Meta-analysis with application to estimating combined estimators of effect sizes in biomedical research.

UNIVERSITY OF
KWAZULU - NATAL

INYUVESI
YAKWAZULU-NATALI

Mhlengi Corrigan Mgaga

November, 2018

Meta-analysis with application to estimating combined estimators of effect sizes in biomedical research.

by

Mhlengi Corrigan Mgaga

A thesis submitted to the

University of KwaZulu-Natal

in fulfilment of the requirements for the degree

of

MASTER OF SCIENCE

in

STATISTICS

Thesis Supervisor: Prof Henry Mwambi

Thesis Co-supervisor: Prof Ding-Geng Chen

UNIVERSITY OF KWAZULU-NATAL

SCHOOL OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE

PIETERMARITZBURG CAMPUS, SOUTH AFRICA

Declaration - Plagiarism

I, Mhlengi Corrigan Mgaga, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowlegded as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then
   (a) their words have been re-written but the general information attributed to them has been referenced, or
   (b) where their exact words have been used, then their writing has been placed in italics and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the reference sections.

Mhlengi Corrigan Mgaga (Student)	Date
Prof Henry Mwambi (Supervisor)	Date
Prof Ding-Geng Chen (Co-supervisor)	Date

Disclaimer

This document describes work undertaken as a Masters programme of study at the University of KwaZulu-Natal (UKZN). All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institution.

Abstract

Meta-analysis is a statistical analysis that combines results from different independent studies. In meta-analysis a number of statistical methods are currently used for combining effect sizes of different studies. The simplest of these methods is based on a fixed-effects model, which assumes that all studies in the meta-analysis share a common true effect size and that the effect sizes in our meta-analysis differ only because of sampling error. Another statistical method that is used in meta-analysis, is the random-effects model, which assumes sampling variation due to fixed-effects model assumptions and random variation because the effect sizes themselves are sampled from a population of effect sizes. These models are compared to determine which model is appropriate and under what circumstances is the model appropriate. We illustrate these models by applying each model to a collection of 3 studies examining the effectiveness of new drug versus placebo to treat patients with duodenal ulcers and meta-analysis of 9 studies of the use of diuretics during pregnancy to prevent the development of pre-eclampsia. Results indicated that the choice between the two model depends on the question of which model fits the distribution of effect sizes better and takes account of the relevant source(s) of error. We further study the meta-analysis of longitudinal studies where effect sizes are reported at multiple time points. Univariate meta-analysis is a statistical approach which may be used to study effect sizes reported at multiple time point. The problem with this approach is that it ignores correlation between the effect sizes, which might increase the standard error of the point estimates. We used the linear mixed-effects model, which borrows ideas from multivariate meta-analysis. One of the advantages of the linear mixed-effects model is that it accounts for correlation between effect sizes both within and between studies. The independence model where separate univariate meta-analysis is done at each of the time points was compared against models where correlation was accounted for different alternatives; including random study effects, correlated random time effects and/or correlated within-study errors, or unstructured covariance structures. We implemented these methods through an example of meta-analysis of 16 randomized clinical trials of radiotherapy and chemotherapy versus radiotherapy alone for the post-operative treatment of patients with malig-

nant gliomas, where in each trial, survival is evaluated at 6, 12, 18 and 24 months post randomization. The results revealed that models that accounted for correlations had better fit.

Keywords: meta-analysis, fixed-effects model, random-effects model, heterogeneity, publication bias, linear mixed-effects model.

Acknowledgements

Firstly, I would like to thank my supervisor Prof. Henry Mwambi for his presence, guidance and invaluable advice throughout this research. I again thank him for reading and correcting my errors in this research. Secondly I would like to thank my co-supervisor Prof. Ding-Geng Chen for willing to work with me. Much thanks for his guidance, help and responses whenever I consulted. Many thank to him for correcting my research.

I am grateful for the scholarship awarded to me by DELTAS Africa, without which I would have been unable to support myself during the year of my research. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Co-ordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [grant 107754/Z/15/Z-DELTAS Africa Sub-Saharan Africa Consortium for Advanced Bio-statistics (SSACAB) programme] and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government.

Finally, I wish to express my appreciation to my family and friends, for their support and encouragement. I may not count every one but I am grateful to everyone who helped me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Meta-analysis is a statistical analysis that combines the findings from different independent studies. The term meta-analysis is given to retrospective investigations in which data from all know studies of a particular clinical issue are assembled and evaluated collectively and quantitatively [1]. The objective of meta-analysis or other methods of quantitative research synthesis is the use of data from a series of studies to obtain information about the effect size for a treatment on various constructs. It is most often applied to treatment effects in randomized clinical trials since meta-analysis provides an objective way of combining information from independent studies looking at the same clinical questions [1]. In medical research it is becoming increasingly popular as a result of the information on efficacy of a treatment that is available from a number of clinical studies with similar treatment protocols [2]. The advantage of combining the findings across such studies represent an attractive alternative to strengthen the evidence about the treatment efficacy [2]. Moreover combing the results of several studies through the techniques of meta-analysis can provide stronger evidence for or against a treatment effect than one can derive from any of the individual studies because it produces a more precise estimate of the effect (i.e an estimate with smaller standard error or a narrower confidence interval) [3]. Meta-analysis is most often used to determine the effectiveness of clinical interventions, providing an estimate of the treatment effect while taking into consideration the weight of individual studies [4]. For instance, the aim for health professionals is to use the best treatment for their patients. But it is crucial to identify the most effective and the least harmful available interventions [4]. Hence meta-analysis inform clinical practice guidelines that make treatment recommendations based on evidence about the least harmful and most effective interventions [4]. In addition, today it is evident from the Cochrane collaboration and the high volume of publications for meta-analysis that clinical decision-making rely heavily on meta-analysis [5]. This statistical method provides both the clinician and the medical investigator

with quantitative summaries of the results of several studies, usually of evaluated therapies or of diagnostics methods [3]. The results strengthen our knowledge beyond what can be contributed even by multiple single studies and may guide diagnosis and treatment of patients as well as suggest directions for future research [3]. In all areas of clinical practice, meta-analysis of well designed and executed randomized controlled trials have the potential to provide high levels of evidence to support therapeutic interventions [5]. Regardless of the potential of the outcome of such trials to guide decision making, they may sometimes fail to produce credible conclusive results or may disagree with multiple independent trial that investigate the same clinical question [5]. In such cases, a meta-analysis of the trial results has the potential to combine conceptually similar and independent studies with the purpose of deriving more reliable statistical conclusions (based on a much larger sample data) than any of the individual studies [5]. Meta-analysis of systematically searched randomized controlled trials are often ranked as the highest available category of evidence because systematic review is theoretically less susceptible to bias [4]. Meta-analysis cannot prevent bias as such, but it can enhance the precision of the estimated treatment effects and reduce the probability of false negative results [4].

Meta-analysis usually involves obtaining an estimate of effect sizes from each study and averaging these estimates to obtain an estimate of the average effect size across studies [6]. In addition, the researcher may be interested in finding whether any characteristics of the studies are systematically related to effect size. Some writers in this research area have cited reasons for the position that different studies of the same treatment might yield quite different result [7]. Grouping studies with similar characteristics may resolve many contradictions in research evidence, Light and Smith argued [7]. They point out that studies with the same characteristic are more likely to yield similar results and many contradictions among research results arise from differences in the characteristics of studies. Pillemer and Light [8] have argued that study characteristic is an important step in assessing the range of generalizations of a research finding and examining the relationship of variations in study outcomes. For example, if a treatment produces essentially the same effect in a wider variety of settings with a variety of people, we are more confident in the generalization of finding of a treatment effect related to effect size.

1.0.1 Effect size

Effect size is defined as the strength of the relationship between the independent and dependent variables [9]. Meta-analyses that deal with medical intervention of-

ten refer to effect size as a treatment effect, and this term is sometimes assumed to refer to odds ratio, relative risk and risk difference [9]. The most challenging step in conducting a meta-analysis is extracting effect sizes from primary research reports. The reports from the studies fail to provide sufficient information for computing effect sizes, as they mostly include the results of significance tests of probability values other than those needed by the meta-analyst. Moreover, the same question can be addressed using different experimental designs of a series of primary research studies. Despite the fact that it is not commonly recognized, effect sizes from different experimental designs often estimate different population parameters [10]. If the adjustments for the designs are made they can be directly compared otherwise they cannot be directly compared [11, 12]. The studies with repeated measures also consist of independent groups (experimental arms) and the problem is thus the combination of results of studies with independent groups but with cross-sectional versus repeated measures design. Almost in every treatment of meta-analysis, the calculation of effect sizes is straight forward when the research depends entirely on independent groups design [13, 14]. Nevertheless, when conducting meta-analysis, in many research areas both repeated measures and independent groups designs. The studies involved often come from different representative populations. It is important to note that this is a general problem in meta-analysis and is not specific to both independent groups and repeated measures design. Unless a set of studies consist of perfect replications, differences in design may result in studies that do not estimate the same population effect size [15]. The only way that the effect sizes can be combined is when these studies provide estimates of the same population parameter, on other hand the effect sizes from different design should not be combined, since it will estimate different parameters [15]. From different research studies, statistical methods have been used to combine information. The work on combining the results of agricultural experiments is one of the examples of this work, other examples can be found in biomedical research and other fields [16]. Cochran developed a more defined weighting methods for combining effect sizes for agricultural experiments [17].

Glass proposed the use of statistical methods in research reviews and from that moment there has been intense interest in quantitative methods for research synthesis. The method that Glass proposed involves the use of effect sizes [6]. The estimates of the effect sizes across all studies are calculated using the proposed method. The average of the effect sizes across studies is used as a treatment effect of the overall effect size across studies [18, 19]. The statistical theory for estimating the effect size was addressed by Hedges [18, 19], He also derived the sampling distribution of effect size estimators and show how to construct the confidence interval for the effect sizes when a series of studies share a common population effect size [20]. An

effect size refers to the magnitude of the effect observed in a study, depending on the size of a relationship between variables [21]. In meta-analysis the basic principle is to calculate the effect size of the individual studies and combine them to obtain average effect size [19].

The general strategy that was recommended by Glass was that of coding the characteristic of studies as a vector of predictor variables and regressing the effect size estimates on these predictors to determine the relationship between characteristics of studies and their effect size [22]. Smith and Glass [23] in their meta-analysis of psychotherapy outcome studies, used ordinary linear regression to determine the relationship between characteristics of studies (e.g.,type of therapy, duration of treatment, interval validity of the study) and effect size [23]. Hail [24] used the same method in quantitative syntheses of gender effects in decoding nonverbal cues. Uguroglu [25] studied the relationship between academic achievement and motivation and the effect of goal structures on achievement. Johnson [26] studied the relations between study characteristics and effect size that are found to be consistent in these analyses. One explanation of these relations derived from the proposal by Cronbach [27]. Cronbach [27] suggested that evaluation studies should consider a model in which each treatment site is a sample realization from a universe of related treatments. Thus when evaluators look at replications of a treatment across sites, they observe many differences, each sampled from some universe of possible treatments [27]. Variations in the true population effect of the treatment would be expected and these variations in treatment are more or less effective in producing the outcome [27]. The relationship between a fixed characteristic (such as age or sex of subject) and the outcome variable might be expected to be attenuated by such variations [27]. Note that this model implies that there is no single true or population effect of the treatment across studies. Rather, there is a distribution of true effect. Each treatment site has its own unique true effect [27]. One may speak of the average true effect of the treatment as an index of overall efficacy. Without some measure of the variation in the true effect of the treatment, the average true effects is not very meaningful [27]. One might find, for example, that the average of the true effect was larger than zero, but the true effect of the treatment was negative in nearly half the implementations. The fact that the true effect size is not known, the problem of estimating the variability in the true effects is complicated. We must estimate the true effect from sample data, and that estimate will itself be subject to sampling changes [27]. The underlying population effect sizes will not be constant across a series of studies that replicate the same treatment. It is these random-effects models that leads to quantitative research synthesis. The test of homogeneity of effect sizes was developed by Hedges [19]. Hedge's procedure tests whether the observed

estimates of effect size vary by more than expected if all studies shared a common population effect size. This procedure was used by Giaconia and Hedges in one quantitative research synthesis [28].

1.0.2 Objectives of the study

The main objectives of the study is to understand methodologies in meta-analysis for both cross-sectional and longitudinal studies. The specific objectives are to:

- Investigate and understand methods for combining effect sizes from stratified analyses.
- Investigate and understand methods for cross-sectional meta-analysis.
- Investigate and understand methods for longitudinal meta-analysis.
- Demonstrate the understanding of these methods by applying them into real data.

1.0.3 Outline of the study

The thesis is organized as follows: Chapter 2 is about an overview of the basic statistical concepts and approaches in meta-analysis. In Chapter 3, we discuss methods for meta-analysis of longitudinal studies. In Chapter 4, we illustrate how to apply these statistical methods in practice and discuss the results found. In Chapter 5, we present a general conclusion for the overall thesis, after which the References and Appendices are presented.

Chapter 2

Overview of basic statistical concepts and approaches in meta-analysis

2.1 Preliminary concepts

In this section we showed how to compute the variance for specific effect sizes such as risk difference, log relative risk and log odds ratio. We discussed the important of these effect sizes in meta-analysis. Moreover, we highlighted reasons for the selection of effect size in the analysis.

2.2 Effect sizes based on binary data (2 $\times$ 2 tables)

Consider L independent studies yielding 2×2 tables of the form (a, b, c, d). Where $(a$,$b)$ and $(c$,$d)$ are the number of positives and negatives which are measured for treatment group and control group respectively. Let π_1 denote the probability of the event in the treatment group, which is estimated as $\hat{\pi}_1 = \frac{a}{n_1}$ and π_2 denote the probability of the event in the control group, which is estimated as $\hat{\pi}_2 = \frac{b}{n_2}$. The null hypothesis is that $H_0 : \pi_1 = \pi_2$ and the alternative hypothesis is $H_1 : \pi_1 \neq \pi_2$. Define the marginal totals by $n_1 = a + c$, $n_2 = b + d$, $m_1 = a + b$, $m_2 = c + d$ and the grand total by $N = n_1 + n_2 = m_1 + m_2$. We define the table as follows

Table 2.1: A table of follow up randomisation clinical trial

	Treatment Group	Control Group	
Positive	a	b	m_1
Negative	c	d	m_2
Total	n_1	n_2	N

2.2.1 Risk difference

The risk difference (RD) refers to the simple algebraic difference between the probabilities of the positive response in the two groups with a domain of [-1,1]. The asymptotic distribution of the risk difference $\widehat{RD} = p_1 - p_2$ follows directly from that of the sample proportions themselves, where $p_1 = \frac{a}{n_1}$ and $p_2 = \frac{b}{n_2}$. Since the $\widehat{RD}$ is a simple linear contrast of two independent proportions, of which each is asymptotically normally distributed, thus $RD = \pi_1 - \pi_2$. We shall derive the risk difference through multivariate δ-method, starting with the asymptotic bivariate distribution of p_1 and p_2. It is because the two groups are independent, then $\mathbf{p} = (p_1, p_2)'$ is asymptotically distributed as bivariate normal with mean vector $\boldsymbol{\pi} = (\pi_1, \pi_2)'$ and covariance matrix under the alternative hypothesis as

$$\boldsymbol{\Sigma_y} = \begin{bmatrix} \frac{\pi_1(1-\pi_1)}{n_1} & 0 \\ 0 & \frac{\pi_2(1-\pi_2)}{n_2} \end{bmatrix}. \tag{2.1}$$

The risk difference is $G(\boldsymbol{\pi}) = \pi_1 - \pi_2$, with the corresponding matrix of partial derivatives

$$\mathbf{H}(\boldsymbol{\pi}) = \begin{bmatrix} \frac{\partial G(\boldsymbol{\pi})}{\partial \pi_1} \\ \frac{\partial G(\boldsymbol{\pi})}{\partial \pi_2} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{2.2}$$

or

$$\mathbf{H}(\boldsymbol{\pi}) = \begin{bmatrix} 1 & -1 \end{bmatrix}'. \tag{2.3}$$

Applying the multivariate δ-method, the asymptotic variance of the $\widehat{RD}$ under the

alternative hypothesis is given by

$$\begin{aligned}\sigma_1^2 = \mathrm{V}[\widehat{RD}] &\cong \mathbf{H}(\boldsymbol{\pi})'\boldsymbol{\Sigma}_{\boldsymbol{y}}\mathbf{H}(\boldsymbol{\pi}) \\ &= \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{\pi_1(1-\pi_1)}{n_1} & 0 \\ 0 & \frac{\pi_2(1-\pi_2)}{n_2} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{\pi_1(1-\pi_1)}{n_1} & \frac{-\pi_2(1-\pi_2)}{n_2} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ &= \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}.\end{aligned} \tag{2.4}$$

Since $p_1 = \frac{a}{n_1}$ and $p_2 = \frac{b}{n_2}$, the asymptotic variance can be consistently estimated as

$$\begin{aligned}\hat{\sigma}_1^2 = \widehat{V}(\widehat{RD}) &= \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} \\ &= \frac{ac}{(a+c)^3} + \frac{bd}{(b+d)^3}.\end{aligned} \tag{2.5}$$

The asymptotic distribution under the alternative leads to the usual expression for the large sample $1-\alpha$ level confidence interval for the population risk difference based on the estimate of the variance under the alternative

$$(\hat{\theta}_l, \hat{\theta}_u) = \hat{\theta} \pm Z_{1-\alpha/2}\hat{\sigma}_1. \tag{2.6}$$

Although it is common in practice, these confidence limits are not necessarily bounded by -1 and 1, and in rare circumstances limits are obtained outside these bounds [29]. Unlike the case of a single proportion, there is no convenient function that may be used to yield asymmetric confidence limits on the risk difference that are then bounded by (-1,1).

2.2.2 Relative risk

The relative risk (RR) is the ratio of the risk probabilities of the two groups. It has a domain consisting of the positive real line and a null value of one. To provide a symmetric distribution under the null hypothesis it is customary to use the log of each. We consider the distribution of the $\log(\widehat{RR})$, which is the difference in the logs of the two probabilities. Then the relative risk is expressed as

$$\mathrm{RR} = \frac{\pi_1}{\pi_2} \tag{2.7}$$

and the log relative risk as

$$\log(RR) = \log(\pi_1) - \log(\pi_2). \tag{2.8}$$

Now, we shall derive the distribution of the log relative risk through multivariate δ-method, starting with the asymptotic bivariate distribution of p_1 and p_2. If the two groups are independent, then $\mathrm{p} = (p_1, p_2)'$ is asymptotically distributed as bivariate normal with mean vector $\boldsymbol{\pi} = (\pi_1, \pi_2)'$ and covariance matrix under the alternative hypothesis as

$$\boldsymbol{\Sigma_y} = \begin{bmatrix} \frac{\pi_1(1-\pi_1)}{n_1} & 0 \\ 0 & \frac{\pi_2(1-\pi_2)}{n_2} \end{bmatrix}. \tag{2.9}$$

The log relative risk is $G(\boldsymbol{\pi}) = \log(\pi_1) - \log(\pi_2)$, with the corresponding matrix of partial derivatives

$$\mathbf{H}(\boldsymbol{\pi}) = \begin{bmatrix} \frac{\partial G(\boldsymbol{\pi})}{\partial \pi_1} \\ \frac{\partial G(\boldsymbol{\pi})}{\partial \pi_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\pi_1} \\ \frac{-1}{\pi_2} \end{bmatrix} \tag{2.10}$$

or

$$\mathbf{H}(\boldsymbol{\pi}) = \begin{bmatrix} \frac{1}{\pi_1} & \frac{-1}{\pi_2} \end{bmatrix}'. \tag{2.11}$$

Applying the multivariate δ-method, the asymptotic variance of the $\log(\widehat{RR})$ under the alternative hypothesis is given by

$$\begin{aligned} \sigma_1^2 = \mathrm{V}[\log\widehat{RR}] &\cong \mathbf{H}(\boldsymbol{\pi})'\boldsymbol{\Sigma_y}\mathbf{H}(\boldsymbol{\pi}) \\ &= \mathbf{H}(\boldsymbol{\pi}) = \begin{bmatrix} \frac{1}{\pi_1} & \frac{-1}{\pi_2} \end{bmatrix} \begin{bmatrix} \frac{\pi_1(1-\pi_1)}{n_1} & 0 \\ 0 & \frac{\pi_2(1-\pi_2)}{n_2} \end{bmatrix} \begin{bmatrix} \frac{1}{\pi_1} \\ \frac{-1}{\pi_2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{(1-\pi_1)}{n_1} & \frac{-(1-\pi_2)}{n_2} \end{bmatrix} \begin{bmatrix} \frac{1}{\pi_1} \\ \frac{-1}{\pi_2} \end{bmatrix} \\ &= \frac{(1-\pi_1)}{n_1\pi_1} + \frac{(1-\pi_2)}{n_2\pi_2}. \end{aligned} \tag{2.12}$$

Since $p_1 = \frac{a}{n_1}$ and $p_2 = \frac{b}{n_2}$, the asymptotic variance can be consistently estimated as

$$\begin{aligned}\hat{\sigma}_1^2 = \widehat{V}[\log(\widehat{RR})] &= \frac{(1-p_1)}{n_1 p_1} + \frac{(1-p_2)}{n_2 p_2} \\ &= \frac{(1-p_1)}{a} + \frac{(1-p_2)}{b} \\ &= \frac{1}{a} - \frac{1}{n_1} + \frac{1}{b} - \frac{1}{n_2}.\end{aligned} \tag{2.13}$$

Further, asymptotically

$$\log(\widehat{RR}) \approx N[\log(RR), V[\log(\widehat{RR})]] \tag{2.14}$$

hence,

$$\frac{\log(\widehat{RR}) - \log(RR)}{\sqrt{\widehat{V}[\log(\widehat{RR})]}} \approx N(0,1). \tag{2.15}$$

This distribution under the alternative hypothesis is used to derive the large sample confidence limits on $\theta = \log(RR)$ as

$$(\hat{\theta}_l, \hat{\theta}_u) = \hat{\theta} \pm Z_{1-\alpha/2}\hat{\sigma}_1. \tag{2.16}$$

The asymmetric confidence limits for relative risk are obtained as

$$(\widehat{RR}_l, \widehat{RR}_u) = \exp[\hat{\theta} \pm Z_{1-\alpha/2}\hat{\sigma}_1] = \exp(\hat{\theta}_l, \hat{\theta}_u). \tag{2.17}$$

that are contained within $[0, \infty)$.

2.2.3 Odds ratio

The odds ratio (OR) is the ratio of the odds of the outcome of interest in the two groups. Which is given as

$$\text{OR} = \frac{\frac{\pi_1}{1-\pi_1}}{\frac{\pi_2}{1-\pi_2}} \tag{2.18}$$

and the log odds ratio is obtained as follows,

$$\begin{aligned}\log(\text{OR}) &= \log\left(\frac{\frac{\pi_1}{1-\pi_1}}{\frac{\pi_2}{1-\pi_2}}\right) \\ &= \log\left(\frac{\pi_1}{1-\pi_1}\right) - \log\left(\frac{\pi_2}{1-\pi_2}\right).\end{aligned} \tag{2.19}$$

The distribution of the log odds ratio is obtained as that of a linear combination of two normally distributed covariates. Within each group, the log odds is simply the logit of the probability. In the following, however, we shall derive the distribution of the log odds ratio through multivariate δ-method, starting with the asymptotic bivariate distribution of p_1 and p_2.

$$\boldsymbol{\Sigma_y} = \begin{bmatrix} \frac{\pi_1(1-\pi_1)}{n_1} & 0 \\ 0 & \frac{\pi_2(1-\pi_2)}{n_2} \end{bmatrix}. \tag{2.20}$$

The log odds ratio is $G(\boldsymbol{\pi}) = \log\left(\frac{\frac{\pi_1}{1-\pi_1}}{\frac{\pi_2}{1-\pi_2}}\right) = \log\left(\frac{\pi_1}{1-\pi_1}\right) - \log\left(\frac{\pi_2}{1-\pi_2}\right)$, with the corresponding matrix of partial derivatives

$$\mathbf{H}(\boldsymbol{\pi}) = \begin{bmatrix} \frac{\partial G(\boldsymbol{\pi})}{\partial \pi_1} \\ \frac{\partial G(\boldsymbol{\pi})}{\partial \pi_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} \\ \frac{-1}{\pi_2(1-\pi_2)} \end{bmatrix} \tag{2.21}$$

or

$$\mathbf{H}(\boldsymbol{\pi}) = \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} & , & \frac{-1}{\pi_2(1-\pi_2)} \end{bmatrix}'. \tag{2.22}$$

Applying the multivariate δ-method, the asymptotic variance of the $\log(\widehat{OR})$ under the alternative hypothesis is obtained as follows

$$\begin{aligned}\sigma_1^2 = \text{V}[\log\widehat{OR}] &\cong \mathbf{H}(\boldsymbol{\pi})'\boldsymbol{\Sigma_y}\mathbf{H}(\boldsymbol{\pi}) \\ &= \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} & \frac{-1}{\pi_2(1-\pi_2)} \end{bmatrix} \begin{bmatrix} \frac{\pi_1(1-\pi_1)}{n_1} & 0 \\ 0 & \frac{\pi_2(1-\pi_2)}{n_2} \end{bmatrix} \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} \\ \frac{-1}{\pi_2(1-\pi_2)} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{n_1} & \frac{-1}{n_2} \end{bmatrix} \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} \\ \frac{-1}{\pi_2(1-\pi_2)} \end{bmatrix} \\ &= \frac{1}{n_1\pi_1(1-\pi_1)} + \frac{1}{n_2\pi_2(1-\pi_2)}.\end{aligned} \tag{2.23}$$

Since $p_1 = \frac{a}{n_1}$ and $p_2 = \frac{b}{n_2}$, the asymptotic variance can be consistently estimated as,

$$\begin{aligned}\hat{\sigma}_1^2 = \widehat{V}[\log(\widehat{OR})] &= \frac{1}{n_1 p_1(1-p_1)} + \frac{1}{n_2 p_2(1-p_2)} \\ &= \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}.\end{aligned} \tag{2.24}$$

This is Woolf's [30] estimate of the variance of the log odds ratio. From Slutsky's theorem (Appendix A.2) it follows that asymptotically

$$\log(\widehat{OR}) \approx N(\log(OR), \sigma_1^2) \tag{2.25}$$

and that,

$$\frac{\log(\widehat{OR}) - \log(OR)}{\sqrt{\widehat{V}[\log(\widehat{OR})]}} \approx N(0,1). \tag{2.26}$$

This yields large sample confidence limits on $\theta = \log(OR)$ as

$$(\hat{\theta}_l, \hat{\theta}_u) = \hat{\theta} \pm Z_{1-\alpha/2}\hat{\sigma}_1 \tag{2.27}$$

and asymmetric confidence interval limits for the odds ratio are expressed as follows,

$$(\widehat{OR}_l, \widehat{OR}_u) = \exp[\hat{\theta} \pm Z_{1-\alpha/2}\hat{\sigma}_1] = \exp(\hat{\theta}_l, \hat{\theta}_u) \tag{2.28}$$

and these are again bounded by -1 and 1. The odds ratio is a very popular measure of treatment effect for meta-analysis. The advantage of the odds ratio is that, it is valid regardless of the type of sampling used, which is not the case for other comparative measures for binary data. The choice of an effect size index in the analysis depend on the following considerations. Firstly, the effect size from the different studies should be comparable to one another in the sense that they measure atleast approximately the same thing. That is, the effect size should not depend on aspects of study design that may vary from study to study (such as sample size or whether covariates are used). Secondly, the effect size should be computable from the information that is likely to be reported in published research reports. That is, it should not require the re-analysis of raw data (unless these are known to be available). Finally, the effect size should have good technical properties. For example, its distribution should be known so that variances and confidence intervals can be computed.

2.3 Fixed-effects model

In this section we introduce the fixed-effects model. We discuss the assumptions of this model and show how these are reflected in the formulas used to compute a summary effect, and in the meaning of the summary effect.

2.3.1 Model description

This approach provides an adjusted estimator that is a minimum variance linear estimator (MVLE) of an unknown parameter θ as a measure of association on any scale $\theta = G(\pi_1, \pi_2)$ for some smooth function $G(\cdot, \cdot)$. Where π_1 and π_2 denote the probability of the event in the treatment group and control group respectively. Thus, within the class of linear estimators these estimates are asymptotically efficient. Since the estimates of θ within each study is consistent then the MVLE is also a consistent estimator with asymptotic variance. Using the framework of weighted least squares, we have a vector of random variables $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1 \cdots \hat{\theta}_L)'$, where the assumed model specifies that a common θ applies to all the studies such that $E(\hat{\theta}_i) = \theta$ for $i = 1, \cdots, L$. Furthermore, the variance of the estimate within the i^{th} study is $V(\hat{\theta}_i) = E(\hat{\theta}_i - \theta)^2 = \sigma_i^2$ for each measure of association such as the risk difference, log relative risk and log odds ratio. For now, assume that the $\{\sigma_i^2\}$ are fixed. Note that we are not assuming a common variance across the L strata. In other words we have a case of *heteroscedasticity*, that is different variances for the different levels of the stratifying variable(s). Under this model

$$\hat{\boldsymbol{\theta}} \cong \begin{bmatrix} 1 \\ . \\ . \\ . \\ 1 \end{bmatrix} \theta + \epsilon = \boldsymbol{J}\theta + \epsilon, \tag{2.29}$$

where $\boldsymbol{J}$ is a $L \times 1$ unit vector of ones, and ϵ is a $L \times 1$ vector where

$$E(\epsilon) = \mathbf{0}, \quad V(\epsilon) = \text{diag}(\sigma_1^2, \cdots, \sigma_L^2) = \boldsymbol{\Sigma_\epsilon}. \tag{2.30}$$

Since the estimates $\hat{\theta}_i$ of each strata provides the *consistent estimator* that converges in probability to the assumed common parameter θ for all i, then the weighted least squares (WLS) estimate of the common parameter is given as

$$\hat{\theta} = (\boldsymbol{J'}{\boldsymbol{\Sigma_\epsilon}}^{-1}\boldsymbol{J})^{-1}(\boldsymbol{J'}{\boldsymbol{\Sigma_\epsilon}}^{-1}\hat{\boldsymbol{\theta}}). \tag{2.31}$$

Since $\boldsymbol{\Sigma_\epsilon}^{-1} = \text{diag}(\sigma_1^{-2}, \cdots, \sigma_L^{-2})$, the estimator in (2.31) can be expressed as a weighted average of the stratum-specific estimate where

$$\hat{\theta} = \frac{\sum_i \sigma_i^{-2}\hat{\theta}_i}{\sum_i \sigma_i^{-2}} = \sum_i^L \omega_i \hat{\theta}_i, \tag{2.32}$$

also

$$\omega_i = \frac{\sigma_i^{-2}}{\sum_l \sigma_l^{-2}} = \frac{\tau_i}{\sum_l \tau_l} \tag{2.33}$$

$\tau_i = \sigma_i^{-2}$and$\sum_i \omega_i = 1$. Moreover the variance of the estimate is

$$V(\hat{\theta}) = \sigma^2 = (\boldsymbol{J'\Sigma_\epsilon}^{-1}\boldsymbol{J})^{-1} = \frac{1}{\sum_i \sigma_i^{-2}}. \tag{2.34}$$

To compute the estimate we use the estimated weight $\{\hat{\omega}_i^2\}$, obtained by substituting the large estimate of the stratum-specific variance in equation(2.32) $\{\hat{\sigma}_i^2\}$ so that

$$\hat{\theta} = \sum_i \hat{\omega}_i \hat{\theta}_i. \tag{2.35}$$

To compute the estimate of the variance is obtained by substituting the stratum-specific variance estimate $\{\sigma_i^2\}$ into equation(2.34) so that

$$\hat{V}(\hat{\theta}) = \frac{1}{\sum_i \hat{\sigma}_i^{-2}}. \tag{2.36}$$

In summary, the *minimum variance linear estimator* is also known as the *fixed-effects model*, since we assume that there is a common value θ for all the strata. Also we assume that all the variation between the values of the observed parameter $\hat{\theta}_i$ is caused by the random sampling variation about a common value θ.

2.4 Multivariate Test of Hypotheses

2.4.1 Multivariate Null Hypothesis

Consider a stratified analysis with L strata. Then, we wish to conduct a test for a vector of L random variables $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1 \cdots \hat{\theta}_L)'$. The stratum-specific estimates θ_i can be measured in any scale function such as $\theta_i = G(\pi_{1i}, \pi_{2i})$ for some smooth

function $\theta_i = G(\cdot, \cdot)$. Where π_{1i} and π_{2i} denote the probability of the event in the i^{th} stratum of the treatment group and control group respectively. The vector of L random variables is assumed to be normally distributed

$$\hat{\boldsymbol{\theta}} \backsim N_L(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}), \tag{2.37}$$

where the mean and the variance are $E(\hat{\boldsymbol{\theta}}) = \boldsymbol{\theta}$ and $V(\hat{\boldsymbol{\theta}}) = \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}} = \text{diag}(\sigma_1^2, \cdots, \sigma_L^2)$ for $i = 1, \cdots, L$.
Let θ be the null value of any scale function such as $\theta = G(\pi_i, \pi_i)$ for all i. Then, the null hypothesis for the value of the estimate of the study is

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_L = \theta \quad \text{or} \quad H_0 : \boldsymbol{\theta} = \boldsymbol{J}\theta \quad \text{for} \quad i = 1, \cdots, L, \tag{2.38}$$

where $\boldsymbol{J}$ is a $L \times 1$ unit vector of ones.

2.4.2 Tests for Homogeneity

The null hypothesis of *homogeneity* specifies that the components of $\boldsymbol{\theta}$ share a common value of θ. Thus, the test for homogeneity among the measures of association $\{\theta_i\}$ on some scale function $\theta_i = g(\pi_{1i}, \pi_{2i})$ is given by

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_L = \theta \tag{2.39}$$

against an alternative hypothesis that, at least two components of $\boldsymbol{\theta}$ are not equal,

$$H_1 : \theta_i \neq \theta_k \quad \text{for} \quad \text{some} \quad i \neq k, \;\; 1 \leqslant i, k \leqslant L. \tag{2.40}$$

Note that, in the case where H_0 true, this means that the stratum-specific estimate $\{\theta_i\}$ share a common value of θ. On the other hand, where H_0 is false, this means that the stratum-specific estimate $\{\theta_i\}$ differs among specific strata and this is referred to as *heterogeneity* among strata, or an interaction between the group and study effects.

2.4.3 Contrast Test for Homogeneity

The null hypothesis for homogeneity in equation(2.39) means that the difference between any two studies is zero in the following way

$$\begin{aligned}
\theta_1 - \theta_2 &= 0 \\
\theta_2 - \theta_3 &= 0 \\
\theta_3 - \theta_4 &= 0 \\
&\vdots \\
\theta_i - \theta_{i+1} &= 0
\end{aligned} \tag{2.41}$$

for $i = 1, 2, \cdots, L$. Now we can write the null hypothesis in the following matrix equation

$$\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_L \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

hence,

$$\boldsymbol{C'\theta} = \mathbf{0} \tag{2.42}$$

where $\boldsymbol{C'}$ is a $(L-1) \times L$ contrast matrix and $\boldsymbol{\theta}$ is a $L \times 1$. Therefore we can define the null hypothesis as

$$H_0 : \boldsymbol{C'\theta} = \mathbf{0}, \tag{2.43}$$

and the alternative hypothesis as

$$H_1 : \boldsymbol{C'\theta} \neq \mathbf{0}. \tag{2.44}$$

The test for homogeneity is provided by the T^2-like wald statistics, define as the quadratic form

$$\chi^2_H = (\boldsymbol{C'\theta})'(\boldsymbol{C'}\hat{\sum}_{\hat{\boldsymbol{\theta}}}\boldsymbol{C})^{-1}\boldsymbol{C'\hat{\theta}} \backsim \chi^2_{L-1}, \tag{2.45}$$

where $\hat{\theta}$ is defined in equation (2.37) as asymptotically normally distributed and the estimate of the covariance matrix is defined in equation (2.37). To obtain the MVLE of the assumed common value for measures of association $\{\theta_i\}$ on the scale $\theta_i = G(\pi_{1i}, \pi_{2i})$. Under the null hypothesis of homogeneity in equation (2.39),

the stratum-specific estimate $\{\theta_i\}$ for the i^{th} strata is asymptotically normally distributed as

$$\hat{\theta} - \theta \backsim N(0, \sigma_i^2). \tag{2.46}$$

Since the hypothesis of homogeneity does not require that $\pi_{1i} = \pi_{2i}$, the variance of $\hat{\theta}_i$ is evaluated assuming that $\pi_{1i} \neq \pi_{2i}$. For now assume that the variances, and their inverses ($\tau_i = \sigma_i^{-2}$) are known. Since $\hat{\theta}$ is consistent for θ, its follows from Slutsky's theorem (Appendix A.2) that

$$\sqrt{\hat{\tau}_j}(\hat{\theta}_i - \theta) \rightarrow \sqrt{\hat{\tau}_i}(\hat{\theta} - \theta) \backsim N(0, 1). \tag{2.47}$$

Since $\hat{\sigma}_i^2$ is consistent for σ_i^2 and $\hat{\tau}_i$ is consistent for τ_i then

$$\hat{\tau}_i(\hat{\theta}_i - \theta)^2 \backsim \chi_1^2. \tag{2.48}$$

Therefore,

$$\chi_{H,C}^2 = \sum_i \hat{\tau}_i(\hat{\theta}_i - \theta)^2 \backsim \chi_{(L-1)}^2 \tag{2.49}$$

is distributed asymptotically as chi-square $L - 1$ degrees of freedom. We estimate θ as a linear combination of the L stratum-specific estimates. Algebraically, Cochran's test is equivalent to the contrast test χ_H^2. The χ_H^2 test has lower power especially when the number of studies is small, as a comprehensive test of heterogeneity [31]. Conversely it has too much power if the number of studies is large [32].

