STATISTICS AND STATISTICAL SOFTWARE

We performed meta-analyses with R, version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria: www.R-project.org) [1] and the packages meta [2], metapoweR [3], tidyr [4], and ggplot2 [5] in RStudio, version 1.1.463 (RStudio: Integrated Development for R. RStudio, Inc., Boston, MA: http://www.rstudio.com/) [6]. In a meta-analysis, the sample size of each study and the between-studies heterogeneity influence the weight of the study on the pooled estimate. Because of expected heterogeneity between the populations of different studies (for example, in age, etiology, neurodevelopment, etc.), we considered a priori a random-effects model. This a priori choice was supported by the results of between-study heterogeneity as measured by the I^2 index [7].

Publication bias occurs when the probability of a study being published depends on the direction or strength of the study results. Often, studies with statistically significant results are more likely to be published [8]. When publication bias occurs, the published set of data may not be a representative sample of reality. A meta-analysis always yields a mathematically accurate synthesis of the studies included in the meta-analysis [8]. However, if the available studies are not an accurate representation of reality because some studies are published and others are not, the meta-analysis results may suffer from this publication bias [8]. To find all existing relevant literature, we systematically searched not only PubMed, but also EMBASE and Web of Science, which yield "gray literature" (abstracts, theses, etc.) that cannot be found in Pubmed.

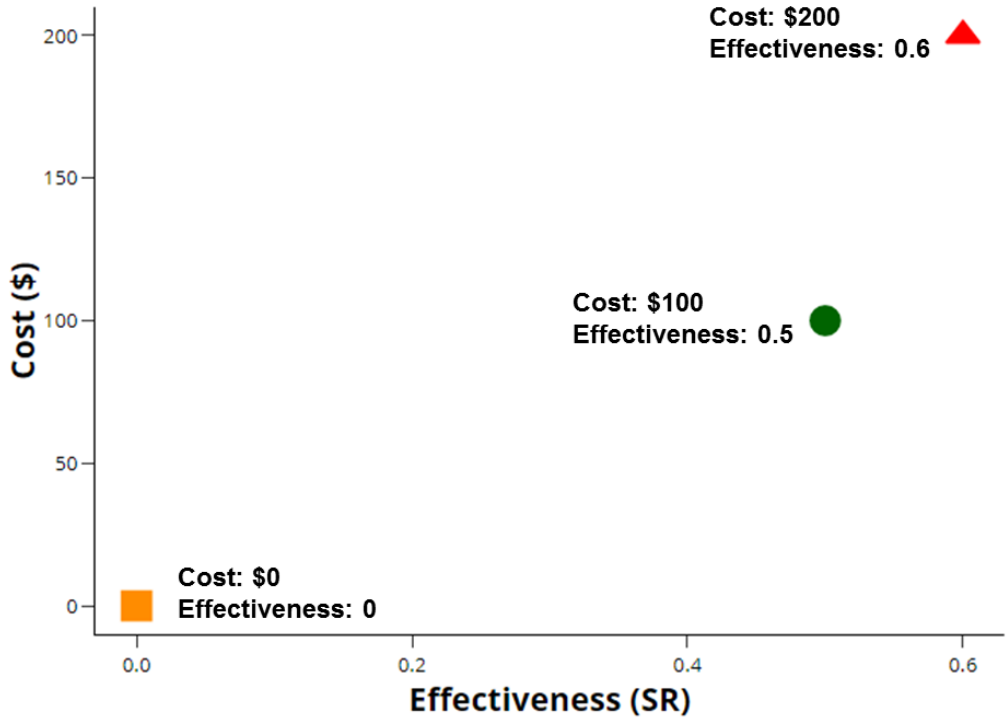
Even when all searchable literature has been included, it is possible that some studies were never published, not even as abstracts or theses. Funnel plots are the most common method for evaluating potential publication bias. In a funnel plot, the effect size is on the x axis and the sample size (or an equivalent measure such as variance or standard error) is on the y axis. The larger the study, the closer it will be to the real effect size. Smaller studies have more random error variation in the effect size. These

smaller studies appear toward the bottom of the funnel plot and tend to be spread across a broad range of values of effect sizes. If there is no publication bias the studies will be distributed symmetrically around the real effect size, since the sampling error is random and all small studies, regardless of their statistical significance or direction of effect size, will be published. In contrast, when there is publication bias, there is an asymmetry with more "missing" studies towards the bottom of the funnel plot (small studies tend to be published only if statistically significant or only in a certain direction of effect size). Visually evaluating funnel plots for publication bias is open to subjective judgments.

