Causal alternatives to meta-analysis

Clément Berenfeld (Premedical)

Joint work with A. Boughdiri[1], B. Colnet , W. van Amsterdam[2], A. Bellet[1], R. Kellaf[1], E. Scornet[3] and J. Josse[1]

[1]Inria-Inserm Premedical [2]UMC Utrecht [3]U. Sorbonne

CIRC Seminar - 13.10.2025

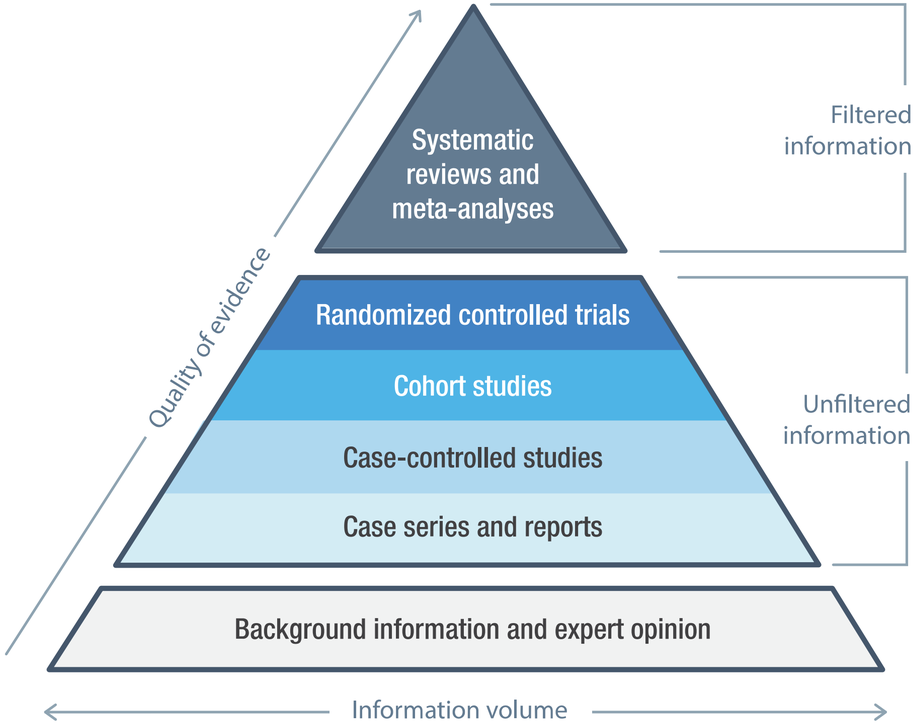
Introduction

What is a meta-analysis?

A **meta-analysis** is a statistical method that combines the results from multiple studies to derive a more precise and reliable overall estimate

→ stands at the **top of the pyramid of evidence** in clinical research

→ used by **regulatory bodies** to formulate recommendations or fix drug prices



source: https://openmd.com/guide/levels-of-evidence

Introduction

What is a meta-analysis?

Let us consider K different **RCTs** measuring the effect of the same binary treatment $A \in \{0,1\}$ and the same binary outcome $Y \in \{0,1\}$.

The results are publicly available in a contingency table of the form

	$Y = 1$	$Y = 0$
$A = 1$	$n_{11}(k)$	$n_{10}(k)$
$A = 0$	$n_{01}(k)$	$n_{00}(k)$

Introduction

What is a meta-analysis?

From these tables, one can compute a desired **treatment effect** $\hat{\theta}_k$ (e.g. risk difference, risk ratio, odds ratio, etc), along with a **standard error** $\hat{\sigma}_k$.

Ex: for the **log-risk ratio**

$$\hat{\theta}_k = \log \frac{n_{11}(k)/n_1(k)}{n_{01}(k)/n_0(k)} \qquad \hat{\sigma}_k^2 = \frac{n_{10}(k)}{n_{11}(k)n_1(k)} + \frac{n_{00}(k)}{n_{01}(k)n_0(k)}$$

where $\quad n_a(k) := n_{a0}(k) + n_{a1}(k)$

Introduction

What is a meta-analysis?

Fixed-effect model: Gaussian model $\hat{\theta}_k \sim \mathcal{N}(\theta^*, \sigma_k^2)$

→ **no heterogeneity** between studies

→ **maximum likelihood** estimator is given by

$$\hat{\theta} = \sum_{k=1}^{K} \omega_k \hat{\theta}_k \qquad \text{where} \qquad \omega_k \propto \frac{1}{\hat{\sigma}_k^2}$$

→ **final variance** estimator is given by $\hat{\sigma}^2 = \left(\sum_{k=1}^{K} \frac{1}{\hat{\sigma}_k^2} \right)^{-1}$

Introduction

What is a meta-analysis?

Random-effects model: Hierarchical model

$$\hat{\theta}_k \,|\, \theta_k \sim \mathcal{N}(\theta_k, \sigma_k^2)$$
$$\theta_k \sim \mathcal{N}(\theta^*, \tau^2)$$

→ **heterogeneity** between studies

→ many **different methods** to estimate τ (e.g, DerSimonian and Laird, Paule-Mandel, etc)

→ **final estimator** is given by

$$\hat{\theta} = \sum_{k=1}^{K} \omega_k \hat{\theta}_k \quad \text{where} \quad \omega_k \propto \frac{1}{\hat{\sigma}_k^2 + \hat{\tau}^2} \quad \text{and} \quad \hat{\sigma}^2 = \left(\sum_{k=1}^{K} \frac{1}{\hat{\sigma}_k^2 + \hat{\tau}^2} \right)^{-1}$$

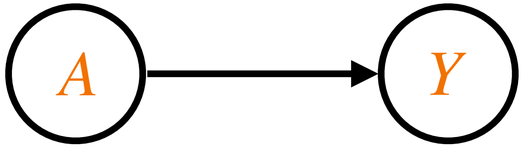
Introduction

What is causal inference?

