38, 0.75]	0.57; [0.41, 0.74]	0.70; [0.39, 1.26]	Moderate
Parenteral AC (LMWH vs placebo)	Minor Bleeding 1.32; [1.02, 1.71]	2.03; [1.49, 2.76]	1.31; [1.00, 1.70]	1.30; [0.98, 1.75]	1.30; [0.96, 1.76]	1.29; [0.93, 1.77]	1.26; [0.90, 1.76]	1.23; [0.84, 1.80]	1.45; [1.13, 1.86]	0.98; [0.63, 1.54]	Moderate
	Symptomatic VTE 0.56; [0.42, 0.74]	0.41; [0.29, 0.57]	0.56; [0.42, 0.73]	0.54; [0.41, 0.71]	0.55; [0.42, 0.73]	0.57; [0.42, 0.76]	0.59; [0.43, 0.81]	0.63; [0.44, 0.90]	0.69; [0.50, 0.96]	0.90; [0.51, 1.58]	High
	Symptomatic DVT 0.49; [0.28, 0.86]	0.29; [0.18, 0.48]	0.49; [0.29, 0.84]	0.48; [0.27, 0.83]	0.48; [0.27, 0.83]	0.49; [00.27, 0.88]	0.51; [0.28, 0.93]	0.55; [0.29, 1.04]	0.64; [0.37, 1.08]	0.94; [0.37, 2.42]	High
	Symptomatic PE 0.59; [0.37, 0.96]	0.37; [0.15, 0.96]	0.59; [0.37, 0.96]	0.55; [0.34, 0.89]	0.57; [0.36, 0.91]	0.59; [0.37, 0.93]	0.59; [0.37, 0.93]	0.62; [0.40, 0.98]	0.92; [0.69, 1.23]	2.06; [0.85, 5.01]	High

AC for CVC (LMWH vs placebo)	Symptomatic DVT 0.48; [0.27,0.86]	0.20; [0.12, 0.32]	0.48; [0.26, 0.86]	0.47; [0.26, 0.84]	0.48; [0.26, 0.89]	0.48; [0.26, 0.89]	0.50; [0.25, 1.00]	0.53; [0.25, 1.15]	0.80; [0.54, 1.20]	1.91; [0.57, 6.38]	Low
AC for CVC (VKA vs placebo)	Asymptomatic DVT 0.43; [0.30, 0.62]	0.34; [0.22, 0.53]	0.45; [0.31, 0.65]	0.38; [0.24, 0.61]	0.38; [0.27, 0.54]	0.40; [0.29, 0.55]	0.45; [0.34, 0.61]	0.54; [0.41, 0.70]	0.55; [0.42, 0.72]	0.76; [0.57, 1.02]	Low
AC for CVC (LMWH vs VKA)	Asymptomatic DVT 1.74; [1.20, 2.52]	2.69; [1.62, 4.46]	1.80; [1.23, 2.63]	1.44; [1.03, 2.01]	1.33; [0.97, 1.84]	1.27; [0.93, 1.74]	1.08; [0.81, 1.45]	0.89; [0.68, 1.16]	1.32; [1.03, 1.68]	0.86; [0.65, 1.15]	Low
	Thrombocytopenia 3.73; [2.26, 6.16]	4.34; [2.64, 7.13]	3.71; [2.24, 6.14]	3.75; [2.27, 6.17]	3.53; [2.17, 5.73]	2.85; [1.18, 6.87]	2.85; [1.18, 6.87]	2.82; [1.30, 6.12]	2.43; [1.22, 4.85]	1.04; [0.10, 10.46]	Low