2.5 Heterogeneity

Heterogeneity in meta-analysis is concerned with the variation in result across studies. Results may vary across studies because of random error, even when all studies are measuring the same underlying average effect. Nevertheless probability alone cannot explain the differences in the results across studies [33, 34]. Meta-analysis often includes studies that are different from each other in important ways. Hence heterogeneity may also be due to differences in study design or patient characteristics across studies [35]. Meta-analysis is not only important for pooling studies to increase the power for estimating average treatment effect, but also crucial for investigating potential sources of heterogeneity (exploratory meta-analysis) that may re-

veal important effect modifiers [34, 35, 36, 33, 37]. This project reviews methods that can be used in meta-analysis for exploring heterogeneity. After a brief discussion about the causes of heterogeneity in meta-analysis, statistical tests for homogeneity of study results are discussed.

2.5.1 Causes of heterogeneity in meta-analysis

Possible causes of heterogeneity in meta-analysis include random sampling error, definition of treatment effect, some design features and other factors [38]. The results of different studies will vary even when all the studies are estimating the same underlying effect size as a result of random sampling error [33, 34]. Such random variation is greater in smaller studies and is less of a problem in larger studies [38]. If meta-analysis includes a larger number of studies and the difference across studies is purely due to random variation, then the results of studies will be distributed around an average and there will be fewer studies whose results are far away from the average [34]. Random errors can be estimated using statistical methods. Random variation alone cannot explain the observed heterogeneity across studies in meta-analysis.

Clinical causes of heterogeneity are characteristics of study participants and interventions [37, 36]. The results of different studies may vary because the patients included in different studies varied and so responded differently to a treatment. The identification of patient characteristics (e.g severity of illness, diagnosis, age, gender, etc) as a cause of heterogeneity is clinically important to identify who may benefit more or less from treatment [38]. This may allow better tailoring of treatment to patients. Similarly, considerable differences in the results across studies may be cause by variations in settings and interventions [33, 34]. For example, the level of training and experience of care givers may differ. Design quality factors, such as the method of patient allocation, blinding and length of follow-up may also be important causes of difference in the results of studies. Study design factors are important methodological causes of heterogeneity and sometimes may also be clinically meaningful. For example, studies with different periods of follow-up may report different results. The association between the results and the duration of follow-up may indicate how long a treatment should be given, when a treatment may become effective or how long a treatment effect can last [33].

2.6 Random-effects model

In this section we introduce the random-effects model. We discuss the assumptions of this model, and show how these are reflected in the formulas used to compute a summary effect, and in the meaning of the summary effect. This model specification is equivalent to the null hypothesis of homogeneity in (2.39). There are two possible reason for heterogeneity to take place. The first reason could be that the fixed-effects model is misspecified in some way, or perhaps homogeneity is present on some scale beside the specified analysis. On the other hand, to yield for homogeneity the covariate must be adjusted. The second reason is that the fixed-effects model may no hold, which means that there is some extra variation or over-dispersion due to random differences among strata, and this leads to the formation of a random-effects model.

2.6.1 Model description

In conducting a meta-analysis for medical research, how to deal with heterogeneity between studies is an important problem. The simplest and the most popular method is to use the normal random-effects model, where a treatment effect in each study is assumed to be randomly selected from a normal distribution [2]. Suppose that we have L independent studies, of which the i^{th} study has estimated effects sizes $\hat{\theta}_i$ and true effect size θ_i. A standard model in meta-analysis assumes that $\hat{\theta}_i$ is normally distributed with mean θ_i,

$$\hat{\theta}_i = \theta_i + \epsilon_i, \quad \epsilon_i \backsim N(0, \sigma_i^2) \tag{2.50}$$

where σ_i^2 is the within study variance, describing the extent of estimation error for θ_i. Any measure can be used for $\hat{\theta}_i$ as long as the normality assumption is atleast approximately appropriate. The within study variance σ_i^2 is unknown in practice, but an estimated value from each study is usually used instead by ignoring the effect of estimation. We follow this convention and make no distinction between true and estimated σ_i^2. The normal random-effects model assume that

$$\theta_i = \theta + \varepsilon_i, \quad \varepsilon_i \backsim N(0, \tau^2) \tag{2.51}$$

for $i = 1, \cdots, L$. By combining equation(2.50) and equation(2.51)

$$\hat{\theta}_i = \theta + \epsilon_i + \varepsilon_i, \quad \epsilon_i \backsim N(0, \sigma_i^2), \quad \varepsilon_i \backsim N(0, \tau^2), \tag{2.52}$$

where τ^2 is the between-study variance, describing the extent of heterogeneity of the effect size between studies. The two random errors $\epsilon_i \backsim N(0, \sigma_i^2)$ and $\varepsilon_i \backsim N(0, \tau^2)$ are assumed to be independent, hence

$$\hat{\theta}_i \backsim N(\theta, \tau^2 + \sigma_i^2). \tag{2.53}$$

The parameter θ, is the overall average effect size which is our main interest. The variance components can be expressed as follows

$$\tau^2 = E(\theta_i - \mu_\theta)^2 = V\left[E(\hat{\theta}_i|\theta_i)\right] \tag{2.54}$$

$$\sigma_i^2 = E(\hat{\theta}_i - \theta_i)^2 = E\left[V(\hat{\theta}_i|\theta_i)\right], \tag{2.55}$$

therefore the $unconditional$ variance of each $\hat{\theta}_i$ is

$$V(\hat{\theta}_i) = \tau^2 + \sigma_i^2. \tag{2.56}$$

The fixed-effects model is appropriate if $\tau^2 = 0$, otherwise if $\tau^2 > 0$ then the there is over-dispersion relative to the fixed-effects model. A test of homogeneity in effect provides a test of the null hypothesis $H_{0H} : \tau^2 = 0$ versus alternative $H_{1H} : \tau^2 > 0$. If the test is significant, then a proper analysis using the two stage random-effects model requires that we estimate the between-stratum variance component τ^2. This is readily done using a simple moment estimator derived from the test of homogeneity. The Cochran's test of homogeneity $\chi^2_{H,C}$ can be expressed as a weighted sum of squares $\sum_i \tau_i(\hat{\theta}_i - \hat{\mu}_\theta)^2$, where $\hat{\mu}_\theta$ is the MVLE of the mean measure of association obtained under the fixed-effects model and $\hat{\tau}_i$ is the inverse of the estimated variance of the estimate. The sum of squares of each estimate about the overall mean can be partitioned about the estimated means as

$$\sum_i \tau_i(\hat{\theta}_i - \mu_\theta)^2 = \sum_i \tau_i(\hat{\theta}_i - \hat{\mu}_\theta)^2 + 2\sum_i \tau_i(\hat{\theta}_i - \mu_\theta)(\hat{\mu}_\theta - \mu_\theta) + \sum_i \tau_i(\hat{\mu}_\theta - \mu_\theta)^2 \tag{2.57}$$

$$\chi^2_{H,C} = \sum_i \tau_i(\hat{\theta}_i - \hat{\mu}_\theta)^2 = \sum_i \tau_i(\hat{\theta}_i - \mu_\theta)^2 - 2\sum_i \tau_i(\hat{\theta}_i - \mu_\theta)(\hat{\mu}_\theta - \mu_\theta) - \sum_i \tau_i(\hat{\mu}_\theta - \mu_\theta)^2, \tag{2.58}$$

since,

$$2E\left\{\sum_i \tau_i(\hat{\theta}_i - \mu_\theta)\,(\hat{\mu}_\theta - \mu_\theta)\right\} = 2\sum_i E\left\{\tau_i\,\theta_i\,\mu_\theta - \tau_i\,\theta_i\,\mu_\theta - \mu_\theta\,\hat{\mu}_\theta\,\tau_i - \mu_\theta^2\,\tau_i\right\}$$
$$= 2\sum_i E\left\{\mu_\theta\,\hat{\mu}_\theta\,\tau_i - \mu_\theta^2\,\tau_i\right\} = 0 \tag{2.59}$$

and $V(\hat{\theta}_i) = E(\hat{\theta}_i - \mu_\theta)^2$, then the expected value of the test statistic in equation(2.58) is

$$E(\chi^2_{H,C}) = \sum_i \tau_i V(\hat{\theta}_i) - V(\hat{\mu}_\theta)(\sum_i \tau_i). \tag{2.60}$$

Since $V(\hat{\theta}_i) = \tau^2 + \sigma_i^2$, is the unconditional variance for each $\hat{\theta}_i$, then

$$\sum_i \tau_i V(\hat{\theta}_i) = \sum_i \tau_i(\tau^2 + \sigma_i^2). \tag{2.61}$$

Note that the MVLE is obtained as $\hat{\mu}_\theta = \sum_i \hat{\omega}_i\hat{\theta}_i$ using the MVLE weights $\omega_i = \frac{\tau_i}{\sum_l \tau_l}$, where $\tau = \sigma_i^{-2}$ is assumed known (fixed). Again using the unconditional variance of each $\hat{\theta}_i$, then the other term is

$$V(\hat{\mu}_\theta) = \frac{\sum_i \tau_i^2(\tau^2 + \sigma_i^2)}{(\sum_i \tau_i)^2}. \tag{2.62}$$

Hence,

$$V(\hat{\mu}_\theta)\sum_i \tau_i = \frac{\sum_i \tau_i^2(\tau^2 + \sigma_i^2)}{\sum_i \tau_i} \tag{2.63}$$

therefore expected value is obtatined as,

$$E(\chi^2_{H,C}) = \sum_i \tau_i(\tau^2 + \sigma_i^2) - \frac{\sum_i \tau_i^2(\tau^2 + \sigma_i^2)}{\sum_i \tau_i}. \tag{2.64}$$

Since $\tau = \sigma_i^{-2}$, then simplifying equation(2.64) we obtained

$$E(\chi^2_{H,C}) = (L-1) + \tau^2 \left[\sum_i \tau_i - \frac{\sum_i \tau_i^2}{\sum_i \tau_i} \right]. \tag{2.65}$$

Then the consistent moment estimate for τ^2 is given by

$$\hat{\tau}^2 = \max \left[0, \frac{\chi^2_{H,C} - (L-1)}{\sum_i \hat{\tau}_i - \frac{\sum_i \hat{\tau}_i^2}{\sum_i \hat{\tau}_i}} \right]. \tag{2.66}$$

If the solution of the estimate is a negative value, therefore it is set to zero. Using the unconditional variance of the estimate within each stratum, we can update the estimate $\hat{\theta}$, when given the estimate of $\hat{\tau}^2$ between the strata.
The *first-step iterative estimate* for the weights are

$$\hat{\omega}_i^{(1)} = \frac{\hat{\tau}_i^{(1)}}{\sum_l \hat{\tau}_l^{(1)}} = \frac{\hat{V}(\hat{\theta}_i)^{-1}}{\sum_l \hat{V}(\hat{\theta}_l)^{-1}} = \frac{(\sigma_i^2 + \tau^2)^{-1}}{\sum_l (\sigma_l^2 + \tau^2)^{-1}}, \tag{2.67}$$

and the *first-step iterative estimate* for the MVLE mean of the strata is

$$\hat{\mu}_\theta^{(1)} = \sum_j \hat{\omega}_i^{(1)} \hat{\theta}_i \tag{2.68}$$

with the estimated variance

$$\hat{V}(\hat{\mu}_\theta^{(1)}) = \sum_i (\hat{\omega}_i^{(1)})^2 (\sigma_i^2 + \tau^2). \tag{2.69}$$

To recalculate the test for homogeneity the reweighted estimate of the mean $\hat{\mu}_\theta^{(1)}$ would be used. The reweighted mean is also used to update the estimate of the variance between the strata $(\hat{\tau}^2)^{(2)}$. The updated weights $\hat{\omega}_i^{(2)}$, mean $\hat{\mu}_\theta^{(2)}$ and so on they are obtain by the updated variance.The iterative procedure continues until both the mean $\hat{\mu}_\theta^{(m)}$ and variance $\hat{V}(\hat{\mu}_\theta^{(m)})$ converges to constants.

Remark
The addition of the nonzero variance component between the strata, $\hat{\tau}^2$ to the variance in equation(2.56) has the effect of adding a constant to all of the weights. Thus, the random-effects shrinks the weight, so that the resulting estimate is close to the unweighted mean of the fixed-effects model.

2.7 Choice between fixed-effects and random-effects models

Hedges et.al[39, 40] developed both fixed-effects and random-effects models for combining effect sizes. Meta-analysis is used as a way of trying to find the true effect size (i.e the effect size in a population) by combining effect sizes from individual studies. Fixed-effects and random-effects models are the two ways to explain this concept of meta-analysis. The difference between these two models is explained by Hedge [39, 40]. In the fixed-effects model the effect sizes in the population are fixed but unknown constants and the effect sizes in the population is assumed to be the same for all studies [41]. This situation is called the homogeneous case. The other possibility is that the population effect sizes vary from study to study. In this case each study in a meta-analysis comes from a population that is likely to have a different effect size to any other study in the meta-analysis. This case is called heterogeneous case, where the population effect size is sampled from a finite population [39, 42].

To simplify the concept of fixed-effects model, the studies in the analysis share a common true effect. While in the random-effects model, studies in the meta-analysis are assumed to be only a sample of all possible studies that could be done on a given topic [41]. The calculation of standard errors associated with the combined effect sizes is the main difference between these models. The fixed-effects model uses only within study variability in their standard error term because all other unknowns in the model are assumed not to affect the effect sizes [39, 40]. In the random-effects model it is necessary to account for the random errors associated with sampling variation from populations that themselves have been sampled from a finite population. As such the error term contains two components, within study variability and variability, arising from differences between studies [40]. If effect sizes are heterogeneous, then the resulting standard errors in the random-effects model are usually much larger than in the fixed-effects model and therefore significance tests of combined effect sizes are more stable. In reality the random-effects model is probably more realistic than the fixed-effects model on the majority of occasions (especially when the researcher wishes to make general conclusions about the research domain as a whole and not restrict their findings to the studies included in the meta-analysis). Despite this fact, the National Research Council reports that fixed-effects are the rule rather than the expectation [43]. Osburn and Callender [44] have also noted that real world data is likely to have heterogeneous population effect sizes even in the absence of known moderator variables[45]. Despite these observation, Hunter and Schmidt [41] reviewed the meta-analysis studies reported in psychology and found 21 studies reporting fixed-effects meta-analysis but none using random-

effects model. Although fixed-effects model have attracted considerable attention [39, 46]. The choice of a model depends largely on the type of inferences that the researcher wishes to make [40]. The random-effects model facilitate inferences that generalize beyond the studies included in the meta-analysis. While the fixed-effects model is appropriate only for inferences that extend only to the studies included in the meta-analysis. Random-effects model are appropriate in the real world data since researchers usually wish to make inferences that generalize beyond the studies included in the meta-analysis [40].

2.8 The likelihood method

Maximum likelihood is widely used for estimation and inference. In this section we review a likelihood method to obtain the estimates of the two parameters of interest in the random-effects model, θ and τ^2.

2.8.1 Estimating θ and τ^2 using maximum likelihood

Recall that the standard random-effects model has $\hat{\theta}_i \sim N(\theta, \tau^2 + \sigma_i^2)$, $i = 1, 2, \cdots, L$ and that the σ_i^2 is treated as a known constant. The density function for the random variable $\hat{\theta}_i$ is

$$f(\hat{\theta}, \tau^2) = \frac{1}{\sqrt{(\sigma_i^2 + \tau^2)2\pi}} \exp^{-} \frac{1}{2} \frac{(\hat{\theta}_i - \theta)^2}{\sigma_i^2 + \tau^2}, \tag{2.70}$$

and the likelihood function is given as

$$\begin{aligned} L(\theta, \tau^2) &= \prod_{i=1}^{L} f(\theta, \tau^2) \\ &= \prod_{i=1}^{L} \frac{1}{\sqrt{(\sigma_i^2 + \tau^2)2\pi}} \exp^{-} \frac{1}{2} \frac{(\hat{\theta}_i - \theta)^2}{\sigma_i^2 + \tau^2} \\ &= \prod_{i=1}^{L} \left(2\pi(\sigma_i^2 + \tau^2)\right)^{-\frac{1}{2}} \exp^{-} \frac{1}{2} \sum_{i=1}^{L} \frac{(\hat{\theta}_i - \theta)^2}{\sigma_i^2 + \tau^2}. \end{aligned} \tag{2.71}$$

Finally the log-likelihood function is

$$\ell(\theta, \tau^2) = -\frac{1}{2} \sum_{i=1}^{L} \log(2\pi(\sigma_i^2 + \tau^2)) - \frac{1}{2} \sum_{i=1}^{L} \frac{(\hat{\theta}_i - \theta)^2}{\sigma_i^2 + \tau^2}, \quad \theta \in \mathbb{R}, \quad \tau^2 \geqslant 0. \tag{2.72}$$

Maximum likelihood estimates for θ and τ^2 can be found by maximizing equation

(2.72) with respect to θ and τ^2 . We first differentiate equation (2.72) with respect to θ and equate the differential to zero to obtain the maximum likelihood estimate for θ

$$\begin{aligned}\frac{\partial \ell(\theta,\tau^2)}{\partial \theta} &= -\frac{1}{2}\times 2\sum_{i=1}^{L}\frac{(\hat{\theta}_i-\theta)}{\sigma_i^2+\tau^2}\times -1\\ &= \sum_{i=1}^{L}\frac{(\hat{\theta}_i-\theta)}{\sigma_i^2+\tau^2}\\ &= \sum_{i=1}^{L}\frac{\hat{\theta}_i}{\sigma_i^2+\tau^2}-\sum_{i=1}^{L}\frac{\theta}{\sigma_i^2+\tau^2},\end{aligned} \tag{2.73}$$

equating (2.73) to zero and re-arranging to obtain a maximum likelihood estimate for θ

$$\hat{\theta}=\sum_{i=1}^{L}\frac{\hat{\theta}_i}{\sigma_i^2+\tau^2}\Big/\sum_{i=1}^{L}\frac{1}{\sigma_i^2+\tau^2}. \tag{2.74}$$

Maximizing (2.72) to obtain maximum likelihood for τ^2

$$\begin{aligned}\frac{\partial \ell(\theta,\tau^2)}{\partial \tau^2} &= -\frac{1}{2}\sum_{i=1}^{L}\frac{1}{2\pi(\sigma_i^2+\tau^2)}\times 2\pi-\frac{1}{2}\sum_{i=1}^{L}-\frac{(\hat{\theta}_i-\theta)^2}{(\sigma_i^2+\tau^2)^2}\\ &= -\frac{1}{2}\sum_{i=1}^{L}\frac{1}{(\sigma_i^2+\tau^2)}+\frac{1}{2}\sum_{i=1}^{L}\frac{(\hat{\theta}_i-\theta)^2}{(\sigma_i^2+\tau^2)^2},\end{aligned} \tag{2.75}$$

setting equation(2.75) to zero then, we obtain an estimate for τ^2

$$\begin{aligned}\sum_{i=1}^{L}\frac{(\hat{\theta}_i-\theta)^2}{(\sigma_i^2+\hat{\tau}^2)^2} &= \sum_{i=1}^{L}\frac{1}{(\sigma_i^2+\hat{\tau}^2)}\\ \sum_{i=1}^{L}\frac{(\hat{\theta}_i-\theta)^2}{(\sigma_i^2+\hat{\tau}^2)^2} &= \sum_{i=1}^{L}\frac{1}{(\sigma_i^2+\hat{\tau}^2)}\times\frac{\sigma_i^2+\hat{\tau}^2}{\sigma_i^2+\hat{\tau}^2}\\ \sigma_i^2+\hat{\tau}^2 &= \sum_{i=1}^{L}\frac{(\hat{\theta}_i-\theta)^2}{(\sigma_i^2+\hat{\tau}^2)^2}\Big/\sum_{i=1}^{L}\frac{1}{(\sigma_i^2+\hat{\tau}^2)^2}\\ \hat{\tau}^2 &= \sum_{i=1}^{L}\frac{(\hat{\theta}_i-\theta)^2-\sigma_i^2}{(\sigma_i^2+\hat{\tau}^2)^2}\Big/\sum_{i=1}^{L}\frac{1}{(\sigma_i^2+\hat{\tau}^2)^2}.\end{aligned} \tag{2.76}$$

The large body of asymptotic theory existing for estimators, is the one major advantage of maximum likelihood estimation. In regular cases a maximum likelihood es-

timator from sample of L independent and identically distributed random variables has a normal distribution. The L variables $\hat{\theta}_i, i = 1, 2, \cdots, L$ from a meta-analysis are independent but not identically distributed since, $\text{Var}(\hat{\theta}_i) = \sigma_i^2 + \tau^2$. The standard assumption will still apply in any realistic meta-analysis with large L. However it is possible to construct a confidence interval for θ, using this asymptotic distribution, since the asymptotic variance of $\hat{\theta}$ depends on the unknown τ^2. This is only an approximate interval. Therefore the variance of $\hat{\theta}$ is given by

$$\begin{aligned} \text{Var}(\hat{\theta}) &= \text{Var}\left(\sum_{i=1}^{L} \frac{\hat{\theta}_i}{\sigma_i^2 + \tau^2} \Big/ \sum_{i=1}^{L} \frac{1}{\sigma_i^2 + \tau^2}\right) \\ &= \frac{1}{\sum_{i=1}^{L} (\sigma_i^2 + \tau^2)^{-1}}. \end{aligned} \tag{2.77}$$

Under the assumption of asymptotic normality we therefore have

$$\hat{\theta} \sim N\left(\theta, \frac{1}{\sum_{i=1}^{L} (\sigma_i^2 + \tau^2)^{-1}}\right). \tag{2.78}$$

This distribution is used for $\hat{\theta}$ even though the likelihood estimate of τ^2 may lie on the boundary of the parameter space, namely, $\tau^2 = 0$. Note that the variance of $\hat{\theta}$ is of the same form as that for the DerSimonian and Laird random-effects model. In this case $\text{Var}(\hat{\theta})$ is estimated using $\hat{\tau}^2$ without any modification to the distribution of $\hat{\theta}$. Therefore the confidence interval for θ is

$$\hat{\theta} \pm \text{Z}_{1-\frac{\alpha}{2}} \sum_{i=1}^{L} \frac{1}{\sigma_i^2 + \tau^2}. \tag{2.79}$$

This method is referred to as the simple likelihood method. A test of homogeneity and a confidence interval may be derived using the generalized likelihood ratio statistic Λ_L together with the fact that $\lambda_L = -2log(\Lambda_L)$ is, under the homogeneity hypothesis $\tau^2 = 0$, asymptotically distributed as χ_1^2 [47]. An asymptotic $100(1-\alpha)\%$ confidence interval for τ^2 is given by the set

$$C_{1-\alpha} = \{\tau^2 : \lambda_L(\tau^2) \leq \chi_1^2; 1-\alpha\}, \tag{2.80}$$

where $\chi_1^2; 1-\alpha$ is the $100(1-\alpha)^{th}$ percentile point of the χ_1^2 distribution [47]. An alternative asymptotic confidence interval for τ^2 can be found by arguing that the asymptotic distribution of the MLE of τ^2 is normally distributed with mean θ and

variance equal to the inverse of the fisher information and it has been shown its readily shown that this variance is $\frac{2}{\sum(\sigma_i^2+\tau^2)^{-2}}$ [48, 49], so that a $100(1-\alpha)\%$ asymptotic symmetric confidence interval for τ^2 is

$$\hat{\tau}^2 - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{2}{\sum(\sigma_i^2+\tau^2)^{-2}}}; \hat{\tau}^2 + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{2}{\sum(\sigma_i^2+\tau^2)^{-2}}}, \tag{2.81}$$

where $Z_{1-\frac{\alpha}{2}}$ is the $100(1-\frac{\alpha}{2})^{th}$ percentile point of the normal distribution.

2.9 Publication bias

Publication bias occurs when results of published studies are systematically different from the results of unpublished studies [50]. Since published studies are more likely to find their way into a meta-analysis, any bias in the literature is likely to be reflected in the meta-analysis [9]. If the findings of published studies are systematically different from those unpublished studies, then the evidence base for clinical and health-policy decisions will be questionable. As a result, the published studies will not be a valid representation of all studies conducted [50]. If studies that are included in the meta-analytic analysis are biasedly sampled from all relevant studies, then the mean effect computed by meta-analysis will reflect this bias [9].

In general, studies with statistically significant or positive results are more likely to be published than those with nonsignificant or negative results [50]. Also for any given sample size the results are more likely to be statistically significant if the effect size is larger. It follows that if there is a population of studies that looked at the magnitude of a relationship and the observed effects are distributed over a range of values, the studies with effects towards the higher end of that range are more likely to be statistically significant and therefore to be published [9]. If studies have relatively small sample size, this tendency has the potential to produce very large biases, in the magnitude of the relationship [9]. Rothstein [51] reviewed the 95 meta-analytic reviews published in psychological bulletin between 1995 and 2005 to see whether they included unpublished research, she found that 23 of the 95 clearly did not include any unpublished data. Clarke and colleagues [52] studied the references from health care protocols and reviews published in the Cochrane library in 1999. They found that about 92% of references to studies included in reviews were from journal articles. Of the remaining 8%, about 4% were from conference proceedings, about 2% were from unpublished material, and slightly over 1% were from book chapters. Furthermore, they also looked at the sources of unpublished literature included in the first 100 Cochrane systematic reviews and found that nearly half of

them did not include any data from unpublished sources [9].

Dissemination bias describes all forms of biases such as language bias (English-language databases and journals are more likely to be searched, which leads to an over-sampling of statistically significant studies) [53, 54], Citation bias (Whereby studies with systematically significant results are more likely to be cited by others and therefore easier to identity) [55, 56], availability bias (selective inclusion of studies that are easily accessible to the researcher), cost bias (selective inclusion of studies that are available free or at low cost), familiarity bias (selective inclusion of studies only from ones own discipline), duplication bias (studies with statistically significant results are more likely to be published more than once) [57] and outcome bias (selective reporting by the author of a primary study of some outcomes but not others, depending on the direction and statistical significance of the results).

All these biases lead to the same consequence namely that the literature located by a systematic reviewer will be unrepresentative of the population of completed studies. Hence all present the same threat of review validity of systematic reviews and reduce our ability to produce valid conclusions based on a body of evidence. This review highlights that no empirical studies of current interventions have shown that they reduce this bias. Publication bias will result in misleading estimates of treatment effects and associations between study variables [50]. Clinical trials often use the results of base medical research. Clinical trials may waste limited resources and fail to confirm the published results of basic studies, if the results of basic research are falsely positive due to biased selection for publication [58]. For example, publication bias may be used to explain the observed discrepancy in results between animal studies and clinical trials regarding the neuroprotective efficacy of nicotinamide for focal cerebral ischemia [59]. Over a wide range of health risk factors, results of observational studies are often highly contradictory, which may be partly due to publication bias [60]. For example, publication bias may cause highly contradictory results observed in early published studies of genetic associations [61]. Publication bias in clinical trials has a direct impact on patients's and population's health. When the relative efficacy of a treatment is overestimated because of publication bias, health resources can be wasted by purchasing more expensive interventions, instead of cheaper alternatives, without corresponding improvement in outcome. There are also many reported cases in which patients have received ineffective or harmful treatment [50]. For example, biased reporting of trial results delayed the detection of increased mortality risk of rofecoxib for alzheimer's disease and cognitive impairment and more than 10 million patients had used rofecoxib before its withdrawal in 2004 [62]. In a meta-analysis, it is possible that the studies

may overestimate the true effect size because they are based on a biased sample of the target population of studies. The way to deal with this concern is to compare effects in the published studies formally with effects in the unpublished studies [9]. If we had the access to the unpublished studies we would no longer be concerned. Nevertheless, the best approach would be for the reviewer to perform a truly comprehensive search of the literature, on hopes of minimizing bias. In fact, there is evidence that this approach is somewhat effective [9]. Moreover funnel plots and statistical methods can be used to indicate the presence or absence of publication bias. Although these can be unreliable in many circumstances [50].

A funnel plot, is a scatter diagram used to visually represent the relationship between the effect of an intervention/treatment (x-axis) and study precision (y-axis), have been proposed as a means of detecting publication bias in meta-analysis. From a statistical point of view, the precision of the estimated intervention effect will increase with increase in study size (presuming the event rate was the same across all studies) [63]. The dots in the funnel plot represent the studies, the larger studies will appear at the top of the funnel, while in contrast, effect estimates from smaller studies may be expected to scatter near the bottom of the funnel. Publication and selection biases in meta-analysis are more likely to affect small studies, which also tends to be of lower methodological quality. This may lead to small study effects where the smaller studies in a meta-analysis show larger treatment effects [64]. One reason for this may be due to publication bias, where the chances of a small study being published is higher if the study shows a statistically significant effect [65]. Moreover small study effects may also arise because of between trial heterogeneity [64]. Although it should be noted that the term small study effect for this case is somewhat misleading, since large studies can be imprecise, just as small studies can be precise. The imprecise study effects may be a better description. In the absence of publication bias, the graph resembles a symmetrical inverted funnel because the treatment effect estimates from smaller studies scatter more widely at the bottom of the graph , with the spread narrowing with increasing precision among larger studies. On the other hand if there is a presence of publication bias, there will be asymmetry as though a bite has been taken out of the funnel.

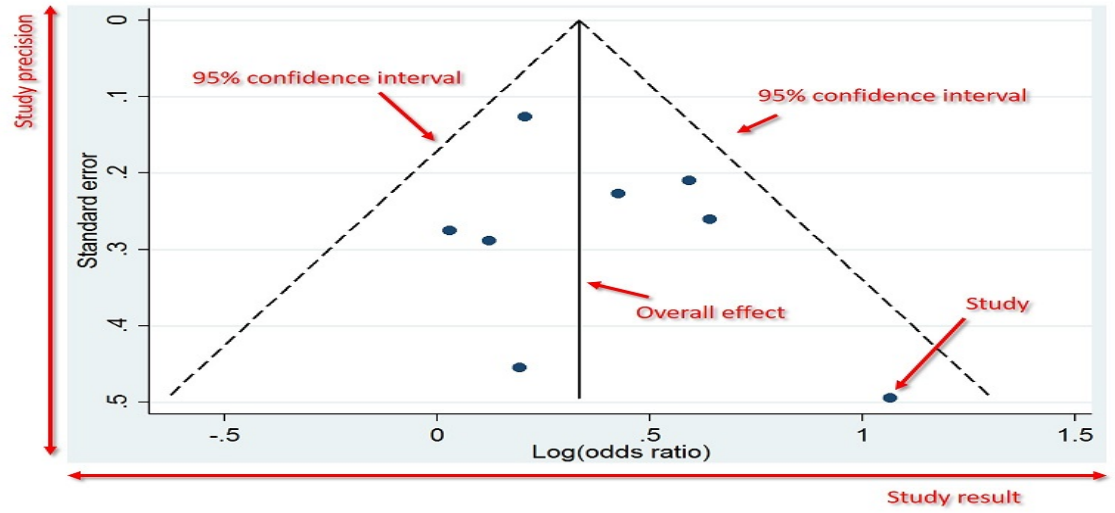

Figure 2.1 – Example of the funnel plot.

Source:https://toptipbio.com/funnel-plot/funnel-plot-annotated/

Publication bias and true heterogeneity in intervention effects are two factors which contribute to funnel plot asymmetry [63]. Asymmetry could also result from the overestimation of treatment effects in smaller studies of inadequate methodological quality. For example, a significant effect may only be seen in high risk patients and these patients were more likely to be recruited into the smaller, early trials [63]. Larger, multi-centre interventions involving hundreds of patients may be more difficult to implement when compared with smaller trials and in addition, there may be methodological differences between centres. In such a case, the data from the smaller, better controlled study may be more precise than the larger and perhaps less forcefully implemented study involving a more heterogeneous group of participants. Furthermore, heterogeneity of treatment effects will lead to funnel plot asymmetry if the true treatment effect is larger in the smaller trials. For example, if a combined outcome is considered then substantial benefit may be seen only in patients at high risk for component of the combined outcome which is affected by the intervention. Trials conducted in high risk patients will also tend to be smaller, because of the difficulty in recruiting such patients. Funnel plot shapes can also be influenced by the statistic used to measure the effect size [63].

2.10 Forest Plot

A forest plot is a graphical method that is commonly used in meta-analysis. It is used for the visual representation of the trials included in the analysis and the results for each of the trials. Moreover, to display point estimates and corresponding confidence interval for all the studies in the analysis. The summary estimate is more than the weighted mean of the individual effects. However, the mechanism used to assign the weights (and therefore the meaning of the summary effect) depends on our assumptions about the distribution of effect sizes from which the studies were sampled. Under the fixed-effects model, we assume that all studies in the analysis share the same true effect size, and the summary effect is our estimate of this common effect size. Under the random-effects model, we assume that the true effect size varies from study to study, and the summary effect is our estimate of the mean of the distribution of effect sizes. Each study is represented by a horizontal line. However in a case where there are no events, then that specific study would not be represented by a line. Such studies will be excluded from the meta-analysis. There is a box in the line for each study. The mid-point of the box represents the point effect estimate that is, the treatment effect for each study. The size of the box represents the weight given to the study. For instance, if more weight is given to the study, then the size of the box will be big. But if less weight is given to the study, then the size of the box will be small. This is designed so that eyes are drawn towards the studies that are given more weights.

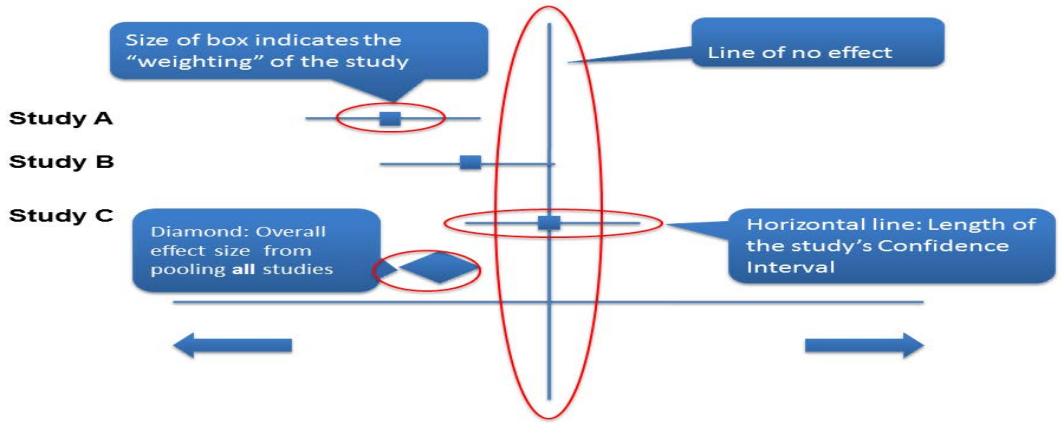

Figure 2.2 – Example of the forest plot.

Source:https://guides.lib.monash.edu/systematic-review/synthesis/quantitativedata

The diamond below the plot represents the overall effect. The width of the lines shows the confidence intervals of the effect estimate of individual studies and the mid-point of the diamond is the point estimate of the treatment effect. Pooling many precise large studies will results in a large long thin diamond while a meta-analysis that contains few small and imprecise studies will results in a small wide diamond. In addition there is a vertical line which corresponds to the value zero in the plot shown. This is the line of no effect. Note that the left of the vertical line it favours drug group and on the right of the vertical line it favours placebo group. All effect size commonly used as effect measures in meta-analysis are relative measures. In this case, zero indicates no effect. If zero is included in the 95% confidence intervals, it indicates that there is no statistical significance at 5% significance levels. If zero is not included in the 95% confidence intervals, the results are statistically significant at 5% significance levels. This is applicable for effect estimates for the individual study level and for the overall estimate. Whether the intervention is beneficial or harmful depends upon the context.

Chapter 3

Meta-analysis for longitudinal studies

In this chapter, we introduce the linear mixed-effects models for meta-analysis of longitudinal studies. We discuss the assumptions of this model, and show how these are reflected in the formulas used to compute a summary effect.

