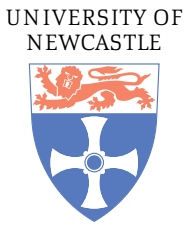
Meta-analysis and Sensitivity analysis for selection bias in Multi-arm trials

Hathaikan Chootrakool

A thesis submitted to the degree of Doctor of Philosophy



School of Mathematics and Statistics

Newcastle University

Acknowledgements

I am grateful to my supervisor, Dr.Jian Qing Shi and would like to thank him for his excellent supervision, support, encouragement and patience over the last few years. His guidance and advice has proven invaluable. This thesis would not have had been possible without him.

I would like to thank Rajhabhat Suan Dusit University, Thailand for funding my studies. I wish to thank Assoc. Prof. Dr. Sukhum Chaleysub for encoragement and help during my time here. Also thanks to all the staff and friends in the school for support and comfort, especially Enikő, Entisar, Yuki and Nasr. Further, I wish to thank Martin for continuous support, help and proof-reading this thesis.

Finally, I would like to thank my family for their infinite love and giving me strength throughout this research. I wish to dedicate this work to my father who inspired me to study for this degree.

Abstract

Meta-analysis of multi-arm trials has been used increasingly in recent years, the aims of which are to combine evidence from all possible similar studies and draw inferences about the effectiveness of multiple compared-treatments. Antiplatelet therapy is a pharmacologic therapy which aims to inhibit platelet activation and aggregation in the setting of arterial thrombosis. Throughout the thesis we use binary data from antiplatelet therapy to apply the model and sensitivity analysis. The normal approximation model using empirical logistic transform has been employed to compare different treatments in multi-arm trials, allowing studies of both direct and indirect comparisons. The issue of direct-indirect comparison is studied in detail, borrowing the strength from the indirect comparisons and making inferences about appropriately chosen parameters. Additionally, a hierarchical structure of the model addresses the problem of heterogeneity among different studies. However the model requires a large sample size of each individual study. When the sample size is small, an exact logistic regression model is introduced. Both unconditional and conditional maximum likelihood approaches are performed to make inferences for the logistic regression model. We use Gaussian-Hermite quadrature to approximate the integral involved in the likelihood functions. Both approaches have been examined to different cases in the simulation study.

Studies with statistically significant results (positive results) are potentially more likely to be submitted or selected more rapidly than studies with non-significant results (negative results). This leads to false-positive results or an incorrect, usually over-optimistic, conclusion, a problem known as selection bias in the meta-analysis. A funnel plot is a graphical tool which is used to detect selection bias in this research. We apply the idea of a sensitivity analysis by defining a selection model to the available data of a meta-analysis, by allowing different amounts of selection bias in the model and investigate how sensitive the main interest parameter is when compared to the estiamtes of the standard model. We also examine the sensitivity analysis by the simulation study.

Contents

1 Introduction **1**

1.1 A brief history and basic concepts of meta-analysis 1

1.2 Measure of treatment effect 4

1.2.1 Comparative binary outcome 4

1.3 Significance problems in meta-analysis 5

1.3.1 Heterogeniety 5

1.3.2 Selection bias 8

1.4 Multi-arm trials 10

1.4.1 Methods of meta-analysis 11

1.5 Gaussian quadrature approximation 14

1.5.1 Gauss-Hermite integration 15

1.6 Outline of the thesis 16

2 Antiplatelet data **19**

2.1 Overview of antiplatelet therapy 19

2.2 Antiplatelet data: maintaining vascular patency (W1) 21

2.3 Antiplatelet data: reduction in venous thrombosis and pulmonary embolism (W2) 23

3 Meta-analysis of multi-arm trials using normal approximation approach **27**

3.1 Introduction 27

3.2 The data structure of multi-arm trials 28
3.3 Normal approximation model based on empirical logistic transform 29
3.4 Empirical log-odds model 30
3.4.1 Meta-analysis of multi-arm trials 31
3.4.2 Meta-analysis of multi-arm trials with both direct and indirect comparisons 34
3.5 Empirical log-odds ratio model 37
3.5.1 Meta-analysis of multi-arm trials 37
3.5.2 Meta-analysis of multi-arm trials with both direct and indirect comparison 38
3.6 Maximum likelihood estimation 40
3.7 Standard error of parameter estimation 41
3.8 Application to antiplatelet therapy data (W1) 42
3.8.1 The model 42
3.8.2 Maximum likelihood estimation 46
3.8.3 Numerical results 47
3.9 Discussion 49

4 Meta-analysis of multi-arm trials using binomial approach 53

4.1 Introduction 53
4.2 Fitting the logistic regression model 55
4.3 Unconditional maximum likelihood approach 56
4.3.1 Probability functions 57
4.3.2 The unconditional likelihood 57
4.3.3 Asymptotic variance-covariance matrix 59
4.4 Conditional maximum likelihood approach 61
4.4.1 Conditional likelihood 62
4.4.2 Estimation 63

4.5 Application to antiplatelet therapy data (W2) 64
4.5.1 Unconditional inference 64
4.5.2 Conditional inference 67
4.6 Discussion 69

5 Simulation study **71**
5.1 Introduction 71
5.2 Simulated data 72
5.3 The models 73
5.3.1 The empirical log-odds ratio model 73
5.3.2 The logistic regression model 74
5.4 Comparison of models 75
5.5 Simulation details 77
5.6 Results 79
5.6.1 Scenario 1 79
5.6.2 Scenario 2 81
5.7 Discussion 82

6 Sensitivity analysis to bivariate normal approximation model **89**
6.1 Identifying selection bias in multi-arm trials 91
6.2 Selection bias 92
6.2.1 Assumption for population model 94
6.2.2 Selection model 94
6.2.3 Relating mathematical consequences 96
6.3 Likelihood 99
6.4 Goodness of fit 102
6.5 Sensitivity analysis 105
6.5.1 The possible range of (a_1, b_1) and (a_2, b_2) (Step 1) 105

6.5.2 Estimation and goodness-of-fit test (Step 2) 108
6.5.3 Sensitivity analysis (Step 3) 113
6.6 Simulation study 115
6.7 Some theorems of mathematical consequences 117
6.8 Discussion 127

7 Sensitivity analysis to logistic regression model 129
7.1 Introduction 129
7.2 Simulated data 130
7.3 Multi-arm trials with the conditional probability 131
7.4 Detecting the selection bias 133
7.4.1 Conditional variance 133
7.4.2 Conditional mean 134
7.4.3 Funnel plot 135
7.4.4 Standard error 135
7.5 Selection bias 137
7.6 Likelihood 140
7.7 Goodness of fit 141
7.8 Sensitivity analysis 142
7.9 Discussion 146

8 Conclusions and further development 149
8.1 Conclusions 149
8.2 Further devolopment 150

List of Figures

2.1 Platelet aggregation 20

5.1 The Q-Q plot: the trial effects for M =27 84

5.2 The Q-Q plot: the trial effects for M =54 84

6.1 The funnel plots:(a) $Y_{i,AC}$ against $s_{i,AC}$ for G_1; (b) $Y_{i,BC}$ against $s_{i,BC}$ for G_1;(c) $Y_{i,AC}$ against $s_{i,AC}$ for G_3;(d) $Y_{i,BC}$ against $s_{i,BC}$ for G_4. 93

