Statistics and Its Interface Volume 13 (2020) 423–424

Editorial

Research in meta-analysis is rapidly expanding and diversifying (e.g. meta-analysis of multiple outcomes; network meta-analysis of multiple treatments; network meta-analysis multiple diagnostic tests; and meta-analysis with individual participant data), making it increasingly popular and widely used in many fields including medicine, biology, public health, epidemiology, engineering, finance, economics, environmental sciences, and social sciences. Meta-analysis combines and contrasts multiple studies into a form of evidence that can be used to underpin guidelines, decision aids, and other products. The evidence from meta-analysis is commonly considered to be at the top of evidence pyramid. In 2018, the U.S. Food and Drug Administration (FDA) released a draft guidance for industry entitled "Meta-Analyses of Randomized Controlled Clinical Trials to Evaluate the Safety of Human Drugs or Biological Products", which demonstrates the importance of meta-analysis in the development of new drugs.

Taking advantage of this precious opportunity, the two co-guest editors and the two co-editors-in-chief announced the Call for Papers and advertised it in several major professional societies including ASA, ICSA, SRSM and ISBA. This special issue reflects the state-of-art of research on the development of statistical methods and computational algorithms at the frontier of this vital and rapidly developing area. All submissions went through a regular review process and two co-guest editors handled the peer review of all invited and contributed submissions. Finally, a total of ten high-quality articles are published in this special issue, which cover a broad range of topics related to meta-analysis, including two overview articles, one article on proteomic data, one article on few studies, two articles on network meta-analysis, two articles on individual patient data, and two articles on non-normal data.

Quantifying the heterogeneity is an important issue in meta-analysis when synthesizing the individual studies, and meta-analysis with binary rare events is another practical issue that attracts more attention in the recent literature. Taking the two issues together, **Zhang, Chen and Wang** summarize eleven descriptive measures, twenty-three estimators, and sixteen confidence intervals for the amount of heterogeneity. They further categorize these methods according to their key features, and evaluate their performance based on simulation studies under various realistic scenarios for rare binary events. To conclude, they provide some useful practical guidelines based on empirical evidences on which the methods is consistently better in the context of rare binary events.

In addition to the aforementioned questions, the small-study effect (SSE) is another critical issue in meta-analysis that may affect decision making. In view of this, **Marks-Anglin and Chen** provide a critical overview on the commonly used methods for detecting and correcting for SSE, including the graph-based methods and the selection models. They further point out that, even with decades of methodological development, it still remains an active research area in statistics with a substantial room for improvement and innovation that can be readily implemented and extended to more complex meta-analytic frameworks, as well as more robust testing procedures.

Meta-analysis also plays an important role in modern genomic research including, for example, combining multiple transcriptomic studies to identify differentially expressed genes, and integrating multiple genomic studies for pathway enrichment analysis. As a key paper in this special issue, **Zhang, Ouyang, Qian, Smith, Wong and Davis** extend the rationale of meta-analysis to model proteomic data from high-throughput shotgun assays. Specifically, they begin with an additive model to obtain peptide-level significance and then adaptively select peptides to make protein-level inference through meta-analysis. Their proposed method, Protein Expression through Adaptive Thresholding (PEAT), is general, flexible and can be adapted to data analysis of other types of shotgun assays as well.

The common-effect model and the random-effects model are the two most popular models for meta-analysis in the literature. Recently, it is recognized that the fixed-effects model is an important alternative method for meta-analysis, especially when the number of studies is small. With this new model, the existing methods are no longer sufficient for model selection in meta-analysis. In view of the demand, **Yang, Kwan, Yu and Tong** propose a novel method for model selection between the fixed-effects model and the random-effects model. Specifically, they apply the Akaike information criterion (AIC) to both models and then select the model with a smaller AIC value. They further propose the generalized AIC to reduce the large variation in the AIC value, and demonstrate its superiority through real data analysis and simulation studies. The authors also claimed that this is the first work in meta-analysis for model selection between the fixed-effects model and the random-effects model.