3.1 Introduction

In longitudinal studies the study participants are measured repeatedly over time, thereby allowing the direct assessment of changes in the response variable over time [66]. Studies that are longitudinal in nature involves multiple or repeated measurements on the same individuals at different times. Furthermore, the studies are designed to investigate changes in response of interest over time on each subject. This type of study is useful for estimating the relationship between risk factors and the development of disease, and the outcomes of the treatment at different points in time. For example, HIV patients may be followed over time and measurements such as CD4 counts, or viral load collected to characterize immune status and disease burden respectively [67]. Special statistical techniques are required for valid analysis and inference for such repeated measures data that are correlated within subjects. Longitudinal studies are in contrast to cross-sectional studies where measurements are obtained at only a single point in time, where it is not possible to assess individual changes on the basis of a single point in time [66]. The primary objectives in longitudinal studies are often to examine the factors that influence heterogeneity among individuals in how individual change throughout the duration of the study [66]. The factors that influence heterogeneity among individuals are genetic, environmental, social, and behaviorial factors [66]. This heterogeneity is natural in terms of how the disease develops and progresses [66]. Longitudinal studies are often use-

ful to characterize normal growth and aging, to assess the effect of risk factors on human health and to evaluate the effectiveness of treatments.

During the course of follow-up, longitudinal studies typically report estimates of the effect of a treatment or exposure at different time points [68]. Meta-analysis of these studies must account for correlations between effect size estimates from the same study [68]. The effects of treatment can be reported by estimates calculated at different times, corresponding to the measurement times in the study [68]. We can analyse longitudinal studies using summary measures in meta-analysis [69]. The effects that are reported by meta-analysis of longitudinal studies in terms of a single summary measure can be handled with standard approaches [2, 70]. Mass [71] describes a mixed-effects model for the meta-analysis of longitudinal effect estimates, where He handle the correlation between observations by allowing random intercepts and linear time effects.

Meta-analysis of longitudinal studies combines effect sizes measured at different time points [72]. In meta-analysis, multiple combined results are required when there are multiple end points of interest across studies, such as multiple outcomes [73], multiple time points [74] and multiple treatments effect [75]. The effect sizes are correlated because the are calculated at multiple time points, from the same group of patients. In such cases there are multiple correlated effect sizes per study. A meta-analyst can choose to perform a separate univariate meta-analysis for each effect size in which individual effect sizes from two or more studies that are combined into a single summary effect size or to perform multivariate meta-analysis where the multiple effect sizes are jointly analysed [72]. Both these approaches can be performed using a standard statistical software such as STATA, R and SAS. The biggest challenge in meta-analysis is to account for correlation between effect sizes both within and between studies when the effect sizes are reported longitudinally. The disadvantage of using separate univariate meta-analysis approach is that it ignores correlation between the effect sizes and this can increase the standard error of point estimates [76]. Furthermore, this might result in bias parameter estimates [76]. Riley [76] studied the effect of ignoring within study correlation and found that ignoring the within study correlation gives poor meta-analysis results with generally inferior statistical properties, for example, it increases the mean-square error and standard error of combined estimates. Moreover, Trikalinos et.al.[77] examined the data with univariate and multivariate models based on discrete and approximate likelihood. They found that both these models were comparable since the summary effects for each outcome were similar with univariate and multivariate meta-analysis. However, the multivariate model with discrete likelihood gave smaller between study

variance estimates and narrower predictive intervals for new studies. Furthermore, Olkin and Gleser [75] have considered that it is generally preferable to analyse the data simultaneously in multivariate meta-analysis framework so as to provide combined conclusions and simultaneous confidence interval for each time points [75]. In contrast to the confidence interval obtained from separate meta-analysis, the idea of simultaneous confidence interval provide correct average probabilities [75]. Multivariate meta-analysis can reduce the impact of bias when compared to univariate meta-analysis in case of outcome reporting bias where some studies in meta-analysis partially report results [78]. Moreover, multivariate meta-analysis allows the joint combination of summing-up effect sizes estimates from multiple end points and accounts for within study and between study correlation. In addition this approach describes the associations between the estimates of effects in order to help make predictions about the true effects of a new study and provide estimates with better statistical properties.

3.2 Theory of the linear mixed-effects model for meta-analysis

Linear mixed-effects models are statistical models for continuous outcome variable in which the residuals are normally distributed but may neither be independent or have constant variance [79]. Recently, in biomedical, economics, education, pharmacological and psychological studies, the application of the linear mixed-effects models to repeated measures data from longitudinal studies has become frequently used due to increasing availability of software that can be used to fit this model [80]. Furthermore, these statistical models can be used to analyse correlated data, that include clustered, longitudinal or repeated measures data that quantifies the relationship between a continuous dependent variable and various predictor variables [79].

In linear mixed-effects models variance-covariance structures are used to describe the correlation present in the response for a given subject [81]. These models provide a flexible and powerful tool for the analysis of data with a complex variance covariance structure, such as correlated data due to a grouping of subject or repeated measurements over time [82]. In addition, linear mixed-effects models combine the information from multiple subjects to improve estimates and inference [80]. The name linear mixed-effects models comes from the fact that these models are linear, and that the covariates or independent variables, may involve a mix of fixed-effects and random-effects. Fixed-effects are the covariates effects that are fixed across subjects in the study sample. These effects are the ones of our particular interest. In a linear mixed-effects model, fixed-effects are unknown constant parameters associated with

either continuous covariates or the levels of categorical factors [79]. In the studies included in a meta-analytical research, fixed-effects analysis models the systematic between study differences and assumes subject level sampling errors [80]. Also, the use of fixed-effects parameters is to describe the relationship of the covariates to the dependent variable for an entire population. Random-effects are in contrast to fixed-effects, which are represented by constant parameters in a linear mixed-effects model [79]. Furthermore, random-effects are represented by unobserved random variables, which are usually assumed to follow a normal distribution. The covariates effects that vary among subjects are the random-effects, since each subject is a random subject drawn from a population. That is, each study specific effect is sampled from the larger population of effects [81]. Hence, each study has its own population effect and inference is made about the larger population of effects [81]. In the random-effects there are two sources of variability. Firstly, variability due to the effect parameters and secondly, sampling variability of experimental units into studies. In other words, random-effects analysis takes into account the true variance in addition to the modeled between study differences and the sampling error in fixed-effects models [81]. Moreover, random-effects are used to model the random variation in the dependent variable at different levels of the data since they are specific to clusters or subjects within a population [79]. The type of correlation present is described by a variance-covariance structure, while random-effects for subject describes only the cause of correlation.

3.2.1 Model description

Consider a meta-analysis of L studies denoted by $i = 1, \cdots, L$. Also consider T longitudinal effect sizes per study denoted by $t = 1, \cdots, T$. So each study i yield T estimated effect sizes $\boldsymbol{Y_i} = (Y_{i1}, \cdots, Y_{it}, \cdots, Y_{iT})$ such that

$$\boldsymbol{Y_{it}} = \boldsymbol{X'_{it}}\boldsymbol{\beta} + \boldsymbol{Z'_{it}}\boldsymbol{u_i} + \boldsymbol{\epsilon_{it}}. \tag{3.1}$$

In this linear model we define our variables as follows,

- $\boldsymbol{X}'_{it}$ is a $1 \times p$ design vector of p fixed-effects.
- $\boldsymbol{\beta}$ is the corresponding responding regression coefficients contained in the $p \times 1$ vector.
- $\boldsymbol{Z}'_{it}$ is a $1 \times q$ design vector of $q(\leq p)$ random-effects.
- $\boldsymbol{u}_i$ is a $q \times 1$ vector of random-effects.
- $\boldsymbol{\epsilon}_{it}$ is the vector of residuals.

By extending the equation(3.1) above we obtain the general form of linear mixed-effects model for $\boldsymbol{Y}_i$ that can account for the correlations between longitudinal effect sizes is given by

$$\boldsymbol{Y_i} = \boldsymbol{X_i\beta} + \boldsymbol{Z_iu_i} + \boldsymbol{\epsilon_i}. \tag{3.2}$$

Where,

- $\boldsymbol{Y}_i$ is the $T \times 1$ vector of effect sizes from a number of L related but independent studies.
- $\boldsymbol{X}_i$ is the $T \times p$ design matrix describing study covariates that influence fixed-effects.
- $\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed-effects parameters.
- $\boldsymbol{Z}_i$ is the subset of $\boldsymbol{X}_i$ is a $T \times q$ design matrix describing the covariates of q random-effects.
- $\boldsymbol{u}_i$ is the $L \times q$ vector of random-effects or residuals on the between study level.
- $\boldsymbol{\epsilon}_i$ is the $T \times 1$ vector of residuals on the within study level.

We represent elements of the $\mathbf{Y}_i$ vector as follows

$$\mathbf{Y}_i = \begin{bmatrix} Y_{i1} \\ \vdots \\ Y_{it} \\ \vdots \\ Y_{iT} \end{bmatrix}.$$

Note that the number of elements in the $\mathbf{Y}_i$ may vary from one study to another. $\mathbf{X}_i$ represent the known values of the p covariates, $X_1, \cdots, X_p$ for each of the T longitudinal effect sizes collected on the i^{th} study

$$\mathbf{X}_i = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{L1} & X_{L2} & \dots & X_{Tp} \end{bmatrix}.$$

The first column would simply be equal to one for all observations, in a model that includes an intercept term. Note that all elements in a column of the $\mathbf{X}_i$ matrix corresponding to a time variant or study specific covariate will be the same. For ease

of presentation, we assume that the $\mathbf{X}_i$ matrices are full rank, that is none of the column or rows is a linear combination of the remaining ones. In general $\mathbf{X}_i$ matrices may not be of full rank, and this may lead to an aliasing or parameter identifiable problem for the fixed-effects stored in the vector $\boldsymbol{\beta}$. The vector of p unknown regression coefficients or fixed-effect parameters associated with the p covariates used in constructing the $\mathbf{X}_i$ matrix and $\boldsymbol{\beta}$ can be represented by

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}.$$

The $\mathbf{Z}_i$ matrix is a design matrix that represents the known values of the q covariates $Z_1, \cdots, Z_q$ for the i^{th} study. This matrix is very much like the $\mathbf{X}_i$ matrix in that it represents the observed values of covariates, however it usually has fewer columns than the $\mathbf{X}_i$ matrix

$$\mathbf{Z}_i = \begin{bmatrix} Z_{11} & Z_{12} & \dots & Z_{1q} \\ Z_{21} & Z_{22} & \dots & Z_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{L1} & Z_{L2} & \dots & Z_{Tq} \end{bmatrix}.$$

The columns in the $\mathbf{Z}_i$ matrix represent observed values for the q predictor variables for the i^{th} study, which have effects on the continuous response variable that vary randomly across studies. In many cases, predictors with effects that vary randomly across studies are represented in both the $\mathbf{X}_i$ and the $\mathbf{Z}_i$ matrix. In a linear mixed-effects model in which only the intercepts are assumed to vary randomly from study to study, the $\mathbf{Z}_i$ matrix would simply be a column of ones.

The $\mathbf{u}_i$ vector for the i^{th} study represents a vector of q random-effects associated with the q covariates on the $\mathbf{Z}_i$ matrix

$$\mathbf{u}_i = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_L \end{bmatrix}.$$

Recall that by definition, random-effects are random variables. We assume that the q random-effects in the $\mathbf{u}_i$ vector follow a multivariate normal distribution, with mean

vector **0** and variance covariance matrix denoted by **D** such that

$$\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{D}). \tag{3.3}$$

Elements along the main diagonal of the **D** matrix represent the variance of each random-effects in $\mathbf{u}_i$, and the off diagonal elements represent the covariance between two corresponding random-effects. Because there are q random-effects in the model associated with the i^{th} study, **D** is a $q \times q$ matrix that is symmetric and positive definite. Elements of this matrix are shown below as follows

$$\mathbf{D} = Var(\mathbf{u}_i) = \begin{bmatrix} Var(u_1) & Cov(u_1, u_2) & \dots & Cov(u_1, u_q) \\ Cov(u_1, u_2) & Var(u_2) & \dots & Cov(u_2, u_q) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(u_1, u_q) & Cov(u_2, u_q) & \dots & Var(u_q) \end{bmatrix}.$$

Finally, $\boldsymbol{\epsilon}_i$ is a vector of L residuals, with each elements in $\boldsymbol{\epsilon}_i$ denoting the residual associated with an observed response at time t for the i^{th} study. It is because some study might have more observations collected than others. The $\boldsymbol{\epsilon}_i$ vector may have different number of elements

$$\boldsymbol{\epsilon}_i = \begin{bmatrix} \epsilon_{i1} \\ \vdots \\ \epsilon_{it} \\ \vdots \\ \epsilon_{iT} \end{bmatrix}.$$

In contrast to the standard linear model, the residuals associated with repeated observations on the same study in an linear mixed-effects model can be correlated. We assume that the $L \times 1$ residuals in the $\boldsymbol{\epsilon}_i$ vector for a given study, i, are random variables. That follows a multivariate normal distribution with a mean vector **0** and a positive definite symmetric covariance matrix $\mathbf{R}_i$

$$\boldsymbol{\epsilon}_i \sim (\mathbf{0}, \mathbf{R}_i). \tag{3.4}$$

We assume that the vectors of residuals $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \cdots, \epsilon_{it}, \cdots, \epsilon_{iT})'$ and random-effects $\mathbf{u}_i = (u_1, u_2, \cdots, u_L)'$ are independent of each other. Effect sizes from different studies are assume to be independent of each other, that is $\text{cov}(\epsilon_{it}, \epsilon_{mt}) = 0$ when $i \neq m$ for time points t, $t' = 1, \cdots, T$. We also assume that residuals and

random-effects are independent, hence cov($\boldsymbol{\epsilon}_i, \boldsymbol{u}_i$) $= 0$. The within study residuals are assumed to be normally distributed. Moreover, they are usually assumed to be distributed in an identical manner within each group. However, for a meta-analysis involving large sample sizes, the within study variances can be considered unknown, and the covariance matrices $\mathbf{R}_i$ are specified as diagonal matrices with known sample variance of the study effect sizes on their diagonals. We represent the general form of the $\mathbf{R}_i$ matrix as shown below

$$\mathbf{R}_i = Var(\boldsymbol{\epsilon}_i) = \begin{bmatrix} Var(\epsilon_1) & Cov(\epsilon_1, \epsilon_2) & \dots & Cov(\epsilon_1, \epsilon_L) \\ Cov(\epsilon_1, \epsilon_2) & Var(\epsilon_2) & \dots & Cov(\epsilon_2, \epsilon_L) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\epsilon_1, \epsilon_L) & Cov(\epsilon_2, \epsilon_L) & \dots & Var(\epsilon_L) \end{bmatrix} \tag{3.5}$$

$$= \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} = \sigma^2 I_L. \tag{3.6}$$

The elements in $\mathbf{D}$ and $\mathbf{R}_i$ are known as variance components and can be written as,

$$\begin{bmatrix} \mathbf{u}_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_i \end{bmatrix} \right). \tag{3.7}$$

The distinction between the conditional and marginal mean of $\mathbf{Y}_i$ in the linear mixed-effects model is given by

$$\boldsymbol{E}[\mathbf{Y_i}|\mathbf{u_i}] = \boldsymbol{X_i\beta} + \boldsymbol{Z_i u_i} \tag{3.8}$$

and the conditional variance covariance of $\mathbf{Y}_i$ given $\mathbf{u}_i$ is

$$\begin{aligned} Var(\mathbf{Y}_i|\mathbf{u}_i) &= Var(\epsilon_i) \\ &= \mathbf{R}_i, \end{aligned} \tag{3.9}$$

hence,

$$\mathbf{Y}_i|\mathbf{u}_i \sim (\boldsymbol{X_i\beta} + \boldsymbol{Z_i u_i}, \mathbf{R}_i). \tag{3.10}$$

The marginal mean of $\mathbf{Y}_i$ when averaged over the distribution of random-effects $\mathbf{u}_i$ is

$$\begin{aligned} E(\mathbf{Y}_i) &= E[E(\mathbf{Y}_i|\mathbf{u}_i)] \\ &= E[\boldsymbol{X_i\beta} + \boldsymbol{Z_i u_i}] \\ &= \boldsymbol{X_i\beta} + \boldsymbol{Z_i}E(\mathbf{u}_i) \\ &= \boldsymbol{X_i\beta} \end{aligned} \tag{3.11}$$

and the marginal variance covariance of $\mathbf{Y}_i$ averaged over the distribution of $\mathbf{u}_i$ is,

$$\begin{aligned} Var(\boldsymbol{Y}_i) &= E[Var(\boldsymbol{Y_i}|\boldsymbol{u_i})] + \text{Var}[E(\boldsymbol{Y_i}|\boldsymbol{u_i})] \\ &= E(\mathbf{R}_i) + Var(\boldsymbol{X_i\beta} + \boldsymbol{Z_i u_i}) \\ &= \mathbf{R}_i + \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i' \\ &= \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i' + \mathbf{R}_i. \end{aligned} \tag{3.12}$$

Hence,

$$\boldsymbol{Y_i} \sim (\boldsymbol{X_i\beta}, \mathbf{Z}_i\mathbf{D}\mathbf{Z}_{\boldsymbol{i}}' + \mathbf{R_i}) \tag{3.13}$$

the observation $\mathbf{Y}_i$ and random-effects $\mathbf{u}_i$ have joint multivariate normal distribution

$$\begin{bmatrix} \mathbf{Y}_i \\ \mathbf{u}_i \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{X_i\beta} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i' + \mathbf{R}_i & \mathbf{Z}_i\mathbf{D} \\ \mathbf{D}\mathbf{Z}_i' & \mathbf{D} \end{bmatrix} \right). \tag{3.14}$$

3.2.2 Estimating fixed-effects for V known

Let $\mathbf{Y}$ be normally distributed and $\boldsymbol{V} = \mathbf{Z}\mathbf{D}\mathbf{Z}'+\mathbf{R}$ be the marginal variance-covariance of $\mathbf{Y}$ averages over the distribution of $\mathbf{u}_i$. Hence,

$$\mathbf{Y} \sim N(\boldsymbol{\mu} = \boldsymbol{X\beta}, \mathbf{V}) \tag{3.15}$$

which has a joint pdf

$$f(\mathbf{Y}) = \frac{1}{(2\pi)^{\frac{1}{2}}|\mathbf{V}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2}(\mathbf{Y}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{\mu}) \right\} \tag{3.16}$$

then the likelihood is given as,

$$L_{ML}(\boldsymbol{\theta}) = \prod_{i=1}^{L} \frac{1}{(2\pi)^{\frac{n}{2}}|\mathbf{V}|^{\frac{L}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{Y}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\right\}. \tag{3.17}$$

Where L is the dimension of $\mathbf{Y}$. Thus, the log likelihood function is given by,

$$l(\boldsymbol{\beta},\boldsymbol{\theta}) = -\frac{L}{2}\log 2\pi - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}(\mathbf{Y}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{\mu}). \tag{3.18}$$

We consider a general parameterization of $\boldsymbol{\mu}$ and $\mathbf{V}$, such that each element of $\boldsymbol{\mu}$ is a function of elements of a parameter vector $\boldsymbol{\theta}$, and similarly each element of $\mathbf{V}$ is a function of elements of a parameter vector $\boldsymbol{\alpha}$ which is unrelated to $\boldsymbol{\theta}$. Thus we write $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $\mathbf{V}$=$\mathbf{V}(\boldsymbol{\alpha})$.
By differentiating the log likelihood with respect to $\boldsymbol{\beta}$ and setting the results expression to zero. This leads to the fixed-effects which is expressed in terms of the variance parameters

$$\frac{\partial l}{\partial \boldsymbol{\theta}} = \frac{\partial \boldsymbol{\mu}'}{\partial \boldsymbol{\theta}}\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{\mu}). \tag{3.19}$$

If we make the following substitution in equation (3.19), let $\boldsymbol{\mu} = \boldsymbol{X\beta}$ and $\boldsymbol{\theta} = \boldsymbol{\beta}$ we obtain the following

$$\begin{aligned}
\frac{\partial l}{\partial \boldsymbol{\beta}} &= \frac{\partial(\boldsymbol{X\beta})'}{\partial \boldsymbol{\beta}}\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{X\beta}) \\
&= \frac{\partial \boldsymbol{\beta}'}{\partial \boldsymbol{\beta}}\mathbf{X}'\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{X\beta}) \\
&= \mathbf{X}'\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{X\beta}) \\
&= \mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} - \mathbf{X}'\mathbf{V}^{-1}\boldsymbol{X\beta}.
\end{aligned} \tag{3.20}$$

Equating $\frac{\partial l}{\partial \boldsymbol{\beta}}$ to zero we obtain the following results

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} - \mathbf{X}'\mathbf{V}^{-1}\boldsymbol{X\beta} = \mathbf{0}, \tag{3.21}$$

transposing $\mathbf{X}'\mathbf{V}^{-1}\boldsymbol{X}\boldsymbol{\beta}$ to right we obtain,

$$\mathbf{X}'\mathbf{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} \tag{3.22}$$

multiplying equation(3.22) by $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$,

$$(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} \tag{3.23}$$

we obtain

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} \tag{3.24}$$

and the variance of $\hat{\boldsymbol{\beta}}$ is obtained as

$$\begin{aligned} \text{Var}(\hat{\boldsymbol{\beta}}) &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\text{Var}(\mathbf{Y})\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{V}\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \end{aligned} \tag{3.25}$$

Since $\boldsymbol{\beta}$ is written as $\hat{\boldsymbol{\beta}}$, it means that $\hat{\boldsymbol{\beta}}$ is the best linear unbiased estimator(BLUE) of $\boldsymbol{\beta}$. Any generalized inverse $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}$ is used instead of $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$, if $\mathbf{X}$ is not full rank in-order to obtain the solution for $\boldsymbol{\beta}$. The solution obtained is not unique and is no longer unbiased. However, $\boldsymbol{X}\hat{\boldsymbol{\beta}}$ is unique and unbiased for $\boldsymbol{X}\boldsymbol{\beta}$, hence

$$\boldsymbol{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} \tag{3.26}$$

therefore $\boldsymbol{X}\boldsymbol{\beta}$ is the maximum likelihood estimator, and so as $\boldsymbol{\lambda}'\boldsymbol{X}\hat{\boldsymbol{\beta}}$ is the maximum likelihood estimator of $\boldsymbol{\lambda}'\boldsymbol{X}\boldsymbol{\beta}$ for any $\boldsymbol{\lambda}$.
Since the Var($\mathbf{Y}$)=$\mathbf{V}$ then

$$\text{Var}(\boldsymbol{X}\hat{\boldsymbol{\beta}}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1\prime}\mathbf{X}'. \tag{3.27}$$

It is Because $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1\prime}$ is a generalized inverse of $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})$, then the invariance property referred to Var($\boldsymbol{X}\hat{\boldsymbol{\beta}}$) reduces to

$$\text{Var}(\boldsymbol{X}\hat{\boldsymbol{\beta}}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'. \tag{3.28}$$

To test the null hypothesis $H_0 : \mathbf{K}'\mathbf{X}\boldsymbol{\beta} = \mathbf{m}$ where $\mathbf{K}'$ is of full row rank ($r_{\mathbf{K}} \leqslant r_{\mathbf{X}}$), we can derive a chi-square statistics using

$$X^2 = (\mathbf{K}'\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{m})' \left[\mathbf{K}'\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{K}\right]^{-1} (\mathbf{K}'\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{m}). \tag{3.29}$$

Under H_0, X^2 has a central χ^2 distribution with $r_{\mathbf{K}}$=rank($\mathbf{K}$) degrees of freedom. Typically $\mathbf{V}$ is a scalar multiple, therefore we can write $\mathbf{V}$ in terms of a weight matrix $\mathbf{W}$, which is the inverse of $\mathbf{V}$ up to a scalar multiple , that is , $\mathbf{V} = \sigma^2\mathbf{W}^{-1}$, where is assumed known. In such case the follow statistic can be derived as the likelihood ratio test and is also the most powerful invariant test

$$\mathrm{F} = \frac{(\mathbf{K}'\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{m})' \left[\mathbf{K}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-}\mathbf{X}'\mathbf{K}\right]^{-1} (\mathbf{K}'\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{m})}{r_{\mathbf{k}}\hat{\sigma}^2}, \tag{3.30}$$

where

$$\hat{\sigma}^2 = \frac{\mathbf{Y}' \left[\mathbf{W} - \mathbf{W}\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\right] \mathbf{Y}}{N - r_{\mathbf{X}}}. \tag{3.31}$$

Under the null hypothesis, F has an $F-$ distribution on $r_{\mathbf{k}}$ and N-$r_{\mathbf{X}}$ degrees of freedom. The null hypothesis is rejected at significant level α when F exceeds $F^{r_{\mathbf{k}}}_{N-r_{\mathbf{X}}}, 1 - \alpha$.

3.2.3 Predicting random-effects for V know

Suppose $\tilde{\mathbf{u}}$ represent an arbitrary predictor of $\mathbf{u}$, then $\tilde{\mathbf{u}}$ is unbiased if $E[\tilde{\mathbf{u}}] = E[\mathbf{u}]$ and $\tilde{\mathbf{u}}$ is the best predictor if it reach the minimum square error

$$E\left[(\tilde{\mathbf{u}} - \mathbf{u})'\mathbf{A}(\tilde{\mathbf{u}} - \mathbf{u})\right] = \iint (\tilde{\mathbf{u}} - \mathbf{u})'\mathbf{A}(\tilde{\mathbf{u}} - \mathbf{u}) f(\mathbf{u}, \mathbf{Y})\, d\mathbf{u}\, d\mathbf{Y}, \tag{3.32}$$

is a minimum where $\mathbf{A}$ is a positive definite symmetric matrix and $f(\mathbf{u}, \mathbf{Y})$ is the joint pdf of $\mathbf{u}$ and $\mathbf{Y}$. The best predictor of $\mathbf{u}$ is the conditional mean of $\mathbf{u}$ given $\mathbf{Y}$, $\tilde{\mathbf{u}} = \mathrm{Bp}(\mathbf{u}) = E(\mathbf{u}|\mathbf{Y})$.
Estimating the best predictor $\tilde{\mathbf{u}}$ requires some knowledge of the joint density of $\mathbf{u}$ and $\mathbf{Y}$.
Let

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{Y} \end{bmatrix} \sim N\left(\begin{bmatrix} E(\mathbf{u}) \\ E(\mathbf{Y}) \end{bmatrix}; \begin{bmatrix} \mathbf{D} & \mathbf{D}\mathbf{Z}' \\ \mathbf{Z}\mathbf{D} & \mathbf{V} \end{bmatrix}\right) \tag{3.33}$$

$$\begin{aligned}\tilde{\mathbf{u}} &= E(\mathbf{u}|\mathbf{Y}) + \mathbf{D}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{Y} - E(\mathbf{Y})) \\ &= \mathbf{D}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{Y} - \boldsymbol{X}\boldsymbol{\beta}).\end{aligned} \tag{3.34}$$

The resulting predictor is a best linear unbiased predictor (BLUP). If $\boldsymbol{\beta}$ is unknown we use $\hat{\boldsymbol{\beta}}$.

$$\begin{aligned}\tilde{\mathbf{u}} &= \mathbf{D}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{D}\mathbf{Z}'\mathbf{P}\mathbf{Y}\end{aligned} \tag{3.35}$$

and the variance is given as

$$\text{Var}(\tilde{\mathbf{u}}) = \mathbf{D}\mathbf{Z}'\mathbf{P}\mathbf{Z}\mathbf{D} \tag{3.36}$$

where $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})\mathbf{X}'\mathbf{V}^{-1}$
Note that

- $\text{Cov}(\hat{\boldsymbol{\beta}}, \hat{\mathbf{u}}) = 0$
- $\text{Var}(\tilde{\mathbf{u}} - \mathbf{u}) = \mathbf{D} - \mathbf{D}\mathbf{Z}'\mathbf{P}\mathbf{Z}\mathbf{D}$
-

$$\begin{aligned}\text{Var}(\mathbf{u}) &= \text{Var}[E(\mathbf{u}|\mathbf{Y})] + E[\text{Var}(\mathbf{u}|\mathbf{Y})] \\ &= \text{Var}(\tilde{\mathbf{u}}) + E[\text{Var}(\mathbf{u}|\mathbf{Y})]\end{aligned}$$

therefore

$$\text{Var}(\tilde{\mathbf{u}}) \leqslant \text{Var}(\mathbf{u}). \tag{3.37}$$

Hence $\tilde{\mathbf{u}}$ has a smaller mean squared error than other estimates based on assuming the random effects were fixed-effects. They also have less variability and are sometimes called shrinkage estimators as $\frac{\text{Cov}(\mathbf{Y},\mathbf{u})}{\text{Var}(\mathbf{Y})} \leqslant 1$

$$\begin{aligned}\tilde{\mathbf{u}} &= \frac{\text{Cov}(\mathbf{Y}, \mathbf{u})}{\text{Var}(\mathbf{Y})}(\mathbf{Y} - E(\mathbf{Y})) \\ &= \frac{\text{Cov}(\mathbf{Y}, \mathbf{u})}{\text{Var}(\mathbf{Y})}(\bar{\mathbf{Y}}_i - \bar{\mathbf{Y}})\end{aligned} \tag{3.38}$$

is shrunk compared to the corresponding fixed effects $\hat{\boldsymbol{\alpha}}_i = (\bar{\mathbf{Y}}_i - \bar{\mathbf{Y}})$.

3.2.4 Predicting random-effects for V unknown

We use $\hat{\mathbf{D}}$ and $\hat{\mathbf{V}}$, when $\mathbf{D}$ and $\mathbf{V}$ are unknown in equation(3.35), giving what could be called the estimated best predictor, to be denoted $\hat{\mathbf{u}}$

$$\begin{aligned}\hat{\mathbf{u}} &= \hat{\mathbf{D}}\mathbf{Z}'\hat{\mathbf{V}}^{-1}(\mathbf{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\ &= \hat{\mathbf{D}}\mathbf{Z}'\hat{\mathbf{P}}\mathbf{Y}.\end{aligned} \tag{3.39}$$

3.2.5 Maximum likelihood estimation

A method known for obtaining estimates of unknown parameters by optimizing a likelihood function, in general, is known as maximum likelihood estimation [79]. Based on distributional assumptions, in order to apply the maximum likelihood estimation specified in the model the likelihood is constructed as a function of the unknown parameters [79]. The likelihood function is defined using the density function of the observations and measures the likelihood of the model parameters given the data [83]. Further, the likelihood function is the product of the density functions for each observation when the observation are assumed independent [83]. Nevertheless, the likelihood function of a linear mixed-effects model needs to be based on a multivariate density function for the observation, since the observation are not independent [83].
Let the vector of all variance and covariance parameters denoted by $\boldsymbol{\alpha}$ found in $\mathbf{V}_i(\boldsymbol{\alpha}) = \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i' + \mathbf{R}_i$ and the vector of all parameters in the marginal model $\mathbf{Y}_i$ be $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\alpha}')$. Then the likelihood is given as

$$L_{ML}(\boldsymbol{\theta}) = \prod_{i=1}^{L}(2\pi)^{-\frac{L}{2}}|\mathbf{V}_i|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\mathbf{Y}_i - \boldsymbol{X_i}\boldsymbol{\beta})'\mathbf{V}_i^{-1}(\mathbf{Y}_i - \boldsymbol{X_i}\boldsymbol{\beta})\right\} \tag{3.40}$$

and the log-likelihood function $l(\boldsymbol{\theta})$ is given by

$$l(\boldsymbol{\theta}) = -\frac{L}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{V}_i| - \frac{1}{2}\sum_{i=1}^{L}(\mathbf{Y}_i - \boldsymbol{X_i}\boldsymbol{\beta})'\mathbf{V}_i^{-1}(\mathbf{Y}_i - \boldsymbol{X_i}\boldsymbol{\beta}). \tag{3.41}$$

The maximum likelihood estimator of $\boldsymbol{\beta}$, obtained by optimizing the likelihood function which is given by

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\alpha}) = \left(\sum_{i=1}^{L}\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i\right)^{-1}\sum_{i=1}^{L}\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{Y}_i \tag{3.42}$$

$\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}(\boldsymbol{\alpha})$ follows a multivariate normal distribution with mean vector $\boldsymbol{\beta}$ which is

given by as

$$\begin{aligned} \mathrm{E}[\hat{\boldsymbol{\beta}}(\boldsymbol{\alpha})] &= \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \mathrm{E}(\boldsymbol{Y}_i) \\ &= \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i \boldsymbol{\beta} \\ &= \boldsymbol{\beta}. \end{aligned} \tag{3.43}$$

Provide that $\mathrm{E}(\boldsymbol{Y}_i) = \boldsymbol{X}_i\boldsymbol{\beta}$. It is sufficient that the mean of the response is correctly specified, in order for $\hat{\boldsymbol{\beta}}$ to be unbiased. Hence the variance-covariance of $\boldsymbol{\beta}$ is

$$\begin{aligned} \mathrm{Var}(\hat{\boldsymbol{\beta}}) &= \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \mathrm{Var}(\boldsymbol{Y}_i) \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \\ &= \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \mathbf{V}_i \mathbf{V}_i^{-1} \boldsymbol{X}_i\right) \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \\ &= \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right) \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1} \\ &= \left(\sum_{i=1}^{L} \boldsymbol{X}_i' \mathbf{V}_i^{-1} \boldsymbol{X}_i\right)^{-1}. \end{aligned} \tag{3.44}$$

Since, the covariance matrix $\mathrm{V}(\boldsymbol{Y}_i)$ is correctly modelled as $\mathbf{V}_i = \boldsymbol{Z}_i \mathbf{D} \boldsymbol{Z}_i' + \boldsymbol{R}_i$. Therefore this is called a sandwich estimator $\mathbf{V}(\boldsymbol{\beta})$, it obtained by replacing $\mathrm{Var}(\boldsymbol{Y}_i) = \mathbf{V}_i = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' + \boldsymbol{R}_i$. In practice, the covariance matrix $\mathbf{V}(\boldsymbol{\beta})$ is estimated by replacing $\boldsymbol{\alpha}$ by it ML or REML estimator. This approach would yield an estimator for the covariance matrix of $\hat{\boldsymbol{\beta}}$ which would take into account the extra variability. The inverse fisher information matrix can be used to obtain the standard errors.

3.2.6 Estimating fixed-effects for V unknown

The log likelihood function has to be maximized with respect to $\mathbf{V}$, since $\mathbf{V}$ is unknown but not being a function of $\boldsymbol{\beta}$. For $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$

$$l(\boldsymbol{\beta}, \boldsymbol{\theta}) = -\frac{L}{2}\log 2\pi - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}(\mathbf{Y} - \boldsymbol{\mu})' \mathbf{V}^{-1} (\mathbf{Y} - \boldsymbol{\mu}). \tag{3.45}$$

The maximum likelihood equations for $\mathbf{V}$ are obtained from equating to $\mathbf{0}$ the fol-

lowing expression

$$\frac{\partial l}{\partial \boldsymbol{\alpha_k}} = -\frac{1}{2}\left[\text{trace}(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}) - (\mathbf{Y}-\boldsymbol{\mu})'\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\right]. \tag{3.46}$$

Using $\boldsymbol{\alpha}$ for each parameter in $\mathbf{V}$. In doing this $\boldsymbol{\mu}$ is replaced by $\mathbf{X}\hat{\boldsymbol{\beta}}$.
Since $\boldsymbol{\beta}$ is a function of $\mathbf{V}$. Hence we evaluate the profile likelihood for $\mathbf{V}$ denoted by l_p, which is the likelihood for a given value of $\mathbf{V}$ with the maximizing value of $\boldsymbol{\beta}$ for that $\mathbf{V}$ is inserted

$$l_p = -\frac{1}{2}\mathbf{YPY} - \frac{1}{2}\log|\mathbf{V}| - \frac{L}{2}\log(2\pi) \tag{3.47}$$

where

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}. \tag{3.48}$$

To obtain the maximum likelihood estimate $\boldsymbol{\alpha_k}$, we set $\frac{\partial l}{\partial \boldsymbol{\alpha_k}}$ to zero.
The iterative methods must be use in order to obtain the estimate of $\boldsymbol{\alpha_k}$. The conventional optimization methods which require first and second derivatives may be applied.
Information matrix

$$\begin{aligned}\frac{\partial^2 l}{\partial \boldsymbol{\theta}\boldsymbol{\theta}'} &= \frac{\partial}{\partial \boldsymbol{\theta}}\left(\frac{\partial l}{\partial \boldsymbol{\theta}'}\right) = \frac{\partial}{\partial \boldsymbol{\theta}}\left(\left[\frac{\partial \boldsymbol{\mu}'}{\partial \boldsymbol{\theta}}\mathbf{V}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\right]'\right)\\ &= \frac{\partial}{\partial \boldsymbol{\theta}}\left[(\mathbf{Y}-\boldsymbol{\mu})'\mathbf{V}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}'}\right]\\ &= -\frac{\partial \boldsymbol{\mu}'}{\partial \boldsymbol{\theta}}\mathbf{V}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}'} + (\mathbf{Y}-\boldsymbol{\mu})'\mathbf{V}^{-1}\frac{\partial^2 \boldsymbol{\mu}}{\partial \theta_i \partial \boldsymbol{\theta}'}\end{aligned} \tag{3.49}$$

and so

$$-E(\frac{\partial^2 l}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}}) = \frac{\partial \boldsymbol{\mu}'}{\partial \boldsymbol{\theta}}\mathbf{V}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}}. \tag{3.50}$$

Also

$$\begin{aligned}\frac{\partial^2 l}{\partial \boldsymbol{\alpha_k}\partial \boldsymbol{\theta}'} &= \frac{\partial}{\partial \boldsymbol{\alpha_k}}\left[(\mathbf{Y}-\boldsymbol{\mu})'\mathbf{V}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}'}\right] = (\mathbf{Y}-\boldsymbol{\mu})'\frac{\partial \mathbf{V}^{-1}}{\partial \boldsymbol{\alpha_k}}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}'}\\ &= -(\mathbf{Y}-\boldsymbol{\mu})'\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}\mathbf{V}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}'},\end{aligned} \tag{3.51}$$

since $E[\mathbf{Y}-\boldsymbol{\mu}]=\mathbf{0}$

$$-E\left[\frac{\partial^2 l}{\partial \boldsymbol{\alpha_k}\partial \boldsymbol{\theta}'}\right]=\mathbf{0}. \tag{3.52}$$

If we differentiating equation(3.46) with respect to $\boldsymbol{\alpha_s}$, we obtain

$$\begin{aligned}\frac{\partial^2 l}{\partial \boldsymbol{\alpha_s}\partial \boldsymbol{\alpha_k}} = & \\ & -\frac{1}{2}\{\text{trace}\left(-\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_s}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}+\mathbf{V}^{-1}\frac{\partial^2 \mathbf{V}}{\partial \boldsymbol{\alpha_s}\partial \boldsymbol{\alpha_k}}\right) \\ & +(\mathbf{Y}-\boldsymbol{\mu})'\Big[(-1)\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_s}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}\mathbf{V}^{-1}+\mathbf{V}^{-1}\frac{\partial^2 \mathbf{V}}{\partial \boldsymbol{\alpha_s}\partial \boldsymbol{\alpha_k}}\mathbf{V}^{-1} \\ & -\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_s}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}\mathbf{V}^{-1}\Big](\mathbf{Y}-\boldsymbol{\mu})\}\end{aligned} \tag{3.53}$$

.