Causal inference pertains to the process of understanding the relationships between a cause and its effects.

Ex: What is the effect of a given **treatment** on a given **outcome**?

A Y



Counterfactual variables: $Y(0)$ and $Y(1)$ are the outcome if the patient has, **possibly contrary to the fact**, taken treatment $A = 0$ or $A = 1$.

Introduction

What is causal inference?

A causal effect is a measure of how $Y(1)$ and $Y(0)$ differ in a **given population of interest**

Ex: the **risk difference** among

→ the study population (ATE) $\theta = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$

→ the treated population (ATT) $\theta = \mathbb{E}[Y(1) \mid A = 1] - \mathbb{E}[Y(0) \mid A = 1]$

→ the control population (ATC) $\theta = \mathbb{E}[Y(1) \mid A = 0] - \mathbb{E}[Y(0) \mid A = 0]$

In a **RCT**, thanks to randomization, **all these quantities coincide**.

Introduction

The big question

Do usual meta-analysis methods target an estimand which is a causal effect, in the sense that it pertains to the effect of the treatment in a specific population?

→ would have **big implications** in term of **interpretability of meta-analysis results!**

1. Causal meta-analysis with aggregated data (AD)

2. Causal meta-analysis with individual data (ID)

B. C., Boughdiri, A., Colnet, B., van Amsterdam, W. A., Bellet, A., Khellaf, R., Scornet, E., & Josse, J. (2025). Causal meta-analysis: rethinking the foundations of evidence-based medicine. *arXiv preprint arXiv:2505.20168*.

Boughdiri, A., B. C., Josse, J., & Scornet, E. (2025). A unified framework for the transportability of population-level causal measures. *NeuRIPS 2025.*

1. Causal meta-analysis with AD

1. Causal meta-analysis with AD

Notations: A patient's data is typically of the form (A, Y, X, H) where

- A is the **treatment variable**
- $Y = AY(1) + (1 - A)Y(0)$ is the **individual outcome**
- $X \in \mathcal{X}$ is the **patient covariate**
- $H \in [K]$ is the **study membership**

1. Causal meta-analysis with AD

In the **aggregated data** setting, we have access to the **aggregated values**

$$n_{ay}(k) := \#\{i \mid H_i = k, A_i = a, Y_i = y\}, \quad a \in \{0,1\}, y \in \{0,1\}$$

for all studies $k \in [K]$.

We can also have access to **summary information** about the covariate distribution in each study:

$$S_k := \mathcal{S}_k(\hat{P}_k)$$

where $\hat{P}_k$ is the **empirical distribution** of X in study K, and $\mathcal{S}_k : \mathscr{P}(\mathscr{X}) \to \mathbb{R}^{d_k}$ is a **summary map** (e.g., means, standard deviations, quantiles, etc)

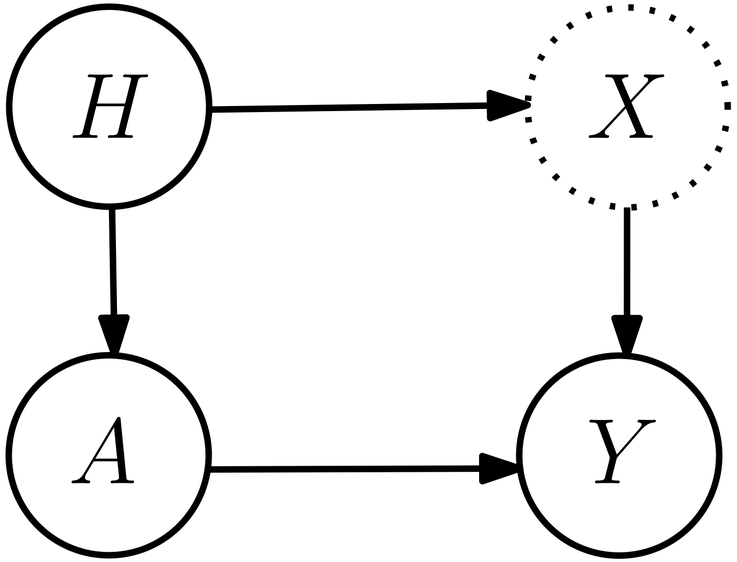
1. Causal meta-analysis with AD

We assume that

- each study is a **RCT** (no arrow from X to A)
- there is **no center effect** (no arrow from H to Y)

$$\forall k, \ell \in [K],$$

$$\underbrace{\mathbb{E}[Y(a) \,|\, X, H = k] = \mathbb{E}[Y(a) \,|\, X, H = \ell]}_{=:\, \mu_a(X)}$$



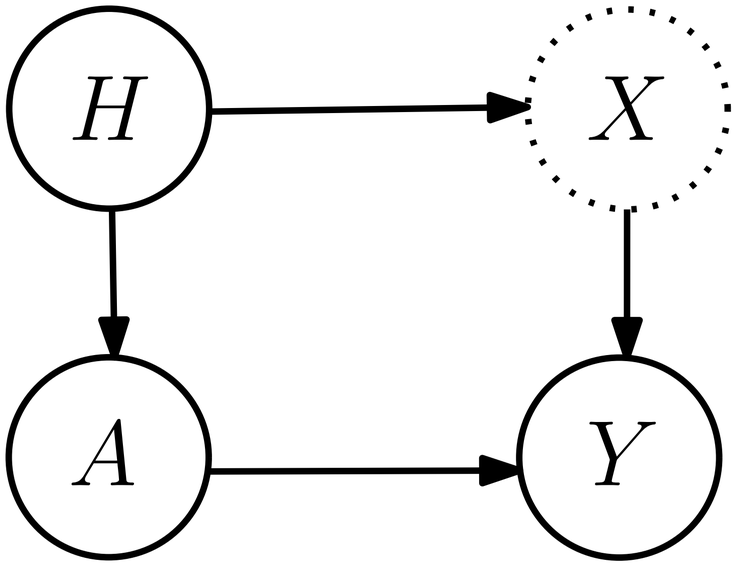
1. Causal meta-analysis with AD

We also assume that θ_k only depends on $\mathbb{E}[Y(a) \mid H = k]$ through

$$\theta_k = \phi(\mathbb{E}[Y(1) \mid H = k], \mathbb{E}[Y(0) \mid H = k])$$

Ex:

- **risk difference:** $\phi(a, b) = a - b$
- **risk ratio:** $\phi(a, b) = a/b$
- **odds ratio:** $\phi(a, b) = \dfrac{a}{1-a}\dfrac{1-b}{b}$
- etc



1. Causal meta-analysis with AD

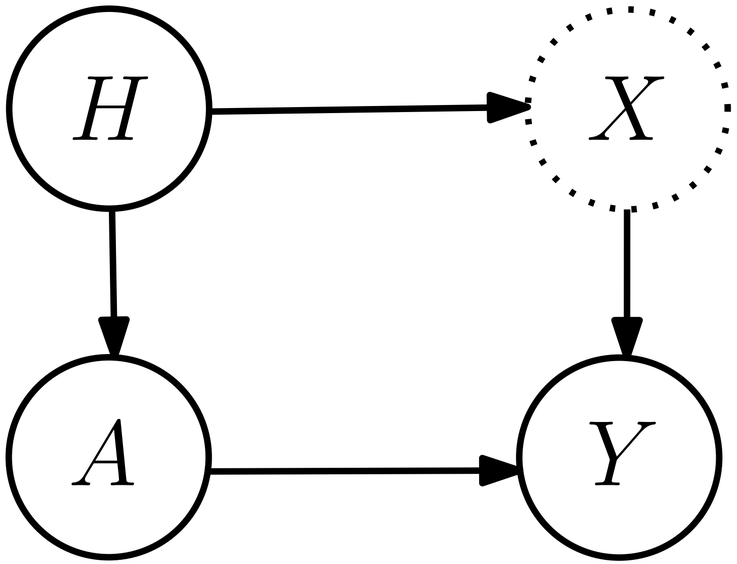
Under these assumptions, θ_k only depends on $P_k = \mathscr{L}(X \mid H = k)$ through

$$\mathbb{E}[Y(a) \mid H = k] = \mathbb{E}[\mathbb{E}[Y(a) \mid X, H = k] \mid H = k] = \int \mu_a(x) \mathrm{d}P_k(x)$$

Defining

$$\theta(P) := \phi\left(\int \mu_1(x) \mathrm{d}P(x), \int \mu_0(x) \mathrm{d}P(x) \right),$$

we find that $\theta_k = \theta(P_k)$



1. Causal meta-analysis with AD

Given a meta-analysis aggregate $\hat{\theta}$, we denote by θ_∞ the values towards which it converges as $n \to \infty$ (if it exists)

Ex: For the **random effect model**, $\hat{\theta}_k \to \theta_k$, $\hat{\sigma}_k \to 0$ and $\hat{\tau} \to \tau$ so that, when $\tau \neq 0$, it holds that

$$\theta_\infty = \frac{1}{K} \sum_{k=1}^{K} \theta_k$$

A meta-analysis effect $\hat{\theta}$ is ***causal*** if, for all covariate distributions $P_1, \ldots, P_K$, there exists P^*, independent from μ_1 and μ_0, such that $\theta_\infty = \theta(P^*)$

$\to$ we say that P^* is the target population

1. Causal meta-analysis with AD

Theorem.

1. If the link function ϕ is **non-linear**, then both the random-effects and the fixed-effect estimator are **not causal**

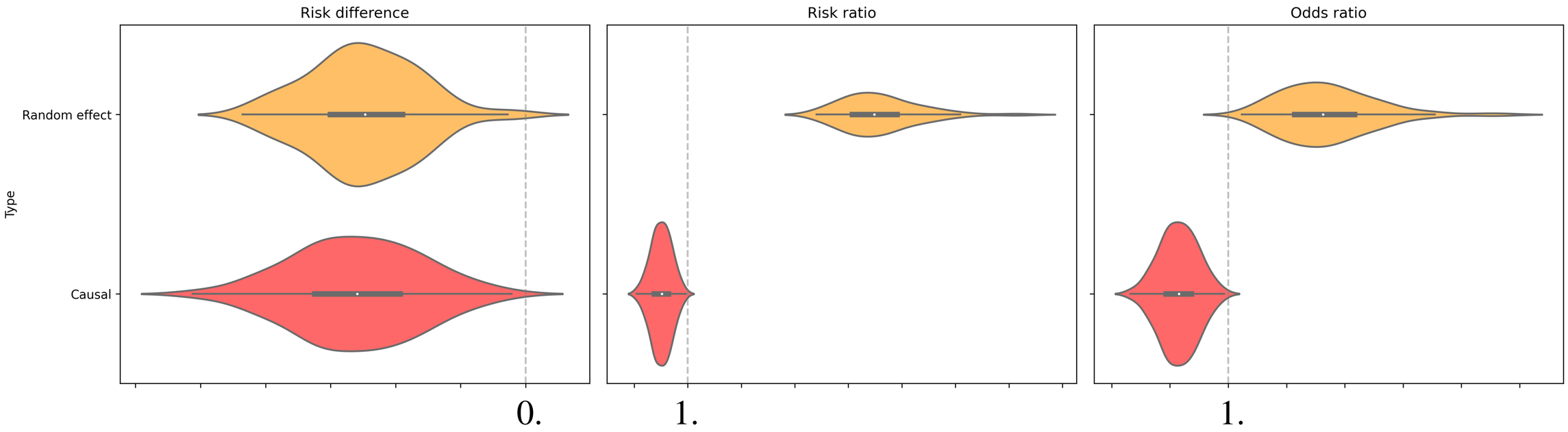
2. If the link function ϕ is **linear** then the random-effects estimator is **causal**

3. If the link function ϕ is **linear**, and if the ratios $\hat{\sigma}_k^2/\hat{\sigma}_{k'}^2$ all converges towards a value in $[0,\infty]$, then the fixed-effects estimator is **causal**

Ex:

- random effects on **risk ratios** → **not causal**
- random effects on **risk differences** → **causal**

1. Causal meta-analysis with AD



A violin plot

1. Causal meta-analysis with AD

So how do we construct causal meta-analysis estimands?

