

Method

Cite this article: Canestaro WJ, Devine EB, Bansal A, Sullivan SD, Carlson JJ (2022). Adjustment of publication bias using a cumulative meta-analytic framework. *International Journal of Technology Assessment in Health Care*, **38**(1), e61, 1–8
<https://doi.org/10.1017/S0266462322000435>

Received: 29 November 2021

Revised: 28 May 2022

Accepted: 14 June 2022


Key words:

Publication bias; Cumulative meta-analysis;
Statistical adjustment

Author for correspondence:

*W. J. Canestaro,
E-mail: canestaro@gmail.com

Adjustment of publication bias using a cumulative meta-analytic framework

W. J. Canestaro^{1,2*} , E. B. Devine¹, A. Bansal¹, S. D. Sullivan¹ and J. J. Carlson¹

¹School of Pharmacy, University of Washington, Seattle, WA, USA and ²Washington Research Foundation, Seattle, WA, USA

Abstract

Objectives: Publication bias has the potential to adversely impact clinical decision making and patient health if alternative decisions would have been made had there been complete publication of evidence.

Methods: The objective of our analysis was to determine if earlier publication of the complete evidence on rosiglitazone's risk of myocardial infarction (MI) would have changed clinical decision making at an earlier point in time. We tested several methods for adjustment of publication bias to assess the impact of potential time delays to identifying the MI effect. We then performed a cumulative meta-analysis (CMA) for both published studies (published-only data set) and all studies performed (comprehensive data set). We then created an adjusted data set using existing methods of adjustment for publication bias (Harbord regression, Peter's regression, and the nonparametric trim and fill method) applied to the limited data set. Finally, we compared the time to the decision threshold for each data set using CMA.

Results: Although published-only and comprehensive data sets did not provide notably different final summary estimates [OR = 1.4 (95 percent confidence interval [CI]: .95–2.05) and 1.42 (95 percent CI: 1.03–1.97)], the comprehensive data set reached the decision threshold 36 months earlier than the published-only data set. All three adjustment methods tested did not show a differential time to decision threshold versus the published-only data set.

Conclusions: Complete access to studies capturing MI risk for rosiglitazone would have led to the evidence reaching a clinically meaningful decision threshold 3 years earlier.

Evidence-based medicine (EBM) is “the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients” (1) and one of the most powerful concepts in modern health care. The process of EBM naturally involves three steps: generation, synthesis, and practice (2). The medium for this entire process is peer-reviewed scientific literature, but the pace of publishing these data has intensified. Recent bibliometric analysis has suggested that global scientific output as measured by the number of publications grows at a rate of 8–9 percent annually (3), meaning that the number of scientific publications more than doubles every 10 years. In health care alone, the National Library of Medicine's bibliographic database, MEDLINE, adds more than 800,000 new citations every year (4).

This productivity creates an unmanageable amount of information for healthcare practitioners to incorporate into the care of patients. It is ultimately the second step of EBM, the synthesis of this large volume of information via systematic reviews and meta-analyses, that translates the results of clinical research into actionable information for healthcare practitioners. The insight that we gain from the literature is compromised by forms of bias in the selection of studies or reporting of these data. Additionally, heterogeneity in the sample of studies might limit the usefulness of the analysis. Although heterogeneity in the sample of studies may be able to be quantified by either the Q or I^2 statistic, systematically missing studies may create a sample that is both internally consistent and systematically misdirected.

Publication bias represents a threat to the effectiveness of evidence synthesis and undermines one of the core tenets of EBM—that a systematic review of published evidence can create an accurate estimate of the known safety and efficacy of an intervention. In fact, many studies go unpublished, and those studies that do go unpublished are often systematically different from those that are published (5–7). This phenomenon is extensive in medicine: Nearly half of studies monitored by Institutional Review Boards go unpublished after data collection is completed (8–14), and almost 60 percent of trials submitted to the U.S. Food and Drug Administration (FDA) for approval of new treatments are never published in a peer-reviewed journal (15). This bias can arise for several reasons. For example, we know that publication is strongly correlated with the study treatment's effect size, direction, and significance (5–7), as well as sponsorship (16).

In the case of publication bias, cumulative meta-analysis (CMA) can be used to evaluate evidence accumulation by comparing a summary measure's level of significance over time for a comprehensive data set versus a published-only data set. Under a scenario where all studies were available, there would presumably be less uncertainty about the summary measure, and the CMA

would reach a significant finding earlier or with fewer studies completed than if there was only a partially reported set of studies. It is during this window of time that there is a potential for publication bias to have a real and measurable impact on patients and the healthcare system. Although the effect of publication bias on real-world decision making has been shown at a single time point (17), it has not yet been evaluated through time to determine what potential effect this could have on clinical decisions in the window of time during which the evidence existed but was not available.

Our objectives were to: (i) determine if there was a difference in time to clinically meaningful evidence for a notable historical case of publication bias, rosiglitazone for patients with diabetes, using the methods of CMA, and (ii) evaluate if available methods of adjustment for publication bias could be useful in arriving at clinically meaningful evidence sooner under our CMA framework.

