

Inconsistent results for Peto odds ratios in multi-arm studies, network meta-analysis and indirect comparisons

Guido Schwarzer¹  | Orestis Efthimiou²  | Gerta Rücker¹ 

¹Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany

²Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

Correspondence

Guido Schwarzer, Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany.
Email: sc@imbi.uni-freiburg.de

Present address

Guido Schwarzer, Stefan-Meier-Straße 26, D-79104 Freiburg, Germany.

Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: RU1747/1-2; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Ambizione Grant No., Grant/Award Number: 180083

Abstract

The Peto odds ratio is a well-known effect measure in meta-analysis of binary outcomes. For pairwise comparisons, the Peto odds ratio estimator can be severely biased in the situation of unbalanced sample sizes in the two treatment groups or large treatment effects. In this publication, we evaluate Peto odds ratio estimators in the setting of multi-arm studies and in network meta-analysis using illustrative examples. We observe that Peto odds ratio estimators in a multi-arm study are inconsistent if the observed event probabilities are different or the sample sizes of treatment groups are unbalanced. The same problem emerges in network meta-analysis including only two-arm studies and translates to indirect comparisons of pairwise meta-analyses. We conclude that the Peto odds ratio should not be used as effect measure in network meta-analysis or indirect comparisons of pairwise meta-analyses.

KEYWORDS

binary outcome, inconsistency, indirect comparison, multi-arm study, network meta-analysis, Peto odds ratio

Highlights

What is already known

The Peto odds ratio estimand in a single study is different from the odds ratio if the true treatment effect is large or group sample sizes are unbalanced. This problem transfers to the Peto odds ratio estimator and can thus impact meta-analysis results.

What is new

Peto odds ratio estimators from a multi-arm study are inconsistent if the observed event probabilities are different or sample sizes are unbalanced across treatment groups. This inconsistency also impacts network meta-analyses and indirect comparisons of pairwise meta-analyses.

Potential impact for RSM readers outside the authors' field

The Peto odds ratio should not be considered as an effect measure in network meta-analysis or indirect comparisons of pairwise meta-analyses.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Research Synthesis Methods* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

The Peto method¹ is a traditional method for meta-analysis of binary outcomes. It is a variant of the generic inverse variance method² using the logarithm of the Peto odds ratio and its standard error from each study. The Peto method is implemented in common software for meta-analysis: RevMan Web from Cochrane, Comprehensive Meta-Analysis, Stata, and several R packages including meta and metafor.

Yusuf et al¹ used likelihood theory to define the Peto odds ratio in a study, based on the difference between the observed and expected number of events in the experimental treatment group under the assumption of no treatment difference. The logarithm of the Peto odds ratio estimator corresponds to the first step of the Newton–Raphson algorithm away from the null effect to the maximum likelihood of the log odds ratio. Accordingly, the Peto method is sometimes called the one-step method.³

The popularity of the Peto method received a major setback with the publication by Greenland and Salvan³ demonstrating—through some hypothetical examples—that the Peto odds ratio estimator in a study can be severely biased in the situation of unbalanced sample sizes in the two treatment groups or large treatment effects (i.e., away from the null effect) impacting the result of the meta-analysis. The method regained some of its reputation after a large simulation study⁴ in meta-analysis with rare binary events. The Peto method was the least biased and most powerful meta-analysis method for event rates below 1 percent, however, this was only in situations with no substantial imbalance in group sample sizes and no excessive treatment effects.

It has been argued that the Peto odds ratio in a study should be considered an alternative effect measure.⁵ Using the delta method, the authors derived the limit of the expected Peto odds ratio. This estimand contains the ratio of the group sample sizes as a component. This peculiar property of the Peto odds ratio is not shared by other estimands for binary data. In addition, extensive simulations of single studies with rare binary events by the same group⁶ showed that the Peto odds ratio estimator does not outperform the usual odds ratio estimator in all performance measures, even in the optimal setting of equal group sample sizes and modest treatment effects.

