Statistical Primer: Heterogeneity, random- or fixed-effect model analyses?

Short title: Heterogeneity in meta-analysis

Authors: **Fabio Barili [1], M.D. Ph.D., Alessandro Parolari [2], M.D. Ph.D., Pieter A. Kappetein [3], M.D. Ph.D., Nick Freemantle [4]**, **Ph.D.**

Institutions:

[1] Department of Cardiac Surgery, S. Croce Hospital, Cuneo, Italy.

[2] Unit of Cardiac Surgery and Translational Research, IRCCS Policlinico S. Donato, Italy.

[3] Thoraxcenter, Erasmus MC, Rotterdam, Netherlands.

[4] Department of Primary Care and Population Health, University College London, London, UK.

Meeting presentation: this was presented at the European Association for Cardio-Thoracic Surgery Annual Meeting, Vienna, Austria, 7-11 October 2017

Corresponding Author: **Fabio Barili, M.D., PhD,**

Department of Cardiac Surgery, S. Croce Hospital

Via M. Coppino 26, 12100 Cuneo, Italy

Tel: +39 0171642571 Fax: +39 0171642064

Email: fabarili@libero.it barili.f@ospedale.cuneo.it

Visual abstract

- Key question: What is heterogeneity in meta-analysis?
- Key findings: Heterogeneity is the true difference in effect sizes, due to intrinsic factors of the studies included in meta-analysis.
- Take-home message: Heterogeneity can be assessed and quantified through random-effects model.

SUMMARY

Heterogeneity in meta-analysis describes differences in treatment effects between trials which exceed those we may expect through chance alone. Accounting for heterogeneity drives different statistical methods for summarizing data and, if heterogeneity is anticipated, a random-effects model will be preferred to the fixed effects model.

Random-effects models assume that there may be different underlying true effects estimated in each trial which are distributed about an overall mean. The confidence intervals around the mean include both within-study and between-study components of variance (uncertainty). Summary effects provide an estimation of the average treatment effect and the confidence interval depicts the uncertainty around this estimate.

There are five statistics that are computed to identify and quantify heterogeneity. They have different meaning and give complementary information: Q statistic and its p-value simply test if effect sizes depart from homogeneity, T^2 and T quantify the amount of heterogeneity, I^2 expresses the proportion of dispersion due to heterogeneity. The point estimate and confidence intervals for random effects models describe the practical implications of the observed heterogeneity, and may usefully be contrasted with the fixed effects estimates.

Max. 200 words

INTRODUCTION

Meta-analysis is the statistical synthesis of data from related studies and the results summarise a body of research. Unlike the narrative review, meta analysis calculates a weighted average treatment effect and its uncertainty [1]. The central unit of meta-analysis is the treatment effect or effect size, a measure of the relationship between two groups [2]. The effect size can vary across related studies and the principal goal of the synthesis is the estimation of a summary effect, which is simply a weighted mean of the individual effects. It is also critical to evaluate the robustness of the summary effect, including some expectation on variability among studies and subsequently quantifying it. The observed dispersion of the estimated effect sizes is partly spurious as it always includes a random (or sampling) error inherent in each study but it may also include a true variation of the effects sizes in each study, namely heterogeneity.

Heterogeneity is the true difference in effect sizes, related to intrinsic factors of the studies included in meta-analysis [2, 3]. Differences in cohorts' characteristics and in treatments options, together with other reasons, lead to assume that studies will not share a common effect size but will have heterogeneous underlying effects. This assumption on heterogeneity is a critical point when conducting a meta-analysis, as it drives different statistical methods for summarizing data and also different interpretation of results. If our understanding is that all studies share the same common effect, we will choose a fixed effect model; otherwise, if heterogeneity is expected, a random-effect model will be preferred (Figure 1) [3].

Fixed-effects model. The fixed effect model assumes that all studies considered in the meta-analysis share the same common true effect size (hence, the term fixed) (Figure 1A). Differences among observed effects are related to sampling error (ε_i; i stands for study i). and factors influencing the effect size are assumed to be the same in all the studies. There is no heterogeneity ($\zeta_i = 0$) and the variance is completely due to spurious dispersion (within-study variance). The summary effect is the estimate of a common true effect and the confidence intervals depict the uncertainty around this estimate.

