International Journal of Mathematical Sciences in Medicine (2013) 1–5

Kernel Density Estimation for Random-effects Meta-analysis

Branko Miladinovic[1], Ambuj Kumar[1] and Benjamin Djulbegovic[1,2]

[1]University of South Florida, Center for Evidence-based Medicine,
Tampa, FL 33612, USA
[2]H. Lee Moffitt Cancer Center and Research Institute, Department of Hematology,
Tampa, FL 33612, USA

ABSTRACT: Random-effects meta-analyses are commonly performed to combine estimates of treatment effect from different studies in the presence of heterogeneity. The method incorporates between-study variance in the overall estimate of summary effect and its standard error. In addition to calculating the summary effect which relates to the average treatment effect across all trials, prediction intervals have been recommended to give a range for the predicted parameter in a new study. As both the calculation of summary effects and prediction intervals rely on the assumption that the effects underlying different studies are normally distributed, in this manuscript we demonstrate how distribution assumption-free weighted kernel density estimation can be used to construct a probability distribution of observed effect sizes, thus gaining insight into the variability of summary effects. In our study, the weighted kernel density estimates are calculated using the Gaussian kernel and the adaptive bandwidth selection process. The weights are incorporated based on five different methods for estimating between-trial heterogeneity and sampling errors from all studies.

1 INTRODUCTION

The Institute of Medicine has defined generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor or improve the delivery of care as one of the key components of comparative effectiveness research (Sox, 2010). The report specifically recommends systematic reviews and meta-analyses for combining treatment effects from randomized clinical trials because they provide the highest level of scientific evidence to enable health care providers to offer better care for patients (Altman et al., 2001; Higgins & Green, 2008).

In general, by synthesizing treatment effects across independent studies any meta-analysis should address three issues related to overall treatment effects: central tendency, prediction, and variability. Central tendency relates to calculating the expected magnitude of effects across studies and the desired confidence interval around the overall estimate. Prediction relates to the need to explain the variability (or difference) of effect sizes across studies in terms of study level moderators, which is commonly addressed in meta-regression and subgroup analysis. Variability itself is generally addressed with tests of the homogeneity of effect sizes in random-effects meta-analysis (REMA), which averages treatment effects across studies by incorporating between-trial heterogeneity into the estimates of overall treatment effect and its standard error.

In this manuscript we present kernel density estimation (KDE) for visualizing evidence in the context of REMA. The traditional role of KDE for estimating the pattern of random variation in the data can be applied to REMA in order to present a non-parametric distribution of treatment effects. This is particularly important for the prediction of treatment effects in a new study where REMA commonly assumes that individual trial treatment effects vary across studies around a normal distribution. KDE can offer distribution-free insight into the pattern of random variation in treatment effects across studies.

2 Methods for REMA

2.1 Estimation of the Overall Population Parameter

Once the treatment effect sizes from independent studies have been extracted, the primary interest in any meta-analysis is the estimation of the overall population parameter. Given individual study summary effects θ_1,..., θ_n (mean, mean difference, relative risk, risk difference, odds ratio, hazard ratio, etc) for n studies, and their corresponding variances ${\sigma_1}^2$,..., ${\sigma_n}^2$, for between-study variance τ^2 a weighted estimator of the overall effect θ is given by

$$\hat{\theta} = \frac{\sum_{i=1}^{n} \hat{W}_i \hat{\theta}_i}{\sum_{i=1}^{n} \hat{W}_i} \tag{1}$$

where

$$\hat{W}_i = \frac{1}{(\hat{\tau}^2 + \hat{\sigma_i}^2)}.$$

Received August 15, 2013

An approximate 100(1-α)% confidence interval for θ may be obtained as

$$\hat{\theta} \pm Z_{\frac{\alpha}{2}} \hat{SE}(\hat{\theta})$$

and an approximate 100(1-α)% prediction interval for θ may be obtained as

$$\hat{\theta} \pm t_{\alpha,k-2} \sqrt{\hat{\tau}^2 + \hat{SE}(\hat{\theta})^2}$$