Now for any $\mathbf{A}$
$E[(\mathbf{Y}-\boldsymbol{\mu})'\mathbf{A}(\mathbf{Y}-\boldsymbol{\mu})]=\text{trace}\{\mathbf{A}E[(\mathbf{Y}-\boldsymbol{\mu})(\mathbf{Y}-\boldsymbol{\mu})']\}=\text{trace}(\mathbf{A}V)$
therefore

$$\begin{aligned}-E\left(\frac{\partial^2 l}{\partial \boldsymbol{\alpha_s}\partial \boldsymbol{\alpha_k}}\right) &= \frac{1}{2}\{\text{trace}\left(-\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_s}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}+\mathbf{V}^{-1}\frac{\partial^2 \mathbf{V}}{\partial \boldsymbol{\alpha_s}\partial \boldsymbol{\alpha_k}}\right) \\ &\quad +\text{trace}\left[\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_s}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}-\mathbf{V}^{-1}\frac{\partial^2 \mathbf{V}}{\partial \boldsymbol{\alpha_s}\partial \boldsymbol{\alpha_k}}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_s}}\right]\} \\ &= \frac{1}{2}\text{trace}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_s}}\right).\end{aligned} \tag{3.54}$$

Hence, the information matrix is

$$-E\begin{bmatrix} \frac{\partial^2 l}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} & \frac{\partial^2 l}{\partial \boldsymbol{\theta}\partial \boldsymbol{\alpha}'} \\ \left(\frac{\partial^2 l}{\partial \boldsymbol{\theta}\partial \boldsymbol{\alpha}'}\right)' & \frac{\partial^2 l}{\partial \boldsymbol{\alpha}\partial \boldsymbol{\alpha}'} \end{bmatrix}$$

$$=-E\begin{bmatrix} \frac{\partial \boldsymbol{\mu}'}{\partial \boldsymbol{\theta}}\mathbf{V}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}\{\text{trace}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \boldsymbol{\alpha_k}}\right)\} \end{bmatrix}.$$

Newton Raphson method :
Consider the log-likelihood function $l(\boldsymbol{\alpha})$ for which we want to find the maximum at $\boldsymbol{\alpha}$. The Newton-Raphson method uses the first-order expansion of the score function around the current estimate $\boldsymbol{\alpha_{(t)}}$ to produce the next estimate $\boldsymbol{\alpha_{(t+1)}}$. Hence

each Newton-Raphson iteration requires the calculation of the score function and it's derivatives

$$\frac{\partial l}{\partial \boldsymbol{\alpha_{(k)}}} = \frac{\partial l}{\partial \boldsymbol{\alpha_{(0)}}} + \frac{\partial^2 l}{\partial \boldsymbol{\alpha_{(k)}} \partial \boldsymbol{\alpha_{(k)}}'} \left(\boldsymbol{\alpha_{(k)}} - \boldsymbol{\alpha_{(0)}}\right). \tag{3.55}$$

Setting $\frac{\partial l}{\partial \boldsymbol{\alpha}}$ to zero, and solving we obtained

$$\frac{\partial l}{\partial \boldsymbol{\alpha_{(0)}}} + \frac{\partial^2 l}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'} \left(\boldsymbol{\alpha} - \boldsymbol{\alpha_{(0)}}\right) = \mathbf{0}, \tag{3.56}$$

re-arranging equation(3.56) to obtain

$$\begin{aligned} \frac{\partial^2 l}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'} \boldsymbol{\alpha} &= \frac{\partial^2 l}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'} \boldsymbol{\alpha_{(0)}} - \frac{\partial l}{\partial \boldsymbol{\alpha_{(0)}}} \\ \boldsymbol{\alpha} &= \boldsymbol{\alpha_{(0)}} - \left[\frac{\partial^2 l}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'}\right]^{-1} \frac{\partial l}{\partial \boldsymbol{\alpha_{(0)}}}. \end{aligned} \tag{3.57}$$

Therefore the estimate of the maximum on the $(t+1)th$ iteration can be obtained iteratively by this equation

$$\boldsymbol{\alpha_{(t+1)}} = \boldsymbol{\alpha_{(t)}} - \left[\frac{\partial^2 l}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'}\right]^{-1} \frac{\partial l}{\partial \boldsymbol{\alpha_{(t)}}}. \tag{3.58}$$

Fisher Scoring algorithm :
The Fisher Scoring algorithm replaces the observed information matrix by the expected information matrix

$$\boldsymbol{\alpha_{(t+1)}} = \boldsymbol{\alpha_{(t)}} - E\left[\frac{\partial^2 l}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'}\right]^{-1} \frac{\partial l}{\partial \boldsymbol{\alpha_{(t)}}} \tag{3.59}$$

3.2.7 Restricted maximum likelihood estimation

Optimizing the likelihood function we obtain the restricted maximum likelihood estimators(REML) for the variance components of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ [84]. Moreover Harville

[85] has shown that the likelihood function of the error contrasts can be written as

$$\begin{aligned} L(\boldsymbol{\alpha}) &= (2\pi)^{-\frac{(L-p)}{2}} |\sum_{i=1}^{L} \mathbf{X}_i' \mathbf{X}_i|^{\frac{1}{2}} \\ &\times |\sum_{i=1}^{L} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i|^{-\frac{1}{2}} \prod_{i=1}^{L} |\mathbf{V}_i|^{-\frac{1}{2}} \\ &\times \exp\left\{ -\frac{1}{2} \sum_{i=1}^{L} (\mathbf{Y}_i - \boldsymbol{X_i}\hat{\boldsymbol{\beta}})' \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{X_i}\hat{\boldsymbol{\beta}}) \right\}. \end{aligned} \tag{3.60}$$

Where $\hat{\boldsymbol{\beta}}$ is given by $\hat{\boldsymbol{\beta}}(\boldsymbol{\alpha}) = \left(\sum_{i=1}^{L} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{L} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{Y}_i$,
finally the likelihood function is equals

$$L(\boldsymbol{\alpha}) = C |\sum_{i=1}^{L} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i|^{-\frac{1}{2}} L_{ML}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}). \tag{3.61}$$

Where C is a constant not depending on $\boldsymbol{\alpha}$ and $L_{ML}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}) = L_{ML}(\boldsymbol{\theta})$. The REML estimator for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can also be found by maximizing the REML likelihood function

$$L_{REML}(\boldsymbol{\theta}) = |\sum_{i=1}^{L} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i|^{-\frac{1}{2}} L_{ML}(\boldsymbol{\theta}) \tag{3.62}$$

with respect to all parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ simultaneously.

3.3 Inference

3.3.1 The likelihood ratio test(LR)

The likelihood ratio test is based on comparing the values of likelihood functions for two nested models, which are full and reduced models. When two models are nested we mean that the parameter space for the reduced model is a subspace of the full model. Furthermore, the parameters in the reduced model can be obtained by imposing certain constraints on the parameters of the full model. The reduced model is denoted as $\hat{l}_{reduced}$ and the full model as $\hat{l}_{full}$, where the likelihood ratio test is given by

$$LR = -2 l_L \lambda_L = 2(\hat{l}_{full} - \hat{l}_{reduced}) \sim \chi^2_{df}. \tag{3.63}$$

The likelihood ratio test statistic follows a χ^2 distribution, in which the number of degrees of freedom df, is obtained by subtracting the number of parameters in the

reduced model from the number of parameters in the full model. If the difference between the reduced model and full model is sufficiently large, there is evidence against the reduced model in favor of the full model. However if the difference is small, we have evidence in favor of the reduced model.

3.3.2 Wald test

A wald test is often used for fixed-effects parameters. If we consider an hypotheses about $\boldsymbol{\beta}$. For each parameters $\boldsymbol{\beta}_j$ in $\boldsymbol{\beta}$, $j, \cdots, p$. Therefore we can test the hypothesis $H_0 : \boldsymbol{\beta}_j = 0$ against $H_a : \boldsymbol{\beta}_j \neq 0$. The corresponding wald test is calculated as follows

$$Z = \frac{\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j}{Se(\hat{\boldsymbol{\beta}}_j)}. \tag{3.64}$$

Suppose that L is a single row vector then $L\text{Cov}(\hat{\boldsymbol{\beta}})L'$ is a single value and its square roots provides an estimates of the standard error for $L\hat{\boldsymbol{\beta}}$. Hence, an approximate 95% confidence interval is given by

$$L\hat{\boldsymbol{\beta}} \pm 1.96\sqrt{L\text{Cov}(\hat{\boldsymbol{\beta}})L'}. \tag{3.65}$$

The hypotheses test for the estimates is given by

$$H_0 : L\boldsymbol{\beta} = 0 \quad \text{vs} \quad H_a : L\boldsymbol{\beta} \neq 0, \tag{3.66}$$

where $\hat{\boldsymbol{\beta}}$ is asymptotically normal with mean $\boldsymbol{\beta}$ and covariance matrix, for any know matrix L. Then the wald statistic is given by

$$Z = \frac{L\hat{\boldsymbol{\beta}}}{\sqrt{L\text{Cov}(\hat{\boldsymbol{\beta}})L'}}. \tag{3.67}$$

If the Z is a standard normal random variable then Z^2 has a χ^2 distribution with one degree of freedom, hence

$$W = \hat{\boldsymbol{\beta}}' L' \left[L\text{Cov}(\hat{\boldsymbol{\beta}})L'\right] L\hat{\boldsymbol{\beta}}. \tag{3.68}$$

W follows an asymptotic χ^2 distribution with rank L degrees of freedom. However wald test do not take into account the variability from replacing $\boldsymbol{\alpha}$ by some estimate.

Hence in sufficiently large samples, wald test will only provide valid inferences. This is often resolved by replacing the χ^2 distribution by an appropriate F distribution for testing hypotheses about $\boldsymbol{\beta}$. In other words, an F-statistic is an alternative test statistic to the wald test. The F-test statistic for the null hypothesis H_0 stated above is given by

$$F = \frac{\hat{\boldsymbol{\beta}}' L' \left[L\text{Cov}(\hat{\boldsymbol{\beta}}) L' \right]^{-1} L\hat{\boldsymbol{\beta}}}{rank(L)}. \tag{3.69}$$

The numerator degrees of freedom of the F statistic above are equal to rank L and denominator degrees of freedom are estimated from the data. Such methods as containment, Satterthwaite approximation, Kenward and Roger approximation are use to estimate the denominator degree of freedom for the F-statistic above. The p-values for all methods are the same when estimating the degree of freedom. As result of the assumption for most application in longitudinal data that different individual contribute independent information. Which result in numbers of degrees of freedom which are large [84]. The F-test reduces to a t-test for univariate hypotheses and the rank L=1.

3.3.3 Estimating the random-effects

It is often useful to calculate estimates for the random-effects $\mathbf{u}_i$, since they reflect how much the subject specific profiles deviate from the overall average profile. However, in practice one is usually primarily interested in estimating the parameters in the marginal linear mixed-effects model, the fixed-effects $\boldsymbol{\beta}$ and the variance components $\mathbf{D}$ and σ^2. Since such estimates can be interpreted as residuals which may be helpful for detecting outlying individuals who are behaving differently over time. Furthermore, whenever the interest is in prediction of subject specific evolutions, estimates for the random-effects are needed [84].

It is no longer sufficient to assume that the data can be represented well by the marginal model $\text{N}(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{V}_i)$. We assume conditional interpretation since $\mathbf{Y}_i|\mathbf{u}_i \sim N(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i, \mathbf{D})$. It is because random-effects represent natural heterogeneity between the subjects. The justification of this assumption is when the between subjects variability is large in comparison to the within subject variability. Therefore it is most natural to estimate them using Bayesian techniques. To explore the inference for random-effects. We denote density function of $\mathbf{Y}_i$ conditional on $\mathbf{u}_i$ and the prior density function of $\mathbf{u}_i$ by $f(\mathbf{y}_i|\mathbf{u}_i)$ and $f(\mathbf{u}_i)$ respectively. Therefore the posterior

density function of $\mathbf{u}_i$ given $\mathbf{Y}_i = \mathbf{y}_i$ is given by

$$\begin{aligned} f(\mathbf{u}_i|\mathbf{y}_i) &\equiv f(\mathbf{u}_i|\mathbf{Y}_i = \mathbf{y}_i) \\ &= \frac{f(\mathbf{y}_i|\mathbf{u}_i)f(\mathbf{u}_i)}{\int f(\mathbf{y}_i|\mathbf{u}_i)f(\mathbf{u}_i)d\mathbf{u}_i}. \end{aligned} \tag{3.70}$$

The estimate of $\mathbf{u}_i$ is often the mean of the posterior distribution. The estimate is given by

$$\begin{aligned} \hat{\mathbf{u}}_i(\boldsymbol{\theta}) &= E(\mathbf{u}_i|\mathbf{Y}_i = \mathbf{y}_i) \\ &= \int \mathbf{u}_i f(\mathbf{u}_i|\mathbf{y}_i)d\mathbf{u}_i \\ &= \mathbf{D}\mathbf{Z}_i'\mathbf{V}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}) \end{aligned} \tag{3.71}$$

and the covariance matrix of the corresponding estimator equals

$$\text{Var}(\hat{\mathbf{u}}_i) = \mathbf{D}\mathbf{Z}_i' \left[\mathbf{V}_i^{-1} - \mathbf{V}_i^{-1}\mathbf{X}_i \left(\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i \right)^{-1} \mathbf{X}_i'\mathbf{V}_i^{-1} \right] \mathbf{Z}_i\mathbf{D}. \tag{3.72}$$

Nonetheless inference for $\mathbf{u}_i$ should account for the variability in $\mathbf{u}_i$. Therefore the inference for $\mathbf{u}_i$ is usually based on

$$\text{Var}(\hat{\mathbf{u}}_i(\boldsymbol{\beta}) - \mathbf{u}_i) = \mathbf{D} - \text{Var}(\hat{\mathbf{u}}_i(\boldsymbol{\beta})). \tag{3.73}$$

The estimates ML or REML obtained from fitting the marginal model can be used to replace the parameter in $\boldsymbol{\theta}$. The empirical Bayes estimate for $\mathbf{u}_i$ is $\hat{\mathbf{u}}_i(\hat{\boldsymbol{\theta}})$. The inference for $\mathbf{u}_i$ is often based on approximate t-test or F-test, which are similar to fixed-effect than the wald test. The following inequality hold for any linear combination $\lambda\mathbf{u}_i$ of the random-effects

$$\text{Var}(\lambda'\hat{\mathbf{u}}_i) \leq \text{Var}(\lambda'\mathbf{u}_i). \tag{3.74}$$

It follows that the empirical Bayes estimate show less variability than actually present in the random-effects population [84]. This is often called shrinkage estimates.
The linear combination of fixed-effects in $\boldsymbol{\beta}$ and random-effects in $\mathbf{u}_i$ are often parameters of interest. For example the subject specific slope in the sum of the average slope for subjects with the same covariate values, and the subject specific random

slope for that subject. Hence

$$\mu = \lambda'_{\boldsymbol{\beta}}\boldsymbol{\beta} + \lambda'_{\mathbf{u}}\hat{\mathbf{u}}_i \tag{3.75}$$

is of interest then

$$\mu = \lambda'_{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}} + \lambda'_{\mathbf{u}}\hat{\mathbf{u}}_i. \tag{3.76}$$

$\mathbf{Y}_i$ is unbiased for $\mathbf{u}$, therefore $\hat{\mu}$ is the best linear unbiased predictor(BLUP) and has minimum variance among all unbiased linear estimators. Consider the prediction of the evolution of the ith subject then

$$\begin{aligned}
\hat{\mathbf{Y}}_i &= \mathbf{X}_i\hat{\boldsymbol{\beta}} + \mathbf{Z}_i\mathbf{u}_i \\
&= \mathbf{X}_i\hat{\boldsymbol{\beta}} + \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i'\mathbf{V}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}) \\
&= (I_{ni} - \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i'\mathbf{V}_i^{-1})\mathbf{X}_i\hat{\boldsymbol{\beta}} + \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i'\mathbf{V}_i^{-1}\mathbf{y}_i \\
&= \sum_i \mathbf{V}_i^{-1}\mathbf{X}_i\hat{\boldsymbol{\beta}} + (I_{ni} - \sum_i \mathbf{V}_i^{-1})\mathbf{y}_i.
\end{aligned} \tag{3.77}$$

The weights of $\mathbf{Y}_i$ are $\hat{\sum}_i \mathbf{V}_i^{-1}$ and $I_{Li} - \hat{\sum}_i \mathbf{V}_i^{-1}$ respectively and $\mathbf{Y}_i$ is defined as the weighted mean for the population average profile $\mathbf{X}_i\hat{\boldsymbol{\beta}}$ and the observed data $\mathbf{y}_i$. The regression coefficient in a random-effects model have a subject specific interpretation. If the within subject variability is high then the bigger weight goes to the overall population mean, while more weight is given to $\mathbf{y}_i$ if between subject variability is large.

3.4 Model selection

In this section, we overview statistical methods used to compare models in the analysis.

3.4.1 Akaike information criteria

Information criteria are another set of tools, that are widely used to compare and select models. Akaike's Information Criteria(AIC) can be used to compare non-nested models. Akaike [86] proposed this method, which is given by

$$, AIC = -2\log L + 2p \tag{3.78}$$

where L is the maximum log-likelihood and p is the number of estimable parameters in the model. If the AIC is smaller, then the model is preferably.

3.4.2 Schwarz criterion

Schwarz Criterion(SC) also known as Bayesian Information Criterion(BIC). The Schwarz Criterion has the same properties as AIC, since they both comparing the non-nested models. However, the AIC aims to find the best approximating model on the true one. While the BIC aims to identify the true model. The method was proposed by Schwarz [87] and is given by

$$SC = -2\log L + p\log(n), \tag{3.79}$$

where n is the sample size and p is the number of estimable parameters in the model. According to Allison [88], the SC produces more critical penalization on the likelihood for estimating more parameters than AIC. The model is preferable if it has the smaller SC than the other model with bigger SC. We can lower the options before comparing model, while doing a model selection. The selection procedure of a variable that enters the model, such as forward, backwards and stepwise selection can be done by building the regression model. Forward selection starts with the reduced model and enters one covariate at a time, that is found to be significant at some level of significance $\boldsymbol{\alpha}$ until all significant variables are added to the model. Backward selection starts with the full model that contains all covariates and drops one at a time, that is, insignificant at some level of significance $\boldsymbol{\alpha}$. This is done until all non-significant variables are removed from the model. The stepwise selection works the same way as the forward selection procedure. Nevertheless, in the stepwise procedure the variables that are already in the model are considered to be excluded in the model each time the new covariate is added in the model, which is the advantage over the forward selection procedure. The stepwise procedure is preferable, in cases where there many covariates since it minimizes the chance of keeping the variables that are no longer needed in the model and leaving out some important ones.

3.5 Checking model assumption(diagnostics)

Model diagnostic is important after fitting a linear mixed-effects model and before making any inference based upon it. In order to check whether distributional assumptions for the residuals are met and whether the fit of the model is sensitive to unusual observations [79]. The aim of checking the distributional assumption is based on the estimated residual errors. Due to the presence of random-effects and different covariance structures, diagnostics for linear mixed-effects model are more difficult to perform and interpret, because of the model itself which is more complex [79]. Throughout the analysis of longitudinal data set, model diagnostics should be

part of the model building process [79].

3.5.1 Residual diagnostics

A residual is defined as the difference between an observed quantity and its estimated value. In a linear model the set of residual plotted against predicted values are used to decide whether, they presents a random pattern or not [79]. The use of residuals against fitted plots plays an important role in verifying model assumptions and to detect outliers and potentially influential observations [79].

3.5.2 Conditional and marginal residuals

The difference between the observed value and the conditional predicted value of the dependent variable is defined as the conditional residuals [89]. Conditional residuals are not well suited for verifying model assumptions and detecting outliers. Even if the true model residuals are uncorrelated and have equal variance, conditional residuals will tend to be correlated and their variances may be different for different subgroups of individuals. The conditional residuals vector equation for a given individual i in a longitudinal data set, is defined as follows

$$\hat{\boldsymbol{\epsilon}}_{(c)\boldsymbol{i}} = \mathbf{Y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}} - \mathbf{Z}_i\hat{\mathbf{u}}_i. \tag{3.80}$$

On the other hand, a marginal residual is defined as the difference between the observed data and the estimated marginal mean [89]. The marginal residuals are in contrast to conditional residuals, since the marginal residuals are based on model which do not include random-effects [79]. The marginal residuals in vector equation are defined as follows

$$\hat{\boldsymbol{\epsilon}}_{(\boldsymbol{m})\boldsymbol{i}} = \mathbf{Y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}. \tag{3.81}$$

The use of raw residuals is to check heterogeneity of the conditional or marginal variance. However, they are less recommended for checking normality assumptions and detecting outlying observations. As a result, the raw residuals will be correlated and their variances will differ.

3.5.3 Standardized and studentized residuals

To avoid the interpretation to depend on the measurement units of the dependent variable, we consider scaling the residuals by dividing by the true or estimated standard deviations. To obtain standardized residuals, it would be preferable to scale the

residuals by their true standard deviations. However, the true standard deviation is rarely known in practice, hence the estimated standard deviation is use to do the scaling instead. Therefore, the residuals that are obtained in this manner are called studentized residuals. The resulting residuals when scaling residuals by dividing them by the estimated standard deviation of the dependent variable are called Pearson residuals. If we assume that variability of $\hat{\beta}$ can be ignored then the Pearson-type scaling is appropriate. We do not consider other scaling choices, although they are possible. The calculation of a studentized residuals may also depend on whether the observation corresponding to the residuals in question is included in the estimation of the standard deviation or not. We refer to it as external studentization, if the observation is excluded.

3.5.4 Influence diagnostics

Maximum likelihood and restricted maximum likelihood method are both sensitive to unusual observations. To identify observations that heavily influence estimates of the parameters in either β or θ, one need formal techniques such as influence diagnostics. Influence diagnostics for linear mixed-effects model is an active area of research. The idea of influence diagnostics for a given observation or subset of observations is to quantify the effect of excluding those observation on the results of the analysis of the entire data set. Influence diagnostics may be used to investigate various aspects of the model fit. The influence of observations on the model fit can clear itself in more varied and complicated ways since linear mixed-effects models are more complicated than standard linear models. In linear mixed-effects model, it is generally recommended to follow a top-down approach when carrying out influence diagnostics. First, check overall influence diagnostics. Assuming that there are influential sets of observation based on the overall influence diagnostics, proceed with other diagnostics to see what aspect of the model a given subset of observations affects: fixed-effects, covariance parameters, the precision of the parameters estimates, or predicted values. Influence diagnostics plays an important role in the interpretation of the result. It is appropriate to interpret the model with respect to prediction, if a given subset of data has a strong influence on the estimates of covariance parameters, but limited impact on the fixed-effects. The precision may be affected by the changes in estimates of covariance parameters to test for fixed-effects and confidence intervals.

3.5.5 Overall influence

An overall influence statistics measures the change in the objective function being minimized. The likelihood distance and likelihood displacement are examples of

an overall influence measure. In the linear mixed-effects model fitted by maximum likelihood and restricted maximum likelihood. Compute the full data parameter estimates $\hat{\psi}$ and estimates based on the reduced data $\hat{\psi}_{(U)}$. Then the likelihood and restricted likelihood distances are obtained as

$$\begin{aligned} LD(U) &= 2\{l(\hat{\psi}) - l(\hat{\psi}_{(U)})\} \\ RLD(U) &= 2\{l_R(\hat{\psi}) - l_R(\hat{\psi}_{(U)})\}. \end{aligned} \tag{3.82}$$

The likelihood distance gives the amount by which the log-likelihood of the full data changes if one were to evaluate it at the reduced data estimates. The important point to note is that $l(\hat{\psi}_{(U)})$ is not the likelihood function obtained by fitting the model to the reduced data set. It is obtained by evaluating the likelihood function based on the full data set containing all n observations at the reduced data estimates.
The likelihood distance is a summary measure expressing the joint influence of the observations in the set U on all parameters in ψ that were subject to updating. Determine the nature of that influence, if the summary measure suggests that the point in U are influential. In particular, the points can affect.

- The estimates of fixed-effects.
- The estimates of the precision of the fixed-effects.
- The estimates of the covariance parameters.
- The estimates of the precision of the covariance parameters.
- fixed and predicted values.

To determine whether data points are actually troublesome it is important to further decompose the initial finding. Simply because they are influential somehow, should not trigger their removal from the analysis or a change in the model. For example, if points primarily affect the precision of the covariance parameters without exerting much influence on the fixed-effects, then their presence in the data may not violate the hypothesis tests or confidence intervals about $\boldsymbol{\beta}$. They will only do so if your inference depends on a estimate of the precision of the covariance parameter estimates.

3.5.6 Change in parameter estimates

The Mixed procedure is statistical command in SAS which fits a variety of mixed linear models to data and enable you to use these fitted model to make statistical inferences about the data. In this context its enables you to compute summary statistics

that capture the change in the entire parameter vector. Since the number of fixed-effects and covariance parameters can be large. These quadratic forms are based on Coook's D and multivariate DFFITS statistics. The multivariate DFFITS use an externalized estimate of the variance of the parameter estimates, while Cook's D does not. For the fixed-effects, the two statistics are

$$\begin{aligned} \text{D}(\boldsymbol{\beta}) &= \frac{(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(\boldsymbol{U})})'\widehat{\text{Var}}(\hat{\boldsymbol{\beta}}_{(\boldsymbol{U})})(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(\boldsymbol{U})})}{\text{rank}(\mathbf{X})} \\ \text{MDFFITS}(\boldsymbol{\beta}) &= \frac{(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(\boldsymbol{U})})'\widehat{\text{Var}}(\hat{\boldsymbol{\beta}}_{(\boldsymbol{U})})(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(\boldsymbol{U})})}{\text{rank}(\mathbf{X})}. \end{aligned} \tag{3.83}$$

Where $(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(\boldsymbol{U})})$, is the difference between the two $p \times 1$ vectors. Large values indicates that the change in the parameter estimate is large, for both statistics. The D($\boldsymbol{\theta}$) and MDFFITS($\boldsymbol{\theta}$) do not involve division by a rank matrix, if the covariance parameters are updated during influence analysis.

3.5.7 Change in precision of estimates

The effect on the point estimate is separate from the effect on the precision of estimates. The hypothesis test and confidence intervals can be affect by the data points that have a small Cook's D. If their influence on the precision of the estimates is large. The trace and determinants of the variance matrices based on the full data and the reduced data estimates are given as follows

$$\begin{aligned} \text{Covtrace}(\boldsymbol{\beta}) &= |\text{trace}(\hat{\text{Var}}(\hat{\boldsymbol{\beta}}))^{-1}\hat{\text{Var}}(\hat{\boldsymbol{\beta}}_{(U)}) - \text{rank}(\mathbf{X})| \\ \text{CovRatio} &= \frac{|\hat{\text{Var}}(\hat{\boldsymbol{\beta}}_{(U)})|}{|\hat{\text{Var}}(\hat{\boldsymbol{\beta}})|}. \end{aligned} \tag{3.84}$$

The benchmarks of no influence are zero for the covariance trace and one for the covariance ratio. If the influence analysis updates the covariance parameters, the Mixed procedure computes similar statistics for $\boldsymbol{\theta}$

$$\begin{aligned} \text{Covtrace}(\boldsymbol{\theta}) &= |\text{trace}(\hat{\text{Var}}(\hat{\boldsymbol{\theta}}) - \hat{\text{Var}}(\boldsymbol{\theta}_U)) - q| \\ \text{CovRatio}(\boldsymbol{\theta}) &= \frac{|\hat{\text{Var}}(\boldsymbol{\theta}_U)|}{|\hat{\text{Var}}(\hat{\boldsymbol{\theta}})|}, \end{aligned} \tag{3.85}$$

where q denotes the rank of $\text{Var}(\hat{\boldsymbol{\theta}})$. The hessian matrix is used to obtain the variance matrix that is used in the computation of covariance and CovRatio for covariance parameters. One can request a listing of this matrix with the asycov option of the

proc Mixed statement.

3.5.8 Effect on fitted and predicted values

The Mixed procedure computes the following statistics to measure influence on fitted and predicted values. The PRESS residuals is the difference between the observed value and the predicted marginal mean where the predicted value is obtained without the observations which is given as

$$\hat{\boldsymbol{\epsilon}}_{i(U)} = \mathbf{Y}_i - \mathbf{X}_i'\hat{\boldsymbol{\beta}}_U. \tag{3.86}$$

The Mixed procedure reports these PRESS residuals. If one compute the influence of individual observations, proc Mixed computes the PRESS statistic, when removing sets of observations. This statistic is the sum of the squared PRESS residuals in a deletion set

$$\text{PRESS}_{(U)} = \sum_{i \in U} \hat{\boldsymbol{\epsilon}}_{i(U)}. \tag{3.87}$$

The DFFITS statistic can measure the effect of observations on fitted values. Due to removal of a single data point, a DFFITS measures the change in predicted values. The DFFITS statistic is obtained, if we change the standardized by the externally estimated standard error of the predicted value in the full data

$$\text{DFFITS}_i = \frac{(\hat{\mathbf{Y}}_i - \hat{\mathbf{Y}}_{i(U)})}{ese(\hat{\mathbf{Y}}_i)}. \tag{3.88}$$

3.5.9 Types of covariance structures

Unstructured.
In general, the unstructured structure allows all the variance and covariance terms to be different. It requires fitting the most parameters and allowing them to choose what structure should be. The less data left to estimate the parameters of linear model and more data are used to assess the covariance structure

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{bmatrix}.$$

Compound symmetry.

There is a correlation between two separate measurements of the same subject. But it is assumed that the correlation that the correlation is constant regardless of how far apart the measurements of the same subject are. The variance and covariance are homogeneous since the observations are from the same subjects. Measurements made apart are further apart are less correlated than consecutive measurements that are highly correlated

$$\begin{bmatrix} \sigma^2 & \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}.$$

Autoregressive(1).
The AR(1) structure consist of homogeneous variances and the correlations that decrease exponentially with distance. Meaning that two measurements that are right next to each other in time are going to be correlated but depending on the value of ρ, but as measurements get further apart they are less correlated

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}.$$

Toeplitz.
The Toeplitz structure is similar to AR(1) in that the correlation is the same for measurements that have that have the same distance. However, the correlation pattern is different from the AR(1). Hence AR(1) is a special case of the toeplitz.

$$\begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{bmatrix}.$$

3.6 Modelling covariance structures

In this section we consider six models with different covariance structures for equation(3.2). We discuss the assumptions of this six models. Also, we assume that $\boldsymbol{X}_i$ consist of only time indicators such that $\boldsymbol{X}_i = \boldsymbol{I}_4$ where $\boldsymbol{I}_4$ is an 4×4 identity matrix, where we ignore intercept terms.

3.6.1 Model 1- Independent random time effects model

In this model, we will perform a separate univariate random effect meta-analysis at each time point. Since the effect sizes at different time points do not depend on each other. This model allows independent random intercept effects at each time point t per study i, $\boldsymbol{u}_{it}$, such that

$$Y_{it} = \beta_t + u_{it} + \epsilon_{it}, \quad t = 1, \cdots, 4, \tag{3.89}$$

where we assume $u_{it} \backsim N(0, \tau_t^2)$ and $\epsilon_{it} \backsim N(0, \sigma_{it}^2)$ to be independent. We can set model by allowing a random-effect at each measurement occasion. That is, we set $\boldsymbol{Z}_i = \boldsymbol{X}_i = \boldsymbol{I}_4$ so that equation(3.2), becomes

$$\boldsymbol{Y}_i = \boldsymbol{\beta} + \boldsymbol{u}_i + \boldsymbol{\epsilon}_i, \tag{3.90}$$

and

$$\mathrm{V}(\boldsymbol{Y}_i) = \mathbf{D} + \boldsymbol{R}_i = \mathrm{diag}(\tau_1^2 + \sigma_{i1}^2, \tau_2^2 + \sigma_{i2}^2, \tau_3^2 + \sigma_{i3}^2, \tau_4^2 + \sigma_{i4}^2). \tag{3.91}$$

Even though, this model ignores within study serial correlation between longitudinal effect size which exists because it is the same individuals who are measured repeatedly at these time points.

3.6.2 Model 2- Random study effects model

This model is a simplest way to account for the correlation between longitudinal effect size that is, $\boldsymbol{R}_i = \mathrm{diag}(\sigma_{i1}^2, \cdots, \sigma_{i4}^2)$. To allow a random-effect that is common to all longitudinal effect size from a given study. This can be thought of as a random intercept model. If for example k=4, we would set $\boldsymbol{Z}_i' = [1111]$, so that $\boldsymbol{u}_i = u_i$ is a scalar and the model is now given by

$$Y_{it} = \beta_t + u_i + \epsilon_{it}, \quad t = 1, \cdots, 4, \tag{3.92}$$

where we assume $u_i \backsim N(0, \tau^2)$ with τ^2 representing the between study variability or heterogeneity. The variance covariance matrix is now given by $\mathrm{V}(\boldsymbol{Y}_i) = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' + \boldsymbol{R}_i$, a 4×4 matrix consisting of diagonal elements set to $\tau^2 \neq \sigma_{it}^2$ and off-diagonal elements all equal to τ^2, where $\boldsymbol{D} = \mathrm{V}(\boldsymbol{u}_i) = \tau^2$. Thus, the correlation between two time points (t, t') is $\mathrm{corr}(Y_{it}, Y_{it'}) = \tau^2 \Big/ \sqrt{(\tau^2 + \sigma_{it}^2)(\tau^2 + \sigma_{it'}^2)}$. Therefore, by including a random study effect, we automatically induce a correlation between

any two effect sizes within a study. Regardless of the time lag between the time points these correlations are assume to be the same for each set of time points. This covariance structure is also known as compound symmetry. This model ignores the serial correlation between effect sizes for instance, effect sizes closer together tends to be more strongly correlated than those measures far apart due to factors such as loss-to-follow-up, however it allows only one random effect for all the longitudinal effect sizes from each study.

3.6.3 Model 3- Correlated random time effects model

In this model, the dependence between effect size is accounted for through the dependence between random time effects, which is the extension of the independent random time effect model. Meanwhile assuming zero within study serial correlations between longitudinal effect sizes, this model imposes heteroscedastic AR(1) covariance structure for the random time effects, that is $\boldsymbol{R}_i = \text{diag}(\sigma^2_{i1}, \cdots, \sigma^2_{i4})$. As a result, the variance covariance matrix is now given by $\text{V}(\boldsymbol{Y}_i) = \boldsymbol{D} + \boldsymbol{R}_i$, with diagonal elements $(\tau^2_1 + \sigma^2_{i1}, \cdots, \tau^2_4 + \sigma^2_{i4})$ and off-diagonal elements $(\rho_\tau^{|t-t'|}\tau_t\tau'_t)$ for time points t and t', where ρ_τ is the correlation between any two adjacent random-time effects. As the lag between effect sizes gets smaller the dependence between effect sizes become stronger. The effect sizes measured far apart have less dependence than those closer to one another, this is possible since in longitudinal studies where loss -to-follow-up increases with time. However, this model assumes independent within study residuals which is not suitable for longitudinal effect sizes. A structure that takes account of the autocorrelation between the effect sizes within a study is more suitable.