Random-effects model. Random-effects models assume that there are different underlying true effects. These true effect sizes are distributed about some mean (Figure 1B) and can be considered as a random sample from a distribution (usually Gaussian) - hence, the term random. Random effects models are preferred when studies cohorts are expected to be different or treatments options are not identical among studies. The variance is accounted by both spurious (within-study variance, ε_i) and real dispersion (between-study variance, ζ_i) and formula are applied to partition it in these two components, as the main focus shifts from the summary effect to the identification and quantification of heterogeneity. Summary effects provide an estimation of the average treatment effect and the confidence interval depicts the uncertainty around this estimate including the component of heterogeneity [2]. In the presence of heterogeneity, the relative weights are more balanced than those assigned under fixed effects as standard random effects methods add a common component of variance to each study weight to account for between study variability in treatment effects. Consequently this double source of variability (within and between study) will lead to wider variance, standard error and confidence interval for the summary effect [2].

For example, we can suppose to conduct a meta-analysis of randomized controlled trials comparing clinical outcomes (30-day mortality and 30-day pace-maker implantation) of adult patients with severe aortic stenosis undergoing either trans-catheter aortic valve implantation (TAVI) or surgical aortic valve replacement (SAVR). Effect sizes can be hypothesized not to be identical across studies, as different risk profiles are included and also different devices were employed. Hence, random effects model would be preferred.

METHODOLOGY

Under the random-effects model, the attention is focused on quantifying heterogeneity and understanding its implications [2]. Specific methodologies are employed to partition the total dispersion, isolate the true variance and give an array of statistics for abstracting interpretation of results (Figure 2).

Q statistic (also know as Cochrane's Q) is the weighted sum of squares; more easily, a measure of the total observed dispersion of the estimated effect sizes. It's a standardized value and it's not affected by the metric of the effect size, hence it's not a measure of dispersion on the same scale of the effect size (not comparable).

Q-df is the part of dispersion related to differences in the true effects (heterogeneity, or excess variation). It is calculated subtracting to Q the degrees of freedom (df), which represent the within-study error. It's also a standardized measure.

Test for Assumption of homogeneity. It is based on Q statistics and tests the null hypothesis that all studies share a common effect size. The test performs badly in the small sample setting and the results are sensitive to the excess of dispersion and the number of studies included, as increase of dispersion moves toward significance and an increased number of studies strengthen the evidence of the test. To be noted, a significant p-value confirms that the true effects vary while a non-significant p-value should be discussed, as it depends not only on robustness of effect sizes but it can account for low power (small number of studies, wide within study variance). Moreover the homogeneity test, as well as Q statistic, cannot be employed as an estimate of amount of heterogeneity and it simply tests the null hypothesis that all effects sizes are consistent.

T^2 and T- Estimates of the variance and standard deviation of the true effect sizes. T^2 is the estimate of the variance of the true effect sizes (τ^2), derived from the observed effects. Differently from Q, it is expressed in the same metric of the summary effect and it represents the amount of true dispersion of the effect sizes. The most common method for estimating the between-studies variance in meta-analysis is the DerSimonian-Laird estimator [4], which is based upon the method of moments and may be biased in some settings. T is the square root of T^2 and represents the estimate of standard deviation of the true effect sizes' normal distribution (τ). It has the same metric of the summary effect. Assuming a normal distribution of the true effect sizes, it can be used to describe the distribution of the effects around their mean, calculating the 95% confidence interval of the summary effect. Increasing T values reflect increased true variance around the mean in the summary estimate.

The I^2 statistic express the proportion of the total dispersion that account for true dispersion, being the ratio between excess of dispersion and total dispersion. It is calculated on Q and hence it is not the estimate of an underlying amount but only a descriptive statistic. It is a measure of inconsistency among the findings of the studies and it's not affected by the number of studies included in the meta-analysis. It was suggested that 25%, 50% and 75% could be considered low, intermediate and high inconsistency [5]*; nonetheless these cut-offs are simply crude guidelines thresholds and the evaluation of I^2 statistic should overcome them.*

In summary, there are five statistics that are computed to identify and quantify heterogeneity. They have different meaning and give complementary information: Q statistic and its p-value simply test if effect sizes are homogeneous, T^2 and T quantify the amount of heterogeneity, I^2 expresses the proportion of dispersion due to heterogeneity. A sixth, and potentially much more useful, statistic describing the effects of heterogeneity is the random effects estimator of the pooled treatment effects.

