Draft: Assessing Outcome Reporting Bias in a Systematic Review Of Venous Thromboembolism Prophilaxis

Blain Morin

March 1, 2019

Introduction

Systematic reviews are a "lens through which evidence is viewed" and aim to provide a scientifically rigorous summary of all available evidence regarding a specified research question (1). Healthcare systematic reviews often help summarize the benefits and harms of medical treatments (2). While a systematic review is a powerful tool, the method is not infallible. For example: publication bias, which is a well-known phenomenon where studies with statistically significant results are more likely to be published, can have a drastic effect on a review's conclusions (3). Since these types of biases threaten the validity of the review, it is therefore vital to formally assess their potential impacts.

In the same vein as publication bias, some outcomes may not be reported within a study if the results are not statistically significant. Outcomes may also be suppressed if their results conflict with the aims of the trial (4). These nonrandom ommissions of information, known as outcome reporting bias (ORB), can have a substantial impact on the meta-analytic conclusions (5).

The direction of ORB depends on whether or not the outcome is a benefit or harm. Existing methods that are designed to correct for ORB assume that benefit outcomes are overstated and harm outcomes are understated (4). For example, it is conceivable that an industry-funded study is incentivized to report outcomes that make its product look favorable and may not report outcomes that make its product look harmful.

There is empirical evidence that statistically significant outcomes are more likely to be reported (6). A systematic review of outcome reporting bias reports that statistically significant outcomes have 2.2 to 4.7 times the odds of being reported (7). Moreover, harm outcomes are less likely to be fully reported than benefit outcomes (8). These empirical studies highlight the need for formally addressing ORB at both the study level and the review level.

ORB is often overlooked in systematic reviews. In a 2010 study, researchers investigated 283 reviews and found that 35% contained a trial with a high risk of ORB, but only 7% mentioned the risk of bias (6). Even when ORB is mentioned, its potential impact on the review's conclusions is often insufficiently addressed (9).

Formally assessing the potential impact of ORB involves scrutinizing the studies that were excluded

from the review because no relevant outcomes were reported, classifying the risk of bias in instances where included studies do not report certain outcomes, and conducting a sensitivity analysis on the meta-analyses' conclusions (7). This paper applies the Outcome Reporting Bias in Trials II (ORBIT II) classification system to a systematic review of Venous Thromboembolism Prophylaxis conducted by Brown University's Center for Evidence Synthesis in Health (CESH). After classifying the risk of bias of the admitted outcomes, the potential impact of ORB is assessed using a model-based sensitivity analysis (4).

Data

The data for this ORB assessment come from the CESH report, "Venous Thromboembolism Prophylaxis in Major Orthopedic Surgery: Systematic Review Update." The review was sponsored by the Agency for Healthcare Research and Quality (AHRQ), whose mission is to "to produce evidence to make health care safer, higher quality, more accessible, equitable, and affordable, and to. make sure that the evidence is understood and used," (10). Users of this review include doctors, patients, and policymakers who wish to have a summary of all the available evidence regarding the benefits and potential harms of venous thromboembolism prophylaxis.

Venous thromboembolism (VTE) is the development of a blood clot in a vein that enters circulation. One type of VTE is deep vein thrombosis (DVT), which refers to a blood clot that develops in an extremity (11). After a major orthopedic leg surgery, such as a total hip replacement (THR), a DVT may develop in the calf (distal DVT) or in the hip (proximal DVT). DVTs may produce symptoms such as pain and swelling but also may be asymptomatic. DVTs are of clinical interest because they are intermediaries for pulmonary embolisms (PE), which occur when a DVT becomes dislodged and enters the lungs. PEs are considered a catastrophic event and may lead to death (12).

The principal outcomes for the CESH review were decided on *a priori* and include various VTE outcomes, major bleeding events, and serious adverse events. The VTE outcomes are the efficacy measures, whereas major bleeding and serious adverse events are the safety measures. Descriptions of the principal outcomes are found in Table 1.

There is considerable variation in the reported outcomes between studies in the VTE prophylaxis review. For example, fatal PE was only reported in 48% of the trials. The clinical judgment of the review authors is that data for VTE and major bleeding outcomes "should have been available for almost all trials; therefore outcomes were excluded selectively," (12). In the review, VTE and major bleeding outcomes that contain data from less than 80% of the available trials were classified as having a "high risk of reporting bias," (12).