→ First, define a **target population** P^*

→ Try to realize P^* as a **convex combination** of $P_1, \ldots, P_K$

$$P^* \approx \sum_{k=1}^{K} \hat{\alpha}_k P_k \quad \text{where} \sum_{k=1}^{K} \hat{\alpha}_k = 1$$

→ Estimate $\theta(P^*)$ with

$$\hat{\theta} = \phi\left(\sum_{k=1}^{K} \hat{\alpha}_k \frac{n_{11}(k)}{n_1(k)}, \sum_{k=1}^{K} \hat{\alpha}_k \frac{n_{01}(k)}{n_0(k)} \right)$$

1. Causal meta-analysis with AD

Depending on the choice of P^*, and on the summary informations on the P_k's, the computation of $\hat{\alpha}_k$ can range to very simple to very complicated

Covariate-free targets:

- **Pooled target:** $P^* = \sum_{k=1}^{K} \mathbb{P}(H = k) P_k$ and $\hat{\alpha}_k = n_k / n$

  $\rightarrow$ θ^* corresponds to the ATE on the population of all studies pooled together

- **Uniform target:** $P^* = \sum_{k=1}^{K} \frac{1}{K} P_k$ and $\hat{\alpha}_k = 1/K$

1. Causal meta-analysis with AD

Are these effects really different from classical approaches, e.g. random-effects model? **Yes**

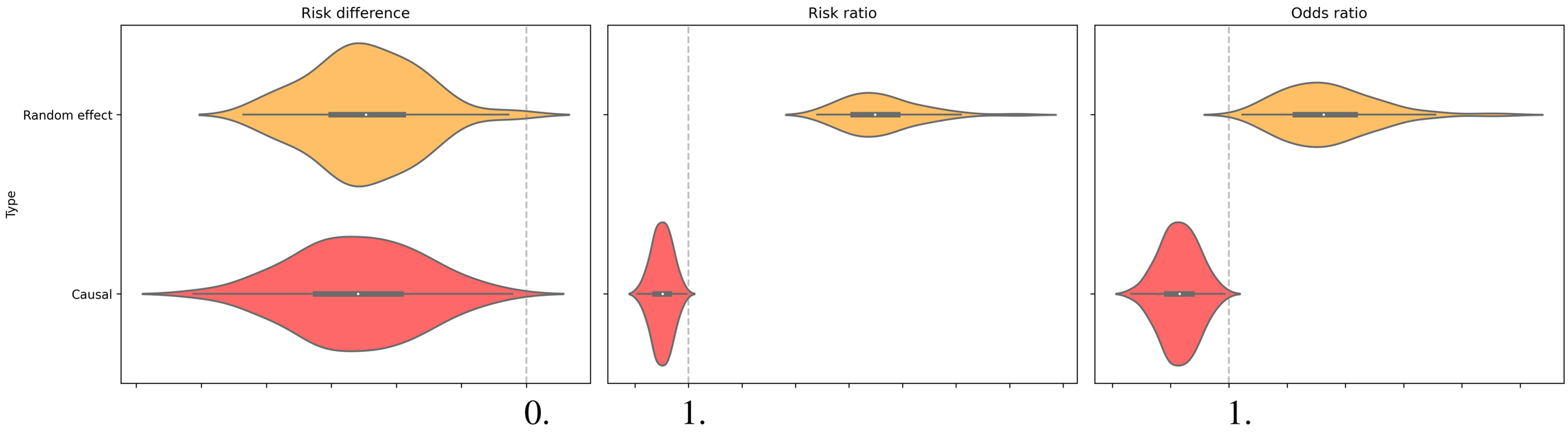
In the large scale limit $n \to \infty$, the random effects estimate converge to, in the case of the **risk ratio**

$$\frac{\prod_{k=1}^{K} P_k(\mu_1)^{1/K}}{\prod_{k=1}^{K} P_k(\mu_0)^{1/K}} \neq \mathrm{RR}(P^*)$$

