Meta-Analysis of summary statistics versus individual participant-level data of trials with binary outcomes

Layla Norman 16013353

WST895 Research Report

Submitted in partial fulfillment of the degree MSc Advanced Data Analytics

Supervisor: Prof D Chen

Department of Statistics, University of Pretoria

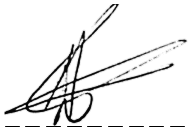
31 October 2022

Abstract

In this essay we will consider the meta-analysis of summary statistics versus individual participant-level data of trials with binary outcomes. Numerous techniques will be considered to essentially determine, both theoretically as well as numerically, by the use of real data analysis, whether or not, in certain instances, analysing individual participant-level data (IPD) from all investigations undoubtedly gains efficiency over combining summary statistics from each study in a binary setting. We will be considering both the one-step as well as two-step meta-analysis, with the focus being mainly on the standard fixed-effects model and partially on the random-effects model. An application of two different clinical studies will be considered. The first study is called the *Cannon (Cannon et al., (2006)[13]) Cardiovascular Clinical Trials* which consists of four clinical studies. The second study is called the *Bacillus Calmette-Guerin Vaccine Study* which consists of a total of 13 clinical studies. We observed that a meta-analysis of summary data is more similar to a meta-analysis of IPD data for the first study, but different for the second study. This is due to the fact that the number of studies are small in both studies, and therefore the asymptotic equivalence should be cautious when the sample size is small and the number of studies is small. In this case, the estimation of the between study variance τ^2 will be very unstable which will lead to a different conclusion between a meta-analysis of summary statistics (SS) and a meta-analysis of individual participant-level data (IPD).

Declaration

I, *Layla Norman,* declare that this essay, submitted in partial fulfillment of the degree *MSc Advanced Data Analytics,* at the University of Pretoria, is my own work and has not been previously submitted at this or any other tertiary institution.



Layla Norman

Prof D Chen

31 October 2022

Date

Acknowledgements

This work is based upon research supported by the National Research Foundation, South Africa (South Africa DST-NRF-SAMRC SARChI Research Chair in Biostatistics to Professor Ding-Geng Chen, Grant number 114613). Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the NRF.

Contents

1 Introduction **8**

1.1 Rationale and Literature review . . . 11

1.2 Research question . . . 14

2 Background Theory **16**

2.1 Research Protocol for Meta-Analyses . . . 16

2.1.1 Defining the Research Objective . . . 16

2.1.2 Finding relevant studies . . . 17

2.1.3 Data Abstraction and Extraction . . . 17

2.1.4 Meta-Analysis Methods . . . 17

2.2 Fixed-Effects and Random-Effects in Meta-Analysis . . . 18

2.2.1 Hypotheses and Effect Size . . . 18

2.2.2 Fixed-Effects Model: The Weighted-Average . . . 20

2.2.3 Random-Effects Model: DerSimonian-Laird . . . 22

2.3 What has been done . . . 25

2.3.1 Comparison of Meta-Analysis Versus Analysis of Variance of Individual Participant-level Data . . . 26

2.3.2 On the Equivalence of Meta-Analysis Using Literature and Using Individual Participant-level Data . . . 32

2.3.3 Comparison of One-Step and Two-Step Meta-Analysis Models Using Individual Participant-level Data . . . 37

3 Meta-Analysis with Binary Data **46**

3.1 Meta-Analysis Methods . . . 47

3.1.1 Analysis with Risk-Ratio . . . 47

3.1.2 Analysis with Risk-Difference . . . 49

3.1.3 Meta-Analysis with Odds-Ratio . . . 50

3.2 Summary Statistics Data Methods . . . 53

3.2.1 Conventional Methods . . . 53

3.2.2 Multilevel Modelling Methods . . . 54

3.2.3 Calculation of Confidence Intervals . . . 54

3.3 Individual Partcipant-level Data Methods . . . 56

3.3.1 Model Description . . . 56

3.3.2 Parameter Estimation . . . 57
3.3.3 Calculation of Confidence Intervals . . . 58
3.4 Individual Partcipant-level Data Methods with Random Trial Effects . . . 58
3.4.1 Model Description . . . 58
3.4.2 Bivariate Approach . . . 59
3.4.3 Coding of Treatment Covariates . . . 59
3.4.4 Model Checking . . . 59

4 Application **60**
4.1 Cannon Cardiovascular Clinical Trials . . . 60
4.1.1 Data Description . . . 60
4.1.2 Meta-analysis of Summary Data . . . 62
4.1.3 Meta-analysis of Individual Participant-Level Data . . . 76
4.2 Bacillus Calmette-Guerin Vaccine Study . . . 79
4.2.1 Data Description . . . 79
4.2.2 Meta-analysis of Summary Data . . . 80
4.2.3 Meta-analysis of Individual Participant-Level Data . . . 84

5 Conclusion **88**

6 Appendix **90**
6.1 Gauss-Markov theorem . . . 90
6.2 The Delta Method . . . 90
6.3 R code: Meta-analysis of summary data - Cannon study of cardiovascular clinical trials . . . 91
6.4 R code: Meta-analysis of Individual participant-level data - Cannon study of cardiovascular clinical trials . . . 104
6.5 R code: Meta-analysis of summary data and Individual participant-level data - Bacillus Calmette-Guerin Vaccine Study . . . 108

7 References **112**

List of Figures

1 Hierarchy of evidence. . . . 10
2 Odds-ratios and corresponding sampling variances for the studies. . . . 63
3 Forest plot of the odds-ratios for the studies . . . 63
4 Funnel plot of the odds-ratios for the studies . . . 64

5 Risk ratio and corresponding sampling variances for the studies. 69
6 Forest plot of the risk ratios for the studies 69
7 Funnel plot of the risk ratios for the studies 70
8 Risk-difference and corresponding sampling variances for the studies. 74
9 Forest plot of the risk-differences for the studies 75
10 Funnel plot of the risk-differences for the studies 75
11 First 10 observations of the IPD data 77
12 Odds-ratios and corresponding sampling variances for the studies for the BCG vaccine study. 81
13 First 10 observations of the IPD data for the BCG vaccine study 86

List of Tables

1 Nomenclature for 2×2 Table of Outcome by Treatment. 51
2 Data for Coronary Death or MI of Statin Use. 61
3 A 2×2 representation of the data for Coronary Death or MI of Statin use for all studies. 62
4 Data from Studies on Efficacy of BCG Vaccine for Preventing Tuberculosis. 80
5 A 2×2 representation of the data for the BCG vaccine for preventing tuberculosis for all studies. 81

1 Introduction

In this dissertation we will consider meta-analysis of summary statistics versus individual participant-level data, particularly for trials with binary outcomes. Over the last few decades, meta-analysis has become increasingly more popular and has played a crucial function in health care, medical research and clinical sciences, and its applications have prompted multitudinous of scientific discoveries. The primary statistical goal in a meta-analysis can be summarised as the evaluation of, or forming conclusions about, a similar/common impact, or numerous common effects, based on data from independent experiments. (Mathew, (2010)[5]). This efficiency of a so called combined estimator or confidence intervals, is one of the key statistical concerns in meta-analysis. Despite the fact that there is a vast array of meta-analysis literature, there are still many important questions that haven't been resolved, one of which is: *whether or not analyzing individual participant data (IPD) from all research studies is more effective than integrating summary statistics from each one?* This specific issue was recently explored by Chen et al. (Chen et al., (2020)[1]) where it was mathematically and conceptually shown that summary-statistics-based analysis is in all things considered as efficient as IPD analysis, with the focus being on the standard random-effects model for various types of responses, on condition that the likelihood function is used (Chen et al., (2020)[1]). It should also be noted that there are relatively few studies which address the fundamental common estimation as well as inference problems from an extensive statistical point of view.

In the big data era which we currently live in, it has become the norm that the information gathered to address the equivalent/comparable scientific questions originate from diverse sources (Chen et al., (2020)[1]). Conclusions about the efficiency of studies or the utility of a theory cannot be based on the conclusions of individual research alone. This is because conclusions typically differ from one research to the next. Meta-analysis is a tool which can be used to address this problem as it integrates the outcomes of multiple studies by applying statistical analysis to data. Meta-analysis additionally has the potential to compare outcomes from numerous studies and recognize models within their results, sources of controversy among the outcomes, or other interesting associations that may arise in the results of various studies.

Data sharing has become increasingly encouraged, and as a result thereof authentic data at the individual level might be available on specific database platforms. It is commonly believed that when individual data are available, the IPD technique, which entails constructing a general model for every individual and inferring from the likelihood utilizing the maximum likelihood technique, could potentially reduce bias and increase efficiency in inference. Regardless of several advantages of analysing IPD, for a variety of reasons, including: (1) The surprisingly time-consuming possibility of the recovery of original

data, summary-statistics-based meta-analysis is nevertheless widely used in practice. It is still debatable whether the advantages of employing the IPD approach can compensate the exorbitant cost of acquiring original data. Due to privacy issues and regulatory considerations like the POPI (Protection of Personal Information) act (Buys, (2017)[16]), the IPD may occasionally be unavailable. Research foundations are discouraged from making individual-level data accessible to the public due to the absence of data sharing norms in several sectors; (2)The ability to analyze all of the data at once is not always implied by access to IPD. When the amount of data becomes too great to ever be broken down on a single machine, as frequently observed when dealing with big data and in software engineering or machine learning issues, the most practical approach is to conduct statistical learning from each study and then integrate the results (for example, summary statistics) towards the end(Chen et al., (2020)[1]). In order to identify the estimators of a linear function of the mean, the statistical task at hand is to compare a linear model for individual participant-level data (IPD) to a linear model for summary (aggregate) data. These are frequently referred to as the one-stage IPD meta-analysis and the two-stage IPD meta-analysis, respectively. As a result, the names of the corresponding estimators, one-step IPD meta-analysis estimator and two-step IPD meta-analysis estimator, are obvious.

The quantitative analysis of meta-data from a systematic review is known as meta-analysis. A systematic review aims to compile empirical evidence that fits prespecified eligibility conditions to respond to a specific research question of interest. The principle qualities of a systematic review are an unmistakably expressed arrangement of objectives with predefined eligibility study standards; a methodology that is reproducible as well as accurate; an efficient search that is able to recognise all the various studies that comply with the requirements; an evaluation that determines whether or not the findings of the included studies are indeed legitimate; along with a systematic explanation and synthesis of the traits and findings from the research used (Haidich, (2010)[2]). The exploration and execution of meta-analyses based on IPD and that of summary statistics has increased in a multitude of contexts, where many of these studies contain binary outcomes. This specific statistical approach was hardly mentioned in enough detail in a prior comprehensive analysis of multiple credible studies published between the period of 1999 - 2001. It was also noticed here that of these studies, 32% of them contained a binary outcome (Thomas et al., (2014)[7]). Numerous papers in the medical literature seem to suggest that there is still a great deal of ambiguity and doubt regarding the validity of a two-step IPD meta-analysis, and even more so in the case of a binary scenario, served as the inspiration for this study. It appears that literature has not completely employed a comprehensive statistical framework to directly evaluate the one-step and two-step IPD meta-analysis estimators for binary data. This elaborates the importance for further research in meta-analysis, specifically in a binary setting.

In medical research fields, evidence-based medicine is defined as the systematic, quantitative, preferentially experimental approach to deal with getting and utilizing clinical data. Consequently, evidence-based medicine is the field where meta-analysis assumes a focal function. In Figure 1 below which illustrates the hierarchy of evidence, it can be seen that meta-analyses take the top position. In the hierarchy, The strength of the freedom that results from numerous biases that impacted the medical study determines the position of the clinical evidence. Conversely, the base of the hierarchy, which have minimal clinical value as verification, is occupied by research about animals, case reports as well as lab studies. (Haidich, (2010)[2])

Figure 1: Hierarchy of evidence.

It should be noted that we purposefully excluded several potential causes of worry, such as publication bias or selection bias, in order to compare the estimators. Despite the fact that these problems would inevitably exist in any real meta-analytic modeling of trials, incorporating them into our comparison would only make things more difficult. However, under this simplifying assumption, by building on past work by Olkin and Sampson (Olkin and Sampson, (1998)[6]) and by Mathew and Nordström (Mathew and Nordström, (1999)[4]), for a big class of linear models, we should be able to derive a complete solution to this problem. The outcome can also be used as a springboard for a more thorough comparison that takes into account factors like publishing or selection bias.

Assume that the goal of the statistical analysis is, from some linear function of the mean vector, to estimate and infer.

Assume moreover that each trial has accessible summary estimators and accompanying covariance matrices, such as maximum likelihood estimators (MLE) or the best (in terms of least variance) linear unbiased estimators (BLUE) under normality assumptions. An estimator based on the two-step IPD meta-analysis may then be created by optimally combining the summary estimators. On the other hand, by providing a linear model for the complete vector of answers, a BLUE (or MLE) may also be generated by merging the data from the many trials. That is, we have a one-step IPD meta-analysis estimator. It

makes sense to inquire as to whether these two estimators generally agree for trials with binary outcomes, and if not, then under what circumstance(s) might one anticipate this to occur.

1.1 Rationale and Literature review

Meta-analysis is a quantitative, formal, epidemiological examination plan which is utilized to systematically evaluate past research studies so as to determine conclusions about that specific research. In comparison to any individual study which contributes to a pooled analysis, the outcomes of a meta-analysis might include far more precise estimations of the treatment effects, or numerous risk factors for disease and numerous other results. One of the most critical outcomes is the assessment of heterogeneity within the outcomes of the studies. Some of the benefits and advantages of meta-analysis include a combined as well as quantitative review of a huge, and frequently complex, usually conflicting, body of writing. The determination of the result and hypotheses that are being tested is critical in conducting a meta-analyses. Though most current research are difficult to identify from one another, which might lead to erroneous results, there exists numerous techniques that have the ability of inspecting data to be able to recognise the potential for studies to be absent; for instance, funnel plot utilization. Specifically in evidence-based medicine, thoroughly conducted meta-analysis are very helpful devices. What guarantees that meta-analytic research is attractive is the need to incorporate findings from numerous studies, and the wide range of research that is now being produced makes it quite possible to do this investigation. (Haidich, (2010)[2])

Olkin and Sampson (Olkin and Sampson, (1998)[6]) clarified that meta-analysis is a technique for combining the findings of several independent investigations. Olkin and Sampson (Olkin and Sampson, (1998)[6]) considered the case wherein there are various treatments in addition to a control, with the goal of evaluating the relative impact of each treatment dependent on continuous results. It has gotten more normal to utilize meta-analytic estimators of treatment contrasts when all the information is accessible, as opposed to just the summary data. An alternative methodology could be to utilize a two-way analysis of variance (ANOVA) model without interactions, where study and treatment are the only factors. Olkin and Sampson (Olkin and Sampson, (1998)[6]) got the astounding outcome, for the unbalanced case, that the estimates of treatment contrasts for the standard meta-analysis approach is virtually indistinguishable to that of the linear model's least squares technique. Regarding data recovery, it could be significantly more expensive to use a meta-analysis of individual participant-level data (IPD), hereon referred to as MA-IPD, as opposed to using a meta-analysis of summary data, hereon referred to as MA-SS. This equivalence accordingly offers a substantially more cost-effective analysis approach.

Mathew and Nordström (Mathew and Nordström, (1999)[4]) reviewed the suspicion made by Olkin and Sampson (Olkin and Sampson, (1998)[6]) that the observations are free with a common variation both within and between studies. Olkin and Sampson (Olkin and Sampson, (1998)[6]) acquired the outcome that the two meta-analytic methods both produce a similar estimator i.e. they are the same. Mathew and Nordström (Mathew and Nordström, (1999)[4]) broadened this examination done by Olkin and Sampson (Olkin and Sampson, (1998)[6]) where a comparable equivalence was established explicitly for the two-way fixed-effects model without interaction. The main assumption for this study was that across the studies, the observations are independent. Regardless of the covariance structure, in a two-way fixed-effects model without interaction, the best linear unbiased estimator (BLUE) may now be explicitly represented for any vector of treatment contrasts. This is only possible on condition that there is independence between the studies. An additional outcome is that, when the studies are independent of each other, paying little mind to covariance structure, for the optimum linear unbiased estimate, a two-way fixed-effects model with unequal variance and no interaction can really be viewed as a collection of separate unbalanced one-way models.

Turner et al. (Turner et al., (2000)[8]) investigated for both summary statistics (SS) data as well as for individual participant-level data (IPD) the capability of multilevel meta-analysis models specifically with binary results. By being included in the multilevel model system, the fixed effects models and the random effects models often offer either maximum likelihood (ML) or restricted maximum likelihood (REML) estimation. To exhibit the different strategies, Turner et al. (Turner et al., (2000)[8]) utilized the outcomes from 22 trials which investigated measures to prevent respiratory tract infections and furthermore compared these findings to a second sample data set with fewer trials. From a certain parametric bootstrap technique or from likelihood-based strategies, we are able to establish confidence intervals for the overall treatment effects as well as the between-trial variance. These are all done within the SS data techniques. It should also be noted that the intervals obtained from the bootstrap approach are more favourable since the assumptions which are needed by the other techniques they relaxed. To give both impartial estimation and correctly located confidence intervals when modelling IPD, one approach that can be considered it to utilize a bias corrected bootstrap. This particular strategy is especially significant for the between-trials variances. Within the IPD models, either random or fixed models might be used to describe the trial effects. An important finding that should be mentioned is that if the random trial effects are utilized, what should be incorporated are both the covariances among these as well as the random treatment effects. This essentially results in a model that is identical to a bivariate meta-analysis approach. Having said this, multilevel modeling's adaptability could be abused in encouraging augmentations to conventional meta-analysis methods.

Mathew and Nordström (Mathew and Nordström, (2010)[5]) stated that the main concern of combining information from different trials is a crucial consideration while setting up a multi-center experiment or doing a meta-analysis. Mathew and Nordström (Mathew and Nordström, (2010)[5]) explained that despite the fact that there seems to be a lot of literature on meta-analysis based on individual participant-level data (IPD), especially a one-step MA-IPD as opposed to a two-step MA-IPD that relies on summary data, ongoing clinical literature articles have shown that the legitimacy of a study that relies on aggregate data, still has a significant amount of ambiguity and confusion. In this article named "Comparison of One-Step and Two-Step Meta-Analysis Models Using Individual Patient Data", Mathew and Nordström (Mathew and Nordström, (2010)[5]) addressed a focal statistical issue by thinking about the linear function of the mean's estimation. Both the IPD and the SS data used in this estimation are based on linear models. The covariance matrix and the best linear unbiased estimator (BLUE) or maximum likelihood (ML) estimate of the specific parameter were anticipated to be present in the SS data from a trial. The arrangement utilized was moderately broad as it considered the existence of covariates and the presence of random effects within the model. It also featured a large number of other models, many of which are regularly used, especially linear models with fixed treatment effects or linear models with fixed or random trial effects. Mathew and Nordström (Mathew and Nordström, (2010)[5]) proceeded to infer a condition whereby the MA-IPD estimators for one-step and two-step coincide, which was represented theoretically and in terms of an application. The functions of balance and heterogeneity was additionally featured. The analysis also showed that when covariates are available, under a specific condition the two estimators coincide. This condition represents homogeneity of the covariate across trials. In practice, it is highly unlikely that this homogeneity condition will hold, and therefore the two estimators will not coincide.

Lin and Zeng (Lin and Zeng, (2010)[3]) communicated that ongoing advances in innovations just as communications have made it progressively attainable to get to the raw data on individual participants, despite the fact that meta-analysis is generally done by joining the SS of relevant studies. In the paper named "On the relative efficiency of using summary statistics versus individual-level data in meta-analysis", Lin and Zeng (Lin and Zeng, (2010)[3]) researched the relative efficiency of combining SS versus examining original data. Considering the three requirements that the parameter of interest has a consistent value across all experiments, the SS depends on maximum likelihood, and the nuisance parameters have diverse values across investigations, they discovered that there is no asymptotic efficiency gain when analyzing the original data. This is true for all widely applied parametric and semiparametric models. Lin and Zeng (Lin and Zeng, (2010)[3]) likewise surveyed the relative efficiency of two strategies, namely, when either all of the studies share similar nuisance parameters or when the parameter of interest

has different values amongst the trials. To affirm the theoretical outcomes and to further give a genetic association study's empirical comparisons, simulations based on these studies and strategies were conducted.

Thomas et al. (Thomas et al., (2014)[7]) clarified that MA-IPD are getting progressively well known over MA-SS. These studies, which frequently have a binary conclusion, are being accepted in an expanding variety of situations. Thomas et al. (Thomas et al., (2014)[7]) investigated various statistical strategies utilized for MA-IPD of binary results and furthermore explored how different studies' heterogeneity is described and also how research investigated the impact of covariates. Thomas et al. (Thomas et al., (2014)[7]) concluded that evidence from this systematic review showed that there has been an increase in the usage of binary outcomes, with random effects logistic regression being the most used method for analysis, in assessing the impacts of medical care concerns. Techniques are still anyway regularly not announcedin sufficient depth. The results also showed that a variety of applications assess the heterogeneity of treatment effects.

Chen et al. (Chen et al., (2020)[1]) stated that a meta-analysis might be carried out by either of the two approached: (1) making use of study-level SS that are synthesized or alternatively (2) if the IPD are available, drawing inference from a thorough model. The latter approach is seen as the "gold standard". It is not fully clear if using IPD for random-effects models actually results in greater efficiency than doing so using SS. Chen et al. (Chen et al., (2020)[1]) considered the relative efficiency of the two methods by making use of a general likelihood inference setting where they showed, both conceptually and quantitatively, that summary-statistics-based analysis is in all respects regarded as being as competent as IPD analysis. This is of course only on condition that when obtaining the SS, maximum likelihood estimation is used, and furthermore that a Gaussian distribution is followed by the random effects. More specifically, Chen et al. (Chen et al., (2020)[1]) showed that (1) the two methods are in fact asymptotically equivalent; and (2) if the sample sizes are too small, summary-statistics-based inference may suffer an evident loss in efficiency. The outcomes are set up with the assumption that the parameter of between-study heterogeneity remains stable, paying little heed to the sample sizes.

1.2 Research question

This dissertation's objective is to present in-depth study on the subject of meta-analysis in a binary environment where an investigation will be done to determine whether, in some cases, compared to integrating SS from each trial, an analysis of IPD from all studies is more effective. Therefore, the aim is to examine the research done by Chen et al. (Chen et al., (2020)[1]) as well as by Mathew and Nordström (Mathew and Nordström, (2010)[5]) in a binary environment. Chen et al. (Chen et al., (2020)[1]) mathematically

as well as conceptually demonstrated that summary-statistics-based analysis is in all things considered as efficient as IPD analysis, with the emphasis being on the standard random-effects model for different sorts of responses, on condition that the likelihood function is utilized. Mathew and Nordström (Mathew and Nordström, (2010)[5]) inferred a condition whereby the one-step as well as the two-step MA-IPD estimators harmonize, which was illustrated both theoretically and in terms of an application. We shall examine both the one-step and two-step meta-analysis in this dissertation, with the focus being on the standard fixed-effects model.

The objectives are to theoretically as well as numerically, by the use of real data analysis, determine whether or not, in certain cases, MA-SS and MA-IPD produce asymptotically equivalent results in a binary setting. This potentially involves: (1) summarizing and integrating outcomes across various individual studies, (2) analysing differences in the outcomes among numerous studies, (3) overcoming the problem of sample sizes that are too small in individual studies to be able to recognize effects of interest, and analyse end points that specifically require sample sizes that are large, (4) in estimating effects, the objective is to increasing precision, (5) evaluating the impact on certain patient groups, (6) determining whether further research is needed to explore a problem and, (7) creating unique hypotheses for future studies. This research will prove to be invaluable in multiple sectors, specifically in the medical research fields where a large percentage of studies indeed have a binary outcome.

All coding for the applications will be done using RStudio software, Version 3.6.3 (2020-02-29), with the main packages used: *meta* (Guido Schwarzer (2020)), *rmeta* (Thomas Lumley (2018)) and *metafor* (Wolfgang Viechtbauer (2020)). When compared to the package *meta*, which can be used for both continuous and binary data, the package *rmeta* is primarily used for binary data.

2 Background Theory

In this section we will briefly discuss the research protocols involved in doing a meta-analysis as well as fixed-effects and random-effects in meta-analysis. This enables the establishment of some notation and expressions required for the meta-analysis with binary data.

2.1 Research Protocol for Meta-Analyses

In the field of meta-analysis, knowing which questions to ask can be a challenging task. One approach to try and alleviate this concern is to work "backwards". What we mean by this is that instead of first formulating a specific research question of interest, trying to find appropriate data to support the research question, then doing a meta-analysis on the data and finally reporting on these findings, it is usually simpler to first make use of SS (with the exception of having complete access to IPD) for a specific set of studies to be combined which have already been meta-analyzed and reported on. From the obtained statistical details and results of the meta-analyses, we are now able to identify the research questions to which the meta-analyses appearing in the literature were potentially directed. (Chen and Peace, (2013)[17])

If we were to have a specific research question that a meta-analysis could potentailly answer, in practise the first and possibly most important step is to construct a protocol. Under this protocol, the entirety of the scientific research effort for the meta-analysis question would be conducted.

2.1.1 Defining the Research Objective

The first step when developing a protocol for a clinical trail of a drug is to clearly formulate the research question that the trial hopes to answer. This objective of a well defined research question when developing a meta-analysis protocol is of no less importance than that of a well defined research quesion for a protocol of a new drug (Chen and Peace, (2013)[17]). Some important things to consider when formulting a research question for a new drug or clinical trial are:

- The drug class.
- Specific measurements e.g what is meant by mild-to-moderate.
- The effectiveness measure or outcome, and how this was assessed.
- The type of control group.
- Study characteristics such as (a) the type of design (parallel, crossover) used, (b) the method of

assignment (e.g. balanced random assignment) of patients to drug and control groups, (c) specific measures to guard against bias (e.g. double-blinded).

- Type of patient.
- Length of study.

These seven points clearly define the main parameters which govern the search to include studies in the meta-analysis to respond to the particular research question or objective. (Chen and Peace, (2013)[17])

2.1.2 Finding relevant studies

There are numerous well known databases that may be utilized when searching for a particular meta-analysis namley: PubMed (MEDLINE), Embase, Web of Science (Science Citation Index), ClinicalTrials.gov and the Cochrane Central Register of Controlled Trials.

The requirements for selecting studies for the meta-analysis are equivalent to the Inclusion/ Exclusion Criteria of a clinical trial protocol. Inclusion implies that the researcher decides which of the studies to keep for the meta-analysis. Exclusion implies that the researcher decides which studies should not be kept. (Chen and Peace, (2013)[17])

2.1.3 Data Abstraction and Extraction

A data extraction form (DAF) should be designed and implemented by the researcher, ensuring that it is clear and straightforward. This DAF will be utilized specifically to extract and record all of the data that is collected from the studies that will be combined in an effort to address the meta-analysis objective. (Chen and Peace, (2013)[17])

Combining all of the individual study findings happens much easier from the DAF than working directly from the publication to analysis file creation. The DAF containing the extracted data generates a documentation record which may be used in a quality assurance process. It can also be used by other researchers who wish to do further research on the topic at hand. What to include on the DAF is guided mainly by the specific criteria for identifying studies that will be analyzed as part of the meta-analysis, as stated in section 2.1.1. (Chen and Peace, (2013)[17])

2.1.4 Meta-Analysis Methods

Within the research protocol, the meta-analysis section will be guided by the specified objective, as well as the type of data to be combined, the appropriate statistical methods and finally the research plan. It

is essential to note how heterogeneity of summary measures across studies will be delt with. (Chen and Peace, (2013)[17])

A variety of decisions are made during the meta-analysis of a group of clinical studies. Whether the meta-analysis has only SS or complete IPD, each stage's choices are largely the same. (Chen and Peace, (2013)[17])

Fixed-effects and random-effects models are the initial alternatives, and for each of these two models, a variety of estimate methods must be chosen. If a random-effects model is fitted, more decisions surface, namely (Turner et al., (2000)[8]):

- how can the estimation of the between-trial variance be adjusted for uncertainty when constructing a confidence interval for the treatment effect;
- how, for the between-trial variance, can confidence intervals be obtained;
- how the model will integrate trial-level covariates;
- and how between-trial heterogeneity will be investigated as mentioned earlier.

