1. Introduction

In the last several decades, systematic reviews and meta-analyses have received increasing attention in comparative effectiveness research and evidence-based medicine. Meta-analysis is a statistical procedure that combines the results of multiple scientific studies. In meta-analysis, small study effects (SSE) is a well-known critical and challenging issue that may threaten the validity of the results (Song et al., 2000; Sutton et al., 2000). "Small-study effects" is a generic term for the phenomenon that smaller studies sometimes show different, often larger, treatment effects than large ones (Sterne and Egger, 2001). One of the most well-known reason for SSE is publication bias (PB), in which case the chance of a small study being published does not depend on its quality, but on its effect size, significance or direction. A possible reason is that authors tend to report significant or positive results or journals tend to publish studies with significant or positive results. Besides PB, outcome reporting bias (ORB) and clinical heterogeneity (i.e. variability in the participants, interventions and outcomes) in small studies are also important sources for SSE.

SSE (including PB) is arguably the greatest threat to the validity of meta-analysis (Schwarzer et al., 2015). Erroneous conclusions can arise from a meta-analysis if SSE is not properly accounted for. For example, conclusions from several meta-analyses were later found to be contradicted by mega-trials (Egger and Smith, 1995). In the last two decades, a great deal of effort has been devoted to better reporting protocols, risk of bias evaluation, and statistical methods to detect and correct for SSE based on reported studies (Rothstein et al., 2006; Bürkner and Doebler, 2014).

Among the many statistical methods on this issue, funnel plots have been commonly used to study SSE. Because the precision of an estimated treatment effect generally increases as the sample size of component studies increases, results from small studies typically scatter widely at the bottom of the funnel plot, while those from larger studies scatter narrowly