6.2 Funnel plot: $Y_{i,AC}$ against φ_i for G_1- the solid line represents the estimate without selectivity $\widehat{\mu_{AC}} = 0.5689296$; the dashed lines represent the fitted values for given (a_1, b_1, a_2, b_2) which $(a_1, b_1, a_2, b_2, \widehat{\mu_{AC}})$ are equal to (0.458,0.063,0.339,0.135, 0.54),(-0.17,0.16,-0.49,0.36,0.52) and (-0.41,0.15,-0.70,0.33,0.50). 111

6.3 Funnel plot: $Y_{i,AC}$ against φ_i for G_1- the solid line represents the estimate without selectivity $\widehat{\mu_{BC}} = 0.6770754$; the dashed lines represent the fitted values for given (a_1, b_1, a_2, b_2), which $(a_1, b_1, a_2, b_2, \widehat{\mu_{AC}})$ are equal to (0.458,0.063,0.339,0.135, 0.58),(-0.17,0.16,-0.49,0.36,0.52) and (-0.41,0.15,-0.70,0.33,0.47). 112

6.4 Funnel plot: $Y_{i,BC}$ against $s_{i,BC}$ for G_4- the solid line represents the estimate without selectivity $\widehat{\mu_{BC}} = 0.6770754$; the dashed lines represent the fitted values for given (a_1, b_1, a_2, b_2), which $(a_1, b_1, a_2, b_2, \widehat{\mu_{BC}})$ are equal to (0.458,0.063,0.339,0.135, 0.58),(-0.17,0.16,-0.49,0.36,0.52) and (-0.41,0.15,-0.70,0.33,0.47). 112

6.5 The W1 data: $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = 0$. 114

6.6 The W1 data: $\widehat{\mu_{BC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = 0$. 114

6.7 The generated data with bias where $\varrho_1 = \varrho_2 = 0.8$ corresponds to (P_{min}(selection), P_{max}(selecti =(0.90,0.10): (a) funnel plot of $Y_{i,AC}$ against $s_{i,AC}$; (b) funnel plot of $Y_{i,BC}$ against $s_{i,BC}$; (c) $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \beta_1 = \beta_2 = 0$; (d)$\widehat{\mu_{BC}}$ against the p-value of $H_0 : \beta_1 = \beta_2 = 0$ 118

6.8 The generated data with bias where $\varrho_1 = \varrho_2 = 0.8$ corresponds to (P_{min}(selection), P_{max}(selecti =(0.80,0.20): (a) funnel plot of $Y_{i,AC}$ against $s_{i,AC}$; (b) funnel plot of $Y_{i,BC}$ against $s_{i,BC}$; (c) $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \beta_1 = \beta_2 = 0$; (d)$\widehat{\mu_{BC}}$ against the p-value of $H_0 : \beta_1 = \beta_2 = 0$ 119

7.1 The funnel plot:$Y_{i,AC}$ against $v_{i,AC}$-the dashed lines represent the conditional mean of $Y_{i,AC}$ given c_i. 136

7.2 The funnel plot:$Y_{i,BC}$ against $v_{i,BC}$-the dashed lines represent the conditional mean of $Y_{i,BC}$ given c_i. 136

7.3 $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = \mathbf{0}$ 145

7.4 $\widehat{\mu_{BC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = \mathbf{0}$ 146

List of Tables

2.1 The W1 data: 31 RCTs of aspirin data 25

2.2 The W2 data: 27 RCTs of aspirin data 26

3.1 The results for the empirical log-odds ratio models on the log-odds ratio (LOR) and odds ratio (LO) scales 48

4.1 The results of the treatment effects for the model using the unconditional method 66

4.2 The trial effects of the model using the unconditional method 67

4.3 The results of the treatment effects for the model using the conditional method 68

5.1 Conclusions of the models 76

5.2 Simulation study results based on the data generated from S1 with M = 27 85

5.3 Simulation study results based on the data generated from S1 with M = 54 86

5.4 Simulation study results based on the data generated from S2 with M = 27 87

5.5 Simulation study results based on the data generated from S2 with M = 54 88

6.1 The pairs (a_1, b_1) for the selection model (6.3) 107

6.2 The pair of (a_2, b_2) for the selection model (6.4) 107

6.3 The W1 data with selection model: summary of outputs 109

6.4 The simulated three-arm data: summary of outputs 116

7.1 The pairs of (a_1, b_1) for the selection model Z_{i1} 143

7.2 The bias-simulated data with selection: summary of outputs 144

Chapter 1

Introduction

1.1 A brief history and basic concepts of meta-analysis

There has been a massive growth in the number of randomised clinical trials (RCTs) since the first RCT was introduced in the well-known streptomycin trial in 1946 (see the disscussion in Hill, 1990). The results of RCTs have been spread over many reports and thousands of medical journals. The available results would be impossible to read individually and difficult to summarize. In making some of this information more readily available, an attempt is made to pull together the existing evidence in a form that can be used by researchers or statisticians; this is called *systematic review*. The aim of systematic reviews is to find and assess for inclusion all possible high quality studies addressing the clinical question of the review. There is an international network of clinicians and methodologists who have formed the *Cochrane Collaboration*. It was founded in 1993 and named after the British epidemiologist, Archie Cochrane. This organization is dedicated to the compilation and registration of RCTs, the combination of appropriate results and the dissemination of findings through a regularly updated electronic database. What does systematic review achieve? It reduces the large quantities of information to a manageable size. The results can often be generalized to a wider population in a broader setting than would be possible from a single study. Also, systematic reviews aim to reduce errors and tend to improve the reliability.

Systematic reviews provide the research evidence input into the process of evidence-based decision making. An important aspect of most reviews is the quantitative synthesis of results; thus *meta-analysis* is the statistical part of systematic review. Other names given to meta-analysis include overview, quantitative overview, pooling, pooled analysis, integrative research review, research integration, research consolidation, data synthesis, quantitative synthesis, and combining studies (Jenicek, 1989). However, a meta-analysis is also possible without doing a systematic review - some studies could be combined without any attempt to be systematic about how the particular studies were chosen. The minimum requirement to produce a meta-analysis is the availability of data from two or more studies, irrespective of whether they are reviewed narratively or systematically (Jadad, 1998, page 83). We can define meta-analysis as a statistical tool that summarizes evidence from multiple studies of a particular topic and attempts to provide an estimate of true effect. The main purpose of meta-analysis is to increase the precision of the conclusions of a review. With statistical perspective, it is able to detect treatment effects with greater power and estimate these effects with greater precision than any single study. In this thesis, we use two meta-analyses from systematic reviews of Antiplatelet Trialists' Collaboration (Collaboration, 1994a,b).