while a covariate-free causal approach will yield

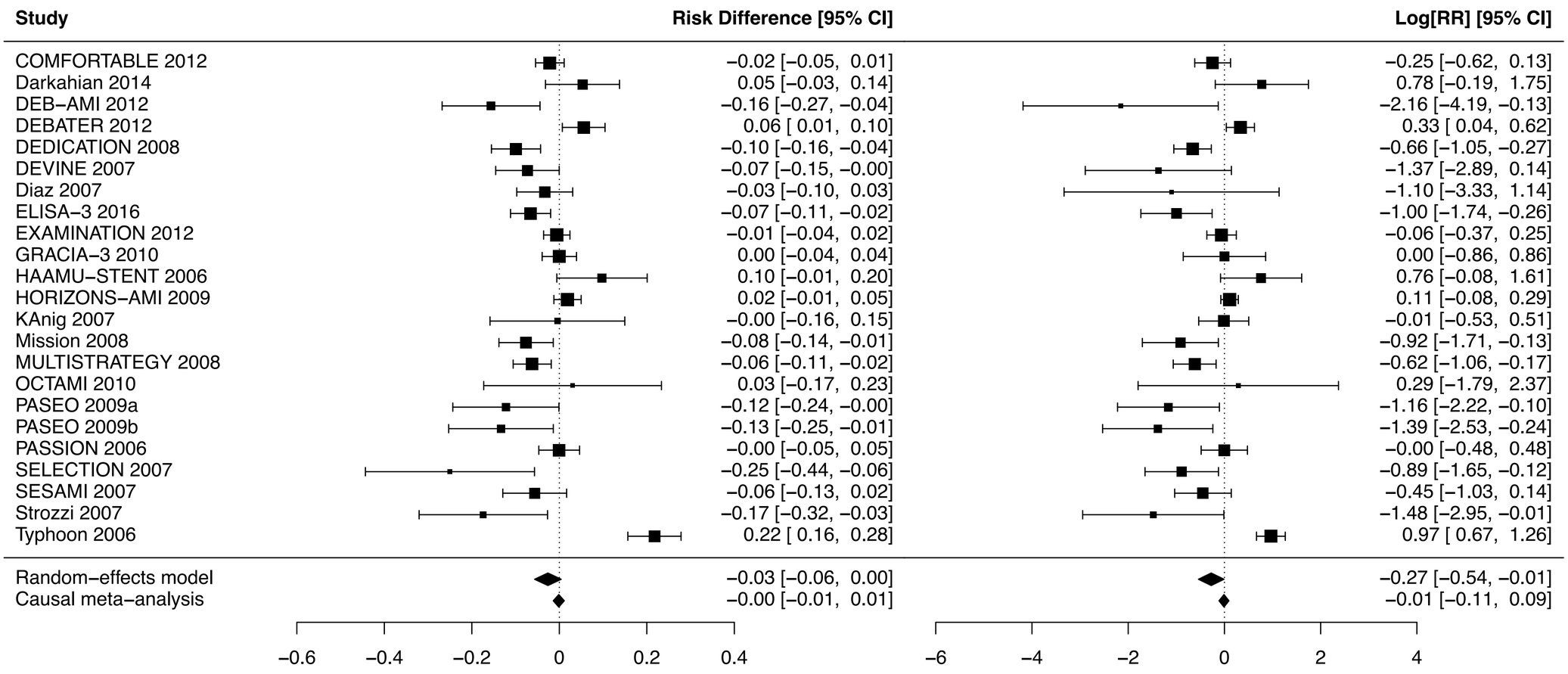
$$\frac{\sum_{k=1}^{K} \alpha_k P_k(\mu_1)}{\sum_{k=1}^{K} \alpha_k P_k(\mu_0)} = \frac{P^*(\mu_1)}{P^*(\mu_0)} = \mathrm{RR}(P^*)$$

1. Causal meta-analysis with AD



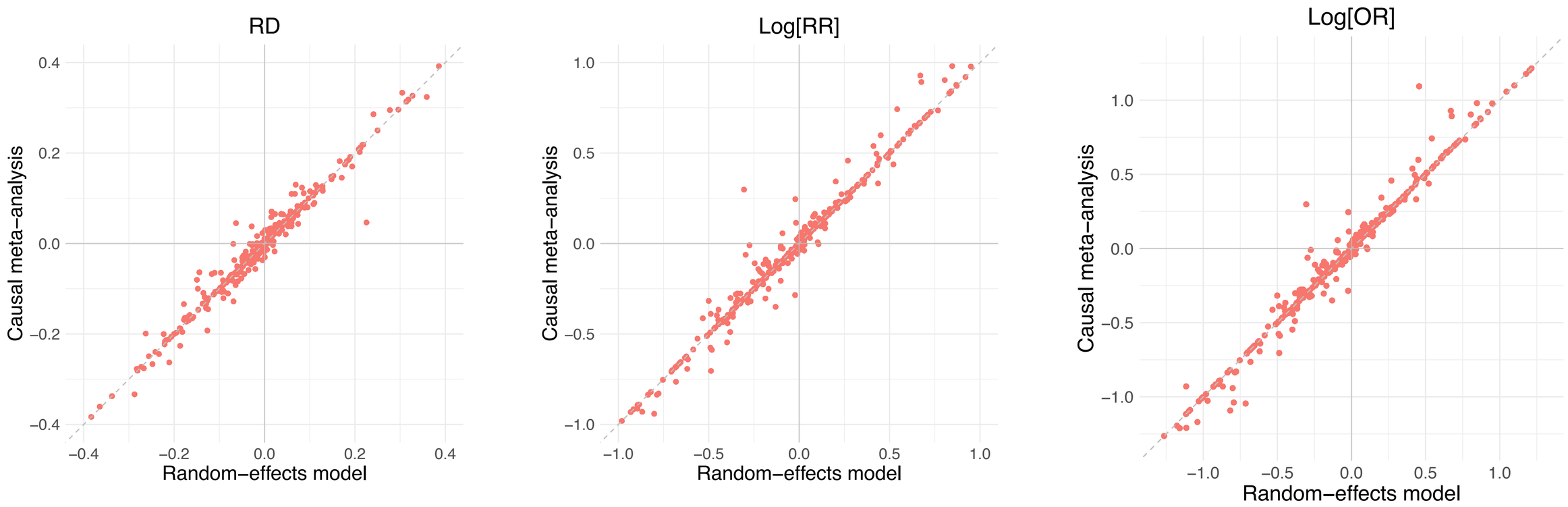
The same violin plot

1. Causal meta-analysis with AD



Reanalysis of Feinberg et al, *Drug-eluting stents versus bare-metal stents for acute coronary syndrome*. Cochrane Database of Systematic Reviews, (8), 2017

1. Causal meta-analysis with AD



Comparison of the random effect models and the pooled meta-analysis for **597 meta-analyses** from the Cochrane Library

1. Causal meta-analysis with AD

Covariate-dependent target: covariate information through $S_k = \mathcal{S}_k(\hat{P}_k)$

- **Linear summaries:** if all $\mathcal{S}_k$ are identical and linear, one can solve

$$\hat{\alpha} \in \operatorname{argmin} \|S^* - S\alpha\| + \Omega(\alpha)$$

where $S = (S_1, \ldots, S_K)$ and $S^* = \mathcal{S}(P^*)$.