Common statistical software and languages have functions to estimate heterogeneity. Fixed and random effect meta-analyses can be implemented in the R packages "Meta", "metafor" "rmeta" "epiR". A tutorial for conducting meta-analysis with R with the package " metaphor" is described in [6]. RevMan 5 is the software developed for preparing and maintaining Cochrane Reviews and it is possible to choice random or fixed effects models while conducting meta-analysis. Macros for conducting meta-analysis in SPSS can be found in the web (for example: http://mason.gmu.edu/~dwilsonb/ma.html). In Stata, Meta and Metan commands has been developed to generate fixed and random-effects meta-analysis. The %METAANAL macro is a SAS version 9 macro that produces the DerSimonian-Laird estimators for random effects or fixed effects models.

REPORTING

Meta-analysis should be reported following published guidelines, such as PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and MOOSE (Meta-analysis of Observational Studies in Epidemiology) reporting guidelines [7, 8].

Authors should explicit the rationale for the choice of the model, underscoring potential sources of variability of the studies included in the meta-analysis. In the results and/or in the forest plot, it should be reported the evaluation of heterogeneity including the Q statistics, the test for assumption of homogeneity, the I^2 statistic and the estimate of the variance of the true effect sizes T^2. The random effects estimator and confidence intervals describe the importance of heterogeneity in the practical setting. In the discussion, Authors should make inference not only on summary effect but also on dispersion.

There are some notes to take in mind. First of all, a very small number of studies can lead to a poor estimate of heterogeneity. Hence, random effect model has been correctly chosen but there is insufficient information for applying it. In this case, one possible option could be to avoid reporting a summary effect, as conclusions on effect size and its confidence interval cannot be drawn, or an alternative could be represented by a different approach, such as a Bayesian one, where the extent of heterogeneity maybe inferred through an informative prior. Moreover, the practice of performing a fixed-effect model and subsequently moving to the random-effect one if the test of homogeneity is significant should be discouraged, as the choice should be based on hypothesis on common effect sizes and not on a statistical test that often suffers of low power. Differences in cohorts' characteristics (for example, different preoperative risk profiles) and in treatments options (such as different devices with potential implementation of interventions), together with other reasons (different ethnicity, geographical variation, etc), lead to assume that studies will not share a common effect size and should be analysed with random–effects model. Further, the standard methods for random effects (DerSimonian and Llaird) include a component of variance to describe the between study variability adaptively, diverging from the fixed effects model when the p value for heterogeneity is significant. If the random effect model is chosen and T^2 was demonstrated to be 0, it reduces directly to the fixed effect, while a significant homogeneity test in a fixed effect model leads to reconsider the motivations at its basis. However the contrast of the fixed and random effects results provides a useful description of the importance of heterogeneity in the results. Finally, interpretation of random effects meta-analysis can be implemented by prediction interval, a measure that provides a predicted range for the true treatment effect in an individual study [3]. It resembles reference ranges usually employed in other areas of medicine, such as those for blood pressure or birth-weight across the population [3].

EXAMPLE

We can aim to meta-analyze randomized controlled trials comparing 30-day mortality and 30-day pacemaker implantation of adult patients with severe aortic stenosis undergoing either TAVI or SAVR. We choose to evaluate the risk difference (RD) of outcomes between treatment and control groups. The seven included trials differ for perioperative risk profiles, as [9, 10] are performed in intermediate-risk, while [11-14] has been performed in high-risk patients. Moreover, also treatment options are different because different TAVI devices have been employed across studies [9-14]. These considerations can lead to assume that heterogeneity (between study differences in treatment effects) is anticipated and the random-effects model is preferred.