Meta-analysis has been widely used in many areas. The term *meta-analysis* was first used by Glass (1976) in education. He distinguished types of statistical analyses in social science and termed the original analysis of a set of data '*primary analysis*'. *Secondary analysis* is a re-analysis of data that has already been collected by another investigator. Some of these analyses are conducted to reaffirm answers to questions raised in the primary analysis, whereas other secondary analyses attempt to answer new questions. In addition, he defined other basic features of meta-analysis as it is known and used today. Hedges and Olkin (1985) published their book 'Statistical methods for meta-analysis', which is the first book in meta-analysis. The idea of meta-analysis can be traced back to Pearson (1904). He developed a method for summarizing correlation coefficients for studies of typhoid vaccina-

tion. Statistical techniques for combining study results were also used by Yates and Cochran (1938) in agriculture. Their technique has led to an increase in development and application of meta-analysis. One of the first meta-analyses in medicine in the modern era was introduced by Chalmers et al. (1977). However, it was not until the mid-1980s that meta-analysis started to be used frequently in the health care field when Yusuf et al. (1985) published their meta-analysis and concluded that the long-term beta blockage following discharge from the coronary care unit after a myocardiac infarction reduced mortality.

Over the last few decades, individual participant data (IPD) of systematic review for meta-analysis has increased rapidly. Jennison and Turnbull (1990); Stewart and Parmar (1993) and Oxman et al. (1995) concluded a number of advantages to IPD meta-analysis. In fact, the disadvantages of performing an IPD meta-analysis are the costs in both time and money. In biostatistics, Van Houwelingen (1997) interestingly listed meta-analysis among his nightmares, which he hoped would not happen in the future. He suggested about analysing summary measures from selective studies and he looked forward to a time when IPD from all studies were available to be synthesized using appropriate random-effects models. Simmonds et al. (2005) argued that the process of systematic review, within which the majority of meta-analyses are now undertaken, has to some extent reduced bias due to selective inclusion of studies, and analyses involving IPD continue to increase in number. Additionally, the results of meta-analysis need to be reported properly. Mother et al. (1999) suggested the guidelines for presenting the results of RCTs in meta-analysis, see more similar suggestions in Mother et al. (2001); Bussuyt et al. (2003) and Von Elm and Egger (2004) .

Meta-analysis has been extended beyond medicine and health to cover various fields from 'astronomy to zoology' (Petticrew, 2001). It has been used in economics (Stanley and Jarrel, 1989, 1998; Stanley, 1998, 2001), and is beginning to be used in political science (Pinello, 1999). In industrial organizational psychology, there have been numerous applications of

meta-analysis (Schmidt, 1988; Schmidt and Hunter, 1981, 1998). A good example of how to explain a meta-analysis is 'mixing apples and oranges', introduced by Moayyedi (2004). Meta-analysis has become important in research in almost every area. Nowadays, it would probably be difficult to find a research area in which meta-analysis is unknown.

1.2 Measure of treatment effect

Before the results of studies can be considered for pooling in a meta-analysis, it is necessary to decide a measure to use for evaluating the efficiency of one treatment relative to another. In clinical trials, the control treatment (or control group) is a standard treatment or a placebo. Various terms have been used for the measure including *'relative efficacy'*, *'efficacy of the (first) treatment'*, and *'treatment difference'*, see e.g. Higgins and Whitehead (1996) and Higgins et al. (2001). The term *'treatment effect'* is preferred and will be used throughout this thesis.

1.2.1 Comparative binary outcome

Measures of outcome need to be calculated for each study in a meta-analysis before they can be quantitatively combined. Outcomes of the data have been categorized into three groups: binary data, continuous data and ordered categorical data. The data being used in this thesis, described in Chapter 2, is a comparative binary outcome, where two possible outcomes - diagnosis/not diagnosis- are compared. In this section, we shall give a measure of the treatment effect, which is the *log-odds ratio*. The other measures can see from Sutton et al. (2000, page 17); for example, mean difference and effect size. To describe the log-odds ratio, suppose that two treatments denoted A and C in Table 2.1 of Chapter 2 represent 'aspirin plus dipyridamole' and 'control group' respectively. Let π_A and π_C be the probabilities of patients that have reocclusion (can be treated as failures) on treatments A and C respectively. The odds ratio (OR) of patients that have reocclusion on treatment A relative to treatment

C can be defined by $\pi_A(1-\pi_C)/\pi_C(1-\pi_A)$. To interpret OR, if an odds ratio estimate is less than one, it would indicate an improvement with treatment A. A ratio of greater than one would imply that treatment A was less effective than the control treatment. For the purpose of combining the studies, it is common to transform the data by taking the natural logarithm of the odds ratio and work with the log-odds ratio, as this should provide a measure which is approximately normally distributed. Thus the log-odds ratio (LOR) can be written by

$$\text{LOR} = \log\left(\frac{\pi_A(1-\pi_C)}{\pi_C(1-\pi_A)}\right). \tag{1.1}$$

The other measures that can be used for this type are *relative risk* (RR), π_A/π_C, or the *risk difference* (RD), $\pi_A - \pi_C$. Each measure has a different clinical meaning.

1.3 Significance problems in meta-analysis

In the first section, we saw how beneficial meta-analysis is and how it has been used in several areas. In the second section, we defined the outcome measure. In meta-analysis, each study involved is different from all the others. Such differences cause statistical problems or difficulties in deciding the appropriateness of pooling. Several problems have arisen in meta-analysis, for example, aggregating studies that include different measuring techniques, different definitions of variables, and subjects that are too dissimilar results in meta-analyses that are uninterpretable because they are from poorly designed studies (Hedges and Olkin, 1985). Thus, if meta-analysis is used or analysed improperly, it can lead to erroneous conclusions regarding to treatment effect. Here we will focus on two major problems, heterogeneity and selection bias, described as follows.

1.3.1 Heterogeniety

Heterogeneity may be defined as the variation that arises due to differences across studies in populations, interventions, outcomes, and designs. Even when all studies are measuring

the same underlying average effect, the results may vary across studies because of random errors. What causes the heterogeneity in a meta-analysis? Bailey (1987) suggested the possible causes of heterogeneity can be categorized as (1) due to chance; (2) spurious, due to the scale used to measure the treatment effect; (3) due to treatment characteristics; (4) due to individual data; (5) characteristics of the design and conduct of the studies; (6) unexplainable, if none of the above account for it. How do we know whether there is heterogeneity or not? A chi-squared test is traditionally undertaken to determine whether there is statistically significant evidence against a null hypothesis of no heterogeneity or not. The null hypothesis is that the true treatment effects are the same in all studies, $H_0 : \delta_1 = \delta_2 = \ldots = \delta_M$ versus the alternative that at least one of the treatment effects differs from the remainder. The δ_i's are the underlying true treatment effects corresponding to the ith study, which is defined in (1.1) for $i = 1, \ldots, M$ where M is a number of studies being combined in a meta-analysis. One test statistic is defined by

$$Q = \sum_{i=1}^{M} w_i T_i^2 - \frac{(\sum_{i=1}^{M} w_i T_i)^2}{\sum_{i=1}^{M} w_i},$$

where T_i is the treatment effect estimate of δ_i and w_i is the weight in the ith study. The weight is usually the reciprocal of the variance of the outcome estimate. We omit the detail here, more discussion and an example can be found in Sutton et al. (2000, page 39). The statistic Q is approximately distributed as a χ^2 distribution on $M - 1$ degrees of freedom under the null hypothesis H_0. If the null hypothesis is not significant then there is assumed to be no heterogeneity between studies. An analysis may be performed by a fixed effect model where the treatment effect is considered to be the same for all studies. The standard error estimate in each study is based on the sampling variation of the study. The model may provide a useful summary of the results. However, the fixed effect models are specific to the particular studies included in the meta-analysis and may not be realistic. Different studies with differing designs will not necessarily estimate the same quantity (Matthews, 2005, page 134). In contrast to the above hypothesis, if the null hypothesis is rejected then

the random effect model would be more appropriate. The model allows the between-study variability to be accounted for the overall estimate and, more particularly, its standard error.