The meta-analyses that are extracted from the VTE prophylaxis review compare the benefits and harms

for two classes of pharmacological treatments, low weight molecular heparin (LWMH) and factor Xa inhibitors (FXaI), after THR surgery. THR surgery is selected because it is the surgery with the highest number of studies. The LWMH versus FXaI comparison is selected because it contains the highest number of meta-analyses (12).

All primary outcomes except for serious adverse events are used in this ORB assessment. Serious adverse events are dropped from the analysis because they were often not well defined within trials. The lack of a formal and consistent definition makes it difficult to judge the risk of bias.

Table 1: Outcome Descriptions

Outcome	Type	Description	p*
Total VTE	Efficacy	PE + Symptomatic and Asymptomatic DVT	15%
Symptomatic VTE	Efficacy	PE + Symptomatic DVT	18%
Total PE	Efficacy	Fatal + Nonfatal PE; Symptomatic and Asymptomatic PE	52%
Fatal PE	Efficacy	Death caused by PE	48%
Symptomatic PE	Efficacy	PE with clinical symptoms	17%
Total DVT	Efficacy	Symptomatic + Asymptomatic; Proximal and Distal	82%
Symptomatic DVT	Efficacy	DVT with clinical symptoms	40%
Proximal DVT	Efficacy	DVT above the knee	66%
Major Bleeding	Safety	Any major bleeding event	52%

*proportion of studies in the review that reported the outcome
VTE = Venous Thromboembolism
DVT = Deep Vein Thrombosis
PE = Pulmonary Embolism

Methods

First, studies that are excluded from the review are examined. If the exclusion reason is due to the lack of relevant outcome data, the trial is analysed to see if the relevant data was measured or likely to have been measured. If so, the trial is documented and included in the ORB assessment (7).

Next, an outcome matrix is constructed where the principle outcomes are the columns and each trial is a row. Each principal outcome that is reported per trial is recorded in the matrix. In the instances where the outcome is not reported, the ORBIT II classification is recorded. The ORBIT II method uses different classifications for benefit and harm outcomes (4). The decision trees for the classifications can be seen in Figures 1 and 2. The benefit outcome tree is used for efficacy outcomes. The harm tree is used for safety outcomes.

Figure 1: Benefit Outcome Tree

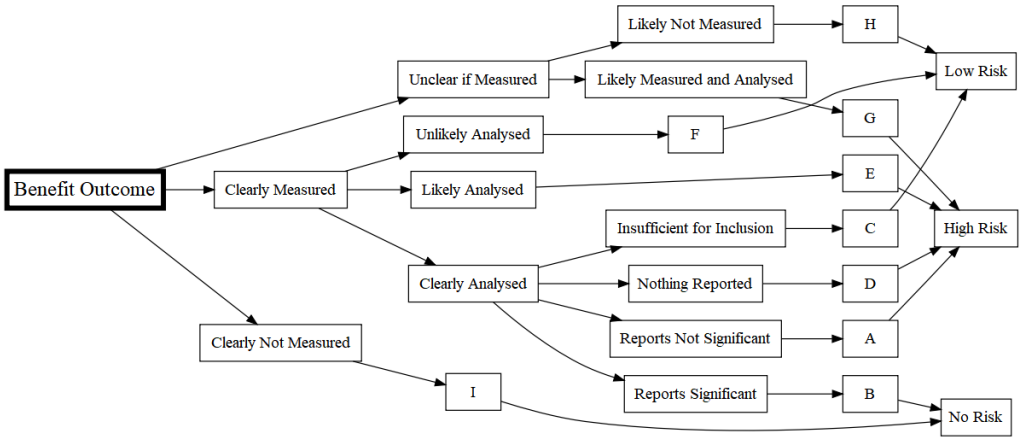
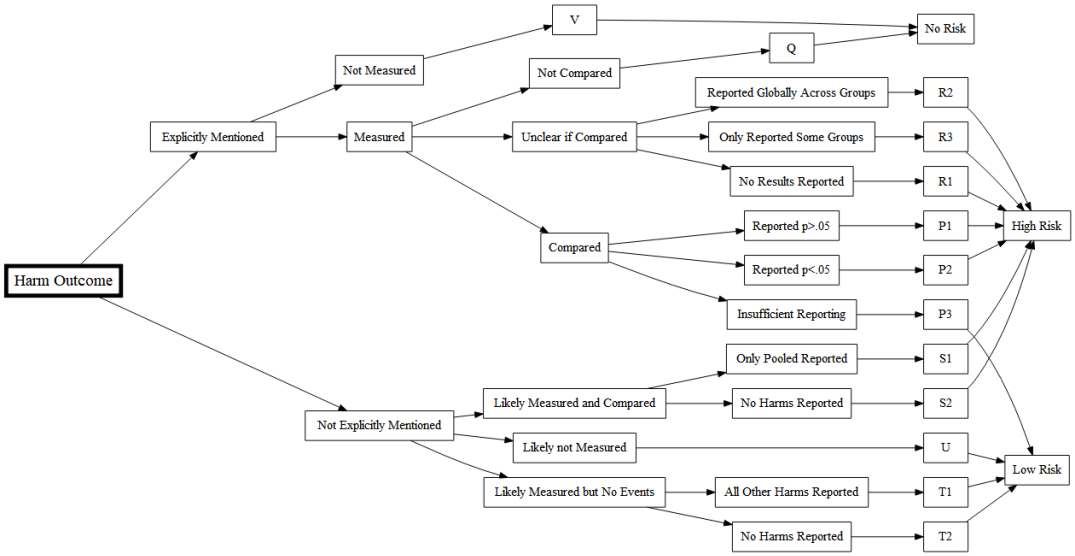