Network meta-analysis has been attracting more and more attention in evidence-based medicine for synthesizing both direct and indirect evidence from multiple treatments. The Bayesian hierarchical model is a popular method to implement network meta-analysis, yet how much improvement it can achieve over a pairwise meta-analysis has never been studied theoretically. To fill the gap, **Lin, Chu and Hodges** show that such improvement depends highly on ev-

idence cycles in the treatment network in the contrast-based network meta-analysis. When all treatment comparisons are assumed to have different heterogeneity variances, a network meta-analysis produces posterior distributions identical to separate pairwise meta-analyses for treatment comparisons that are not contained in any evidence cycles. However, this equivalence does not hold under the commonly-used assumption of a common heterogeneity variance for all comparisons. As pointed out by the authors, the results on the evidence cycles also provide some useful insights for journal editors, reviewers and investigators to conduct and evaluate future network meta-analyses.

Sample size calculation is a practical and important problem in the design of most clinical research. Nevertheless, even though systematic reviews are considered the pinnacle of evidence-based medicine, current sample size calculations usually do not take into account the existing body of evidence. To advocate the idea that sample size calculations should be conducted in the context of a systematic review and meta-analysis of the existing body of evidence, **DeSantis and Hwang** present an interesting framework to estimate the sample size and power for a future study, based on a prospective multivariate network meta-analysis (MNMA) of randomized clinical trials (RCTs). They further apply their approach to a systematic review of pharmacologic treatments for adult acute manic disorder, and suggest that new trials should be designed/powered within the context of either a multivariate or univariate network meta-analysis.

With increasingly accessible individual patient data (IPD) in the era of big data, it is practically feasible to conduct more precise and informative meta-analysis for better decision making. And consequently, new and novel methods for the IPD meta-analysis are also increasingly desired. Motivated by twenty-six pivotal Merck clinical trials, **Kim, Chen, Ibrahim, Shah and Lin** propose a flexible class of multivariate meta-regression models for IPD. Their proposed multivariate meta-regression models allow for different skewness parameters and different degrees of freedom for the multivariate outcomes from different trials under a general class of skew t-distributions. To conclude, the work provides a novel extension of the multivariate skew meta-regression model and can serve as an effective modeling tool for explaining heterogeneity between trials, synthesizing evidence across studies, investigating individual-level interactions, and/or identifying subgroups.

For the IPD meta-analysis, it is also known that simultaneously combining multiple related parameters across heterogeneous studies can be challenging because each parameter from each study has a specific interpretation within the context of the study and other covariates in the model. **Jiao, Mun, Trikalinos and Xie** propose a novel mapping method to combine within-study estimates of multiple related parameters across heterogeneous studies, which ensures valid inference at all levels by combining sample-dependent functions known as confidence distributions. They further propose a mapping method and provide a data application for a multivariate random-effects meta-analysis model. The mapping method is also shown to provide a robust methodological solution when combining complex evidence using IPD.

Most existing methods or models for meta-analysis assume that the random effects follow a normal distribution. In contrast, meta-analysis for non-normal data can be much more challenging. Motivated by the self-thinning meta-data with missing sample sizes and outliers, **Ma, Chen and Tang** propose a random-effects meta-analysis model with unknown precision parameters with a truncated Poisson regression model for missing sample sizes. The random effects are assumed to follow a heavy-tailed distribution to accommodate outlying aggregate values in the response variable. They further apply the logarithm of the pseudo-marginal likelihood for model comparison, and also develop a plausibility index to determine which self-thinning law is more supported by the meta-data.

To conduct a meta-analysis, the mean and variance from each individual study are often required; whereas in certain studies, researchers may instead report the five-number summary. To transform the five-number summary back to the mean and variance, a few popular methods have emerged in the literature under the normality assumption. In consideration that the normality assumption may be violated, **Shi, Tong, Wang and Genton** propose a three-step method for estimating the mean and variance from the five-number summary of a log-normal distribution. They also propose a bias-corrected method to further improve the estimation of the mean and variance, and demonstrate through simulation studies that their new estimators perform better than the normal-based estimators in most settings.

To conclude, meta-analysis is a rapidly growing research field for synthesizing multiple independent studies for decision making. We hope that this special issue promotes further research on more efficient statistical and computational methods for meta-analytical models, in particular those for complex data or big data. We also hope that this special issue makes Statistics and Its Inference (SII) a friendly home to many more exciting developments and innovations on research related to meta-analysis and network meta-analysis.

Haitao Chu (Co-Guest Editor)
University of Minnesota, Twin Cities

Tiejun Tong (Co-Guest Editor)
Hong Kong Baptist University

Ming-Hui Chen (Co-Editor-in-Chief)
University of Connecticut

Yuedong Wang (Co-Editor-in-Chief)
University of California, Santa Barbara