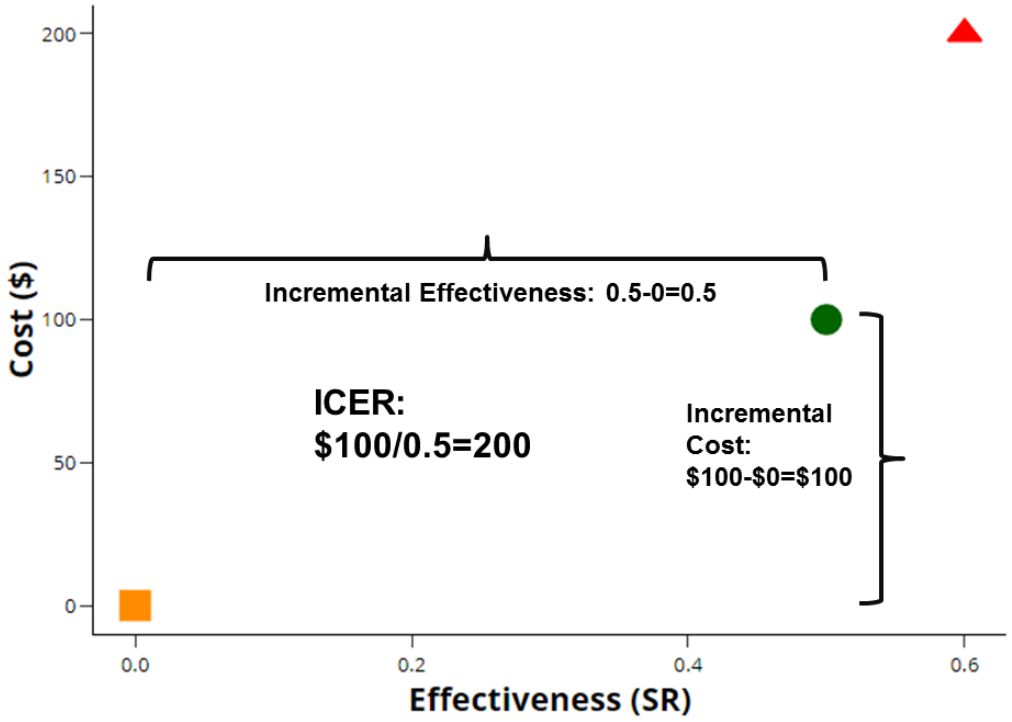
In contrast, Duval and Tweedie trim and fill method is an objective algorithm to adjust for publication bias and iteratively calculate the unbiased effect size [9]. The trim and fill method automates the detection and correction for publication bias in two steps. Initially, studies that cause funnel plot asymmetry are removed (trimmed). The resulting plot is largely unaffected by publication bias. Then, based on that plot and that subset of studies, the initially trimmed studies are brought back with their "filled" counterparts. We used the trim and fill algorithm by Duval and Tweedie with the L_0 estimator [9], and identified the direction of potentially missing studies with the Egger method [10]. For the trim and fill method, we used a fixed-random model, which is a fixed-effect model (less likely to be influenced by small biased studies: publication bias) to trim and fill studies and a random-effects model (more appropriate for studies coming from heterogeneous populations) to estimate the pooled effect of the studies once they are trimmed and filled [11]. In the Duval and Tweedie trim and fill method, the fixed-random model performs better than the fixed-fixed model and no worse and marginally better in certain situations than the random-random model [11].

The nput parameters for effectiveness were estimated from the result of our meta-analysis and represented with a beta distribution. The input parameters for cost were estimated from publicly available data and represented with a triangular distribution. We used TreeAge Pro 2015 (TreeAge Software, Inc., Williamstown, MA) [12] for cost-effectiveness models. We used R packages shiny [13], ggplot2 [5], and plotly [14] for the interactive versions of the meta-analyses.

