ANESTHESIA & ANALGESIA® Statistical Minute

Related Article, see p 1092

OPEN

Meta-Analysis in Clinical Research

Patrick Schober, MD, PhD, MMedStat,* and Thomas R. Vetter, MD, MPH†

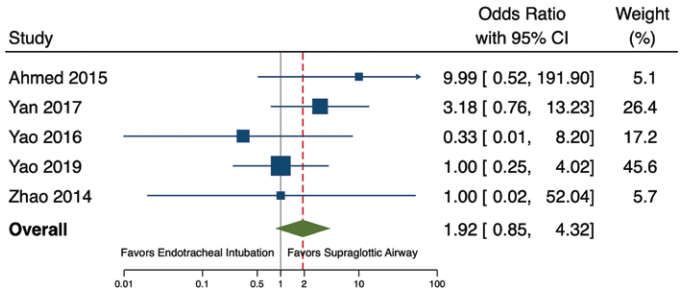

Figure. Forest plot of first-attempt success of supraglottic airway versus endotracheal tube intubation, redrawn based on data shown in Figure 3 from White et al.[1] Note that in contrast to the original figure, the x-axis is on the logarithmic scale, as commonly done when presenting ORs or RRs. An OR of 0.1 (= 1/10) represents the same absolute magnitude of effect as an OR of 10, and 0.5 (= 1/2) the same as 2. On the logarithmic scale, these values have the same distance from 1 (the null value indicating no effect), and moreover, the CIs (and the diamond representing the pooled or summary effect) are symmetric. This facilitates reading the graph, in particular, for values around 1. CI indicates confidence interval; OR, odds ratio; RR, relative risk.

KEY POINT: A meta-analysis is used to combine quantitative data from different studies addressing the same or similar research question.

In this issue of *Anesthesia & Analgesia*, White et al[1] report the results of a systematic review and meta-analysis of studies comparing first-pass success rate and time to insertion for supraglottic airway devices with endotracheal intubation for elective cesarean delivery.

A meta-analysis is used to combine quantitative data from different studies addressing the same or similar research question.[2,3] As in the study by White et al[1], meta-analyses most commonly synthesize effects of treatments or interventions across comparative studies in which effect sizes[4] (eg, mean differences between treatment groups for continuous outcomes like time to insertion, relative risks, or odds ratios for binary outcomes like first-pass success) are reported. However, the principles can essentially be applied to any point estimate for which a variance can be computed, including proportions, hazard ratios,[5] or regression coefficients.[6]

The pooled or summary effect value reported in a meta-analysis is a weighted average of the effect sizes from the individual studies with weights assigned based on how precisely each study estimates the effect size. The precision of a study is primarily driven by its sample size; therefore, in a meta-analysis, larger studies generally receive more weight and thus contribute more information than smaller studies. The exact method for assigning this relative weight depends specifically on whether a fixed-effect or random-effects model is used.

The fixed-effect model assumes that there is one true treatment effect or, more generally, true effect size or point estimate that all the included studies are estimating, and that any observed variation in effect sizes across the studies represents sampling error. With a fixed-effect meta-analysis, the pooled or summary effect estimates this common true effect size.

In contrast, under the random-effects model, the true effect of the treatment is assumed to vary from study to study, and observed variation between studies is a combination of true variation and sampling error. This is usually a more realistic assumption because patient and disease characteristics, as well as how the treatment is administered and how outcomes are assessed, usually vary from study to study. Under the random-effects model, the pooled or summary effect estimates the mean of the distribution of true effect sizes. Note, however, that a random-effects analysis can be problematic when there are only a few included studies—even though there is no consensus on what represents "few"—because, given that scenario, the between-study variation cannot be estimated with sufficient precision.

The meta-analysis allows for not only computing a summary effect with its confidence interval and statistical significance but also quantifying study heterogeneity. Statistical heterogeneity refers to the variation in true effect sizes between studies rather than variation due to sampling error. Statistical heterogeneity provides important but often undervalued information. For instance, when assessing a treatment effect, it is an important distinction whether the treatment (ie, intervention) consistently

From the *Department of Anesthesiology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands; and †Department of Surgery and Perioperative Care, Dell Medical School at the University of Texas at Austin, Austin, Texas.

Address correspondence to Patrick Schober, MD, PhD, MMedStat, Department of Anesthesiology, Amsterdam UMC, Vrije Universiteit Amsterdam, De Boelelaan 1117, 1081 HV Amsterdam, the Netherlands. Address e-mail to p.schober@amsterdamumc.nl.

reduces an adverse outcome by 10% across all study conditions or whether the treatment reduces the risk by 30% under some conditions and increases the risk by 10% under other conditions. Such heterogeneity can be further explored by subgroup analyses and meta-regression.

Several measures of heterogeneity are available, and researchers tend to focus on the I^2 statistic, which estimates the percentage of variation that is real rather than by chance. While this provides important information, it is also important to realize that the I^2 statistic is a relative measure. Like the Q statistic that is used to test whether there is evidence for a significant heterogeneity, the I^2 statistic does not provide any information on the absolute magnitude of the heterogeneity. The estimate of the between-study variance, τ^2, quantifies the absolute dispersion of the effect sizes around their mean. From this variance, a 95% prediction interval can be derived, which is an estimate of the dispersion of the true effect sizes reported in the same unit of measurement as the effect size itself. In 95% of cases, the true effect size of a new study is expected to fall inside this prediction interval.

Results of a meta-analysis are commonly presented as a forest plot (Figure) in which the individual effect sizes are displayed for each study as a square for its point estimate and a horizontal line for the range of its confidence interval. The area of this square is proportional to the assigned study weight. The pooled or summary effect is shown as a diamond with its vertical center representing the pooled point estimate and the width of the diamond representing the range of its confidence interval.

REFERENCES

1. White LD, Thang C, Hodsdon A, et al. Comparison of supraglottic airway devices with endotracheal intubation in low-risk patients for cesarean delivery: systematic review and meta-analysis. *Anesth Analg*. 2020;131:1092–1101.
2. Vetter TR. Systematic review and meta-analysis: sometimes bigger is indeed better. *Anesth Analg*. 2019;128:575–583.
3. Schober P, Vetter TR. Systematic review in clinical research. *Anesth Analg*. 2020;131:1090–1091.
4. Schober P, Vetter TR. Effect size measures in clinical research. *Anesth Analg*. 2020;130:869.
5. Schober P, Vetter TR. Survival analysis and interpretation of time-to-event data: the tortoise and the hare. *Anesth Analg*. 2018;127:792–798.
6. Vetter TR, Schober P. Regression: the apple does not fall far from the tree. *Anesth Analg*. 2018;127:277–283.