AC= AntiCoagulation; *BCS*= Best Case Scenario; *CCA* = Complete Case Analysis; *CI*= Confidence Interval; *CVC*= Central Venous Catheter; *DVT*= Deep Vein Thrombosis; *GRADE*= Grading of Recommendations Assessment, Development and Evaluation; *LMWH*= Low Molecular Weight Heparin; *MPD*= Missing Participant Data ;*PE*= Pulmonary Embolism; *QoE*= Quality of Evidence; *RR*= Relative Risk; *SR*= Systematic Review; *UFH*= UnFractionated Heparin; *VKA*= Vitamin K Antagonist; *VTE*= Venous ThromboEmbolism; *WCS*= Worst Case Scenario

$RI_{LTFU/FU}$ refers to the event incidence among those lost to follow-up (LTFU) relative to the event incidence among those followed up (FU). The two values between brackets refer to the combination of $RI_{LTFU/FU}$ in the two trial groups. Going from left to right, the combinations increasingly challenge the statistical significance of the effect estimates.

Effect remained significant	**Effect became non-significant**	**Effect Changed direction and became not significant**	**Effect changed direction and became significant**

Table 7: Multiple analysis exploring association of "percentage of effect estimates losing significance under $RI_{5/1}$ assumption" and other general and methodological characteristics

	Effect estimate losing significance under $RI_{5/1}$		
Variables	**Adjusted OR**	**p-Value**	**95% CI**
Number of trials included	0.85	0.30	-0.28,0.91
Rate of observed events	0.36	0.40	-0.18,0.44
Rate of MPD	0.85	0.16	-0.13,0.75

Figure 1: Conventional and cumulative meta-analysis of 17 RCTs of intravenous streptokinase vs placebo or no drug for acute myocardial infarction (Antman et al., 1992)