2.2 Fixed-Effects and Random-Effects in Meta-Analysis

The typical fixed-effects and random-effects models for meta-analysis are now discussed. Fixed-effects is also known as the weighted mean method whereas the DerSimonian-Laird random-effects model describes random effects. (Chen and Peace, (2013)[17])

As we have previously mentioned, the practice of combining the findings from several studies that address a group of related research hypotheses is known as meta-analysis. Usually, to achieve this, a common measure of effect size is found and then modeled using meta-regression. It should be emphasized that when research characteristics are taken into account, the resultant general averages may be regarded as meta-effect sizes. Compared to estimates of the true effect size that come from a single research with a single set of variables and assumptions, these meta-effect sizes are more accurate. (Chen and Peace, (2013)[17])

2.2.1 Hypotheses and Effect Size

When performing a clinical research to determine if a new drug (D) is successful in treating a certain condition, the overarching goal is to show that the new medicine is, in fact, effective in treating the disease. We can rewrite this in the form of a hypothesis, as follows (Chen and Peace, (2013)[17]):

$$H_0 : \text{Effect of D is no different from that of control (placebo} = \text{P)}$$

$$H_a : \text{Effect of D is better than that of P}$$

Within each treatment group, the treatment effet size is the main comparative function of the response effectiveness. If the response is binary, this comparison function may be the difference in proportions; if the response is continuous, it may be the difference in means. These are not, however, the only comparative functions as another comparative functions of effect size, specifically for binary data, are the log-odds ratio or relative risk (Chen and Peace, (2013)[17]). Let us denote the treatment effect size by δ. The hypothesis can then be represented as:

$$H_0 : \delta = 0$$

$$H_a : \delta > 0$$

H_0 and H_a are equal for each study when we have numerous randomized, controlled, effective studies of a certain drug. To ensure valid and unbiased estimates of treatment effect within studies, randomly allocating patients to treatment groups within studies is crucial as well as conduct the study in a quality and blinded manner.(Chen and Peace, (2013)[17])

The treatment effect estimates across the centers that are analyzed in a meta-analysis is essentially no different to a design-based analysis strategy. Specifically (Chen and Peace, (2013)[17]):

- Firstly, we calculate the estimates for the treatment effects $\hat{\delta}_i$ as well as the within variance $\hat{\sigma}_i^2$ for the treatment effects at each study, $i = 1, 2, \ldots, K$.
- Then, meta-analyze $\hat{\delta}_i$ over all of the studies.

The estimations from each study are meta-analyzed to be able to achieve an estimate of the general effectiveness of the drug over all of the studies as well as to provide a conclusion about the statistical significance of the overall effect (Chen and Peace, (2013)[17]). There are two main appraches in meta-analysis:

- Fixed-effects, in which we assume that we have an estimated treatment effect $\hat{\delta}_i$ for each clinical study i and its within variance $\hat{\sigma}_i^2$. Each of the treatment effects $\hat{\delta}_i$ is an estimate of the general global effect of δ over all of the studies. We need to combine the $\hat{\delta}_i$ using some weighting scheme in order to meta-analyze them.
- Random-effects, where it is presumable that each of the $\hat{\delta}_i$ is essentially an estimate of its own true effect δ_i. These estimates are each a single realization from the global effect δ. As a result,

the random-effects meta-analysis model is able to incorporate both within-variability as well as between-variability.

2.2.2 Fixed-Effects Model: The Weighted-Average

The fundamental assumption of the fixed-effects model is that, in a meta-analysis, all studies share the same true effect size δ and are equally influenced by all other risk variables. Using this assumption, the true effect size is equivalent for all of the studies. All of the observed effect sizes $\hat{\delta}_i$ for the fixed-effects model may vary among the studies as a result of the random errors from each study. Additionally, it is also assumed to be an estimate of the underlying global overall effect δ.(Chen and Peace, (2013)[17])

Using the fixed-effects model, we assume that all of the factors that are able to impact the effect size are in fact the same for all of the studies, and that

$$\hat{\delta}_i = \delta + \epsilon_i \tag{1}$$

where it is assumed that ϵ_i is normally distubuted with $N(0, \hat{\sigma}_i^2)$. Therefore

$$\hat{\delta}_i \sim N(\delta, \hat{\sigma}_i^2) \tag{2}$$

The underlying global overall effect δ is estimated using a weighting scheme to combine each individual estimates. In other words, $\hat{\delta}_i$ is weighted for each study i by an appropriate weight w_i. Next, we determine the weighted mean, also known as the pooled estimate $\hat{\delta}$ of treatment effect as well as its variance $\hat{\sigma}^2$, where (Chen and Peace, (2013)[17])

$$\hat{\delta} = \sum_{i=1}^{K} w_i \hat{\delta}_i \tag{3}$$

$$\hat{\sigma}^2 = \text{Var}(\hat{\delta}) = \sum_{i=1}^{K} w_i^2 \hat{\sigma}_i^2 (\text{under independence of the K studies}) \tag{4}$$

Now, utilizing the weighted mean from equation (3) as well as its variance in equation (4), we are able to construct an appropriate 95% confidence interval (CI) for δ:

$$\hat{\delta} \pm 1.96 \times \sqrt{\hat{\sigma}^2} \tag{5}$$

Furthermore, we may also formulate a t -type of test as:

$$T = \frac{\hat{\delta} - \delta}{\sqrt{\hat{\sigma}^2}} \tag{6}$$

which may be used as a method of testing the hypothesis $H_0 : \delta = 0$. The test statistic given in equation (6) may be applied to build confidence intervals on the overall global effect of δ. (Chen and Peace, (2013)[17])

2.2.2.1 The Weighting Schemes

In equation (3), the weighted mean requires the following condition to be valid: $\sum_{i=1}^{K} w_i = 1$. The following are some of the most common choices of w_i (Chen and Peace, (2013)[17]):

1. Weighting based on the quantity of studies:

$$w_i = \frac{1}{K} \tag{7}$$

where K is the predetermined number of studies utilized. The unweighted mean, commonly referred to as the arithmetic average, of the estimates of treatment effect across all trials is produced by this weighting system.

2. Weighting based on the quantity of participants in each study:

$$w_i = \frac{N_i}{N} \tag{8}$$

where Ni represents the overall patient count for each study ($i = 1, 2, \ldots, K$), and N represents the total patient count ($N = \sum_{i=1}^{K} N_i$). This weighting scheme yields the average of the estimates of treatment effect across all studies, which are weighted based on the quantity of participants in each study. It should be noted that equation (8) reduces to equation (7) if all of the studies are balanced.

3. Weighting based on the number of participants from each study as well as each treatment:

$$w_i = \frac{N_{iD} N_{iP}}{N_{iD} + N_{iP}} \times \frac{1}{w} \tag{9}$$

where N_{iD} represent the number of patients in the new drug treatment group (D), N_{iP} represent the number of patients in the Placebo (P) group and $w = \sum_{i=1}^{K} w_i$, in each study i. This weighting scheme also yields the average of the treatment effect estimates from all studies however, they are weighted to take into account treatment group imbalance in each study. Furthermore, equation (9) may be reduces to equation (8) on condition that the treatment groups are evenly distributed among studies.

4. Weighting according to inverse variance:

$$w_i = \frac{1}{\hat{\sigma}_i^2} \times \frac{1}{w} \tag{10}$$

where $w = \sum_{i=1}^{K} w_i$. This weighting system averages treatment effect estimates from many studies, with estimates weighted inversely to their variance. Almost all fixed-effects models employ this specific method. It is important to note that equation (10) reduces equation (7) if the $\hat{\sigma}_i^2$ are equivalent (true homogeneity) across all studies.

For binary data, a two-by-two table can be used to summarize the results from each study as follows (Chen and Peace, (2013)[17]):

- as columns, responders against nonresponders,
- as rows, treatment groups.

In each research, we use Oi to indicate the number of participants in the pivotal cell of the two-by-two table. $E(Oi)$ and $\mathrm{Var}(Oi)$ are also used to calculate the expected value and variance of O_i, using the hypergeometric distribution. (Chen and Peace, (2013)[17])

2.2.3 Random-Effects Model: DerSimonian-Laird

Recall that when we meta-analyze numerous studies' effect sizes, for the fixed-effects model the underlying assumption is that the true effect size is equivalent for all of the studies (Chen and Peace, (2013)[17]). This is not always a very practical assumption. It is exceedingly impracticable to demand that every study have the exact same true effect size when integrating a collection of studies using meta-analysis. We do however expect that all of these studies have enough similar information in order to synthesize the information which will be used for statistical inference. (Chen and Peace, (2013)[17])

The random-effects model assumes that the treatment effect $\hat{\delta}_{iR}$ from each study $i = 1, 2, \ldots, K$ is in fact an estimate of its own underlying true treatment effect δ_{iR} which has variance σ_i^2. Furthermore, the δ_{iR} from all of the K studies follows some overall general distribution denoted by $N(\delta, \tau^2)$ (Chen and Peace, (2013)[17]). This random-effects model can be written as follows:

$$\hat{\delta}_{iR} \sim N(\delta_{iR}, \sigma_i^2)$$

$$\delta_{iR} \sim N(\delta, \tau^2) \tag{11}$$

Simply put, the random-effects model is an expansion of the fixed-effects model given in equation (1) as:

$$\hat{\delta}_{iR} = \delta + \nu_i + \epsilon_i \tag{12}$$

where the between-study variation is described by $\nu_i \sim N\left(0, \tau^2\right)$. Equation (11)'s random-effects model may be rewritten as follows if we assume that νi and ϵi are independent:

$$\hat{\delta}_{iR} \sim N(\delta, \sigma_i^2 + \tau^2) \tag{13}$$

The extra parameter τ^2 from the calculation above denotes the variability in the underlying global treatment effect δ between studies. It is now simple to demonstrate, using this formulation, that the weighted mean, which is comparable to that of the fixed-effects model given in equation (3), also estimates the global δ:

$$\hat{\delta}_R = \frac{\sum_{i=1}^{K} w_{iR}\hat{\delta}_{iR}}{\sum_{i=1}^{K} w_{iR}} \tag{14}$$

with standard error estimated as:

$$\text{se}\left(\hat{\delta}_R\right) = \sqrt{\frac{1}{\sum_{i=1}^{K} w_{iR}}} \tag{15}$$

where the weights now are given by:

$$\hat{w}_{iR} = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}^2} \tag{16}$$

$$\text{se}\left(\hat{\delta}_R\right) = \sqrt{\frac{1}{\sum_{i=1}^{K} w_{iR}}} \tag{17}$$

As a result, a 95% CI may be formulated in order to provide statistical inference which is similar to that of the fixed-effects model. (Chen and Peace, (2013)[17])

There are numerous well known methods which may be used to estimate $\hat{\tau}^2$. The most commonly used estimate is from DerSimonian and Laird (DerSimonian and Laird, (1986)[9]) and is derived by using method of moments. The actual procedure is commonly referred to as the Cochran-DerSimonian-Laird procedure as a result of the work that was done by Cochran on synthesizing data from multiple experiments (Chen and Peace, (2013)[17]). The estimate is given as follows:

$$\hat{\tau}^2 = \frac{Q - (K-1)}{U} \tag{18}$$

on condition that $Q > K - 1$, otherwise, $\hat{\tau}^2 = 0$ where

$$Q = \sum_{i=1}^{K} w_i \left(\hat{\delta}_i - \hat{\delta}\right)^2 \tag{19}$$

$$U = \sum_{i=1}^{K} w_i - \frac{\sum_{i=1}^{K} w_i^2}{\sum_{i=1}^{K} w_i} \tag{20}$$

The statistic Q is utilized for the testing of statistical significance of heterogeneity across all studies. (Chen and Peace, (2013)[17])

As previously mentioned, both within-study and between-study variability may be accounted for by the random-effects model, which may be a significant source of heterogeneity for meta-analyses. The random-effects model is therefore more cautious because $w_{iR} \leq w_i$ which leads to (Chen and Peace, (2013)[17])

$$\text{se}\left(\hat{\delta}_R\right) = \sqrt{\frac{1}{\sum_{i=1}^{K} w_{iR}}} \geq \sqrt{\frac{1}{\sum_{i=1}^{K} w_i}} = \text{se}(\hat{\delta}) \tag{21}$$

2.2.3.1 Derivation of DerSimonian-Laird Estimator of τ^2

As previously mentioned, the DerSimonian-Laird Estimator of τ^2 given in equation (18) is derived using method of moments. This is done by simply equating the sample statistic of Q to the corresponding expected value. (Chen and Peace, (2013)[17])

Remember that the sole distinction between fixed-effects and random-effects is the weighting factors in the weighted average. In the random-effects case, $\hat{w}_{iR} = \frac{1}{\hat{\sigma}_i^2 + \hat{\tau}^2}$ and in the fixed-effects case $\hat{w}_i = \frac{1}{\hat{\sigma}_i^2}$ (Chen and Peace, (2013)[17]). The meta-estimators which correspond to th weighting factors are $\hat{\delta}_R = \frac{\sum_{i=1}^{K} w_{iR}\hat{\delta}_{iR}}{\sum_{i=1}^{K} w_{iR}}$ for random-effects and $\hat{\delta} = \frac{\sum_{i=1}^{K} w_i\hat{\delta}_i}{\sum_{i=1}^{K} w_i}$ for fixed-effects. (Chen and Peace, (2013)[17])

It can be shown that both estimators of $\hat{\delta}^*$ and δ are unbiased with expected value of δ. The variances are

$$\text{var}(\hat{\delta}) = \frac{1}{\sum_{i=1}^{K} w_i} \tag{22}$$

$$\text{var}\left(\hat{\delta}_R\right) = \frac{1}{\sum_{i=1}^{K} w_{iR}} \tag{23}$$

Using the above, we are now able to derive the DerSimonian-Laird estimator. The decomposition of Q

is as follows:

$$
\begin{aligned}
Q &= \sum_{i=1}^{K} w_i \left(\hat{\delta}_i - \hat{\delta}\right)^2 = \sum_{i=1}^{K} w_i \left[\left(\hat{\delta}_i - \delta\right) - (\hat{\delta} - \delta)\right]^2 \\
&= \sum_{i=1}^{K} w_i \left(\hat{\delta}_i - \delta\right)^2 - 2\sum_{i=1}^{K} w_i \left(\hat{\delta}_i - \delta\right)(\hat{\delta} - \delta) + \sum_{i=1}^{K} w_i (\hat{\delta} - \delta)^2
\end{aligned} \tag{24}
$$

Therefore the expected values of Q under random-effects model in equation (11) can be shown as follows:

$$
\begin{aligned}
E(Q) &= \sum_{i=1}^{K} w_i E\left(\hat{\delta}_i - \delta\right)^2 - \left(\sum_{i=1}^{K} w_i\right) E(\hat{\delta} - \delta)^2 \\
&= \sum_{i=1}^{K} w_i \operatorname{var}\left(\hat{\delta}_i\right) - \left(\sum_{i=1}^{K} w_i\right) \operatorname{var}(\hat{\delta}) \\
&= \sum_{i=1}^{K} w_i \operatorname{var}\left(\hat{\delta}_i\right) - \left(\sum_{i=1}^{K} w_i\right) \operatorname{var}\left[\frac{\sum_{i=1}^{K} w_i \hat{\delta}_i}{\sum_{i=1}^{K} w_i}\right] \\
&= \sum_{i=1}^{K} w_i \operatorname{var}\left(\hat{\delta}_i\right) - \left(\sum_{i=1}^{K} w_i\right) \frac{\sum_{i=1}^{K} w_i^2 \operatorname{var}\left(\hat{\delta}_i\right)}{\left(\sum_{i=1}^{K} w_i\right)^2} \\
&= \sum_{i=1}^{K} w_i \left(w_i^{-1} + \tau^2\right) - \left(\sum_{i=1}^{K} w_i\right) \frac{\sum_{i=1}^{K} w_i^2 \left(w_i^{-1} + \tau^2\right)}{\left(\sum_{i=1}^{K} w_i\right)^2} \\
&= (K-1) + \tau^2 \left[\sum_{i=1}^{K} w_i - \frac{\sum_{i=1}^{K} w_i^2}{\sum_{i=1}^{K} w_i}\right]
\end{aligned} \tag{25}
$$

For the method of moments, let $E(Q) \approx Q$ to estimate the τ^2 as

$$
\hat{\tau}^2 = \frac{Q - (K-1)}{\sum_{i=1}^{K} w_i - \frac{\sum_{i=1}^{K} w_i^2}{\sum_{i=1}^{K} w_i}} \tag{26}
$$

This is the well-known DerSimonian-Laird method of moments for τ^2 given in equation (18). It should be noted that the estimated variance $\hat{\tau}^2$ can be lower than zero even though the true variance of τ^2 can never be. This condition happens when $Q < df = K - 1$ and when this happens, the estimated $\hat{\tau}^2$ is set to zero. (Chen and Peace, (2013)[17])

2.3 What has been done

We will now consider three main papers which from the foundation and inspiration behind this dissertation.

2.3.1 Comparison of Meta-Analysis Versus Analysis of Variance of Individual Participant-level Data

The first paper to consider is titled *Comparison of meta-analysis versus analysis of variance of individual patient data*, by Olkin, I. & Sampson, A. (Olkin and Sampson, (1998)[6])

As stated in the literature review, the objective that Olkin and Sampson (Olkin and Sampson, (1998)[6]) had was to be able to estimate the relative treatment effect for continuous outcome data when there are several treatments and a control. Olkin and Sampson suitably state that when only the summary data from the studies are available, meta-analytic procedures are the most appropriate, and when the individual participant-level data are available then we should consider analysis of variance (ANOVA) techniques (Olkin and Sampson, (1998)[6]). This led to the obvious question of whether utilizing summary data in a meta-analysis yields different results when compared to a meta-analysis of individual participant-level data. More specifically, what models will provide the same outcomes if we obtain summary data as well as individual participant-level data from the same collection of studies? This equivalence of results is crucial specifically when developing new pharmaceutical drugs.

The main focus in this paper by Olkin, I. & Sampson, A. (Olkin and Sampson, (1998)[6]) is to evaluate the efficacy of employing meta-analysis when there is access to individual participant-level data. Olkin, I. & Sampson, A. also purposely avoid and excluded issues such as tests for homogeneity and whether to use random or fixed effects.

Consider a fixed effects model for treatment and study. Let $y_{ij\alpha}$ denote the αth response on the ith treatment for the jth study, for $\alpha = 1, \ldots, n_{ij}; i = 1, \ldots, m; j = 1, \ldots, k$, where $y_{ij\alpha}, \alpha = 1, \ldots, n_{ij}$, are independently and identically distributed (IID) with common (within study) variance σ^2 (Olkin and Sampson, (1998)[6]). Furthermore, assume homogeneity of the treatment differences across all of the studies, i.e. $\mathrm{E}\left(y_{ij\alpha} - y_{i'j\alpha'}\right) = \tau_i - \tau_{i'}$ for all j, α, and α'.

For simplicity, assume that the within-study estimates of variances are unavailable (Olkin and Sampson, (1998)[6]). This assumption is valid since the primary concern involves estimating these contrasts in main effects, rather than interactions. For additional simplicity, all treatments are required to appear in all studies, i.e., $n_{ij} > 0, i = 1, \ldots, m;\ j = 1, \ldots, k$.

Let the mth treatment serve as the control or baseline treatment. Hence, the concern is with an analysis that centres around the pairwise treatment contrasts relative to the mth 'control' treatment

(Olkin and Sampson, (1998)[6]), that is

$$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_{m-1})' \equiv (\tau_1 - \tau_m, \tau_2 - \tau_m, \ldots, \tau_{m-1} - \tau_m)' \tag{27}$$

Suppose that only the summary data for each study, i.e., $\bar{y}_{ij\cdot}, i = 1, \ldots, m, j = 1, \ldots, k$, were available. The approach is then to use specific meta-analytic procedures in order to estimate $\boldsymbol{\delta}$. This so-called appropriate estimator is in fact a matrix-weighted average of individual study contrast estimators where the matrices are dependent on the sample sizes. If we have access to the individual patient, $\boldsymbol{\delta}$ can be estimated using the least squares estimators (LSE) coming from a two-way additive ANOVA model. (Olkin and Sampson, (1998)[6])

It should be mentioned that the findings of a meta-analysis and an ANOVA of the individual participant-level data yield the same contrast estimates in the case where there is a common or proportionate number of observations for each treatment-study combination i.e. the balanced case (Olkin and Sampson, (1998)[6]). An unexpected result which was obtained is that, meta-analysis and the individual patient analysis provide an equivalent estimator of the contrasts and that the meta-analytic approach produces an effortlessly computed estimator of contrasts for the linear model, for the unbalanced case. (Olkin and Sampson, (1998)[6])

2.3.1.1 The Meta-Analytic Procedure

In 1994, Gleser and Olkin considerd a more general approach to the above model, whereby not all of the treatments are used in every study. This method yields the following meta-analytic procedure which is used to combine treatment comparisons across studies. (Olkin and Sampson, (1998)[6])

We estimate the vector $\boldsymbol{\delta}$ of contrasts within the jth study by

$$\begin{aligned} \mathbf{d}^{(j)} &= (\bar{y}_{1j\cdot} - \bar{y}_{mj\cdot}, \ldots, \bar{y}_{m-1,j\cdot} - \bar{y}_{mj})' \\ &= \bar{y}_{mj\cdot} \begin{bmatrix} \frac{\bar{y}_{1j\cdot}}{\bar{y}_{mj\cdot}} - 1 \\ \vdots \\ \frac{\bar{y}_{m-1j\cdot}}{\bar{y}_{mj\cdot}} - 1 \end{bmatrix} \end{aligned} \tag{28}$$

A core point to note is that the within-study estimates of the contrasts are correlated by means of having a common mean $\bar{y}_{mj\cdot}$. The covariance matrix is easily obtained as follows:

Recall that the within-study variances are all the same i.e $y_{ij\alpha}$ are IID with common variance σ^2. Therefore,

$$\begin{aligned}
\text{cov}\left(\mathbf{d}^{(j)}\right) &= \sigma^2 cov\left(\begin{bmatrix} \frac{\bar{y}_{1j\cdot}}{\bar{y}_{mj\cdot}} - 1 \\ \vdots \\ \frac{\bar{y}_{m-1j\cdot}}{\bar{y}_{mj\cdot}} - 1 \end{bmatrix}\right) \\
&= \sigma^2 \left[\begin{pmatrix} \frac{1}{n_{1j}} & 0 & \dots & 0 \\ 0 & \frac{1}{n_{2j}} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \frac{1}{n_{m-1j}} \end{pmatrix} + \frac{1}{n_{mj}}\mathbf{e}\mathbf{e}'\right] \\
&= \sigma^2 \left[diag(\frac{1}{n_{1j}}, \frac{1}{n_{2j}}, \dots, \frac{1}{n_{m-1j}}) + \frac{1}{n_{mj}}\mathbf{e}\mathbf{e}'\right] \\
&\equiv \sigma^2 \Sigma_j
\end{aligned} \tag{29}$$

where $\mathbf{e}$ is the unit vector. Now, if we use a regression format we are able to obtain the estmate of $\boldsymbol{\delta}$ across all of the studies as follows:

Let

- $\mathbf{H} = (\mathbf{I}_{m-1}, \dots, \mathbf{I}_{m-1})'$ with $\mathbf{I}_{m-1}$ being the identify matrix of order $m-1$
- $\boldsymbol{\Sigma} = \sigma^2 diag(\Sigma_1, \dots, \Sigma_k)$
- $\mathbf{d} = (\mathbf{d}^{(1)}, \dots, \mathbf{d}^{(k)})$

Now

$$\mathbf{H}\boldsymbol{\delta} = \mathbf{d}$$

Multiplying both sides with $\mathbf{H}'\boldsymbol{\Sigma}^{-1}$ we obtain

$$\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H}\boldsymbol{\delta} = \mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{d}$$

Therefore, the estimate of $\boldsymbol{\delta}$ across all studies is

$$\hat{\boldsymbol{\delta}} = (\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H})^{-1}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{d} \tag{30}$$

In equation (30), $\mathbf{H}'\mathbf{\Sigma}^{-1}\mathbf{H} = \Sigma_1^{-1} + \cdots + \Sigma_k^{-1}$, therefore the solution in equation (30) simplifies to

$$\hat{\boldsymbol{\delta}} = (\Sigma_1^{-1} + \cdots + \Sigma_k^{-1})^{-1}(\Sigma_1^{-1}\mathbf{d}^{(1)} + \cdots + \Sigma_k^{-1}\mathbf{d}^{(k)}) \tag{31}$$

Equation (31) represents a matrix-weighted linear combination of the individual study estimates, accounting for the corresponding sample sizes. It therefore follows that

$$\begin{aligned} E(\hat{\boldsymbol{\delta}}) &= E((\Sigma_1^{-1} + \cdots + \Sigma_k^{-1})^{-1}(\Sigma_1^{-1}\mathbf{d}^{(1)} + \cdots + \Sigma_k^{-1}\mathbf{d}^{(k)})) \\ &= \boldsymbol{\delta} \end{aligned} \tag{32}$$

Thus, $\hat{\boldsymbol{\delta}}$ is an unbiased estimator of $\boldsymbol{\delta}$. Furthermore,

$$\begin{aligned} cov(\hat{\boldsymbol{\delta}}) &= cov[(\Sigma_1^{-1} + \cdots + \Sigma_k^{-1})^{-1}(\Sigma_1^{-1}\mathbf{d}^{(1)} + \cdots + \Sigma_k^{-1}\mathbf{d}^{(k)})] \\ &= \sigma^2(\Sigma_1^{-1} + \cdots + \Sigma_k^{-1})^{-1} \end{aligned} \tag{33}$$

2.3.1.2 The Procedure for the Individual Participant-level Data Analysis

Redescribing the method in a linear model format, we additionally have the effects of studies which in turn results in the familiar two-way ANOVA model with no interactions. This absence of the interaction terms is consistent with the meta-analysis assumption of homogeneity of treatment contracts among all of the studies. The unbalanced model is where all of the studies and treatments sample sizes are arbitrary. The unbalanced model has the rebuke that if we disregard special cases, solutions to the normal equations have no simple form. A suitable recommendation in order to numerically solve the normal equations is surely sufficient for data analysis, however causes challenges when evaluating the efficiency in the linear model of the estimator $\hat{\boldsymbol{\delta}}$ given by equation (31). As a result therefore, Olkin, I. & Sampson, A. (Olkin and Sampson, (1998)[6]) developed a variation of the analysis of the two-way model.