The analysis of heterogeneity for 30-day mortality demonstrates that trials are homogeneous (Figure 3A), being the test for assumption of homogeneity (see Methodology) p-value=0.50 and the percentage of heterogeneity on total variability (I^2) of 0%, suggesting the variability in study estimates is entirely due to chance. The estimate of the variance of the true effect sizes (T^2) is 0. In this case with no source of heterogeneity and only within-study variance, the random effect model coincides with the fixed effect one, as shown in Figure 4A, and the summary risk difference (-0.009; 95%CI: -0.0191 and 0.0011) is the estimate of a common true effect size. The point estimate thus suggests that average mortality under TAVI is 0.9% lower than under SAVR, but the 95% confidence intervals include a reduction of 1.9% or an increase of 0.1%.

The analysis of heterogeneity for 30-day pacemaker implantation shows significant heterogeneity across studies, with the test for assumption of homogeneity p-value < 0.0001 (Figure 3B) and high inconsistency (I^2 96.16%). The estimate of the variance of the true effect sizes (T^2) is 0.0094. The summary risk difference (0.11; 95%CI 0.03-0.19) is the estimation of the mean of distribution of the effects. As the confidence interval does not contain zero, there is good evidence that on average TAVR is related to increased incidence of 30-day pacemaker implantation. Figure 4B shows the implication of model choice; in random effect, the relative weights are more balanced and the double source of variability led to wider variance, standard error and confidence interval for the summary effect.

CONCLUSIONS

Summarizing, heterogeneity assessment is an important step in meta-analysis, as in many cases the assumption of same true effect across studies is implausible. Thus random effects meta-analysis, which accounts for unexplained heterogeneity, will continue to be prominent in the medical literature [3].

REFERENCES

[1] Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res 1993;2;121-145.

[2] Borenstein M, Hedges LV, Higghins JPT, Rothstein HR. Introduction to Meta-Analysis. John Wiley 2009.

[3] Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011;342:d549.

[4] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986 Sep;7(3):177-88.

[5] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.

[6] Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. J Stat Soft 2010;36 (3):1-48.

[7] Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the
PRISMA statement. Int J Surg. 2010;8:336-41. [PMID: 20171303] doi:10.1016/j.ijsu.2010.02.007

[8] Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. Journal of the American Medical Association 2000, 283: 2008–12.

[9] Leon MB, Smith CR, Mack MJ, Makkar RR, Svensson LG, Kodali SK, et al. Transcatheter or Surgical Aortic-Valve Replacement in Intermediate-Risk Patients. N Engl J Med 2016;374(17):1609-20.

[10] Reardon MJ, Van Mieghem NM, Popma JJ, Kleiman NS, Søndergaard L, Mumtaz M, et al. Surgical or Transcatheter Aortic-Valve Replacement in Intermediate-Risk Patients. N Engl J Med. 2017;376(14):1321-1331.

[11] Smith CR, Leon MB, Mack MJ, Miller DC, Moses JW, Svensson LG, et al. Transcatheter versus surgical aortic-valve replacement in high-risk patients. N Engl J Med 2011;364(23):2187-98.

[12] Thyregod HG, Steinbrüchel DA, Ihlemann N, Nissen H, Kjeldsen BJ, Petursson P, et al. Transcatheter Versus Surgical Aortic Valve Replacement in Patients With Severe Aortic Valve Stenosis: 1-Year Results From the All-Comers NOTION Randomized Clinical Trial. J Am Coll Cardiol 2015;65(20):2184-94.

[13] Adams DH, Popma JJ, Reardon MJ, Yakubov SJ, Coselli JS, Deeb GM, et al. Transcatheter aortic-valve replacement with a self-expanding prosthesis. N Engl J Med 2014;370(19):1790-8.

[14] Nielsen HH, Klaaborg KE, Nissen H, Terp K, Mortensen PE, Kjeldsen BJ, et al. A prospective, randomised trial of transapical transcatheter aortic valve implantation vs. surgical aortic valve

replacement in operable elderly patients with aortic stenosis: the STACCATO trial. EuroIntervention 2012;8(3):383-9.

FIGURES LEGEND

Figure 1. Schematic diagram of fixed and random effect models' assumption. In fixed effect model, there is no heterogeneity and the variance is completely due to spurious dispersion. Summary effect is the estimate of the true effect (μ). In random effect model, the true effect sizes are different and consequently there is between-studies variance. The summary effect is the estimate of the mean of the true effect sizes distribution, with an estimated variance of T^2.

Figure 2. Flow chart of the array of statistics for abstracting interpretation of results.