where $t_{\alpha,k-2}$ is the 100(1-$\frac{\alpha}{2}$) percentile of the t-distribution with k - 2 degrees of freedom. The approximate standard error of $\hat{\theta}$ is given by

$$\hat{SE}(\hat{\theta}) = (\sum_{i=1}^{n} \hat{W}_i)^{-0.5}$$

The normality assumption for individual trial treatment effects $\hat{\theta}_i \sim N(\theta_i, \sigma_i^2)$ and for the effect in the future study $\theta_{new} \sim N(\hat{\theta}, \tau^2 + SE(\hat{\theta})^2)$ allow for the calculation of confidence intervals and prediction intervals. However, empirical evidence suggests that the normality assumption for θ_i may not hold. In a cohort of all National Cancer Institute cooperative oncology groups (n = 614 trials), log hazard ratios for overall survival were not normally distributed (Djulbegovic et al., 2008). Statistical inference in the context of REMA based solely on the estimated average treatment effect across studies can be misleading and the accounting for τ^2 is as important.

2.2 Estimation of the overall between-study variance τ^2

Five different methods for the estimation of between-trial heterogeneity τ^2 have been discussed in literature, namely the Cochran ANOVA estimator (Cochran, 154), Paule and Mandel (Paule & Mandel, 1982) , DerSimonian-Laird (DerSimonian & Laird, 1986) and two versions of the DerSimonian and Kacker estimates (DerSimonian & Kacker, 2007). All estimates rely on equating the expression for the general moment generating estimate of τ^2, which is obtained by solving

$$\sum_{i=1}^{n} a_i(\theta_i - \theta)^2 = E(\sum_{i=1}^{n} a_i(\theta_i - \theta)^2) = \tau^2(\sum_{i=1}^{n} a_i - \frac{\sum_{i=1}^{n} {a_i}^2}{\sum_{i=1}^{n} a_i}) + (\sum_{i=1}^{n} a_i {\sigma_i}^2 - \frac{\sum_{i=1}^{n} {a_i}^2 {\sigma_i}^2}{\sum_{i=1}^{n} a_i}) \quad (2)$$

By substituting sample variances ${s_1}^2, ..., {s_k}^2$ for ${\sigma_1}^2, ..., {\sigma_k}^2$, the general method of moment estimator of τ^2 is (Kacker & Harville, 1984):

$$\hat{\tau}^2 = \frac{\sum_{i=1}^{n} a_i(\theta_i - \hat{\theta})^2 - \sum_{i=1}^{n} a_i {s_i}^2 + \frac{\sum_{i=1}^{n} {a_i}^2 {s_i}^2}{\sum_{i=1}^{n} a_i}}{\sum_{i=1}^{n} a_i - \frac{\sum_{i=1}^{n} {a_i}^2}{\sum_{i=1}^{n} a_i}} \quad (3)$$

For example, to derive the Cochrane ANOVA estimator of τ^2 we assume $a_i = \frac{1}{k}$ and for the DerSimonian-Laird estimator we assume $a_i = \frac{1}{{s_i}^2}$. The DerSimonian-Laird estimator has been the most widely used in practice.

3 Weighted Kernel Density Estimation

Kernel density methodology aims to estimate the density function of a random variable θ from a random sample θ_i without assuming that the function belongs to a known parametric family. By constructing a non-parametric density of treatment effects, we can gain insight into the natural variability of treatment effects across studies. Kernel density estimation (KDE) has been applied meta-analytically in neuro-imaging studies to synthesize complex evidence of brain function (Wager et al., 2003; Wager et al., 2007), but not in systematic reviews and meta-analysis of treatment effectiveness. Given n number of studies and individual study treatment effects θ_1,..., θ_n, a weighted kernel density estimator of treatment effects is defined as

