

## Background

The presence of publication bias is often verified in meta-analyses by visual inspection of the funnel plot. Statistical tests may estimate the association between the reported effect size and their standard error (Egger's test, **ER**), total sample size (Macaskill's test, **MR**) or inverse of the total sample size (Peters' test, **PR**). Although these tests have been evaluated for pooling odds ratios, their application may be less appropriate for survival data where censoring influences statistical significance (and thus selective reporting) of the hazard ratio.

## Results for scenario (a.i)

| HR   | m <sup>pb</sup> | $\tau_\lambda = 0.00$ |      |      |      | $\tau_\lambda = 0.01$ |      |      |      | $\tau_\lambda = 0.02$ |      |      |      |
|------|-----------------|-----------------------|------|------|------|-----------------------|------|------|------|-----------------------|------|------|------|
|      |                 | 10                    | 20   | 50   | 100  | 10                    | 20   | 50   | 100  | 10                    | 20   | 50   | 100  |
| 0.50 | ER              | 0.10                  | 0.11 | 0.11 | 0.13 | 0.10                  | 0.11 | 0.11 | 0.13 | 0.10                  | 0.11 | 0.11 | 0.13 |
|      | MR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | PR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T1              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T2              | 0.10                  | 0.11 | 0.12 | 0.14 | 0.10                  | 0.11 | 0.11 | 0.13 | 0.10                  | 0.10 | 0.10 | 0.11 |
| 0.75 | ER              | 0.10                  | 0.10 | 0.10 | 0.11 | 0.10                  | 0.10 | 0.10 | 0.11 | 0.10                  | 0.10 | 0.10 | 0.11 |
|      | MR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | PR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T1              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T2              | 0.10                  | 0.10 | 0.10 | 0.11 | 0.10                  | 0.10 | 0.10 | 0.11 | 0.10                  | 0.10 | 0.10 | 0.10 |
| 1.00 | ER              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | MR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | PR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T1              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T2              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
| 5.00 | ER              | 0.11                  | 0.12 | 0.15 | 0.19 | 0.11                  | 0.12 | 0.15 | 0.19 | 0.11                  | 0.12 | 0.15 | 0.19 |
|      | MR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | PR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T1              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T2              | 0.11                  | 0.12 | 0.14 | 0.20 | 0.10                  | 0.11 | 0.13 | 0.16 | 0.10                  | 0.10 | 0.11 | 0.13 |

Type-I error (false positive) rates in the absence of publication bias. Results are based on 100 000 simulated meta-analyses with  $\lambda = 0.03$  and no censoring of events.

## Details of the simulation study

We generated a survival time for each subject in each trial according to a Weibull distribution with shape parameter 0.65 and scale parameter 0.03. We simulated between-study heterogeneity in baseline survival by randomly sampling the scale parameter from  $[0.03 - \tau_\lambda, 0.03, 0.03 + \tau_\lambda]$ , with  $\tau_\lambda$  equal to 0, 0.01 or 0.02. We subsequently introduced non-informative right-censoring using a common censoring time point for all subjects from the same trial. This time point was defined a priori to yield a specific proportion of censored observations ( $\pi^{\text{cens}} = 0$  or  $\pi^{\text{cens}} = 0.30$ ). For each scenario, we generated 100 000 meta-analyses of 10, 20, 50 and 100 studies. Cox regression was applied in each study to estimate the log hazard ratio and corresponding standard error. We initially assess the type-I error rate of each approach in each scenario; no publication bias was thus introduced in these simulation studies. Results from publication bias tests were considered as indicative of publication bias presence when yielding a two-sided p-value below 0.10.

## Contact



Thomas Debray, PhD  
Julius Center  
University Medical Center Utrecht  
Tel: +31 (0) 88 75 68026  
Email: t.debray@umcutrecht.nl

## Methods

We propose and evaluate two new publication bias tests for survival data that are based on the total number of events (**T1**) and the total survival time (**T2**). We compare their performance to existing tests in an extensive simulation study where more than 20,000,000 meta-analyses were generated. Here, we varied the true hazard ratio (HR=0.50-5.00), the number of available studies (N=10-100), the censoring proportion (cp=0.00-0.30) and the degree of between-study heterogeneity in the scale of the hazard distribution ( $\lambda = 0.01-0.05$ ). We used a set of predefined parameters reflecting meta-analyses of randomized clinical trials in the medical literature.

## Results for scenario (a.ii)

| HR   | m <sup>pb</sup> | $\tau_\lambda = 0.00$ |      |      |      | $\tau_\lambda = 0.01$ |      |      |      | $\tau_\lambda = 0.02$ |      |      |      |
|------|-----------------|-----------------------|------|------|------|-----------------------|------|------|------|-----------------------|------|------|------|
|      |                 | 10                    | 20   | 50   | 100  | 10                    | 20   | 50   | 100  | 10                    | 20   | 50   | 100  |
| 0.50 | ER              | 0.10                  | 0.10 | 0.11 | 0.11 | 0.10                  | 0.10 | 0.10 | 0.11 | 0.10                  | 0.10 | 0.11 | 0.11 |
|      | MR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | PR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T1              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T2              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
| 0.75 | ER              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | MR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | PR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T1              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T2              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
| 1.00 | ER              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | MR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | PR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T1              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T2              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
| 5.00 | ER              | 0.11                  | 0.11 | 0.13 | 0.16 | 0.10                  | 0.10 | 0.11 | 0.11 | 0.11                  | 0.11 | 0.11 | 0.13 |
|      | MR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | PR              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T1              | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.10 |
|      | T2              | 0.10                  | 0.10 | 0.11 | 0.12 | 0.10                  | 0.10 | 0.10 | 0.10 | 0.10                  | 0.10 | 0.10 | 0.11 |

Type-I error (false positive) rates in the absence of publication bias. Results are based on 100 000 simulated meta-analyses with  $\lambda = 0.03$  and  $\pi^{\text{cens}} = 0.30$ .

## Results

Results demonstrate that Egger's test has problematic type-I error rates, particularly in meta-analyses with many studies (20) or in the presence of between-study heterogeneity ( $\tau_\lambda > 0$ ). T2 also yields problematic type-I error rates in the presence of between-study heterogeneity, but to a lesser extent than ER. PR, MR and T1 yield ideal type-I error rates under all circumstances (note that T1 corresponds to PR when there is no censoring of events). Further research is needed to investigate the power of these tests.