Materials and methods

Our process to achieve this objective was threefold. First, we determined if there was a systematic difference in summary estimate via traditional meta-analysis between what was available in the published-only data set for rosiglitazone and what was available to the manufacturer in the comprehensive data set. Second, we measured how the time to our clinically meaningful decision (CMD) threshold would be affected by having access to a comprehensive data set rather than published-only data set using a CMA framework. Finally, we evaluated whether currently available methods for adjustment of publication bias in traditional meta-analysis when used under a CMA framework would affect the time to our CMD threshold. For our purposes, we defined clinically meaningful evidence via a measure we are naming the CMD threshold. The CMD threshold was defined as the level of evidence (as measured via treatment effect size of a healthcare outcome) at which either stated or revealed preferences (18) of healthcare providers would change their prescription patterns. This incorporates both the absolute measure of minimal clinically important difference (19;20) and uncertainty around this absolute value, and will vary from analysis to analysis based on the outcome, treatment, and empirically defined prescriber preferences.

Case Study

The case study used for this analysis was rosiglitazone (trade name, Avandia), an insulin sensitizer that works by sensitizing fat cells to make them more responsive to insulin. The drug was first approved by the U.S. FDA in 1999 based on the laboratory measure of blood sugar control via glycated hemoglobin (HbA1c) levels without sufficient numbers of patients to detect differences in clinical events such as heart attacks (Figure 1). Rosiglitazone ultimately rose to sales of more than \$3B in 2006 before a meta-analysis by Nissen and Wolski showed that it was associated with a higher rate of myocardial infarction (MI) and death from cardiovascular causes (odds ratios of 1.43 and 1.64, respectively) (21). This association is notable considering that greater than 50 percent of people with diabetes die from cardiovascular causes (22). The Nissen meta-analysis included forty-two qualifying studies, of which twenty-six were previously unpublished. Following the release of this information, sales for rosiglitazone dropped significantly, although the drug was never pulled from the market (23;24). Although a risk mitigation strategy was implemented for rosiglitazone, this was ultimately lifted in 2015 after the results of the Rosiglitazone evaluated for

cardiovascular outcomes in oral agent combination therapy for type 2 diabetes (RECORD) trial failed to show a risk of MI associated with the drug (25).

Data Source

We included the studies from the meta-analysis performed by Nissen and Wolski (21). Responding to calls to increase transparency following a suspected risk of cardiovascular events, GlaxoSmithKline (GSK), the manufacturer of rosiglitazone, created an online database of all studies for the agent that they had sponsored. Nissen and Wolski used this resource to conduct their meta-analysis. These data represent the most comprehensive list available and include both published and unpublished studies. Additionally, if there were intentional publication bias occurring, it would be performed by the manufacturer and captured in this database. For our analysis, we have replicated the analysis by Nissen and Wolski, which caused a notable shift in prescription patterns (23;24), and subsequently have used MIs as our outcome of interest. To normalize between studies that were published and those that were unpublished, we have used the study completion date as provided in the GSK database as the time at which results were available.

Traditional Meta-Analysis

As a first step, we conducted traditional meta-analyses with both the published-only and comprehensive data sets. This analysis was designed to replicate the analysis by Nissen and Wolski (26) to validate our results. To calculate the odds ratio of MI effect size, we used each arm's sample size as the denominator and performed no adjustment for arms with zero events. This effect was pooled across studies via the Peto method to faithfully recreate the original analysis (27). Nissen and Wolski suggest that the use of the Peto method was appropriate given the outcome of MI is relatively rare occurring in less than 1 percent of patients and the arms of the trials were relatively balanced in size (28;29). Although more recent statistical simulation studies (CITE BROCKHAUS 2016 paper) show that the Peto method might not be the most appropriate choice for this analysis, this insight would not have been available to the original authors, and so when given the option of updating their analysis given current understanding or remaining faithful to their choice of methods, we side with the authors original choice of method. For any trial with more than two arms, such as a trial comparing control to multiple doses of interventional drug, we combined intervention arms. Additionally, as there was significant variability in the duration of trials, we conducted a sensitivity analysis pooling with a person-year, rather than population denominator for effect.