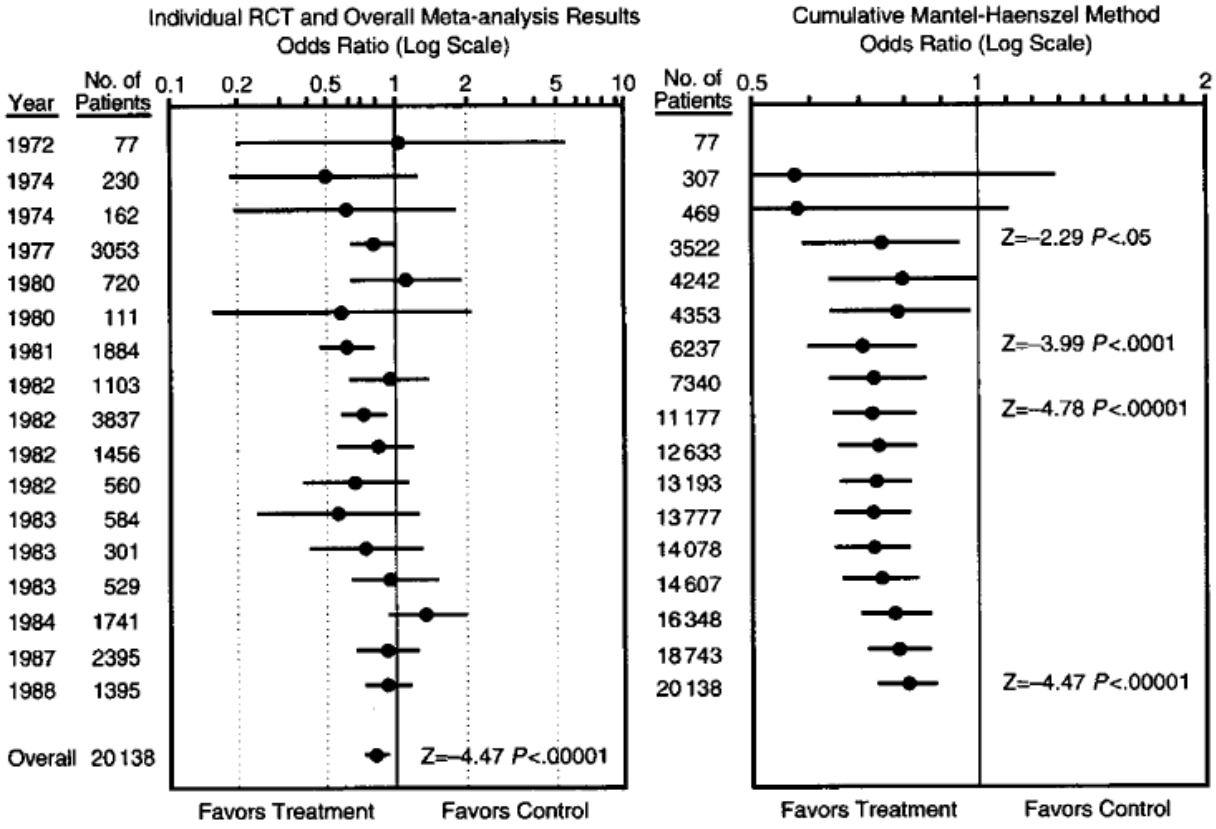

Figure 2: Example flow diagram from a published trial

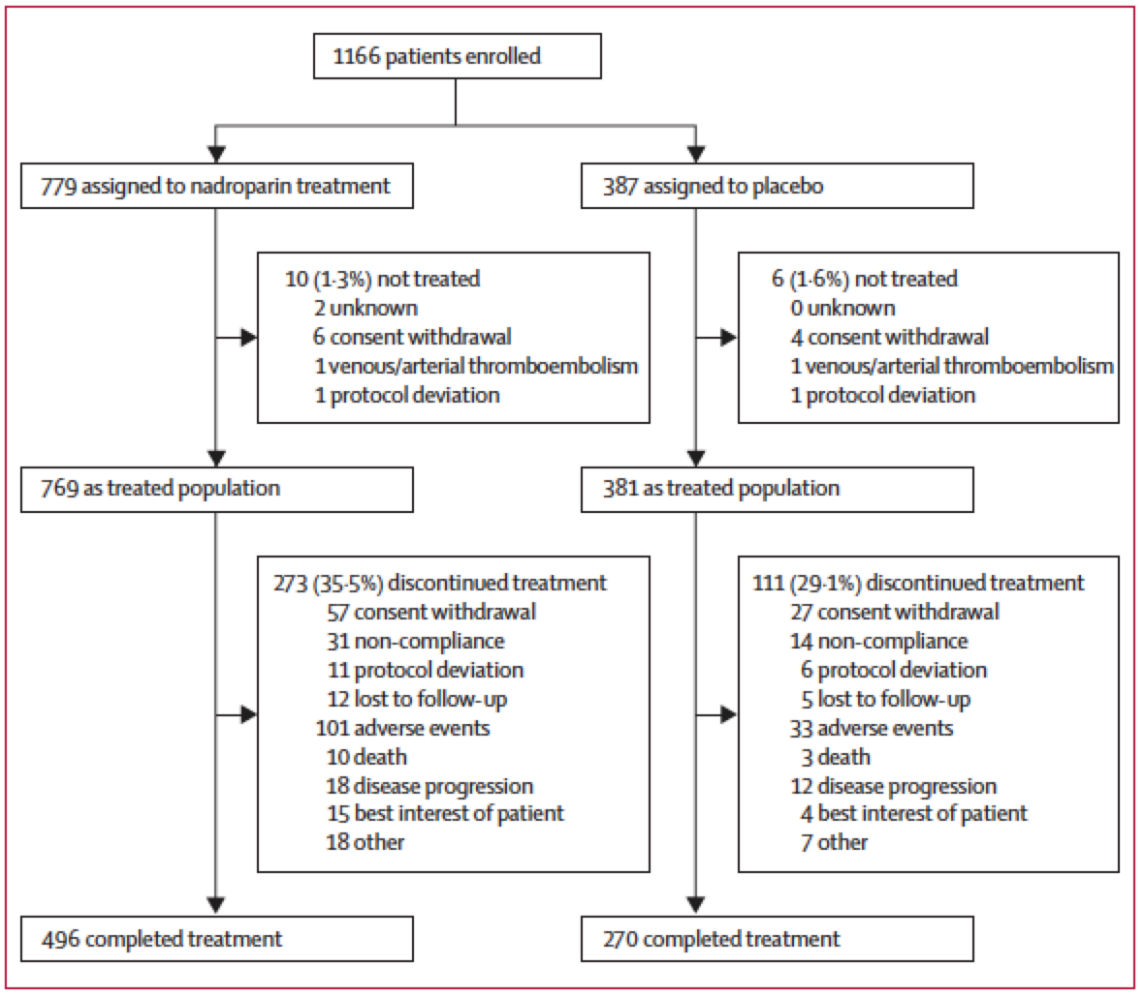

Figure 3: Example outcomes frequency table from a published trial

	Nadroparin (N=769)	Placebo (N=381)
Overall thromboembolic events	15 (2·0)	15 (3·9)
Deep-vein thrombosis	8 (1·0)	8 (2·1)
Pulmonary embolism	3 (0·4)	3 (0·8)
Visceral venous thrombosis	1 (0·1)	1 (0·3)
Stroke and peripheral thrombosis	3 (0·4)	3 (0·8)
Thromboembolic event by cancer site		
Lung	7/199 (3·5)	7/80 (8·8)
Gastrointestinal	4/272 (1·5)	4/148 (2·7)
Pancreas	3/36 (8·3)	1/17 (5·9)
Other	1/262 (0·4)	3/136 (2·2)

Data are n (%).

Table 3: Thromboembolic events by treatment group and cancer site

Appendix 1: Hierarchy of outcomes relative to patient importance

I. Mortality
    1. all - cause mortality
    2. disease specific mortality

II. Morbidity
    1. cardiovascular major morbid events
    2. other major morbid events (e.g. loss of vision, seizures, fracture, revascularization)
    3. onset/recurrence/relapse/remission of cancer and other chronic diseases (e.g. COPD exacerbation, new onset of diabetes)
    4. renal failure requiring dialysis
    5. hospitalization, medical and surgical procedures (e.g. placement of a pacemaker, and cardioversion)
    6. infections
    7. dermatological/ rheumatologic disorders

III. Symptoms/Quality of life/Functional status (e.g. failure to become pregnant, successful nursing/breastfeeding, depression)