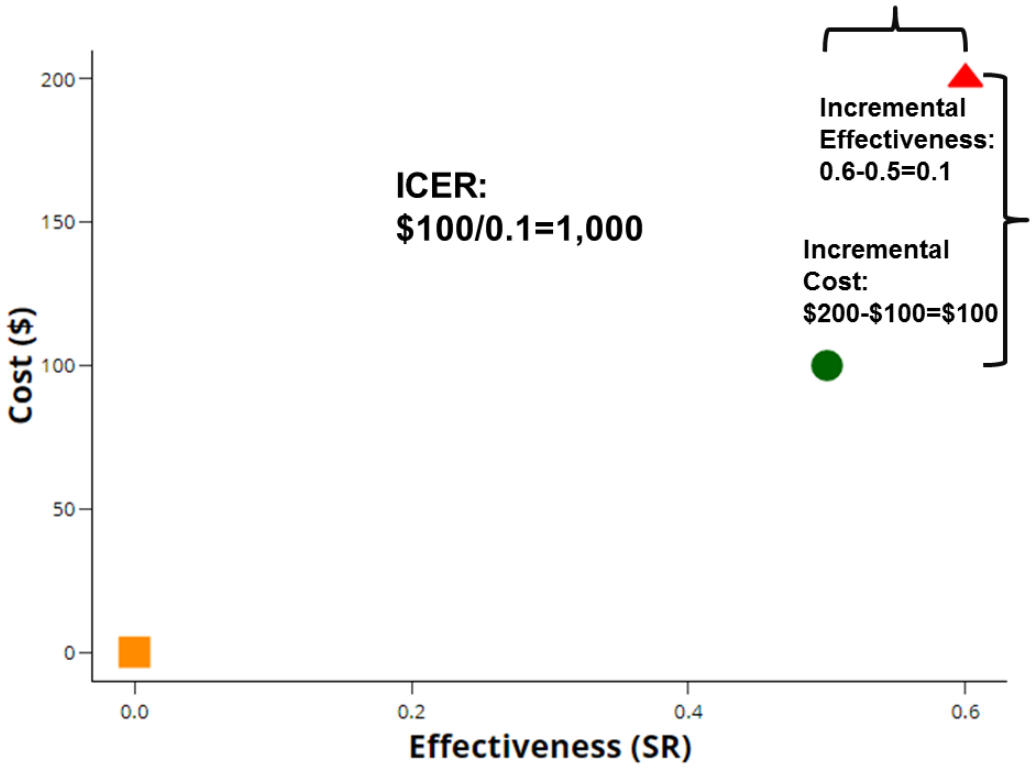
We used the incremental cost-effectiveness ratio to quantify cost-effectiveness. The ICER measures the relative cost-effectiveness of competing treatments by dividing the incremental cost of a treatment (compared to the next most cost-effective treatment) by the incremental effectiveness (compared to the next most cost-effective treatment) [15]. Although detailed explanations of ICER, the efficiency frontier, and dominance are beyond the scope of this study, we provide a graphical explanation to roughly grasp these concepts and we refer the interested reader elsewhere to more in-depth explanations and calculations [15]. First, the ICER is a comparative measure between treatments, not a characteristic of any individual treatment. For illustration purposes, let's calculate the ICER for three treatments with simulated data.



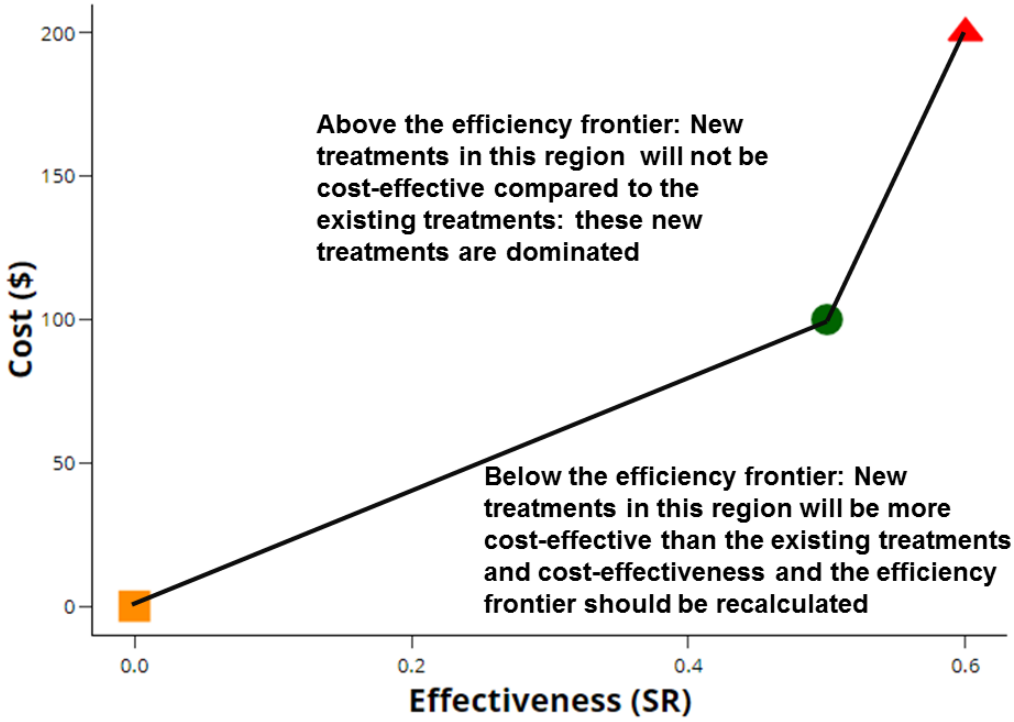
The baseline "no treatment" is the orange square with $0 cost and 0 probability of spasm resolution (SR); the green circle represents "treatment A" with $100 cost and 0.5 probability of spasm resolution; and the red triangle is "treatment B" with $200 cost and 0.6 probability of spasm resolution.