The linear model is written as

$$y_{ij\alpha} = \mu + \tau_i + \nu_j + \epsilon_{ij\alpha}, \quad \alpha = 1, \ldots, n_{ij}, i = 1, \ldots, m, j = 1, \ldots, k \tag{34}$$

where $\tau_1, \ldots, \tau_m$ are the fixed treatment effects, $\nu_1, \ldots, \nu_k$ are the fixed study effects, and where the $n = \Sigma n_{ij}$ random variables $\{\epsilon_{ij\alpha}\}$ are independently and identically distributed with mean 0 and common variance σ^2. We write $\Sigma_{\alpha=1}^{n_{ij}} y_{ij\alpha} = y_{ij\cdot}$ and $y_{ij\cdot}/n_{ij} = \bar{y}_{ij\cdot}$ and so forth. The normal equations for this specific model contain a total of $m + k + 1$ variables but are of rank $m + k - 1$. Using a common

generalized inverse to $X'X$, where X is the corresponding $n \times (m+k+1)$ design matrix, we are able to obtain simplified normal equations and furthermore gives the variance of the treatment contrasts in a nonsimplified form. (Olkin and Sampson, (1998)[6])

To obtain the covariance matrix of the least squares estimators (LSE) of $\boldsymbol{\delta}$ in a useful form for the above argument, consider the following calculations.

Firstly, set $\mu = 0$ and $\tau_m = 0$ which parameterize the model. Denote the remaining parameters by $\tau_1^*, \ldots, \tau_{m-1}^*, \nu_1^*, \ldots, \nu_k^*$. For $i = 1, 2, \ldots, m-1$ the parameter τ_i^* now represents $\tau_i - \tau_m$, which is estimable with LSE $\hat{\tau}_i^* - \hat{\tau}_m^* = \hat{\tau}_i^* - 0 = \hat{\tau}_i^*$.

We therefore obtain the following model, which is of full rank:

$$\operatorname{cov}\left(\hat{\tau}_1^*, \ldots, \hat{\tau}_{m-1}^*, \hat{\nu}_1^*, \ldots, \hat{\nu}_k^*\right) = \sigma^2 \begin{pmatrix} D_{1,} & N^* \\ N^{*'} & D_2 \end{pmatrix}^{-1} \tag{35}$$

because,

$$\operatorname{cov}\left(\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}'\right) = \begin{pmatrix} \operatorname{cov}\left(\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\tau}}'\right) & \operatorname{cov}\left(\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}'\right) \\ \operatorname{cov}\left(\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}'\right) & \operatorname{cov}\left(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\nu}}'\right) \end{pmatrix}$$

where $\left(\hat{\tau}_1^*, \ldots, \hat{\tau}_{m-1}^*, \hat{\nu}_1^*, \ldots, \hat{\nu}_k^*\right)$ are the LSEs for the reparameterized model,

$$D_1 = \begin{pmatrix} n_{1.} & 0 & \ldots & 0 \\ 0 & n_{2.} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & n_{m-1.} \end{pmatrix}$$

$$= \operatorname{diag}\left(n_{1.}, \ldots, n_{m-1.}\right)$$

$$D_2 = \begin{pmatrix} n_{.1} & 0 & \ldots & 0 \\ 0 & n_{.2} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & n_{.k} \end{pmatrix}$$

$$= \operatorname{diag}\left(n_{.1}, \ldots, n_{.k}\right)$$

$$N^* = \begin{pmatrix} n_{11} & \dots & n_{1k} \\ \vdots & \vdots & \vdots \\ n_{m-1,1} & \dots & n_{m-1,k} \end{pmatrix}$$
$$\equiv (\mathbf{n}_1^*, \cdots, \mathbf{n}_k^*)\,.$$

In particular,

$$\operatorname{cov}\left(\hat{\tau}_1^*, \ldots, \hat{\tau}_{m-1}^*\right) = \sigma^2 \left(D_1 - N^* D_2^{-1} N^{*'}\right)^{-1} \tag{36}$$

2.3.1.3 Equivalence of the Meta-Analysis and Pooled Analysis Estimators

A direct method which can be used in order to show the equivalence of the two estimators is to simply solve the normal equations in equation (37) and (38). The true LSE, $\hat{\tau}_1^*, \ldots, \hat{\tau}_{m-1}^*, \hat{\nu}_1^*, \ldots, \hat{\nu}_k^*$, are given by first solving (37) below:

$$c_{ii}\hat{\tau}_i^* + \sum_{\substack{i'=1 \\ i' \neq i}}^{m-1} c_{ii'}\hat{\tau}_{i'}^* = r_i \tag{37}$$

where

$$r_i = y_{i\cdot} - \sum_{j=1}^{k} n_{ij}\bar{y}_{\cdot j\cdot}$$
$$c_{ii} = n_{i\cdot} - \sum_{j=1}^{k} \frac{n_{ij}^2}{n_{\cdot j}}$$

and for $i \neq i'$,

$$c_{ii'} = -\sum_{j=1}^{k} n_{ij} n_{i'j} / n_{\cdot j}$$

and then substituting into and solving

$$\hat{\nu}_j^* = \bar{y}_{\cdot j\cdot} - \frac{1}{n_{\cdot j}} \sum_{i=1}^{m-1} n_{ij}\hat{\tau}_i^*, \quad j = 1, \ldots, k \tag{38}$$

for the $\left\{\hat{\nu}_j^*\right\}$. This approach can be algebraically strenuous however, using the Gauss-Markov theorem (Appendix, 6.1) we are able to show this equivalence in a strainghforward way. Now, to show that $\hat{\boldsymbol{\delta}} = (\Sigma_1^{-1} + \cdots + \Sigma_k^{-1})^{-1}(\Sigma_1^{-1}\mathbf{d}^{(1)} + \cdots + \Sigma_k^{-1}\mathbf{d}^{(k)})$ is the LSE for the ANOVA model, we observe by the Gauss-Markov theorem:

1. $\mathrm{E}(\hat{\boldsymbol{\delta}}) = (\tau_1 - \tau_m, \ldots, \tau_{m-1} - \tau_m)'$, i.e. $\hat{\boldsymbol{\delta}}$ is an unbiased estmator of $\boldsymbol{\delta}$.

2. $\hat{\boldsymbol{\delta}}$ is a linear estimator in $\bar{y}_{11\cdot}, \ldots, \bar{y}_{mk\cdot}$

3. The diagonal elements of $\operatorname{cov}(\hat{\boldsymbol{\delta}})$ and the covariance matrix's diagonal elements of the LSE are the

same. This furthermore implies that the least squares variance is obtained by the meta-analysis estimator that is both linear as well as unbiased.

To prove point 3, it is sufficient to show that

$$\Sigma_1^{-1} + \cdots + \Sigma_k^{-1} = D_1 - N^* D_2^{-1} N^{*'} \tag{39}$$

We know from equation (29) that

$$\Sigma_j = diag(\frac{1}{n_{1j}}, \frac{1}{n_{2j}}, \ldots, \frac{1}{n_{m-1j}}) + \frac{1}{n_{mj}} \mathbf{e}\mathbf{e}'$$

Remark 1: $\left(A + \underset{\text{NN}}{b}\, b'\right)^{-1} = A^{-1} - A^{-1} \underset{\text{NN}}{b}\, b' A^{-1} / (a + \underset{\text{N}}{b'} A^{-1} \underset{\text{N}}{b})$

Therefore, using the remark above we are able to represent the Σ_j^{-1} as follows:

$$\Sigma_j^{-1} = \text{diag}\,(n_{1j}, \ldots, n_{m-1,j}) - \mathbf{n}_j^* \mathbf{n}_j^{*'} / n \cdot j \tag{40}$$

Remark 2:

$$N^* D_2^{-1} N^{*'} = \sum_1^k \mathbf{n}_j^* \mathbf{n}_j^{*'} / n \cdot j \tag{41}$$

which results in the following from equation (39),

$$\begin{aligned} \Sigma_1^{-1} + \cdots + \Sigma_k^{-1} &= \text{diag}\,(n_1, \ldots, n_{m-1,}) - \sum_1^k \mathbf{n}_j^* \mathbf{n}_j^{*'} / n \cdot j \\ &= D_1 - N^* D_2^{-1} N^{*'} \end{aligned} \tag{42}$$

which was what we aimed to prove.

2.3.2 On the Equivalence of Meta-Analysis Using Literature and Using Individual Participant-level Data

The next paper to consider is titled *On the Equivalence of Meta-Analysis Using Literature and Using Individual Patient Data*, by Mathew, T. & Nordström, K. (Mathew and Nordström, (1999)[4])

This paper extends the previous work done by Olkin and Sampson (Olkin and Sampson, (1998)[6]) to a much more general case. Recall that Olkin and Sampson (Olkin and Sampson, (1998)[6]) successfully obtained the conclusion that the two meta-analytical methods are equivalent in that the same estima-

tor is produced by both. This conclusion was established under the assumption that all observations have a common variance and are independent within and between studies. Therefore the aim of this paper by Mathew and Nordström is to demonstrate a much more robust setting in which the result is true, specifically where it is not necessary for the within study observations to be independent, and various covariance structures may exist between studies' observations. The key discovery demonstrates that the meta-analysis estimate derived from best linear unbiased estimators (BLUEs) from numerous studies is comparable to the BLUE produced from a two-way model without interaction for individual participant-level data, especially when the studies are independent. The result is presented in a more generic and robust environment where each study's covariance matrices are arbitrary positive definite matrices.

2.3.2.1 The Problem and the Two-Way Model

Let $y_{ij\alpha}$ follow the same notation as given in the previous paper by Olkin and Sampson (Olkin and Sampson, (1998)[6]). Recall that the mth treatment was chosen to serve as the control, m represents the treatments, and lastly k is the studies.

Recall also that $\tau_1, \tau_2, \ldots, \tau_m$ represents the treatment effects, so that τ_m is the effect due to the control. Assume that $\mathrm{E}(y_{ij\alpha}) = \mu_j + \tau_i$, for all i, j and α, i.e., the observations follow a one-way classification model, within the j th study, and the μ_j 's are unknown parameters.

Note:

- $\mathbf{y}_{ij.} = \left(y_{ij1}, \ldots, y_{ijn_{ij}}\right)'$
- $\mathbf{y}_{.j.} = \left(\mathbf{y}'_{1j.}, \ldots, \mathbf{y}'_{mj.}\right)'$
- $n_{.j} = \Sigma_{i=1}^{m} n_{ij}$
- $\tau = (\tau_1, \ldots, \tau_m)'$

Let $\mathbf{1}_p$ denote the vector of ones for p components and also let $\Delta_j = \mathrm{diag}\left(1_{n_{1j}}, \ldots, 1_{n_{mj}}\right)$. Suppose that we have an arbitrary covariance matrix V_j for $\mathbf{y}_{.j.}$, where the V_j 's are $n_{.j} \times n_{.j}$ positive definite matrices. Recall that in the paper by Olkin and Sampson (Olkin and Sampson, (1998)[6]), it was assumed that $V_j = \sigma^2 I_{n_j}$, for $j = 1, 2, \ldots, k$. Thus

$$\mathrm{E}(\mathbf{y}_{.j.}) = \mu_j \mathbf{1}_{n_{.j}} + \Delta_j \boldsymbol{\tau}, \quad \mathrm{cov}(\mathbf{y}_{.j.}) = V_j \tag{43}$$

where the $\mathbf{y}_{.j.}$'s are independent for $j = 1, \ldots, k,$, suggesting that the observations from different studies

are independent. This paper's biggest challenge is estimating the parameter vector δ, specifically for the differences of the treatment-control. This parameter vector δ is given by:

$$\boldsymbol{\delta} = (\tau_1 - \tau_m, \tau_2 - \tau_m, \ldots, \tau_{m-1} - \tau_m)' = L\boldsymbol{\tau} \tag{44}$$

where $L = (I_{m-1}, -1_{m-1})$.

In addition to the aforementioned model, we must additionally consider the meta-analysis model, specifically using individual participant-level data, which is directly related to meta-analysis utilizing summary data. For the $y_{ij\alpha}$'s, take into account the later model's two-way (non-interaction) model, which is represented by

$$y_{ij\alpha} = \mu + \tau_i + \nu_j + \epsilon_{ij\alpha} \tag{45}$$

for $\alpha = 1, 2, \ldots, n_{ij}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, k$, where all effects are taken to be fixed effects and ν_j reflects the effects resulting from the studies. This essentially indicates that (45) reflects an unbalanced fixed-effects two-way model, but the covariance structure provided in (43) is assumed.

2.3.2.2 Meta-Analysis of Summary Data

The BLUE of δ provided in equation (44), together with its covariance matrix, may be used to concisely summarize the results from the jth study. The BLUE of $\boldsymbol{\delta}$ is represented by $\mathbf{d}^{(j)}$, which is based on information from the jth study, or in other terms $\mathbf{y}_{.j.}$. Using model (43), we are able to compute the BLUE. Without losing generality, we may further assume that $\mu_j = 0$ in (43). Mathew and & Nordström (Mathew and Nordström, (1999)[4]) therefore obtained the following:

$$\begin{aligned} \mathbf{d}^{(j)} &= L\left(\Delta_j' V_j^{-1} \Delta_j\right)^{-1} \Delta_j' V_j^{-1} \mathbf{y}_{.j.} \\ \operatorname{cov}\left(\mathbf{d}^{(j)}\right) &= L\left(\Delta_j' V_j^{-1} \Delta_j\right)^{-1} L' = \Sigma_j \end{aligned} \tag{46}$$

for $j = 1, 2, \ldots, k$, where the matrix L is given in equation (44). As seen in Olkin and Sampson (Olkin and Sampson, (1998)[6]), the combined estimator $\hat{\boldsymbol{\delta}}$ from the meta-analytic approach is provided by equation (31) and (33) where the Σ_j's are given in (46).

2.3.2.3 Meta-Analysis of Individual Participant-level Data

We now consider the model based on equation (45) under assumption given in equation (44). We denote the BLUE of $\boldsymbol{\delta}$ by $\widetilde{\boldsymbol{\delta}}$. It should be noted that a representation for $\widetilde{\boldsymbol{\delta}}$ was not obtained since it was eventually show that $\tilde{\delta}$ coincides with $\hat{\delta}$ given in (31). A representation for cov($\tilde{\boldsymbol{\delta}}$) was however obtained

utilizing a reparameterization developed by Olkin and Sampson (Olkin and Sampson, (1998)[6]).

Therefore, assume that $\tau_m = 0$ and $\mu = 0$ in (45) and denote the remaining parameters by $\tau_1^*, \tau_2^*, \ldots, \tau_{m-1}^*$ and $\nu_1^*, \nu_2^*, \ldots, \nu_k^*$. In terms of the τ_i 's, we then have that $\tau_i^* = \tau_i - \tau_m, i = 1, 2, \ldots, m-1$. In respect of this reparameterization, the mean of $\mathbf{y}_{.j.}$ obtained from equation (45), can be written as $\mathrm{E}(\mathbf{y}_{.j.}) = \Delta_j^* \boldsymbol{\tau}^* + \mathbf{1}_{.nj}\nu_j^*$, with $\tau^* = \left(\tau_1^*, \tau_2^*, \ldots, \tau_{m-1}^*\right)'$. Also,

$$\Delta_j^* = \begin{pmatrix} \mathrm{diag}\left(\mathbf{1}_{n_{1j}}, \mathbf{1}_{n_{2j}}, \ldots, \mathbf{1}_{n_{m-1,j}}\right) \\ 0 \end{pmatrix}$$

where the 0 in the expression for Δ_j^* represents an $n_{mj} \times (m-1)$ matrix of zeros. Now write,

$$\begin{aligned} \mathbf{y} &= \left(\mathbf{y}'_{.1.}, \mathbf{y}'_{.2.}, \ldots, \mathbf{y}'_{.k.}\right)' \\ \Delta^* &= \left(\Delta_1^{*'}, \Delta_2^{*'}, \ldots, \Delta_k^{*'}\right)' \\ \Lambda &= \mathrm{diag}\left(1_{n_{.1}}, \mathbf{1}_{n,2}, \ldots, 1_{n_{.k}}\right) \\ \boldsymbol{\nu}^* &= \left(\nu_1^*, \nu_2^*, \ldots, \nu_k^*\right)' \end{aligned} \tag{47}$$

As a result we therefore have the following linear model

$$\begin{aligned} \mathrm{E}(\mathbf{y}) &= \Delta^* \boldsymbol{\tau}^* + \Lambda \boldsymbol{\nu}^* \\ \mathrm{cov}(\mathbf{y}) &= \mathrm{diag}\left(V_1, V_2, \ldots, V_k\right) = V \end{aligned} \tag{48}$$

If $\left(\tilde{\tau}_1^*, \tilde{\tau}_2^*, \ldots, \tilde{\tau}_{m-1}^*, \tilde{\nu}_1^*, \tilde{\nu}_2^*, \ldots, \tilde{\nu}_k^*\right)$ is the BLUE of $\left(\tau_1^*, \tau_2^*, \ \ldots, \tau_{m-1}^*, \nu_1^*, \nu_2^*, \ldots, \nu_k^*\right)$, computed based on the model in equation (48) the BLUE of $\boldsymbol{\delta}$ is the quantity $\tilde{\boldsymbol{\delta}} = \left(\tilde{\tau}_1^*, \tilde{\tau}_2^*, \ldots, \tilde{\tau}_{m-1}^*\right)'$. We also note that

$$\begin{aligned} &\mathrm{cov}\left(\tilde{\tau}_1^*, \tilde{\tau}_2^*, \ldots, \tilde{\tau}_{m-1}^*, \tilde{\nu}_1^*, \tilde{\nu}_2^*, \ldots, \tilde{\nu}_k^*\right) \\ &= \begin{pmatrix} \Delta^{*'} V^{-1} \Delta^* & \Delta^{*'} V^{-1} \Lambda \\ \Lambda' V^{-1} \Delta^* & \Lambda' V^{-1} \Lambda \end{pmatrix}^{-1} \end{aligned}$$

Hence,

$$\begin{aligned} \mathrm{cov}(\tilde{\boldsymbol{\delta}}) &= \mathrm{cov}\left(\tilde{\tau}_1^*, \tilde{\tau}_2^*, \ldots, \tilde{\tau}_{m-1}^*\right) \\ &= \left(\Delta^{*'} V^{-1} \Delta^* - \Delta^{*'} V^{-1} \Lambda \left(\Lambda' V^{-1} \Lambda\right)^{-1} \Lambda' V^{-1} \Delta^*\right)^{-1} \end{aligned} \tag{49}$$

2.3.2.4 The Equivalence Result

From the previous section, it is clear to see that the BLUE $\tilde{\boldsymbol{\delta}}$ and the estimator $\hat{\boldsymbol{\delta}}$ are both unbiased estimators of $\boldsymbol{\delta}$. To prove that $\tilde{\boldsymbol{\delta}} = \hat{\boldsymbol{\delta}}$, it is sufficient to prove that $\text{cov}(\hat{\boldsymbol{\delta}}) = \text{cov}(\tilde{\boldsymbol{\delta}})$. To do this, we firstly need to simplify $\text{cov}(\tilde{\boldsymbol{\delta}})$, which is given equstion (49). Using direct computations as well as the expressions for Δ^*, Λ, and V shown in equations (47) and (48) result in

$$\begin{aligned}
\Delta^{*'} V^{-1} \Delta^* &= \Sigma_{j=1}^{k} \Delta_j^{*'} V_j^{-1} \Delta_j^* \\
\Delta^{*'} V^{-1} \Lambda &= \left(\Delta_1^{*'} V_1^{-1} \mathbf{1}_{n.1}, \Delta_2^{*'} V_2^{-1} \mathbf{1}_{n.2}, \ldots, \Delta_k^{*'} V_k^{-1} \mathbf{1}_{n.k} \right) \\
\Lambda' V^{-1} \Lambda &= \text{diag} \left(\mathbf{1}'_{n.1} V_1^{-1} \mathbf{1}_{n.1}, \mathbf{1}'_{n.2} V_2^{-1} \mathbf{1}_{n_2 2}, \ldots, \mathbf{1}'_{n.k} \; V_k^{-1} \mathbf{1}_{n.k} \right)
\end{aligned}$$

Hence,after simplification of equation (49) we obtain

$$\text{cov}(\tilde{\boldsymbol{\delta}}) = \left[\sum_{j=1}^{k} \Delta_j^{*'} \left(V_j^{-1} - \frac{V_j^{-1} \mathbf{1}_{n,j} \mathbf{1}'_{n,j} V_j^{-1}}{\mathbf{1}'_{n,j} V_j^{-1} \mathbf{1}_{n,j}} \right) \Delta_j^* \right]^{-1} \tag{50}$$

A more robust term-by-term equality will be established

$$\Sigma_j^{-1} = \Delta_j^{*'} \left(V_j^{-1} - \frac{V_j^{-1} \mathbf{1}_{n,j} \mathbf{1}'_{n,j} V_j^{-1}}{\mathbf{1}'_{n,j} V_j^{-1} \mathbf{1}_{n,j}} \right) \Delta_j^* \tag{51}$$

Considering equation (50), the covariance matrices $\text{cov}(\hat{\boldsymbol{\delta}})$ and that provided in equation (49) are equivalent after equation (51) is verified.

To establish (51), recall that Σ_j is the covariance matrix of $\mathbf{d}^{(j)}$, where $\mathbf{d}^{(j)}$ is the BLUE of $\boldsymbol{\delta} = L\boldsymbol{\tau}$ dependant on the model given by equstion (43). We need to show that

$$\left[\Delta_j^{*'} \left(V_j^{-1} - \frac{V_j^{-1} \mathbf{1}_{n.j} \mathbf{1}'_{n.j} V_j^{-1}}{\mathbf{1}'_{n.j} V_j^{-1} \mathbf{1}_{n.j}} \right) \Delta_j^* \right]^{-1}$$

is furthermore the covariance matrix for the BLUE of δ assuming that the model in equation (43)'s parameterization is different. It is plain to see that model (43) is equivalent to

$$\text{E}\left(\mathbf{y}_{.j.}\right) = \mu_j^* \mathbf{1}_{n.j} + \Delta_j^* \boldsymbol{\tau}^*, \quad \text{cov}\left(\mathbf{y}_{.j.}\right) = V_j \tag{52}$$

where $\mu_j^* = \mu_j + \tau_m$. Recall that $\delta = \tau^*$. Furthermore, we can demonstrate through direct calculations

that if $\tau^{*(j)}$ indicates the BLUE of τ^* from model (52), then

$$\operatorname{cov}\left(\tau^{*(j)}\right) = \left[\Delta_j^{*'}\left(V_j^{-1} - \frac{V_j^{-1}\mathbf{1}_{n,j}\mathbf{1}'_{n.j}V_j^{-1}}{\mathbf{1}'_{n.j}V_j^{-1}\mathbf{1}_{n,j}}\right)\Delta_j^*\right]^{-1}$$

The inverse of the covariance matrix of the BLUE of δ, generated under model (43) for a given j is represented by both sides of (51). As a result, equation (51)'s equality is valid, proving that $\operatorname{cov}(\hat{\boldsymbol{\delta}})$ and $\operatorname{cov}(\tilde{\boldsymbol{\delta}})$ are equal.

It is important to note that irrespective of which covariance matrix is utilized in model (43), the equivalence result remains, even though Olkin and Sampson (Olkin and Sampson, (1998)[6]) have assumed $V_j = \sigma^2 I_{n.j}$. Some other well known choices of V_j are:

$$\begin{aligned} V_j &= \sigma_j^2 I_{n.j} \\ V_j &= \sigma_j^2 I_{n.j} + \rho_j \mathbf{1}_{n.j}\mathbf{1}'_{n\cdot j} \\ V_j &= \operatorname{diag}\left(\sigma_{11}^2 I_{n_{1j}}, \sigma_{21}^2 I_{n_{2j}}, \ldots, \sigma_{m1}^2 I_{n_{mj}}\right) \end{aligned} \tag{53}$$

for $j = 1, 2, \ldots, m$.

2.3.3 Comparison of One-Step and Two-Step Meta-Analysis Models Using Individual Participant-level Data

The final paper we consider is titled *Comparison of One-Step and Two-Step Meta-Analysis Models Using Individual Patient Data*, by Mathew, T. & Nordström, K. (Mathew and Nordström, (2010)[5])

According to Mathew and Nordström (Mathew and Nordström, (2010)[5]), the main statistical challenge in a multicenter experiment or meta-analysis is estimating and/or inferring a common effect (or many common effects) based on the results of independent trials. As a result, in both meta-analysis and multicenter trials, the effectiveness of a combination estimator or that of a combination of confidence intervals and tests is a critical statistical problem.

It is possible to estimate a linear function of the mean using both a linear model for individual participant-level data (IPD) and a linear model for summary data. The statistical goal is to compare these estimators. The corresponding estimators are referred to as the one-step IPD meta-analysis estimator and the two-step IPD meta-analysis estimator, respectively. It is inevitable that several potential causes of concern, including selection bias or publishing bias, will be present in any practical meta-analytic modelling of trials, it is well known that thier inclusion would significantly complicate matters.

In order to avoid such complications, Mathew and Nordström (Mathew and Nordström, (2010)[5]) chose to exclude all sources of concern.

The motivation behind this article is that there doesn't appear to be a common statistical paradigm employed in literature that allows for a direct comparison of the IPD meta-analysis estimators for both the one-step and two-step.

Assume that the objective is to estimate and deduce from a linear function of the mean vector. This is demonstrated by the entire vector of treatment-control differences in an analysis of variance (ANOVA) configuration. Assume moreover that each of the trials' related covariance matrices and summary estimators are available. These include the best linear unbiased estimators (BLUE) or maximum likelihood estimators (MLE) under normality assumptions. In order for the two-step IPD meta-analysis estimator to be created in an ideal way, the summary estimators may be combined. On the other hand, if data from all of the trials were combined, an MLE or BLUE could be obtained by establishing a linear model for the full vector of responses, i.e. we have an IPD meta-analysis estimator, specifically for the one-step (Mathew and Nordström, (2010)[5]). Whether these two estimators generally coincide is a natural issue to pose in this situation, and if they do not, under what circumstance(s) might one anticipate this to occur.

Mathew and Nordström (Mathew and Nordström, (2010)[5]) consider two examples.

Example 1: Consider the following example which is taken from Whitehead (Whitehead, (2003)[10]). In this study, the recovery durations of patients who undergo short surgical procedures are compared between two distinct anesthetics, A and B. The mean and standard deviations (SD) of the log-transformed recovery times (in minutes) are taken into account. The data are collected from nine distinct centers.

The population mean log-recovery times for both anesthetics A and B are represented by μ_A and μ_B. The objective is to therefore estimate $\mu_A - \mu_B$.

2.3.3.1a One-step IPD meta-analysis estimator

The following model is used to calculate the one-step IPD meta-analysis estimator of $\mu_{\mathrm{A}} - \mu_{\mathrm{B}}$. The sample mean log-recovery times for anaesthetics A and B for the jth centre, are represented by y_{A_j} and y_{B_j}. The variance of the recovery times for the jth centre is represented by σ_j^2. Lastly the sample sizes (i.e. the number of patients) for anaesthetics A and B at the jth centre are represented by n_{A_j} and n_{B_j}.