One of the controversies surrounding meta-analysis has concerned the choices between a fixed effect model and a random effect model for providing an overall estimate of the treatment effect. Many authors have exploited the heterogeneity and the fixed-random effect model. The popular DerSimonian-Laird approach to random-effects meta-analysis uses a simple estimate of within-study variance, and does not incorporate uncertainty in the variance estimate when making inference on the mean of the random-effects distribution (DerSimonian and Laird, 1986). According to the use of test Q, when the sample sizes in each study are very large, the null hypothesis may be rejected even if the individual treatment effect estimates are not very different (Shadish and Haddock, 1987). If the number of combined studies is small then the statistical power of tests are, in most cases, very low (Boissel et al., 1989). The alternative way to deal with heterogeneity is to use a one-way analysis of variance (ANOVA) to investigate heterogeneity between and within groups of studies, where the groups are categorized by study characteristics (Hedges and Olkin, 1985, page 12). Since the formal Q statistic (in most cases) has a low power, there are a number of graphical informal tests: a plot of normalized scores, a forest plot, a Radial plot (Galbraith diagram) and a L'Abbé plot (Sutton et al., 2000, chapter 7). To assess heterogeneity, Thompson and Sharp (1999) compared a number of methods used to investigate whether a particular covariate, with a value defined for each study in the meta-analysis, explained any heterogeneity. The random-effects method has also long been associated with the problems due to poor estimation of among-study variance when there is little information (Hardy and Thompson, 1996; Ziegler et al., 2001). Song et al. (2001) reviewed the methods used in meta-analysis for exploring heterogeneity. Glasziou and Sanders (2002) addressed the cause of heterogeneity in a system review. Recently, Hedges and Pigott (2001) and Jackson (2006) discussed theoretically the power of the test for heterogeneity. In this thesis, we assume all treatment effects in the

model to be random effects to avoid the problem of heterogeneity and also we do not believe that the results from different studies and different designs can have the same treatment effect.

1.3.2 Selection bias

It has long been accepted that studies or researches with statistically significant results (positive results) are potentially more likely to be written up, submitted, selected or published more rapidly than studies with non-significant results (negative results), which leads to false-positive results. In meta-analysis, combining only the identified published studies uncritically may lead to an incorrect, usually over-optimistic conclusion. This problem is known as *publication bias* or *selection bias*. For example, several studies (Greenwald, 1975; Coursol and Wagner, 1986; Sommer, 1987) have surveyed authors, and found that, generally, studies with non-significant results are less likely to be submitted for publication compared to those with statistically significant results. Various tools such as the funnel plot, the rank correlation test, the linear regression test and trim and fill to identify publication bias are briefly described below.

Funnel plots are a primary visual tool for the investigation of publication bias in meta-analysis. They are simple scatter plots of the treatment effects, estimated from individual studies against a measure of study size. The axis of the treatment effect can be log-odds ratio, log risk ratio or risk difference. The other axis can be one of these choices: the standard error, the inverse of standard error, the variance, the inverse of variance, the sample size, log sample size. They can be used in different circumstances (see Sterne and Egger, 2001). Generally, the treatment effect estimates from individual studies are often plotted against their standard errors (or the inverse of the standard error), instead of the corresponding sample size. The log-odds ratio and standard error are the best choices in most cases (Rothstein et al., 2005, page 86). The name 'funnel plot' is based on the fact that the

precision in the estimation of the underlying treatment effect increases as the sample size of the studies increases. In this thesis, the measure of study size is plotted on the horizontal axis and the treatment effect estimate on the vertical axis. The results from smaller studies will scatter widely on the right-hand side of the graph, with the spread narrowing among larger studies. In the absence of bias, the plot will resemble a symmetrical funnel. Asymmetry in the funnel plots may indicate publication bias in meta-analysis. Funnel plots were first introduced in educational research and psychology by Light and Pillemer (1984). In 1995, Egger and Davey Smith (1995) used funnel plots for a meta-analysis that might have alerted investigators to the unreliability of small studies on the effect of magnesium treatment for myocardial infarction that found no or little evidence that magnesium treatment reduced mortality.

The 'rank correlation test', described by Begg and Mazumdar (1994), examined the association between the treatment effect estimates and their variances, to exploit the fact that publication bias will tend to induce a correlation between the two factors, and constructs the rank-ordered sample on the basis of one of them. The test is a distribution-free method, which involves no modelling assumptions, but it suffers from a lack of power, and so the possibility of publication bias cannot be ruled out even when the test is non-significant. To test the asymmetry of a funnel plot, Egger et al. (1997) suggested a method, called the 'linear regression test' based on a regression analysis of Galbraith's radial plot (Galbraith, 1988).

To address the problem of publication bias, the 'trim and fill' method was developed by Duval and Tweedie (2000a,b) to adjust a meta-analysis for the impact of missing studies. The method relies on the scrutiny of one side of a funnel plot for asymmetry, assumed to be due to publication bias. It appears to give results that match the subjective visual assessment of a funnel plot. This method is based on a strong assumption of symmetry. Copas and Shi (2001, 2002) argued that some parameters linked to selection bias are inestimable

since the number of unselected studies is impossible to know. They proposed a sensitivity analysis with which different patterns of selection bias can be tested against the fit of the funnel plot. In a similar way, they discussed the sensitivity analysis for the meta-analysis of 2×2 tables using the exact conditional distributions (Shi and Copas, 2002). A Markov chain Monte carlo EM algorithm was used to calculate maximum likelihood estimates.

Group dose measures in epidemiological studies have been another problem for meta-analysis. Shi and Copas (2004) proposed a model that allows for an arbitrarily aggregated dose level, and indicated that the resulting estimates and standard errors can be quite different from those given by the usual method.

1.4 Multi-arm trials

Most meta-analysis has focused on summarizing treatment effect measures based on the comparison of two treatments (called *'arms'*, sometimes also called *'interventions'* or *'exposures'*). In this comparison, two groups of individual studies are exposed to two different treatments. Standard two-arm RCTs are frequently used in clinical research due in part to its relative simplicity of design and interpretation. At its most basic, one power, one significance level and one magnitude of difference are analyzed for two-arm comparisons. Conclusions are straightforward: either the two arms are shown to be different or they are not. The implementation for the model is not complicated. When more than two arms are included in meta-analysis, complexity ensues. For example, suppose that two treatments A and C are considered in meta-analysis and the treatment C is a control group. A new treatment B is included which can be compared with the control group (C) or a standard active treatment (A). We can obtain the effectiveness of treatment A versus C, treatment B versus C and treatment A versus B. These types of dataset are called *multi-arm trials* although some authors call it *mixed treatment comparison* (MTC) (Lu and Ades, 2004, 2006). Eddy et al. (1992) said of mixed comparisons, 'when there are several interventions that

can be applied to a particular problem, the available evidence can compare different pairs of interventions'. In this thesis, we focus on meta-analysis for multi-arm trials. Two data sets of meta-analysis comparing three arms are given in Tables 2.1 and 2.2 of Chapter 2. We will use both data sets to demonstrate the method we propose.