Accordingly, the Peto method is nowadays not recommended as the default approach for meta-analysis.^{6–8}

All publications summarized so far evaluated the Peto odds ratio either in the setting of a single study with two treatment groups, that is, a single two-by-two table, or in a pairwise meta-analysis, that is, one two-by-two table per study. In this publication, we will look at the Peto odds ratio in the setting of multi-arm studies and in

network meta-analysis. Like Greenland and Salvan,³ we will use illustrative examples to show that another problem exists for Peto odds ratio estimators in multi-arm studies as well as network meta-analysis and indirect comparisons of pairwise meta-analyses.

2 | ODDS RATIO AND PETO ODDS RATIO IN MULTI-ARM STUDIES

Let us assume we have data from a single study comparing three treatment groups with respect to a binary outcome, that is, the number of events in each group follows a binomial distribution,^{9,p. 39}. The true (but unknown) probabilities of the event of interest are denoted by p_1 , p_2 , and p_3 , respectively. Estimates of p_1 , p_2 , and p_3 are given by $\hat{p}_i = x_i/n_i$, for $i = 1, 2, 3$, with number of events x_1 , x_2 , x_3 and group sample sizes n_1 , n_2 , n_3 .

2.1 | Odds ratio

The odds ratio OR_{12} comparing group 1 and 2 is defined as

$$OR_{12} = \left(\frac{p_1}{1-p_1} \right) / \left(\frac{p_2}{1-p_2} \right).$$

Odds ratios OR_{13} , OR_{23} , OR_{21} , OR_{31} and OR_{32} are defined accordingly.

By construction, the odds ratios in a three-arm study are consistent which means, for example, that the product of $OR_{12} \cdot OR_{23} \cdot OR_{31} = 1$. Or, alternatively, that $OR_{12} = OR_{13}/OR_{23}$ which means that the odds ratio of the direct comparison of treatment 1 versus 2 is equal to the odds ratio of the indirect comparison, that is, OR_{13}/OR_{23} .

2.2 | Peto odds ratio

The estimated Peto odds ratio \widehat{POR}_{12} comparing treatment 1 and 2 is given by

$$\widehat{POR}_{12} = \exp \left(\frac{x_1 - E(x_1 | \dots; OR_{12} = 1)}{\text{Var}(x_1 | \dots; OR_{12} = 1)} \right)$$

where $E(x_1 | \dots; OR_{12} = 1)$ and $\text{Var}(x_1 | \dots; OR_{12} = 1)$ are the mean and variance of x_1 under the hypergeometric distribution with “...” denoting the four (observed) cell margins. Under this distribution, we have

$$E(x_1 | \dots; OR_{12} = 1) = \frac{n_1}{n_1 + n_2} (x_1 + x_2)$$

and

$$\text{Var}(x_1 | \dots; \text{OR}_{12} = 1) = \frac{n_1 n_2 (x_1 + x_2) (n_1 + n_2 - x_1 - x_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}.$$

Estimated Peto odds ratios $\widehat{\text{POR}}_{13}$, $\widehat{\text{POR}}_{23}$, $\widehat{\text{POR}}_{21}$, $\widehat{\text{POR}}_{31}$, and $\widehat{\text{POR}}_{32}$ are defined accordingly.

It is not straightforward to determine if and when the set of estimated Peto odds ratios is consistent in the way the odds ratios are, as shown in section 2.1. Therefore, we decided to evaluate this using real and artificial examples.

2.3 | Illustrative examples

For illustration, we will use data from the two three-arm studies included in a network meta-analysis to prevent bleeding in cirrhosis.¹⁰ The actual data is listed in Table 1. We use R function pairwise from R package netmeta¹¹ to conduct all analyses for a single three-arm study. The R code is available as a supplement.

