Metrologia **57** (2020) 064004 (13pp) https://doi.org/10.1088/1681-7575/abb064

Bayesian inference for heterogeneity in meta-analysis

Olha Bodnar[1], **Rebecca Nalule Muhumuza**[1,2,3] **and Antonio Possolo**[3]

[1] Unit of Statistics, School of Business, Örebro University, Örebro, Sweden
[2] Department of Mathematics, Busitema University, Box 236, Tororo, Uganda
[3] Department of Mathematics, Makerere University, Box 7062 Kampala, Uganda
[4] National Institute of Standards and Technology, Gaithersburg, Maryland, United States of America

E-mail: olha.bodnar@oru.se, beckynalule@yahoo.com and antonio.possolo@nist.gov

Received 2 June 2020, revised 7 August 2020
Accepted for publication 18 August 2020
Published 30 October 2020

Abstract

A generalized marginal random effects model is described that enables exact Bayesian inference using either the Jeffreys or Berger-Bernardo non-informative prior distributions without the need for Markov Chain Monte Carlo sampling, requiring only numerical integrations. This contribution focuses on inference for the heterogeneity parameter, often called "dark uncertainty" and denoted τ in this contribution. The proposed models are used for consensus building in meta-analyses of measurement results for the Newtonian constant of gravitation, G, and for the effectiveness of anti-retroviral pre-exposure prophylaxis in preventing HIV infection.

The estimates of τ that seventeen alternative different methods produce, including those that we propose, were also compared. The relative impact (gauged in terms of the ratio of the range of estimates to their median) that model choice had on the estimate of τ was very substantial: 79 % for G and 87 % for HIV prophylaxis. For G, the estimates of τ ranged from $0.0009 \times 10^{-10} m^3 kg^{-1} s^{-2}$ to $0.0013 \times 10^{-10} m^3 kg^{-1} s^{-2}$. For *Truvada* they ranged from 0.49 to 0.92.

Since the estimate of τ impacts the quality of the estimate of the measurand substantially, we recommend the Bayesian approaches to estimate τ because they take the whole posterior distribution of τ into account, hence the corresponding uncertainty, rather than using a single value and pretending that it is known with certainty.

In the case of the measurement results for G, we found that the model with Student's t random effects and the Jeffreys prior provides the best fit, while for the *Truvada* data the normal marginal random effects model, also with Jeffreys prior, produces an estimate of τ closest to classical estimators like DerSimonian-Laird, but offers the advantage of recognizing and propagating the uncertainty associated with τ, which the classical procedures ignore.

Keywords: marginal distribution, random effects, dark uncertainty, non-informative prior, shades of dark uncertainty, Jeffreys, Berger-Bernardo

(Some figures may appear in colour only in the online journal)

1. Introduction

Random effects meta-analysis is a widely used statistical tool to combine the results of clinical studies [1–7], and for the reduction of data from interlaboratory studies in metrology [8–10]. Models of this kind have also been used to evaluate the uncertainty associated with fundamental physical constants [11–14], and to estimate laboratory biases [15]. [16] presented an extension of the random effects model to key comparisons where two reference standards are measured by two different sets of national metrology institutes having at least one institute in common.

In its most common application, random effects meta analysis aims to draw inferences about the common mean μ of the