Some issues have arisen in meta-analysis as follows.

- Direct-indirect comparison: direct comparison exists in treatment comparison but it might not provide enough information for a statistical analysis. We may need to 'borrow strength' from an indirect comparison (Higgins and Whitehead, 1996). This issue will be described in detail in Chapter 3.
- The consistency of multi-arm trials should be considered, particularly, with indirect comparison (Lu and Ades, 2006).
- Analyses in multi-arm trials need a large number of studies to achieve the good results (Green et al., 1997, Chapter 4).

1.4.1 Methods of meta-analysis

We have presented an overview of meta-analysis in the first section and described particular problems such as heterogeneity and selection bias in Section 1.3. In this section, we will review the methods that have been used in meta-analysis of two-arm and multi-arm comparisons.

Pagliaro et al. (1992) used RCTs, comparing beta-blockers or sclerotherapy with a nonactive treatment (control group) to assess the effectiveness of those treatments in the prevention of first bleeding and the reduction of mortality in patients with cirrhosis and esophagogastric varices. The Mantel-Haenszel-Peto method is applied for statistical evaluation of heterogeneity and for pooling of the results. They estimated the treatment effects of beta-blockers

and the control group, sclerotherapy and the control group separately. The results show that no heterogeneity was found and the incidence of bleeding in the use of beta-blockers was significantly reduced.

Indirect comparison has been an important issue of two and multi-arm comparisons. Higgins and Whitehead (1996) presented a random effect meta-analysis for binary data and introduced an idea of 'borrowing strength' from an indirect comparison. A three-arm comparison was also considered in the meta-analysis to improve the inference with both heterogeneity and the treatment difference. Two approaches, namely the *general parameter approach* and the *exact binomial approach*, were used to estimate parameters of interest in a meta-analysis. We apply the idea of 'borrowing strength' in the thesis. Bucher et al. (1997) presented a model for making indirect comparisons of the magnitude of treatment effects that preserved the randomization of the originally assigned patient group. They illustrated the model with an example that compared two experimental prophylactic regimens against the standard prophylaxis for the prevention of pneumocystis carinii pneumonia in HIV infected patients. Similarly, Song et al. (2003) examined the validity of adjusted indirect comparisons by using data from 44 published meta-analyses (from 28 systematic reviews) of RCTs. Lumley (2002) used 'incoherence' in networks of pairwise comparisons to estimate the treatment differences of indirect comparisons. His model is

$$Y_{ijk} \sim N(\mu_i - \mu_j + \eta_{ik} + \eta_{jk} + \xi_{ij}, \sigma^2_{ijk});\ \eta_{ij} \sim N(0, \tau^2),\ \xi_{ij} \sim N(0, \omega^2).$$

The Y_{ijk} is the treatment difference of treatment i and j in the kth randomized trial and its standard error is σ^2_{ijk}. The parameters μ_i and μ_j represent the true average effects of the treatment i and j respectively. Random effects η_{ik} and η_{jk} with variance τ^2 represent the difference between the average effects of treatments i and j and their effects in the study; they capture the heterogeniety of treatment effect. The ξ_{ij} represents the change in the effect of treatment i when it is compared with treatment j and captures the inconsistency.

However, the network needs a large number of different treatment comparisons and it does not guarantee that the conclusions are reliable and generalizable. There is progress in this area: see e.g. Hasselblad (1998); Party et al. (2003); Yazdanpanah et al. (2004) and Glenny et al. (2005) along with texts of Eddy et al. (1992); Whitehead (2002).

Many authors have considered a Bayesian approach to meta-analysis. Domenici et al. (1999) constructed a hierarchical Bayesian grouped random-effect model to synthesis existing evidence from RCTs of which treatments were most effective and of quantifying the remaining uncertainty about treatment effectiveness. They applied their models to migraine headache treatments to incorporate explicitly the relationship between the different classes of treatments and creating a common scale by using a latent variable to combine information from studies that had a difference in results. Ades (2003) introduced the idea of a 'chain of evidence' structure to mixed treatment comparisons by using the Bayesian Markov Chain Monte Carlo (MCMC) method to fit his models. Lu and Ades (2004) proposed a range of Bayesian hierarchical models using the MCMC to represent meta-analysis of multi-arm trials. They extended the Bayesian hierarchical model for two-arm comparisons proposed by Smith et al. (1995) to a general model for multi-arm trials of K-arm comparisons. As mentioned earlier, the consistency of structure evidence of multi-arm trials should be taken into account. Lu and Ades (2006) examined inconsistency using a Bayesian hierarchical model with fixed effects or random effects for fitting multi-arm trials. It is made under the assumption that the available evidence sources were consistent in estimating all treatment contrasts. There is a series of articles attempting to investigate evidence consistency in a variety of different evidence structures, see e.g. Ades and Cliffe (2002); Ades (2003); Welton and Ades (2005). Some issues about the use of Bayesian methods in meta-analysis are related to sensitivity of prior distribution, estimation of posterior distribution, and comparison of classical and Bayesian approaches (Sutton et al., 2000, page 179).

Chootrakool and Shi (2008) propose normal approximation models using an empirical logistic transform to compare different treatments in multi-arm trials, allowing studies of both direct and indirect comparisons. Additionally a hierarchical structure is introduced in the model to address the problem of heterogeneity among different studies. The proposed models are performed with the antiplatelet therapy data.

1.5 Gaussian quadrature approximation

Our approaches in this thesis involve calculation of integrals in the likelihood. We will use Gaussian quadrature approximation to estimate those integrals throughout the thesis. This approximation is a well-known and efficient technique for numerically evaluating integrals of the type $\int_{-1}^{1} f(x)\,dx$ and has been used in many statistical applications. By using Gaussian quadrature, see Abramowitz and Stegun (1972), an approximation of the definite integral of a function $f(x)$ can be given by

$$\int_{-1}^{1} f(x)dx \approx \sum_{n=1}^{l} w_n f(x_n) \tag{1.2}$$

where x_n is a particular node with weight w_n and l is the number of nodes and weights. An l-point Gaussian quadrature rule, named after Carl Friedrich Gauss, is a quadrature rule constructed to yield an exact result for the polynomials of degree $2l-1$, by a suitable choice of the l points and weights. The domain of integration for such a rule is conventionally taken as [-1, 1]. However, the Gaussian quadrature in (1.2) can be expressed in a slightly more general way by introducing a positive weight function g into the integrand and allowing an interval other than [-1, 1]. That is

$$\int_{a}^{b} g(x)f(x)dx, \tag{1.3}$$

where the interval (a, b) and the weight function $g(x)$ can be several choices. For instance, if the interval $(a, b) = (-1, 1)$ and $g(x) = (1 - x^2)^{1/2}$ then this quadrature is called the *Chebyshev-Gauss quadrature*. The details of other choices of (a, b) and $g(x)$ can be found in Abramowitz and Stegun (1972, page 875) and Scheid (1988, page 136).

