DOCTORAL THESIS

Some new developments in data transformation and meta-analysis with small number of studies

Lin, Enxuan

Date of Award:
2019

Abstract

Meta-analysis is an important statistical tool for systematic reviews and evidence-based medicine. Extracting the observed effect sizes, assessing the magnitude of heterogeneity, choosing the suitable statistical model, and interpreting the summary effect size are four key steps in meta-analysis. It is known that each of the above steps has its own unique characteristics and may require some specific attention. As an example, the observed effect sizes from individual studies may not be reported in the same scale and hence cannot be combined directly. Another example is on selecting a model for meta-analysis from the common-effect model and the random-effects model. When a meta-analysis contains only few studies, the common-effect model and the random-effects model will often lead to misleading or unreliable results. In the first part of the thesis, we give a brief introduction on evidence-based medicine, systematic reviews and meta-analysis. We will also show their practical importance, display their relationships, and present a motivating example for conducting a meta-analysis.

In Chapter 2, we first review the common effect sizes in meta-analysis for both continuous data and binary data. How to combine different categories of effect sizes is a critical issue after extracting the observed effect sizes from the clinical studies in the literature. For continuous data, researchers have recently proposed methods that transform the five number summary to the sample mean and standard deviation (Hozo et al., 2005; Wan et al., 2014; Luo et al., 2018). For binary data, the transformation from the odds ratio (OR) to the relative risk (RR) in the cohort study was proposed by Zhang and Yu (1998). To the best of our knowledge, however, there is little work in the literature that converts OR to RR in the case-control study. In view of this, we establish a new formula for this transformation to fulfill the gap. The performance of the new method will be examined through simulations and real data analysis. Our method and formulas can not only handle meta-analyses with different effect sizes, but also offer some insights for medical researchers to further understand the meaning of OR and RR in both cohort and case-control studies.

In Chapter 3, we first give a brief introduction on the three available models in meta-analysis: the common-effect model, the random-effects model, and the fixed-

effects model. When a meta-analysis contains only few studies, the common-effect model and the random-effects model will often lead to misleading or unreliable results. In contrast, the fixed-effects model is capable to provide a good compromise between the existing two models. In this chapter, we propose to further improve the estimation accuracy of the *average effect* in the fixed-effects model by assigning different weight for each study as well as fully utilizing the information in the within-study variances. Through theory and simulation, we demonstrate that the fixed-effects model can serve as the most convincing model for meta-analysis with few studies. And most importantly, with a total of three models, we expect that meta-analysis can be conducted more flexibly, more meaningfully, and more accurately.

In Chapter 4, we first give a brief introduction on the heterogeneity in meta-analysis. We then review the methods for quantifying heterogeneity in three directions as follows: the tests for heterogeneity, the estimates of the between-study variance, and the measures of the impact of heterogeneity. Note that most existing methods were derived under the assumption of known within-study variances. In practice, however, a direct use of the reported within-study variance estimates may largely reduce the power of the tests and also lower the accuracy of the estimates, especially when the sample sizes in some studies are not sufficiently large. To overcome this problem, we propose a family of shrinkage estimators for the within-study variances that are able to borrow information across the studies, and derive the optimal shrinkage parameters under the Stein loss function. We then apply the new estimates of the within-study variances to some well known methods for measuring heterogeneity. Simulation studies and real data examples show that our shrinkage estimators can dramatically reduce the estimation bias and hence improve the exiting literature.

Keywords: Common-effect model, Effect size, Fixed-effects model, Heterogeneity, Meta-analysis, Odds ratio, Random-effects model, Relative risk, Risk ratio

Table of Contents