Assessment of Heterogeneity and Publication Bias

To quantify the extent of heterogeneity in our analysis, we calculated the I^2 statistic (30) for each of the models evaluated. Additionally, to qualitatively assess publication bias or potential systematic missingness, we performed a funnel plot analysis for all our models. The funnel plot is a graph that displays studies included on the two axes of effect size and variance (31). Publication bias is then evaluated visually by looking for the characteristic asymmetrical pattern and missingness in the quadrant of the plot representing small sample sizes and effects (32). For our analysis, as we had both published and comprehensive sets of studies, we

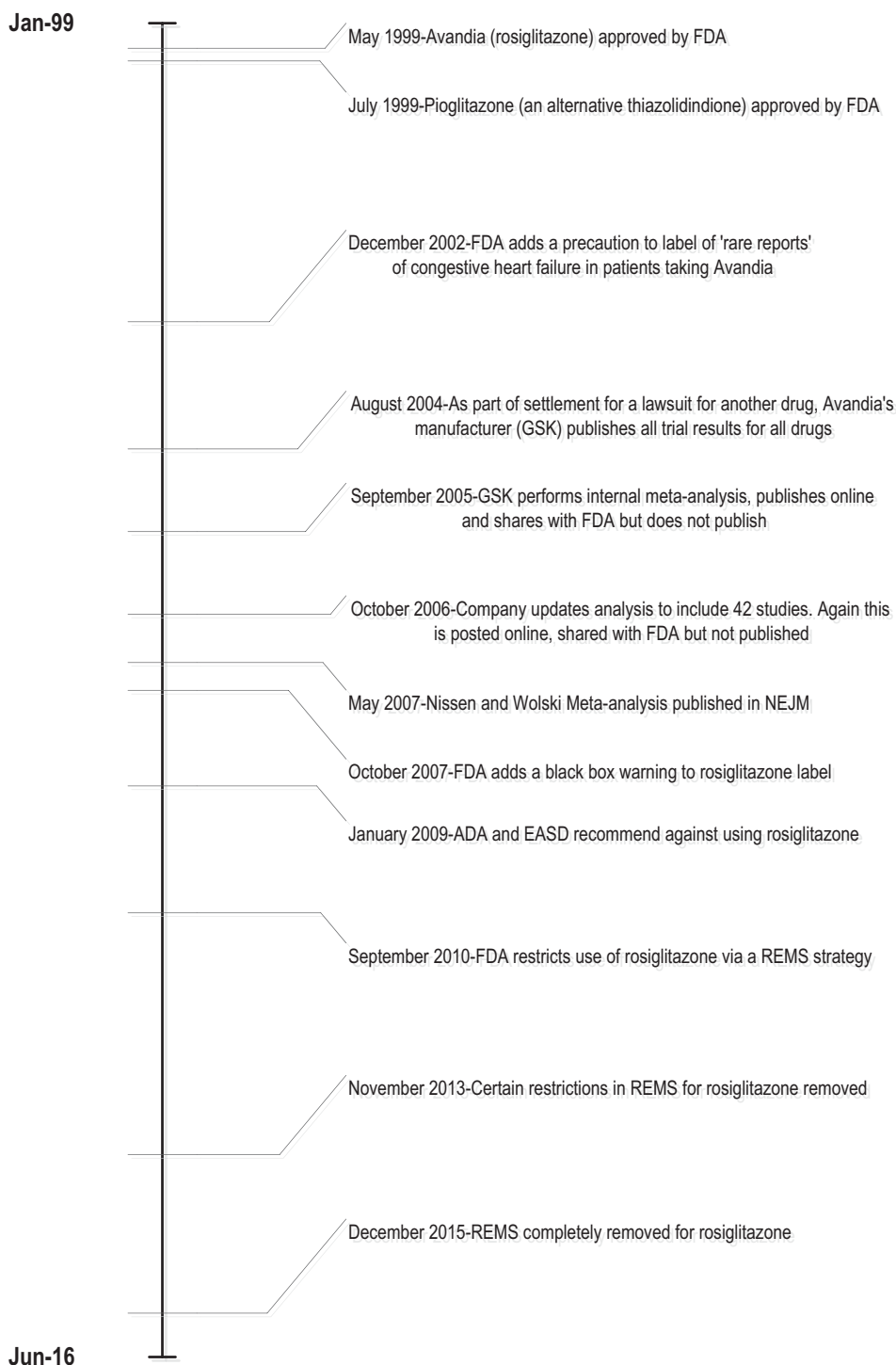


Figure 1. Timeline for rosiglitazone. ADA, American Diabetes Association; EASD, European Association for the Study of Diabetes; FDA, Food and Drug Administration; GSK, GlaxoSmithKline; NEJM, New England Journal of Medicine; REMS, risk evaluation and mitigation strategy.

evaluated the patterns of these plots side by side to see if the unpublished studies were within the generally predicted location or more dispersed.

Cumulative Meta-Analysis

Our analysis compares decisions made through time with a published-only data set to what would have been possible with

a comprehensive data set. The key method for performing this analysis is CMA. For uniformity, we have applied all the same methods from our traditional meta-analysis under a CMA framework. We defined the CMD threshold as the date at which the confidence interval for the odds ratio of rosiglitazone's risk of MI excluded the null. We chose this threshold as it resembled the level of evidence presented in the Nissen and Wolski analysis (26), which resulted in a distinct shift in

prescriber patterns. A claims-based analysis of prescriptions for diabetes medications showed that, following the Nissen and Wolski analysis (26) publication, there was a greater than 50 percent reduction in the number of rosiglitazone prescriptions (23;24). This gives us validation that the level of evidence presented by Nissen and Wolski was clinically meaningful and enough to shift decision making. Underlying this analysis is the assumption that, had this same level of evidence been available sooner, it would have produced a similar response among prescribers.