1.5.1 Gauss-Hermite integration

If the interval (a, b) in (1.3) is equal to $(-\infty, \infty)$ and the weight function $g(x) = e^{-x^2}$ then the quadrature is called *Gauss-Hermite Quadrature*. Gauss-Hermite quadrature is often used for numerical integration in statistics because of its relation to a normal density. The quadrature is defined in term of an integral of the form

$$\int_{-\infty}^{\infty} f(x)e^{-x^2}dx. \tag{1.4}$$

Using Gauss-Hermite quadrature, the integral (1.4) is approximated by $\sum_{n=1}^{l} w_n f(x_n)$, where the nodes x_n are roots of the lth order Hermite polynomial and the w_n are suitably corresponding weights. Tables of (x_n, w_n) for $l = 1, 2 \ldots, 10, 12, 16, 20$ are given by Abramowitz and Stegun (1972, page 924) and for $l > 20$, computation formulae are given by Golub and Welsch (1969). Suppose that a parameter δ is a random effect and approximately distributed by $N(\mu, \tau^2)$ and an integral of Gauss-Hermite quadrature can be in the form of

$$\int_{-\infty}^{\infty} f(\delta)\phi(\delta; \mu, \tau^2)d\delta, \tag{1.5}$$

where $\phi(\delta; \mu, \tau^2)$ is the density function of a normal distribution: $e^{-(\delta-\mu)^2/2\tau^2}/(2\pi)^{1/2}$. The sampling nodes are then at $\delta_n = \mu + 2^{1/2}\tau x_n$ and the weights are modified to $w_n/\sqrt{\pi}$. Using the approximation of Gauss-Hermite quadrature, the integral (1.5) is approximated by

$$\int_{-\infty}^{\infty} f(\delta)\phi(\delta : \mu, \tau^2)d\delta \approx \sum_{n=1}^{l} \frac{w_n}{\sqrt{\pi}} f(\mu + 2^{1/2}\tau x_n). \tag{1.6}$$

Similarly, if the integral (1.5) involves a multivariate normal distribution of $N_k(\boldsymbol{\mu}, \boldsymbol{\Omega})$,

$$\int_{-\infty}^{\infty} f(\boldsymbol{\delta})\phi(\boldsymbol{\delta}; \boldsymbol{\mu}, \boldsymbol{\Omega})d\boldsymbol{\delta}. \tag{1.7}$$

Then, this integral can be approximated by

$$\int_{-\infty}^{\infty} f(\boldsymbol{\delta})\phi(\boldsymbol{\delta}; \boldsymbol{\mu}, \boldsymbol{\Omega})d\boldsymbol{\delta} \approx \pi^{-k/2} \sum_{n_1=1}^{l_1} w_{n_1}^{(1)} \ldots \sum_{n_k=1}^{l_k} w_{n_k}^{(k)} f\left(\boldsymbol{\mu} + \sqrt{2}\boldsymbol{\Omega}^{1/2}\mathbf{d}_n\right).$$

The sampling nodes are at $\boldsymbol{\mu} + \sqrt{2}\boldsymbol{\Omega}^{1/2}\mathbf{d}_n$ and $\mathbf{d}_n = (x_{n_1}^{(1)}, \ldots, x_{n_k}^{(k)})$.

Liu and Pierce (1994) considered Gauss-Hermite quadrature in numerical integration and also examined its effectiveness in Laplace approximation. Crouch and Spiegelman (1990) evaluated the integral form (1.4) to the logistic normal model.

1.6 Outline of the thesis

Earlier in this chapter, we provided an overview of meta-analysis for multi-arm trials and existing methods to make inferences on the treatment effect. Gaussian quadrature approximation has also been described. As reviewed in Section 1.4.1, most existing methods for meta-analysis of multi-arm trials use the logistic regression model with unconditional likelihood approach, see e.g. Lu and Ades (2004, 2006). In this thesis, we propose the normal approximation model using empirical logistic transform (e.g. empirical log-odds ratio model) when the sample size is relatively large and also introduce the logistic regression model with conditional likelihood approach. The trial effects are eliminated in both models, thus our models give a precise estimate and make the computation more stable. More details are given in Chapter 3 and Chapter 4. A main important objective of the thesis is to use a sensitivity analysis with the models by allowing different amounts of selection bias.

Chapter 2 gives a brief introduction to antiplatelet therapy, which has been used for patients with a history of coronary artery disease, heart attacks, angina (chest pain) and peripheral artery disease. Two data sets of RCTs: antiplatelet therapy with maintenance of vascular graft or arterial patency (W1) and antiplatelet therapy with reduction in venous thrombosis and pulmonary embolism (W2), are presented in this chapter.

Chapter 3 first introduces statistically the structure of multi- arm trials. We propose normal approximation models using empirical logistic transform to make inferences on treatment effects of multi-arm comparison. The treatment effect and the trial effect are also explained in detail. The indirect comparison plays an important role in multi-arm trials, particularly if there is little or no evidence from a direct comparison provided in meta-analysis. Our models allow an indirect comparison by using the idea of 'borrowing strength' from indirect comparisons. Additionally, we address the correlation structure of the covariance matrix. The proposed models in this chapter are applied to the W1 data.

Chapter 4 employs the logistic regression model for the exact binomial distribution. Two alternative approaches, based on *unconditional* and *conditional* likelihoods, are performed to estimate the unknown parameters in the model. All treatment effects of the model are assumed to be random and they are normally distributed. This causes the likelihood function to involve integrals. We use Gaussian-Hermite quadrature to approximate the integral. The logistic regression models for both approaches are illustrated with the W2 data.

Chapter 5 investigates the performance of the maximum likelihood estimation (MLE) for the normal approximation model and the logistic regression model using unconditional and conditional approaches with the simulated data. In comparison of the different cases, we exploit two scenarios to generate the data. The simulated data is used to draw inferences on various different models in order to analyse their MLEs. We specially focus an attention on

MLEs for the logistic regression model using the unconditional and conditional approaches.

Chapter 6 begins by describing the funnel plot to identify selection bias in multi-arm trials. We use the normal approximation model for the W1 data as a standard model in this chapter. Our main purpose here is to develop inferences about parameters of interest. We employ the idea of a sensitivity analysis by using a selection model to the normal approximation model, allowing different amounts of selection bias. We then analyze how the parameter of interest changes when compared to the results of normal approximation model. Goodness-of-fit tests are used to check whether taking the selection model into account is appropriate or not for the treatment effect estimates. We also examine the performance of the method for sensitivity analysis by the simulation study.

Finally, Chapter 7 extends the work of Chapter 6 to the logistic regression model using the conditional method. The idea of a selection model in Chapter 6 is adapted to the probability of selection in the likelihood function. This chapter is structured in a similar way to the preceding one.