Figure 2: Harm Outcome Tree



Some of the ORBIT II branches require clinical judgment. For example, the "G" classification on the benefit branch in Figure 1 requires a judgment call to be made as to whether or not the outcome was "likely" to have been measured and analyzed. For these instances, the clinicians who conducted the review were contacted for their assessment.

The ORBIT II matrix is used to conduct a sensitivity analysis that produces ORB-adjusted estimates. The ORB-adjusted estimates are calculated through a random-effects maximum likelihood function proposed by Copas 2017 (4). The random effects model is chosen because trials pulled from the literature cannot be assumed to share a common underlying population (13). The likelihood functions are different for benefit and harm outcomes.

Both benefit and harm functions start with the classic random effects unadjusted likelihood function, $L_U(\theta, \tau^2)$:

$$L_U(\theta, \tau^2) = -\frac{1}{2} \sum_{Reported} \left\{ log(\sigma_i^2 + \tau^2) + \frac{(y_i - \theta)^2}{\sigma_i^2 + \tau^2} \right\}$$

The benefit likelihood, $L_B(\theta, \tau^2)$, adds a term that accounts for the high risk of bias benefit outcomes being not statistically significant:

$$L_B(\theta, \tau^2) = L_U(\theta, \tau^2) + \sum_{High} log \left[\Phi \left\{ \frac{z_\alpha \sigma_i - \theta}{\sqrt{\sigma_i^2 + \tau^2}} \right\} - \Phi \left\{ \frac{-z_\alpha \sigma_i - \theta}{\sqrt{\sigma_i^2 + \tau^2}} \right\} \right]$$

where Φ is the cummulative density function of the normal distribution and z_α is the appropriate z score for the significance level, *alpha*.

The harm likelihood, $L_H(\theta, \tau^2)$, starts with the unadjusted likelihood function and adds a term that accounts for the high risk of bias harm outcomes being statistically significant:

$$L_H(\theta, \tau^2) = L_U(\theta, \tau^2) + \sum_{High} log \left\{ \Phi \left(\frac{\theta}{\sqrt{\sigma_i^2 + \tau^2}} \right) \right\}$$

Results

Analysis of Excluded Studies

There were 290 excluded studies. Appendix B in the VTE prophylaxis lists the excluded studies along with the exclusion reason. Table 2 presents a tally of all the exclusion reasons.

Table 2: Tally of Rejection Reasons

Rejection Reason	n
Combined TKR and THR	32
Comparator mixed interventions	2
Duplicate publication (no additional data)	50
Insufficient results data reported	2
No abstract or full text available	1
No analysis by intevention	2
No comparison of interest	21
No intervention of interest	25
No outcome of interest	1
No primary data	13
No results reported	3
Not population of interest	5
Not primary study	1
Not surgery of interest	7
nRCS N<750	15
nRCS placebo comparison	2
Pre-2010 (presumably excluded by UConn)	95
SR or MA without references (conference abstract)	12

The rejection reasons in Table 2 that are possibly relevant to ORB are: "No outcome of interest", "Insufficient results data reported", and "No results reported." Each study with these rejection reasons was examined to see if the is relevant to the LWMH and FXaI comparison in THR. No excluded studies compare LWMH and FXaI treatments and thus none are added to the ORBIT matrix. Relevancy conclusions are listed in Table 3.

