1. PREFACE

Most published meta-analyses include a plot that displays the effect size for each study, followed by a row that displays the summary effect size. This is called a forest plot, since it allows us to see "the forest and the trees" – the forest being the larger picture, and the trees being the individual studies.

Beginning in the early 1990s, a substantial number of papers referred to these as "Forrest" plots rather than "forest" plots, and one might ask why. It turns out that the eminent researcher Richard Peto had given a talk at a conference, and he noticed that Dr. Christine Forrest was in the audience. As a joke, Dr. Peto remarked that the plot had been named in honor of Dr. Forrest. Someone in the audience took this remark seriously, and used the spelling "Forrest" in their next publication. This spelling was copied once, then again, and soon became codified as the correct spelling in some journals (Lewis & Clarke, 2001).

In this anecdote, it is simple to see how the mistake originated and perpetuated itself. By contrast, there are many other mistakes that appear on a regular basis in the meta-analysis literature, where the origin of the mistake is less clear. Some of these have no more basis than the one cited above, yet have become the default way to do things. Unlike the Forrest mistake, some of these have profound implications for the analyses.

For example –

- Many researchers believe that if the impact of an intervention is statistically significant, we can assume that the intervention is effective in all populations. In fact, the intervention may be helpful in some populations and harmful in others.
- Most researchers report I^2 as an index of heterogeneity, and assume that I^2 values of 25%, 50%, and 75% reflect low, moderate, and high amounts of dispersion. This represents a fundamental misunderstanding of what I^2 actually tells us.
- Many researchers think that the confidence interval in a random-effects analysis tells us how widely the effects vary. It does not.
- Many researchers use a test for heterogeneity to decide whether to use a fixed-effect or random-effects model. This is invariably a mistake, and reflects a misunderstanding of what a statistical model means.
- Most researchers assume that if a test for publication bias is statistically significant, we can assume that bias exists. In fact, there are many reasons

that the test might be statistically significant, and bias is only one of several plausible explanations.

Of course, the incorrect interpretation of the various statistics is not intentional. Rather, as in the Forrest example, it is simply what happens when people model their own reports on prior publications, and perpetuate mistakes. My goal in this volume is to call attention to mistakes that appear repeatedly in the meta-analysis literature. I explain *what* the mistake is, why it *is* a mistake, and *how to avoid* the mistake.

In each section I start with an overview that provides context for understanding the issue discussed in that section, and list the common mistakes in that area. Then, each of the mistakes is addressed in detail, along with examples and solutions. Those who are looking for a simple fix will find it, and those who want more detail can study the examples. There is also an appendix with additional information on the statistical issues.

At the end of the volume I have included a section with examples of how I would transcribe the results section of an analysis. In these examples I implement the suggestions outlined in the prior chapters. These can serve as templates for your own analyses.

To illustrate the various mistakes, I have included examples from the published literature. I sincerely hope that I have not offended people by using their papers in this way, and would like to make the following points.

- If I included an example of a mistake, it is NOT an isolated mistake. Every mistake cited in this volume has been repeated *many* times in the published literature.
- I avoided picking examples from meta-analyses that were performed poorly. Rather, I tried to pick examples from analyses that were generally of a high standard but nevertheless included the mistake in question. If your analysis is included here, it is one of the better analyses I have seen.

If anyone still feels offended, I do apologize. I made many of these same mistakes myself when I started working in this field. I would still be making those mistakes to this day, except that I was fortunate enough to have colleagues who provided guidance and clarity. I hope this book serves that same purpose for others.

Michael Borenstein

New York City, 2019