IV. Surrogate outcomes (e.g. viral load, physical activity, weight loss, cognitive function, recurrent polyps, adherence to medication)

Appendix 2: Illustration of calculation of main analysis and sensitivity analyses of one RCT included in wound hematoma outcome in *perioperative* systematic review

Table 1: Numerical information from one trial included the eligible meta-analysis of wound hematoma outcome in *perioperative* systematic review

	Intervention group			Control group		
Trial	# Randomized	# Participant with MPD	# Observed events	# Randomized	# Participant with MPD	# Observed events
Bergqvist 1990	311_A	19_B	36_c	326_E	19_F	47_g

Refer to original table 1 for what values letters.

Table 2: Numerical details of different methods to be used in the sensitivity analyses of one trial included the eligible meta-analysis of wound hematoma outcome in *perioperative* systematic review

Analytic method	Intervention group		Control group	
	Numerator	Denominator	Numerator	Denominator
Complete case analysis	36_c	311_A-19_B= 292	47_g	326_E-19_F= 307
None of participants with MPD had the outcome of interest	36_c	311_A	47_g	326_E
All participants with MPD had the outcome of interest	19_B + 36_c= 55	311_A	19_F + 47_g= 66	326_E
Best case scenario	36_c	311_A	19_F + 47_g= 66	326_E
Worst case scenario	19_B + 36_c= 55	311_A	47_g	326_E
$RI_{LTFU/FU}$ = 2:1	[19_B . ***2*** . 36_c /(311_A -19_B)] + 36_c= 41	311_A	[19_F . ***1*** . 47_g /(326_E -19_F)] + 47_g= 50	326_E