Using this method, we can estimate the time interval during which decision making may have been affected by systematically biased information. To concisely capture the results of these analyses, we have recorded for each CMA the date at which the odds ratio for MI crossed our predefined CMD threshold for both comprehensive and published-only data sets.

Methods of Adjustment Applied via Cumulative Meta-Analysis

Next we evaluated if any of the currently available methods for adjustment for publication bias altered the time to evidence crossing a CMD threshold. To evaluate this, we used three of the most commonly used methods for publication bias adjustment: Harbord's regression, Peter's regression, and the Trim and Fill. More detail on these methods and how they adjust for missingness can be found in the Technical Appendix in the Supplementary Material.

Measurements of Performance

To measure the relative performance of each method of adjustment in a CMA framework, we considered how closely the adjusted data set matched the timing of effect estimates for the comprehensive data set crossing our predefined CMD threshold. For example, in a hypothetical example, if a CMA for the effect of Drug A on Outcome B using a comprehensive set of studies, published studies with adjustment, or only the published set of studies crossed the predefined CMD threshold in months 1, 4, and 12, respectively, then that method for adjustment would receive a score of 8 months since it identified the meaningful effect 8 months sooner than would have been possible with no adjustment.

Results

Traditional Meta-Analyses

Our meta-analysis of the odds ratio of MI in rosiglitazone with population denominators showed little difference in the point estimate between our published and comprehensive data sets with estimates of 1.40 (95 percent confidence interval (CI): .95–2.05) and 1.42 (95 percent CI: 1.03–1.97), respectively (Figure 2). Importantly, the confidence interval for the effect estimate in the comprehensive data set excludes the null and creates a statistically significant result. Our results were not affected by the use of a person-year denominator (Supplementary Table 2). For all of our models, the I^2 statistic was 0 percent, suggesting extremely low