Table 3: Papers Omitted Because of Missing Data

Pubmed ID	Rejection Reason	Relevant?	Reason
NCT01720108	No results reported	No	Compares Rivaroxaban to Aspirin
21305339	No outcome of interest	No	Outcomes relate to length of hospital stay and prophylaxis patterns
Abstract 903	Insufficient results data reported	No	Not RCT, no sample sizes reported
Abstract page 21	Insufficient results data reported	No	RCT of FXaI versus Enoxaparin
noPMID 01	No results reported	No	Not RCT, pro and con discussion of prophylaxis options
23355673	No results reported	No	Examines total knee replacement, not THR

ORBIT II Classification Matrix

There were 12 RCT that compared LWMH and FXaI in THR. Only two of the trials, Eriksson 2010 and Fuji 2015, reported all principal outcomes. There was substantial missingness in many of the outcomes. For example, only 50% of the trials reported Total PE. The instances with a high risk of bias are almost all "E" classifications, which means that they were clearly measured and likely analyzed but no results were reported. Many of the low risk of bias "F" classifications were due to no events occurring. One article, Zhang 0214, was in Chinese and only had a translation for its abstract.

There were many instances where it was unclear as to why some trial outcomes were not included in certain meta-analyses. For example, in Kim2016, it seems that Total VTE could be calculated from the asymptomatic DVT, symptomatic DVT, and PE events reported. Instead of classifying these instances of high risk, it may be better to include the inferred data in the meta-analysis. These cases, where data can be inferred without the need for an ORBIT II classification are documented as "Y" in Table 4.

Table 4: ORBIT II Matrix: LWMH vs FXaI in THR

Study.Name	Total.VTE	Symptomatic.VTE	Total.PE	Fatal.PE	Symptomatic.PE	Total.DVT	Symptomatic.DVT	Proximal.DVT	Major.Bleeding
eriksson2010	✓	✓	✓	✓	✓	✓	✓	✓	✓
fuji2014A	✓	✓	F[1]	✓	✓	✓	✓	✓	✓
fuji2014D	✓	Y[2]	Y[3]	F[4]	✓	✓	✓	✓	✓
fuji2015	✓	✓	✓	✓	✓	✓	✓	✓	✓
lassen2010A	✓	E[5]	✓	✓	Y[6]	✓	✓	✓	✓
raskob2010	✓	E[7]	Y[8]	F[9]	✓	E[10]	E[11]	✓	✓
eriksson2014	✓	E[12]	E[13]	✓	E[14]	E[15]	E[16]	E[17]	Y[18]
yokote2011	Y[19]	✓	✓	✓	F[20]	✓	✓	✓	✓
zhang2014	H[21]	H[22]	H[23]	H[24]	H[25]	✓	H[26]	H[27]	U[28]
turpie2002	Y[29]	✓	✓	✓	Y[30]	✓	✓	✓	✓
lassen2002	Y[31]	✓	✓	✓	Y[32]	✓	✓	✓	✓
kim2016	Y[33]	Y[34]	Y[35]	✓	✓	✓	✓	G[36]	✓

✓= Outcome was reported and included in the review Meta Anlaysis

[1] No Events

[2] Can calculate symptomatic VTE (symptomatic DVT + Symptomatic PE) from Table 2 page 204.

[3] Symptomatic PE was measured and there were no deaths from any cause. Therefore nonfatal + fatal PE is known.

[4] No Events.

[5] The outcome that was reported and analysed is "Symptomatic VTE and Deaths from VTE." Symptomatic VTE was measured and likely analyzed.

[6] Can calculate symptomatic PE from Non-fatal and fatal PE from Table 2 on pg 2494.

[7] Symptomatic DVT and PE clearly were measured, likely analyzed. See pg 603.

[8] There was one symptomatic PE, and no Fatal PE. Therefore, we know total PE.

[9] No Events.

[10] Total DVT was measured and likely analysed. See pg 603.

[11] Symptomatic DVT was measured and likely analysed. See pg 603.

[12] Symptomatic VTE measured as secondary outcome but not reported, was likely analysed.

[13] See figure 2 and 3. Clear that PE was measured, but results underreported (bar graph, missing counts).

[14] Per appendix, was measured. Results reported unclearly in a bar graph, was analysed.

[15] Per appendix, total DVT was measured. Not reported in paper.

[16] Was measured as part of symptomatic VTE. Was likely analysed.

[17] Per appendix, was measured. Not reported in paper.