2.3.1 | Original study data

The estimated odds ratios for study 1 in Table 1 are $\widehat{\text{OR}}_{12} = 0.1789$, $\widehat{\text{OR}}_{13} = 0.1051$ and $\widehat{\text{OR}}_{23} = 0.5874$. Accordingly, the indirect estimate for the comparison of treatment 1 versus 2 is $\widehat{\text{OR}}_{13}/\widehat{\text{OR}}_{23} = 0.1051/0.5874 = 0.1789$ which is identical to the direct estimate. We can also calculate the product of the loop 1-2-3 as $0.1789 \cdot 0.5874 \cdot (1/0.1051) = 0.9999$ (which is slightly different from 1 due to rounding errors). We see that the estimated odds ratios are consistent.

Next we estimate the three Peto odds ratios for study 1: $\widehat{\text{POR}}_{12} = 0.2296$, $\widehat{\text{POR}}_{13} = 0.1616$ and $\widehat{\text{POR}}_{23} = 0.5938$, respectively. The indirect Peto estimate for the comparison of treatment 1 versus 2 is $\widehat{\text{POR}}_{13}/\widehat{\text{POR}}_{23} = 0.1616/0.5938 = 0.2722$ which obviously differs from the

direct Peto odds ratio. The magnitude of the discrepancy is more visible looking at the direct and indirect estimates for the comparison of treatment 2 versus 1: $1/0.2296 = 4.36$ (direct) versus $1/0.2722 = 3.67$ (indirect).

For study 2 in Table 1, the estimated odds ratios are $\widehat{\text{OR}}_{12} = 0.9890$, $\widehat{\text{OR}}_{13} = 0.9725$ and $\widehat{\text{OR}}_{23} = 0.9833$. The Peto odds ratios are very similar: $\widehat{\text{OR}}_{12} = 0.9891$, $\widehat{\text{OR}}_{13} = 0.9727$ and $\widehat{\text{OR}}_{23} = 0.9834$. Accordingly, the direct and indirect Peto odds ratios are also very similar for study 2, however, not identical: 0.98909 (direct) versus 0.98910 (indirect).

In the following subsections, to further elucidate the properties of the Peto odds ratio, we explore a set of artificially created data taking study 1 as our starting point.

2.3.2 | Variant 1: Same effect and equal sample sizes in active treatment groups

In Table 2 (top), we assume that the number of events and sample sizes in study 1 are the same under the two active treatments (beta blockers and sclerotherapy).

The estimated odds ratios for this hypothetical example are $\widehat{\text{OR}}_{12} = 1.0000$, $\widehat{\text{OR}}_{13} = 0.1051$ and $\widehat{\text{OR}}_{23} = 0.1051$, respectively. Accordingly, the indirect estimate for the comparison of treatment 1 versus 2 is $\widehat{\text{POR}}_{13}/\widehat{\text{POR}}_{23} = 0.1051/0.1051 = 1.0000$ which, again, is consistent.

The Peto odds ratios are $\widehat{\text{POR}}_{12} = 1.0000$, $\widehat{\text{POR}}_{13} = 0.1616$ and $\widehat{\text{POR}}_{23} = 0.1616$, respectively. The indirect Peto estimate for the comparison of treatment 1 versus 2 is $\widehat{\text{POR}}_{13}/\widehat{\text{POR}}_{23} = 0.1616/0.1616 = 1.0000$. In this setting we have consistent Peto odds ratios.

2.3.3 | Variant 2: Same effect and different sample sizes in active treatment groups

Next, we assume again that the event probability is the same under the two active treatments, however, that the sample size is twice as large for the second active treatment (Table 2, variant 2).