$$\hat{f}(\theta) = \sum_{i=1}^{n} w(\theta_i) K_h(\theta - \theta_i) \quad (4)$$

where $K_h(.)$ is a kernel function and h is the bandwidth which controls the smoothness of the density estimate. Essentially, KDE is a continuous histogram whose blocks are centralized in each of the data points from where the density is estimated. The kernel function defines the shapes of the peaks of the observed data so that the estimator is the sum of the peaks. Properties of the kernel function K(u) partially determine the properties of KDE, such as differentiability and

continuity, so $K_h(.)$ is usually chosen to be a symmetric density function; w(.) is a re-weighting function which is used to control the roles of different θ_i. In the context of REMA, w(.) is formulated to incorporate study-level sampling errors σ_i and between-study heterogeneity τ^2, as the true variance of individual study effects is $\sigma_i^2 + \tau^2$. This is equivalent to weighting in standard meta-analysis as previously described.

The most important issue in KDE is the bandwidth selection. There have been several approaches to finding the optimal bandwidth. We consider the approach based on minimizing the mean integrated square error (Rosenblatt, 1956; Silverman, 1982) under which the approximate optimal bandwidth is given by

$$h = 0.9min(s_w, IQR_w/1.34)n^{-0.20} \tag{5}$$

for s_w = sample standard deviation and IQR_w = sample inter-quartile range. A drawback of using a fixed bandwidth is that where the data are dense will be masked or spurious noise will appear where the data are sparse. To account for data clustering in estimating the bandwidth, the adaptive kernel density estimates can be calculated using the following three step procedure (Silverman, 1986):

1. Find a pilot estimate $\hat{f}$ using the bandwidth in (5)
2. Define the local bandwidth factor $L_i = (\frac{g}{\hat{f}})^{\alpha}$, where log g = $\frac{log\hat{f}}{n}$
3. Define the adaptive kernel estimate $\hat{f}$ by

$$\hat{f} = \sum_{i=1}^{n} \frac{w(\theta_i)}{hL_i} K_h(\frac{\theta - \theta_i}{hL_i})$$

Literature has shown that the pilot estimate in Step 1 is not that crucial.

4 Example: Do community health worker interventions improve rates of screening mammography?

As part of a systematic review to study the effectiveness of community health workers in improving screening mammography rates (Wells et al., 2011), REMA was performed to calculate the overall intervention effect. Community health

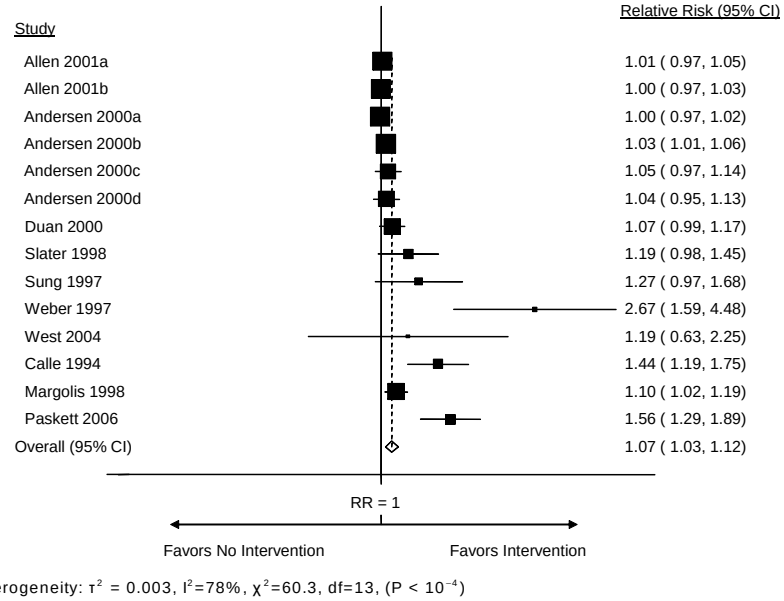

Figure 1: Forest plot of comparison: intervention versus no intervention for outcome receipt of mammography. Heterogeneity is presented under DerSimonian and Laird method