- **Parametric proxies:** one can solve $\mathcal{S}_k(Q_k) \approx S_k$ under a parametric model $Q_k \in \Sigma_k$ and adjust

$$\hat{\alpha} \in \operatorname{argmin} \operatorname{dist}\left(P^*, \sum_{k=1}^{K} \alpha_k Q_k\right) + \Omega(\alpha)$$

2. Causal meta-analysis with ID

(a.k.a. *generalization*)

2. Causal meta-analysis with ID

Imagine now that we have access to all the individual data

$$(X_i, A_i, Y_i, H_i) \text{ for } i \in [n]$$

Given a target population P^*, we wish to estimate $\theta(P^*)$

$\rightarrow$ having access to all the covariates $\{X_i\}_{i \in [n]}$ allows to adjust much more precisely to the covariate distribution

2. Causal meta-analysis with ID

A slight change of setting:

- We consider a **single source study** (rather than K) for which $H = 1$
- We have **a sample from the target population**, denoted by $H = 0$
- Data is collapsed to the form

$$(X_i, H_i A_i, H_i Y_i, H_i), i \in [N]$$

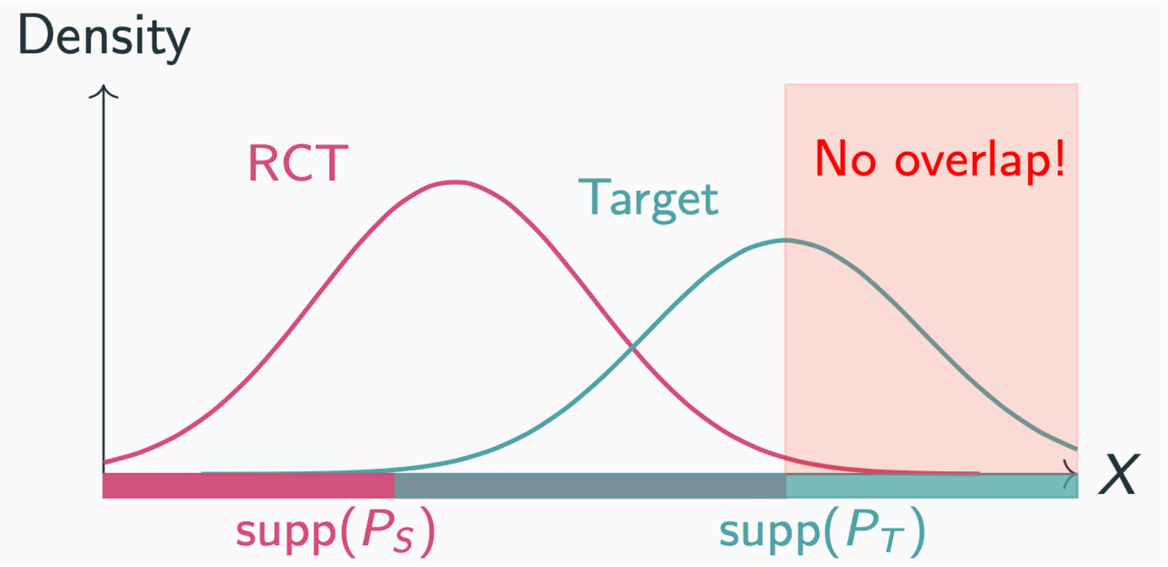
- We let $n = \#\{i, H_i = 1\}$ and $m = \#\{i, H_i = 0\}$.

2. Causal meta-analysis with ID

We let T (resp. S) denotes distribution conditional to $H = 0$ (resp. $H = 1$)

Overlap assumption:

$$\mathrm{supp} P_{\mathrm{T}} \subset \mathrm{supp} P_{\mathrm{S}}$$



Exchangeability in mean:

$$\mathbb{E}_{\mathrm{T}}[Y(a) \,|\, X] = \mathbb{E}_{\mathrm{S}}[Y(a) \,|\, X] \qquad (= \mu_a(X))$$

2. Causal meta-analysis with ID

Goal: estimate the effect in the target population $\theta_{\mathrm{T}} = \theta(P_{\mathrm{T}})$

Identifiability formulae: letting $r(X) := \dfrac{\mathrm{d}P_{\mathrm{T}}}{\mathrm{d}P_{\mathrm{S}}}(X)$, it holds

$$
\begin{aligned}
\mathbb{E}_{\mathrm{T}}[Y(a)] &= \mathbb{E}_{\mathrm{T}}[\mathbb{E}_{\mathrm{T}}[Y(a) \,|\, X]] \\
&= \mathbb{E}_{\mathrm{T}}[\mu_a(X)] && \textbf{(1)} \\
&= \mathbb{E}_{\mathrm{S}}[r(X)\mu_a(X)] \\
&= \mathbb{E}_{\mathrm{S}}[r(X)Y(a)] \\
&= \mathbb{E}_{\mathrm{S}}[r(X)Y \,|\, A = a] && \textbf{(2)}
\end{aligned}
$$

2. Causal meta-analysis with ID

G-formula:

$$\mathbb{E}_{\mathrm{T}}[Y(a)] = \mathbb{E}_{\mathrm{T}}[\mu_a(X)] \quad \textbf{(1)}$$

1. **Fit two (or one) models** on the **source data**

$$Y_{A=a} = \mu_a(X) + \varepsilon$$

2. Predict the counterfactual outcomes in the target population

$$\hat{Y}_i(a) = \hat{\mu}_a(X_i)$$

3. Average over the **target population**

$$\hat{\theta} = \phi\left(\frac{1}{m}\sum_{H_i=0}\hat{\mu}_1(X_i), \frac{1}{m}\sum_{H_i=0}\hat{\mu}_0(X_i)\right)$$

2. Causal meta-analysis with ID

Reweighted Neyman:

$$\mathbb{E}_{\mathrm{T}}[Y(a)] = \mathbb{E}_{\mathrm{S}}[r(X)Y \mid A = a] \quad \textbf{(2)}$$