Chapter 2

Antiplatelet data

2.1 Overview of antiplatelet therapy

Platelets are remnants of cells circulating in the blood that are necessary for blood clots to form. Platelets initiate the formation of blood clots by clumping together, a process called *platelet aggregation*, presented in Figure 2.1. Clumps of platelets are further bound together by a protein (fibrin) formed from clotting factors present in the blood. The clumps of platelets and fibrin make up the blood clot. Blood clots are important because they restrict the amount of bleeding when we get cut. However, if a blood clot forms inside an artery, it can block the flow of blood to the tissue that the artery supplies and can damage the tissue. For example, a blood clot that forms in a coronary artery supplying blood to the heart muscle can cause a heart attack, and a blood clot that forms in an artery supplying blood to the brain can cause a stroke. Antiplatelet drugs are a group of powerful medications that help to prevent the formation of blood clots. They are effective in the arterial circulation, where anticoagulants have little effect. Aspirin is the most widely used antiplatelet drug and is in a group of medications called *salicylates*. Aspirin is cheap and relatively safe, despite a possible side effect of gastric irritation or bleeding. Aspirin is also given to patients with coronary heart disease to reduce the risk of a heart attack. It remains the most commonly used long-term antiplatelet therapy. Other antiplatelet drugs have been introduced such as

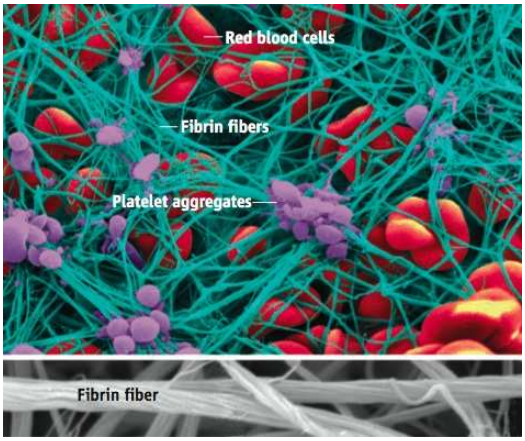


Figure 2.1: Platelet aggregation

ticlopidine and clopidogrel. These have a similar antiplatelet effect of blocking the clotting pathway, though they do this in a slightly different way to aspirin. They seem to have fewer side effects of gastric discomfort or bleeding. Ticlopidine or clopidogrel are prescribed, in the short term, with aspirin for patients undergoing stent implantation with angioplasty, to reduce the extra risk of blood clotting after the procedure. Dipyridamole is often used with other drugs to reduce the risk of blood clots. It was originally introduced in 1959 as an anti-anginal medication: it has coronary vasodilator properties through increasing coronary blood flow without affecting myocardial oxygen consumption. Its effectiveness as an antithrombotic agent was subsequently demonstrated in the rabbit (Emmons et al., 1965). Antiplatelet drugs may be prescribed for patients with a history of: coronary artery disease, heart attacks, angina (chest pain), and peripheral artery disease (PAD). They are often prescribed after angioplasty and stent placement and after heart bypass surgery.

Throughout the thesis, we use two collections of antiplatelet data: antiplatelet therapy with maintenance of vascular graft or arterial patency (W1) given in Table 2.1, and antiplatelet therapy with reduction in venous thrombosis and pulmonary embolism (W2) given in Table 2.2. The W1 data will be applied to the multi-arm trials model using the normal approxi-

mation approach in Chapter 3 and also will be used with a sensitivity analysis in Chapter 6. The multi-arm trials model using exact binomial distribution will be undertaken with the W2 data in Chapter 4 and Chapter 7. Additionally in Chapter 5, the W2 data will be used for generating data to compare the performance of estimations.

2.2 Antiplatelet data: maintaining vascular patency (W1)

After coronary artery revascularisation, whether by coronary artery bypass grafting or by percutaneous transluminal coronary angioplasty, angiographic studies show substantial rates of reocclusion (Gillum, 1987). For example, about one fifth of coronary artery bypass grafts occlude during the first postoperative year (Fuster and Chesebro, 1986) and a few per cent per year occlude thereafter (Campeau et al., 1984). These occlusions are often subclinical, though some may produce clinical signs of myocardial infarction. Occlusion or reocclusion is also seen after peripheral artery revascularisation, though many such occlusions are also subclinical. Experimental and clinical evidence suggests that antiplatelet therapy may help prevent vascular graft or arterial occlusions, particularly during the period soon after vascular procedures, before any intimal damage has healed (Pirk et al., 1990; Bonchek et al., 1982).

Collaboration (1994a) analyzed 46 RCTs of antiplatelet therapy versus the control group and 14 RCTs comparing one antiplatelet regimen with another by setting RCTs that could have been available by March 1990 and in which vascular graft or arterial patency was to be studied systematically. Several treatments are involved in RCTs such as high dose aspirin, medium aspirin, aspirin plus dipyridamole, aspirin alone, sulphinpyrazone, ticlopidine and the control group. The objective is to determine the efficacy of antiplatelet therapy in maintaining vascular patency in patients. The total number of about 8000 patients at varying degrees of risk of vascular occulusion (by virtue of disease or of having some vascular

procedure) were in trials of antiplatelet therapy versus control and 4000 such patients were in trials directly comparing different antiplatelet regimens.

A forest plot (see the detail in Lewis and Clarke, 2001) was used to present the results of the meta-analysis. The treatment effect estimate of each study (odds ratio) and respective confidence interval were plotted on one set of axes. They concluded that antiplatelet therapy (aspirin plus dipyridamole (A) or aspirin alone (B)) produced a highly significant ($2p \leq 0.00001$) reduction in vascular occlusion in a wide range of patients comparing to the control group (C). The odds of vascular graft or arterial occlusion were reduced by about 40% while treatment continued.

Collaboration (1994a) used a forest plot in their systematic review. We will re-analyse the data by using a normal approximation model based on empirical logistic transform in Chapter 3. The problem of selection bias will be addressed in Chapter 6. The data used in this thesis consists of 31 RCTs of three-arm trials. We shall call this data set 'W1'. The studies compare three treatments: aspirin plus dipyridamole (A), aspirin alone (B) and the control group (C). Six trials compare aspirin plus dipyridamole, aspirin alone and the control group (i.e. comparing all A, B and C), four trials compare aspirin plus dipyridamole and aspirin alone (i.e. comparing A and B), thirteen trials compare aspirin plus dipyridamole and the control group (i.e. comparing A and C), and seven trials compare aspirin alone and control group (i.e. comparing B and C). The W1 data is given in Table 2.1. The 'event' in the table represents the number of patients who have reocclusion on those treatments and the 'total' represents the number of patients in total to enter in those groups.

2.3 Antiplatelet data: reduction in venous thrombosis and pulmonary embolism (W2)

During prolonged general anaesthesia or any other period of limited mobility thrombus formation may be initiated in the deep veins of the legs. Specific tests disclose deep venous thrombosis in about a quarter of all patients who have had general surgery and in about half of those who have had orthopaedic surgery (Kakkar, 1981). Most such thromboses are subclinical and resolve completely when mobility is restored (though a few produce permanent valvular damage and chronic venous insufficiency), but some may embolise to the lungs, producing slight, substantial, or fatal effects. Venous thrombosis and pulmonary embolism remain an important cause of morbidity and mortality both in surgical patients and in immobilised medical patients. Various thromboprophylactic treatments have therefore been devised to prevent or limit thromboembolism (Dalen et al., 1986). An overview of randomised trials of perioperative subcutaneous heparin showed that among surgical patients such treatment can roughly halve the risk not only of deep venous thrombosis but, more importantly, of pulmonary embolism. Subcutaneous heparin is now widely recommended for surgical or medical patients at high risk of venous occlusion, but antiplatelet therapy still is not (Gent M., 1986; Collins et al., 1988).