3.6.4 Model 4- Correlated within-study effect sizes model

The dependence between effect size in this model is accounted for through the dependence in effect sizes within the same study, hence this is an extension of the independent random time effects model. Meanwhile assuming independent random time effect, that is $\boldsymbol{D} = \text{diag}(\tau^2_1, \cdots, \tau^2_4)$, this model imposes heteroscedastic AR(1) covariance structure for the within study of longitudinal effect sizes. As a result, the variance covariance matrix is now give by $\text{V}(\boldsymbol{Y}_i) = \boldsymbol{D} + \boldsymbol{R}_i$ with diagonal elements $(\tau^2_1 + \sigma^2_{i1}, \cdots, \tau^2_4 + \sigma^2_{i4})$ and off-diagonal elements $(\rho_s^{|t-t'|}\sigma_{it}\sigma_{it'})$ for time points t and t', where ρ_s is the correlation between any adjacent within study effect sizes. The purpose of including this model is to assess which covariance structure results in a more imposed model between the within study covariance matrix $(\boldsymbol{R}_i)$ and between study covariance $(\boldsymbol{D})$.

3.6.5 Model 5- Correlated within-study effect sizes and correlated random time effects

In this model, the dependence between effect sizes is accounted for through the dependence in both effect sizes within the same study and random time effects, which is the extension of the independent random time effects model. It is a combination of the above two models, where the heteroscedastic AR(1) covariance structures are imposed on both $\boldsymbol{R}_i$ and $\boldsymbol{D}$. The variance covariance matrix now given by $\mathrm{V}(\boldsymbol{Y}_i) = \boldsymbol{D} + \boldsymbol{R}_i$, with diagonal elements $(\tau_1^2 + \sigma_{i1}^2, \cdots, \tau_4^2 + \sigma_{i4}^2)$ and off-diagonal elements $(\rho_\tau^{|t-t'|}\tau_t\tau_{t'} + \rho_s^{|t-t'|}\sigma_{it}\sigma_{it'})$ for time point t and t'. This model requires estimation of one more parameter compared to each of the above models. However this model accounts for any dependence between effect sizes, both within and between studies.

3.6.6 Model 6-Correlated random time effects(unstructured) and correlated within-study effect sizes

In this model, where the dependence between effect sizes is accounted for through the dependence in both effect sizes within the same study and random time effects, this is an extension of the independent random time effects model. We assume an heteroscedastic AR(1) covariance structure for the within study longitudinal effect sizes an unstructured covariance structure for the random time effects. The variance covariance structure is now given by $\mathrm{V}(\boldsymbol{Y}_i) = \boldsymbol{D} + \boldsymbol{R}_i$, with diagonal elements $(\tau_1^2 + \sigma_{i1}^2, \cdots, \tau_4^2 + \sigma_{i4}^2)$ and off-diagonal elements $(\rho_\tau^{|t-t'|}\tau_t\tau_{t'} + \rho_s^{|t-t'|}\sigma_{it}\sigma_{it'})$ for time points t and t'. The unstructured covariance matrix is quite a superior covariance structure although its requirement for higher number of parameters many compromise parsimony and convergence in some cases.

Chapter 4

Application and results

4.1 Examples of a univariate meta-analysis

In this section we show how to apply fixed-effects and random-effects models in real data. As mentioned above these current standard approaches combines the estimates in meta-analysis from independent studies.

4.1.1 Meta-analysis of clinical trial in duodenal ulcers

Our example consists of 3 clinical trials of the effectiveness of new drug versus placebo, which was used by Lachin [29]. In this example we asses the effectiveness of a new drug for the treatment of duodenal ulcers where the drug is expected to promote healing of the ulcers by retarding the excretion of gastric juices that leads to ulceration of the duodenum. The following tables describe the association between the stratification covariate ulcers type and treatments group and the association between the covariate and the likelihood of healing. Here the ulcer type takes the role of a study for illustration purposes. The data is not the real data its an contrived setup. The aim was to find the data, which the fixed-effects model holds.

Table 4.1: Clinical Trial in Duodenal ulcers.

	Drug			**Placebo**		
Study	a_i	c_i	n_{1i}	b_i	d_i	n_{2i}
1	16	26	42	20	27	47
2	9	3	12	4	5	9
3	28	18	46	16	28	44

4.1.2 Risk difference

Table 4.2: Fixed-effects computations on risk difference.

Study	$\hat{\theta}_i =$ **Risk difference**	$\hat{\sigma}_i^2$	$\hat{\sigma}_i^{-2}$	$\hat{\omega}_i$	95%C.I.
1	-0.0446	0.0108	92.59	0.4369	(-0.2484,0.1593)
2	0.3056	0.0431	23.20	0.1095	(-0.1012,0.7123)
3	0.2451	0.0104	96.15	0.4537	(0.0448,0.4453)
Total			211.95	1	

Table 4.3: Random-effects computations on risk difference.

Study	$\hat{V}(\hat{\theta}_i)$	$\hat{\tau}_i^{(1)} = \hat{\sigma}_i^{-2}$	$\hat{\omega}_i^{(1)}$
1	0.0204	49.02	0.4154
2	0.0527	18.98	0.1608
3	0.0200	50.00	0.4237
Total		117.99	1

Table 4.4: Results of randomised controlled trials of effect of duodenal ulcers from two methods of meta-analysis.

Method	**Fixed-effects**	**Random-effects**
RD(95%)CI	0.13(-0.01,0.26)	0.13(-0.05,0.32)
τ^2		0.0096
P-value	0.0909	0.0909
Q(df=8)	4.797	4.797
I^2		37.54%
H^2		1.60
AIC	0.076	1.50
BIC	-0.825	-0.298
AICc	4.076	13.50
Deviance	4.797	4.225

Table 4.5: Results of the influence diagnostics for the clinical trial in duodenal ulcers using risk difference as the measure of association.

study	**rstudent**	**dffits**	**cook.d**	**cov.r**	**tau2.del**	**QE.del**	**hat**	**weight**	**dfbs**	**inf**
1	-2.1745	-1.7519	1.0185	0.6481	0.0000	0.0684	0.3977	39.7703	-1.6709	*
2	0.6628	0.3368	0.1301	1.6180	0.0313	3.9472	0.1999	19.9916	0.3211	
3	0.5931	0.4284	0.2553	2.2010	0.0344	2.2755	0.4024	40.2381	0.4357	

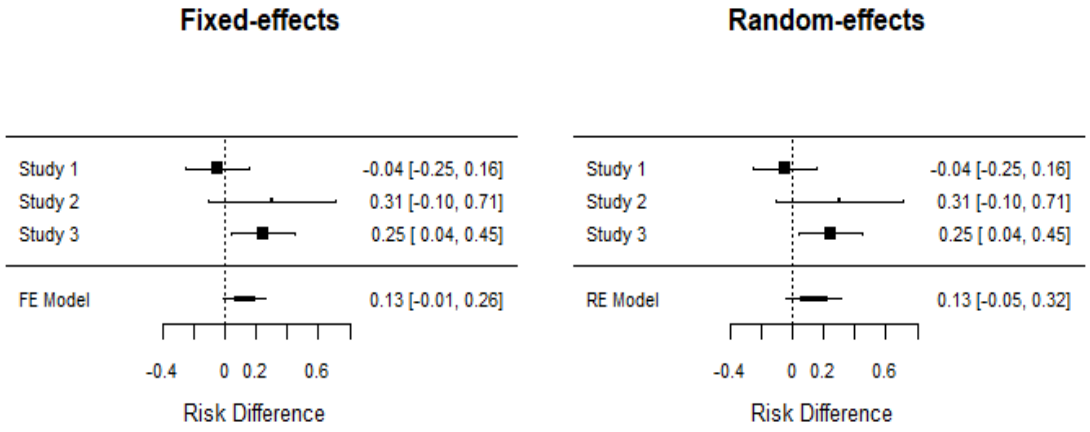

Figure 4.1 – Forest plot showing the results of three studies examining the effectiveness of new drug versus placebo. The figure shows the risk difference of effectiveness of new drug for treatment of duodenal ulcers versus the placebo group with corresponding confidence intervals in the individual studies and based on fixed-effects and random-effects models.

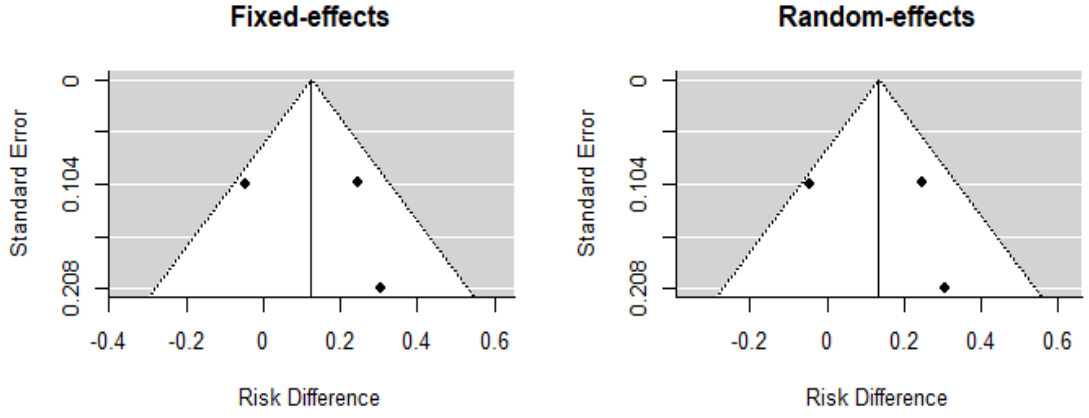

Figure 4.2 – Funnel plot shows the risk difference of three studies examining the effectiveness of new drug versus placebo to treat patients with duodenal ulcers. The points corresponds to the treatment effects from individual trials and the diagonal or curved lines show the expected 95% confidence intervals around the summary estimate based on fixed-effects and random-effects models.

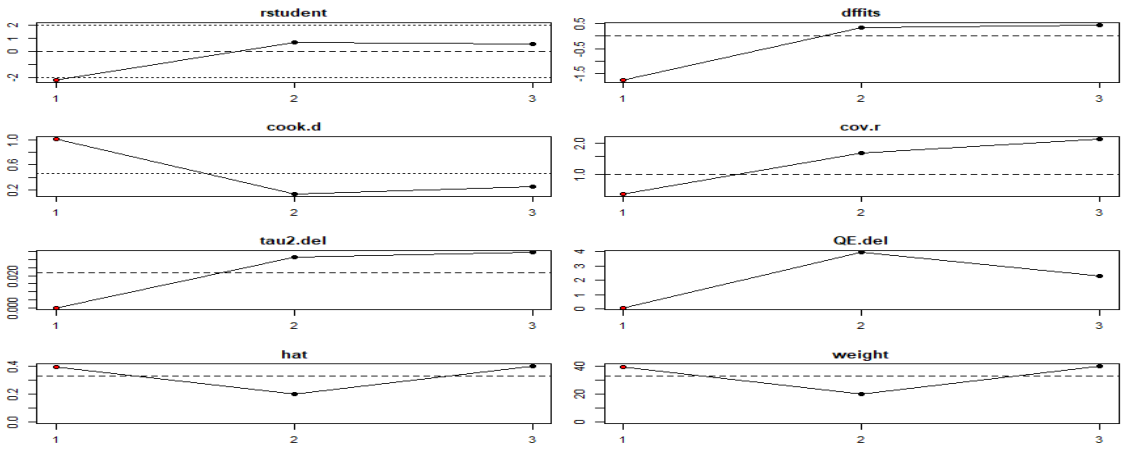

Figure 4.3 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for the 3 studies examining the effectiveness of the new drug versus placebo.

Consider Figure 4.1 above, the pooled estimate of risk difference is 0.13(-0.01,0.26) by using the fixed-effects model. The pooled estimate of risk difference is 0.13(-0.05,0.32), under random-effects model. Under both models, study 3 zero is not included in the 95% confidence interval, it indicate that the study was statistically significant at 5% level of significance. In studies 1 and 2 zero was included in the 95% confidence interval, indicating that there studies were not statically significant at 5% level of significance. In study 1, the box is on the left of the line of no effect, which indicate that the patients favours new drug. Suggesting that, using the new drug to treat patients with duodenal ulcers is beneficial. But using placebo is harmful. In studies 2 and 3, the box was on the right of the line of no effect, which indicates that the patients favours placebo. Suggesting that, using placebo to treat patients with duodenal ulcers is beneficial. Under the fixed-effects model the summary estimate does includes zero in the confidence interval, indicate that there studies were not statistically significant at 5% level of significance. Under the random-effects model, the summary estimate does include zero in the confidence interval indicating that there studies were not statistically significant at 5% level of significance. In both models the diamond touches the line of no effect, which indicates that using both placebo and the new drug to treat patients with duodenal ulcers will promote healing. In Table 4.4, I^2 was obtained to be ($I^2 = 37.54\%$, p=0.0909), we found the presence of moderate heterogeneity in this meta-analysis [90]. The AIC is smaller in the fixed-effects model than in the random-effects model, hence the fixed-effects model had better fit than the random-effects model. Which means that our studies

are close replications of each other (they use same procedures and measures). The funnel plot in Figure 4.2 does not necessary informs us wether publication bias exist or not, because the studies are small (less than ten). When the studies are small, the funnel could not tell if publication bias exists or not. In Table 4.5 and Figure 4.3, study 1 is identified as potential outlier and also to be a influential case. The value of the covariance ratio for this study also suggest that precision could be gained by its removal [91].

4.1.3 Relative risk

Table 4.6: Fixed-effects computations on log relative risk.

Study	$\hat{\theta}_i =$ **log relative risk**	$\hat{\sigma}_i^2$	$\hat{\sigma}_i^{-2}$	$\hat{\omega}_i$	95%C.I.
1	-0.1107	0.0674	14.84	0.3760	(-0.6196,0.3982)
2	0.5232	0.1667	6.00	0.1520	(-0.2769,1.3234)
3	0.5152	0.0537	18.62	0.4719	(0.0608,0.9696)
Total			39.46	1	

Table 4.7: Random-effects computations on log relative risk.

Study	$\hat{V}(\hat{\theta}_i)$	$\hat{\tau}_i^{(1)} = \hat{\sigma}_i^{-2}$	$\hat{\omega}_i^{(1)}$
1	0.0914	10.94	0.3766
2	0.1907	5.244	0.1805
3	0.0777	12.87	0.4430
Total		29.05	1

Table 4.8: Results of randomised controlled trials of effect of duodenal ulcers from two methods of meta-analysis.

Method	**Fixed-effects**	**Random-effects**
log RR(95%)CI	0.28(-0.03,0.59)	0.28(-0.08,0.64)
τ^2		0.0240
P-value	0.1614	0.1614
Q(df=8)	3.65	3.65
I^2		22.48%
H^2		1.29
AIC	3.75	5.60
BIC	2.85	3.71
AICc	7.75	17.60
Deviance	3.65	3.41

Table 4.9: Results of the influence diagnostics for the clinical trial in duodenal ulcers using log relative risk as the measure of association.

study	rstudent	dffits	cook.d	cov.r	tau2.del	QE.del	hat	weight	dfbs	inf
1	-1.9098	-1.4704	1.0509	0.7888	0.0000	0.0003	0.3714	37.1445	-1.4613	*
2	0.4905	0.2788	0.0986	1.8982	0.1353	3.2326	0.2165	21.6513	0.2623	
3	0.8039	0.6570	0.4751	1.8310	0.0839	1.7167	0.4120	41.2042	0.6613	*

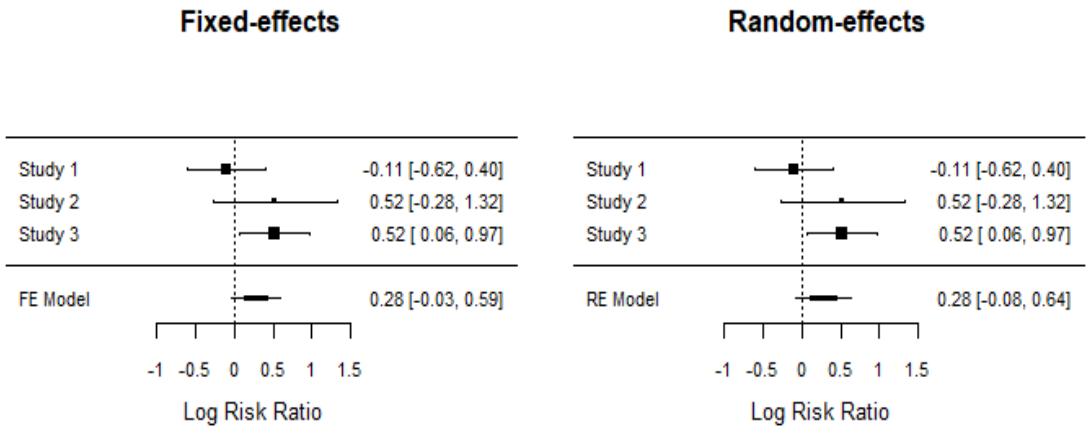

Figure 4.4 – Forest plot showing the results of three studies examining the effectiveness of new drug versus placebo. The figure shows the log relative risk of effectiveness of new drug for treatment of duodenal ulcers versus the placebo group with corresponding confidence intervals in the individual studies and based on fixed-effects and random-effects models.

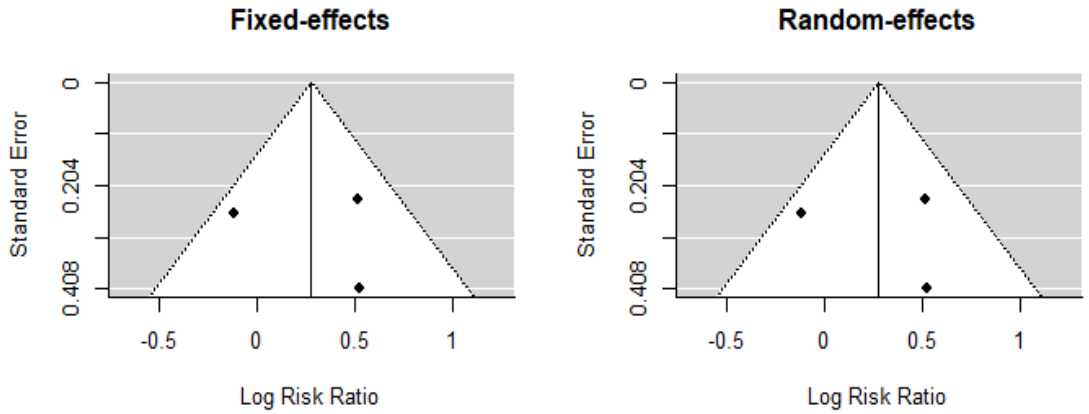

Figure 4.5 – Funnel plot shows the log relative risk of three studies examining the effectiveness of new drug versus placebo to treat patients with duodenal ulcers. The points corresponds to the treatment effects from individual trials and the diagonal or curved lines show the expected 95% confidence intervals around the summary estimate based on fixed-effects and random-effects models.

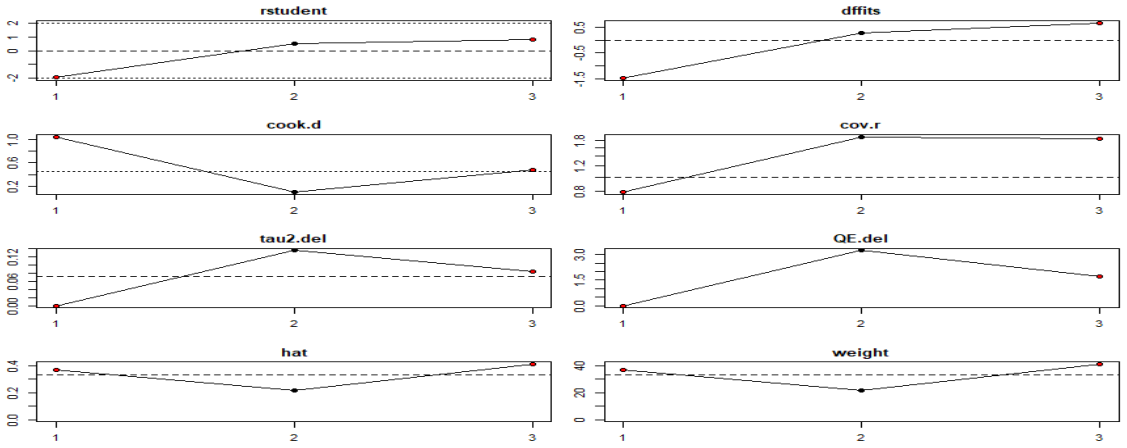

Figure 4.6 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for the 3 studies examining the effectiveness of the new drug versus placebo.

Consider Figure 4.4 above, the pooled estimate of log relative risk is 0.28(-0.03,0.59) by using the fixed-effects model, hence the relative risk of the summary effect is $e^{0.28} = 1.32$, indicating that the risk that a patient being treated with a new drug will heal the duodenal ulcers is 32% lower than patients treated with placebo. The pooled

estimate of log relative risk is 0.23(-0.08,0.64) under the random-effects model, hence the relative risk of the summary effect is $e^{0.28} = 1.32$, indicating that the risk of a patient being treated with a new drug to heal the duodenal ulcers is 32% lower than patients treated with placebo. Under both models, study 3 zero was not included in the 95% confidence interval, indicating that the study was statistically significant at 5% level of significance. In studies 1 and 2 zero was included in the 95% confidence interval, indicating that there studies were not statically significant at 5% level of significance. In study 1, the box is on the left of the line of no effect, which indicates that the patients favours new drug. Suggesting that, using the new drug to treat patients with duodenal ulcers is beneficial. But using placebo is harmful. In studies 2 and 3, the box was on the right of the line of no effect, which indicates that the patients favours placebo. Suggesting that, using placebo to treat patients with duodenal ulcers is beneficial. Under the fixed-effects model the summary estimate does includes zero in the confidence interval, indicate that there studies were not statistically significant at 5% level of significance. Under the random-effects model the summary estimate does include zero in the confidence interval indicating that there studies were not statistically significant at 5% level of significance. In both models the diamond touches the line of no effect, which indicates that using both placebo and the new drug to treat patients with duodenal ulcers will promote healing. In Table 4.8, I^2 was obtained to be ($I^2 = 22.48\%$, p=0.1614), we found the presence of no heterogeneity in this meta-analysis [90]. The AIC is smaller in the fixed-effects model than in the random-effects model, hence the fixed-effects model had better fit than the random-effects model. Which means that our studies are close replications of each other (they use same procedures and measures). The funnel plot in Figure 4.5 does not necessary informs us wether publication bias exist or not, because the studies are small (less than ten). When the studies are small, the funnel could not tell if publication bias exists or not. In Table 4.9 and Figure 4.6, studies 1 and 3 are identified as potential outliers and also to be a influential cases. The covariance ratio values of these studies also suggest that precision could be gained by their removal [91].

4.1.4 Odds ratio

Table 4.10: Fixed-effects computations on log odds ratio.

Study	$\hat{\theta}_i =$ **log odds ratio**	$\hat{\sigma}_i^2$	$\hat{\sigma}_i^{-2}$	$\hat{\omega}_i$	95%C.I.
1	-0.1854	0.1880	5.319	0.4541	(-1.0352,0.6644)
2	1.3218	0.8944	1.118	0.0954	(-0.5319,3.1754)
3	1.0014	0.1895	5.277	0.4505	(0.1483,1.8546)
Total			11.714	1	

Table 4.11: Random-effects computations on log odds ratio.

Study	$\hat{V}(\hat{\theta}_i)$	$\hat{\tau}_i^{(1)} = \hat{\sigma}_i^{-2}$	$\hat{\omega}_i^{(1)}$
1	0.3450	2.899	0.4303
2	1.0514	0.9511	0.1412
3	0.3465	2.886	0.4285
Total		6.736	1

Table 4.12: Results of randomised controlled trials of effect of duodenal ulcers from two methods of meta-analysis.

Method	**Fixed-effects**	**Random-effects**
log OR(95%)CI	0.49(-0.08,1.07)	0.54(-0.22,1.29)
τ^2		0.1570
P-value	0.1013	0.1013
Q(df=8)	4.58	4.58
I^2		34.85%
H^2		1.53
AIC	8.65	10.16
BIC	7.75	8.36
AICc	12.65	22.16
Deviance	4.58	4.09

Table 4.13: Results of the influence diagnostics for the clinical trial in duodenal ulcers using log odds ratio as the measure of association.

study	**rstudent**	**dffits**	**cook.d**	**cov.r**	**tau2.del**	**QE.del**	**hat**	**weight**	**dfbs**	**inf**
1	-2.1179	-1.7442	1.0165	0.6783	0.0000	0.0947	0.4097	40.9716	-1.6569	*
2	0.6888	0.3279	0.1195	1.5276	0.5156	3.7316	0.1816	18.1646	0.3126	
3	0.5862	0.4232	0.2489	2.2252	0.5945	2.0985	0.4086	40.8638	0.4302	

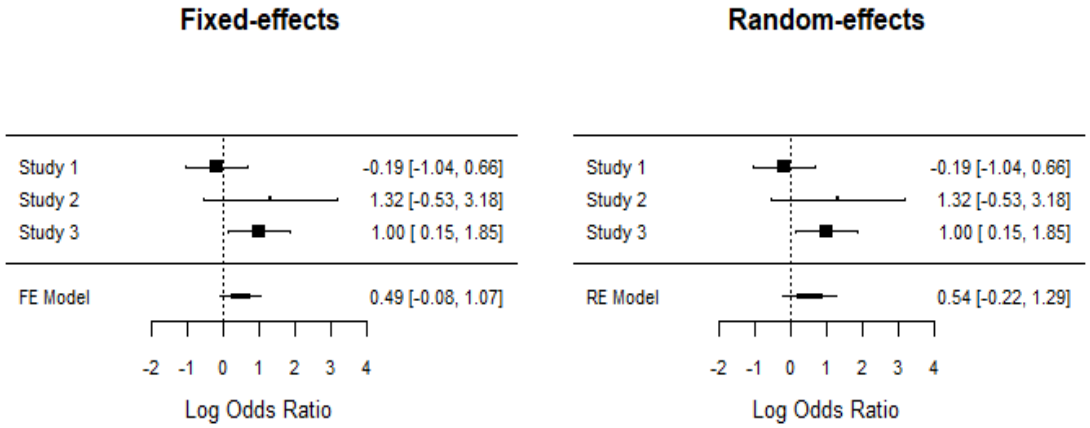

Figure 4.7 – Forest plot showing the results of three studies examining the effectiveness of new drug versus placebo. The figure shows the log odds ratio of effectiveness of the new drug for treatment of duodenal ulcers versus the placebo group with corresponding confidence intervals in the individual studies and based on fixed-effects and random-effects models.

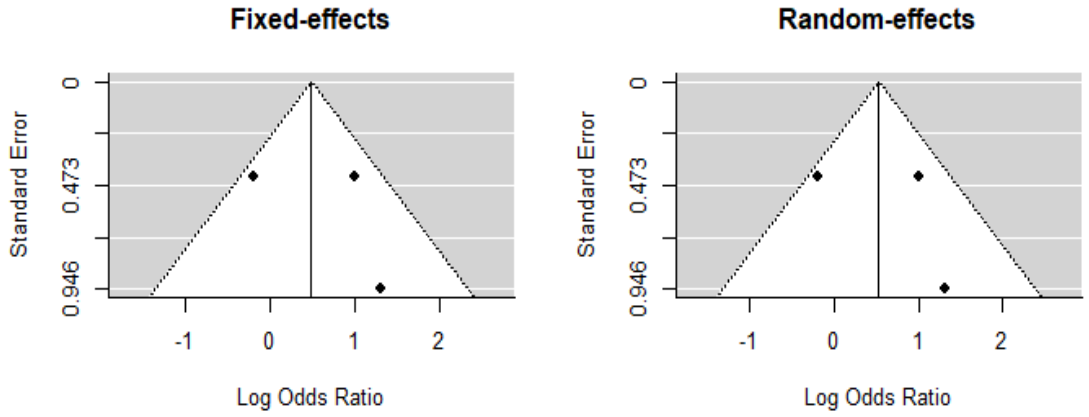

Figure 4.8 – Funnel plot shows the log odds ratio of three studies examining the effectiveness of new drug versus placebo to treat patients with duodenal ulcers. The points corresponds to the treatment effects from individual trials and the diagonal or curved lines show the expected 95% confidence intervals around the summary estimate based on fixed-effects and random-effects models.

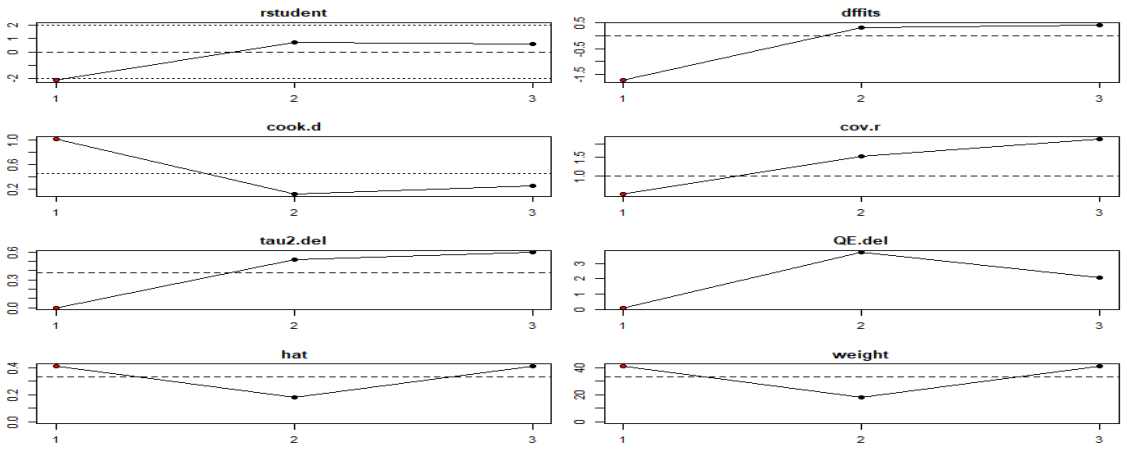

Figure 4.9 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for the 3 studies examining the effectiveness of the new drug versus placebo.

Consider Figure 4.7 above, the pooled estimate of log odds ratio is 0.49(-0.08,1.07) by using the fixed-effects model, hence the odds ratio of the summary effect is $e^{0.49} = 1.63$, indicating that the odds of a patient being treated with a new drug to heal the duodenal ulcers is 63% lower than patients treated with placebo. The pooled estimate of log odds ratio is 0.54(-0.22,1.29) under the random-effects model, hence the odds ratio of the summary effect is $e^{0.54} = 1.72$, indicating that the odds of a patient being treated with a new drug to heal the duodenal ulcers is 72% lower than patients treated with placebo. Under both models, study three confidence interval does not include zero, indicating that the study was statistically significant at 5% level of significance. In studies 1 and 2 zero is included in the 95% confidence interval, indicating that there studies were not statically significant at 5% level of significance. In study 1, the box is on the left of the line of no effect, which indicate that the patients favours diuretics. Suggesting that, using the new drug to treat patients with duodenal ulcers is beneficial. But using placebo is harmful. In studies 2 and 3, the box was on the right of the line of no effect, which indicate that the patients favours placebo. Suggesting that, using placebo to treat patients with duodenal ulcers is beneficial. Under the fixed-effects model the summary estimate does includes zero in the confidence interval, indicating that there studies were not statistically significant at 5% level of significance. Under the random-effects model the summary estimate does include zero in the confidence interval indicating that there studies were not statistically significant at 5% level of significance. In both models, the diamond touches the line of no effect which indicates that using both placebo and the

new drug to treat patients with duodenal ulcers will promote healing. In Table 4.12, I^2 was obtained to be ($I^2 = 34.85\%$, p=0.1013), we found the presence of moderate heterogeneity in this meta-analysis [90]. The AIC is smaller in the fixed-effects model than in the random-effects model, hence the fixed-effects model had better fit than the random-effects model. Which means that our studies are close replications of each other (they use same procedures and measures). The funnel plot in Figure 4.8 does not necessary informs us wether publication bias exist or not, because the studies are small (less than ten). When the studies are small, the funnel could not tell if publication bias exists or not. In Table 4.13 and Figure 4.9, study 1 is identified as potential outlier and also to be a influential case. The value of the covariance ratio for this study also suggest that precision could be gained by its removal [91].

4.1.5 Meta-analysis of effects of diuretics on pre-eclampsia

We use the example by Yusuf [92], of a meta-analysis of nine clinical trials investigating the effect of taking diuretics during pregnancy on the risk of pre-eclampsia.

Table 4.14: Meta-analysis of nine trials of effects of diuretics on pre-eclampsia.

		Diuretics Group			**Placebo Group**		
study	a_i	n_{1i}	p_{1i}	b_i	n_{2i}	p_{2i}	OR_i
Weseley 1962	14	131	0.107	14	136	0.103	1.043
Flowers 1962	21	385	0.055	17	134	0.127	0.397
Menzies 1964	14	57	0.246	24	48	0.500	0.326
Fallis 1964	6	38	0.158	18	40	0.450	0.229
Cuadros 1964	12	1011	0.012	35	760	0.046	0.249
Landesman 1965	138	1370	0.101	175	1336	0.131	0.743
Kraus 1966	15	506	0.030	20	524	0.038	0.770
Tervila 1971	6	108	0.056	2	103	0.019	2.971
Campbell 1975	65	153	0.425	40	102	0.392	1.145

4.1.6 Risk Difference

Table 4.15: Fixed-effects computations on risk difference.

Study	$\hat{\theta}_i =$ **Risk difference**	$\hat{\sigma}_i^2$	$\hat{\sigma}_i^{-2}$	$\hat{\omega}_i$	95%C.I.
1	0.0039	0.0014	714.286	0.025	(-0.0696,0.0775)
2	-0.0723	0.0010	1000	0.035	(-0.1331,-0.0116)
3	-0.2544	0.0085	117.647	0.004	(-0.4346,-0.0741)
4	-0.2921	0.0097	103.093	0.004	(-0.4850,-0.0992)
5	-0.0342	0.0001	10000	0.349	(-0.0505,-0.0179)
6	-0.0303	0.0002	5000	0.175	(-0.0544,-0.0061)
7	-0.0085	0.0001	10000	0.349	(-0.0306,0.0136)
8	0.0361	0.0007	1428.571	0.050	(-0.0146,0.0869)
9	0.0327	0.0039	256.410	0.009	(-0.0903,0.1556)
Total			28620.007	1	

Table 4.16: Random-effects computations on risk difference.

Study	$\hat{V}(\hat{\theta}_i^{(1)})$	$\hat{\tau}_i^{(1)} = \hat{\sigma}_i^{-2}$	$\hat{\omega}_i^{(1)}$
1	0.0016	625	0.0506
2	0.0010	1000	0.0809
3	0.0087	114.94	0.0093
4	0.0099	101.01	0.0082
5	0.0003	3333.33	0.2696
6	0.0004	2500	0.2022
7	0.0003	3333.33	0.2696
8	0.0009	1111.11	0.0899
9	0.0041	243.90	0.0197
Total		12362.63	1

Table 4.17: Results of randomised controlled trials of effect of diuretics on pre-eclampsia from two methods of meta-analysis.

Method	**Fixed-effects**	**Random-effects**
RD(95%)CI	-0.03(-0.04,-0.01)	-0.02(-0.04,-0.01)
τ^2		0.0002
P-value	0.0009	0.0009
Q(df=8)	26.46	26.46
I^2		35.32%
H^2		1.55
AIC	-18.11	-16.61
BIC	-17.91	-16.30
AICc	-17.54	-14.61
Deviance	26.46	25.88

Table 4.18: Results of the influence diagnostics for the trials of effects of diuretics on pre-eclampsia using risk difference as the measure of association.

study	**rstudent**	**dffits**	**cook.d**	**cov.r**	**tau2.del**	**QE.del**	**hat**	**weight**	**dfbs**	**inf**
1	0.6429	0.5685	0.5121	2.2150	0.0022	25.8333	0.0861	8.6101	0.5089	*
2	-1.3187	-0.5304	0.2176	0.7126	0.0005	24.0984	0.1072	10.7177	-0.5792	
3	-2.4478	-0.3782	0.1366	0.6538	0.0004	20.2421	0.0210	2.0991	-0.4603	
4	-2.6716	-0.3822	0.1401	0.6415	0.0004	19.0967	0.0185	1.8549	-0.4702	
5	0.0555	0.3927	0.6152	3.4988	0.0037	24.4937	0.2093	20.9330	0.4562	*
6	0.1232	0.4295	0.6881	3.4725	0.0036	26.2691	0.1925	19.2468	0.4834	*
7	0.5168	0.5524	1.0313	3.1673	0.0032	23.4822	0.1972	19.7202	0.6211	*
8	2.2086	0.2691	0.0384	0.3704	0.0001	20.5331	0.1274	12.7402	0.3359	
9	0.8725	0.3688	0.1493	1.4328	0.0013	25.5946	0.0408	4.0780	0.3313	

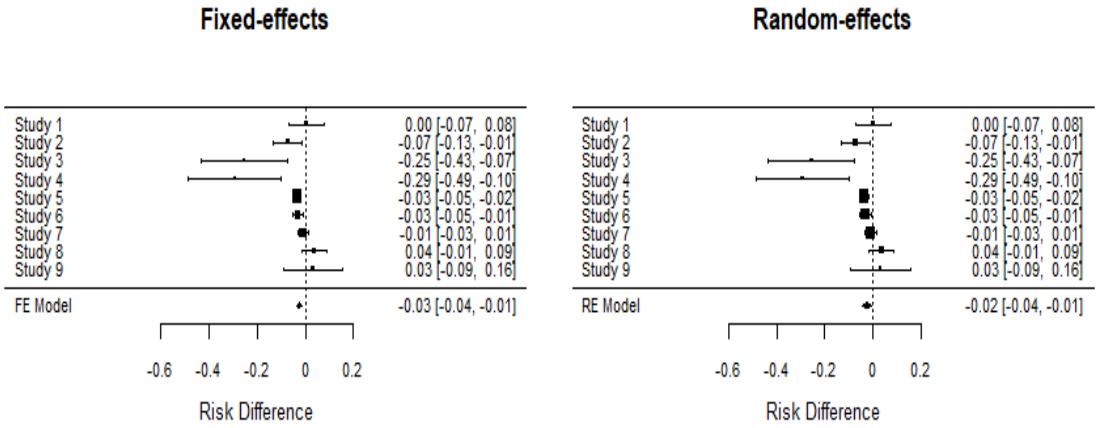

Figure 4.10 – Forest plot showing the results of nine studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia. The figure shows the risk difference of pre-eclampsia among those treated with diuretics versus the placebo group with corresponding confidence intervals in the individual studies and based on fixed-effects and random-effects models.