BIBLIOGRAPHY

1. Agnelli, G., Gussoni, G., Bianchini, C., Verso, M., Mandala, M., Cavanna, L., Barni, S., Labianca, R., Buzzi, F., Scambia, G., Passalacqua, R., Ricci, S., Gasparini, G., Lorusso, V., Bonizzoni, E., Tonato, M., & Investigators, P. (2009). Nadroparin for the prevention of thromboembolic events in ambulatory patients with metastatic or locally advanced solid cancer receiving chemotherapy: a randomised, placebo-controlled, double-blind study. *Lancet Oncol*, *10*(10), 943-949. doi: 10.1016/S1470-2045(09)70232-3
2. Akl, E. A., Briel, M., You, J. J., Lamontagne, F., Gangji, A., Cukierman-Yaffe, T., Alshurafa, M., Sun, X., Nerenberg, K. A., Johnston, B. C., Vera, C., Mills, E. J., Bassler, D., Salazar, A., Bhatnagar, N., Busse, J. W., Khalid, Z., Walter, S., Cook, D. J., Schunemann, H. J., Altman, D. G., & Guyatt, G. H. (2009). LOST to follow-up Information in Trials (LOST-IT): a protocol on the potential impact. *Trials*, *10*, 40. doi: 10.1186/1745-6215-10-40
3. Akl, E. A., Briel, M., You, J. J., Sun, X., Johnston, B. C., Busse, J. W., Mulla, S., Lamontagne, F., Bassler, D., Vera, C., Alshurafa, M., Katsios, C. M., Zhou, Q., Cukierman-Yaffe, T., Gangji, A., Mills, E. J., Walter, S. D., Cook, D. J., Schunemann, H. J., Altman, D. G., & Guyatt, G. H. (2012). Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. *BMJ*, *344*, e2809. doi: 10.1136/bmj.e2809
4. Akl, E. A., Johnston, B. C., Alonso-Coello, P., Neumann, I., Ebrahim, S., Briel, M., Cook, D. J., & Guyatt, G. H. (2013). Addressing dichotomous data for participants excluded from

trial analysis: a guide for systematic reviewers. *PLoS One, 8*(2), e57132. doi: 10.1371/journal.pone.0057132

5. Akl, E. A., Kahale, L., Barba, M., Neumann, I., Labedi, N., Terrenato, I., Sperati, F., Muti, P., & Schunemann, H. (2014). Anticoagulation for the long-term treatment of venous thromboembolism in patients with cancer. *Cochrane Database Syst Rev, 7*, CD006650. doi: 10.1002/14651858.CD006650.pub4
6. Akl, E. A., Kahale, L., Sperati, F., Neumann, I., Labedi, N., Terrenato, I., Barba, M., Sempos, E. V., Muti, P., Cook, D., & Schunemann, H. (2014). Low molecular weight heparin versus unfractionated heparin for perioperative thromboprophylaxis in patients with cancer. *Cochrane Database Syst Rev, 6*, CD009447. doi: 10.1002/14651858.CD009447.pub2
7. Akl, E. A., Kahale, L., Terrenato, I., Neumann, I., Yosuico, V. E., Barba, M., Sperati, F., & Schunemann, H. (2014). Oral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation. *Cochrane Database Syst Rev, 7*, CD006466. doi: 10.1002/14651858.CD006466.pub5
8. Akl, E. A., Kahale, L. A., Agarwal, A., Al-Matari, N., Ebrahim, S., Alexander, P. E., Briel, M., Brignardello-Petersen, R., Busse, J. W., Diab, B., Iorio, A., Kwong, J., Li, L., Lopes, L. C., Mustafa, R., Neumann, I., Tikkinen, K. A., Vandvik, P. O., Zhang, Y., Alonso-Coello, P., & Guyatt, G. (2014). Impact of missing participant data for dichotomous outcomes on pooled effect estimates in systematic reviews: a protocol for a methodological study. *Syst Rev, 3*, 137. doi: 10.1186/2046-4053-3-137
9. Akl, E. A., Kahale, L. A., Ballout, R. A., Barba, M., Yosuico, V. E., van Doormaal, F. F., Middeldorp, S., Bryant, A., & Schunemann, H. (2014). Parenteral anticoagulation in

ambulatory patients with cancer. *Cochrane Database Syst Rev, 12*, CD006652. doi: 10.1002/14651858.CD006652.pub4