[18] "Of the patients with major bleeding events (darexaban, n=25; enoxaparin, n=8)," (219). Data is reported.

[19] Primary efficacy outcome appears to be total VTE. See description on pg 251.

[20] No Events.

[21–28] Only have abstract. Article is in Chinese.

[29] Primary efficacy outcome appears to be total VTE. See description on pg 1722.

[30] Table 3: Could sum symptomatic fatal PE and symptomatic non-fatal PE to get PE symptomatic.

[31] Primary efficacy outcome appears to be total VTE. See description on pg 1716.

[32] Table 3: Could sum symptomatic fatal PE and symptomatic non-fatal PE to get PE symptomatic.

[33] Can calculate total VTE (asymptomatic + symptomatic DVT + PE) from table 6, page 605.

[34] Can calculate symptomatic VTE (symptomatic DVT + PE) from table 6, page 605.

[35] Can calculate total PE (Fatal + Nonfatal) from table 6, pg 605.

[36] Duplex Doppler USG scans were used on the legs. Proximal versus distal likely measured and analysed.

Sensitivity Analysis

Using the ORBIT II matrix in Table 4, the sensitivity analysis is conducted in two stages. First, the "Y" event data is added to the meta-analyses and effect sizes are re-estimated. Then, the bias-adjusted estimates are calculated using the Copas 2017 random effects model (4).

The following tables are preliminary trial runs with the fixed effects code. "Y" event data not included. (This section is still in progress)

Table 5: ORB Adjustments For Benefit Outcomes

Outcome	Unadjusted OR	Unadj CI	Adjusted OR	Adj CI
Total VTE	2.233	(1.617 , 3.083)	1.976	(1.456 , 2.628)

Note: These are preliminary fixed effect results.

OR greater than 1 means favors FXaI. ORB adjustment towards the null.

Table 6: ORB Adjustments For Harm Outcomes

Outcome	Unadjusted OR	Unadj CI	Adjusted OR	Adj CI
Major Bleeding	0.739	(.547 , .997)	0.787	(.589, .787)

Note: These are preliminary fixed effect results

Direction of the OR estimate is unexpected. However, the CI is narrower and more significantly favors LMWH.

Discussion

References

1. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. BMJ Evid-Based Med. 2016 Aug 1;21(4):125-7.
2. Higgings J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011.
3. Rothstein HR, Sutton AJ, Borenstein M. Publication Bias in Meta-Analysis Prevention, Assessment and Adjustments [Internet]. New York, NY: John Wiley & Sons; 2006 [cited 2019 Feb 19]. Available from: http://nbn-resolving.de/urn:nbn:de:101:1-201411164755
4. Copas J, Marson A, Williamson P, Kirkham J. Model-based sensitivity analysis for outcome reporting bias in the meta analysis of benefit and harm outcomes. Stat Methods Med Res. 2017 Jan 1;962280217738546.

5. Copas J, Dwan K, Kirkham J, Williamson P. A model-based correction for outcome reporting bias in meta-analysis. Biostat Oxf Engl. 2014 Apr;15(2):370-83.
6. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ. 2010 Feb 15;340:c365.
7. Dwan K, Gamble C, Kolamunnage-Dona R, Mohammed S, Powell C, Williamson PR. Assessing the potential for outcome reporting bias in a review: a tutorial. Trials. 2010 May 12;11(1):52.
8. Saini P, Loke YK, Gamble C, Altman DG, Williamson PR, Kirkham JJ. Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews. BMJ. 2014 Nov 21;349:g6501.
9. Schmid CH. Outcome Reporting Bias: A Pervasive Problem in Published Meta-analyses. Am J Kidney Dis. 2017 Feb 1;69(2):172-4.
10. About AHRQ [Internet]. [cited 2019 Feb 26]. Available from: https://www.ahrq.gov/cpi/about/index.html
11. What Is Venous Thromboembolism? [Internet]. WebMD. [cited 2019 Feb 26]. Available from: https://www.webmd.com/dvt/what-is-venous-thromboembolism
12. Balk EM, Ellis AG, Di M, Adam GP, Trikalinos TA. Venous Thromboembolism Prophylaxis in Major Orthopedic Surgery: Systematic Review Update [Internet]. Agency for Healthcare Research and Quality (AHRQ); 2017 Jun [cited 2019 Feb 19]. Available from: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2480
13. Borenstein M. Introduction to meta-analysis: Michael Borenstein ... [et al. Chicester: Wiley; 2009.