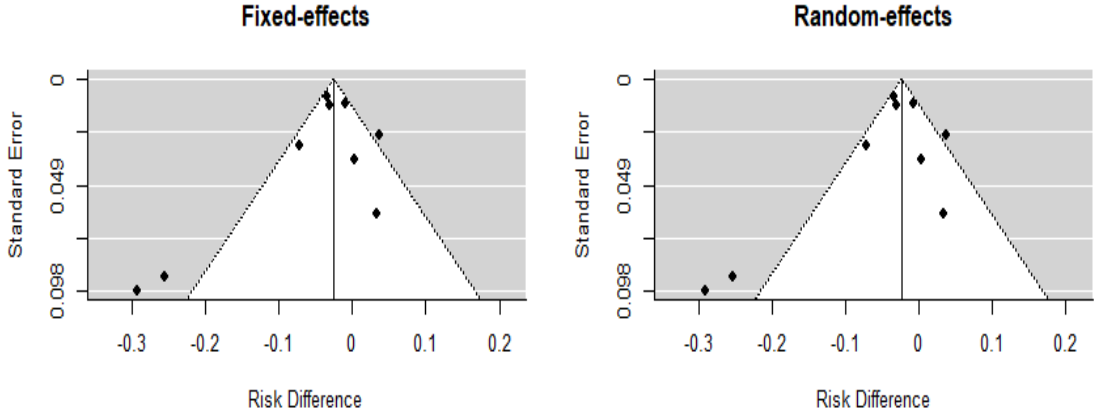

Figure 4.11 – Funnel plot shows the risk difference of nine studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia. The points corresponds to the treatment effects from individual trials and the diagonal or curved lines show the expected 95% confidence intervals around the summary estimate and based on fixed-effects and random-effects models.

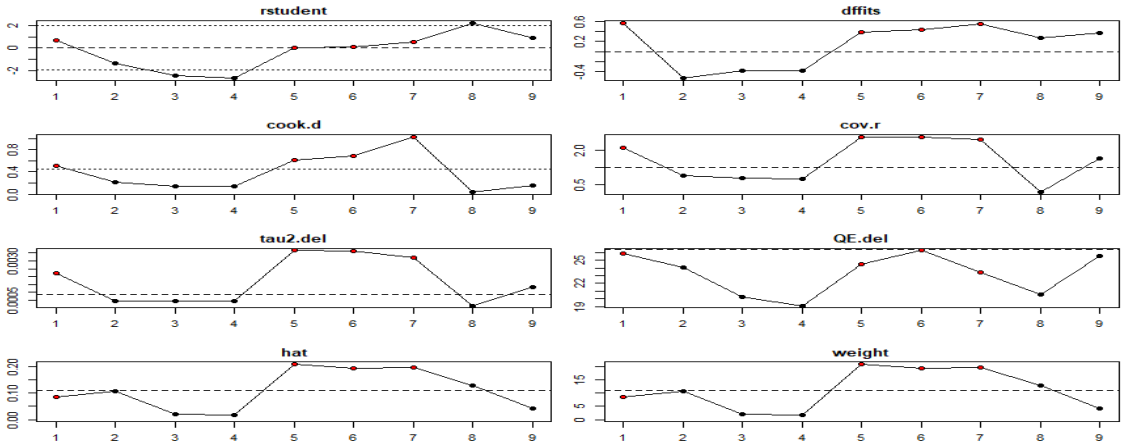

Figure 4.12 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for the 9 studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia.

Consider Figure 4.10 above, the pooled estimate of risk difference is -0.03(-0.04,-0.01) by using the fixed-effects model and the pooled estimate of risk difference is -0.03(-0.04,-0.01) using the random-effects model. The confidence interval for study 7 is noticeably narrower than all the studies in both fixed-effects and random-effects models, reflecting the fact that it has greater precision compared to other studies. Under the fixed-effects model in Table 4.15, studies 5, 6 and 7 are assigned relatively high weight, while somewhat less weight is assigned to studies 1, 2, 3, 4, 8 and 9. As one would expect, there is a relationship between a study's precision and study's weight in the analysis. Studies with relatively good precision are studies 5, 6 and 7 are assigned more weight, while studies with relatively poor precision are studies 1, 2, 3, 4, 8 and 9 are assigned less weights. On the other hand, under the random-effects model in Table 4.16, studies 2, 5, 6, 7 and 8 are assigned relatively high weights, while somewhat less weight is assigned to studies 1, 3, 4 and 9. In studies 2, 5, 6, 7 and 8 has relatively good precision since they are assigned more weight while studies 1, 3, 4 and 9 are assigned less weight thus they had relatively poor precision. Under both models, in studies 2, 3, 4, 5 and 6 zero was not included in the 95% confidence interval, indicating that there studies were statistically significant at 5% level of significance. In studies 1, 7, 8 and 9 zero was included in the 95% confidence interval, it indicating that there studies were not statically significant at 5% level of significance. In studies 2, 3, 4, 5 and 6, the box was on the left of the line of no effect, which indicate that the patients favours diuretics. Suggesting

that, using the diuretics in preventing the development of pre-eclampsia in pregnant woman is beneficial. But using placebo is harmful. Studies 1 and 9 the box touches the line of no effect, which means that the patients in these studies had no effect, when treated with both diuretics and placebo to prevent the development of pre-eclampsia in pregnant woman. In study (8 and 9), the box is on the right of the line of no effect, which indicate that the patients favours placebo. Suggesting that, using placebo in preventing the development of pre-eclampsia in pregnant woman is beneficial. But using diuretics is harmful. Under the fixed-effects model the summary estimate does not include zero in the confidence interval, indicate that there studies were statistically significant at 5% level of significance. Under the random-effects model the summary estimate does include zero in the confidence interval indicating that there studies were not statistically significant at 5% level of significance. In both models, the summary estimate, indicated by the diamond is in the left side of the line of no effect, meaning that the final conclusion is that the use of diuretics during pregnancy to prevent the development of pre-eclampsia is beneficial. In Table 4.17, I^2 was obtained to be ($I^2 = 35.32\%$, p=0.0009), we found the presence of moderate heterogeneity in this meta-analysis [90]. The AIC is smaller in the random-effects model than in the fixed-effects model, hence the random-effects model had better fit than the fixed-effects model. Which means that there is some extra variation or over-dispersion due to random differences among the studies. The funnel plot in Figure 4.11 does not necessary informs us wether publication bias exist or not, because the studies are small (less than ten). When the studies are small, the funnel could not tell if publication bias exists or not. In Table 4.18 and Figure 4.12, shows that removing study (1, 5, 6 and 7) would yield little change in the amount of residual heterogeneity, but their influence on the model fit is more considerable. Study 8 introduce some additional residual heterogeneity into the model, removing this study in turn would yield considerably smaller estimate of τ^2, but only have a modest influence on the fit of the model [93].

4.1.7 Relative Risk

Table 4.19: Fixed-effects computations on log relative risk.

Study	$\hat{\theta}_i =$ **log relative risk**	$\hat{\sigma}_i^2$	$\hat{\sigma}_i^{-2}$	$\hat{\omega}_i$	95%C.I.
1	0.0375	0.1279	7.819	0.0421	(-0.6634,0.7383)
2	-0.8441	0.0964	10.373	0.0559	(-1.4526,-0.2356)
3	-0.7108	0.0747	13.387	0.0721	(-1.2466,-0.1751)
4	-1.0473	0.1709	5.851	0.0315	(-1.8576,-0.2371)
5	-1.3558	0.1096	9.124	0.0492	(-2.0047,-0.7070)
6	-0.2627	0.0115	86.957	0.4685	(-0.4727,-0.0526)
7	-0.2527	0.1128	8.865	0.0478	(-0.9109,0.4055)
8	1.0512	0.6477	1.544	0.0083	(-0.5262,2.6286)
9	0.0800	0.0240	41.667	0.2245	(-0.2239,0.3840)
Total			185.587	1	

Table 4.20: Random-effects computations on log relative risk.

Study	$\hat{V}(\hat{\theta}_i^{(1)})$	$\hat{\tau}_i^{(1)} = \hat{\sigma}_i^{-2}$	$\hat{\omega}_i^{(1)}$
1	0.2867	3.4880	0.1005
2	0.2552	3.9185	0.1129
3	0.2335	4.2827	0.1234
4	0.3297	3.0331	0.0874
5	0.2684	3.7258	0.1073
6	0.1703	5.8720	0.1692
7	0.2716	3.6819	0.1061
8	0.8065	1.2399	0.0357
9	0.1828	5.4705	0.1576
Total		34.7122	1

Table 4.21: Results of randomised controlled trials of effect of diuretics on pre-eclampsia from two methods of meta-analysis.

Method	**Fixed-effects**	**Random-effects**
log RR(95%)CI	-0.31(-0.45,-0.16)	-0.44(-0.77,-0.10)
τ^2		0.1588
P-value	0.0004	0.0004
Q(df=8)	28.62	28.62
I^2		72.46%
H^2		3.63
AIC	25.38	19.60
BIC	25.58	20.00
AICc	25.96	21.60
Deviance	28.62	20.84

Table 4.22: Results of the influence diagnostics for the trials of effects of diuretics on pre-eclampsia using log relative risk as the measure of association.

study	**rstudent**	**dffits**	**cook.d**	**cov.r**	**tau2.del**	**QE.del**	**hat**	**weight**	**dfbs**	**inf**
1	0.8644	0.2958	0.0908	1.1642	0.2113	27.6640	0.1031	10.3139	0.2949	
2	-0.7662	-0.2751	0.0807	1.2092	0.2190	25.4296	0.1141	11.4127	-0.2746	
3	-0.5184	-0.1932	0.0427	1.2993	0.2385	26.2477	0.1232	12.3154	-0.1935	
4	-1.0520	-0.3327	0.1100	1.0914	0.1970	25.2939	0.0911	9.1146	-0.3330	
5	-2.1305	-0.7622	0.3967	0.7119	0.1012	18.0285	0.1092	10.9242	-0.7878	
6	0.3700	0.1514	0.0291	1.4156	0.2558	28.3266	0.1601	16.0122	0.1556	
7	0.3312	0.1172	0.0156	1.2969	0.2427	28.5966	0.1081	10.8128	0.1164	
8	1.6686	0.3215	0.1012	0.9713	0.1809	25.7581	0.0398	3.9831	0.3290	
9	1.3132	0.5694	0.2677	1.0197	0.1602	20.6717	0.1511	15.1111	0.5582	

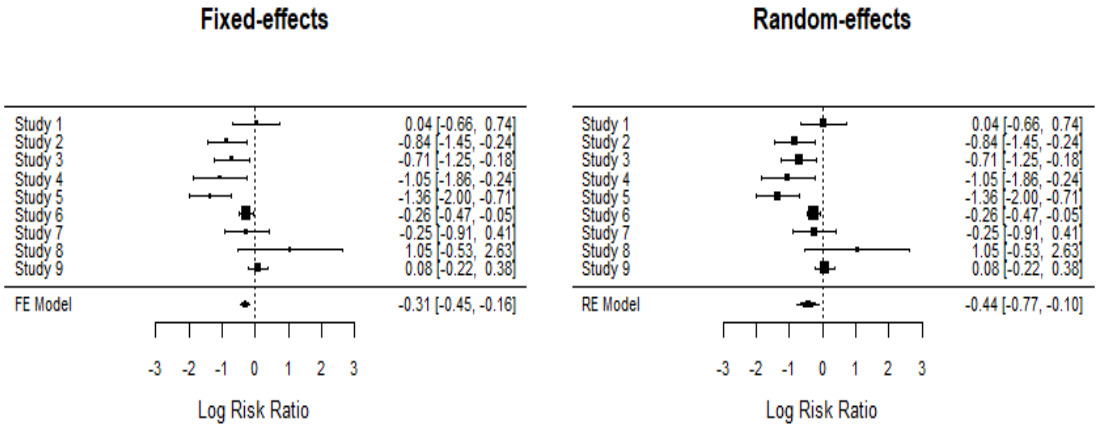

Figure 4.13 – Forest plot showing the results of nine studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia. The figure shows the log relative risk of pre-eclampsia among those treated with diuretics versus the placebo group with corresponding confidence intervals in the individual studies and based on fixed-effects and random-effects models.

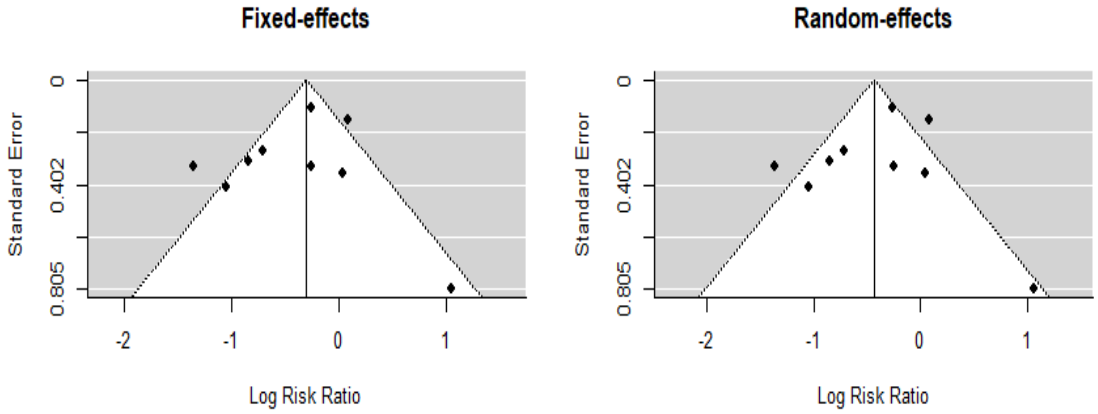

Figure 4.14 – Funnel plot shows the log relative risk of nine studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia. The points corresponds to the treatment effects from individual trials and the diagonal or curved lines show the expected 95% confidence intervals around the summary estimate and based on fixed-effects and random-effects models.

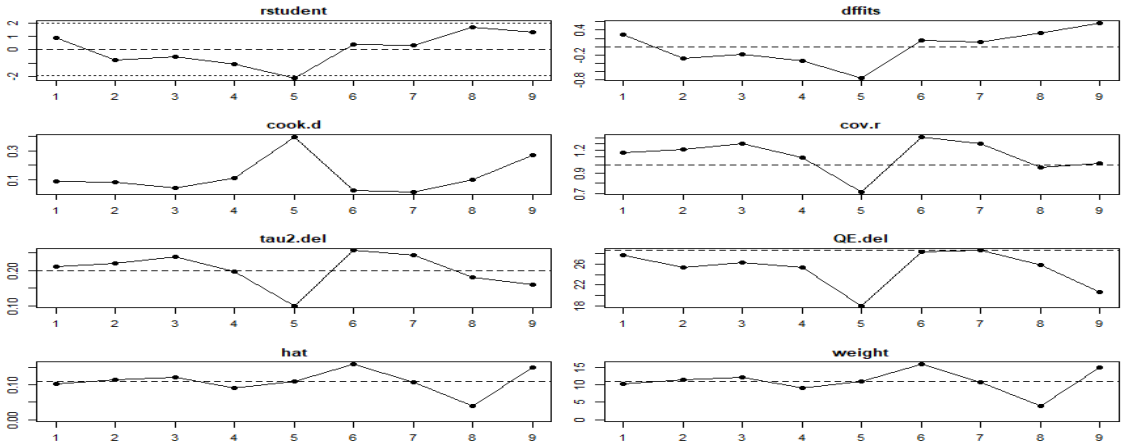

Figure 4.15 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for the 9 studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia.

In epidemiologic meta-analysis, random-effects summary effect estimates are more conservative than fixed-effects summaries [94]. This view is clearly seen, when there is evidence of appreciable heterogeneity among the result from the individual studies. As a result, random-effects summaries have high estimated variances and, consequently, wider confidence interval than fixed-effects summaries [94]. In such instances, however, the random-effects point estimates are not always closer to the null value nor are their p-values always larger than those of fixed-effects summaries. Moreover they can appear more strongly supportive of causation or prevention than fixed-effects summaries. In Figure 4.13, the pooled estimate of log relative risk is -0.31(-0.45,-0.16) by using the fixed-effects model and the pooled estimate of log relative risk is -0.44(-0.77,-0.11) using the random-effects model. As expected the random-effects model provides a wider confidence interval of the pooled estimate than the fixed-effects model. Therefore the random-effects summary effect estimates are more conservative than fixed-effects summaries. Hence the probability of a type I error in the meta-analyses is minimized. The confidence interval for study 6 is noticeably narrower than all the studies in both fixed-effects and random-effects models, reflecting the fact that it has greater precision compared to other studies. The solid squares that are used to outline each of the studies vary in size, with the size of each square reflecting the weight that is assigned to the corresponding study when we compute the summary effect. Under the fixed-effects model in Table 4.19, studies 6 and 9 are assigned relatively high weight, while somewhat less weight is assigned

to studies 1, 2, 3, 4, 5, 7 and 8. As one would expect, there is a relationship between a study's precision and that study's weight in the analysis. Studies with relatively good precision studies 6 and 9 are assigned more weight while studies with relatively poor precision studies 1, 2, 3, 4, 5, 7 and 8 are assigned less weight. On the other hand under the random-effects model in Table 4.20, studies 1, 2, 3, 5, 6, 7 and 9 are assigned relatively high weights, while somewhat less weight is assigned to studies 4 and 5. In studies 1, 2, 3, 4, 5, 7 and 8 has relatively good precision since they are assigned more weight while studies 4 and 5 are assigned less weight thus they had relatively poor precision. Under the fixed-effects model we assume that the true effect size for all studies is identical, and the only reason the effect sizes varies between studies is sampling error (error in estimating the effect sizes). Therefore, when assigning weights to different studies we can largely ignore the information in the smaller studies since we have better information about the same effect size in the larger studies. By contrast, under the random-effects model the goal is not to estimate one true effect, but to estimate the mean of a distribution of effects. Since each study provides information about a different effect size, we want to be convinced that all these effect sizes are represented in the summary estimate. This means that we cannot discount a small study by giving it a very small weight (the way we would in a fixed-effects analysis). The estimate provided by that study may be imprecise, but it is information about an effect that no other study has estimated. By the same logic we cannot give too much weight to a very large study(the way we might in a fixed-effects analysis). Our goal is to estimated the mean effect in a range studies, and we do not want that overall estimate to be overly influenced by any one of them. In Figure 4.13, the weight assigned to each study is reflected in the size of the box (specifically the area) for that study. Under fixed-effects model there is a wide of weights (as reflected in the size of the boxes). Whereas under random-effects model the weights fall in a relatively narrow range. For example, compare the weight assigned to the largest study 6 with that assigned to the smallest study 8 under both models. Under the fixed-effects model study 6 has high weight than study 8. Under random-effects model study 6 has slightly high weight than study 8.
In Figure 4.13 above the is a forest plot of a meta-analysis of nine clinical trials investigating the effect of taking diuretics during pregnancy on the risk of pre-eclampsia. The treatment effect is measured in terms of the log relative risk. Figure 4.13 consist of two plot the first plot we use the fixed-effects analysis and the second plot we use the random-effects analysis. There plots provide a context for the discussion that follows. In both models studies 1, 7, 8 and 9, zero was included in the 95% confidence interval, indicating that there studies were not statistically significant at 5% level of significance. Also in studies 2, 3, 4, 5 and 6, zero was not included in the 95% confidence interval, it indicate that there studies were statistically significant

at 5% level of significance. In studies 2, 3, 4, 5, 6 and 7, the box was on the left of the line of no effect, which indicates that the patients favours diuretics. Suggesting that, using the diuretic drug in preventing the development of pre-eclampsia in pregnant woman is beneficial. But using placebo is harmful. Studies 1 and 9 the box touches the line of no effect, which means that the patients in these studies had no effect, when treated with both diuretics and placebo to prevent the development of pre-eclampsia in pregnant woman. On the plot summary effect shown on the bottom line. Under the fixed-effects model in this example the summary relative risk is $e^{-0.31} = 0.733$, indicating that the risk of a patient being treated with diuretics to prevent the development of pre-eclampsia in pregnant woman is 27% lower than patients treated with placebo. Under the random-effects model the relative risk is $e^{-0.44} = 0.644$, indicating that the risk of a patient being treated with diuretics to prevent the development of pre-eclampsia in pregnant woman is 35.6% lower than patients treated with placebo. In conclusion, under both models the diamond is in the left side of the line of no effect, meaning that using diuretics to prevent the development of pre-eclampsia in pregnant woman has less risk compared to placebo. In Table 4.21 , the value of I^2 was found to be ($I^2 = 72.46\%$, P=0.0004) and we found the presence of great heterogeneity in this meta-analysis [90]. The AIC is smaller in the random-effects model than in the fixed-effects model, hence the random-effects model is preferably. Which means that there is some extra variation or over-dispersion due to random differences among the studies. The funnel plot in Figure 4.14 does not necessary informs us wether publication bias exist or not, because the studies are small (less than ten). When the studies are small, the funnel could not tell if publication bias exists or not. In Table 4.13 and Figure 4.15, studies 5 and 8, are identified as potential outliers and also to be a influential cases. The values of the covariance ratios for these studies also suggest that precision could be gained by its removal [91].

4.1.8 Odds Ratio

Table 4.23: Fixed-effects computations on log odds ratio.

Study	$\hat{\theta}_i =$ **log odds ratio**	$\hat{\sigma}_i^2$	$\hat{\sigma}_i^{-2}$	$\hat{\omega}_i$	95%C.I.
1	0.0418	0.1596	6.266	0.0500	(-0.7412,0.8249)
2	-0.9237	0.1177	8.496	0.0677	(-1.5962,-0.2512)
3	-1.1221	0.1780	5.618	0.0448	(-1.9491,-0.2952)
4	-1.4733	0.2989	3.346	0.0267	(-2.5449,-0.4017)
5	-1.3910	0.1143	8.749	0.0697	(-2.0536,-0.7284)
6	-0.2969	0.0146	68.493	0.5460	(-0.5340,-0.0598)
7	-0.2615	0.1207	8.285	0.0660	(-0.9424,0.4193)
8	1.0888	0.6864	1.457	0.0116	(-0.5350,2.7125)
9	0.1353	0.0679	14.728	0.1174	(-0.3753,0.6459)
Total			125.437	1	

Table 4.24: Random-effects computations on log odds ratio.

Study	$\hat{V}(\hat{\theta}_i^{(1)})$	$\hat{\tau}_i^{(1)} = \hat{\sigma}_i^{-2}$	$\hat{\omega}_i^{(1)}$
1	0.3982	2.5113	0.1069
2	0.3563	2.8066	0.1195
3	0.4166	2.4004	0.1022
4	0.5375	1.8605	0.0792
5	0.3529	2.8337	0.1206
6	0.2532	3.9494	0.1681
7	0.3593	2.7832	0.1185
8	0.9250	1.0811	0.0460
9	0.3065	3.2626	0.1389
Total		23.489	1

Table 4.25: Results of randomised controlled trials of effect of diuretics on pre-eclampsia from two methods of meta-analysis.

Method	**Fixed-effects**	**Random-effects**
log OR(95%)CI	-0.40(-0.57,-0.22)	-0.52(-0.92,-0.11)
τ^2		0.2386
P-value	0.0006	0.0006
Q(df=8)	27.26	27.26
I^2		71.44%
H^2		3.50
AIC	27.32	22.94
BIC	27.52	23.33
AICc	27.90	24.94
Deviance	27.26	20.88

Table 4.26: Results of the influence diagnostics for the trials of effects of diuretics on pre-eclampsia using log odds ratio as the measure of association.

study	**rstudent**	**dffits**	**cook.d**	**cov.r**	**tau2.del**	**QE.del**	**hat**	**weight**	**dfbs**	**inf**
1	0.8565	0.3003	0.0940	1.1737	0.3201	25.9889	0.1086	10.8631	0.2998	
2	-0.6331	-0.2332	0.0606	1.2634	0.3485	24.7472	0.1195	11.9497	-0.2334	
3	-0.9105	-0.3120	0.0999	1.1508	0.3135	24.1809	0.1045	10.4452	-0.3114	
4	-1.3212	-0.3904	0.1447	1.0103	0.2705	23.2907	0.0834	8.3394	-0.3949	
5	-1.6831	-0.6395	0.3035	0.8438	0.1938	17.9886	0.1205	12.0490	-0.6395	
6	0.3778	0.1523	0.0295	1.4148	0.3870	25.7280	0.1586	15.8556	0.1567	
7	0.3899	0.1419	0.0234	1.31853	0.3696	27.0997	0.1187	11.8660	0.1419	
8	1.6909	0.3661	0.1287	0.9519	0.2615	24.0065	0.0507	5.0663	0.3768	
9	1.2253	0.4909	0.2152	1.0508	0.2614	22.5164	0.1357	13.5657	0.4873	

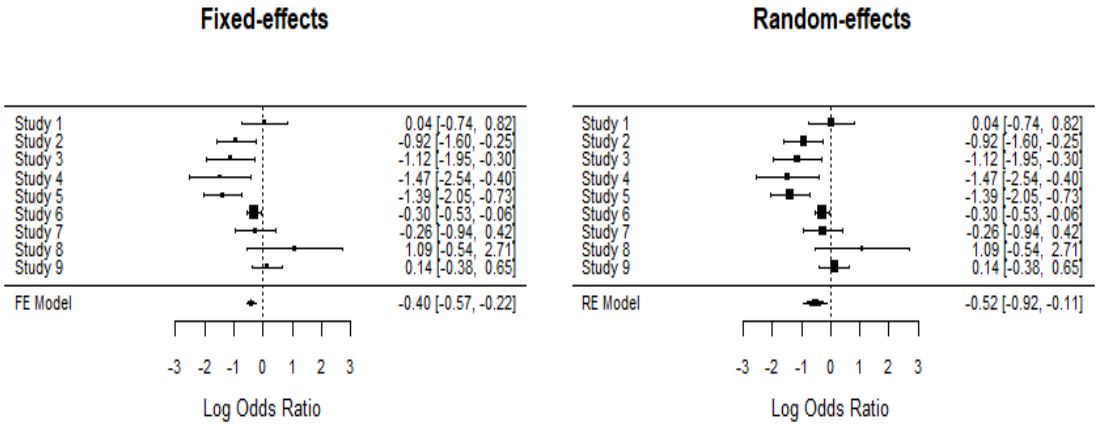

Figure 4.16 – Forest plot showing the results of nine studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia. The figure shows the log odds ratio of pre-eclampsia among those treated with diuretics versus the placebo group with corresponding confidence intervals in the individual studies and based on fixed-effects and random-effects models.

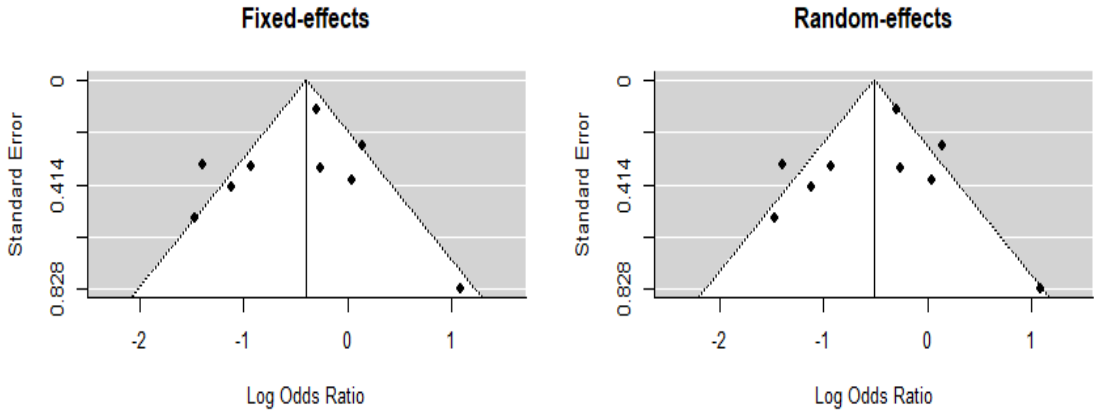

Figure 4.17 – Funnel plot shows the log odds ratio of nine studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia. The points corresponds to the treatment effects from individual trials and the diagonal or curved lines show the expected 95% confidence intervals around the summary estimate and based on fixed-effects and random-effects models.

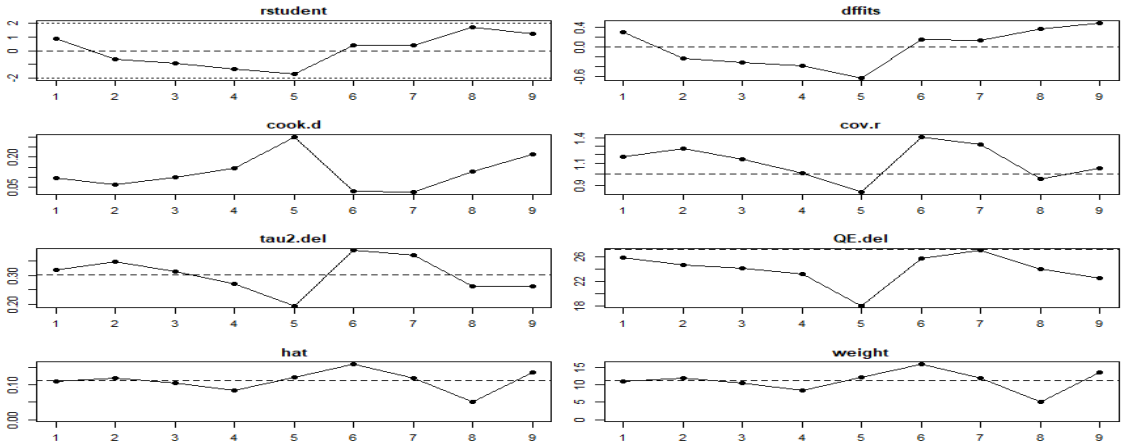

Figure 4.18 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for the 9 studies examining the use of diuretics during pregnancy to prevent the development of pre-eclampsia.

Consider the forest plots in Figure 4.16. Both plots show a summary effect on the bottom line, but the meaning of this summary effect is different in the two models. In the fixed-effects analysis we assume that the true effect size is the same in all studies, and the summary effect is our estimate of this common effect size. In random-effects analysis we assume that the true effect size varies from one study to the next, and that the studies in our analysis represent a random sample of effect size that could have observed. The summary effect is our estimate of the mean of these effects. Under the fixed-effects model we assume that the true effect size for all studies is identical, and the only reason the effect size varies between studies is sampling error (error in estimating the effect size). Therefore, when assigning weights to the different studies we can largely ignore the information about the same effect size in the larger studies. By contrast, under the random-effects model the goal is not to estimate one true effect, but to estimate the mean of a distribution of effects. Since each study provides information about a different effect size, we want to be sure that all these effect sizes are represented in the summary estimate. This means that we cannot discount a small study by giving it a very small weight (the way we would in a fixed-effects analysis). The estimate provided by that study may be imprecise, but it is information about an effect that no other study has estimated.

The pooled estimate of log odds ratio is -0.40(-0.57,-0.22) using the fixed-effects model and the pooled estimate of log odds ratio is -0.52(-0.92,-0.1)by using the random-

effects model. As expected the random-effects model provides a wider confidence interval of the pooled estimate than the fixed-effects model. By examining the weight allocated to each individual study Table (4.23 and 4.24) it can be seen that large studies gain relatively smaller weight in the random-effects model compared to that in the fixed-effects model (vice-versa). For example the largest study (study 6) in the meta-analysis captures more than half of the total weight (0.5460) in the fixed-effects model but a much smaller proportion of the total weight (0.1681) in the random-effects model (to illustrate the influence of the choice of different models, both the fixed-effects model and random-effects model estimates for this meta-analysis will be used through the project). The random-effects model considers between study variation but is not an ideal solution to the problem of heterogeneity because the model accommodates rather than explains the excess variability between studies. The assumption that underlying true effects to different studies are normally distributed is a rather bold hypothesis and may not be true in many cases. It is because of the wider confidence intervals when calculating random-effects estimates it has been suggested that random-effects estimates may be conservative than fixed-effects estimates [95]. However the example presented earlier shows the random-effects estimates may not be always conservative enriching the similar observation by others [94]. In addition by giving relatively larger weight to smaller studies the random-effects model may be more vulnerable to report bias than fixed-effects model [36].

Figure 4.16 shows that point estimates of treatment effects (log odds ratio of pre-eclampsia) are different across nine randomized controlled trials. When the outcome is measure using an odds ratio or other effect size, log scale is often used in the forest plot. On a log scale an odds ratio or other effect size and its reciprocal are symmetrical around 1. A unit change in log odds ratio or other effect size has the same interpretation. Over the scale corresponding to multiplying the odds by the same factor [96]. The effect size for each study is bounded by a confidence interval, reflecting the precision with which the effect size has been estimated in that study. The confidence interval for study 6 is noticeably narrower than the confidence interval of study 8, reflecting the fact that study 6 has greater precision. Also mean that study 6 has least effect than all the studies. Since the box is near the line of no action. If the box is far away from the line of no effect it means the is a large effect in that study. In studies 1, 7, 8 and 9, zero was included in the 95% confidence interval, it indicate that there studies are not statistical significance at 5% level of significance. In Studies 2, 3, 4, 5 and 6, zero was not included in the 95% confidence interval, indicating that there studies are statistical significance at 5% level of significance. In both models the overall mean of all studies does not include zero in the 95% confidence interval, which indicate that the studies were statistical significant at 5% level

of significance. Which indicate that the use of diuretics during pregnancy to prevent the development of pre-eclampsia is beneficial.

In studies 1, 8 and 9, the box is on the right of the line of no effect, which indicate that it favours placebo. Suggesting that, using the diuretic in preventing pregnancy to prevent the development of pre-eclampsia is harmful. In studies 2, 3, 4, 5, 6, 7 and 8 the box was in the left side of the line of no effect, which indicates that using the diuretics during pregnancy to prevent the development of pre-eclampsia is beneficial. On the plot the summary effect is shown on the bottom line. Under the fixed-effects model the summary odds ratio is $e^{-0.40}$=0.67, indicating that the risk of a patient being treated with diuretics to prevent the development of pre-eclampsia in pregnant woman is 33% lower than patients treated with placebo. Under the random-effects model the summary odds ratio is $e^{-0.52}$=0.59, indicating that the risk of a patient being treated with diuretics to prevent the development of pre-eclampsia in pregnant woman is 41% lower than patients treated with placebo. In both models the overall mean, indicate by the diamond is in the left side of the line of no effect, meaning that the final conclusion is that the use of diuretics during pregnancy to prevent the development of pre-eclampsia is beneficial. In Table 4.25 , we found the I^2 to be ($I^2 = 71.44\%$, p=0.0006). which indicate that there is great heterogeneity in this meta-analysis [90]. The AIC is smaller in the random-effects model compared to the fixed-effects model, which means that there the random-effects model is preferably. Hence there is some extra variation or over-dispersion due to random differences among the studies. The funnel plot in Figure 4.17 does not necessary informs us wether publication bias exist or not, because the studies are small (less than ten). When the studies are small, the funnel could not tell if publication bias exists or not. In Table 4.13 and Figure 4.18, studies 5 and 8, are identified as potential outliers and also to be a influential cases. The values of the covariance ratios for these studies also suggest that precision could be gained by its removal [91].

4.2 Application of the mixed-effects model for meta-analysis

Our example consists of a meta-analysis of randomized controlled trials comparing postoperative radiation therapy with and without adjuvant chemotherapy in patients with malignant gliomas. This example was given by [97], also used by [77, 72]. The outcome of interest is the number of patients surviving at 6, 12, 18 and 24 months. We use this example to present the longitudinal meta-analysis models described above to illustrate it efficiency. For multivariate meta-analysis of effect sizes reported at multiple time points. In the original data that was described and use by [97, 77, 72]. There were missing data for study 17 at months 6 and 18. In

this thesis we did not cover missing studies, as a result the study that was missing was not include in the analysis. We were quite aware of the advantage of the linear mixed-effects model, that it's can handle missing studies. Notice that in studies 10 and 15 in months 6, when computing odds ratio you get zero or infinity. We used a continuity correct of 0.5 for the entries that had zero. There reason we used odds ratio in the following analysis, it is because of its statistical properties. One of the statistical properties is that its valid regardless of the type of sampling used, which is not the case for other comparative measures for binary data.