10. Akl, E. A., Ramly, E. P., Kahale, L. A., Yosuico, V. E., Barba, M., Sperati, F., Cook, D., & Schunemann, H. (2014). Anticoagulation for people with cancer and central venous catheters. *Cochrane Database Syst Rev, 10*, CD006468. doi: 10.1002/14651858.CD006468.pub5
11. Alper, B. S., Hand, J. A., Elliott, S. G., Kinkade, S., Hauan, M. J., Onion, D. K., & Sklar, B. M. (2004). How much effort is needed to keep up with the literature relevant for primary care? *J Med Libr Assoc, 92*(4), 429-437.
12. Alshurafa, M., Briel, M., Akl, E. A., Haines, T., Moayyedi, P., Gentles, S. J., Rios, L., Tran, C., Bhatnagar, N., Lamontagne, F., Walter, S. D., & Guyatt, G. H. (2012). Inconsistent definitions for intention-to-treat in relation to missing outcome data: systematic review of the methods literature. *PLoS One, 7*(11), e49163. doi: 10.1371/journal.pone.0049163
13. Altman, D. G. (2009). Missing outcomes in randomized trials: addressing the dilemma. *Open Med, 3*(2), e51-53.
14. Antman, E. M., Lau, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (1992). A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA, 268*(2), 240-248.
15. Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K. F., Simel, D., & Stroup, D. F. (1996). Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA, 276*(8), 637-639.

16. Bergqvist, D., Burmark, U. S., Frisell, J., Guilbaud, O., Hallbook, T., Horn, A., Lindhagen, A., Ljungner, H., Ljungstrom, K. G., Matzsch, T., & et al. (1990). Thromboprophylactic effect of low molecular weight heparin started in the evening before elective general abdominal surgery: a comparison with low-dose heparin. *Semin Thromb Hemost, 16 Suppl*, 19-24.
17. Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis.*
18. Campbell, H. (2014). PLOS One Output Drops - AAAS Cackles Like Scrooge McDuck. from http://www.science20.com/science_20/plos_one_output_drops_aaas_cackles_like_scrooge_mcduck-138206#ixzz3Wi5Ff2JR
19. Cochrane, A. (1973). *Effectiveness and efficiency: random reflections on health services.* London: Nuffield Provincial Hospitals Trust.
20. Cochrane Community. (2014). Evidence-based health care and systematic reviews. from http://community.cochrane.org/about-us/evidence-based-health-care
21. Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med, 126*(5), 376-380.
22. Dickersin, K., & Berlin, J. A. (1992). Meta-analysis: state-of-the-science. *Epidemiol Rev, 14*, 154-176.
23. Ebrahim, S., Akl, E. A., Mustafa, R. A., Sun, X., Walter, S. D., Heels-Ansdell, D., Alonso-Coello, P., Johnston, B. C., & Guyatt, G. H. (2013). Addressing continuous data for

participants excluded from trial analysis: a guide for systematic reviewers. *J Clin Epidemiol, 66*(9), 1014-1021 e1011. doi: 10.1016/j.jclinepi.2013.03.014