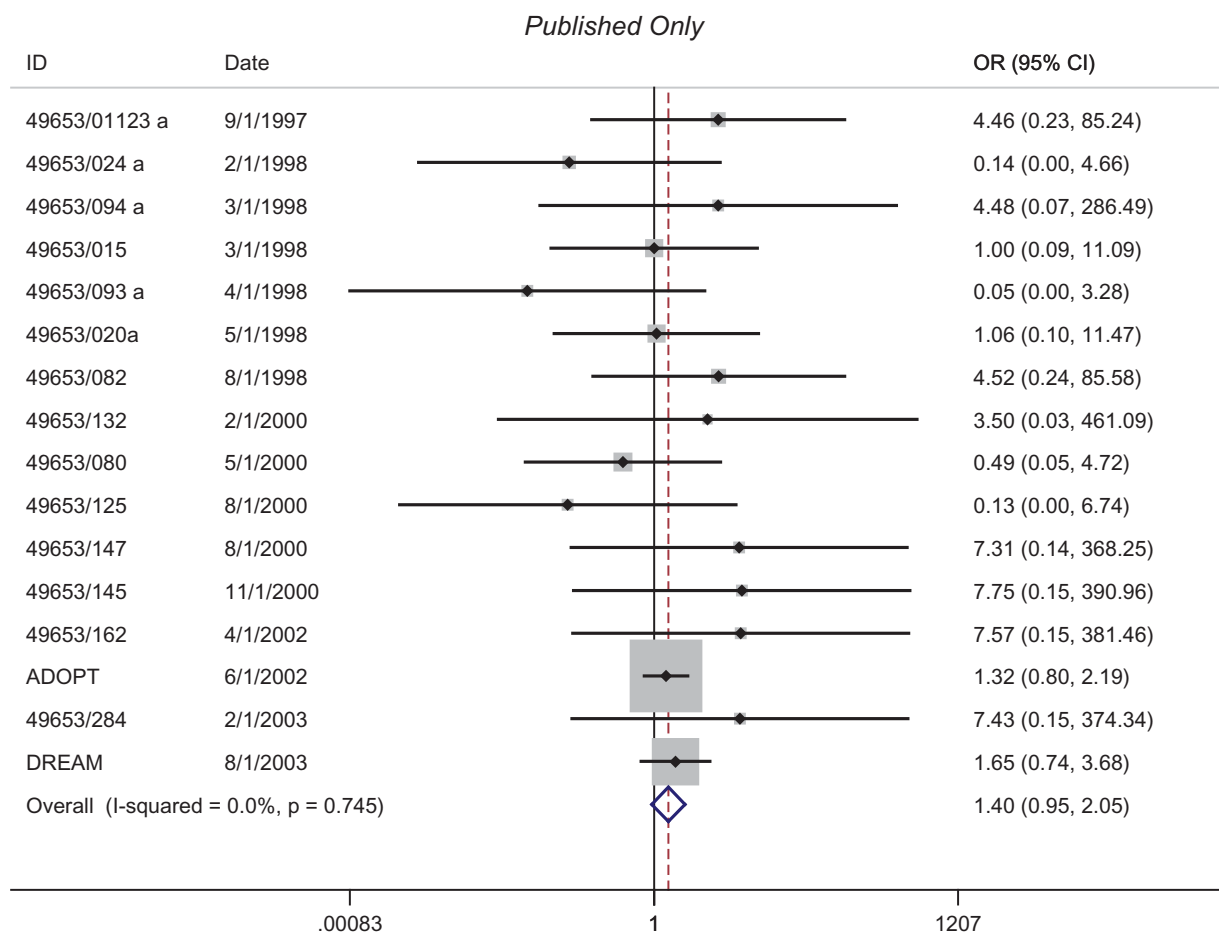


Figure 2. Meta-analysis of myocardial infarction events in rosiglitazone trials.

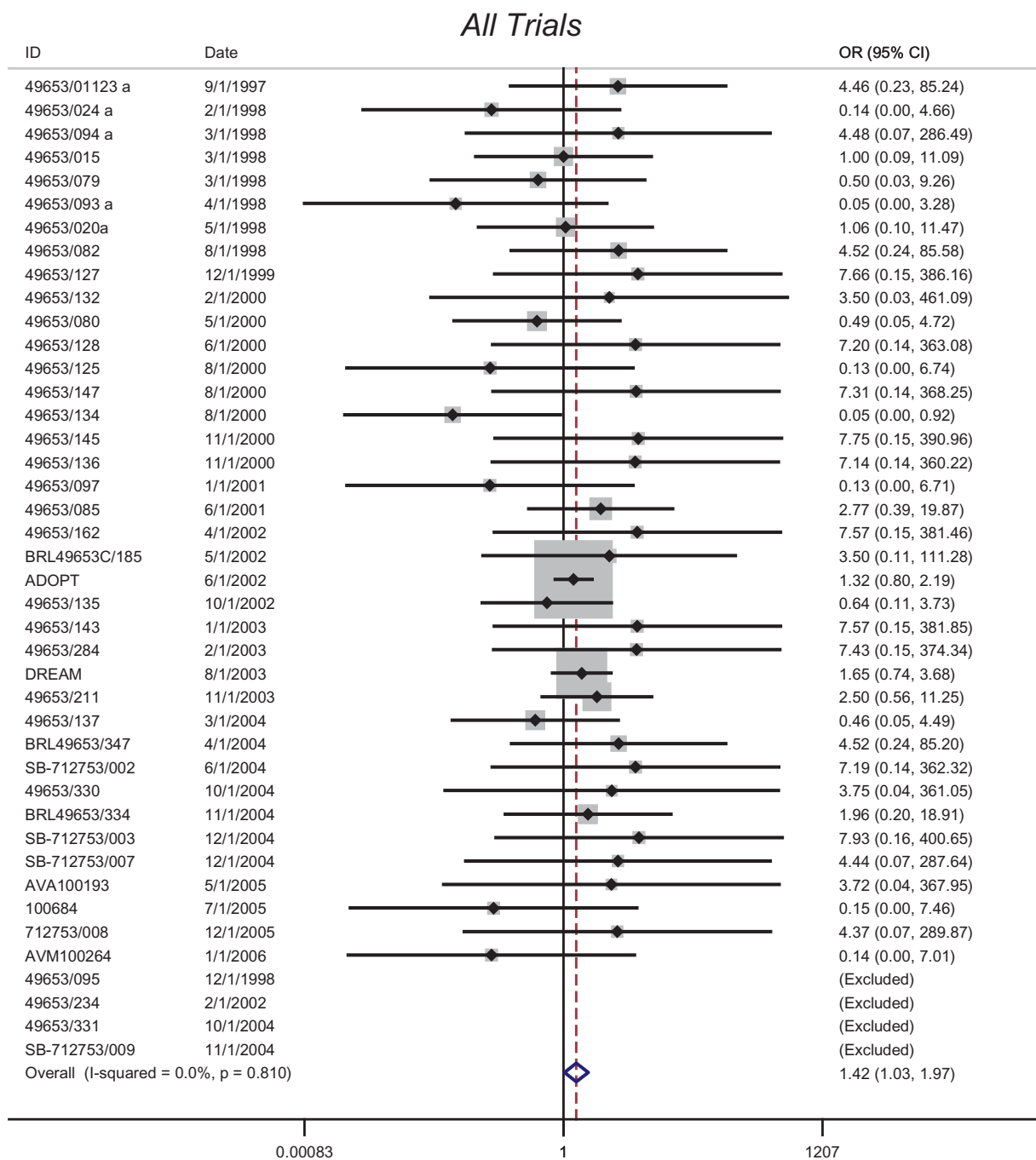


Figure 2. Continued.

levels of heterogeneity (30). Finally, a visual comparison of the funnel plots for published versus all evidence did not suggest the characteristic pattern of missingness associated with publication bias. This pattern was similar across both population and person-year denominators (Supplementary Figure 1).