To calculate the ICER of "treatment A" we have to reference cost and effectiveness to the next most cost-effective alternative, which is "no treatment". The incremental cost is $100-$0 = $100 and the incremental effectiveness is 0.5-0 = 0.5. The ICER is $100/0.5 = 200, that is, $200 will have to be spent per each case of infantile spasms resolved with "treatment A".



To calculate the ICER of "treatment B" we have to reference cost and effectiveness to the next most cost-effective alternative, which is "treatment A". The incremental cost is $200-$100 = $100 and the incremental effectiveness is 0.6-0.5 = 0.1. The ICER is $100/0.1 = 1,000, that is, $1,000 will have to be spent per each case of infantile spasms resolved with "treatment B". It may seem more intuitive to compare the cost and effectiveness of "treatment B" with the baseline of "no treatment", but that comparison is incorrect in a competing choice problem because it does not take into account that one can obtain an effectiveness of 0.5 at a cost of $100 with "treatment A".



The efficiency frontier is determined by the line that joints the most cost-effective treatments: new treatments above and to the left of the efficiency frontier are not cost-effective compared to the existing treatments. In contrast, new treatments below and to the right of the efficiency frontier are more cost-effective than the existing treatments and, in case that these new treatments appear, a new calculation of cost-effectiveness and of the efficiency frontier should be performed.

For more details on how ICER is calculated, the difference between a competing choice and a shopping spree problem, implications of the efficiency frontier, and the difference between absolute dominance and extended dominance, we refer the reader to a more in-depth textbook on decision analysis [15].

We also provided intuitive measures of effect such as the number needed to treat and its clinical meaning and the ratio of ICER of the compared options and their clinical meaning.

REFERENCES

1. Team RC. R: A language and environment for statistical computing, version 3.5.1. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. 2018.
2. Schwarzer G. meta: An R package for meta-analysis, R News, 7(3), 40-45. 2007.
3. Griffin JW. metapoweR: an R package for computing meta-analytic statistical power. R package version 0.2.0. URL: https://CRAN.R-project.org/package=metapower. 2020.
4. Wickham H, Henry L. tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions. R package version 0.8.3. URL: https://CRAN.R-project.org/package=tidyr. 2019.
5. Wickham H. ggplot2: Elegant Graphics for Data. Analysis. Springer-Verlag New York, 2016. URL http://ggplot2.org. 2016.
6. RStudio Team. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. URL: http://www.rstudio.com/. 2016.
7. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Bmj 2003;327:557-60.
8. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Publication bias. In: Borenstein M, Hedges LV, Higgins JPT, Rothstein HR, editors. Introduction to Meta-Analysis. United Kingdom: Wiley; 2009. p. 277-92.
9. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455-63.
10. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Bmj 1997;315:629-34.
11. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. Statistics in medicine 2007;26:4544-62.
12. TreeAge Software. TreeAge Pro 2015. Williamston, MA. URL: https://www.treage.com/ 2015.
13. Chang W, Cheng J, Allaire J, Xie Y, McPherson J. shiny: Web Application Framework for R. R package version 1.0.5. https://CRAN.R-project.org/package=shiny. 2017.
14. Sievert C, Parmer C, Hocking T, Chamberlain S, Ram K, Corvellec M, et al. plotly: Create Interactive Web Graphics via 'plotly.js'. R package version 4.7.1. https://CRAN.R-project.org/package=plotly. 2017.
15. Hunink MMG, Weinstein MC, Wittenberg E, Drummond MF, Pliskin JS, Wong JB, et al. Decision making in health and medicine. Integrating evidence and values. Second ed. Cambridge: Cambridge University Press; 2014.