TABLE 1 Original data of study 1 and 2 from network meta-analysis to prevent bleeding in cirrhosis¹⁰

Study 1		Bleeding	No bleeding	Total
Beta-blockers	(Treatment 1)	2	41	43
Sclerotherapy	(Treatment 2)	9	33	42
Control	(Treatment 3)	13	28	41
Study 2		Bleeding	No bleeding	Total
Beta-blockers	(Treatment 1)	12	56	68
Sclerotherapy	(Treatment 2)	13	60	73
Control	(Treatment 3)	13	59	72

Variant 1		Bleeding	No bleeding	Total
Beta-blockers	(Treatment 1)	2	41	43
Sclerotherapy	(Treatment 2)	2	41	43
Control	(Treatment 3)	13	28	41
Variant 2		Bleeding	No bleeding	Total
Beta-blockers	(Treatment 1)	2	41	43
Sclerotherapy	(Treatment 2)	4	82	86
Control	(Treatment 3)	13	28	41
Variant 3		Bleeding	No bleeding	Total
Beta-blockers	(Treatment 1)	2	4298	4300
Sclerotherapy	(Treatment 2)	4	8596	8600
Control	(Treatment 3)	13	4087	4100
Variant 4		Bleeding	No bleeding	Total
Beta-blockers	(Treatment 1)	2	39	41
Sclerotherapy	(Treatment 2)	4	37	41
Control	(Treatment 3)	13	28	41

Note: Variant 1: Same event probabilities and sample sizes in active treatment groups; Variant 2: Same event probabilities and unequal sample sizes in active treatment groups; Variant 3: Same event probabilities, unequal sample sizes and rare events; Variant 4: Equal sample sizes, but different event probabilities.

The estimated odds ratios for this hypothetical example are unchanged: $\widehat{OR}_{12} = 1.0000$, $\widehat{OR}_{13} = 0.1051$ and $\widehat{OR}_{23} = 0.1051$. Accordingly, the indirect estimate for the comparison of treatment 1 versus 2 is the same (and consistent).

The Peto odds ratio for the comparison of the two active treatments is still the same: $\widehat{POR}_{12} = 1.0000$. However, the Peto odds ratios for the comparisons of the two active treatments with control differ substantially: $\widehat{POR}_{13} = 0.1616$ and $\widehat{POR}_{23} = 0.0987$, respectively. Accordingly, the indirect Peto estimate for the comparison of treatment 1 versus 2 is $\widehat{POR}_{13}/\widehat{POR}_{23} = 0.1616/0.0987 = 1.6367$ which is dramatically different from the direct estimate.

2.3.4 | Variant 3: Rare events, same effect and different sample sizes

In Table 2, variant 3, we additionally assume that events are rare by multiplying the group sample sizes by 100. The estimated odds ratios for this hypothetical example are consistent (results not shown). However, the Peto odds ratios comparing the two active treatments with placebo are again substantially different: $\widehat{POR}_{13} = 0.2192$ and $\widehat{POR}_{23} = 0.1322$, respectively. The inconsistency between direct and indirect Peto odds ratio for the comparison of treatment 1 versus 2 is slightly even more pronounced than in the previous example: $\widehat{POR}_{13}/\widehat{POR}_{23} = 0.2192/0.1322 = 1.6581$.

TABLE 2 Variants of study 1 from network meta-analysis to prevent bleeding in cirrhosis¹⁰

2.3.5 | Variant 4: Different effects and equal sample sizes

Finally, we consider a situation with equal group sample sizes, but different treatment effects (Table 2, variant 4). The estimated odds ratios are $\widehat{OR}_{12} = 0.4744$, $\widehat{OR}_{13} = 0.1105$ and $\widehat{OR}_{23} = 0.2328$ which are consistent. The estimated Peto odds ratios are $\widehat{POR}_{12} = 0.4914$, $\widehat{POR}_{13} = 0.1698$ and $\widehat{POR}_{23} = 0.2673$. The Peto odds ratio for the indirect comparison of treatments 1 and 2, 0.6353, is again substantially different from the direct estimate.