1. Notice that

$$r(X) = \frac{\mathbb{P}(H = 1 \mid X)\mathbb{P}(H = 1)}{\mathbb{P}(H = 0 \mid X)\mathbb{P}(H = 0)} = \frac{\alpha\rho(X)}{1 - \rho(X)} \quad \text{where} \quad \begin{cases} \rho(X) = \mathbb{P}(H = 1 \mid X) \\ \alpha = \mathbb{P}(H = 1)/\mathbb{P}(H = 0) \end{cases}$$

2. **Fit a model** for ρ (using e.g. a logistic regression) on **the whole data**

$$H = \rho(X) + \varepsilon$$

3. Average over the **source population**

$$\hat{\theta} = \phi\left(\frac{1}{n_1}\sum_{H_i=1} A_i\hat{r}(X_i)Y_i, \frac{1}{n_0}\sum_{H_i=1}(1 - A_i)\hat{r}(X_i)Y_i\right) \qquad n_a := \#\{i, A_i = a, H_i = 1\}$$

2. Causal meta-analysis with ID

Both approach require to fit a model (r or μ_a)

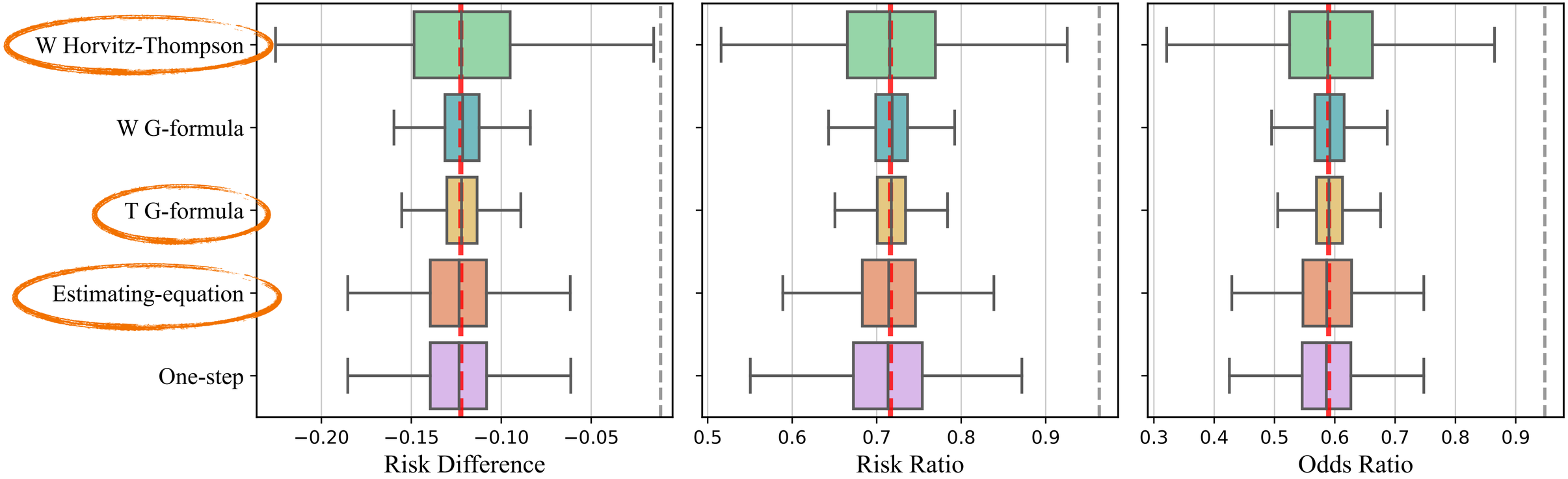
→ one can combine both for a **double-robust estimator**

$$\underbrace{\frac{1}{m}\sum_{H_i=0}\hat{\mu}_a(X_i)}_{\text{G-formula}} + \underbrace{\frac{1}{n_a}\sum_{H_i=1}\mathbb{I}\{A_i = a\}\hat{r}(X_i)(Y_i - \hat{\mu}_a(X_i))}_{\text{corrective term}}$$

→ consistent estimation of θ^* as soon as $\hat{\mu}_a$ **or** $\hat{r}$ **is well specified**