Figure 3. Random effects meta-analysis of 6 trials that examine the effect of TAVR vs SAVR on 30-day incidence of mortality (Panel A) and pacemaker implantation (Panel B). In forest plot for 30-day mortality, there is no heterogeneity and the random effects analysis reduces to fixed effects analysis. In Panel B, heterogeneity is significant and the summary effect is an estimates of the true effects sizes.

Figure 4. Comparison between random and fixed effects models in the example. Both fixed and random effects model were applied to the example in order to underscore the differences on estimation. Fixed effects model is reported in red, random effect model is depicted in black. In Panel A, there is coincidence between the two models, as heterogeneity is null and random effects model is reduced to fixed-effect model. In the second outcome (Panel B), there is a significant heterogeneity and hence different estimates are obtained applying fixed or random effect model, as fixed effect model does not consider the between-studies variance and summary estimate is performed forcing $T^2=0$, although it is significant (red). The appropriate choice of random effects model (black) leads to more balanced relative weights and to wider variance, standard error and confidence interval for the summary risk difference.

Central picture. Schematic differences between of fixed and random effect models' assumptions.

Figure 1.

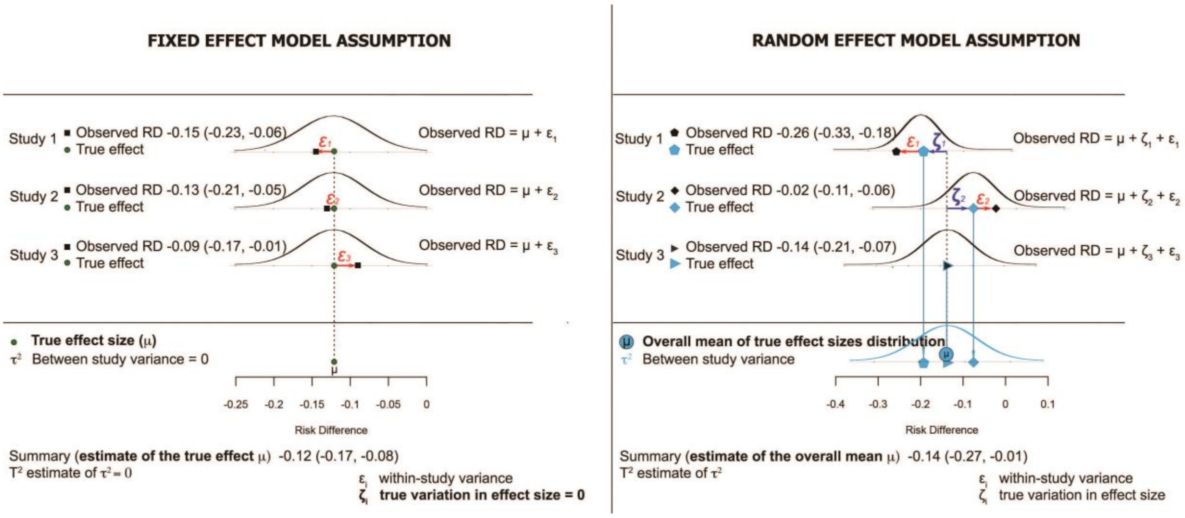
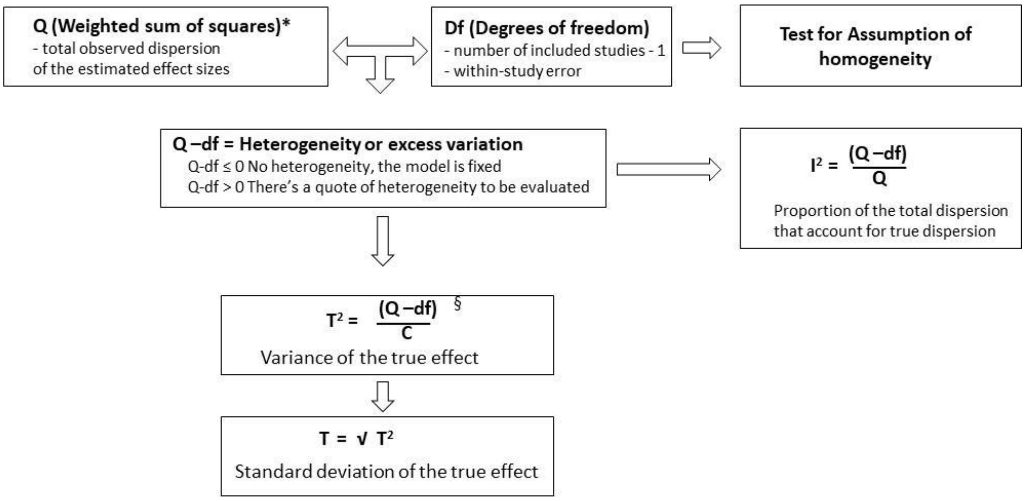