For the mean log-recovery times at the jth center, let's assume the following linear model.

$$\mathrm{E}\begin{pmatrix} y_{\mathrm{A}_j} \\ y_{\mathrm{B}_j} \end{pmatrix} = \begin{pmatrix} \mu_{\mathrm{A}} \\ \mu_{\mathrm{B}} \end{pmatrix}, \quad \mathrm{Cov}\begin{pmatrix} y_{\mathrm{A}_j} \\ y_{\mathrm{B}_j} \end{pmatrix} = \sigma_j^2 \begin{pmatrix} 1/n_{\mathrm{A}_j} & 0 \\ 0 & 1/n_{\mathrm{B}_j} \end{pmatrix} = V_j \tag{54}$$

Rewrite $\mathbf{y}_j = \left(y_{\mathrm{A}_j}, y_{\mathrm{B}_j}\right)'$, the IPD model is created by stacking the respective mean log-recovery times from each center into a single column vector as $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_9)'$

$$\mathrm{E}(\mathbf{y}) = W \begin{pmatrix} \mu_{\mathrm{A}} \\ \mu_{\mathrm{B}} \end{pmatrix}, \quad \mathrm{Cov}(\mathbf{y}) = \mathrm{diag}\left(V_1, \ldots, V_9\right) = V \tag{55}$$

where the Kronecker product $W = 1_9 \otimes I_2$, and the column vector of nine ones, or 1_9, represents the number of centers. From the model mentioned above, we derive the one-step IPD meta-analysis estimator of $\mu_{\mathrm{A}} - \mu_{\mathrm{B}}$.

2.3.3.1b Two-step IPD meta-analysis estimator

It is also possible to estimate the mean difference $\mu_A - \mu_B$. This is accomplished by utilizing the corresponding estimators from each of the centers together with their standard errors, then combining these estimators of $\mu_A - \mu_B$ into a weighted linear combination. This is what is known as the the two-step IPD meta-analysis estimate of $\mu_A - \mu_B$. When the summary data from the multiple centers includes the estimators of $\mu_A - \mu_B$ as well as the accompanying standard errors, this particular estimator may be computed with ease.

As previously mentioned, a natural question to ask here is if the one-step and two-step IPD meta-analysis estimators coincide, and if not, how much they will diverge.

Example 2: This example is given in Bower et al. (Bower et al., (2003)[11]), where the costs associated with counseling are taken into account, specifically in primary care. Data on the expenditures for patients who received counseling vs those who continued receiving treatment from a general practitioner were gathered from a wide range of research (trials). The data are obtained from four studies and the mean as well as standard deviations (SD) for each are considered.

The issue therefore becomes the estimation of $\mu_1 - \mu_2$ if μ_1 and μ_2 represent the population average cost for patients treated by counselors and for patients who continued under the supervision of a general practitioner, respectively. By taking into account a model that is comparable to the one in the prior example i.e. Example 1, the $\mu_1 - \mu_2$ estimates from the one- and two-step IPD meta-analysis are available to us.

The potential loss of efficiency that follows from the utilization of the two-step IPD meta-analysis estimator is clearly of importance in both Example 1 and Example 2. A general application of the primary finding shows that two estimators in fact do not coincide unless the proportion of observations belonging to the first treatment is constant across trials. It should be noted that this portion of observations is almost the same for Example 1, however this is not the case in Example 2. In contrast to Example 2, the efficiency loss can be expected to be minimal in Example 1.

2.3.3.2 The Model and the Meta-Analysis Problem

Take into account k separate trials, and use $\mathbf{y}_j$ to represent the vector of n_j replies in the jth trial. Therefore, $\mathbf{y}_1, \ldots, \mathbf{y}_k$ contains the IPD. Assume that a linear model is used for the responses that have the following form

$$\mathrm{E}\left(\mathbf{y}_j\right) = W_j\boldsymbol{\beta} + Z_j\boldsymbol{\delta}_j, \quad \mathrm{Cov}\left(\mathbf{y}_j\right) = V_j \tag{56}$$

for $j = 1, \ldots, k$, where $\boldsymbol{\beta}$ with dimension p and $\boldsymbol{\delta}_j$ with dimension q represent vectors of unknown parameters. The design matrices are given by W_j and Z_j. For example, in model (54) of Example 1 above, δ is absent, $W_j = I_2$ and $\boldsymbol{\beta} = (\mu_{\mathrm{A}}, \mu_{\mathrm{B}})'$. The covariance matrix, Vj, of the responses in (56), typically contains unidentified parameters. Instead, assume that Vj has no unknown parameters and that these parameters have been estimated. This assumption is very common in meta-analysis literature. It is important to note that this design does support both fixed and random effects, with the latter appearing as variance components in V_j. Therefore, there is no between-study variance.

The interested parameter is a linear function of the trial-specific main parameter, shown by $\boldsymbol{\beta}$. In order to account for trial-specific quantities, the model includes what are referred to as nuisance parameters, represented by $\boldsymbol{\delta}_j s$. The nuisance parameters that are present in all of the trials are included in the parameter $\boldsymbol{\beta}$. Since these nuisance factors are both unique to the trials and shared by all trials, the model is sufficiently broad to incorporate them. Let's now assume that the matrix $X_j = (W_j, Z_j)$ has the full rank $p + q_j$. Based on the information from the jth trial, the weighted least squares estimator for $\left(\boldsymbol{\beta}', \boldsymbol{\delta}_j\right)'$, is provided as follows:

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}^{(j)} \\ \hat{\boldsymbol{\delta}}_j \end{pmatrix} = \left(X_j' V_j^{-1} X_j\right)^{-1} X_j' V_j^{-1} \mathbf{y}_j$$

This estimate is called the BLUE estimator, or best (i.e., minimal covariance) linear unbiased estima-

tor, of $\left(\boldsymbol{\beta}', \boldsymbol{\delta} j'\right)'$, and is based on $\mathbf{y}_j$. Assuming normality for the responses, it is then also the MLE.

Suppose we are interested in the parameter $\boldsymbol{\theta} = L\boldsymbol{\beta}$, with L an $s \times p$ matrix which consists of full row rank. If we employ the model presented in equation (56), together with the known covariance matrix Vj, then the BLUE of $\boldsymbol{\theta}$ which comes from the jth trial and the accompanying covariance matrix are then obtained by:

$$\hat{\boldsymbol{\theta}}^{(j)} = (L, 0) \begin{pmatrix} \hat{\boldsymbol{\beta}}^{(j)} \\ \hat{\boldsymbol{\delta}}_j \end{pmatrix} \quad \text{and} \quad \text{Cov}\left(\hat{\boldsymbol{\theta}}^{(j)}\right) = (L, 0)\mathcal{I}^{(j)^{-1}}(L, 0)', \tag{57}$$

where $\mathcal{I}^{(j)} = X_j' V_j^{-1} X_j$.The (information) matrix $\mathcal{I}^{(j)}$ can be partitioned as follows, if we use the partition of X_j, which corresponds to the primary and nuisance parameter provided in model (56):

$$\mathcal{I}^{(j)} = \begin{pmatrix} \mathcal{I}_{11}^{(j)} & \mathcal{I}_{12}^{(j)} \\ \mathcal{I}_{21}^{(j)} & \mathcal{I}_{22}^{(j)} \end{pmatrix}$$

Furthermore, letting

$$\mathcal{I}_{11\cdot 2}^{(j)} = \mathcal{I}_{11}^{(j)} - \mathcal{I}_{12}^{(j)} \mathcal{I}_{22}^{(j)^{-1}} \mathcal{I}_{21}^{(j)} \tag{58}$$

it follows that

$$\text{Cov}\left(\hat{\boldsymbol{\theta}}^{(j)}\right) = L\mathcal{I}_{11\cdot 2}^{(j)^{-1}} L' \tag{59}$$

$\hat{\boldsymbol{\theta}}^{(j)}$ is therefore the summary estimator of $\boldsymbol{\theta}$ from the jth trial, with covariance matrix given by (57) or (59).

2.3.3.2a The one-step IPD meta-analysis estimator

Suppose we synthesize the IPD $\mathbf{y}_1, \ldots, \mathbf{y}_k$ into a single linear model which will be utilized to estimate $\boldsymbol{\theta}$. Consequently, by letting $\mathbf{y} = (\mathbf{y}_1', \ldots, \mathbf{y}_k')'$, $W = (W_1', \ldots, W_k')'$, $Z = \text{diag}(Z_1, \ldots, Z_k)$, $\boldsymbol{\delta} = \left(\boldsymbol{\delta}_1', \ldots, \boldsymbol{\delta}_k'\right)'$ and $V = \text{diag}(V_1, \ldots, V_k)$, one can easily obtain the following linear model

$$\text{E}(\mathbf{y}) = W\boldsymbol{\beta} + Z\boldsymbol{\delta}, \quad \text{Cov}(\mathbf{y}) = V \tag{60}$$

Using the above model given by equation (60), the BLUE or MLE of $\boldsymbol{\theta}$ as well as the corresponding

covariance matrix are found to be given by

$$\hat{\boldsymbol{\theta}} = (L, 0)\mathcal{I}^{-1}X'V^{-1}\mathbf{y}, \quad \text{and} \quad \text{Cov}(\hat{\boldsymbol{\theta}}) = (L, 0)\mathcal{I}^{-1}(L, 0)', \tag{61}$$

where $\mathcal{I} = X'V^{-1}X$ and $X = (W, Z)$. Accordingly, the one-step IPD meta-analysis estimator of $\boldsymbol{\theta}$ is given by $\hat{\boldsymbol{\theta}}$.

If we manipulate the structure of the matrices, it becomes trivial that the covariance matrix in given in equation (61) obtains the following form

$$\text{Cov}(\hat{\boldsymbol{\theta}}) = L\left(\sum_{j=1}^{k}\mathcal{I}_{11\cdot 2}^{(j)}\right)^{-1}L' \tag{62}$$

with $\mathcal{I}_{11\cdot 2}^{(j)}$ as defined in equation (58).

2.3.3.2b The two-step IPD meta-analysis estimator

The two-step IPD meta-analysis estimator of $\boldsymbol{\theta}$, which is denoted by $\tilde{\boldsymbol{\theta}}$, is based on the estimators $\hat{\boldsymbol{\theta}}^{(j)}$ given in equation (57), as well as their covariance matrices given in equation (59). The associated covariance matrix and the two-step IPD meta-analysis estimator of $\boldsymbol{\theta}$, are shown as:

$$\tilde{\theta} = \left(\sum_{j=1}^{k}\text{Cov}\left(\hat{\boldsymbol{\theta}}^{(j)}\right)^{-1}\right)^{-1}\sum_{j=1}^{k}\text{Cov}\left(\hat{\boldsymbol{\theta}}^{(j)}\right)^{-1}\hat{\boldsymbol{\theta}}^{(j)}$$

and

$$\text{Cov}(\tilde{\boldsymbol{\theta}}) = \left(\sum_{j=1}^{k}\text{Cov}\left(\hat{\boldsymbol{\theta}}^{(j)}\right)^{-1}\right)^{-1} = \left(\sum_{j=1}^{k}\left(L\mathcal{I}_{11\cdot 2}^{(j)\,-1}L'\right)^{-1}\right)^{-1} \tag{63}$$

The aforementioned equations can be obtained by combining the estimators from each trial in a weighted manner.

The question that has to be answered is which linear model leads to a coincident value for the two estimators $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$. In the sections that follow, conditions for this equality are obtained. It is also addressed how to measure the precision loss caused by employing the two-step IPD meta-analysis estimator $\tilde{\boldsymbol{\theta}}$ in models where the two estimators are different. Special cases of this comparson, namely those done by Olkin and Sampson (Olkin and Sampson, (1998)[6]) and by Mathew and Nordström (Mathew and Nordström, (1999)[4]) have already been delt with and discussed.

2.3.3.3 The Meta-Analysis Result

Because $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$ are both linear unbiased estimators of $\boldsymbol{\theta}$, with $\hat{\boldsymbol{\theta}}$ the BLUE of $\boldsymbol{\theta}$, the corresponding covariance matrices given in equations (62) and (63) need to satisfy the following inequality:

$$\left(\sum_{j=1}^{k}\left(L{\mathcal{I}_{11\cdot 2}^{(j)}}^{-1}L'\right)^{-1}\right)^{-1} \geq L\left(\sum_{j=1}^{k}\mathcal{I}_{11\cdot 2}^{(j)}\right)^{-1}L' \tag{64}$$

where $A \geq B$ indicates that the matrix $A - B$ is nonnegative definite for the matrices A and B. We are therefore able to conclude that in order for the two estimators $\hat{\boldsymbol{\theta}}$ and $\tilde{\theta}$ to coincide, the covariance matrices for them must be equal. The next result provides a precise condition under which this equality will hold.

Result:

The two estimators $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$ coincide if and only if

1. the matrices $\left(L{\mathcal{I}_{11\cdot 2}^{(j)}}^{-1}L'\right)^{-1}L{\mathcal{I}_{11\cdot 2}^{(j)}}^{-1}$ are equal for every $j = 1, \ldots, k$, or equivalently,
2. the matrices $\left(L_1\mathcal{I}_{11\cdot 2}^{(j)}L_1'\right)^{-1}L_1\mathcal{I}_{11\cdot 2}^{(j)}L'$ are equal for every $j = 1, \ldots, k$, where L_1 represents a $(p-s)\times p$ matrix of rank $(p-s)$ satisfying $L_1L' = 0$.

The evidence for this finding is provided as such.

Proof. For the purpose of notational convenience, we write

$$\mathcal{I}_{11\cdot 2}^{(j)} = A_j, \quad j = 1, \ldots, k \tag{65}$$

The A_j's are thus $p \times p$ matrices. Remember that L is a $s \times p$ matrix with rank s. Given that both $\hat{\theta}$ and $\tilde{\theta}$ are unbiased estimators of $\boldsymbol{\theta}$, with $\hat{\boldsymbol{\theta}}$ being the BLUE, it follows that $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$ will agree if and only if the accompanying covariance matrices are identical. Therefore, we must demonstrate that

$$\left(\sum_{j=1}^{k}\left(LA_j^{-1}L'\right)^{-1}\right)^{-1} = L\left(\sum_{j=1}^{k}A_j\right)^{-1}L' \tag{66}$$

This is only true if and only if L_1 is a $(p-s)\times p$ matrix that has a rank of $p-s$ satisfying $L_1L' = 0$ and there is equality of the matrices $\left(LA_j^{-1}L'\right)^{-1}LA_j^{-1}$. Alternatively, this is also true if and only if equality exists between matrices $\left(L_1A_jL_1'\right)^{-1}L_1A_jL'$ for $j = 1, \ldots, k$.

First off, keep in mind that we may make the assumption that $LL' = Is$ without losing generality. This is inferred from the finding that, if equation (66) is valid for a given L, equation (66) is likewise valid

when L is substituted by $\tilde{L} = (LL')^{-1/2} L$ fulfilling $\tilde{L}\tilde{L}' = Is$. We shall therefore proceed by making the basic assumption that $LL' = Is$ from this point onward. Suppose L_1 is a $(p-s) \times p$ matrix of rank $p-s$, which satisfies $L_1 L' = 0$, such that $L_0' = (L', L_1')$ is a $p \times p$ orthogonal matrix. Define

$$B_j = L_0 A_j L_0' = \begin{pmatrix} LA_jL' & LA_jL_1' \\ L_1A_jL' & L_1A_jL_1' \end{pmatrix} = \begin{pmatrix} B_{11}^{(j)} & B_{12}^{(j)} \\ B_{21}^{(j)} & B_{22}^{(j)} \end{pmatrix} \tag{67}$$

for $j = 1, \ldots, k$.

Now we are able to re-express equation (66) as

$$\left(\sum_{j=1}^{k} \left(LL_0' B_j^{-1} L_0 L' \right)^{-1} \right)^{-1} = LL_0' \left(\sum_{j=1}^{k} B_j \right)^{-1} L_0 L' \ ,$$

or, equivalently, as

$$\left(\sum_{j=1}^{k} \left((I_s, 0)\, B_j^{-1}\, (I_s, 0)' \right)^{-1} \right)^{-1} = (I_s, 0) \left(\sum_{j=1}^{k} B_j \right)^{-1} (I_s, 0)' \tag{68}$$

Given the orthogonality of the matrix L_0, we exploited the fact that $LL_0' = (I_s, 0)$ to arrive at the outcome stated in equation (67). Given that equation (67)'s matrix B_j has a partitioned form, the terms in equation (68) may be represented as follows using the notation for a partitioned matrix's inverse:

$$(I_s, 0)\, B_j^{-1}\, (I_s, 0)' = \left(B_{11}^{(j)} - B_{12}^{(j)} B_{22}^{(j)^{-1}} B_{21}^{(j)} \right)^{-1}$$

$$(I_s, 0) \left(\sum_{j=1}^{k} B_j \right)^{-1} (I_s, 0)' = \left\{ \sum_{j=1}^{k} B_{11}^{(j)} - \left(\sum_{j=1}^{k} B_{12}^{(j)} \right) \left(\sum_{j=1}^{k} B_{22}^{(j)} \right)^{-1} \left(\sum_{j=1}^{k} B_{21}^{(j)} \right) \right\}^{-1}$$

Therefore, equation (68) can be written as

$$\sum_{j=1}^{k} \left(B_{11}^{(j)} - B_{12}^{(j)} B_{22}^{(j)^{-1}} B_{21}^{(j)} \right) = \sum_{j=1}^{k} B_{11}^{(j)} - \left(\sum_{j=1}^{k} B_{12}^{(j)} \right) \left(\sum_{j=1}^{k} B_{22}^{(j)} \right)^{-1} \left(\sum_{j=1}^{k} B_{21}^{(j)} \right)$$

or, equivalently, as

$$\sum_{j=1}^{k} B_{12}^{(j)} B_{22}^{(j)^{-1}} B_{21}^{(j)} = \left(\sum_{j=1}^{k} B_{12}^{(j)} \right) \left(\sum_{j=1}^{k} B_{22}^{(j)} \right)^{-1} \left(\sum_{j=1}^{k} B_{21}^{(j)} \right) \tag{69}$$

Thus, in order to obtain a condition under which equation (66) holds, we need to obtain a condition under which equation (69) holds, with $B_{11}^{(j)}, B_{12}^{(j)}, B_{21}^{(j)}$, and $B_{22}^{(j)}$, the blocks of the partitioned matrix

defined in (67).

Now using a result found in (Kiefer, (1959)[12]) (Lemma 3.2), which states that, given a set of scalars $\lambda_j, j = 1, \ldots, k$, satisfying $0 < \lambda_j < 1$ and $\sum_{j=1}^{k} \lambda_j = 1$, the difference

$$\sum_{j=1}^{k} \lambda_j B_{12}^{(j)} B_{22}^{(j)^{-1}} B_{21}^{(j)} - \left(\sum_{j=1}^{k} \lambda_j B_{12}^{(j)} \right) \left(\sum_{j=1}^{k} \lambda_j B_{22}^{(j)} \right)^{-1} \left(\sum_{j=1}^{k} \lambda_j B_{21}^{(j)} \right) \tag{70}$$

is a nonnegative definite matrix which is equal to the zero matrix if and only if the matrices $B_{22}^{(j)^{-1}} B_{21}^{(j)}$ are all equal for $j = 1, \ldots, k$. If we choose $\lambda_j = 1/k$, for $j = 1, \ldots, k$, we are therefore able to conclude that equation (69) holds if and only if the matrices $B_{22}^{(j)} B_{21}^{(j)}$ are all equal for $j = 1, \ldots, k$. By using the formulas in (67), we can show that this criterion is equivalent to the matrices $(L_1 A_j L_1')^{-1} L_1 A_j L'$ being equal for $j = 1, \ldots, k$. This is our primary result's second (2) condition.

Utilizing the observation the follows, we demonstrate that condition (2) is identical to condition (1):

$$L_1' \left(L_1 A_j L_1' \right)^{-1} L_1 = A_j^{-1} - A_j^{-1} L' \left(L A_j^{-1} L' \right)^{-1} L A_j^{-1}$$

Using the property $LL' = I$, the above identity implies

$$L_1' \left(L_1 A_j L_1' \right)^{-1} L_1 A_j L' = L' - A_j^{-1} L' \left(L A_j^{-1} L' \right)^{-1}$$

It now follows that if the matrices $(L_1 A_j L_1')^{-1} L_1 A_j L'$ are equal, then so are the matrices $\left(L A_j^{-1} L' \right)^{-1} L A_j^{-1}$ for $j = 1, \ldots, k$, and conversely. The evidence of the main finding is now complete.

It should be noted that the matrix inequality in equation (64), when L is a known matrix with full row rank and A is positive definite, may also be inferred from the concavity of the matrix function $g(A) = \left(L A^{-1} L' \right)^{-1}$. □

It's interesting to note that the equality equation in (64) holds if L is the identity matrix, which means that the condition (1) (or (2)) clearly holds if one is estimating the entire vector $\boldsymbol{\beta}$. As a result, the two estimators $\tilde{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}$ always coincide. However, there is no requirement for common nuisance parameters between trials. The aforementioned conditions may not always have an intuitive meaning.

The following is an illustration of a measure for the reduction in precision that occurs when the

one-step IPD meta-analysis estimator $\hat{\boldsymbol{\theta}}$ is substituted with the two-step IPD meta-analysis estimator $\tilde{\boldsymbol{\theta}}$.

$$\frac{1}{s}\,\mathrm{tr}\left\{\left(\sum_{j=1}^{k}\left(L\mathcal{I}_{11\cdot 2}^{(j)^{-1}}L'\right)^{-1}\right)^{-1} - L\left(\sum_{j=1}^{k}\mathcal{I}_{11\cdot 2}^{(j)}\right)^{-1}L'\right\} \tag{71}$$

We get the above quantity by simply dividing the trace of the difference between the matrices in equation (64) by the dimension of the matrices. Due to the fact that the difference between the matrices in equation (64), the difference always has is nonnegative definite and has a nonnegative trace. Therefore, the amount found in (71) is zero if and only if the estimators agree. We may infer that the accuracy loss when using $\tilde{\boldsymbol{\theta}}$ instead of $\hat{\boldsymbol{\theta}}$, increases with the magnitude of the quantity in (71), and vice versa. It's also crucial to remember that the matrices $\mathcal{I}_{11.2}^{(j)}$ must be accessible in order to compute the loss of precision using the criterion in (71). If all that is known about each trial's summary data is the estimate $\hat{\boldsymbol{\theta}}^{(1)}$ and the related covariance matrix $L\mathcal{I}_{11.2}^{(j)} - 1L'$, then it is not possible to calculate (71).

Other functional metrics, omitting a criteria based on the trace, can be used to evaluate the efficiency loss of the two-step IPD meta-analysis estimator $\tilde{\boldsymbol{\theta}}$. For instance, the determinants' ratio

$$\left|\sum_{j=1}^{k}\left(L\mathcal{I}_{11\cdot 2}^{(j)^{-1}}L'\right)^{-1}\right|^{-1/2} \div \left|L\left(\sum_{j=1}^{k}\mathcal{I}_{11\cdot 2}^{(j)}\right)^{-1}L'\right|^{1/2} \tag{72}$$

is a standard that could be useful in real-world situations. The ratio in (72) is really the ratio of the estimated volumes of the confidence areas for the $\boldsymbol{\theta}$, based on $\tilde{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}$, respectively under normality. Similar to this, one may examine the power functions of F-tests based on the one-step and two-step IPD meta-analysis estimators that are used to test linear hypotheses related to $\boldsymbol{\theta}$.

3 Meta-Analysis with Binary Data

With the conceptual introduction convered in section 2, we now come to a more detailed discussion on meta-analysis for a specific data type. In this section, we will be discussing numerous meta-analysis techniques which are used when dealing specifically with binary data.The ultimate goal of this paper is to advance earlier research done by Olkin, I. & Sampson, A. (Olkin and Sampson, (1998)[6]), Mathew, T. & Nordström, K. (Mathew and Nordström, (1999)[4]), and Mathew and Nordström (Mathew and Nordström, (2010)[5]), which delt with meta-analysis techniques for continuous data. We will start with an investigation into the underlying theory for meta-analysis with binary data, which will be used to extend existing works in an attempt to answer to following research question: *" Does a meta-analysis of summary statistics and a meta-analysis of IPD produce equivalent results, and if so, under which condi-*

tions is this true."

Meta-analysis of binary or binomial data is very common, specifically in the fields of health care, medical research and clinical sciences. Since we are considering binary data, it is critical to keep in mind that the number of successes is taken into account in a sequence of independent investigations, each of which measures success using probability p (Chen and Peace, (2013)[17]).

3.1 Meta-Analysis Methods

The primary objective of meta-analysis is to synthesize the estimates of treatment effect or effect sizes (ESs) across studies which are similar. In some cases the treatment effect estimates or the effect sizes are not provided, however, the number of patients who responded relative to all those who were evaluated for treatment and control is accessible. In such instances, for each study we will need to calculate the effect size and thereafter combine the effect sizes in order to evaluate the consistency of the effect across all studies as well as compute a summary effect. (Chen and Peace, (2013)[17])

The most popular ESs for binomial data include the risk ratio, risk-difference, and odds-ratio.

3.1.1 Analysis with Risk-Ratio

3.1.1.1 Definition

Suppose we have two treatments, one being Experimental (E) and the other being the Control (C). We then define the effect size for the risk-ratio (RR) of the two treatments as:

$$ES = \frac{p_E}{p_C} = \frac{x_E/n_E}{x_C/n_C} \tag{73}$$

where the risk probability p_E , for the experimental treatment (E), is equal to the sum of all occurrences (x_E) divided by all patients (n_E), i.e. $p_E = \frac{x_E}{n_E}$, similary for the control treatment (C). We can interpret the risk-ratio as follows (Chen and Peace, (2013)[17]):

- If the risk ratio is 1, It implies that there is no difference in risk between the two groups i.e. the experimental group (E) and the control group (C).
- A risk ratio > 1 implies a higher risk of that outcome in the experimental group (E).
- A risk ratio < 1 implies a lower risk in the experimental group (E).