workers are individuals trained to act as intermediaries between the patients and their health care providers and services. The pooled relative risk of receiving screening mammography was based on 14,159 mammography events (7,107 in intervention and 7,052 in control group) in 10 randomized controlled trials (14 comparisons). Based on the results from DerSimonian-Laird REMA of 10 randomized controlled trials (Figure 1 forest plot), the authors concluded that the

intervention was effective and associated with statistically significant increase in the receipt of screening mammography (Relative Risk RR = 1.07; 95% CI: 1.03-1.12, P = 5 x 10^{-4}). There was a statistically significant between-study heterogeneity among included trials (Heterogeneity index I^2 = 78%, P $< 10^{-6}$).

We performed REMA under all five estimators for τ^2. The 95 percent confidence intervals, prediction intervals and probability coverage under KDE are given in Figure 2.

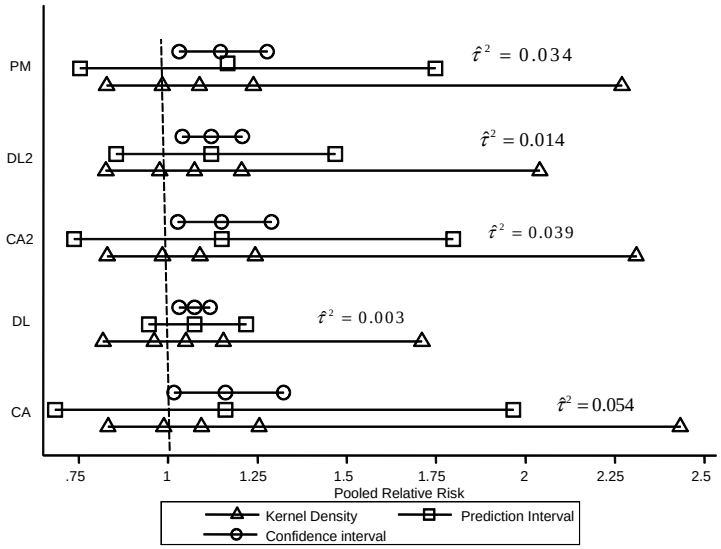

Figure 2: 95% prediction intervals, 95% confidence intervals, and 95% kernel density coverage under Cochrane ANOVA (CA), DerSimonian and Laird (DL), Cochrane ANOVA two-step (CA2), and DerSimonian and Laird two-step (DL2), and Paule and Mandel (PM) estimates of between-study variance $\hat{\tau}^2$.

All five confidence intervals were to the right of RR = 1 (point of no treatment difference), indicating the intervention was effective in this cohort of 14 comparisons under all five estimates $\hat{\tau}^2$; however all five prediction intervals contained RR = 1 indicating likelihood that the next intervention may not be effective. The KDE coverage was consistently the widest, correctly reflecting the uncertainty and between-study heterogeneity induced by the observed intervention effects in favor of community health workers effectiveness. The smallest estimated $\hat{\tau}^2$ was under the DerSimonian-Laird estimation. We note that the distribution of 14 observed treatment effects on the natural log scale was not normally distributed (Shapiro-Wilk P = 6 x 10^{-4}).

5 Conclusion

In this manuscript we presented how KDE can be implemented to compliment confidence interval and prediction interval estimation to account for uncertainty in observed treatment effects across studies in random-effects meta-analysis. In the absence of evidence of normality, a number of solutions have been proposed, including the t-distribution or skewed distributions in order to reduce the effect of outlier studies (Lee & Thompson, 2008), non-parametric maximum likelihood (Bohning, 2005) and Bayesian semi-parametric MCMC methods (Burr & Doss, 2005; Ohlssen et al., 2007). Published literature (Higgins et al., 2009) has rightly cautioned against the use of the empirical distribution of individual trial effects θ_i in the form of a histogram on the grounds that the results may be misleading due to the substantial differences among trial sampling errors. The weighted kernel density estimation as applied to REMA correctly gives consideration to the whole distribution of treatment effects without relying on the normality assumption for random effects.