FIXED EFFECT MODEL ASSUMPTION
Study 1 ■ Observed RD -0.15 (-0.23, -0.06)
● True effect
Observed RD = μ + ε_1
ε_1
Study 2 ■ Observed RD -0.13 (-0.21, -0.05)
● True effect
Observed RD = μ + ε_2
ε_2
Study 3 ■ Observed RD -0.09 (-0.17, -0.01)
● True effect
Observed RD = μ + ε_3
ε_3
● True effect size (μ)
τ^2 Between study variance = 0
μ
-0.25
-0.2
-0.15
-0.1
-0.05
0
Risk Difference
Summary (estimate of the true effect μ) -0.12 (-0.17, -0.08)
T^2 estimate of $\tau^2 = 0$
ε_i within-study variance
ζ_i true variation in effect size = 0
RANDOM EFFECT MODEL ASSUMPTION
Study 1 ⬟ Observed RD -0.26 (-0.33, -0.18)
⬟ True effect
ε_1
ζ_1
Observed RD = μ + ζ_1 + ε_1
Study 2 ◆ Observed RD -0.02 (-0.11, -0.06)
◆ True effect
ζ_2
ε_2
Observed RD = μ + ζ_2 + ε_2
Study 3 ▸ Observed RD -0.14 (-0.21, -0.07)
▶ True effect
Observed RD = μ + ζ_3 + ε_3
μ Overall mean of true effect sizes distribution
τ^2 Between study variance
μ
-0.4
-0.3
-0.2
-0.1
0
0.1
Risk Difference
Summary (estimate of the overall mean μ) -0.14 (-0.27, -0.01)
T^2 estimate of τ^2
ε_i within-study variance
ζ_i true variation in effect size

Figure 2.


Q (Weighted sum of squares)*
- total observed dispersion of the estimated effect sizes
Df (Degrees of freedom)
- number of included studies - 1
- within-study error
Test for Assumption of homogeneity
Q –df = Heterogeneity or excess variation
Q-df ≤ 0 No heterogeneity, the model is fixed
Q-df > 0 There's a quote of heterogeneity to be evaluated
$I^2 = \frac{(Q-df)}{Q}$
Proportion of the total dispersion that account for true dispersion
$T^2 = \frac{(Q-df)}{C}$ §
Variance of the true effect
T = √ T²
Standard deviation of the true effect

* $Q = \sum^{k} W_i (Y_i - M)^2$

§ $C = \sum W_i - \frac{\sum W_i^2}{\sum W_i}.$

Where:
i = study number
k = total number of studies included in the meta-analysis
Y_i = effect size for the study i
W_i = weight for the study i (inverse of its variance)
M = summary effect

Figure 3.