The ES is first converted using the natural logarithm (i.e.ln) to provide an estimated confidence interval based on the normal distribution, and then employing the delta-method where $\ln ES = \ln(ES)$

(see appendix, section 6.2, pg 89-90). The variance for $\ln ES$ can be shown to be

$$\begin{aligned}\text{Var}(\ln ES) &= \text{Var}(\ln \frac{x_E/n_E}{x_C/n_C}) \\ &= \frac{1}{x_E} - \frac{1}{n_E} + \frac{1}{x_C} - \frac{1}{n_C}\end{aligned} \tag{74}$$

The above is achieved by representing $\ln ES$ as the linear terms in a Taylor series expansion about its expected value as:

$$\begin{aligned}\ln ES &= \ln ES\left(x_E, x_C\right) = \ln\left(p_E\right) - \ln\left(p_C\right) \\ &= \ln\left(x_E\right) - \ln\left(n_E\right) - \ln\left(x_C\right) + \ln\left(n_C\right) \\ &= \ln ES\left(\mu_{x_E}, \mu_{x_C}\right) + \left(\frac{\partial \ln ES}{\partial x_E}\right)_{\mu_E}\left(x_E - \mu_{x_E}\right) \\ &\quad + \left(\frac{\partial \ln ES}{\partial x_C}\right)_{\mu_C}\left(x_C - \mu_{x_C}\right)\end{aligned}$$

Then

$$\begin{aligned}\text{Var}(\ln ES) &= \left(\frac{\partial \ln ES}{\partial x_E}\right)^2_{\mu_E} \text{Var}\left(x_E\right) + \left(\frac{\partial \ln ES}{\partial x_C}\right)^2_{\mu_C} \text{Var}\left(x_C\right) \\ &= \left(\frac{1}{n_E p_E}\right)^2 n_E p_E\left(1 - p_E\right) + \left(\frac{1}{n_C p_C}\right)^2 n_C p_C\left(1 - p_C\right) \\ &= \frac{1}{n_E p_E}\left(1 - p_E\right) + \frac{1}{n_C p_C}\left(1 - p_C\right) \\ &= \frac{1}{n_E p_E}\left(1 - p_E\right) - \frac{1}{n_E} + \frac{1}{n_C p_C}\left(1 - p_C\right) - \frac{1}{n_C}\end{aligned}$$

which further simplifies to be equation (74). Consequently, the approximated standard error is

$$SE_{\ln ES} = \sqrt{\text{Var}(\ln ES)} \tag{75}$$

With this, the 95% CI for $\ln ES$ can be expressed as

$$\left(\ln ES - 1.96 \times SE_{\ln ES}, \ln ES + 1.96 \times SE_{\ln ES}\right) \tag{76}$$

The risk-ratio(RR) scale is then transformed back to its original form as follows:

$$\begin{aligned}RR &= \exp(\ln ES) \\ L_{RR} &= \exp\left(\ln ES - 1.96 \times SE_{\ln ES}\right) \\ U_{RR} &= \exp\left(\ln ES + 1.96 \times SE_{\ln ES}\right)\end{aligned}$$

3.1.1.2 The Risk-Ratio Meta-Analysis: Step-by-Step

In order to increase both precision and statistical power which are used to detect a statistically significant effect size, a meta-analysis of the effect sizes covering the individual studies needs to be conducted. (Chen and Peace, (2013)[17])

Precision in meta-analysis can be addressed in two ways:

1. The confidence interval.
2. The standard error of the ES.

According to the accuracy with which the ES has been estimated in each of the studies under consideration, the ES will be bounded by a confidence interval for each study. A study may have higher precision if its confidence interval is noticeably narrower than that of the other studies. Conversely, if the confidence interval for a study is noticeably wider than that for the other studies, it suggests that the study has little precision (Chen and Peace, (2013)[17]). The alternative method is to simply observe the size of the variances or standard errors, where the smallest variance indicates the greatest precision. Studies that have greater precision should have larger weights when combining studies. This is achieved by using a fixed-effects meta-analysis model and weighting each study by its inverse variance:

$$w_i = \frac{1}{Var_i} \tag{77}$$

where $i = 1, \ldots, K$. Studies that have a relatively good precision are allocated greater weights whereas studies that have relatively poor precision are allocated lesser weights. Since sample size is the driving factor of precision, studies can therefore be thought of as being weighted by sample size (Chen and Peace, (2013)[17]). We can then derive the fixed-effects weighted average meta-analysis model to combine the studies as:

$$\ln ES = \frac{\sum_{i=1}^{K} \ln ES_i \times w_i}{\sum_{i=1}^{K} w_i}$$

with variance calculated by:

$$\text{Var}(\ln ES) = \frac{1}{\sum_{i=1}^{K} w_i} \tag{78}$$

3.1.2 Analysis with Risk-Difference

3.1.2.1 Definition

In binomial data, although the risk-ratio is most frequently used, the risk difference is an ES, which is

easier to understand (Chen and Peace, (2013)[17]). The definition of risk difference (RD) is unambiguously the difference of the risks between two treatments, which is given as:

$$ES_{RD} = \hat{p}_E - \hat{p}_C = \frac{x_E}{n_E} - \frac{x_C}{n_C} \tag{79}$$

from the notations used in the previous section, equation (73). We can estimate the variance of ES_{RD} as:

$$\text{Var}\left(ES_{RD}\right) = \frac{\hat{p}_E\left(1-\hat{p}_E\right)}{n_E} + \frac{\hat{p}_C\left(1-\hat{p}_C\right)}{n_C} \tag{80}$$

and the standard error (SE) is then simply calculated as $SE_{ES_{RD}} = \sqrt{\text{Var}\left(ES_{RD}\right)}$. We can interpret the risk difference as follows:

- If the risk difference is 0, it implies that there is no difference in risk between the two groups i.e. the experimental group (E) and the control group (C).
- A risk ratio > 0 implies a lower risk in the control group (C).
- A risk ratio < 0 implies a lower risk in the experimental group (E).

3.1.3 Meta-Analysis with Odds-Ratio

3.1.3.1 Data Structure

An odds-ratio (OR) is a statistic that is employed to determine the degree to which two occurrences, A and B, are related. The odds-ratio is defined as the ratio between the odds of A in the presence of B and the odds of A without B, or as the ratio between the odds of B in the presence of A and the odds of B without A (Chen and Peace, (2013)[17]). We can interpret odds-ratios as follows:

- If and only if the OR equals 1, the two events are independent i.e. the probabilities of one occurrence are the same whether the other event occurs or not.
- Events A and B are correlated if the OR is larger than 1, meaning that the presence of B increases the chances of A relative to the absence of B, and the converse is also true: the presence of A increases the odds of B.
- If the OR is less than 1, then A and B have a negative correlation, meaning that having one event lowers the odds of having the other.

Using this defintion of an odds-ratio, we are able to construct a 2×2 table which is usually used for reporting the number of events and non-events in each of the two groups (Chen and Peace, (2013)[17]).

From a total of K studies, the data from the ith study $(i = 1, \ldots, K)$ can be represented as cells A_i, B_i, C_i, and D_i, as shown in Table 1 below.

Table 1: Nomenclature for 2×2 Table of Outcome by Treatment.

	Events	Non-Events	Total Event
Experimental	A_i	B_i	n_1
Control	C_i	D_i	n_2

3.1.3.1 Odds-Ratio: Woolf's Method

We define the odds-ratio (OR) that is associated with an event as the ratio of the odds of the event in one study group to the odds of the event in another study group (Chen and Peace, (2013)[17]). Therefore the treatment/experimental group's odds of the event in the ith study is given by

$$OddsE_i = \frac{p_{E_i}}{1 - p_{E_i}} = \frac{\frac{A_i}{A_i+B_i}}{1 - \frac{A_i}{A_i+B_i}} = \frac{\frac{A_i}{A_i+B_i}}{\frac{B_i}{A_i+B_i}} = \frac{A_i}{B_i} \tag{81}$$

In a similar way, the odds of the event for the control group in the ith study is given by

$$OddsC_i = \frac{p_{C_i}}{1 - p_{C_i}} = \frac{\frac{C_i}{C_i+D_i}}{1 - \frac{C_i}{C_i+D_i}} = \frac{\frac{C_i}{C_i+D_i}}{\frac{D_i}{C_i+D_i}} = \frac{C_i}{D_i} \tag{82}$$

Consequently, for the ith study, the odds-ratio (OR) of the treatment/experimental group to the control group is as follows:

$$OR_i = \frac{OddsE_i}{OddsC_i} = \frac{\frac{A_i}{B_i}}{\frac{C_i}{D_i}} = \frac{A_i D_i}{B_i C_i} \tag{83}$$

The odds-ratio compared to the RR or RD, may appear to be intuitively less appealing. Nevertheless the odds ratio is used extensively in numerous statistically sound methods of analysis, mainly in logistic regression (Chen and Peace, (2013)[17]).The odds-ratio is frequently employed as a measure of the effect size in analyzing categorical data in the form of 2×2 tables, including the meta-analyses of binomial data. It has been discovered that the odds ratio and risk ratio are quite similar when the risk of an event is low.(Chen and Peace, (2013)[17])

In general, odds-ratios are converted to the log scale in order to approximate the normal distribution using them. The log odds-ratio and its standard error are then estimated. The meta-analysis is then performed using these numbers. The results are then transformed back into the original metric. (Chen and Peace, (2013)[17])

The log odds-ratio is given as

$$\log OR = \ln(OR) \tag{84}$$

Using the delta method (see appendix, section 6.2, pg 89-90) we are able to derive an approximate variance from the definition, which is given as follows:

$$OR = \frac{AD}{BC} = \frac{\frac{\hat{P}_E}{1-\hat{P}_E}}{\frac{\hat{P}_C}{1-\hat{P}_C}} \tag{85}$$

and the log odds-ratio is

$$\begin{aligned} \log OR = \ln(OR) &= \ln\left(\frac{\hat{P}_E}{1-\hat{P}_E}\right) - \ln\left(\frac{\hat{P}_C}{1-\hat{P}_C}\right) \\ &= f\left(\hat{P}_E\right) - f\left(\hat{P}_C\right) \end{aligned} \tag{86}$$

where $f(x) = \ln\left(\frac{x}{1-x}\right)$ with derivative as $f'(x) = -\frac{1}{x(1-x)}$.

By expanding (through Taylor series) the log-odds for both the treatment/experimental and the control around their expected values, the approximate variance may be derived using the delta method as follows:

$$\begin{aligned} \log OR &= \ln\left(\frac{\hat{P}_E}{1-\hat{P}_E}\right) - \ln\left(\frac{\hat{P}_C}{1-\hat{P}_C}\right) \\ &\approx \left[f\left(P_E\right) + f'\left(P_E\right)\left(\hat{P}_E - P_E\right)\right] - \left[f\left(P_C\right) + f'\left(P_C\right)\left(\hat{P}_C - P_C\right)\right] \\ &= \left[f\left(P_E\right) - \frac{1}{P_E\left(1-P_E\right)}\left(\hat{P}_E - P_E\right)\right] \\ &\quad - \left[f\left(P_C\right) - \frac{1}{P_C\left(1-P_C\right)}\left(\hat{P}_C - P_C\right)\right] \end{aligned} \tag{87}$$

Now, we are able to obtain the variance as follows:

$$\begin{aligned}
\text{var}(\log OR) = & \left[\frac{1}{P_E\left(1-P_E\right)}\right]^2 \text{var}\left(\hat{P}_E - P_E\right) \\
& + \left[\frac{1}{P_C\left(1-P_C\right)}\right]^2 \text{var}\left(\hat{P}_C - P_C\right) \\
= & \left[\frac{1}{P_E\left(1-P_E\right)}\right]^2 \frac{P_E\left(1-P_E\right)}{n_E} \\
& + \left[\frac{1}{P_C\left(1-P_C\right)}\right]^2 \frac{P_C\left(1-P_C\right)}{n_C} \\
= & \frac{1}{n_E P_E\left(1-P_E\right)} + \frac{1}{n_C P_C\left(1-P_C\right)} \\
= & \frac{1}{n_E \frac{A}{n_E}\frac{B}{n_E}} + \frac{1}{n_C \frac{C}{n_C}\frac{D}{n_C}} \\
= & \frac{n_T}{AB} + \frac{n_C}{CD} \\
& = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}
\end{aligned} \tag{88}$$

Therefore the approximate standard error is:

$$SE_{logOR} = \sqrt{V_{logOR}} \tag{89}$$

We then transform these calculations in the log-scale back to the original scale for odds-ratios (OR) using the following:

$$OR = \exp(\log OR) \tag{90}$$

$$LL_{OR} = \exp\left(LL_{\log OR}\right) \tag{91}$$

$$UL_{OR} = \exp\left(LL_{\log OR}\right) \tag{92}$$

where LL and UL are the lower and upper limits of OR, respectively.

3.2 Summary Statistics Data Methods

3.2.1 Conventional Methods

The common treatment effect θ is estimated by a weighted average of the trial-specific estimates because the fixed effect model for meta-analysis assumes that the true treatment effects are homogeneous across all of the trials. The weights are calculated as the reciprocals of their within-trial variances (Turner et al., (2000)[8]). It is assumed by the conventional random effects model that the true treatment effects vary in a random way between the trials. The model includes a between-trial component of variance, τ^2, which is often estimated by non-iterative methods using a method of moments estimator. The random

effects model can be expressed as follows assuming normality of the observed and actual treatment effects

$$y_i \sim \mathrm{N}\left(\theta + v_i, \sigma_i^2\right)$$
$$v_i \sim \mathrm{N}\left(0, \tau^2\right) \tag{93}$$

where σ_i^2, normally assumed to be known, is the variance of the observed treatment effect y_i in the ith trial (Turner et al., (2000)[8]). A confidence interval for the average treatment effect θ may be calculated, assuming normality for y_i. The log-odds ratio is a frequently used measure of treatment effect in binary data, as was previously explained. Compared to alternative measures such as the risk differerence, the log-odds ratio makes it easier to satisfy the necessary normality assumption.(Turner et al., (2000)[8])

If there is evidence of heterogeneity across the trials, The confidence interval for the random effects estimate will be wider, and the random effects estimate of θ will not be equal to the fixed effect estimate. This difference between estimates for fixed effects and random effects, and the extent thereof, could be due to the presence of publication bias. (Turner et al., (2000)[8])

3.2.2 Multilevel Modelling Methods

When dealing with summary data, the random effects method may be used to derive a multilevel model. The model defined by (93) is viewed as a random effects regression model in this particular approach, where the observed log-odds ratios are viewed as a continuous outcome. The estimations of maximum likelihood (ML) and restricted maximum likelihood (REML) may then be easily produced. (Turner et al., (2000)[8])

The random effects methods used by ML, REML, and the traditional non-iterative method of moments estimation should all yield comparable estimates of θ (Turner et al., (2000)[8]). It should be noted that ML estimates of variance components such as τ^2 neglect to consider the use of the same data when estimating the fixed effects, so are in general downward biased. To get unbiased results, REML estimation employs a modified likelihood. In general, there is little difference between moment and REML estimates and the standard non-iterative method of moments estimator is likewise unbiased for τ^2 (Turner et al., (2000)[8]).

3.2.3 Calculation of Confidence Intervals

Confidence intervals can be computed using the likelihood-based and bootstrap methods for both θ and τ^2, when utilizing the multilevel models framework. If we use asymptotic standard errors, Wald confidence intervals are obtainable. If we consider Wald confidence intervals for θ, τ^2 is taken to be known

rather than estimated, thus the resultant intervals are actually quite narrow. The Wald approach's required normality assumption is almost definitely invalid when confidence intervals for τ^2 are obtained, with the exception of when there are a large number of trials. (Turner et al., (2000)[8])

Likelihood based intervals are preferable to Wald intervals for θ. This is because likelihood based intervals take into consideration the inaccuracy in estimation of τ^2. When a parameter's log-likelihood is abnormally shaped or severely skewed, likelihood based intervals are also favoured (Turner et al., (2000)[8]). Normality is however required by the likelihood approach if the range of values acquired is to be interpreted as a confidence interval in a stringent sense. Therefore, the complications in constructing confidence intervals for τ^2 still remain. The computation of the profile log-likelihood $l(\lambda)$ is necessary for the likelihood based method with a range of values of the parameter of interest, λ. This can also be applied within the framework of multilevel models (Turner et al., (2000)[8]). Likelihood ratio test statistics which are computed from the REML log-likelihood may not be strictly justifiable and therefore require some modification prior to use in testing of fixed effects. As a result, it's possible that the REML profile likelihood confidence intervals may not be constructed effectively for θ. However, REML likelihood ratio test statistics may be directly employed for testing of variance components, such as τ^2. The majority of multilevel software presently does not support the method for constructing intervals for fixed effects based on REML likelihood. Due to this, we only use ML to give profile likelihood confidence intervals. (Turner et al., (2000)[8])

Confidence intervals for θ and τ^2 can be employed by the use of a parametric bootstrap.This approach does not need a normality assumption for the estimate upon which the confidence interval is constructed, making it more practical for small sample sizes, specifically for τ^2 (Turner et al., (2000)[8]). Consider simulating a number of data sets using the initial model's distributional assumptions and that for each data set a bootstrap set of parameter estimates is generated. Using the smoothed percentiles of the bootstrap distributions, confidence intervals may then be based on the bootstrap replications of every relevant parameter.

Considering the differences between the widths of the three Wald confidence intervals for θ using standard random effects, ML and REML estimation may reflect the differences in the estimates of τ^2. For both θ and τ^2 the profile likelihood method produces the widest confidence intervals. The likelihood based interval for θ is moderately wider and somewhat asymmetric about the estimate compared to the ML Wald interval, however, compared to the equivalent Wald intervals, the ML and REML bootstrap intervals are marginally narrower (Turner et al., (2000)[8]). The increase in the confidence interval

width which resulted from using likelihood rather than Wald methods is more significant for τ^2 than for θ. The width of the bootstrap intervals for τ^2 is comparable to that of the Wald intervals, but the likelihood and bootstrap intervals both span ranges of values that are farther from zero than the two Wald intervals do. The forced symmetry about the estimate for τ^2 is unsuitable, nonetheless the Wald method performs adequately for θ. The normality assumption that is necessary for interpreting likelihood support intervals as approximative confidence intervals is relaxed by the bootstrap approach. In light of this, bootstrap intervals are preferred to likelihood intervals. The likelihood intervals and the bootstrap show a considerable discrepancy for τ^2 because this assumption is less sound for τ^2 than for θ. (Turner et al., (2000)[8])

3.3 Individual Partcipant-level Data Methods

3.3.1 Model Description

The fitting of logistic regression models with fixed trial effects, where the log-odds are permitted to vary over the n trials, is a method for meta-analyzing individual participant-level binary data (Turner et al., (2000)[8]). By virtue of a model that includes treatment group and trial effects represented by n dummy variables as explanatory variables, an analysis of the fixed effects may be carried out. If a random effects meta-analysis is to be performed that is analogous with the DerSimonian and Laird model, an additional term v_i is required which represents each trial's true treatment effect (i.e. log-odds ratio) deviation from the average:

$$\begin{aligned} \text{logit}\left(\pi_{ij}\right) &= \left(\theta + v_i\right) x_{ij} + \phi_i \\ v_i &\sim \text{N}\left(0, \tau^2\right) \end{aligned} \tag{94}$$

where π_{ij} is the true response probability for the jth individual in the ith trial, $x_{ij} = 0/1$ is a binary outcome that indicates their control/treatment group, and the set of fixed parameters ϕ_i indicates their trial membership (Turner et al., (2000)[8]).

When trial effects on the log-odds scale are viewed as random rather than fixed, this is known as the alternative random effects method. It is important that the appropriateness of both approaches are carefully considered. When fixed trial effects are taken into account, a nuisance parameter is estimated for each trial that is included, which reduces the amount of data available for estimating any one model parameter. If the trial sample sizes are small, this could result in inconsistent estimates (Turner et al., (2000)[8]). The true treatment effects from a set of trials may not be presumed to be randomly selected from an underlying population, contrary to competing arguments that imply modeling trial effects as

random is largely unsuitable. Conditioning on the marginals of the 2×2 tables as an alternative approach is rather challenging to implement.

3.3.2 Parameter Estimation

Standard logistic regression is all that is required for a fxed effect meta-analysis model for individual partcipant-level data, however more specialized techniques and software is necessary for a model that contains random treatment effects (Turner et al., (2000)[8]). Different iterative estimation techniques for a random effects logistic regression model entail the use of either first-order or second-order Taylor expansion approximations, as well as either marginal quasi-likelihood (MQL) or penalized quasi-likelihood (PQL). In general, the PQL precedes improved estimates of variance components in mixed models, whereas MQL facilitates model convergence more readily. However, even the best estimating method for PQL, when combined with second-order approximations, can result in estimates of the between-trial variance that are downwardly biased, particularly when the probability of events are high or the number of trials is small. (Turner et al., (2000)[8])

The use of a bias corrected parametric bootstrap procedure could remove bias from the estimate of τ^2. To assess the bias that exists in the initial estimators, the first step is to generate a set of bootstrap replications. These values are biased, thus the bootstrap bias estimation procedure is repeated using the new bias corrected estimates as a basis (Turner et al., (2000)[8]). On average, a better estimate of bias is produced by the second process than by the first. Once more, the bootstrap process is used once the original estimates have been adjusted using the second bias estimates. This iterative process between bias estimation and bias correction continues at that point. When convergence is reached, asymptotically consistent and unbiased estimates of the parameter are produced (Turner et al., (2000)[8]). Most statistical and programming software directly supports the bias corrected parametric bootstrap technique, but because it may be computationally demanding, it should only be employed in final analyses to provide unbiased estimates.

It is possible to convert published summary outcome data to individual binary data, which is then represented as a list of ones and zeros. Since individual level covariates are excluded, if the model given in (94) is fitted using ML and REML estimation, the results obtained by fitting model (94) to the grouped binomial data are identical. (Turner et al., (2000)[8])

If bias corrected estimate of τ^2 is larger than the standard REML estimate, this would suggest that the latter estimate is downwardly biased. When summary data methods are used, different estimates of

between-trial variance τ^2 may be expected because the quantities σ_i^2 from model (93) are assumed to be known. The individual data model (94) makes the assumption that the random treatment effects ν_i are normal. Examining the normal plot of the empirical Bayes estimates of ν_i might provide light on the validity of this assumption. (Turner et al., (2000)[8])

3.3.3 Calculation of Confidence Intervals

When using individual partcipant data methods, profile likelihood based intervals are not easily obtainable. This is due to the fact that most software cannot yet produce accurate log-likelihood values for multilevel models fitted to binary data (Turner et al., (2000)[8]). However, 95 percent confidence intervals may be created for both θ and τ^2 using quantiles of the bootstrap distributions that are produced using the parametric bootstrap with bias correction. To compute these intervals, it is recommended to utilize a method that involves a scaling parameter to modify the empirical quantiles. As previously mentioned, Wald intervals for θ may be too narrow and the intervals for τ^2 are likely to be invalid. (Turner et al., (2000)[8])

3.4 Individual Partcipant-level Data Methods with Random Trial Effects

3.4.1 Model Description

In the meta-analysis of the individual binary data, fixed trial effects were utilized in the preceding section to allow the log-odds to vary across trials. Random trial effects, as opposed to fixed trial effects, assume that the log-odds are, in fact, taken from a random distribution (Turner et al., (2000)[8]). Two effects are present in the model of random effects meta-analysis with random trial effects. These effects include the effects of the trial on the log-odds, u_i, as well as the trial's effects on the treatment effect, or v_i. The model is given as:

$$\begin{gathered} \operatorname{logit}\left(\pi_{ij}\right) = \alpha + u_i + \left(\theta + v_i\right) x_{ij} \\ u_i \sim \mathrm{N}\left(0, \sigma^2\right), \quad v_i \sim \mathrm{N}\left(0, \tau^2\right), \quad \operatorname{cov}\left(u_i, v_i\right) = \rho\sigma\tau \end{gathered} \tag{95}$$

As previously stated, $x_{ij} = 0/1$ is a binary outcome that indicates the control/treatment group. The covariance between u_i and v_i must be taken into account. The covariance between u_i and v_i must be taken into account. The variance-covariance matrix linked to the bivariate log-odds parameter is implicitly modeled when random treatment and trial effects are taken into account. The between-trial variance of the log-odds across the control groups is modelled by σ^2, whereas the across experimental groups is modelled by $\sigma^2+\tau^2$, if we consider the $cov(u_i, v_i)$ to be zero.(Turner et al., (2000)[8]). Because of this, the variance across trials for control groups must be lower or equal to the variance across trials for experimental groups.

However, this assumption could be viewed as being inappropriate. Furthermore, it is assumed that the variance of the log-odds in the control groups' between-trial variation and the covariance between control group and experimental group log-odds are both equal. The variance-covariance matrix of the bivariate log-odds parameter estimates will be freely modelled by combinations of the three parameters σ; τ and ρ if the $cov(u_i, v_i)$ is estimated rather than assumed to be zero. (Turner et al., (2000)[8])

3.4.2 Bivariate Approach

The establishment of model (95) requires both the assumption that the set of control group and experimental group log-odds follows a bivariate normal distribution and the complete likelihood for binomial data. The bivariate technique published by van Houwelingen et al. (van Houwelingen et al., 1993[14]) is represented in this model on many levels, where maximum likelihood estimates are derived by employing two approximation methods (of which the second was recommended) in addition to exact likelihood.

3.4.3 Coding of Treatment Covariates

Generally in practice, treatment covariates are coded as 0/1. However, given that estimation of the covariance between two random effects may be challenging or perhaps impossible when fitting a random effects meta-analysis model with random trial effects to data sets with limited degrees of freedom, it may be useful to code the treatment covariates as $\pm 1/2$ (Turner et al., (2000)[8]). The variance of the log-odds in the control group of patients is modeled as being equal to the variance of the log-odds in the experimental group of patients when $cov(ui, vi)$ is assumed to be zero and utilizing z$\pm 1/2$ coding. Additionally, the covariance between control group and experimental group log-odds is modelled separately. If it is not possible for convergence to be achieved when $cov(u_i, v_i)$ is included, the model may need to be simplified and a zero $cov(u_i, v_i)$ model with $\pm 1/2$ coding may be preferred to a zero $cov(u_i, v_i)$ model with the conventional 0/1 coding. (Turner et al., (2000)[8])

3.4.4 Model Checking

Model (94) requires the assumption that the random treatment effects v_1 will be normal, and model (95) additionally requires that the random trial effects u_i are assumed to be normally distributed. When non-normality arises, the individual data model (94) with fixed trial effects would be preferable to model (95). (Turner et al., (2000)[8])

4 Application

In this section we will be investigating two different clinical studies. The first study is called the *Cannon (Cannon et al., (2006)[13]) Cardiovascular Clinical Trials* which consists of four clinical studies. The second study is called the *Bacillus Calmette-Guerin Vaccine Study* which consists of a total of 13 clinical studies.

The purpose of investigating these two different clinical studies is to see how the results of a meta-analysis of summary data versus a meta-analysis of the IPD data comapre and/or differ if we have a small number of studies. We will be investigating each of the clinical studies separately.

4.1 Cannon Cardiovascular Clinical Trials

4.1.1 Data Description

The data used is from Cannon et al. (Cannon et al., (2006)[13]), which is a meta-analysis study of cardiovascular clinical trials that is specifically designed to establish a comparison between moderate statin therapy and intensive statin therapy in reducing cardiovascular outcomes. It has been found in recent clinical studies that treatments involving intensive statin therapy (also commonly known as high-dose statins) are more effective in reducing cardiovascular events as opposed to moderate statin therapy (also known as standard-dose statins). This can be seen in the following clinical studies:

1. PROVE IT-TIMI-22 (Pravastatin or Atorvastatin Evaluation and Infection Therapy-Thrombolysis In Myocardial Infarction-22)

2. TNT (Treating to New Targets)

There are however two other clinical studies, namely:

1. A-to-Z (Aggrastat to Zocor)

2. IDEAL (Incremental Decrease in End Points Through Aggressive Lipid Lowering)

These two studies had shown non-statistically significant results regarding the benefits of the intensive statin therapy. This inevitably brought to light certain questions around the reliability of this observation. A meta-analysis of these four trials is therefore conducted to determine the clinical utility of high-dose versus standard dose statin therapy in a more accurate manner. The dataset is shown in Table 2 below:

Table 2: Data for Coronary Death or MI of Statin Use.

Study	*nhigh*	*evhigh*	*nstd*	*evstd*	*ntot*	*evtot*
Prove It	2099	147	2063	172	4162	319
A-to-Z	2265	205	2232	235	4497	440
TNT	4995	334	5006	418	10001	752
IDEAL	4439	411	4449	463	8888	874

Note the following:

- *nhigh* is the total number of patients randomized to high dose.
- *evhigh* is the number of patients with event (defined as the combined incidence of coronary death or non-fatal myocardial infarction (MI)) in the high dose group.
- *nstd* is the total number of patients randomized to standard dose.
- *evstd* is the number of patients with event in the standard dose group.
- *ntot* is the total number of patients in both treatment groups.
- *evtot* is the total number of patients with events in both treatment groups

All coding is done using RStudio software, Version 3.6.3 (2020-02-29), with the main packages used:

- meta (Guido Schwarzer (2021)) : A mainstream package consisting of usual methods and techniques for meta-analysis such as: meta-analysis for both fixed and random effects; numerous plots including forest and funnel; statistical tests and trim-and-fill method used to evaluate bias in meta-analysis; meta-regression techniques; generalised linear mixed models etc.
- rmeta (Thomas Lumley (2018)) : A package for cumulative meta-analyses and two-sample comparisons using functions for straightforward fixed effects and random effects meta-analysis. Furthermore, the package has the functionality to enable the user to draw funnel plots, common summary plots and also tests for association and heterogeneity.
- metafor (Wolfgang Viechtbauer (2021)) : A powerful package for conducting meta-analyses consisting of an extensive all-inclusive assembly of functions. Included in the package are functions used to: calcualte effect sizes; fit fixed, random and mixed effects models; conduct analyses for meta-regression; and generate numerous plots for meta-analysis. Specifically for meta-analyses of binary data, functions for specialized methods are also available, including Peto's method, Mantel-Haenszel method and multiple generalized linear (mixed-effects) models. Over and above this, the package also provides functionality for fitting meta-analytic multivariate/multilevel models which are used to account for non-independent sampling errors and/or true effects. Correlation structures may also be conducted through network meta-analyses.