Table 4.27: Number of survivors at four time points 6, 12, 18 and 24 months from randomized trials on the treatment of malignant gliomas using radio-therapy plus adjuvant chemotherapy versus radiotherapy alone.

Study	**Sample size, E(C)**	**Number of survivors, E(C)**			
		6 months	**12 months**	**18 months**	**24 months**
1	19(22)	16(20)	11(12)	4(8)	4(3)
2	34(35)	22(22)	18(12)	15(8)	15(6)
3	72(68)	44(40)	21(15)	10(3)	3(0)
4	72(68)	19(12)	14(5)	5(4)	2(3)
5	70(32)	62(27)	42(13)	26(6)	15(5)
6	183(94)	130(65)	80(33)	47(14)	30(11)
7	26(50)	24(30)	13(18)	5(10)	3(9)
8	61(55)	51(44)	37(30)	19(19)	11(15)
9	36(25)	30(17)	23(12)	13(4)	10(4)
10	45(35)	43(35)	19(14)	8(4)	6(0)
11	246(208)	169(139)	106(76)	67(42)	51(35)
12	386(141)	279(97)	170(46)	97(21)	73(8)
13	59(32)	56(30)	34(17)	21(9)	20(7)
14	45(15)	42(10)	18(3)	9(1)	9(1)
15	14(18)	14(18)	13(14)	12(13)	9(12)
16	26(19)	21(15)	12(10)	6(4)	5(1)

4.2.1 Results for the separate univariate random effects meta-analysis

Table 4.28: Meta-analysis results from separate univariate random-effects meta-analyses for the log odds ratio of surviving under the experimental versus the control treatments at month 6, 12, 18 and 24.

	Month 6	**Month 12**	**Month 18**	**Month 24**
log OR(95%)CI	0.22(0.02,0.43)	0.41(0.23,0.60)	0.49(0.27,0.71)	0.49(0.13,0.85)
τ^2	0	0	0	0.15
P-value	0.35	0.87	0.66	0.15
Q(df=8)	16.52	9.19	12.24	20.85
I^2	0.00%	0.00%	0.00%	32.49%
H^2	1	1	1	1.48
AIC	34.23	17.79	26.37	39.59
BIC	35.65	19.20	27.78	41.00
AICc	35.23	18.79	27.37	40.59
Deviance	30.23	13.79	22.37	35.59

Table 4.29: Meta-analysis results of the influence diagnostics for the log odds ratio of surviving under the experimental versus the control treatments at month 6.

study	**rstudent**	**dffits**	**cook.d**	**cov.r**	**tau2.del**	**QE.del**	**hat**	**weight**	**dfbs**	**inf**
1	-0.8797	-0.0955	0.0091	1.0118	0	15.7512	0.0116	1.1644	-0.0955	
2	-0.2895	-0.0620	0.0038	1.0459	0	16.4412	0.0439	4.3850	-0.0620	
3	-0.3850	-0.1228	0.0151	1.1018	0	16.3768	0.0924	9.2429	-0.1228	
4	1.5954	0.2193	0.0481	1.0188	0	13.9797	0.0185	1.8532	0.2193	
5	0.2300	0.0398	0.0016	1.0300	0	16.4721	0.0291	2.9124	0.0398	
6	-0.5151	-0.2114	0.0447	1.1684	0	16.2597	0.1441	14.4100	-0.2114	
7	2.3706	0.3175	0.1008	1.0179	0	10.9053	0.0176	1.7622	0.3175	
8	0.0446	0.0100	0.0001	1.0495	0	16.5230	0.0472	4.7219	0.0100	
9	1.0379	0.1783	0.0318	1.0296	0	15.4477	0.0287	2.8695	0.1783	
10	-1.0423	-0.0700	0.0049	1.0045	0	15.4387	0.0045	0.4493	-0.0700	
11	-0.7918	-0.4832	0.2334	1.3724	0	15.8981	0.2713	27.1309	-0.4832	*
12	-0.2891	-0.1622	0.0263	1.3150	0	16.4415	0.2396	23.9584	-0.1622	*
13	-0.0034	-0.0004	0.0000	1.0126	0	16.5250	0.0125	1.2452	-0.0004	
14	2.1448	0.2801	0.0784	1.0170	0	11.9250	0.0168	1.6760	0.2801	
15	-0.2333	-0.0121	0.0001	1.0027	0	16.4706	0.0027	0.2669	-0.0121	
16	-0.1460	-0.0207	0.0004	1.0200	0	16.5037	0.0195	1.9518	-0.0206	

Table 4.30: Meta-analysis results of the influence diagnostics for the log odds ratio of surviving under the experimental versus the control treatments at month 12.

study	rstudent	dffits	cook.d	cov.r	tau2.del	QE.del	hat	weight	dfbs	inf
1	-0.4427	-0.0660	0.0044	1.0222	0	8.9901	0.0217	2.1717	-0.0660	
2	0.7316	0.1402	0.0196	1.0367	0	8.6507	0.0354	3.5408	0.1402	
3	-0.0997	-0.0245	0.0006	1.0602	0	9.1761	0.0568	5.6758	-0.0245	
4	1.8475	0.2552	0.0651	1.0191	0	5.7727	0.0187	1.8723	0.2552	
5	0.8762	0.1921	0.0369	1.0481	0	8.4183	0.0459	4.5856	0.1921	
6	-0.2083	-0.0790	0.0062	1.1439	0	9.1426	0.1258	12.5830	-0.0790	
7	0.3375	0.0653	0.0043	1.0374	0	9.0721	0.0360	3.6029	0.0653	
8	-0.4443	-0.1133	0.0128	1.0650	0	8.9886	0.0610	6.1047	-0.1133	
9	0.4560	0.0814	0.0066	1.0319	0	8.9781	0.0309	3.0893	0.0814	
10	-0.7151	-0.1484	0.0220	1.0430	0	8.6746	0.0413	4.1260	-0.1484	
11	-0.8210	-0.4517	0.2040	1.3027	0	8.5121	0.2324	23.2384	-0.4517	*
12	0.3950	0.1992	0.0397	1.2542	0	9.0300	0.2027	20.2686	0.1992	*
13	-0.5341	-0.1152	0.0133	1.0466	0	8.9008	0.0445	4.4486	-0.1152	
14	0.8028	0.1057	0.0112	1.0173	0	8.5415	0.0170	1.7025	0.1057	
15	0.7629	0.0603	0.0036	1.0062	0	8.6040	0.0062	0.6200	0.0603	
16	-1.1249	-0.1753	0.0307	1.0243	0	7.9206	0.0237	2.3698	-0.1753	

Table 4.31: Meta-analysis results of the influence diagnostics for the log odds ratio of surviving under the experimental versus the control treatments at month 18.

study	rstudent	dffits	cook.d	cov.r	tau2.del	QE.del	hat	weight	dfbs	inf
1	-1.7656	-0.2822	0.0796	1.0255	0	9.1259	0.0249	2.4909	-0.2822	
2	0.9518	0.2076	0.0431	1.0476	0	11.3375	0.0454	4.5429	0.2076	
3	1.1370	0.1912	0.0365	1.0283	0	10.9507	0.0275	2.7494	0.1912	
4	-0.4338	-0.0656	0.0043	1.0229	0	12.0552	0.0224	2.2370	-0.0656	
5	0.9006	0.2022	0.0409	1.0504	0	11.4323	0.0480	4.7990	0.2022	
6	0.6132	0.2195	0.0482	1.1281	0	11.8674	0.1135	11.3547	0.2195	
7	-0.8928	-0.1683	0.0283	1.0355	0	11.4464	0.0343	3.4300	-0.1683	
8	-1.6885	-0.5029	0.2529	1.0887	0	9.3923	0.0815	8.1483	-0.5029	
9	0.9440	0.1676	0.0281	1.0315	0	11.3523	0.0306	3.0574	0.1676	
10	0.0454	0.0079	0.0001	1.0303	0	12.2414	0.0294	2.9429	0.0079	
11	-0.4908	-0.2863	0.0820	1.3402	0	12.0025	0.2539	25.3858	-0.2863	*
12	0.6894	0.3266	0.1067	1.2245	0	11.7682	0.1833	18.3309	0.3266	
13	-0.3047	-0.0742	0.0055	1.0592	0	12.1506	0.0559	5.5927	-0.0742	
14	0.7000	0.0723	0.0052	1.0107	0	11.7534	0.0106	1.0560	0.0723	
15	0.3796	0.0466	0.0022	1.0151	0	12.0993	0.0149	1.4857	0.0466	
16	-0.5114	-0.0801	0.0064	1.0246	0	11.9819	0.0240	2.3964	-0.0801	

Table 4.32: Meta-analysis results of the influence diagnostics for the log odds ratio of surviving under the experimental versus the control treatments at month 24.

study	rstudent	dffits	cook.d	cov.r	tau2.del	QE.del	hat	weight	dfbs	inf
1	0.0351	-0.0030	0.0000	1.0901	0.1663	20.8437	0.0390	3.8987	-0.0030	
2	1.3236	0.3756	0.1339	1.0057	0.1247	18.2696	0.0708	7.0754	0.3783	
3	0.8174	0.0942	0.0089	1.0159	0.1495	20.1053	0.0132	1.3216	0.0941	
4	-1.0277	-0.1856	0.0345	1.0420	0.1524	19.7185	0.0304	3.0425	-0.1849	
5	-0.1557	-0.0614	0.0040	1.1676	0.1807	20.8355	0.0705	7.0538	-0.0609	
6	-0.1879	-0.0901	0.0095	1.2654	0.1973	20.8185	0.1139	11.3878	-0.0917	
7	-1.2801	-0.2888	0.0827	1.0365	0.1429	18.9263	0.0502	5.0214	-0.2898	
8	-2.0342	-0.5909	0.2721	0.8628	0.0711	15.5674	0.0945	9.4503	-0.5902	
9	0.2778	0.0566	0.0033	1.1248	0.1713	20.7041	0.0568	5.6761	0.0560	
10	1.2356	0.1485	0.0220	1.0026	0.1444	19.1747	0.0140	1.4006	0.1492	
11	-0.5105	-0.2258	0.0626	1.3149	0.1959	19.9011	0.1598	15.9795	-0.2362	
12	2.4796	1.2068	0.7302	0.6034	0.0000	14.7021	0.1114	11.1362	1.1714	*
13	0.1749	0.0334	0.0012	1.1879	0.1833	20.7572	0.0815	8.1521	0.0333	
14	0.6608	0.1026	0.0106	1.0386	0.1535	20.3172	0.0244	2.4421	0.1021	
15	-0.7193	-0.1699	0.0294	1.0874	0.1621	20.2726	0.0468	4.6761	-0.1685	
16	0.8096	0.1230	0.0152	1.0295	0.1508	20.0725	0.0229	2.2859	0.1228	

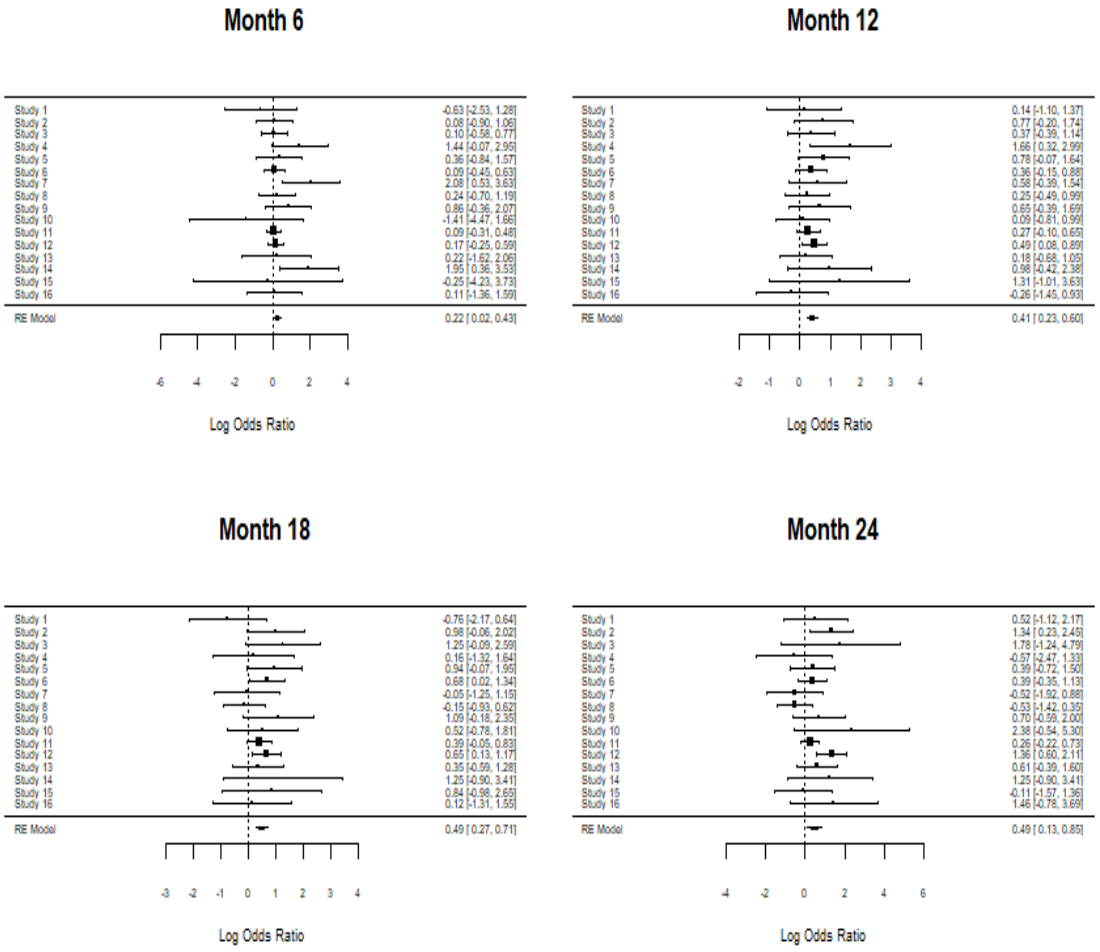

Figure 4.19 – forest plots for month 6, 12, 18 and 24 of the post-operative treatment with either radiotherapy plus chemotherapy or radiotherapy alone in patients with malignant gliomas from 16 studies.

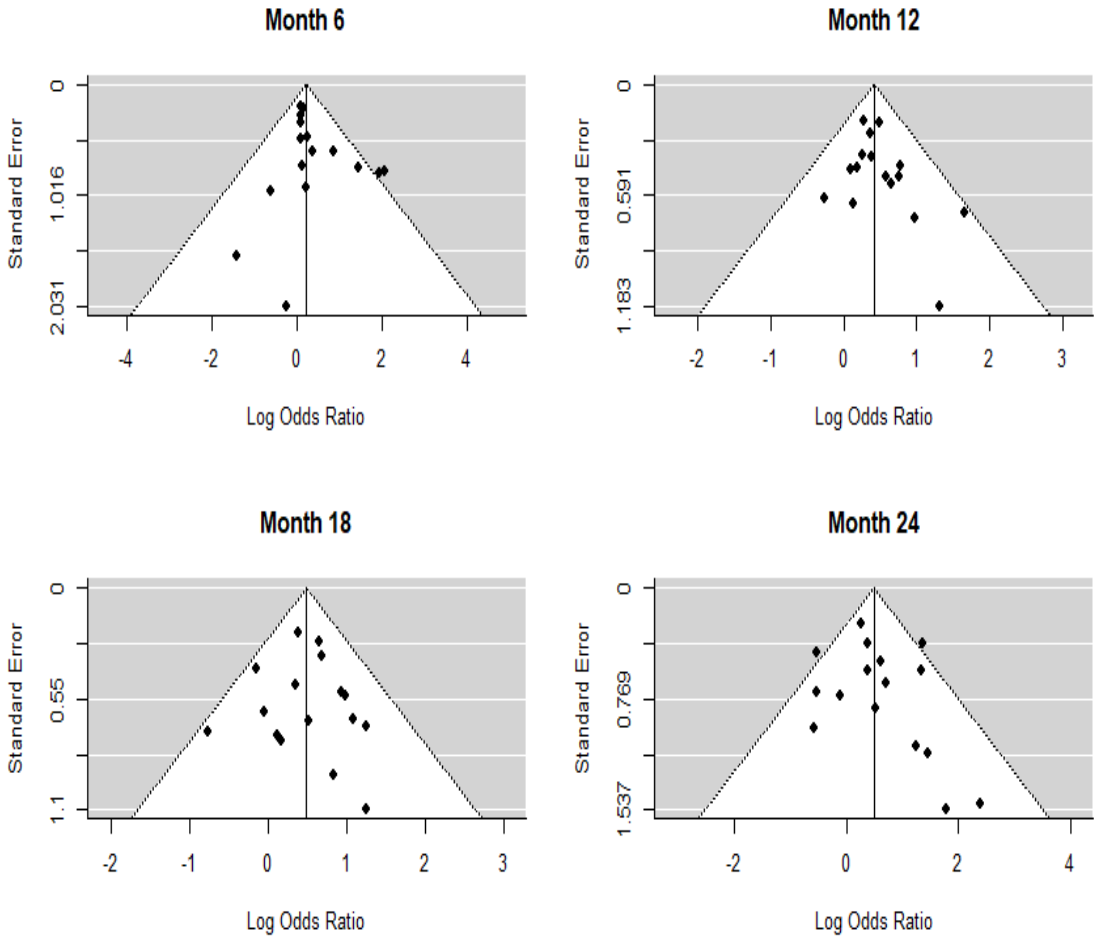

Figure 4.20 – Funnel plot for month 6, 12, 18 and 24 of the post-operative treatment with either radiotherapy plus chemotherapy or radiotherapy alone in patients with malignant gliomas from 16 studies.

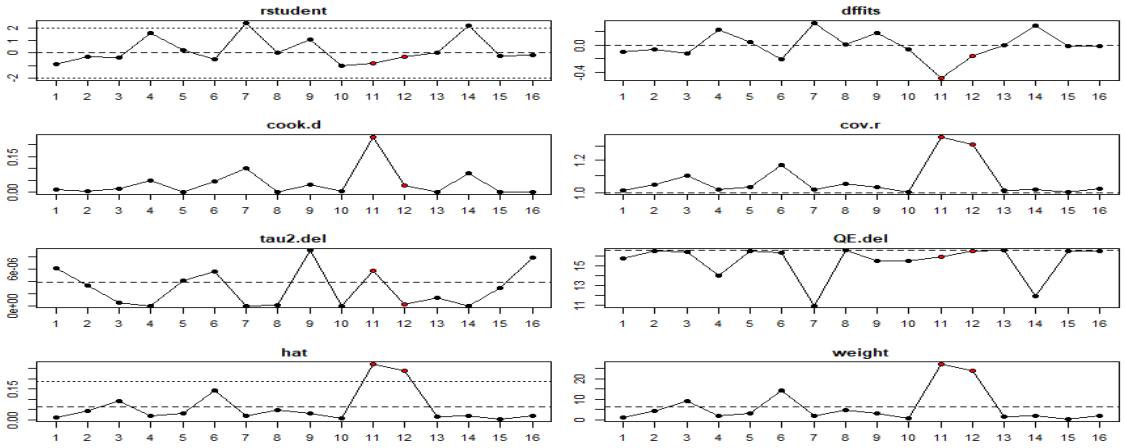

Figure 4.21 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for 16 studies, for month 6 of the post-operative treatment with either radiotherapy plus chemotherapy or radiotherapy alone in patients with malignant gliomas.

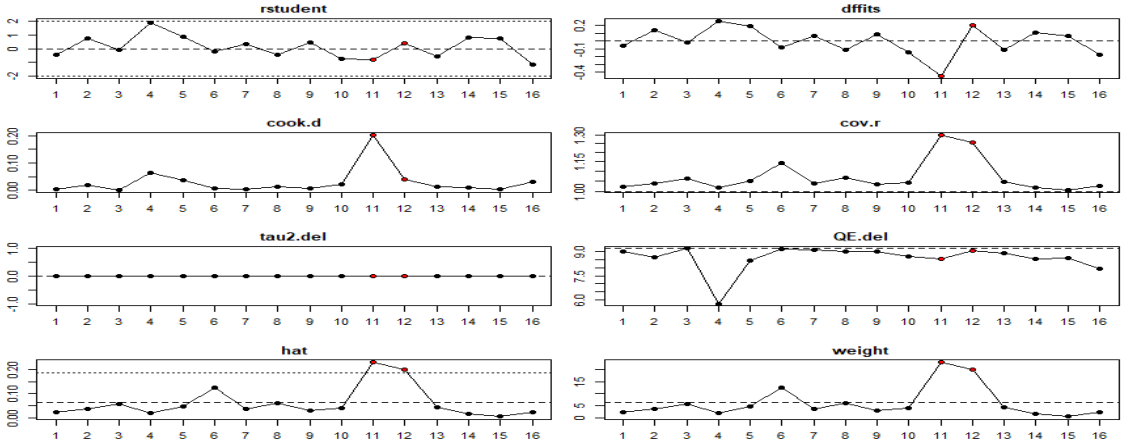

Figure 4.22 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for 16 studies, for month 12 of the post-operative treatment with either radiotherapy plus chemotherapy or radiotherapy alone in patients with malignant gliomas.

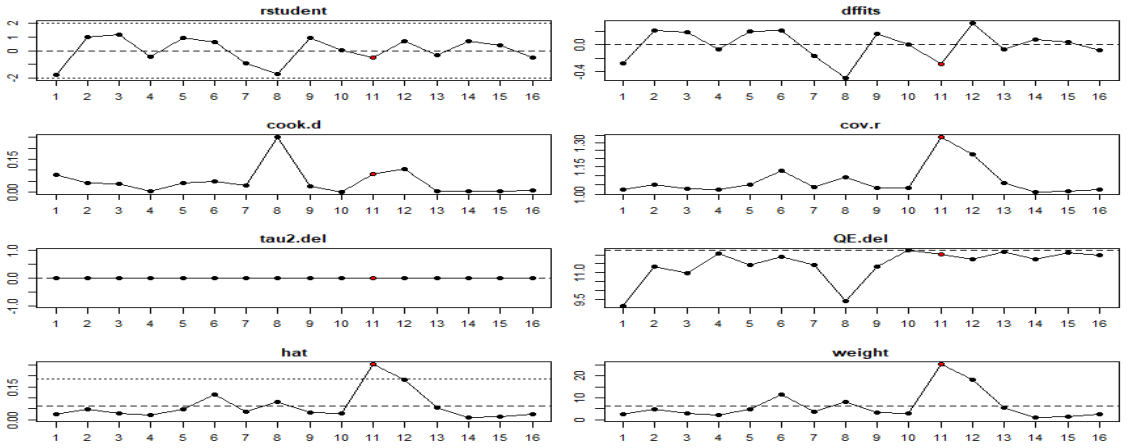

Figure 4.23 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for 16 studies, for month 18 of the post-operative treatment with either radiotherapy plus chemotherapy or radiotherapy alone in patients with malignant gliomas.

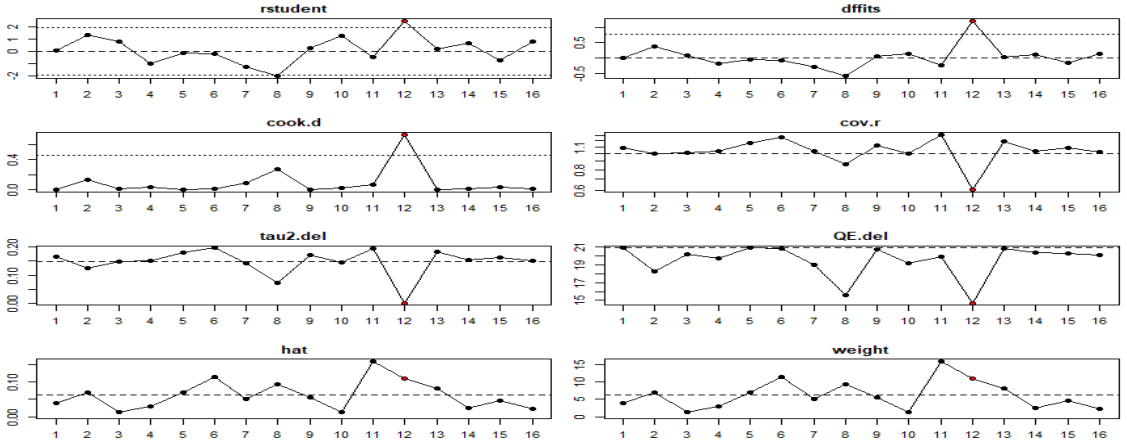

Figure 4.24 – Plot of the externally standardized residuals, DFFITS values, Cook's distance, covariance ratios, estimates of τ^2 and test for residual heterogeneity when each study is removed in turn, hat values, and weights for 16 studies, for month 24 of the post-operative treatment with either radiotherapy plus chemotherapy or radiotherapy alone in patients with malignant gliomas.

Figure 4.19 above show the forest plot of separate univariate random-effects meta-analyses for months 6,12,18 and 24. In month 6 studies 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 15 and 16, zero was included in the confidence interval indicating that the studies were not statistically significant. Moreover since the line of no effect, results

crossing this line cannot show whether the intervention is better or worse than the control. Study (7 and 14) zero was not included in the confidence interval indicating that the studies were statistically significant. In month 12 studies 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15 and 16, zero was included in the confidence interval indicating that the were not statistically significant. In studies 4 and 12 zero was not included in the confidence interval indicating that the were statistically significant. In month 18 studies 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 15 and 16, zero was included in the confidence interval indicating that the studies were not statistically significant. In studies 6 and 12 zero was not included in the confidence interval indicating that the were statistically significant. In month 24 studies 1,2,3,4,5,6,7,8,9,10,11,13,14,15 and 16 zero was included in the confidence interval indicating that the studies were not statistically significant. But study(2 and 12) zero was not included in the confidence interval indicating that the were statistically significant. The result in Table 4.28 clearly shows that the odds of survival were significant higher in the experimental group compared with the control group. This was consistent across all longitudinal time points from month 6 to month 24. All four overall odds ratios at month 6, 12, 18 and 24 months were statistically significant because zero was not included in the confidence interval. The least log odds ratio (0.22) was at month 6 which increased at month 12 (0.41) and at month 18 (0.49), it remain the same at month 24 (0.49).
In month 6 studies 1 and 14, the box is on the left of the line of no effect, which indicate that it favours experimental group. Suggesting that, using the radiotherapy plus chemotherapy in patients with malignant gliomas is beneficial. In studies 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 15 the box was in the right side of the line of no effect, indicated that using the radiotherapy alone in patients with malignant gliomas is beneficial. In month 12 study (16), the box is on the left of the line of no effect, which indicate that it favours the experimental group. Suggesting that, using the radiotherapy plus chemotherapy is beneficial. In studies 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15, the box was on the right of the line of no effect, which indicate that it favours the control group. Suggesting that, using radiotherapy in patients with malignant gliomas is beneficial. In month 18 studies 1, 7 and 8, the box was on the left of the line of no effect, which indicate that it favours experimental group. Suggesting that, using radiotherapy plus chemotherapy in patients with malignant gliomas is beneficial. In studies 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 15 and 16 the box was in the right side of the line of no effect, indicating that using radiotherapy alone in patients with malignant gliomas is beneficial. In month 24 studies 4, 7, 8 and 15, the box was on the left of the line of no effect, which indicate that it favours experimental group. Suggesting that using radiotherapy plus chemotherapy in patients with malignant gliomas is beneficial. In studies 1, 2, 3, 5, 6, 9, 10, 11, 12, 13, 14 and 16, in month 6, 12, 18 and 24 the overall mean, indicated by the diamond was on the left side of

the line of no effect, meaning that the final conclusion is that the use of radiotherapy alone in patients with malignant gliomas beneficial. In Table 4.28, month 6, 12 and 18 we found the I^2 value to be $I^2 = 0\%$, which means that the is no heterogeneity in month 6, 12 and 18 in this meta-analysis. But in month 24, we found the value of I^2 to be $I^2 = 32.49\%$ which means that the is moderate heterogeneity at month 24 in this meta-analysis. Based on the plot (Figure 4.20) there is no visual indication of publication bias, since the plot resembles a symmetrical inverted funnel because the treatment effect estimates from smaller studies scatter more widely at the bottom of the graph, with the spread narrowing with increasing precision among larger studies. In month 6 and 12, in Table 4.29 and Figure 4.23 shows that removing study (11 and 12) would yield little change in the amount of residual heterogeneity, but their influence on the model fit is more considerable [93]. In month 18, in Table 4.31 and Figure 4.23 shows that the removal of study 11 would yield little change in the amount of residual heterogeneity, but their influence on the model fit is more considerable [93]. In month 24, Table 4.32 and Figure 4.24 shows that the influence of study 12 on the model fit is more considerable, also its removal would yield little change in the amount of residual heterogeneity [93].

4.2.2 Results for the linear mixed-effects model for meta-analysis

Table 4.33: Meta-analysis results for models 1 to 3 from the linear mixed-effects model for the log odds ratio of surviving under experimental treatment compared to the control treatment using data for 16 trials.

	Model 1	**Model 2**	**Model 3**
Covariance structures between random time effects($\sum$)	Indep	CS	HAR(1)
Within-study errors ($\boldsymbol{R}_i$)	Indep	Indep	Indep
Log odds ratio estimates			
Month 6	0.22(0.02,0.43)	0.23(0.01,0.46)	0.22(0.02,0.43)
Month 12	0.41(0.23,0.60)	0.43(0.23,0.63)	0.41(0.22,0.59)
Month 18	0.49(0.27,0.71)	0.51(0.27,0.74)	0.49(0.23,0.74)
Month 24	0.49(0.13,0.85)	0.48(0.21,0.76)	0.50(0.13,0.87)
Between study variance estimates			
Month 6	0.00	0.02	0.00
Month 12	0.00		0.00
Month 18	0.00		0.04
Month 24	0.15		0.18
Model Fit			
AIC	117.97	114.95	117.64
BIC	134.73	127.51	136.49
AICc	120.80	116.53	121.24
Deviance	101.97	102.95	99.64

Table 4.34: Meta-analysis results for models 4 to 6 from the linear mixed-effects model for the log odds ratio of surviving under experimental treatment compared to the control treatment using data for 16 trials.

	Model 4	**Model 5**	**Model 6**
Covariance structures between random time effects($\sum$)	Indep	HAR(1)	UN
Within-study errors ($\boldsymbol{R}_i$)	HAR(1)	HAR(1)	HAR(1)
Log odds ratio estimates			
Month 6	0.19(-0.01,0.39)	0.19(-0.01,0.40)	0.21(0.00,0.43)
Month 12	0.38(0.20,0.55)	0.37(0.19,0.55)	0.38(0.20,0.56)
Month 18	0.44(0.22,0.66)	0.43(0.20,0.67)	0.43(0.20,0.66)
Month 24	0.44(0.12,0.77)	0.43(0.06,0.81)	0.42(0.05,0.79)
Between study variance estimates			
Month 6	0.00	0.00	0.01
Month 12	0.00	0.00	0.00
Month 18	0.00	0.02	0.02
Month 24	0.10	0.19	0.20
Model Fit			
AIC	104.67	105.73	115.34
BIC	121.43	124.58	144.66
AICc	107.50	109.33	124.67
Deviance	88.67	87.73	87.34

In Table(4.33 and 4.34) shows the results of applying the linear mixed-effects model in equation(3.2) to the example of the data by Fine [97], using model 1 to 6. The results were the same for the independence model as the one obtained from the separate meta-analysis in Table 4.28. As a result performing univariate meta-analysis at each time point separately is equivalent to a independence model. There is a slightly differences of the log odds ratio estimate between model 1 to 6 in Table (4.33 and 4.34). The log odds ratio estimates from model 2 and 6 the pattern of the results was the same, month 6 was the least it increases in month 12 and 18 then it decreases in month 24. But in model 1, 4 and 5 the estimates were also have the same pattern, in month 6 were the least estimates, they increase in month 12 up till month 18 and stayed constant in month 24. In model 3 the estimates of log odds ratio, increases from month 6 till month 24. In month 6, 12, 18 and 24 in all six model the log odds ratio estimates of surviving were significant higher for the experimental treatment compared to the control treatment since the 95% confidence interval does not include zero at 5% level of significance. Although month 6 in model 4 and 5 show that the log odds ratio estimates were not statistically significant at 5% level of significance, the corresponding p-values were slightly above 0.05. All in all, the likelihood of survival is significantly better under the experimental treatment compared to the

control treatment. The results of the linear mixed-effects model for the between study variances ranges from 0.00 to 0.20 and were not statistically different from zero. Model 4 and 5 had much better fit than the rest of the models, since the values of Akaike Information Criterion(AIC) were the smaller values among the rest which indicate better fit. Model 2, 3 and 6 performed slightly better than the independence model and there were very slight differences in the model fit between these four model.

Chapter 5

Discussion and Conclusion

Under the fixed-effects model, we assumed that the effect sizes in our meta-analysis are different only because of sampling error and that they all share a common mean. Realistically, we know that each study leads to a different effect size, but each effect size is an estimate of common mean. Effects sizes differ from each other only because each study used a different sample of participants and also because of differences in the way studies were conducted. This assumption is possible when studies are close replications of one another (they use same procedures and measures). Under the random-effects model, we assume sampling variation as in our fixed-effects model assumption and random variation because the effect sizes themselves are sampled from a population of effect sizes.

Under the random-effects model, the summary effect is an estimate of the mean of a distribution of true effects. Study weights are more uniform (similar to one another) under the random-effects model than under the fixed-effects model. Large studies are assigned less relative weight and small studies are assigned more relative weight as compared with the fixed-effects model. The standard error of the summary effect and, therefore, the confidence intervals for the summary effect are wider under the random-effects model than under the fixed-effects model. Under the fixed-effects model the only source of variation is the within study estimation error [46]. With a large enough sample size, accumulated across studies, this source of variation will diminish and the common effect size will be estimated precisely. Under the random-effects model there are two sources of variation, namely, within study estimation error variance and between studies variance. With a large enough sample size, accumulated across studies, the effect of first source of variation will diminish. However, if the between studies variance is substantial, the only way to obtain good precision is to increase the number of studies. If we increase the sample size in a few studies we may know the effect sizes in those studies very precisely,

but still not have a precise estimate of the mean across all studies. The selection of a model should be based solely on the question of which model fits the distribution of effect sizes better and thus takes account of the relevant source(s) of error [46]. When studies are gathered from the published literature, the random-effects model is generally a more plausible match. The addition of the nonzero variance component between strata, $\hat{\tau}^2$ to the variance of the estimate to obtain the unconditional variance has the effect of adding a constant to all of the weights. Thus, the random-effects analysis shrinks the weights towards the average, so that the resulting estimate is closer to the unweighted mean of the $\hat{\theta}_i$ than is the fixed-effects model [29]. If the estimate of this variance components is zero or nearly so, the random-effects analysis differs trivially from the fixed-effects analysis. Thus one strategy could be to always adopt a random-effects model because there is usually some extra variation or over-dispersion and if not then $\hat{\tau}^2 = 0$ and the fixed-effects will result. However this could sacrifice power in those cases where a fixed-effects model actually applies because even when $\tau^2 = 0$, the estimate $\hat{\tau}^2 = 0$ will vary and a small value will still inflate the variance of the estimate $\hat{\mu}_{\theta}^{(1)}$ [29]. Thus it is customary to first conduct a test of homogeneity and to conduct a random-effects analysis only if significant heterogeneity is detected. While the random-effects model is often the appropriate model, there are cases where it cannot be implemented properly because there are too few studies to obtain an accurate estimate of the between-studies variance. In Table(4.4, 4.8 and 4.12), the pattern of the results were the same across all three measures of association. We found that the values of the Akaike Information Criterion(AIC), were smaller values which indicate better fit, the results show that the fixed-effects model had much better fit than the random-effects model. Even though we found the presence of heterogeneity as a result of random sampling error[33, 38]. Such random variation is greater in smaller studies and is less of a problem in larger studies [34]. All three effect sizes, suggested that using both new drug and placebo to patients with duodenal ulcers is beneficial. In the example by Yusuf [92], of a meta-analysis of nine clinical trails investigating the effect of taking diuretics during pregnancy on the risk of pre-eclampsia. All three measures, the pooled estimate suggested that treating patients with diuretics in preventing the development of pre-eclampsia in pregnant woman is beneficial. But when treated with placebo is harmful. Random-effects model was the appropriate model for all three measures of association, which means that there is some extra variation or over-dispersion due to random differences among the studies.