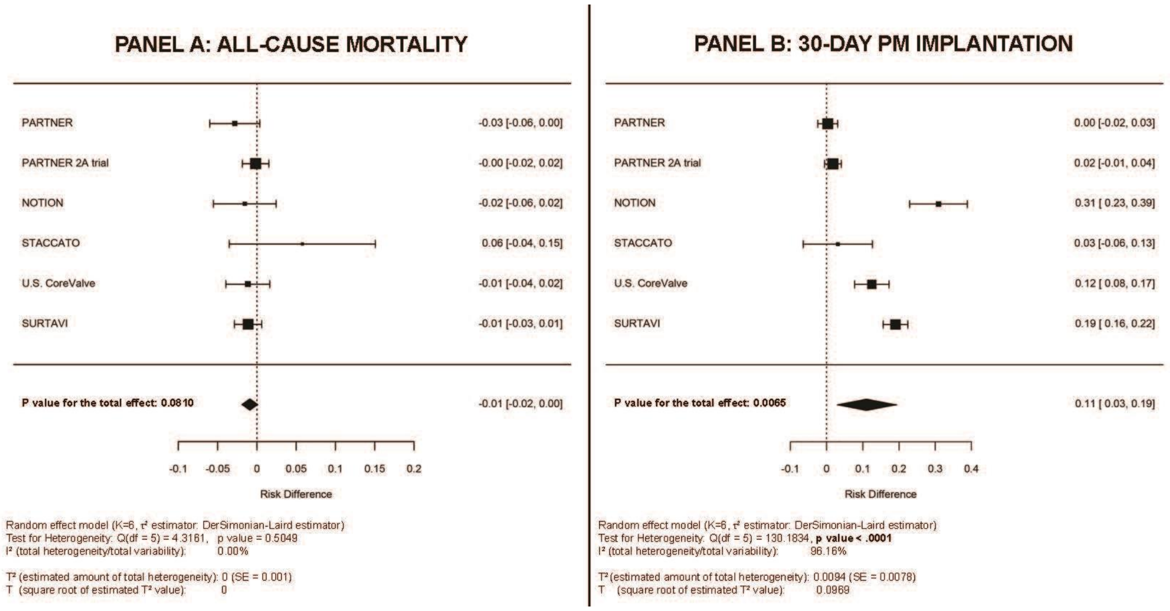
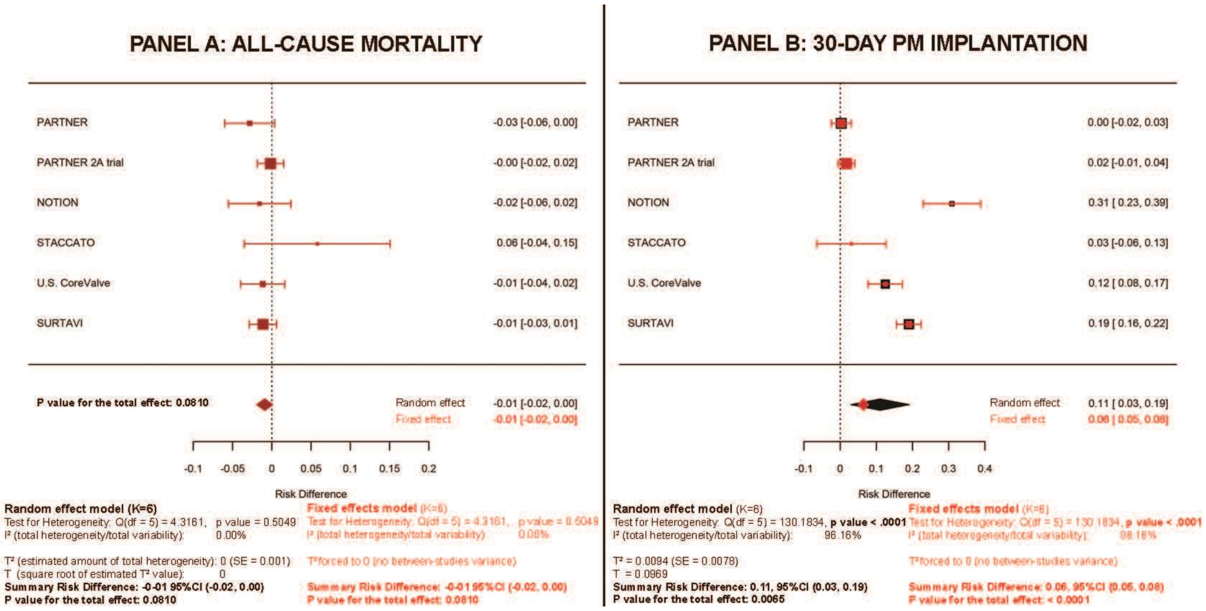