- lme4 (Douglas Bates (2021)): Used to fit linear as well as generalized linear mixed-effects models. S4 classes and methods are used to represent the models and their corresponding components, where the 'Eigen' C++ library for numerical linear algebra and 'RcppEigen' "glue" are used to implement the fundamental computational algorithms.

Note that the complete R-code is provided in the appendix, however certain sections of the code are provided in the following sections for reference purposes.

4.1.2 Meta-analysis of Summary Data

As described in section 3.1.3.1, we can construct a 2×2 table which represents the number of events and non-events in each of the two groups i.e. the treatment group and the control group. This table for the data on cardiovascular clinical trials is given as:

Table 3: A 2×2 representation of the data for Coronary Death or MI of Statin use for all studies.

Study	*tpos*	*tneg*	*cpos*	*cneg*
Prove It	147	1952	172	1891
A-to-Z	205	2060	235	1997
TNT	334	4661	418	4588
IDEAL	411	4028	463	3986

A meta-analysis of this 2×2 table data can be done based on one of numerous outcome measures, namely:

- Odds-ratio.
- Relative risk, also known as the risk ratio.
- Risk-difference.
- The arcsine transformed risk difference.

The above measures are primarily based on the sampling method used. It is also important to note that we take the log of the relative risk and the odds ratio, as this allows the outcome measures to have a symmetric shape around 0 which results in these outcome measure having a distribution that is closer to the normal distribution.

4.1.2.1 Odds-Ratio

In this section we consider the (log) odds-ratio as the outcome measure. We can obtain the values of the outcome measure y_i as well as the corresponding sampling variances v_i using the follwoing R-code:

```
datOR = escalc(measure = "OR", ai = tpos, bi = tneg, ci = cpos, di = cneg,
               data = datNew, append = TRUE)
datOR
```

The results obtained using the above code are given in the output below:

```
          tpos tneg cpos cneg      yi     vi
Prove_It   147 1952  172 1891 -0.1888 0.0137
A_to_Z     205 2060  235 1997 -0.1676 0.0101
TNT        334 4661  418 4588 -0.2401 0.0058
IDEAL      411 4028  463 3986 -0.1296 0.0051
```

Figure 2: Odds-ratios and corresponding sampling variances for the studies.

A forest plot of the odds-ratios can be seen in Figure 3 below, with the corresponding funnel plot shown in Figure 4.

```
y2=metabin(data=dat, evhigh, nhigh, evstd, nstd, studlab = Study, sm="OR")


forest(y2, is.summary=c(TRUE,TRUE,rep(FALSE,8),TRUE),
       leftcols = c("studlab", "evhigh", "nhigh", "evstd", "nstd"),
       rightcols = c("effect","ci", "w.random", "w.fixed"),
       col.square = "red", col.diamond = "blue",
       print.tau2 = F, print.Q = T, print.I2 = T)

funnel(y2)
```

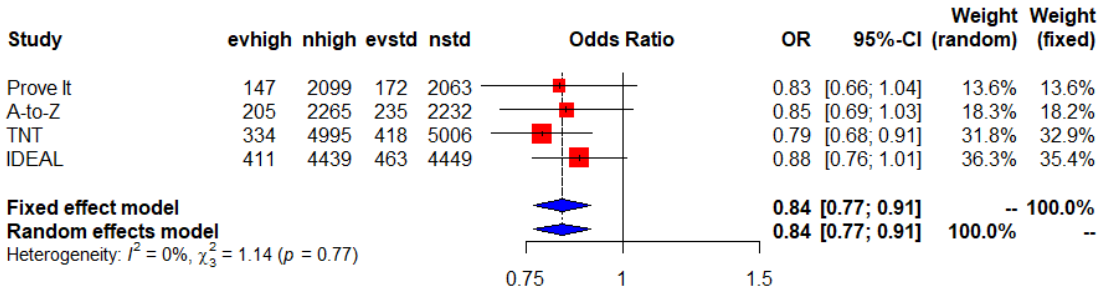


Figure 3: Forest plot of the odds-ratios for the studies

From Figure 3 above, we can see the odds-ratios for each of the four studies and their corresponding 95% confidence intervals, where the point estimates are represented by the red boxes and the horizonal lines represent the 95% confidence intervals. The red box that represents the point estimate becomes larger and the horizontal line shrinks as the study's size, or the number of participants, increases. As a result, the 95% confidence intervals for those studies should have a significantly narrower range, making it less probable that they will go beyond the line of null effect (in this example, the null effect is 1). The red box that represents the point estimate is smaller and the horizontal line is wider in smaller studies. Thus, because the 95% confidence intervals will be substantially larger, it follows that it is more probable that those studies will cross the line of null effect.

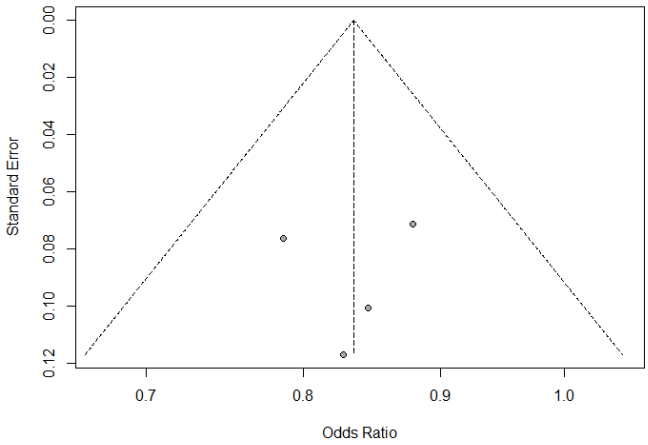


Figure 4: Funnel plot of the odds-ratios for the studies

Using equation (81), (82) and (83) we calculate the (log) odds-ratios for each of the studies as

```
#The odds of MI in the experimental high-dose group is
oddsE = pE/(1-pE)
oddsE

#The odds of MI in the standard low-dose group is
oddsC = pC/(1-pC)
oddsC

#The OR would then be
OR = oddsE/oddsC
OR

#The log odds ratios are
LogOR=log(OR)
LogOR
```

where

- (log) odds-ratios for *Prove It*: -0.1888106
- (log) odds-ratios for *A to Z*: -0.1676355
- (log) odds-ratios for *TNT*: -0.2401263
- (log) odds-ratios for *IDEAL*: -0.1296156

As seen in section 3.1.3, the delta method can be used to find the approxiamte varaince and approximate standard error for each study by expanding (via Taylor series) the log-odds for both treatment and control about their expected values.

Study: TNT

```
#The approximate variance can be calculated as
VLogOR1 = 1/334 + 1/(4995-334) + 1/418 + 1/(5006-418)
VLogOR1

#The standard error
SE.LogOR1 = sqrt(VLogOR1)
SE.LogOR1
```

The approximate variance for *TNT* is 0.005818863 and the corresponding approximate standard error is 0.07628147.

Study: Prove IT

```
#The approximate variance can be calculated as
VLogOR2 = 1/147 + 1/(2099-147) + 1/172 + 1/(2063-172)
VLogOR2


#The standard error
SE.LogOR2 = sqrt(VLogOR2)
SE.LogOR2
```

The approximate variance for *Prove It* is 0.01365779 and the corresponding approximate standard error is 0.1168665.

Study: A-to-Z

```
#The approximate variance can be calculated as
VLogOR3 = 1/205 + 1/(2265-205) + 1/235 + 1/(2232-235)
VLogOR3


#The standard error
SE.LogOR3 = sqrt(VLogOR3)
SE.LogOR3
```

The approximate variance for *A to Z* is 0.01011956 and the corresponding approximate standard error is 0.100596.

Study: IDEAL

```
#The approximate variance can be calculated as
VLogOR4 = 1/411 + 1/(4439-411) + 1/463 + 1/(4449-463)
VLogOR4


#The standard error
SE.LogOR4 = sqrt(VLogOR4)
SE.LogOR4
```

The approximate variance for *IDEAL* is 0.005092057 and the corresponding approximate standard error is 0.07135865.

4.1.2.1a) The Fixed-Effects Meta-Analysis Model

We can calculate the fixed-effects meta-analysis model using the following R-code:

```
# FE
meta.FE = rma(yi, vi, data = dat, method="FE")
# Print the summary
meta.FE
```

The results are given as:

```
Fixed-Effects Model (k = 4)

I^2 (total heterogeneity / total variability):   0.00%
H^2 (total variability / sampling variability):  0.38

Test for Heterogeneity:
Q(df = 3) = 1.1406, p-val = 0.7673

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.1797  0.0430  -4.1779  <.0001  -0.2641  -0.0954  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It is observed from the above model that the overall odds-ratio is -0.1797 with a 95% confidence interval of $(-0.2641, -0.0954)$, indicating a statistically significant overall effect that the high-dose treatment is more effective than the standard-dose treatment. In addition, Cochran's Q test for heterogeneity yielded a p-value of 0.7673 indicating non-statistically significant heterogeneity.

4.1.2.1b) The Random-Effects Meta-Analysis Model: DerSimonian-Laird

The random-effects meta-analysis model, using the DerSimonian-Laird method, can be calculated using the following R-code:

```
# RE-DL
meta.DL = rma(yi, vi, data = dat, method="DL")
# Print the summary
meta.DL
```

The results are given as:

```
Random-Effects Model (k = 4; tau^2 estimator: DL)

tau^2 (estimated amount of total heterogeneity): 0 (SE = 0.0063)
tau (square root of estimated tau^2 value):      0
I^2 (total heterogeneity / total variability):   0.00%
H^2 (total variability / sampling variability):  1.00

Test for Heterogeneity:
Q(df = 3) = 1.1406, p-val = 0.7673

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.1797  0.0430  -4.1779  <.0001  -0.2641  -0.0954  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With degrees of freedom $df = 3$, we obtain the statistical significance i.e. p-value as 0.7673 and we can therefore conclude that there is no statistically significant heterogeneity among these 4 studies. In addition, because the estimated $Q = 1.1406$ is less than the df of 3, we can conclude that random-effects meta-analysis is identical to the fixed-effects model from τ^2 in equation (18) such that the overall odds-ratio is -0.1797 with a 95% confidence interval of $(-0.2641, -0.0954)$, indicating a statistically significant overall effect that the high-dose treatment is more effective than the standard-dose treatment.

4.1.2.1c) The Random-Effects Meta-Analysis Model: Restricted Maximum Likelihood

The random-effects meta-analysis model, using the restricted maximum likelihood method, can be calculated using the following R-code:

```
# RE-REML
meta.REML = rma(yi, vi, method="REML",data = dat)
meta.REML
```

The results are given as:

```
Random-Effects Model (k = 4; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0 (SE = 0.0060)
tau (square root of estimated tau^2 value):      0
I^2 (total heterogeneity / total variability):   0.00%
H^2 (total variability / sampling variability):  1.00

Test for Heterogeneity:
Q(df = 3) = 1.1406, p-val = 0.7673

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.1797  0.0430  -4.1779  <.0001  -0.2641  -0.0954  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With degrees of freedom $df = 3$, we obtain the statistical significance i.e. p-value as 0.7673 and we can therefore conclude that there is no statistically significant heterogeneity among these 4 studies. In

addition, because the estimated $Q = 1.1406$ is less than the df of 3, we can conclude that random-effects meta-analysis is identical to the fixed-effects model from τ^2 in equation (18) such that the overall odds-ratio is -0.1797 with a 95% confidence interval of $(-0.2641, -0.0954)$, indicating a statistically significant overall effect that the high-dose treatment is more effective than the standard-dose treatment. This is also identical to the random-effects meta-analysis model, using the DerSimonian-Laird method.

4.1.2.1d) The Random-Effects Meta-Analysis Model: Maximum Likelihood

The random-effects meta-analysis model, using the maximum likelihood method, can be calculated using the following R-code:

```
# RE-ML
meta.ML = rma(yi, vi, method="ML",data = dat)
meta.ML
```

The results are given as:

```
Random-Effects Model (k = 4; tau^2 estimator: ML)

tau^2 (estimated amount of total heterogeneity): 0 (SE = 0.0049)
tau (square root of estimated tau^2 value):      0
I^2 (total heterogeneity / total variability):   0.00%
H^2 (total variability / sampling variability):  1.00

Test for Heterogeneity:
Q(df = 3) = 1.1406, p-val = 0.7673

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.1797  0.0430  -4.1779  <.0001  -0.2641  -0.0954  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With degrees of freedom $df = 3$, we obtain the statistical significance i.e. p-value as 0.7673 and we can therefore conclude that there is no statistically significant heterogeneity among these 4 studies. In addition, because the estimated $Q = 1.1406$ is less than the df of 3, we can conclude that random-effects meta-analysis is identical to the fixed-effects model from τ^2 in equation (18) such that the overall odds-ratio is -0.1797 with a 95% confidence interval of $(-0.2641, -0.0954)$, indicating a statistically significant overall effect that the high-dose treatment is more effective than the standard-dose treatment. This is also identical to the random-effects meta-analysis model, using the DerSimonian-Laird method and to the random-effects meta-analysis model, using the restricted maximum likelihood method.

4.1.2.2 Risk Ratio

In this section we consider the (log) risk ratio as the outcome measure. We can obtain the values of the outcome measure y_i as well as the corresponding sampling variances v_i using the follwoing R-code:

```
datRR = escalc(measure = "RR", ai = tpos, bi = tneg, ci = cpos, di = cneg,
               data = datNew, append = TRUE)
datRR
```

The results obtained using the above code are given in the output below:

```
          tpos tneg cpos cneg      yi     vi
Prove_It   147 1952  172 1891 -0.1744 0.0117
A_to_Z     205 2060  235 1997 -0.1513 0.0082
TNT        334 4661  418 4588 -0.2221 0.0050
IDEAL      411 4028  463 3986 -0.1169 0.0041
```

Figure 5: Risk ratio and corresponding sampling variances for the studies.

We can interpret the results as (log) risk ratios that have a value less than 0 indicate a more reduced risk of cardiovascular outcomes for the high-dose statin therapy group (treatment group). Therefore, since all of the studies had a risk ratio of less than 0, this means that all of the studies resulted in reduced risk of cardiovascular outcomes for the high-dose statin therapy group (treatment group) and consequently the experimental high-dose treatment is descriptively better than the standard treatment.

A forest plot of the risk ratios can be seen in Figure 6 below, with the corresponding funnel plot shown in Figure 7.

```
y1=metabin(data=dat, evhigh, nhigh, evstd, nstd, studlab = Study, sm="RR")

forest(y1, is.summary=c(TRUE,TRUE,rep(FALSE,8),TRUE),
       leftcols = c("studlab", "evhigh", "nhigh", "evstd", "nstd"),
       rightcols = c("effect","ci", "w.random", "w.fixed"),
       col.square = "red", col.diamond = "blue", print.tau2 = F,
       print.Q = T, print.I2 = T)

funnel(y1)
```

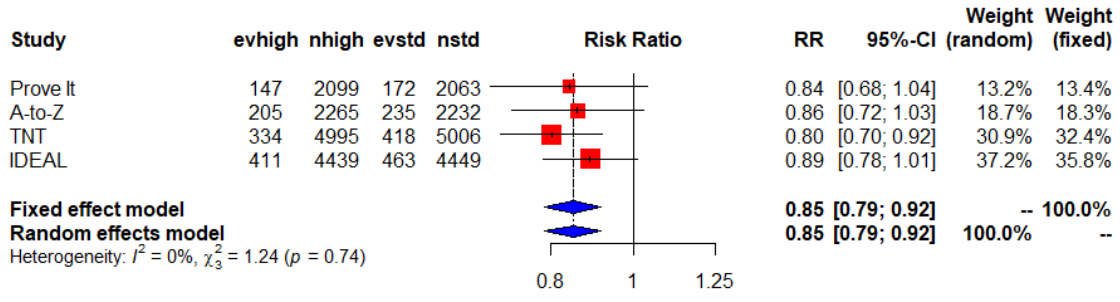


Figure 6: Forest plot of the risk ratios for the studies

From Figure 6 above, we can see the risk ratios for each of the four studies and their corresponding 95% confidence intervals, where the point estimates are represented by the red boxes and the horizonal lines represent the 95% confidence intervals. Since all the risk ratios are less than 1, this means that risk

is lower in the high-dose statin therapy group (treatment group). This corresponds with the results given in Figure 5.

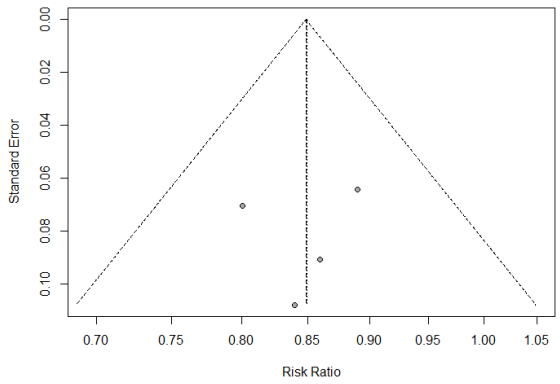


Figure 7: Funnel plot of the risk ratios for the studies

To test whether the ESs are statistically significant or not, we investigate the confidence intervals as well as the p-values for each of the studies. Since we have the (log) risk ratios, the normal approximation can be used, enabling us construct confidence interval of the (log) risk ratios for each of the studies. This is done using the following R-code:

```
# calculate the log risk-ratio
lnES = log(ES)
# print the log risk-ratio
round(lnES,digits=3)

# calculate the variance
vlnES = 1/dat$evhigh - 1/dat$nhigh +1/dat$evstd - 1/dat$nstd
# print the variance
round(vlnES,digits=5)

#We now can construct the 95% CI for lnES as follows:
# upper CI
ciup = lnES + 1.96*sqrt(vlnES)
round(ciup,digits=5)

# lower CI
cilow = lnES - 1.96*sqrt(vlnES)
round(cilow,digits=3)

#We then transform back to the original scale for risk- ratio(RR) as:
cat("The low CI is:", exp(cilow),"\n\n")
cat("The upper CI is:", exp(ciup),"\n\n")
```

The output for the above code is;

```
> cat("The low CI is:", exp(cilow),"\n\n")
The low CI is: 0.6797945 0.7194908 0.6972963 0.7842407

> cat("The upper CI is:", exp(ciup),"\n\n")
The upper CI is: 1.037944 1.027067 0.9196737 1.009315
```

This corresponds with the confidence interval values given by the Forest plot in Figure 6. Looking at the values for the 95% confidence intervals, we can see that study TNT (i.e. the third study in this case) is the only study that is statistically significant because the CI does not include the number one. The other three studies are all not statistically significant.

The p-values for each of the studies is calculated using the following R-code:

```
#We first calculate the z-values as follows:
z = lnES/sqrt(VlnES)
round(z, digits = 2)

#Then the p-values can be calculated as:
pval = 2*(1 - pnorm(abs(z)))
cat("p-values = ", pval,"\n\n")
```

Using the above code, the p-values for each of the studies are found to be:

- p-value for *Prove It* = 0.1063005
- p-value for *A to Z* = 0.09574236
- p-value for *TNT* = 0.001656186
- p-value for *IDEAL* = 0.06937868

Again we see that since study *TNT* is the only study that has a p-value less than 0.05, we conclude that the study is statistically significant at the 5% level.

4.1.2.2a) Risk Ratio Meta-Analysis

As previously mentioned in section 3.1.1, we can address precision in meta-analysis using one of two ways:

1. Confidence interval
2. The standard error of the ES

From Figure 6, we see that the confidence interval for the last two studies (*TNT* and *IDEAL*) are narrower compared to the first two studies (*ProveIt* and *AtoZ*). This implies that the last two studies have

greater precision.

If we consider the variance for the log risk ratio for each of the studies, they are calculated as follows:

```
round(vlnES, digits=5)
```

Using the above code, we obtain the following variances for each of the studies as:

- Variance of *Prove It* = 0.01166
- Variance of *A to Z* = 0.00824
- Variance of *TNT* = 0.00499
- Variance of *IDEAL* = 0.00414

Again, we can see the last two studies have the smallest variances which indicates g greater precision.

Using the fixed-effects meta-analysis model and weighting each study by its inverse variance as described by equation (77), we obtain the inverse weights for each study using the following R-code:

```
w = 1/vlnES
round(w, digits = 1)
```

The output for the above line of code is

```
85.8 121.3 200.5 241.4
```

The last two studies are assigned relatively large weights, while a lower weight is assigned to the first two studies. This implies that there is some relationship between precision and weight of the studies in a combined analysis, where greater weights are assigned to studies that have good precision and studies that have poor precision are assigned smaller weights.

4.1.2.2b) The Fixed-Effects Weighted Average Meta-Analysis Model

We can calculate the fixed-effects weighted average meta-analysis model to combine the studies using the following R-code:

```
#the inverse of variance for each study
fwi = 1/vlnES
round(fwi, digits=1)

# the total weight
fw = sum(fwi)
fw

# the weighted mean
flnES = sum(lnES*fwi)/fw
flnES

# the variance for the weighted mean
var = 1/fw
round(var, digits = 5)
```

The weighted mean is found to be -0.1634293 and the variance of the weighted mean is 0.00154. If we then tranform back to the original scale, we obtain the estimate for the risk ratio and the corresponding confidence interval as:

```
# the RR
exp(flnES)

# the lower and upper CI bounds
up = exp(flnES + 1.96*sqrt(var))
round(up, digits = 3)

low = exp(flnES - 1.96*sqrt(var))
round(low, digits = 3)
```

From the above code, we find that the estimate of the risk ratio is 0.8492265, with corresponding 95% confidence interval $(0.786, 0.917)$. Because the confidence interval is below the risk ratio of 1, this indicates that the high-dose treatment is statistically significantly more effective than the standard-dose treatment.

4.1.2.2c) The Random-Effects Meta-Analysis Model

The estimate of the heterogeneity statistic Q is calculated as follows:

```
# estimate heterogeneity statistic Q
Q = sum(fwi*lnES^2)-(sum(fwi*lnES))^2/fw
round(Q, digits=2)
```

where the value obtained is $Q = 1.24$.

```
# Get the degrees of freedom
K = dim(dat)[1]
K

df=K-1
df

# calculate the statistical significance: p-value
pval4Q = pchisq(Q, df, lower.tail = F)
pval4Q
```

With degrees of freedom $df = 3$, we obtain the statistical significance i.e. p-value as 0.7428365 and we can therefore conclude that there is no statistically significant heterogeneity among these 4 studies. In addition, because the estimated Q is less than the df of 3, we can conclude that random-effects meta-analysis is identical to the fixed-effects model from τ^2 in equation (18).

4.1.2.3 Risk-Difference

In this section we consider the risk difference as the outcome measure. We can obtain the values of the outcome measure y_i as well as the corresponding sampling variances v_i using the follwoing R-code:

```
datRD = escalc(measure = "RD", ai = tpos, bi = tneg, ci = cpos, di = cneg,
               data = datNew, append = TRUE)
datRD
```

The results obtained using the above code are given in the output below:

```
          tpos tneg cpos cneg      yi     vi
Prove_It   147 1952  172 1891 -0.0133 0.0001
A_to_Z     205 2060  235 1997 -0.0148 0.0001
TNT        334 4661  418 4588 -0.0166 0.0000
IDEAL      411 4028  463 3986 -0.0115 0.0000
```

Figure 8: Risk-difference and corresponding sampling variances for the studies.

The output can be interpreted as differences that are less than 0 indicate that the experimental treatment is descriptively better than the standard treatment. Therefore, all four studies resulted in reduced cardiovascular outcomes for the high-dose statin therapy group (treatment group) compared to the standard-dose statin therapy group (control group).

A forest plot of the risk-difference can be seen in Figure 9 below, with the corresponding funnel plot shown in Figure 10.

```
y3=metabin(data=dat, evhigh, nhigh, evstd, nstd, studlab = Study, sm="RD")

forest(y3, is.summary=c(TRUE,TRUE,rep(FALSE,8),TRUE),
       leftcols = c("studlab", "evhigh", "nhigh", "evstd", "nstd"),
       rightcols = c("effect","ci", "w.random", "w.fixed"),
       col.square = "red", col.diamond = "blue",
       print.tau2 = F, print.Q = T, print.I2 = T)

funnel(y3)
```

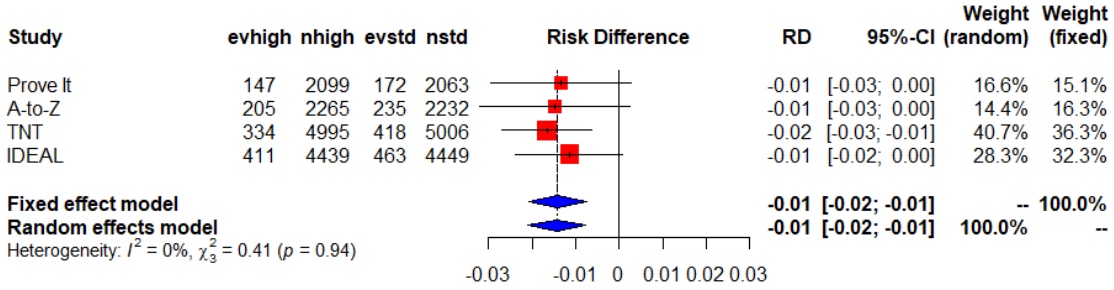


Figure 9: Forest plot of the risk-differences for the studies

From Figure 9 above, we can see the risk-differences for each of the four studies and their corresponding 95% confidence intervals, where the point estimates are represented by the red boxes and the horizonal lines represent the 95% confidence intervals.

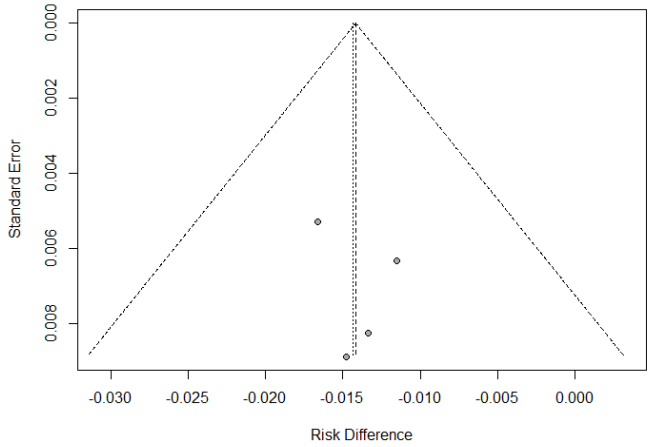


Figure 10: Funnel plot of the risk-differences for the studies

The risk-differences for each of the studies can be calculated using equation (79) as seen in section 3.1.2. This is done in the following R-code:

```
# the risk difference
ESRD = pE - pC
round(ESRD, digits = 4)
```

where the output is given as:

- Risk-difference for *Prove It* : -0.0133
- Risk-difference for *A to Z* : -0.0148
- Risk-difference for *TNT* : -0.0166
- Risk-difference for *IDEAL* : -0.0115

Since the risk-differences are less than 0 for all four studies, we are able to conclude that the high-dose statin treatment is descriptively better than the standard-dose statin treatment.