Cumulative Meta-Analyses and Time to Clinically Meaningful Decision Threshold

Our CMA for both comprehensive and published-only data sets showed a rather large differential in time to crossing our predefined

threshold. In fact, the evidence for MI risk would have never reached statistical significance using solely published data as MI risk did not achieve statistical significance until the comprehensive data set was published by Nissen Wolski in 2006 (Figure 3 and Supplementary Table 3). Having access to all studies gives a clinically meaningful result a full 36 months sooner. Under the published-only scenario, the available evidence only showed a significant effect after the full analysis by Nissen and Wolski. Had all the studies been accessible and published, this same level of evidence would have been reached in June 2004 (Supplementary Table 3).

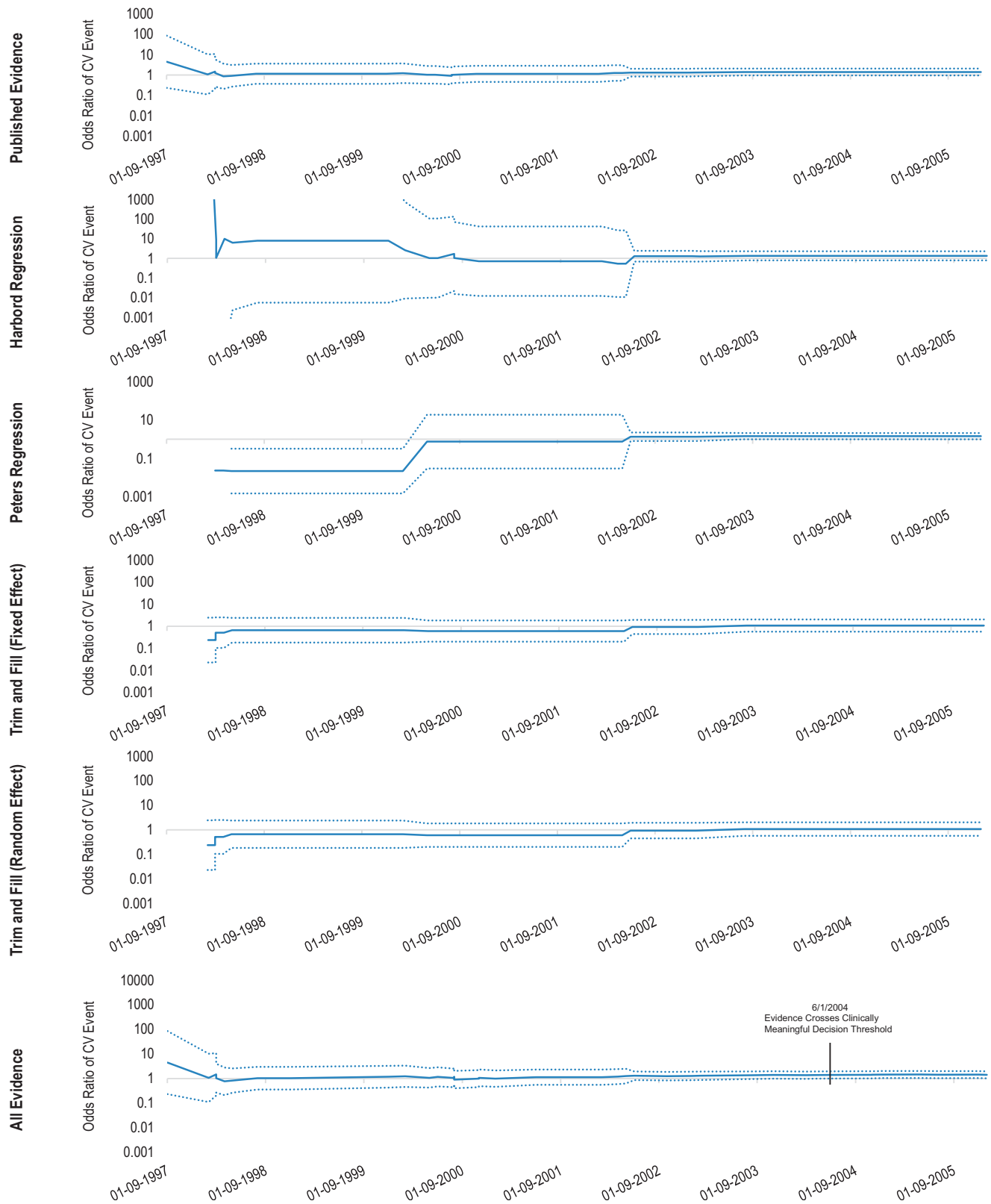


Figure 3. Cumulative meta-analysis for published-only, comprehensive, and adjusted data sets. Each panel displays a cumulative analysis of a data set. The solid line represents the point estimate for the odds ratio of myocardial infarction. The dotted line is the 95 percent confidence interval. The statistical model used is the Peto method with a population denominator. The middle four panels represent adjusted estimates using different methods for estimating comprehensive estimates from the published-only data set.