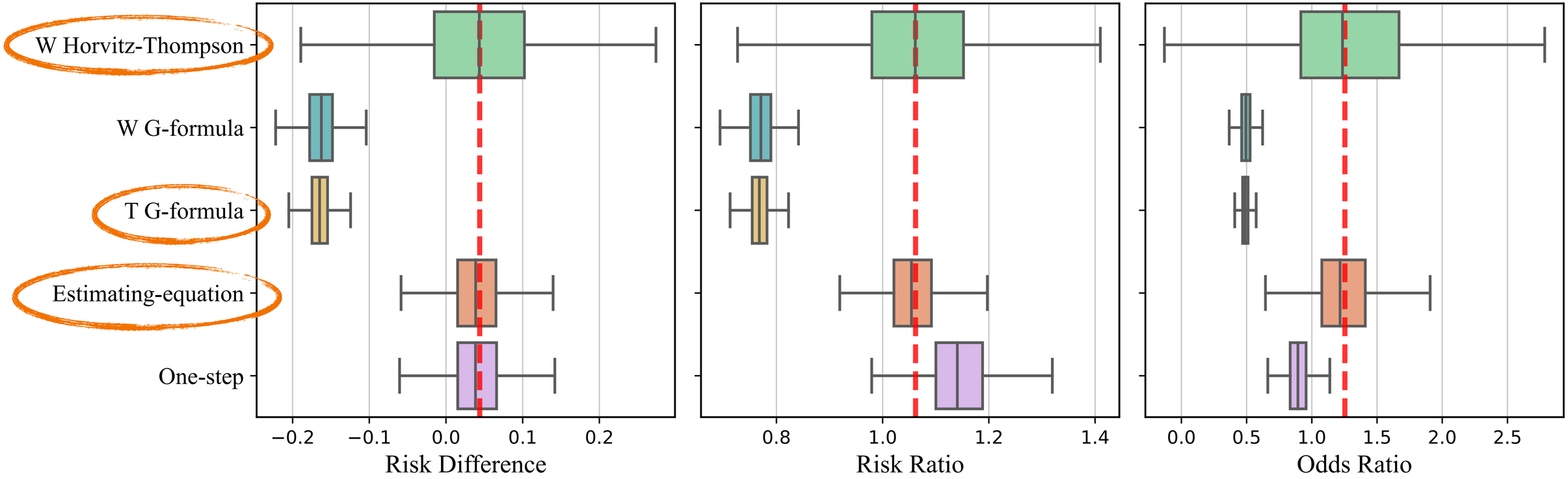
2. Causal meta-analysis with ID

Simulation study: under **well-specification**



2. Causal meta-analysis with ID

Simulation study: under **mis-specification of the treatment response**



2. Causal meta-analysis with ID

Case study: CRASH-3 results generalization to the **Traumabase** population

CRASH-3 trial

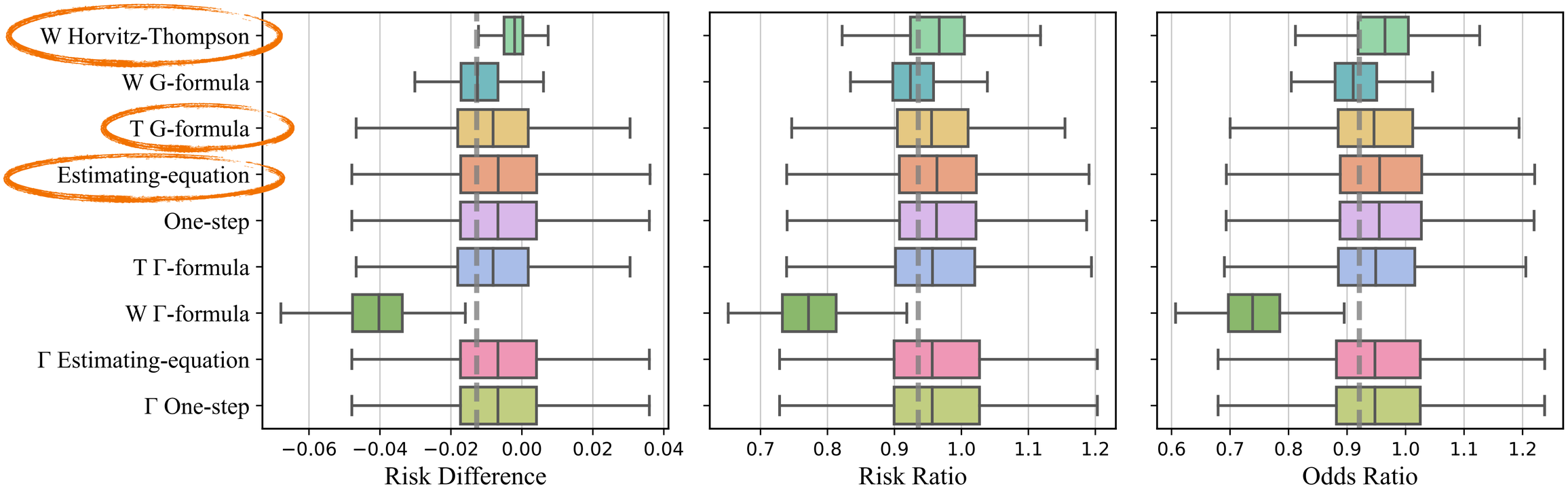
- Randomized trial ($n \approx 12{,}000$)
- Patients with TBI, GCS ≤ 12, within 3h
- Treatment: Tranexamic Acid (TXA)
- Outcome: Head injury-related death at 28 days

Traumabase cohort

- Observational registry ($m \approx 9{,}000$)
- Severe trauma, real-world population
- Selected CRASH-3-eligible patients
- Aim: Apply TXA treatment effect to this cohort

2. Causal meta-analysis with ID

Case study: CRASH-3 results generalization to **Traumabase** population



Grey line corresponds to the effect estimated from the RCT

Conclusion

- Meta-analysis methods depends on the kind of data available (**AD** vs **ID**)

- Traditional methods with **AD** (e.g. **random effects**) do **not** target an estimand which can be interpreted as an **average treatment effect**

- One can alternatively aggregate the data using **weighting strategies** at **the population level** to specifically target a treatment effect

- Under the **ID** setting, one can reweight at **the individual level** or resort to **double robust approaches**

What’s next?

- **Sensitivity analysis** wrt no-study effect assumption
- Implementing and testing **covariate-based weighting strategies for AD**

We created a very simple R package you can play with :)

CaMeA

https://cran.r-project.org/package=camea

CaMeA: Causal Meta-Analysis for Aggregated Data

A tool for causal meta-analysis. This package implements the aggregation formulas and inference methods proposed in Berenfeld et al. (2025) <doi:10.48550/arXiv.2505.20168>. Users can input aggregated data across multiple studies and compute causally meaningful aggregated effects of their choice (risk difference, risk ratio, odds ratio, etc) under user-specified population weighting. The built-in function camea() allows to obtain precise variance estimates for these effects and to compare the latter to a classical meta-analysis aggregate, the random effect model, as implemented in the 'metafor' package <https://CRAN.R-project.org/package=metafor>.

Thank you for your attention!

B. C., Boughdiri, A., Colnet, B., van Amsterdam, W. A., Bellet, A., Khellaf, R., Scornet, E., & Josse, J. (2025). Causal Meta-Analysis: Rethinking the Foundations of Evidence-Based Medicine. *arXiv preprint arXiv:2505.20168*.

Boughdiri, A., B. C., Josse, J., & Scornet, E. (2025). A unified framework for the transportability of population-level causal measures. *NeuRIPS 2025.*