4.1.2.3a) Fixed-Effects and Random-Effects Meta-Analysis Model

Both the fixed-effects model and the random-effects model can be fitted using the following R-code:

```
RD.Statin = metabin(evhigh,nhigh,evstd,nstd,studlab=Study,
                    data=dat, method="Inverse", sm="RD")
# print the analysis summary
summary(RD.Statin)
```

The results are given as:

```
Number of studies combined: k = 4
Number of observations: o = 27548
Number of events: e = 2385

                          RD             95%-CI     z  p-value
Fixed effect model   -0.0144 [-0.0209; -0.0078] -4.27 < 0.0001
Random effects model -0.0144 [-0.0209; -0.0078] -4.27 < 0.0001

Quantifying heterogeneity:
 tau^2 = 0 [0.0000; 0.0001]; tau = 0 [0.0000; 0.0082]
 I^2 = 0.0% [0.0%; 84.7%]; H = 1.00 [1.00; 2.56]

Test of heterogeneity:
    Q d.f. p-value
 0.41    3  0.9379

Details on meta-analytical method:
- Inverse variance method
- DerSimonian-Laird estimator for tau^2
- Jackson method for confidence interval of tau^2 and tau
```

As expected, the results and conlcusions for the risk-difference (RD) are the same as the results and conlcusions that were reached in the meta-analysis of the risk ratio (RR) i.e. the high-dose treatment is statistically more effective than the standard dose treatment.

4.1.3 Meta-analysis of Individual Participant-Level Data

We now consider a meta-analysis of the individual participant-level data, specifically focusing on the odds-ratio as the outcome measure. As previously mentioned, the aim of this analysis is to investigate whether or not, in certain cases, a meta-analysis of individual participant-level data produces asymptotically equivalent results to a meta-analysis of summary data, in a binary setting.

Before we can do any analysis, the data needs to be individualized. This is done using the following R-code:

```
# number of Studies
K = 4
# reformat the SS data into "death" from "total"(=neg+pos), along with "study"+"Treatment"
SS  = data.frame(Deaths=rep(0, 2*K), Total=rep(0,2*K),Study=rep(0, 2*K), Treatment=rep(0,2*K))
# take Treatment: 0=std and 1=high statin
SS$Treatment = c(rep(0, K), rep(1, K))
SS$Study = rep(1:K,2)
# death = pos
SS$Deaths = c(dat$cpos, dat$tpos)
# total= pos+neg
SS$Total = c(dat$cpos+dat$cneg, dat$tpos+dat$tneg)
# check the data to see whether this is reproduced:
SS
# check with the original data
dat
```

```
# some indices for IPD, especially the number of patients since we need to
# create a row for each patients as o or 1
len.SS = dim(SS)[1];len.SS
num.patient = sum(SS$Total);num.patient

# make it as IPD with 4 columns
IPD = matrix(0,nrow=num.patient,ncol=4)
tt=1
for(i in 1:len.SS){
  # 1st col as ID of the patient
  IPD[tt:(tt-1+SS$Total[i]),1]=tt:(tt-1+SS$Total[i])
  # 2nd col as Death = 1 and 0 = live for IPD logistic regression
  IPD[tt:(tt-1+SS$Total[i]),2]=c(rep(1,SS$Deaths[i]),rep(0,SS$Total[i]-SS$Deaths[i]))
  # 3rd as Study index to track the which study to contain those patients
  IPD[tt:(tt-1+SS$Total[i]),3]=rep(SS$Study[i],SS$Total[i])
  # 4th as Treatment index to track the treatment for the patients at that study
  IPD[tt:(tt-1+SS$Total[i]),4]=rep(SS$Treatment[i],SS$Total[i])
  tt=tt+SS$Total[i]
  }  #End of i-loop
# make to data frame
IPD = data.frame(IPD)
names(IPD) = c("ID","TB","Trial","TRT")
summary(IPD)
# see how many patients. You see now we have an IPD data with 27,548 rows, with each row for a patient
dim(IPD)
sum(SS$Total)
# just check the first 100 patients from the 27,548 to see whether we did correct
head(IPD, 100)

# check with the SS:
aggregate(IPD$TB, list(IPD$Trial, IPD$TRT), sum)
```

The first 10 observations for the IPD data can be seen in Figure 11.

	ID	TB	Trial	TRT
1	1	1	1	0
2	2	1	1	0
3	3	1	1	0
4	4	1	1	0
5	5	1	1	0
6	6	1	1	0
7	7	1	1	0
8	8	1	1	0
9	9	1	1	0
10	10	1	1	0

Figure 11: First 10 observations of the IPD data

We now fit a multi-level logistic regression model to the IPD data using the following R-code:

```
# random-intercept
fitIPD2 = glmer(TB~TRT+(1|Trial), data=IPD,family=binomial)
sumIPD2 = summary(fitIPD2)
sumIPD2

meta.ML
meta.REML
```

Consider the output below that shows the results for the generalized linear mixed model fit, using maximum likelihood (Laplace Approximation).

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: TB ~ TRT + (1 | Trial)
   Data: IPD

     AIC      BIC   logLik deviance df.resid
 16185.5  16210.2  -8089.8  16179.5    27545

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.3432 -0.3137 -0.3041 -0.2737  3.6541

Random effects:
 Groups Name        Variance Std.Dev.
 Trial  (Intercept) 0.01784  0.1336
Number of obs: 27548, groups:  Trial, 4

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.27138    0.07341 -30.942  < 2e-16 ***
TRT         -0.17980    0.04300  -4.181  2.9e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
    (Intr)
TRT -0.272
```

We see that the IPD treatment effect is -0.17980 and the standard error for the treatment effect is 0.043. This is very close to the results given by the meta-analysis of summary data based on the odds-ratio outcome measure, using the maximum likelihood method as well as the restricted maximum likelihood method.

We can further see that the results are so close since for this data the τ^2 is estimated as zero, i.e. the random-effects meta-analysis is not needed. It is however important to note that in general, this is not true.

Now, let us consider the more general glm (generalized linear model) with estimated correlation of the random-effects. This is done using the following R-code:

```
fitIPD1 = glmer(TB~TRT+(TRT|Trial), data=IPD,family=binomial)
sumIPD1 = summary(fitIPD1)
sumIPD1
```

The results are given as follows:

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: TB ~ TRT + (TRT | Trial)
   Data: IPD

     AIC      BIC   logLik deviance df.resid
 16188.6  16229.7  -8089.3  16178.6    27543

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.3404 -0.3170 -0.3068 -0.2702  3.7011

Random effects:
 Groups Name        Variance Std.Dev. Corr
 Trial  (Intercept) 0.013167 0.11475
        TRT         0.001594 0.03992  1.00
Number of obs: 27548, groups:  Trial, 4

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.27016    0.06492 -34.968  < 2e-16 ***
TRT         -0.18273    0.04762  -3.837 0.000125 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
    (Intr)
TRT 0.098
optimizer (Nelder_Mead) convergence code: 0 (OK)
boundary (singular) fit: see ?isSingular
```

Here we see that the IPD treatment effect is -0.18273 and the standard error for the treatment effect is 0.04762. This is not close to the results obtained by the meta-analysis of summary data based on the odds-ratio outcome measure, using the maximum likelihood method as well as the restricted maximum likelihood method. The above result further implies that the IPD meta-analysis and the summary data meta-analysis can indeed be different.

4.2 Bacillus Calmette-Guerin Vaccine Study

4.2.1 Data Description

This dataset originates from clinical trials that were conducted in order to assess the impact of a Bacillus Calmette-Guerin (BCG) vaccine in the prevention of tuberculosis (TB). The dataset, which is well known and widely used, was first reported and published by Colditz et al. (Colditz et al., (1994)[15]). The dataset includes 13 clinical trials of the BCG vaccine, each investigating the effectiveness of BCG in the treatment of tuberculosis. The dataset is shown in Table 4 below:

Table 4: Data from Studies on Efficacy of BCG Vaccine for Preventing Tuberculosis.

author	*year*	*tpos*	*tneg*	*cpos*	*cneg*	*ablat*	*alloc*
Aronson	1948	4	119	11	128	44	random
Ferguson and Simes	1949	6	300	29	274	55	random
Rosenthal et al	1960	3	228	11	209	42	random
Hart and Sutherland	1977	62	13536	248	12619	52	random
Frimodt-Moller et al	1973	33	5036	47	5761	13	alternate
Stein and Aronson	1953	180	1361	372	1079	44	alternate
Vandiviere et al	1973	8	2537	10	619	19	random
TPT Madras	1980	505	87886	499	87892	13	random
Coetzee and Berjak	1968	29	7470	45	7232	27	random
Rosenthal et al	1961	17	1699	65	1600	42	systematic
Comstock et al	1974	186	50448	141	27197	18	systematic
Comstock and Webster	1969	5	2493	3	2338	33	systematic
Comstock et al	1976	27	16886	29	17825	33	systematic

Note the following:

- *author* identifies the 13 studies' authors.
- *year* is the year that these 13 studies were published.
- *tpos* is the quantity of TB positive cases among those who received the BCG vaccine.
- *tneg*is the quantity of TB-negative cases among those who received the BCG vaccine.
- *cpos* is the quantity of TB positive patients in the control group.
- *cneg* is the quantity of TB negative patients in the control group.
- *ablat* indicates the study location's exact latitude (in degrees).
- *alloc* indicates a three-level system of random, alternative, or systematic treatment assignment.

Since this dataset is significantly larger than the dataset used in section 4.1, we will only be considering the odds-ratio outcome measure for both the meta-analysis of summary data and the meta-analysis of individual participant-level data.

All coding is done using RStudio software, Version 3.6.3 (2020-02-29), with the same main packages used as defined in section 4.1.1. Note that the complete R-code is provided in the appendix, however certain sections of the code are provided in the following sections for reference purposes.

4.2.2 Meta-analysis of Summary Data

As described in section 3.1.3.1, we can construct a 2×2 table which represents the number of events and non-events in each of the two groups i.e. the treatment group and the control group. This table for the data on the BCG vaccine is given as:

Table 5: A 2×2 representation of the data for the BCG vaccine for preventing tuberculosis for all studies.

Study	*tpos*	*tneg*	*cpos*	*cneg*
Aronson	4	119	11	128
Ferguson and Simes	6	300	29	274
Rosenthal et al	3	228	11	209
Hart and Sutherland	62	13536	248	12619
Frimodt-Moller et al	33	5036	47	5761
Stein and Aronson	180	1361	372	1079
Vandiviere et al	8	2537	10	619
TPT Madras	505	87886	499	87892
Coetzee and Berjak	29	7470	45	7232
Rosenthal et al	17	1699	65	1600
Comstock et al	186	50448	141	27197
Comstock and Webster	5	2493	3	2338
Comstock et al	27	16886	29	17825

A meta-analysis of this 2×2 table data can be done based on numerous outcome measures. As previously mentioned, we will only be considering the odds-ratio outcome measure for this dataset.

The above mentioned outcome measure is primarily based on the sampling method used. It is also important to note that we take the log of the odds ratio, as this allows the outcome measures to have a symmetric shape around 0 which results in these outcome measure having a distribution that is closer to the normal distribution.

4.2.2.1 Odds-Ratio

In this section we consider the (log) odds-ratio as the outcome measure. We can obtain the values of the outcome measure y_i as well as the corresponding sampling variances v_i using the follwoing R-code:

```
dat  = escalc(measure="OR",ai=tpos,bi=tneg,ci=cpos,
              di = cneg, data = dat, append = TRUE)
dat
```

The results obtained using the above code are given in the output below:

```
   trial              author year tpos  tneg cpos  cneg ablat       alloc      yi     vi
       1             Aronson 1948    4   119   11   128    44      random -0.9387 0.3571
       2     Ferguson & Simes 1949    6   300   29   274    55      random -1.6662 0.2081
       3      Rosenthal et al 1960    3   228   11   209    42      random -1.3863 0.4334
       4    Hart & Sutherland 1977   62 13536  248 12619    52      random -1.4564 0.0203
       5 Frimodt-Moller et al 1973   33  5036   47  5761    13  alternate -0.2191 0.0520
       6      Stein & Aronson 1953  180  1361  372  1079    44  alternate -0.9581 0.0099
       7     Vandiviere et al 1973    8  2537   10   619    19      random -1.6338 0.2270
       8           TPT Madras 1980  505 87886  499 87892    13      random  0.0120 0.0040
       9     Coetzee & Berjak 1968   29  7470   45  7232    27      random -0.4717 0.0570
      10      Rosenthal et al 1961   17  1699   65  1600    42 systematic -1.4012 0.0754
      11       Comstock et al 1974  186 50448  141 27197    18 systematic -0.3408 0.0125
      12   Comstock & Webster 1969    5  2493    3  2338    33 systematic  0.4466 0.5342
      13       Comstock et al 1976   27 16886   29 17825    33 systematic -0.0173 0.0716
```

Figure 12: Odds-ratios and corresponding sampling variances for the studies for the BCG vaccine study.

4.2.2.1a) The Fixed-Effects Meta-Analysis Model

We can calculate the fixed-effects meta-analysis model using the following R-code:

```
# FE
meta.FE = rma(yi, vi, data = dat, method="FE")
# Print the summary
meta.FE
```

The results are given as:

```
Fixed-Effects Model (k = 13)

I^2 (total heterogeneity / total variability):   92.65%
H^2 (total variability / sampling variability):  13.60

Test for Heterogeneity:
Q(df = 12) = 163.1649, p-val < .0001

Model Results:

estimate      se      zval    pval    ci.lb    ci.ub
 -0.4361  0.0423  -10.3190  <.0001  -0.5190  -0.3533  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It is observed from the above model that the overall odds-ratio is -0.4361 with a 95% confidence interval of $(-0.5190, -0.3533)$, indicating a statistically significant overall effect that the BCG vaccine is indeed effective in the prevention of tuberculosis (TB). In addition, Cochran's Q test for heterogeneity yielded a p-value of < 0.0001 indicating statistically significant heterogeneity amongst the 13 studies.

4.2.2.1b) The Random-Effects Meta-Analysis Model: DerSimonian-Laird

The random-effects meta-analysis model, using the DerSimonian-Laird method, can be calculated using the following R-code:

```
# RE-DL
meta.DL = rma(yi, vi, data = dat, method="DL")
# Print the summary
meta.DL
```

The results are given as:

```
Random-Effects Model (k = 13; tau^2 estimator: DL)

tau^2 (estimated amount of total heterogeneity): 0.3663 (SE = 0.2659)
tau (square root of estimated tau^2 value):      0.6053
I^2 (total heterogeneity / total variability):   92.65%
H^2 (total variability / sampling variability):  13.60

Test for Heterogeneity:
Q(df = 12) = 163.1649, p-val < .0001

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.7474  0.1923  -3.8873  0.0001  -1.1242  -0.3706  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With degrees of freedom $df = 12$, we obtain the statistical significance i.e. p-value as < 0.0001 and we can therefore conclude that there is statistically significant heterogeneity among these 13 studies. In addition, because the estimated $Q = 163.1649$ is more than the df of 12, we can conclude that random-effects meta-analysis is not identical or even similar to the fixed-effects model from τ^2 in equation (18) such that the overall odds-ratio is -0.7474 with a 95% confidence interval of $(-1.1242, -0.3706)$, indicating a statistically significant overall effect that the BCG vaccine is indeed effective in the prevention of tuberculosis (TB).

4.2.2.1c) The Random-Effects Meta-Analysis Model: Restricted Maximum Likelihood

The random-effects meta-analysis model, using the restricted maximum likelihood method, can be calculated using the following R-code:

```
# RE-REML
meta.REML = rma(yi, vi, method="REML",data = dat)
meta.REML
```

The results are given as:

```
Random-Effects Model (k = 13; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0.3378 (SE = 0.1784)
tau (square root of estimated tau^2 value):      0.5812
I^2 (total heterogeneity / total variability):   92.07%
H^2 (total variability / sampling variability):  12.61

Test for Heterogeneity:
Q(df = 12) = 163.1649, p-val < .0001

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.7452  0.1860  -4.0057  <.0001  -1.1098  -0.3806  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With degrees of freedom $df = 12$, we obtain the statistical significance i.e. p-value as < 0.0001 and we can therefore conclude that there is statistically significant heterogeneity among these 13 studies. In

addition, because the estimated $Q = 163.1649$ is more than the *df* of 12, we can conclude that random-effects meta-analysis is not identical or similar to the fixed-effects model from τ^2 in equation (18) such that the overall odds-ratio is -0.7452 with a 95% confidence interval of $(-1.1098, -0.3806)$, indicating a statistically significant overall effect that the BCG vaccine is indeed effective in the prevention of tuberculosis (TB).

4.2.2.1d) The Random-Effects Meta-Analysis Model: Maximum Likelihood

The random-effects meta-analysis model, using the maximum likelihood method, can be calculated using the following R-code:

```
# RE-ML
meta.ML = rma(yi, vi, method="ML",data = dat)
meta.ML
```

The results are given as:

```
Random-Effects Model (k = 13; tau^2 estimator: ML)

tau^2 (estimated amount of total heterogeneity): 0.3025 (SE = 0.1549)
tau (square root of estimated tau^2 value):      0.5500
I^2 (total heterogeneity / total variability):   91.23%
H^2 (total variability / sampling variability):  11.40

Test for Heterogeneity:
Q(df = 12) = 163.1649, p-val < .0001

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.7420  0.1780  -4.1694  <.0001  -1.0907  -0.3932  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With degrees of freedom $df = 12$, we obtain the statistical significance i.e. p-value as < 0.0001 and we can therefore conclude that there is statistically significant heterogeneity among these 13 studies. In addition, because the estimated $Q = 163.1649$ is more than the *df* of 12, we can conclude that random-effects meta-analysis is not identical or similar to the fixed-effects model from τ^2 in equation (18) such that the overall odds-ratio is -0.7420 with a 95% confidence interval of $(-1.0907, -0.3932)$, indicating a statistically significant overall effect that the BCG vaccine is indeed effective in the prevention of tuberculosis (TB).

4.2.3 Meta-analysis of Individual Participant-Level Data

Again, we now consider a meta-analysis of the individual participant-level data for the BCG vaccine study, specifically focusing on the odds-ratio as the outcome measure. As previously mentioned, the aim of this analysis is to investigate whether or not, in certain cases, a meta-analysis of individual participant-level

data produces asymptotically equivalent results to a meta-analysis of summary data, in a binary setting.

Before we do any analysis, the data needs to be individualized. This is done using the following R-code:

```
# number of Studies
K = 13
# reformat the SS data into "death" from "total"(=neg+pos), along with "study"+"Treatment"
SS  = data.frame(Deaths=rep(0, 2*K), Total=rep(0,2*K),Study=rep(0, 2*K), Treatment=rep(0,2*K))
# take Treatment: 0=std and 1=high statin
SS$Treatment = c(rep(0, K), rep(1, K))
SS$Study = rep(1:K,2)
# death = pos
SS$Deaths = c(dat$cpos, dat$tpos)
# total= pos+neg
SS$Total = c(dat$cpos+dat$cneg, dat$tpos+dat$tneg)
# check the data to see whether this is reproduced:
SS
# check with the original data
dat
```

```
# some indices for IPD, especially the number of patients since we need to
# create a row for each patients as o or 1
len.SS = dim(SS)[1];len.SS
num.patient = sum(SS$Total);num.patient

# make it as IPD with 4 columns
IPD = matrix(0,nrow=num.patient,ncol=4)
tt=1
for(i in 1:len.SS){
  # 1st col as ID of the patient
  IPD[tt:(tt-1+SS$Total[i]),1]=tt:(tt-1+SS$Total[i])
  # 2nd col as Death = 1 and 0 = live for IPD logistic regression
  IPD[tt:(tt-1+SS$Total[i]),2]=c(rep(1,SS$Deaths[i]),rep(0,SS$Total[i]-SS$Deaths[i]))
  # 3rd as Study index to track the which study to contain those patients
  IPD[tt:(tt-1+SS$Total[i]),3]=rep(SS$Study[i],SS$Total[i])
  # 4th as Treatment index to track the treatment for the patients at that study
  IPD[tt:(tt-1+SS$Total[i]),4]=rep(SS$Treatment[i],SS$Total[i])
  tt=tt+SS$Total[i]
}  #End of i-loop
# make to data frame
IPD = data.frame(IPD)
names(IPD) = c("ID","TB","Trial","TRT")
summary(IPD)
# see how many patients. You see now we have an IPD data with 27,548 rows, with each row for a patient
dim(IPD)
sum(SS$Total)
# just check the first 100 patients from the 27,548 to see whether we did correct
head(IPD, 100)

# check with the SS:
aggregate(IPD$TB, list(IPD$Trial, IPD$TRT), sum)
```

The first 10 observations for the IPD data can be seen in Figure 11.

ID	TB	Trial	TRT
1	1	1	0
2	1	1	0
3	1	1	0
4	1	1	0
5	1	1	0
6	1	1	0
7	1	1	0
8	1	1	0
9	1	1	0
10	1	1	0

Figure 13: First 10 observations of the IPD data for the BCG vaccine study

We now fit a multi-level logistic regression model to the IPD data using the following R-code:

```
# random-intercept
fitIPD2 = glmer(TB~TRT+(1|Trial), data=IPD,family=binomial)
sumIPD2 = summary(fitIPD2)
sumIPD2

meta.ML
meta.REML
```

Consider the output below that shows the results for the generalized linear mixed model fit, using maximum likelihood (Laplace Approximation).

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: TB ~ TRT + (1 | Trial)
   Data: IPD

     AIC      BIC   logLik deviance df.resid
 27129.3  27161.6 -13561.6  27123.3   357344

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.5322 -0.0840 -0.0662 -0.0589 28.2869

Random effects:
 Groups Name        Variance Std.Dev.
 Trial  (Intercept) 1.988    1.41
Number of obs: 357347, groups:  Trial, 13

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.20443    0.19193  -21.91   <2e-16 ***
TRT         -0.47746    0.04053  -11.78   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
    (Intr)
TRT -0.009
```

We see that the IPD treatment effect is -0.47746 and the standard error for the treatment effect is 0.04053. As we can see, this is different to the results given by the fixed-effects meta-analysis of summary

data based on the odds-ratio outcome measure.

Now, let us consider the more general glm (generalized linear model) with estimated correlation of the random-effects. This is done using the following R-code:

```
fitIPD1 = glmer(TB~TRT+(TRT|Trial), data=IPD,family=binomial)
sumIPD1 = summary(fitIPD1)
sumIPD1
```

The results are given as follows:

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: TB ~ TRT + (TRT | Trial)
   Data: IPD

     AIC      BIC   logLik deviance df.resid
 26987.9  27041.8 -13488.9  26977.9   357342

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.5888 -0.0756 -0.0755 -0.0606 25.1681

Random effects:
 Groups Name        Variance Std.Dev. Corr
 Trial  (Intercept) 2.4379   1.5614
        TRT         0.3361   0.5797   -0.74
Number of obs: 357347, groups:  Trial, 13

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.1233     0.1574 -26.199  < 2e-16 ***
TRT          -0.7417     0.1124  -6.598 4.18e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
    (Intr)
TRT -0.283
```

Here we see that the IPD treatment effect is -0.7417 and the standard error for the treatment effect is 0.1124. This is not close to the results obtained by the fixed-effects meta-analysis of summary data based on the odds-ratio outcome measure. The above result further implies that the IPD meta-analysis and the summary data meta-analysis can indeed be different.

5 Conclusion

In this dissertation we considered meta-analysis of summary data versus individual participant-level (IPD) data, for trials with binary outcomes. The aim was to determine whether, in some cases, a meta-analysis of IPD data gains efficiency and produces asymptotically equivalent results to a meta-analysis of summary data in a binary setting.

A thorough investigation was done on the theory behind both meta-analysis and meta-analysis with binary data where we explored the fixed-effects model as well as the random-effects model. In these secions 2 and 3, important formulations and models were defined in detail and served as the bases of the application.

Two different clinical studies were investigated in order to determine how the results of a meta-analysis of summary data versus a meta-analysis of IPD data compare and/or differ if the sample sizes of the studies are different. The first study is called the *Cannon (Cannon et al., (2006)[13]) Cardiovascular Clinical Trials* which consists of four clinical studies. The second study is called the *Bacillus Calmette-Guerin Vaccine Study* which consists of a total of 13 clinical studies.

Both fixed-effects meta-analysis methods and random-effects meta-analysis methods were considered. Different outcome measures where also considered, specifically the odds-ratio, risk ratio and risk-difference. It was found that for the smaller study i.e. Cannon (Cannon et al., (2006)[13]) Cardiovascular Clinical Trials, if we consider only the odds-ratio outcome measure, the treatment effect results when doing a meta-analysis of summary data is very close to that of the meta-analysis of IPD data, only differing from the fourth decimal. What was interesting however, is when we considered the larger study i.e. Bacillus Calmette-Guerin Vaccine Study, and again only considering the odds-ratio outcome measure, the treatment effect results when doing a meta-analysis of summary data were different to that of the meta-analysis of IPD data. In general it is expected that as the sample size of the studies increase, the results of the meta-analysis of summary data and individual participant-level (IPD) data would become more similar. However, in both studies used in this dissertation, we illustrated that the number of studies are still small in the principle of mathematic asymptotics. In this case, the estimation of the between study variance τ^2 will be very unstable which will lead to a different conclusion between a meta-analysis of summary statistics (SS) and a meta-analysis of individual participant-level (IPD) data.

Further research can be done in investigating what effect even larger studies will have on a meta-

analysis of summary data versus a meta-analysis of IPD data. These future studies could also consider other outcome measures for both fixed-effects and random-effects to determine if, in any case, a meta-analysis of summary data and a meta-analysis of IPD data will produce asymptotically equivalent results in a binary setting.

6 Appendix

6.1 Gauss-Markov theorem

The Gauss Markov theorem suggests that the ordinary least squares (OLS) estimate for regression coefficients gives you the best linear unbiased estimate (BLUE) provided that a specific set of assumptions are met. These assumptions are as follows:

- The parameters which we are estimating using the OLS method must be linear.
- The data which we are using must be a random sample from the population.
- The regressors being calculated are not perfectly correlated with each other.
- The regressors and the error term are not correlated.
- Constant variance i.e. homoscedasticity

Suppose we have a linear regression nodel which is represented by

$$y_i = x_i\beta + \varepsilon_i$$

Generated by the ordinary least squares (OLS) estimate is the best linear unbiased estimate (BLUE) possible if

- $\mathrm{E}\{\varepsilon_i\} = 0$, for i = 1, . . . , N. sub
- $\{\varepsilon_1 \ldots\ldots \varepsilon_n\}$ and $\{x_1 \ldots\ldots, x_N\}$ are independent.
- $\mathrm{cov}\{\varepsilon_i, \varepsilon_j\} = 0$, for $i, j = 1, \ldots, N$ and $i \neq j$.
- $V\{\varepsilon_i\} = \sigma^2$, for i = 1, . . . N.