24. Ebrahim, S., Johnston, B. C., Akl, E. A., Mustafa, R. A., Sun, X., Walter, S. D., Heels-Ansdell, D., Alonso-Coello, P., & Guyatt, G. H. (2014). Addressing continuous data measured with different instruments for participants excluded from trial analysis: a guide for systematic reviewers. *J Clin Epidemiol, 67*(5), 560-570. doi: 10.1016/j.jclinepi.2013.11.014
25. Gamble, C., & Hollis, S. (2005). Uncertainty method improved on best-worst case analysis in a binary meta-analysis. *J Clin Epidemiol, 58*(6), 579-588. doi: 10.1016/j.jclinepi.2004.09.013
26. Guyatt, G. H., Oxman, A. D., Kunz, R., Brozek, J., Alonso-Coello, P., Rind, D., Devereaux, P. J., Montori, V. M., Freyschuss, B., Vist, G., Jaeschke, R., Williams, J. W., Jr., Murad, M. H., Sinclair, D., Falck-Ytter, Y., Meerpohl, J., Whittington, C., Thorlund, K., Andrews, J., & Schunemann, H. J. (2011). GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol, 64*(12), 1283-1293. doi: 10.1016/j.jclinepi.2011.01.012
27. Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., Alonso-Coello, P., Falck-Ytter, Y., Jaeschke, R., Vist, G., Akl, E. A., Post, P. N., Norris, S., Meerpohl, J., Shukla, V. K., Nasser, M., Schunemann, H. J., & Group, G. W. (2011). GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol, 64*(12), 1303-1310. doi: 10.1016/j.jclinepi.2011.04.014
28. Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., Alonso-Coello, P., Glasziou, P., Jaeschke, R., Akl, E. A., Norris, S., Vist, G., Dahm, P., Shukla, V. K., Higgins, J., Falck-Ytter, Y., Schunemann, H. J., & Group, G. W. (2011). GRADE

guidelines: 7. Rating the quality of evidence--inconsistency. *J Clin Epidemiol, 64*(12), 1294-1302. doi: 10.1016/j.jclinepi.2011.03.017

29. Guyatt, G. H., Oxman, A. D., Montori, V., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., Djulbegovic, B., Atkins, D., Falck-Ytter, Y., Williams, J. W., Jr., Meerpohl, J., Norris, S. L., Akl, E. A., & Schunemann, H. J. (2011). GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol, 64*(12), 1277-1282. doi: 10.1016/j.jclinepi.2011.01.011
30. Guyatt, G. H., Oxman, A. D., Schunemann, H. J., Tugwell, P., & Knottnerus, A. (2011). GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol, 64*(4), 380-382. doi: 10.1016/j.jclinepi.2010.09.011
31. Guyatt, G. H., Oxman, A. D., Sultan, S., Glasziou, P., Akl, E. A., Alonso-Coello, P., Atkins, D., Kunz, R., Brozek, J., Montori, V., Jaeschke, R., Rind, D., Dahm, P., Meerpohl, J., Vist, G., Berliner, E., Norris, S., Falck-Ytter, Y., Murad, M. H., Schunemann, H. J., & Group, G. W. (2011). GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol, 64*(12), 1311-1316. doi: 10.1016/j.jclinepi.2011.06.004
32. Guyatt, G. H., Oxman, A. D., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., Montori, V., Akl, E. A., Djulbegovic, B., Falck-Ytter, Y., Norris, S. L., Williams, J. W., Jr., Atkins, D., Meerpohl, J., & Schunemann, H. J. (2011). GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol, 64*(4), 407-415. doi: 10.1016/j.jclinepi.2010.07.017
33. Higgins, J. P., Altman, D. G., Gotzsche, P. C., Juni, P., Moher, D., Oxman, A. D., Savovic, J., Schulz, K. F., Weeks, L., Sterne, J. A., Cochrane Bias Methods, G., & Cochrane

Statistical Methods, G. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ, 343*, d5928. doi: 10.1136/bmj.d5928