REFERENCES

[1] D. G. Altman, K. F. Schulz, D. Moher, et al., *The revised CONSORT statement for reporting randomized trials: explanation and elaboration*, Ann Internal Medicine, 134 (2001) 663-694.

[2] D. Bohning. *Meta-analysis: a unifying meta-likelihood approach framing unobserved heterogeneity, study covariates, publication bias, and study quality*, Methods Inf Med, 44 (2005) 127-135.

[3] W. G. Cochrane, *The combination of estimates from different experiments* , Biometrics, 10 (1954) 101-129.

[4] D. Burr and H. Doss, *A Bayesian semi-parametric model for random-effects meta-analysis* , J Am Stat Assoc, 100 (2005) 242-251.

[5] R. DerSimonian and N. Laird, *Meta-analysis in clinical trials*, Control Clin Trials, 7 (1986) 177-188.

[6] R. DerSimonian and R. Kacker, *Random-effects model for meta-analysis of clinical trials: an update* , Contemp Clin Trials, 28 (2007) 105-114.

[7] B. Djulbegovic, A. Kumar, H. P. Soares, et al, *Treatment success in cancer: new cancer treatment successes identified in phase III randomized controlled trials conducted by the National Cancer Institute-sponsored cooperative oncology groups, 1955 to 2006* , Arch Internal Med, 168 (2008) 632-642.

[8] J. P. Higgins, S. G. Thompson, and D. J. Spiegehalter, *A re-evaluation of random-effects meta-analysis* , J Roy Stat Soc Ser A, 172 (2009) 137-159.

[9] J. P. Higgins and S. Green, *Cochrane handbook for systematic reviews of interventions* , John Wiley and Sons, 2008.

[10] R. N. Kacker and D. A. Harville, *Approximations for standard errors of estimators of fixed and random effects in mixed linear models*, J Am Stat Assoc, 79 (1984), 853-862.

[11] K. J. Lee and S. G. Thompson, *Flexible parametric models for random-effects distributions* , Stat Med, 27 (2008) 418-434.

[12] D. I. Ohlssen, L. D. Sharples, and D. J. Spiegehalter, *Flexible random-effects models using Bayesian semi-parametric models: applications to institutional comparisons* , Stat Med, 26 (2007) 2088-2112.

[13] R. C. Paule andJ. Mandel, *Consensus values and weighting factors* , J Res National Institute of Standards and Technology, 87 (1982) 377-385.

[14] M. Rosenblatt, *Remarks on some nonparametric estimates of a density function* , Ann Math Stat, 27 (1956) 832-837.

[15] B. W. Silverman, *Density estimation for statistics and data analysis* , Chapman and Hall, London, 1986.

[16] B. W. Silverman, *Kernel density estimation using the fast Fourier transform* , J Roy Stat Soc Ser C, 31 (1982) 93-99.

[17] H. C. Sox, *Defining comparative effectiveness research: the importance of getting it right* , Med Care, 48 (2010) S7-8.

[18] T. D. Wager, M. Lindquist, and L. Kaplan, *Meta-analysis of functional neuroimaging data; current and future directions*, Soc Cogn Affect Neurosci, 2 (2007), 150-158.

[19] T. D. Wager, K. L. Phan, I. Liberzon, and S. F. Taylor, *Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging* , Neuroimage, 19 (2003) 513-531.

[20] K. J. Wells, J. S. Luque, B. Miladinovic, et al., *Do community health worker interventions improve rates of screeing mammography in the United States? A systematic review* , Cancer Epidemiol Biomarkers Prev, 20 (2011) 1580-1598.