In the example by Fine [97] where, we applied a linear mixed-effects model which borrows ideas from multivariate meta-analysis. The simplest approach which does not account for correlation, that is, the independence model where separate meta-

analysis were done at each of the time points was compared against models where correlation was accounted for in different alternatives, including random study effects, correlated random time effects and/or correlated within-study errors, or unstructured covariance structures. From the results above, the random study effect model or the correlated random time effect model that accounts for correlation between effect sizes, yield similar results to the independence model where separate meta-analysis are done at each time point. The results from models 4 and 5 clearly show the benefit of accounting for within study serial correlations between effect sizes. Accounting for correlation using the compound symmetry showed very little benefit compared to the independence model. In our example, the autoregressive covariance structure yielded more precise estimates compared to the compound symmetry covariance structure.Furthermore the results from months 6 to month 24 showed that the odds of survival under the experimental treatment were significantly higher compared to the control treatment. The confidence interval for parameter estimates show that the best performing model 4 had narrowest confidence intervals compared to the other five models at the four time points. In the multivariate example used, Figure 4.20, the subjective impression does support presence of asymmetry, therefore there is no visual indication of publication bias. Since the graphs resembles symmetrical inverted funnel plots the treatment effect estimates from smaller studies scatter more widely at the bottom of the graph, with the spread narrowing with increasing precision among larger studies. Funnel plot is not the only method to detect publication bias in meta-analysis. The is a numerous number of statistical methods to detect and adjunct for publication bias in meta-analysis.

Simulations to confirm whether our findings of the benefit of taking account of within study correlations using autoregressive structure are needed. Since the example that we used cannot generalized other data sets. Our modeling approach was to estimate point estimates at each fixed time point. However, this thesis has potential to be extended in some respect. In order to improve the estimation and for further research purposes. The alternative approach by Ahn and French [98], were they treated time as a continuous covariate and explore both linear and non-linear models can be considered. Moveover, for further research purposes the thesis can be extended where we have multiple effect sizes of different types, at each time point. Such extensions are well suited in the prevailing of longitudinal studies, where a number of outcomes are measured at multiple time points. However, this creates complexity in the modelling structure.

References

[1] Ponnuraja C. An empirical investigation of meta-analysis using randomized controlled clinical trials in a particular centre. *Journal of Applied Quantitative Methods*, 5(1):166–175, 2010.

[2] DerSimonian R. and Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3):177–188, 1986.

[3] Halvorsen K.T., Burdick E., Colditz G.A., Frazier H.S., and Mosteller F. Combining results from independent investigations. *Medical Uses of Statistics*, 2(22):413–426, 1992.

[4] Finckh A. and Tramèr M.R. Primer: Strengths and weaknesses of meta-analysis. *Nature Reviews Rheumatology*, 4(3):146, 2008.

[5] Doi S.A., Barendregt J.J., Khan S., Thalib L., and Williams G.M. Advances in the meta-analysis of heterogeneous clinical trials i: The inverse variance heterogeneity model. *Contemporary Clinical Trials*, 45:130–138, 2015.

[6] Glass G.V. Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10):3–8, 1976.

[7] Light R. and Smith P. Accumulating evidence: Procedures for resolving contradictions among different research studies. *Harvard Educational Review*, 41(4):429–471, 1971.

[8] Pillemer D. and Light R. Synthesizing outcomes: How to use research evidence from many studies. *Harvard Educational Review*, 50(2):179–195, 1980.

[9] Borenstein M., Hedges L.V., Higgins J.P., and Rothstein H.R. *Introduction to meta-analysis*. United Kingdom, John Wiley & Sons, 2011.

[10] Ray J.W. and Shadish W.R. How interchangeable are different estimators of effect size?. *Journal of Consulting and Clinical Psychology*, 64(6):1316, 1996.

[11] Glass G.V., Smith M.L., and McGaw B. *Meta-analysis in Social Research.* Beverly Hills, SAGE Publications, 1981.

[12] Morris S.B. and DeShon R.P. Correcting effect sizes computed from factor analysis of variance for use in meta-analysis. *Psychological Methods*, 2(2):192, 1997.

[13] Schmidt F.L. and Hunter J.E. *Methods of meta-analysis: Correcting error and bias in research findings*. Los Angeles, London, New Delhi, Singapore, Washington DC, SAGE Publications, 2014.

[14] Rosenthal R. *Meta-analytic procedures for social research*. Newbury Park, London, New Delhi, SAGE Publications, 1991.

[15] Morris S.B. and DeShon R.P. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7(1):105, 2002.

[16] Cochran W.G. Problems arising in the analysis of a series of similar experiments. *Supplement to the Journal of the Royal Statistical Society*, 4(1):102–118, 1937.

[17] Yates F. and Cochran W.G. The analysis of groups of experiments. *The Journal of Agricultural Science*, 28(4):556–580, 1938.

[18] Hedges L.V. Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2):107–128, 1981.

[19] Hedges L.V. Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, 92(2):490, 1982.

[20] Hedges L.V. A random effects model for effect sizes. *Psychological Bulletin*, 93(2):388, 1983.

[21] Law K.S., Schmidt F.L., and Hunter J.E. Nonlinearity of range corrections in meta-analysis: Test of an improved procedure. *Journal of Applied Psychology*, 79(3):425, 1994.

[22] Glass G.V. Integrating findings: The meta-analysis of research. *Review of Research in Education*, 5(1):351–379, 1977.

[23] Smith M.L. and Glass G.V. Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(9):752–760, 1977.

[24] Hall J.A. Gender effects in decoding nonverbal cues. *Psychological Bulletin*, 85(4):845, 1978.

[25] Uguroglu M.E. and Walberg H.J. Motivation and achievement: A quantitative synthesis. *American Educational Research Journal*, 16(4):375–389, 1979.

[26] Johnson D.W., Maruyama G., Johnson R., Nelson D., and Skon L. Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin*, 89(1):47, 1981.

[27] Cronbach L.J., Ambron S.R., Dornbusch S.M., Hess R.D., Hornik R.C., Phillips D.C., and Weiner S.S. *Toward reform of program evaluation*. San Francisco, Jossey-Bass, 1980.

[28] Giaconia R.M. and Hedges L.V. Identifying features of effective open education. *Review of Educational Research*, 52(4):579–602, 1982.

[29] Lachin J.M. *Biostatistical methods: The assessment of relative risks*. New York, Chichester, Weinheim, Brisbane, Singapore, Toronto, John Wiley & Sons, 2009.

[30] Woolf B. On estimating the relation between blood group and disease. *Annals of Human Genetics*, 19(4):251–253, 1955.

[31] Gavaghan D.J., Moore R.A., and McQuay H.J. An evaluation of homogeneity tests in meta-analyses in pain using simulations of individual patient data. *Pain*, 85(3):415–424, 2000.

[32] Higgins J.P., Thompson S.G., Deeks J.J., and Altman D.G. Measuring inconsistency in meta-analyses. *British Medical Journal*, 327(7414):557, 2003.

[33] Light R.J. Accumulating evidence from independent studies: What we can win and what we can lose. *Statistics in Medicine*, 6(3):221–228, 1987.

[34] Anello C. and Fleiss J.L. Exploratory or analytic meta-analysis: Should we distinguish between them?. *Journal of Clinical Epidemiology*, 48(1):109–116, 1995.

[35] Davey S.G., Egger M., and Phillips A.N. Meta-analysis. beyond the grand mean?. *British Medical Journal*, 315(7122):1610–1614, 1997.

[36] Greenland S. Invited commentary: A critical look at some popular meta-analytic methods. *American Journal of Epidemiology*, 140(3):290–296, 1994.

[37] Thompson S.G. Why sources of heterogeneity in meta-analysis should be investigated. *British Medical Journal*, 309(6965):1351, 1994.

[38] Bailey K.R. Inter-study differences: How should they influence the interpretation and analysis of results?. *Statistics in Medicine*, 6(3):351–358, 1987.

[39] Hedges L.V. Meta-analysis. *Journal of Educational Statistics*, 17(4):279–296, 1992.

[40] Hedges L.V. and Vevea J.L. Fixed-and random-effects models in meta-analysis. *Psychological Methods*, 3(4):486, 1998.

[41] Hunter J.E and Schmidt F.L. Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, 8(4):275–292, 2000.

[42] Becker B.J. *The generalizability of empirical research results*. Baltimore, MD, US, Johns Hopkins University Press, 1996.

[43] Gaver P.G., Draper D., Goel P.K., Greenhouse J.B., Hedges L.V., Morris C.N., and Waternaux C. *Combining information: Statistical issues and opportunities for research*. Washigton DC, National Academy Press, 1992.

[44] Osburn H.G. and Callender J. A note on the sampling variance of the mean uncorrected correlation in meta-analysis and validity generalization. *Journal of Applied Psychology*, 77(2):115, 1992.

[45] Schmidt F.L. and Hunter J.E. Comparison of three meta-analysis methods revisited: An analysis of johnson, mullen, and salas(1995). *Journal of Applied Psychology*, 84(1):144–148, 1999.

[46] Hedges L.V. *Fixed effects models*. New York, The Russell Sage Foundation, 1994.

[47] Mood F.A. and Boes D.C. *Introduction to the theory of statistics*. New York, McGraw-Hill, 1963.

[48] Orwin R.G., Cooper H., and Hedges L.V. *The handbook of research synthesis*. New York, Russell Sage Foundation, 1994.

[49] Rao P.S, Kaplan J., and Cochran W.G. Estimators for the one-way random effects model with unequal error variances. *Journal of the American Statistical Association*, 76(373):89–97, 1981.

[50] Song F., Hooper L., and Loke Y. Publication bias: What is it? how do we measure it? how do we avoid it?. *Open Access Journal of Clinical Trials*, 2013(5):71–81, 2013.

[51] McDaniel M.A., Rothstein H.R., and Whetzel D.L. Publication bias: A case study of four test vendors. *Personnel Psychology*, 59(4):927–953, 2006.

[52] Clarke M. and Clarke T. A study of the references used in cochrane protocols and reviews: Three bibles, three dictionaries, and nearly 25,000 other things. *International Journal of Technology Assessment in Health Care*, 16(3):907–909, 2000.

[53] Egger M., Smith G.D., and Altman D. *Systematic reviews in health care: Meta-analysis in context*. London, BMJ Publication Group, 2008.

[54] Juni P., Holenstein F., Sterne J., Bartlett C., and Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *International Journal of Epidemiology*, 31(1):115–123, 2002.

[55] Gotzsche P.C. Reference bias in reports of drug trials. *British Medical Journal*, 295(6599):654–656, 1987.

[56] Ravnskov U. Frequency of citation and outcome of cholesterol lowering trials. *British Medical Journal*, 305(6855):717, 1992.

[57] Tramer M.R., Reynolds D.J.M., Moore R.A., and McQuay H.J. Impact of covert duplicate publication on meta-analysis: A case study. *British Medical Journal*, 315(7109):635–640, 1997.

[58] Ter Riet G., Korevaar D.A., Leenaars M., Sterk P.J., Van Noorden C.J., Bouter L.M., and Hooft L. Publication bias in laboratory animal research: A survey on magnitude, drivers, consequences and potential solutions. *Plos One*, 7(9):43404, 2012.

[59] Macleod M.R., O'collins T., Howells D.W., and Donnan G.A. Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke*, 35(5):1203–1208, 2004.

[60] Taubes G. and Mann C.C. Epidemiology faces its limits. *Science*, 269(5221):164, 1995.

[61] Ioannidis J.P. and Trikalinos T.A. Early extreme contradictory estimates may appear in published research: The proteus phenomenon in molecular genetics research and randomized trials. *Journal of Clinical Epidemiology*, 58(6):543–549, 2005.

[62] Psaty B.M. and Kronmal R.A. Reporting mortality findings in trials of rofecoxib for alzheimer disease or cognitive impairment: A case study based on documents from rofecoxib litigation. *The Journal of the American Medical Association*, 299(15):1813–1817, 2008.

[63] Choi S.W. and Lam D.M.H. An alarm for a false alarm. *Anaesthesia*, 71(1):106–108, 2016.

[64] Sterne J.A., Gavaghan D., and Egger M. Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology*, 53(11):1119–1129, 2000.

[65] Thaler K., Kien C., Nussbaumer Van Noord B., Griebler M.G., U., and Klerings I. Inadequate use and regulation of interventions against publication bias decreases their effectiveness: A systematic review. *Journal of Clinical Epidemiology*, 68(7):792–802, 2015.

[66] Fitzmaurice G.M., Laird N.M., and Ware J.H. *Applied longitudinal analysis*. Hoboken, New Jersey, John Wiley & Sons, 2012.

[67] Kaslow R.A., Ostrow D.G., and Detels R. The multicenter aids cohort study: Rationale, organization, and selected characteristics of the participants. *American Journal of Epidemiology*, 126(2):310–318, 1987.

[68] Ishak K.J., Platt R.W., Joseph L., Hanley J.A., and Caro J.J. Meta-analysis of longitudinal studies. *Clinical Trials*, 4(5):525–539, 2007.

[69] Senn S., Stevens L., and Chaturvedi N. Repeated measures in clinical trials: Simple strategies for analysis using summary measures. *Statistics in Medicine*, 19(6):861–877, 2000.

[70] Hardy R.J. and Thompson S.G. A likelihood approach to meta-analysis with random effects. *Statistics in Medicine*, 15(6):619–629, 1996.

[71] Maas C.J.M., Hox J.J., and Lensvelt-Mulders G.J.L.M. Longitudinal meta-analysis. *Quality and Quantity*, 38(4):381–389, 2004.

[72] Musekiwa A., Manda S.O.M., Mwambi H.G., and Chen D.G. Meta-analysis of effect sizes reported at multiple time points using general linear mixed model. *Plos one*, 11(10):0164898, 2016.

[73] Berkey C.S., Antczak-Bouckoms A., Hoaglin D.C., Mosteller E., and Pihlstrom B.L. Multiple-outcomes meta-analysis of treatments for periodontal disease. *Journal of Dental Research*, 74(4):1030–1039, 1995.

[74] Dear K.B.G. Iterative generalized least squares for meta-analysis of survival data at multiple times. *Biometrics*, 50(4):989–1002, 1994.

[75] Olkin I. and Gleser L. Stochastically dependent effect sizes. *The Handbook of Research Synthesis and Meta-analysis*, 2(19):357–376, 2009.

[76] Riley R.D. Multivariate meta-analysis: The effect of ignoring within-study correlation. *Journal of the Royal Statistical Society*, 172(4):789–811, 2009.

[77] Trikalinos T.A. and Olkin I. Meta-analysis of effect sizes reported at multiple time points: A multivariate approach. *Clinical Trials*, 9(5):610–620, 2012.

[78] Kirkham J.J., Riley R.D., and Williamson P.R. A multivariate meta-analysis approach for reducing the impact of outcome reporting bias in systematic reviews. *Statistics in Medicine*, 31(20):2179–2195, 2012.

[79] West B.T., Welch K.B., and Galecki A.T. *Linear mixed models: A practical guide using statistical software*. New York, Chapman and Hall/CRC, 2014.

[80] Morrell C.H. Likelihood ratio testing of variance components in the linear mixed-effects model using restricted maximum likelihood. *Biometrics*, 54(4):1560–1568, 1998.

[81] Sheu C.F. and Suzuki S. Meta-analysis using linear mixed models. *Behavior Research Methods, Instruments and Computers*, 33(2):102–107, 2001.

[82] Aviles A.I. Linear mixed models for longitudinal data. *Technometrics*, 50(4):375–375, 2001.

[83] Brown H. and Prescott R. *Applied mixed models in medicine*. United Kingdom, John Wiley & Sons, 2014.

[84] Verbeke G. and Molenberghs G. *Linear mixed models for longitudinal data, linear mixed models in practice*. New York, Springer, 1997.

[85] Harville D.A. Bayesian inference for variance components using only error contrasts. *Biometrika*, 61(2):383–385, 1974.

[86] Akaike H. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.

[87] Schwarz G. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.

[88] Allison P.D. *Logistic Regression Using SAS: Theory and Application*. Cary, North Carolina, USA, SAS Publishing, 2012.

[89] Galecki A. and Burzykowski T. *Linear mixed-effects models using R: A step-by-step approach*. New York, Heidelberg, Dordrecht, London, Springer Science & Business Media, 2013.

[90] Kumari T., Vachher M., Bansal S., Bamezai R.N.K., and Kumar B. Meta-analysis of mitochondrial t16189c polymorphism for cancer and type 2 diabetes risk. *Clinica Chimica Acta*, 482:136–143, 2018.

[91] Viechtbauer W. and Cheung M.W.L. Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2):112–125, 2010.

[92] Yusuf S., Peto R., Lewis J., Collins R., and Sleight P. Beta blockade during and after myocardial infarction: An overview of the randomized trials. *Progress in Cardiovascular Diseases*, 27(5):335–371, 1985.

[93] Viechtbauer W. Conducting meta-analyses in r with the metafor package. *Journal of statistical software*, 36(3):1–48, 2010.

[94] Poole C. and Greenland S. Random-effects meta-analyses are not always conservative. *American Journal of Epidemiology*, 150(5):469–475, 1999.

[95] Berlin J.A., Laird N.M., Sacks H.S., and Chalmers T.C. A comparison of statistical methods for combining event rates from clinical trials. *Statistics in Medicine*, 8(2):141–151, 1989.

[96] Galbraith R.F. A note on graphical presentation of estimated odds ratios from several clinical trials. *Statistics in Medicine*, 7(8):889–894, 1988.

[97] Fine H.A., Dear K.B.G., Loeffler J.S., Black P.M., and Canellos G.P. Meta-analysis of radiation therapy with and without adjuvant chemotherapy for malignant gliomas in adults. *Cancer*, 71(8):2585–2597, 1993.

[98] Ahn J.E. and French J.L. Longitudinal aggregate data model-based meta-analysis with nonmem: approaches to handling within treatment arm correlation. *Journal of Pharmacokinetics and Pharmacodynamics*, 37(2):179–201, 2010.

Appendix A

A.1 Partition of variation

A useful result is the principle of partitioning of sums of squares that forms the basis for the analysis of variance. Given a set of constant $\{a_i\}$ and values $\{z_i\}$ such that $\sum_i a_i y_i = \sum_i a_i z_i$ and

$$\sum_i a_i (y_i - \bar{y})^2 = \sum_i a_i (y_i - z_i)^2 + \sum_i a_i (z_i - \bar{y})^2. \tag{5.1}$$

However it can be shown that $\sum_i a_i z_i (y_i - z_i) = 0$. The values $\{z_i\}$ may be random observations or constants. A similar result provides the partitioning of variation (mean square error) for an estimate $\hat{\theta}$ of a parameter θ. When the estimate has expectation $E(\hat{\theta}) \neq \theta$, then the mean square of the estimate can be partitioned as

$$E(\hat{\theta} - \theta)^2 = E[\hat{\theta} - E(\hat{\theta})]^2 + E[E(\hat{\theta}) - \theta]^2 = V(\hat{\theta}) + [\text{bias}(\hat{\theta})]^2. \tag{5.2}$$

A related expression is the well-known result

$$V(Y) = E_X[V(Y|x)] + V_X[E(Y|x)]. \tag{5.3}$$

Expressed in terms of the conditional moments of Y given the value of another variable X, integrated with respect to the distribution of X.

A.2 Slutsky's convergence theorem

Slutsky's theorem is a multifaceted result which can be used to establish the convergence in distribution and/or the convergence in probability(consistency) of multidimensional transformations of a vector of statistics. For the purpose here, we shall present these as two results rather than as a single theorem. The theorem is

then used in conjunction with the delta method to obtain asymptotic distribution of transformations of statistics.

A.3 Convergence in distribution

The most common application of slutsky's theorem concerns the asymptotic distribution of linear combination of two sequences of statistics, one that convergence in probability to a constant and another that convergence in distribution to a specified distribution. In this project we are only concerned with functions of statistics that are asymptotically normally distributed, for which the theorem is so described. The result, however, applies more generally to statistics that follow any specified distribution. The theorem also readily generalized to more than two such statistics. Let t_n be a sequence of statistics such that as $n \to \infty$,

$$\sqrt{n}(t_n - \mu) \to N(0, \sigma^2) \tag{5.4}$$

where the variance σ^2 may be a function of the expectation μ. Also, let r_n be a sequence of statistics that converges in probability to a constant ρ, expressed as $r_n \to \rho$. Then

$$\sqrt{n}[(r_n + t_n) - (\rho + \mu)] \to N(0, \sigma^2) \tag{5.5}$$

and

$$\sqrt{n}[(r_n t_n - \rho\mu)] \to N(0, \rho^2\sigma^2). \tag{5.6}$$

A.4 Delta method

A.4.1 Univariate case

Deriving large sample moment is a common problem in statistics. For example the expectation and variance, of a transformation of a statistic, including non-linear transformations. We can obtain these expression by using the $\delta-$delta method. Let t be any statistic for which the two central moments are know, $E(t) = \mu$ and $V(t) = \sigma^2$. We desire the moments of a transformation $y = g(t)$ for some function $g(\cdot)$ which is assumed to be twice differentiable, with derivatives designated as $g'(\cdot)$ and $g''(\cdot)$. A first order Taylor's series expansion of $g(t)$ about μ is

$$g(t) = g(\mu) + g'(\mu)(t - \mu) + R_2(a). \tag{5.7}$$

From the mean value theorem, the remainder for second order is

$$R_2(a) = (t-\mu)^2 g''(a)/2, \tag{5.8}$$

for some value a contained in the interval (t, μ). If the remainder vanishes under specified conditions, such as asymptotically, then

$$g(t) \cong g(\mu) + g'(\mu)(t-\mu) \tag{5.9}$$

so that

$$E(y) = \mu_y = E[g(t)] \cong g(\mu) + g'(\mu)E(t-\mu) = g(\mu) \tag{5.10}$$

and

$$\begin{aligned} V(y) &= E(y-\mu_y)^2 \cong E[g(t)-g(\mu)]^2 \\ &= E[g'(\mu)(t-\mu))]^2 = [g'(\mu)]^2 V(t). \end{aligned} \tag{5.11}$$

We frequently consider the moments of a transformation of a statistic t that is a consistent estimate of μ in such cases, since $t \to \mu$, then the remainder in equation (5.8) vanishes asymptotically, or $R_2(a) \to 0$, and the above results apply to any transformation of t. Furthermore, if $\hat{V}(t)$ is a consistent estimator of $V(t)$, than it follows from Slutsky's theorem convergence theorem, that $\hat{V}(y) = [g'(t)]^2 \hat{V}(t)$ is a consistent estimator of $V(y)$.

A.5 Log(p)

Consider the moments of the natural log of the simple proportion p for which $\mu = \pi$ and $\sigma^2 = \frac{\pi(1-\pi)}{N}$. The Taylor's expansion yields

$$log(p) = log(\pi) + \frac{dlog(\pi)}{d\pi}(p-\pi) + R_2(a), \tag{5.12}$$

where the remainder for some values $a \in (p, \pi)$ is $R_2(a) = (p-\pi)^2 g''(a)/2$. Since p is consistent for π, $p \to \pi$ then asymptotically $R_2(a) \to 0$ and thus

$$E[log(p)] \cong log(\pi) \tag{5.13}$$

and

$$V[log(p)] \cong [\frac{dlog(\pi)}{d\pi}]^2 V(p) = [\frac{1}{\pi^2}]\frac{\pi(1-\pi)}{N} = \frac{1-\pi}{\pi N}. \tag{5.14}$$

A.5.1 Multivariate case

Now consider a transformation of a p-vector $\boldsymbol{T} = (t_1, \ldots, t_p)'$ of statistic with mean vector μ and covariance matrix $\sum_{\boldsymbol{T}}$. Assume that $\boldsymbol{Y} = (y_1 \ldots y_m)' = \boldsymbol{G}(\boldsymbol{T}) = [g_t, \ldots, g_{\boldsymbol{T}}]'$, $m \leqslant p$, where the k^{th} transformation $g_k(\boldsymbol{T})$ is a twice differentiable function of $\boldsymbol{T}$. Applying a first order Taylor's series and assuming that the vector of remainders $\boldsymbol{R}_2(\boldsymbol{A})$ vanishes for values $\boldsymbol{A} \in (\boldsymbol{T}, \boldsymbol{\mu})$ yields

$$\begin{aligned} E(\boldsymbol{Y}) &= \boldsymbol{\mu_Y} \cong \boldsymbol{G}(\boldsymbol{T}) \\ V(\boldsymbol{Y}) &= \sum_{\boldsymbol{Y}} \cong \boldsymbol{H}(\boldsymbol{\mu})' \sum_{\boldsymbol{T}} \boldsymbol{H}(\boldsymbol{\mu}). \end{aligned} \tag{5.15}$$

Where $\boldsymbol{H}(\boldsymbol{\mu})$ is a $p \times m$ matrix with elements

$$\boldsymbol{H}(\boldsymbol{\mu}) = \begin{bmatrix} \frac{\partial g_1(\boldsymbol{T})}{\partial \boldsymbol{T}} \\ \vdots \\ \frac{\partial g_m(\boldsymbol{T})}{\partial \boldsymbol{T}} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1(\boldsymbol{T})}{\partial t_1} & \cdots & \frac{\partial g_1(\boldsymbol{T})}{\partial t_p} \\ & \vdots & \\ \frac{\partial g_m(\boldsymbol{T})}{\partial t_1} & \cdots & \frac{\partial g_m(\boldsymbol{T})}{\partial t_p} \end{bmatrix}$$

evaluated at $T = \mu$ and $\boldsymbol{T}$ is a jointly consistent estimator for $\boldsymbol{\mu}$. Then provides the first two moments of the asymptotic distribution of $\boldsymbol{Y}$. Further from slutsky's theorem if $\hat{\sum}_{\boldsymbol{T}}$ is consistent for $\sum_{\boldsymbol{T}}$, then

$$\hat{\sum_{\boldsymbol{Y}}} = \hat{H}(\boldsymbol{T})' \hat{\sum_{\boldsymbol{T}}} \hat{H}(\boldsymbol{T}) = \hat{H}' \hat{\sum_{\boldsymbol{T}}} \hat{H} \tag{5.16}$$

is a consistent estimator of $\sum_{\boldsymbol{Y}}$.

A.5.2 Multinomial generalized logits

Consider the case of a trinomial where we wish to estimate the mean and variance of the vector of log odds, logits of the second category versus the first, $log(\frac{p_2}{p_1})$ and also category versus the first, $log(\frac{p_3}{p_1})$. Thus $\boldsymbol{P} = (p_1 p_2 p_3)'$ has mean vector $\boldsymbol{\pi} = (\pi_1 \pi_2 \pi_3)'$ and covariance matrix

$$\sum(\boldsymbol{H}) = \frac{1}{N} = \begin{bmatrix} \pi_1(1-\pi_1) & -\pi_1\pi_2 & -\pi_1\pi_3 \\ -\pi_1\pi_2 & \pi_2(1-\pi_2) & -\pi_2\pi_3 \\ -\pi_1\pi_3 & -\pi_2\pi_3 & \pi_3(1-\pi_3) \end{bmatrix}.$$

The transformation is $\boldsymbol{Y} = \boldsymbol{G}(\boldsymbol{P}) = [g_1\boldsymbol{P}g_2\boldsymbol{P}]'$, where $g_1(\boldsymbol{P}) = log(\frac{p_2}{p_1})$ and $g_2(\boldsymbol{P}) = log(\frac{p_3}{p_1})$ and expected value of $\boldsymbol{Y}$ is given as

$$E[\boldsymbol{Y}] = \boldsymbol{\mu_Y} = [log(\frac{\pi_2}{\pi_1})log(\frac{\pi_3}{\pi_1})]'. \tag{5.17}$$

To obtain the asymptotic variance requires the matrix of derivatives, which are

$$\sum(\boldsymbol{H}) = \begin{bmatrix} \frac{\partial g_1(\boldsymbol{\pi})}{\partial \pi_1} & \frac{\partial g_1(\boldsymbol{\pi})}{\partial \pi_2} & \frac{\partial g_1(\boldsymbol{\pi})}{\partial \pi_3} \\ \frac{\partial g_2(\boldsymbol{\pi})}{\partial \pi_1} & \frac{\partial g_2(\boldsymbol{\pi})}{\partial \pi_2} & \frac{\partial g_2(\boldsymbol{\pi})}{\partial \pi_3} \end{bmatrix}' = \begin{bmatrix} -\frac{1}{\pi_1} & \frac{1}{\pi_2} & 0 \\ -\frac{1}{\pi_1} & 0 & \frac{1}{\pi_3} \end{bmatrix}'$$

hence,

$$\begin{aligned} \sum_Y &= \boldsymbol{H}' \sum_{\boldsymbol{T}} \boldsymbol{H} \\ &= \frac{1}{N} \begin{bmatrix} \frac{1}{\pi_1} + \frac{1}{\pi_2} & \frac{1}{\pi_1} \\ \frac{1}{\pi_1} & \frac{1}{\pi_1} + \frac{1}{\pi_3} \end{bmatrix}. \end{aligned} \tag{5.18}$$

Provides the asymptotic covariance matrix of the two logits.

Appendix B

B.1 R code for univariate examples

```
setwd("C:/Users/mhlengi/Desktop/Meta-R")

Load metafor package for meta-analysis
install.packages("metafor")
library("metafor")

Bring in data from Excel spreadsheet
data2<-read.csv("Ulcers.csv")

Print data on screen
data2

Effect size calculation
dat <- escalc(measure="RD", $ai=a_j, bi=n_1j-a_j,
 ci=b_j, di=n_2j-b_j$, data=data2)
summary(dat)
m1<-rma(measure="RD", $ai=a_j, bi=n_1j-a_j, ci=b_j,
 di=n_2j-b_j$, data=dat, method="FE")
summary(m1)

Residuals
res1 <- rma(yi, vi,data=dat)
inf1 <- influence(res1)
inf1
plot(inf1,plotdfb = TRUE)

m2<-rma(measure="RD", $ai=a_j, bi=n_1j-a_j,
 ci=b_j, di=n_2j-b_j$, data=dat, method="ML")
```

```
summary(m2)

Forest plot for risk difference
par(mfrow=c(2,2))
forest(m1, main="Fixed-effects")
forest(m2, main="Random-effects")

Funnel plot for risk difference
par(mfrow=c(2,2))
funnel(m1, main="Fixed-effects")
funnel(m2, main="Random-effects")


setwd("C:/Users/mhlengi/Desktop/Meta-R")

 Load metafor package for meta-analysis
install.packages("metafor")
library("metafor")

Bring in data from Excel spreadsheet
data2<-read.csv("Ulcers.csv")

Print data on screen
data2

Effect size calculation
dat <- escalc(measure="RR", $ai=a_j, bi=n_1j-a_j,
 ci=b_j, di=n_2j-b_j$, data=data2)
summary(dat)
m1<-rma(measure="RR", $ai=a_j, bi=n_1j-a_j,
ci=b_j, di=n_2j-b_j$, data=dat, method="FE")
summary(m1)

Residuals
res1 <- rma(yi, vi,data=dat)
inf1 <- influence(res1)
inf1
plot(inf1,plotdfb = TRUE)
```

```
m2<-rma(measure="RR", $ai=a_j, bi=n_1j-a_j,
 ci=b_j, di=n_2j-b_j$, data=dat, method="ML")
summary(m2)

Forest plot for log relative risk
par(mfrow=c(2,2))
forest(m1, main="Fixed-effects")
forest(m2, main="Random-effects")

 Funnel plot for log relative risk
par(mfrow=c(2,2))
funnel(m1, main="Fixed-effects")
funnel(m2, main="Random-effects")


setwd("C:/Users/mhlengi/Desktop/Meta-R")

Load metafor package for meta-analysis
install.packages("metafor")
library("metafor")

Bring in data from Excel spreadsheet
data2<-read.csv("Ulcers.csv")

Print data on screen
data2

Effect size calculation
dat <- escalc(measure="OR", $ai=a_j, bi=n_1j-a_j,
 ci=b_j, di=n_2j-b_j$, data=data2)
summary(dat)
m1<-rma(measure="OR", $ai=a_j, bi=n_1j-a_j,
ci=b_j, di=n_2j-b_j$, data=dat, method="FE")
summary(m1)

Residuals
res1 <- rma(yi, vi,data=dat)
inf1 <- influence(res1)
inf1
```

```
plot(inf1,plotdfb = TRUE)

m2<-rma(measure="OR", $ai=a_j, bi=n_1j-a_j,
 ci=b_j, di=n_2j-b_j$, data=dat, method="ML")
summary(m2)

Forest plot for log odds ratio
par(mfrow=c(2,2))
forest(m1, main="Fixed-effects")
forest(m2, main="Random-effects")

Funnel plot for log odds ratio
par(mfrow=c(2,2))
funnel(m1, main="Fixed-effects")
funnel(m2, main="Random-effects")
```

B.2 R code for meta-analysis of longitudinal studies

```
%setwd("C:/Users/mhlengi/Desktop/Meta-R")
%install.packages("metafor")
%library("metafor")
%data2<-read.csv("Book4.csv")
%data2
%dat<-escalc(measure="OR",$ai=a_j, bi=c_j, ci=b_j, di=d_j$, data=data2)
%dat
%
%Separate univariate random-effects model
%m6 = rma(yi, vi, data=dat[dat$time==6,],   method="REML")
%m12 = rma(yi, vi, data=dat[dat$time==12,], method="REML")
%m18 = rma(yi, vi, data=dat[dat$time==18,], method="REML")
%m24 = rma(yi, vi, data=dat[dat$time==24,], method="REML")
%
%The summary for log odds ratio
%summary(m6);summary(m12);summary(m18);summary(m24)
%
res1 <- rma(yi, vi,data=dat[dat$time==6,])
inf1 <- influence(res1)
inf1
plot(inf1,plotdfb = TRUE)
```

```
res1 <- rma(yi, vi,data=dat[dat$time==12,])
inf1 <- influence(res1)
inf1
plot(inf1,plotdfb = TRUE)

res1 <- rma(yi, vi,data=dat[dat$time==18,])
inf1 <- influence(res1)
inf1
plot(inf1,plotdfb = TRUE)

res1 <- rma(yi, vi,data=dat[dat$time==24,])
inf1 <- influence(res1)
inf1
plot(inf1,plotdfb = TRUE)

%Forest plot for log odds ratio
%par(mfrow=c(2,2))
%forest(m6, main="Month 6")
%forest(m12, main="Month 12")
%forest(m18, main="Month 18")
%forest(m24, main="Month 24")
%
%Funnel plot for log odds ratio
%par(mfrow=c(2,2))
%funnel(m6, main="Month 6")
%funnel(m12, main="Month 12")
%funnel(m18, main="Month 18")
%funnel(m24, main="Month 24")
%
%Multivariate Meta-Analysis for longitudinal data using the linear
mixed model. Independent random time effects model -MODEL1.
%mvmMODEL1 = rma.mv(yi, vi, mods= ~ as.factor(time)-1,
random = ~ as.factor(time) | as.factor(study),
 struct = "DIAG", data=dat);summary(mvmMODEL1, digits=2)
%
%Random study effects model -MODEL 2  "HCS" for compound symmetry,
%mvmMODEL2 = rma.mv(yi, vi, mods= ~ as.factor(time)-1,
 random = ~ as.factor(time) | as.factor(study),
```

```
  struct = "CS", data=dat);summary(mvmMODEL2, digits=2)
%
%
%HAR(1) correlated random time effects model -MODEL 3 "HAR"
for a heteroscedastic AR(1) autoregressive structure
%mvmMODEL3 = rma.mv(yi, vi, mods= ~ as.factor(time)-1,
 random = ~ as.factor(time) | as.factor(study),
 struct = "HAR", data=dat);summary(mvmMODEL3, digits=2)
%
%Note: For models 4 to 6 we created a variance-covariance
matrix for the within-study errors
%Load data in wide format with variables: study yi6 vi6
yi12 vi12 yi18 vi18 yi24 vi24
%dat=read.csv("data4.csv")
%dat
%
%Reshape to long
%dat.long <- reshape(dat, direction = "long", idvar = "study",
 v.names = c("yi","vi"), varying = list(c(2,4,6,8),c(3,5,7,9)))
%dat.long <- dat.long[order(dat.long$study, dat.long$time),]
%rownames(dat.long) <- 1:nrow(dat.long)
%
%Construct the full (block diagonal) V matrix with an AR(1) structure
%rho.within <- .59
%V <- lapply(split(with(dat, cbind(vi6,vi12,vi18,vi24)), dat$study), diag)
%V <- lapply(V, function(v) sqrt(v)
 %*% toeplitz(ARMAacf(ar=rho.within, lag.max = 3)) %*% sqrt(v))
%V <- bldiag(V)
%
%Independent random time effects and HAR(1)
 correlated within-study effects -MODEL 4
%resMODEL4 <- rma.mv(yi, V , mods= ~ as.factor(time)-1,
 random =~ as.factor(time) | study, struct = "DIAG",
  data=dat.long);summary(resMODEL4, digits=2)
%
%
%HAR(1) correlated random time effects and HAR(1)
correlated within-study effects -MODEL5
```

```
%resMODEL5 <- rma.mv(yi, V , mods= ~ as.factor(time)-1,
random = ~ as.factor(time) | study, struct = "HAR",
data=dat.long);summary(resMODEL5, digits=2)
%
%
%Unstructured random time effects and HAR(1)
 correlated within-study effect -MODEL6
%resMODEL6 <- rma.mv(yi, V , mods= ~ as.factor(time)-1,
random = ~ as.factor(time) | study, struct = "UN",
 data=dat.long);summary(resMODEL6, digits=2)
```