PANEL A: ALL-CAUSE MORTALITY
PARTNER
-0.03 [-0.06, 0.00]
PARTNER 2A trial
-0.00 [-0.02, 0.02]
NOTION
-0.02 [-0.06, 0.02]
STACCATO
0.06 [-0.04, 0.15]
U.S. CoreValve
-0.01 [-0.04, 0.02]
SURTAVI
-0.01 [-0.03, 0.01]
P value for the total effect: 0.0810
-0.01 [-0.02, 0.00]
-0.1
-0.05
0
0.05
0.1
0.15
0.2
Risk Difference
Random effect model (K=6, τ² estimator: DerSimonian-Laird estimator)
Test for Heterogeneity: Q(df = 5) = 4.3161, p value = 0.5049
I² (total heterogeneity/total variability): 0.00%
T² (estimated amount of total heterogeneity): 0 (SE = 0.001)
T (square root of estimated T² value): 0
PANEL B: 30-DAY PM IMPLANTATION
PARTNER
0.00 [-0.02, 0.03]
PARTNER 2A trial
0.02 [-0.01, 0.04]
NOTION
0.31 [0.23, 0.39]
STACCATO
0.03 [-0.06, 0.13]
U.S. CoreValve
0.12 [0.08, 0.17]
SURTAVI
0.19 [0.16, 0.22]
P value for the total effect: 0.0065
0.11 [0.03, 0.19]
-0.1
0
0.1
0.2
0.3
0.4
Risk Difference
Random effect model (K=6, τ² estimator: DerSimonian-Laird estimator)
Test for Heterogeneity: Q(df = 5) = 130.1834, p value < .0001
I² (total heterogeneity/total variability): 96.16%
T² (estimated amount of total heterogeneity): 0.0094 (SE = 0.0078)
T (square root of estimated T² value): 0.0969

Figure 4.


PANEL A: ALL-CAUSE MORTALITY
PARTNER
-0.03 [-0.06, 0.00]
PARTNER 2A trial
-0.00 [-0.02, 0.02]
NOTION
-0.02 [-0.06, 0.02]
STACCATO
0.06 [-0.04, 0.15]
U.S. CoreValve
-0.01 [-0.04, 0.02]
SURTAVI
-0.01 [-0.03, 0.01]
P value for the total effect: 0.0810
Random effect
-0.01 [-0.02, 0.00]
Fixed effect
-0.01 [-0.02, 0.00]
-0.1
-0.05
0
0.05
0.1
0.15
0.2
Risk Difference
Random effect model (K=6)
Test for Heterogeneity: Q(df = 5) = 4.3161, p value = 0.5049
I² (total heterogeneity/total variability): 0.00%
T² (estimated amount of total heterogeneity): 0 (SE = 0.001)
T (square root of estimated T² value): 0
Summary Risk Difference: -0-01 95%CI (-0.02, 0.00)
P value for the total effect: 0.0810
Fixed effects model (K=6)
Test for Heterogeneity: Q(df = 5) = 4.3161, p value = 0.5049
I² (total heterogeneity/total variability): 0.00%
T² forced to 0 (no between-studies variance)
Summary Risk Difference: -0-01 95%CI (-0.02, 0.00)
P value for the total effect: 0.0810
PANEL B: 30-DAY PM IMPLANTATION
PARTNER
0.00 [-0.02, 0.03]
PARTNER 2A trial
0.02 [-0.01, 0.04]
NOTION
0.31 [0.23, 0.39]
STACCATO
0.03 [-0.06, 0.13]
U.S. CoreValve
0.12 [0.08, 0.17]
SURTAVI
0.19 [0.16, 0.22]
Random effect
0.11 [0.03, 0.19]
Fixed effect
0.06 [0.05, 0.08]
-0.1
0
0.1
0.2
0.3
0.4
Risk Difference
Random effect model (K=6)
Test for Heterogeneity: Q(df = 5) = 130.1834, p value < .0001
I² (total heterogeneity/total variability): 96.16%
T² = 0.0094 (SE = 0.0078)
T = 0.0969
Summary Risk Difference: 0.11, 95%CI (0.03, 0.19)
P value for the total effect: 0.0065
Fixed effects model (K=6)
Test for Heterogeneity: Q(df = 5) = 130.1834, p value < .0001
I² (total heterogeneity/total variability): 96.16%
T² forced to 0 (no between-studies variance)
Summary Risk Difference: 0.06, 95%CI (0.05, 0.08)
P value for the total effect: < 0.0001