3 | PETO ODDS RATIO IN NETWORK META-ANALYSIS AND INDIRECT COMPARISONS

The three estimated treatment effects in a three-arm study are not independent as the results of the first and second comparisons determine the third result. Two approaches exist to account for the dependency in order to include multi-arm studies in network meta-analysis.¹² We can either only consider comparisons with a reference treatment (called basic parameters) or we can consider all pairwise comparisons but increase the standard error of each comparison. Higgins et al¹⁰ used the first approach while R package netmeta implements the second approach.

Based on our results described in the previous section, we would argue against using the Peto odds ratio in a

TABLE 3 Depicting data from first three-arm study¹⁰ as three independent two-arm studies

Study 1		Bleeding	No bleeding	Total
Beta-blockers	(Treatment 1)	2	41	43
Control	(Treatment 3)	13	28	41
Study 2		Bleeding	No bleeding	Total
Sclerotherapy	(Treatment 2)	9	33	42
Control	(Treatment 3)	13	28	41
Study 3		Bleeding	No bleeding	Total
Beta-blockers	(Treatment 1)	2	41	43
Sclerotherapy	(Treatment 2)	9	33	42

network meta-analysis including multi-arm studies. For the cirrhosis dataset, a network meta-analysis including multi-arm studies by only considering comparisons with a reference treatment would not notice that the Peto odds ratios of study 1 are inconsistent.¹⁰ A network meta-analysis of this dataset with R package netmeta results in an error stating that treatment estimates of study 1 are inconsistent (see supplementary R code S1).

Another issue with the Peto odds ratio is that the inconsistency problem also affects network meta-analyses including only studies with two treatments. In Table 3 we present the data of study 1 from the network meta-analysis to prevent bleeding in cirrhosis¹⁰ as if they come from three independent studies. A fixed-effect network meta-analysis of these three studies using the odds ratio as effect measure results in the same odds ratios as for the data in Table 1 and the Q statistic for inconsistency is equal to 0. A fixed effect network meta-analysis using the Peto odds ratio as effect measure results in slightly different network estimates (0.2467, 0.1528, 0.6196) and the Q statistic for inconsistency is equal to 0.0297. The network estimates must be different from the original Peto odds ratio estimators as the underlying network model assumes consistency of treatment effects.

Finally, the inconsistency problem also transfers to indirect comparisons of pairwise meta-analyses. To illustrate this, we only consider data from study 1 and 2 in Table 3. In this situation, the indirect treatment estimate of beta-blockers versus sclerotherapy is calculated in the same way as described in subsection 2.3.1 resulting in an inconsistent indirect Peto odds ratio estimator.

4 | DISCUSSION

In this article we show that the use of Peto odds ratio estimators can lead to inconsistent treatment estimates in settings comparing more than two treatments. While we first noticed this problem for a single multi-arm study and a network meta-analysis including multi-arm

studies, we later recognized that the same problem exists in network meta-analysis of two-arm studies. One reviewer pointed out that this problem also translates to indirect comparisons of pairwise meta-analyses. To our knowledge, this is the first time that this issue of the Peto odds ratio estimator is described in the literature. We assume that it has not been recognized so far as the usual approach to include multi-arm studies in network meta-analysis is to only consider comparisons with a common comparator. Accordingly, the inconsistency is not noticed but ascribed to differences in the individual studies.

We did not comprehensively evaluate the inconsistency of Peto odds ratio estimators, but considered some real and fictitious examples. The aim of this work is to inform about this problem instead of thoroughly investigating it. The Peto odds ratio estimators in a multi-arm study are only very rarely consistent in a strict mathematical sense, for example, if the event numbers and group sample sizes in two of three studies are the same; see variant 1 where odds ratio and Peto odds ratio estimators were rather different, however, both consistent. We would speculate that Peto odds ratio estimands will in general be inconsistent if event probabilities differ across treatments in a study or group sample sizes are unequal.

One may argue—like in pairwise meta-analysis—that the Peto method could be an option if events are rare, treatment effects are small or modest, and group sample sizes are (almost) balanced. However, the problem is whether these conditions are met in a network meta-analysis. As other statistical methods for network meta-analysis of rare binary outcomes are available,^{13,14} we would not recommend the use of the Peto method in network meta-analysis.