Performance of Adjustment Methods via a Cumulative Framework

Following our CMA of published-only and cumulative data sets, we performed four adjustment methods in a cumulative framework to assess the degree to which each available method of adjustment corrected the summary estimate in the published-only data set toward the comprehensive estimate (Technical Appendix in the Supplementary Material). This process was only conducted for our primary statistical model: the Peto method of pooling odds ratio with a population denominator.

None of our methods for adjustment shortened the time differential to our CMD threshold (Figure 3). In terms of adjustment at the time of the final study, Peter's regression (OR = 1.44 [95 percent CI: .99–2.10]) appeared to perform better than the Harbord (OR = 1.34 [95 percent CI: .79–2.29]), which slightly underestimated the effect, although both were reasonably close to the effect estimate for all studies (OR = 1.42 [95 percent CI: 1.03–1.97]). Under the Trim and Fill, using either a fixed or random-effects model gave the same effect, which was a notable underestimate of the true effect. Again, all methods showed the characteristic narrowing of the confidence interval for a CMA, although none had a significant effect.

Discussion

Publication bias is a pervasive problem in health care, which limits the accurate synthesis of clinical trials. Although the release of previously unpublished trials has been found to influence clinical decisions, the impact of this evidence over time has not yet been evaluated. Our objective was to determine if access to a comprehensive data set would reduce time to a CMD threshold, a significant risk of MI, relative to a published-only data set, and if any currently available methods for adjustment could shorten this differential. Previous analyses have evaluated these adjustment methods in simulated (33) and real (17) data sets, but none have tested them in the context of CMA, which is more useful in quantifying the potential impact through the time of unpublished evidence on clinical practice.

The difference in time to crossing CMD threshold in the published-only versus comprehensive data sets was 36 months. Known methods of adjustment for publication bias in traditional meta-analysis were not able to shorten this 36-month differential. Specifically, we found that these adjustment methods, although somewhat helpful at indicating an adjusted point estimate, widely expand the confidence interval and do not appear to be as useful in adjustment and reducing uncertainty under a CMA framework.

The common definition of publication bias is that the set of published studies is systematically different or unrepresentative of the whole set of completed studies. This would be visible on a funnel plot and could be corrected by known adjustment methods. *A priori*, we sought to see if there was evidence of this traditional publication bias and how adjustment methods might have influenced the time to an important decision. Our evidence showed that the unpublished studies were not systematically different and therefore did not fit this traditional. That being said, they added enough statistical power to allow greater certainty at an earlier time point, thereby shortening the time to this clinically important decision.

The difference in timing between when evidence of an effect could have been found from a comprehensive set of studies and

when it was available for decision-makers is potentially meaningful and impactful. It is during this window of time that patients are receiving treatment with less than complete evidence for risk or benefit that could influence their decision if they had known about it. Unfortunately, none of the methods of adjustment for publication bias shortened this time frame. This suggests, albeit from a single yet high-profile case study, that adjustment methods may be insufficient to address publication bias if it is suspected or even established. In fact, the only way to shorten the time to more informed clinical decision making is a complete and timely publication of all evidence.

Our analysis has several notable limitations. First, we used a historical example; therefore, we did not perform a comprehensive simulation to assess the performance of these methods under alternative forms of publication bias. For example, all the methods that we used adjusted for result size and significance, but would not correct for any bias due solely to factors such as sponsorship. Additionally, although we have evaluated our case retrospectively with the full benefit of hindsight, this same method may be less generalizable when performed prospectively as a form of sequential testing. As has been shown in the monitoring of clinical trials, sequential testing of a data set through time with a static threshold for decision making has the potential to lead to a higher number of false-positive results (34). This suggests that caution should be taken in applying these methods to prospective cases.

Additionally, the benefit of hindsight suggests that the results of these analyses should be evaluated in a historical context. In particular, although we found an effect of rosiglitazone on MI as early as 2004, later results of the RECORD trial failed to show a risk of MI associated with the drug (25), resulting in the FDA removing its requirement for a risk mitigation strategy on the part of the manufacturer. Although early safety signals were contradicted by later evidence, this does not totally negate their usefulness. Had the safety signal been identified and acted upon sooner, postmarketing studies such as RECORD could have been launched sooner.

Conclusion

In a well-known historical case, incomplete publication of clinical trial results resulted in a 3-year delay to a clinically meaningful evidence threshold. During this 3 year period of time many more clinicians were potentially prescribing rosiglitazone than would have if the complete set of studies been made available. Although currently available methods of adjustment for publication bias exist, they did not reduce this time to a CMD threshold.

Supplementary Materials. To view supplementary material for this article, please visit <http://doi.org/10.1017/S0266462322000435>.

Conflicts of interest. The authors have no conflict of interest.

References

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS (1996) Evidence based medicine: What it is and what it isn't. *BMJ* 312, 71–72.
2. Sackett DL, Rosenberg WM (1995) The need for evidence-based medicine. *J R Soc Med* 88, 620–624.
3. Bornmann L, Mutz R (2015) Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *J Assoc Inf Sci Technol* 66, 2215–2222.