6.2 The Delta Method

Suppose we wish to consider the asymptotic distribution of some function of $\bar{X}_n$. Typically, the answer would depend on results which are already known: Consider a linear function $g(t) = at + b$ for some arbitrary constants a and b. Because $\mathrm{E}[\bar{X}_n] = \mu$, we are able to conclude by the linearity of the expectation operator that $\mathrm{E}[g(\bar{X}_n)] = a\mu + b = g(\mu)$. Consequently, a reasonable question to ask is whether $\sqrt{n}\left[g\left(\bar{X}_n\right) - g(\mu)\right]$ tends to some distribution as $n \to \infty$. Using the linearity of $g(t)$, we are able to write

$$\sqrt{n}\left[g\left(\bar{X}_n\right) - g(\mu)\right] = a\sqrt{n}\left(\bar{X}_n - \mu\right)$$

Theorem 1 : Suppose that $f(x)$ is a continuous function.

(a) If $X_n \xrightarrow{P} X$, then $f(X_n) \xrightarrow{P} f(X)$.

(b) If $X_n \xrightarrow{d} X$, then $f(X_n) \xrightarrow{d} f(X)$.

By Theorem 1, we are therefore able to conclude that

$$\sqrt{n}\left[g\left(\bar{X}_n\right) - g(\mu)\right] \xrightarrow{d} a\sigma Z$$

where the distribution of the right hand side above is $N\left(0, a^2\sigma^2\right)$.

The preceding development that a linear transformation of the random variable $\bar{X}_n$ alters its asymptotic distribution by a constant multiple is quite obvious. Now, what if the function $g(t)$ is nonlinear? A good understanding of the argument above is essential in this nonlinear case. Suppose we have the case where $\bar{X}_n$ is consistent for μ, then we know that $\bar{X}_n$ will be very close to μ for large values of n. As a result, the only meaningful aspect of $g(t)$ is its behavior in a small neighborhood of μ. In a small neighborhood of μ and using a first-order Taylor expansion, $g(\mu)$ may be considered to be roughly a linear function. In particular, we may approximate

$$g(t) \approx g(\mu) + g'(\mu)(t - \mu)$$

for t in a small neighborhood of μ. We see that $g'(\mu)$ is the multiple of t, and so the logic of the linear case above suggests the following

$$\sqrt{n}\left\{g\left(\bar{X}_n\right) - g(\mu)\right\} \xrightarrow{d} g'(\mu)\sigma^2 Z \tag{96}$$

Expression (74) is a special case of the powerful theorem known as the delta method.

Theorem 2 : The Delta method

If $g'(a)$ exists and $n^b\left(X_n - a\right) \xrightarrow{d} X$ for $b > 0$, then

$$n^b\left\{g\left(X_n\right) - g(a)\right\} \xrightarrow{d} g'(a)X$$

6.3 R code: Meta-analysis of summary data - Cannon study of cardiovascular clinical trials

```
# Load the library
```

```
library(gdata)
library(tidyverse)
library(fs)
library(readxl)
library(metafor)
library(CompQuadForm)
library(meta)

#dataset 1: Statin Clinical Trials
# Get the excel data book
datfile1 =
"C:/Users/Layla Norman/Documents/Masters ADA 2020/Research/Books/R4MetaBook1/dat4Meta.xls"
# Call "read.xls" to read the specific data
dat  = read_excel(datfile1, 25)

#nhigh - the total number of patients randomized to high dose
#evhigh - the number of patients with event (defined as the combined incidence of coronary death
or non-fatal myocardial infarction (MI)) in the high dose group
#nstd - the total number of patients randomized to standard dose
#evstd - the number of patients with event in the standard dose group
#ntot - the total number of patients in both treatment groups
#evtot - the total number of patients with events in both treatment groups

dat1 = dat[which(dat$Study == "Prove It"), ]
dat1

dat2 = dat[which(dat$Study == "A-to-Z"), ]
dat2

dat3 = dat[which(dat$Study == "TNT"), ]
dat3

dat4 = dat[which(dat$Study == "IDEAL"), ]
dat4
```

```
tpos1 = dat1$evhigh
tpos1

tpos2 = dat2$evhigh
tpos2

tpos3 = dat3$evhigh
tpos3

tpos4 = dat4$evhigh
tpos4

tneg1 = dat1$nhigh - tpos1
tneg1

tneg2 = dat2$nhigh - tpos2
tneg2

tneg3 = dat3$nhigh - tpos3
tneg3

tneg4 = dat4$nhigh - tpos4
tneg4

cpos1 = dat1$evstd
cpos1

cpos2 = dat2$evstd
cpos2

cpos3 = dat3$evstd
cpos3
```

```
cpos4 = dat4$evstd
cpos4

cneg1 = dat1$nstd - cpos1
cneg1

cneg2 = dat2$nstd - cpos2
cneg2

cneg3 = dat3$nstd - cpos3
cneg3

cneg4 = dat4$nstd - cpos4
cneg4

#Table
Prove_It = c(tpos1, tneg1, cpos1, cneg1)
A_to_Z = c(tpos2, tneg2, cpos2, cneg2)
TNT = c(tpos3, tneg3, cpos3, cneg3)
IDEAL = c(tpos4, tneg4, cpos4, cneg4)

# Make a data frame
datNew = as.data.frame(rbind(Prove_It, A_to_Z, TNT, IDEAL))
colnames(datNew) = c("tpos", "tneg", "cpos", "cneg")
# Print the data
datNew

#Commonly used ESs for binomial data are the risk ratio, the risk-difference, and the odds-ratio

#RR
datRR = escalc(measure = "RR", ai = tpos, bi = tneg, ci = cpos, di = cneg,
data = datNew, append = TRUE)
datRR
```

```
#OR
datOR = escalc(measure = "OR", ai = tpos, bi = tneg, ci = cpos, di = cneg,
data = datNew, append = TRUE)
datOR

#RR

datRD = escalc(measure = "RD", ai = tpos, bi = tneg, ci = cpos, di = cneg,
data = datNew, append = TRUE)
datRD
#-----------------------------------------------------------------------------------------------------
#Create a data frame called y2
#Odds Ratio
y2=metabin(data=dat, evhigh, nhigh, evstd, nstd, studlab = Study, sm="OR")

#plot the forest
forest(y2, is.summary=c(TRUE,TRUE,rep(FALSE,8),TRUE),leftcols = c("studlab", "evhigh",
"nhigh", "evstd", "nstd"), rightcols = c("effect","ci", "w.random", "w.fixed"),
col.square = "red", col.diamond = "blue", print.tau2 = F, print.Q = T, print.I2 = T)

funnel(y2)

#Create a data frame called y3
#Risk difference
y3=metabin(data=dat, evhigh, nhigh, evstd, nstd, studlab = Study, sm="RD")

#plot the forest
forest(y3, is.summary=c(TRUE,TRUE,rep(FALSE,8),TRUE),leftcols = c("studlab", "evhigh",
"nhigh", "evstd", "nstd"), rightcols = c("effect","ci", "w.random", "w.fixed"),
col.square = "red", col.diamond = "blue", print.tau2 = F, print.Q = T, print.I2 = T)

funnel(y3)

#Create a data frame called y1
```

```
#Risk Ratio
y1=metabin(data=dat, evhigh, nhigh, evstd, nstd, studlab = Study, sm="RR")

#plot the forest
forest(y1, is.summary=c(TRUE,TRUE,rep(FALSE,8),TRUE),
leftcols = c("studlab", "evhigh", "nhigh", "evstd", "nstd"),
rightcols = c("effect","ci", "w.random", "w.fixed"),
col.square = "red", col.diamond = "blue", print.tau2 = F, print.Q = T, print.I2 = T)

funnel(y1)
#-------------------------------------------------------------------------------------------------
#-------------------------------------------------------------------------------------------------
#4.2.1 Analysis with Risk-Ratio
#4.2.1.1 Definition

#Create a data frame called y1
#Risk Ratio
y1=metabin(data=dat, evhigh, nhigh, evstd, nstd, studlab = Study, sm="RR")

#plot the forest
forest(y1, is.summary=c(TRUE,TRUE,rep(FALSE,8),TRUE),leftcols = c("studlab", "evhigh",
"nhigh", "evstd", "nstd"), rightcols = c("effect","ci", "w.random", "w.fixed"),
col.square = "red", col.diamond = "blue", print.tau2 = F, print.Q = T, print.I2 = T)

funnel(y1)

#pE - risk (or risk probability) for the experimental high-dose group
# The events
xE = dat$evhigh
# the total number
nE = dat$nhigh
# the risk
pE = xE/nE
# print the risk
```

```
round(pE,digits=4)

#pC - risk (or risk probability) for the standard low-dose group
# The events
xC = dat$evstd
# the total number
nC = dat$nstd
# the risk
pC = xC/nC
# print the risk
round(pC,digits=4)

#Then the risk-ratios are calculated as:
# The risk-ratios
ES=pE/pC
# print the RR
round(ES,digits=3)

# A risk- ratio less than 1.0 would mean that the risk was lower in the experimental
high-dose group.
# Therefore, since the RR is less than 1 for all four studies, this would imply that the
experimental high-dose treatment is descriptively better than the standard treatment.
#-------------------------------------------------------------------------------------------
#4.2.1.2 Statistical Significance
#1. Confidence Interval Approach

# calculate the log risk-ratio
lnES = log(ES)
# print the log risk-ratio
round(lnES,digits=3)

# calculate the variance
VlnES = 1/dat$evhigh - 1/dat$nhigh +1/dat$evstd - 1/dat$nstd
# print the variance
```

```
round(VlnES,digits=5)

#We now can construct the 95% CI for lnES as follows:
# upper CI
ciup = lnES + 1.96*sqrt(VlnES)
round(ciup,digits=5)

# lower CI
cilow = lnES - 1.96*sqrt(VlnES)
round(cilow,digits=3)

#We then transform back to the original scale for risk- ratio(RR) as:
cat("The low CI is:", exp(cilow),"\n\n")
cat("The upper CI is:", exp(ciup),"\n\n")

#We can see from the 95% CI that only the study TNT (i.e. the third study) is statistically
significant, bacuase the interval does not contain 1, and the other three are not,
because the interval does contain 1.

#2. The p-value Approach

#We first calculate the z-values as follows:
z = lnES/sqrt(VlnES)
round(z, digits = 2)

#Then the p-values can be calculated as:
pval = 2*(1 - pnorm(abs(z)))
cat("p-values = ", pval,"\n\n")

#Again, only the study TNT is statistically significant at the 5% level.

#3. The Post-Power Approach

#Consider the sample sizes for each study
```

```
#The largest sample size would have greater power to detect statistically significant effect.

library(pwr)

# calculate the power and print it
pow.study = pwr.2p2n.test(ES.h(pE,pC),n1=dat$nhigh,
n2=dat$nstd,sig.level=0.05)
pow.study

#The associated statistical power for Study TNT is 0.885 and the rest of the three studies are all
about 40%, indicating these studies are low-powered statistically.
#-------------------------------------------------------------------------------------------------
#4.2.1.3 The Risk-Ratio Meta-Analysis: Step-by-Step

#The precision in meta-analysis can be addressed in two ways with one as the confidence interval
and another as the standard error of the ES.

round(VlnES, digits=5)

#The last two studies have the smallest variances indicating greater precision.

#Studies with greater precision should have larger weights when combining studies.
#This may be done by using a fixed-effects meta-analysis model and weighting each study by its
inverse variance as:

w = 1/VlnES
round(w, digits = 1)

#The fixed-effects weighted average meta-analysis model

#the inverse of variance for each study
fwi = 1/VlnES
round(fwi, digits=1)
```

```
# the total weight
fw = sum(fwi)
fw

# the weighted mean
flnES = sum(lnES*fwi)/fw
flnES

# the variance for the weighted mean
var = 1/fw
round(var, digits = 5)

#We then transform back to the original scale to estimate the RR as well as the CI as:

# the RR
exp(flnES)

# the lower and upper CI bounds
up = exp(flnES + 1.96*sqrt(var))
round(up, digits = 3)

low = exp(flnES - 1.96*sqrt(var))
round(low, digits = 3)

#This CI is below the RR of 1 which indicates that the high- dose regimen is statistically
significantly more effective than the standard dose regimen.

#To illustrate the random-effects meta-model, we first need to estimate the heterogeneity
statistic Q

# estimate heterogeneity statistic Q
Q = sum(fwi*lnES^2)-(sum(fwi*lnES))^2/fw
round(Q, digits=2)
```

```
# Get the degrees of freedom
K = dim(dat)[1]
K

df=K-1
df

# calculate the statistical significance: p-value
pval4Q = pchisq(Q, df, lower.tail = F)
pval4Q

#From this p-value, we can conclude that there is no statistically significant heterogeneity
among these 4 studies.
#-------------------------------------------------------------------------------------------
#4.2.1.4 Risk-Ratio Meta-Analysis: R package meta
# call "metabin" for RR meta-analysis

RR.Statin = metabin(evhigh,nhigh,evstd,nstd,studlab=Study,
data=dat, method="Inverse", sm="RR")
RR.Statin

forest(RR.Statin)
#-------------------------------------------------------------------------------------------
#-------------------------------------------------------------------------------------------
#4.2.2 Analysis with Risk-Difference
#4.2.2.1 Definition

# the risk difference
ESRD = pE - pC
round(ESRD, digits = 4)
#The differences are less than 0 for all four studies indicating that the experimental treatment
is descriptively better than the standard treatment.

# calculate the variance
```

```
VarRD = pE*(1 - pE)/nE + pC*(1 - pC)/nC
VarRD

# calculate the standard error
SERD = sqrt(VarRD)
SERD
#-------------------------------------------------------------------------------------------------
#4.2.2.2 Implementation in R Package meta

# call "metabin" for RD meta-analysis

RD.Statin = metabin(evhigh,nhigh,evstd,nstd,studlab=Study,
data=dat, method="Inverse", sm="RD")
# print the analysis summary
summary(RD.Statin)

#The conclusions from the meta-analysis of the risk ratio (RR) are reached from the meta-analysis
yof the risk dierence (RD); i.e. the high-dose regimen is statistically more effective
than the standard dose regimen.

forest(RD.Statin)
#-------------------------------------------------------------------------------------------------
#-------------------------------------------------------------------------------------------------
#4.2.3 Meta-Analysis with Odds-Ratio
#4.2.3.2 Odds-Ratio: Woolf 's Method

#The odds of MI in the experimental high-dose group is
oddsE = pE/(1-pE)
oddsE

#The odds of MI in the standard low-dose group is
oddsC = pC/(1-pC)
oddsC
```

```
#The OR would then be
OR = oddsE/oddsC
OR

#The log odds ratios are
LogOR=log(OR)
LogOR

#Study: TNT
#The approximate variance can be calculated as
VLogOR1 = 1/334 + 1/(4995-334) + 1/418 + 1/(5006-418)
VLogOR1

#The standard error
SE.LogOR1 = sqrt(VLogOR1)
SE.LogOR1

#Study: Prove It
#The approximate variance can be calculated as
VLogOR2 = 1/147 + 1/(2099-147) + 1/172 + 1/(2063-172)
VLogOR2

#The standard error
SE.LogOR2 = sqrt(VLogOR2)
SE.LogOR2

#Study: A-to-Z
#The approximate variance can be calculated as
VLogOR3 = 1/205 + 1/(2265-205) + 1/235 + 1/(2232-235)
VLogOR3

#The standard error
SE.LogOR3 = sqrt(VLogOR3)
SE.LogOR3
```

```
#Study: IDEAL
#The approximate variance can be calculated as
VLogOR4 = 1/411 + 1/(4439-411) + 1/463 + 1/(4449-463)
VLogOR4

#The standard error
SE.LogOR4 = sqrt(VLogOR4)
SE.LogOR4
#-------------------------------------------------------------------------------------------------------
#4.2.3.3 R Implementation with R Package meta

OR.Statin = metabin(evhigh,nhigh,evstd,nstd,studlab=Study,
data=dat, method="Inverse", sm="OR")
summary(OR.Statin)

forest(OR.Statin)
```

6.4 R code: Meta-analysis of Individual participant-level data - Cannon study of cardiovascular clinical trials

```
#################
# all libraries
###################
library(nlme)
library(lme4)
library(metafor)
library(gee)

#### Get the Statin data from metafor
data("dat.cannon2006", package = "metafor")
# for easy typing, just name it as dat
dat = dat.cannon2006;
# rename the data to confirm with Table 2 in your writing and keep the rest variable intact
colnames(dat) = c("Study","pop","nhigh","nstd", "evhigh","evstd", "ep2t","ep2c","ep3t",
```

```
"ep3c","ep4t","ep4c","ep5t","ep5c","ep6t","ep6c")
# add more variables for Table 3
dat$tpos = dat$evhigh
dat$tneg = c(1952,2060,4661,4028)
dat$cpos = dat$evstd
dat$cneg = c(1891, 1997,4588,3986)
# check the data: all ok
dat
############################################################
#I am only illustrating the logOR since OR is used in logistic regression for the IPD data
################################################
# Calculate the ES of log-odds using "escalc"
dat  = escalc(measure="OR",ai=tpos,bi=tneg,ci=cpos,
di = cneg, data = dat, append = TRUE)
dat

# FE
meta.FE = rma(yi, vi, data = dat, method="FE")
# Print the summary
meta.FE

# RE-DL
meta.DL = rma(yi, vi, data = dat, method="DL")
# Print the summary
meta.DL

# RE-REML
meta.REML = rma(yi, vi, method="REML",data = dat)
meta.REML

# RE-ML
meta.ML = rma(yi, vi, method="ML",data = dat)
meta.ML
################################################
```

```
# Now the individualization:
# ######################3
# 1. the data reformat
############################

# number of Studies
K = 4
# reformat the SS data into "death" from "total"(=neg+pos), along with "study"+"Treatment"
SS  = data.frame(Deaths=rep(0, 2*K), Total=rep(0,2*K),Study=rep(0, 2*K), Treatment=rep(0,2*K))
# take Treatment: 0=std and 1=high statin
SS$Treatment = c(rep(0, K), rep(1, K))
SS$Study = rep(1:K,2)
# death = pos
SS$Deaths = c(dat$cpos, dat$tpos)
# total= pos+neg
SS$Total = c(dat$cpos+dat$cneg, dat$tpos+dat$tneg)
# check the data to see whether this is reproduced:
SS
# check with the original data
dat

# some indices for IPD, especially the number of patients since we need to
# create a row for each patients as o or 1
len.SS = dim(SS)[1];len.SS
num.patient = sum(SS$Total);num.patient

# make it as IPD with 4 columns
IPD = matrix(0,nrow=num.patient,ncol=4)
tt=1
for(i in 1:len.SS){
# 1st col as ID of the patient
IPD[tt:(tt-1+SS$Total[i]),1]=tt:(tt-1+SS$Total[i])
# 2nd col as Death = 1 and 0 = live for IPD logistic regression
IPD[tt:(tt-1+SS$Total[i]),2]=c(rep(1,SS$Deaths[i]),rep(0,SS$Total[i]-SS$Deaths[i]))
```

```
# 3rd as Study index to track the which study to contain those patients
IPD[tt:(tt-1+SS$Total[i]),3]=rep(SS$Study[i],SS$Total[i])
# 4th as Treatment index to track the treatment for the patients at that study
IPD[tt:(tt-1+SS$Total[i]),4]=rep(SS$Treatment[i],SS$Total[i])
tt=tt+SS$Total[i]
}  #End of i-loop
# make to data frame
IPD = data.frame(IPD)
names(IPD) = c("ID","TB","Trial","TRT")
summary(IPD)
# see how many patients. You see now we have an IPD data with 27,548 rows,
# with each row for a patient
dim(IPD)
sum(SS$Total)
# just check the first 100 patients from the 27,548 to see whether we did correct
head(IPD, 100)

# check with the SS:
aggregate(IPD$TB, list(IPD$Trial, IPD$TRT), sum)
############################################################
# 2.multi-level logistic regression using IPD data
#########################################################
# random-intercept
fitIPD2 = glmer(TB~TRT+(1|Trial), data=IPD,family=binomial)
sumIPD2 = summary(fitIPD2)
sumIPD2
# the IPD trt effect is -0.1798, SE =0.043
# check it with the MA-SS from ML and REML
meta.ML
meta.REML
# you see that they are so close since for this data the tau^2 is estimated as
#zero, i.e. the random-effects MA is not needed.
# but this is not true for general case.
```

```
# more general glm with estimated correlation of random-effects
fitIPD1 = glmer(TB~TRT+(TRT|Trial), data=IPD,family=binomial)
sumIPD1 = summary(fitIPD1)
sumIPD1
# check this result with the MA-SS using REML and ML: we can see that the
# IPD and SS can be different
meta.REML
meta.ML
```

6.5 R code: Meta-analysis of summary data and Individual participant-level data - Bacillus Calmette-Guerin Vaccine Study

```
#################
# all libraries
#################
library(nlme)
library(lme4)
library(metafor)
library(gee)

#### Get the BCG data from metafor
data("dat.bcg", package = "metafor")
# for easy typing, just name it as dat
dat = dat.bcg;
dat
############################################################
# Calculate the ES of log-odds using "escalc"

dat  = escalc(measure="OR",ai=tpos,bi=tneg,ci=cpos,
di = cneg, data = dat, append = TRUE)
dat

# FE
meta.FE = rma(yi, vi, data = dat, method="FE")
# Print the summary
```

```
meta.FE

# RE-DL
meta.DL = rma(yi, vi, data = dat, method="DL")
# Print the summary
meta.DL

# RE-REML
meta.REML = rma(yi, vi, method="REML",data = dat)
meta.REML

# RE-ML
meta.ML = rma(yi, vi, method="ML",data = dat)
meta.ML
##################################################
# Now the individualization:
# ######################3
# 1. the data reformat
############################

# number of Studies
K = 13
# reformat the SS data into "death" from "total"(=neg+pos), along with "study"+"Treatment"
SS  = data.frame(Deaths=rep(0, 2*K), Total=rep(0,2*K),Study=rep(0, 2*K), Treatment=rep(0,2*K))
# take Treatment: 0=std and 1=high statin
SS$Treatment = c(rep(0, K), rep(1, K))
SS$Study = rep(1:K,2)
# death = pos
SS$Deaths = c(dat$cpos, dat$tpos)
# total= pos+neg
SS$Total = c(dat$cpos+dat$cneg, dat$tpos+dat$tneg)
# check the data to see whether this is reproduced:
SS
# check with the original data
```

```
dat

# some indices for IPD, especially the number of patients since we need to
# create a row for each patients as o or 1
len.SS = dim(SS)[1];len.SS
num.patient = sum(SS$Total);num.patient

# make it as IPD with 4 columns
IPD = matrix(0,nrow=num.patient,ncol=4)
tt=1
for(i in 1:len.SS){
# 1st col as ID of the patient
IPD[tt:(tt-1+SS$Total[i]),1]=tt:(tt-1+SS$Total[i])
# 2nd col as Death = 1 and 0 = live for IPD logistic regression
IPD[tt:(tt-1+SS$Total[i]),2]=c(rep(1,SS$Deaths[i]),rep(0,SS$Total[i]-SS$Deaths[i]))
# 3rd as Study index to track the which study to contain those patients
IPD[tt:(tt-1+SS$Total[i]),3]=rep(SS$Study[i],SS$Total[i])
# 4th as Treatment index to track the treatment for the patients at that study
IPD[tt:(tt-1+SS$Total[i]),4]=rep(SS$Treatment[i],SS$Total[i])
tt=tt+SS$Total[i]
}  #End of i-loop
# make to data frame
IPD = data.frame(IPD)
names(IPD) = c("ID","TB","Trial","TRT")
summary(IPD)

dim(IPD)
sum(SS$Total)

head(IPD, 100)

# check with the SS:
aggregate(IPD$TB, list(IPD$Trial, IPD$TRT), sum)
##########################################################
```

```
# 2.multi-level logistic regression using IPD data
#######################################################
# random-intercept
fitIPD2 = glmer(TB~TRT+(1|Trial), data=IPD,family=binomial)
sumIPD2 = summary(fitIPD2)
sumIPD2

# check it with the MA-SS from ML and REML
meta.ML
meta.REML

# more general glm with estimated correlation of random-effects
fitIPD1 = glmer(TB~TRT+(TRT|Trial), data=IPD,family=binomial)
sumIPD1 = summary(fitIPD1)
sumIPD1
# check this result with the MA-SS: we can see that the IPD and SS can be different
meta.REML
meta.ML
```

7 References

[1] Chen, D.G.; Liu, D.; Min, X. & Zhang, H. *Relative efficiency of using summary versus individual data in random-effects meta-analysis.* Biometrics, Wiley Online Library, 2020.

[2] Haidich, A.B. *Meta-analysis in medical research.* Hippokratia, Hippokratio General Hospital of Thessaloniki, 14, 29, 2010.

[3] Lin, D. & Zeng, D. *On the relative efficiency of using summary statistics versus individual-level data in meta-analysis.* Biometrika, Oxford University Press, 97, 321-332, 2010.

[4] Mathew, T. & Nordström, K. *On the equivalence of meta-analysis using literature and using individual patient data.* Biometrics, Wiley Online Library, 55, 1221-1223, 1999.

[5] Mathew, T. & Nordström, K. *Comparison of one-step and two-step meta-analysis models using individual patient data.* Biometrical Journal: Journal of Mathematical Methods in Biosciences, Wiley Online Library, 52, 271-287, 2010.

[6] Olkin, I. & Sampson, A. *Comparison of meta-analysis versus analysis of variance of individual patient data.* Biometrics, JSTOR, 317-322, 1998.

[7] Thomas, D.; Radji, S. & Benedetti, A. *Systematic review of methods for individual patient data meta-analysis with binary outcomes.* BMC medical research methodology, Springer, 14, 79, 2014.

[8] Turner, R. M.; Omar, R. Z.; Yang, M.; Goldstein, H. & Thompson, S. G. *A multilevel model framework for meta-analysis of clinical trials with binary outcomes.*Statistics in medicine, Wiley Online Library, 19, 3417-3432, 2000.

[9] DerSimonian, R. & N. Laird. *Meta-analysis in clinical trials.* Controlled Clinical Trials, 7, 177-188, 1986.

[10] Whitehead, A. *Meta-analysis of Controlled Clinical Trials.* Biometrical Journal, Wiley, 2003.

[11] Bower, P., Byford, S., Barber, J., Beecham, J., Simpson, S., Friedli, K., Corney, R., King, M. and Harvey, I. *Meta-analysis of data on costs from trials of counseling in primary care: using individual patient data to overcome sample size limitations in economic analyses.* British Medical Journal 326, 1247 - 1250, 2003.

[12] Kiefer, J. *Optimum experimental designs.* Journal of the Royal Statistical Society, Series B 21, 272-319, 1959.

[13] Cannon, C. P., B. A. Steinberg, S. A. Murphy, J. L. Mega, and E. Braunwald. *Meta-analysis of cardiovascular outcomes trials comparing intensive versus moderate statin therapy.* Journal of the American College of Cardiology 48, 438-445, 2006.

[14] van Houwelingen HC, Zwinderman KH, Stijnen T. *A bivariate approach to meta-analysis.* Statistics in Medicine 12:2273-2284, 1993.

[15] Colditz, G. A., T. F. Brewer, C. S. Berkey, M. E. Wilson, E. Burdick, H. V.Fineberg, and F. Mosteller. *Efficacy of bcg vaccine in the prevention of tuberculosis: Meta-analysis of the published literature.* Journal of the American Medical Association 271, 698-702, 1994.

[16] Buys, M. *Protecting personal information: Implications of the Protection of Personal Information (POPI) Act for healthcare professionals.* South African Medical Journal 107.11, 954-956, 2017.

[17] Chen, D.G & Peace, K.E. *Applied Meta-Analysis with R.* crc press, 2013.