Admittedly, our results do not have a real-world impact on applications of network meta-analyses as the Peto odds ratio is rarely used. However, (informal) indirect comparisons of multiple pairwise meta-analyses are quite common. For example, Cochrane reviews often report results of several pairwise meta-analyses comparing active

treatments with placebo as head-to-head comparisons of active treatments are scarce. Furthermore, the Peto odds ratio is a popular effect measure in Cochrane reviews. An empirical investigation of the impact of using Peto odds ratio estimators in such indirect comparisons could provide additional insights on the inconsistency problem.

ACKNOWLEDGMENTS

We would like to thank three reviewers for their valuable comments on the manuscript. One reviewer made us aware of the problem in indirect comparisons of pairwise meta-analyses. GR acknowledges the support by the Deutsche Forschungsgemeinschaft (DFG), grant number RU1747/1-2. OE was supported by Ambizione grant No. 180083 from the Swiss National Science Foundation (SNSF).

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

AUTHOR CONTRIBUTIONS

Guido Schwarzer noticed the inconsistency of Peto odds ratios and discussed it with the co-authors, conducted the analyses and wrote the initial draft of the manuscript. Orestis Efthimiou and Gerta Rücker revised the manuscript. All authors have read and approved the final version of the manuscript.


SUPPORTING INFORMATION

The following supporting information is available as part of the online article: R code S1. This file contains R code for producing all examples in this article.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Guido Schwarzer  <https://orcid.org/0000-0001-6214-9087>

Orestis Efthimiou  <https://orcid.org/0000-0002-0955-7572>

Gerta Rücker  <https://orcid.org/0000-0002-2192-2560>

REFERENCES

1. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis*. 1985;27:335-371.
2. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods*. 2010;1:97-111.
3. Greenland S, Salvan A. Bias in the one-step method for pooling study results. *Stat Med*. 1990;9:247-252.
4. Bradburn MJ, Deeks JJ, Berlin JA, Localio AR. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med*. 2007;26:53-77.
5. Brockhaus AC, Bender R, Skipka G. The Peto odds ratio viewed as a new effect measure. *Stat Med*. 2014;33:4861-4874.
6. Brockhaus AC, Grouven U, Bender R. Performance of the Peto odds ratio compared to the usual odds ratio estimator in the case of rare events. *Biom J*. 2016;58:1428-1444.
7. Sharma T, Götzsche PC, Kuss O. The Yusuf-Peto method was not a robust method for meta-analyses of rare events data from antidepressant trials. *J Clin Epidemiol*. 2017; 91:129-136.
8. Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd ed. Chichester, UK: John Wiley & Sons; 2019.
9. Agresti A. *Categorical Data Analysis*. 2nd ed. New York, NY: John Wiley & Sons; 2002.
10. Higgins JPT, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med*. 1996;15:2733-2749.
11. Rücker G, Krahn U, König J, Efthimiou O, Schwarzer G. netmeta: network meta-analysis using frequentist methods. CRAN. R package version 1.3-1; 2021.
12. Rücker G, Schwarzer G. Reduce dimension or reduce weights? Comparing two approaches to multi-arm studies in network meta-analysis. *Stat Med*. 2014;33:4353-4369.
13. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making*. 2013;33:607-617.
14. Efthimiou O, Rücker G, Schwarzer G, Higgins JPT, Egger M, Salanti G. Network meta-analysis of rare events using the mantel-Haenszel method. *Stat Med*. 2019;38:2992-3012.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Schwarzer G, Efthimiou O, Rücker G. Inconsistent results for Peto odds ratios in multi-arm studies, network meta-analysis and indirect comparisons. *Res Syn Meth*. 2021;12(6):849–854. <https://doi.org/10.1002/jrsm.1503>