34. Higgins, J. P., White, I. R., & Wood, A. M. (2008). Imputation methods for missing outcome data in meta-analysis of clinical trials. *Clin Trials, 5*(3), 225-239. doi: 10.1177/1740774508091600
35. Higgins, J. P., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester (UK): John Wiley & Sons.
36. Juni, P., & Egger, M. (2009). PRISMAtic reporting of systematic reviews and meta-analyses. *Lancet, 374*(9697), 1221-1223. doi: 10.1016/S0140-6736(09)61765-7
37. Mavridis, D., Chaimani, A., Efthimiou, O., Leucht, S., & Salanti, G. (2014). Addressing missing outcome data in meta-analysis. *Evid Based Ment Health, 17*(3), 85-89. doi: 10.1136/eb-2014-101900
38. Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M., Altman, D. G., & Consort. (2012). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg, 10*(1), 28-55. doi: 10.1016/j.ijsu.2011.10.001
39. Montori, V., & Guyatt, G. (2001). Intention-to-treat principle. *Canadian Medical Association Journal, 165*(10).
40. Mulrow, C. D. (1994). Rationale for systematic reviews. *BMJ, 309*(6954), 597-599.
41. National Library of Medicine, U. S. (2014). Finding medical information in MEDLINE. from http://www.nlm.nih.gov/services/usemedline.html

42. Rennie, D. (2001). CONSORT revised--improving the reporting of randomized trials. *JAMA, 285*(15), 2006-2007.

43. Savovic, J., Weeks, L., Sterne, J. A., Turner, L., Altman, D. G., Moher, D., & Higgins, J. P. (2014). Evaluation of the Cochrane Collaboration's tool for assessing the risk of bias in randomized trials: focus groups, online survey, proposed recommendations and their implementation. *Syst Rev, 3*, 37. doi: 10.1186/2046-4053-3-37

44. Shah, H. M., & Chung, K. C. (2009). Archie Cochrane and his vision for evidence-based medicine. *Plast Reconstr Surg, 124*(3), 982-988. doi: 10.1097/PRS.0b013e3181b03928

45. Suresh, K. (2011). An overview of randomization techniques: An unbiased assessment of outcome in clinical research. *J Hum Reprod Sci, 4*(1), 8-11. doi: 10.4103/0974-1208.82352

46. Sylvestre, Y. (2011). CONSORT: missing missing data guidelines, the effects on HTA monograph reporting. *Trials, 12*(Suppl 1), A61. doi: 10.1186/1745-6215-12-s1-a61

47. The Nordic Cochrane Centre. (2014). Review Manager. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.

48. Unnebrink, K., & Windeler, J. (2001). Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med, 20*(24), 3931-3946.

49. Vale, C. L., Rydzewska, L. H., Rovers, M. M., Emberson, J. R., Gueyffier, F., Stewart, L. A., & Cochrane, I. P. D. M.-a. M. G. (2015). Uptake of systematic reviews and meta-analyses based on individual participant data in clinical practice guidelines: descriptive study. *BMJ, 350*, h1088. doi: 10.1136/bmj.h1088

50. Van Gemert, A. (2010). PLoS ONE Publishes 10,000th Manuscript! The PLOS ONE Community Blog. Retrieved from http://blogs.plos.org/everyone/2010/04/02/plos-one-publishes-10000th-article/

51. White, I. R., Higgins, J. P., & Wood, A. M. (2008). Allowing for uncertainty due to missing data in meta-analysis--part 1: two-stage methods. *Stat Med, 27*(5), 711-727. doi: 10.1002/sim.3008

52. White, I. R., Horton, N. J., Carpenter, J., & Pocock, S. J. (2011). Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ, 342*, d40. doi: 10.1136/bmj.d40

53. White, I. R., Welton, N. J., Wood, A. M., Ades, A. E., & Higgins, J. P. (2008). Allowing for uncertainty due to missing data in meta-analysis--part 2: hierarchical models. *Stat Med, 27*(5), 728-745. doi: 10.1002/sim.3007

54. Wood, A. M., White, I. R., & Thompson, S. G. (2004). Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clin Trials, 1*(4), 368-376.