4. **Medicine USNLo** (2017) *Detailed indexing statistics: 1965–2016*. Available at: https://www.nlm.nih.gov/bsd/index_stats_comp.html. Accessed 2017.
5. **Ioannidis JP** (1998) Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA* **279**, 281–286.
6. **Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K** (2009) Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev* **2009**, MR000006.
7. **Hopewell S, Clarke M, Stewart L, Tierney J** (2007) Time to publication for results of clinical trials. *Cochrane Database Syst Rev* **2007**, MR000011.
8. **Blumle A, Antes G, Schumacher M, Just H, von Elm E** (2008) Clinical research projects at a German medical faculty: Follow-up from ethical approval to publication and citation by others. *J Med Ethics* **34**, e20.
9. **von Elm E, Rollin A, Blumle A, et al** (2008) Publication and non-publication of clinical trials: Longitudinal study of applications submitted to a research ethics committee. *Swiss Med Wkly* **138**, 197–203.
10. **Stern JM, Simes RJ** (1997) Publication bias: Evidence of delayed publication in a cohort study of clinical research projects. *BMJ (Clin Res Ed)* **315**, 640–645.
11. **Pich J, Carne X, Arnaiz JA, et al** (2003) Role of a research ethics committee in follow-up and publication of results. *Lancet* **361**, 1015–1016.
12. **Decullier E, Lheritier V, Chapuis F** (2005) Fate of biomedical research protocols and publication bias in France: Retrospective cohort study. *BMJ (Clin Res Ed)* **331**, 19.
13. **Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR** (1991) Publication bias in clinical research. *Lancet* **337**, 867–872.
14. **Dickersin K, Min YI, Meinert CL** (1992) Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA* **267**, 374–378.
15. **Lee K, Bacchetti P, Sim I** (2008) Publication of clinical trials supporting successful new drug applications: A literature analysis. *PLoS Med* **5**, e191.
16. **Lundh A, Sismondo S, Lexchin J, Busuioic OA, Bero L** (2012) Industry sponsorship and research outcome. *Cochrane Database Syst Rev* **12**, MR000033.
17. **Moreno SG, Sutton AJ, Turner EH, et al** (2009) Novel methods to deal with publication biases: Secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications. *BMJ (Clin Res Ed)* **339**, b2981.
18. **Mark TL, Swait J** (2004) Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Health Econ* **13**, 563–573.
19. **Wells G, Beaton D, Shea B, et al** (2001) Minimal clinically important differences: Review of methods. *J Rheumatol* **28**, 406–412.
20. **Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC** (2007) Understanding the minimum clinically important difference: A review of concepts and methods. *Spine J* **7**, 541–546.
21. **Nissen SE, Wolski K** (2007) Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med* **356**, 2457–2471.
22. **Morrish NJ, Wang SL, Stevens LK, Fuller JH, Keen H** (2001) Mortality and causes of death in the WHO multinational study of vascular disease in diabetes. *Diabetologia* **44**, S14–S21.
23. **Starner CI, Schafer JA, Heaton AH, Gleason PP** (2008) Rosiglitazone and pioglitazone utilization from January 2007 through May 2008 associated with five risk-warning events. *J Manag Care Pharm* **14**, 523–531.
24. **Morrow RL, Carney G, Wright JM, et al** (2010) Impact of rosiglitazone meta-analysis on use of glucose-lowering medications. *Open Med* **4**, e50–e59.
25. **Home PD, Pocock SJ, Beck-Nielsen H, et al** (2009) Rosiglitazone evaluated for cardiovascular outcomes in oral agent combination therapy for type 2 diabetes (RECORD): A multicentre, randomised, open-label trial. *Lancet* **373**, 2125–2135.
26. **Nissen SE, Wolski K** (2007) Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med* **356**, 2457–2471.
27. **Yusuf S, Peto R, Lewis J, Collins R, Sleight P** (1985) Beta blockade during and after myocardial infarction: An overview of the randomized trials. *Prog Cardiovasc Dis* **27**, 335–371.
28. **Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A** (2007) Much ado about nothing: A comparison of the performance of meta-analytical methods with rare events. *Stat Med* **26**, 53–77.
29. **Higgins JP, Green S** (2011) *Cochrane handbook for systematic reviews of interventions*, vol. 4. Chichester, UK: Wiley.
30. **Higgins J, Thompson SG** (2002) Quantifying heterogeneity in a meta-analysis. *Stat Med* **21**, 1539–1558.
31. **Sterne JA, Egger M** (2001) Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *J Clin Epidemiol* **54**, 1046–1055.
32. **Song F, Khan KS, Dinnes J, Sutton AJ** (2002) Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* **31**, 88–95.
33. **Moreno SG, Sutton AJ, Ades AE, et al** (2009) Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Med Res Methodol* **9**, 2.
34. **Emerson SS, Fleming TR** (1990) Interim analyses in clinical trials. *Oncology* **4**, 126–133; Discussion 134, 136.