Collaboration (1994b) analysed 53 trials (total 8400 patients) of an average of two weeks of antiplatelet therapy versus control in general or orthopaedic surgery; nine trials (600 patients) of antiplatelet therapy versus control in other types of immobility; 18 trials (1000 patients) of one antiplatelet regimen versus another. Many treatments are involved in RCTs such as high dose aspirin, medium aspirin, aspirin plus dipyridamole, aspirin alone, aspirin plus hydroxychloroquine, ticlopidine and the control group. The objective was to determine the efficacy of antiplatelet therapy as prophylaxis against deep venous thrombosis or pulmonary embolism in surgical and high risk medical patients. It had previously been

supposed that antiplatelet therapy did not influence venous thromboembolism, and many surgeons and physicians do not use it routinely for thromboprophylaxis, even for patients who are at substantial risk of deep venous thrombosis or pulmonary embolism.

Collaboration (1994b) used a forest plot to present the results of the meta-analysis. They concluded that antiplatelet therapy - either alone or, for greater effect, in addition to other proved forms of thromboprophylaxis (such as subcutaneous heparin) - should be considered. Also antiplatelet therapy produced a highly significant ($2p \leq 0.00001$) reduction in deep venous thrombosis by about 67%.

As shown in Table 2.2, the sample sizes for many studies are quite small. An exact logistic regression model will therefore be used with both unconditional likelihood approach, see the details in Chapter 4. In the thesis, we will investigate 27 RCTs from systematic reviews of Antiplatelet Trialists' Collaboration (Collaboration, 1994b) in total. We shall call this data set 'W2'. The studies compare three treatments: aspirin plus dipyridamole (A), aspirin alone (B) and control group (C), where seven trials compare aspirin plus dipyridamole, aspirin alone and control group (i.e. comparing all A, B and C), ten trials compare aspirin plus dipyridamole and control group (i.e. comparing A and C) and ten trials compare aspirin alone and control group (i.e. comparing B and C) . The W2 data is given in Table 2.2. The 'event' in the table represents the number of patients in whom deep venous thrombosis was detected by systematic fibrinogen scans or venography, or both, after general and orthopaedic surgery and in high risk medical patients. The 'total' represents the number of patients controlled in each group.

Table 2.1: The W1 data: 31 RCTs of aspirin data

Study number	Number of patients		
	Aspirin + Dipyridamole (A) event/total	Aspirin (B) event/total	Control (C) event&total
1	15/49	10/47	18/51
2	35/162	37/155	47/153
3	83/368	85/373	114/371
4	23/100	16/100	39/100
5	6/16	2/16	12/17
6	0/100	6/100	12/100
7	20/60	22/64	
8	26/313	27/317	
9	10/41	6/40	
10	8/55	15/55	
11	33/160		37/160
12	37/202		81/205
13	4/18		9/30
14	17/62		20/63
15	8/61		24/64
16	13/47		27/46
17	21/34		14/35
18	11/72		15/68
19	6/187		13/189
20	86/286		86/263
21	4/33		15/32
22	15/50		12/50
23	7/22		19/31
24	15/132		13/67
25		15/71	16/71
26		6/29	15/31
27		7/68	17/69
28		24/215	47/213
29		19/148	28/150
30		6/19	18/25
31		2/47	11/45

Table 2.2: The W2 data: 27 RCTs of aspirin data

Study number	Number of patients		
	Aspirin + Dipyridamole (A) event/total	Aspirin (B) event/total	Control (C) event&total
1	3/31	7/30	13/35
2	6/12	6/9	4/9
3	3/30	9/32	13/34
4	0/100	4/100	5/100
5	6/18	8/16	8/25
6	1/11	2/10	4/11
7	0/11	2/14	1/14
8	13/75		35/75
9	12/85		24/75
10	3/38		14/66
11	1/30		11/36
12	20/32		21/32
13	10/20		8/20
14	8/21		8/22
15	3/13		6/15
16	1/19		7/19
17	6/40		14/40
18		42/153	33/150
19		5/702	11/679
20		9/56	11/49
21		9/357	32/357
22		16/50	12/50
23		7/138	17/140
24		27/66	29/63
25		16/44	20/44
26		7/26	4/25
27		11/58	23/59

Chapter 3

Meta-analysis of multi-arm trials using normal approximation approach

3.1 Introduction

As described in Chapter 1, in standard two-arm comparison, evidences from two treatments have been combined directly in meta-analysis. In multi-arm trials, we aim to summarize the studies providing more than two arms to estimate the overall treatment effects from the pair-wise treatment comparison. Some studies in multi-arm trials might give useful information on indirect comparison in a situation where the treatments have not been directly compared to the control group. Treatment comparisons in meta-analysis have been divided into two types (Glenny et al., 2005). One is to compare two treatments directly, called *direct comparison*, or *head-to-head comparison*. The other is to use information from *indirect comparisons*. For example, from antiplatelet data given in Table 2.1 of Chapter 2, there are three treatment comparisons available: treatments A, B and C; the control group of meta-analysis is treatment C. Three groups of studies compare treatment A versus C, treatment B versus C, and treatment A versus B, respectively. If our aim is to compare treatment A versus B then the studies comparing treatment A versus C and treatment B versus C provide the

indirect comparison for treatment A versus B. The direct and indirect comparisons for RCTs in meta-analysis have been explored by several authors (Bucher et al., 1997; Lumley, 2002; Song et al., 2003; Lu and Ades, 2004, 2006). This chapter proposes the model for multi-arm trials approximated by a normal approximation model (Chootrakool and Shi, 2008).

The chapter is organized as follows. We begin by introducing the data structure of multi-arm trials in Section 3.2. Section 3.3 discusses the normal approximation model using the empirical logistic transform. The model on a log-odds scale is performed in Section 3.4, and the direct and indirect comparisons are given. Section 3.5 describes the model on a log-odds ratio scale including both comparisons. The maximum likelihood method and its properties are illustrated in Section 3.6. We give the standard errors of MLEs in Section 3.7. In Section 3.8, the proposed models in the chapter are applied with the W1 data, given in Chapter 2. The last section concludes the ideas of this chapter and gives some comments.

3.2 The data structure of multi-arm trials

Suppose that M RCTs of a meta-analysis make multi-arm comparisons between $K+1$ treatments. The indices $i = 1, \ldots, M$ and $j = 0, 1 \ldots, K$ stand for the studies and the treatments respectively, where the index $j = 0$ stands for the control group. For the ith study, let r_{ij} represent the number of an unsuccessful outcome on treatment j and let n_{ij} denote the number of observation in the corresponding group. Let π_{ij} be the probability of an unsuccessful outcome of a patient given the treatment j (treated as a failure) in the ith study. The r_{ij} has a binomial distribution

$$r_{ij} \sim Bin(\pi_{ij}, n_{ij}); \;\; i = 1, \ldots, M \text{ and } j = 0, 1 \ldots, K. \tag{3.1}$$

Some studies might not have all the treatments available. For example, from the W1 data, treatment C is not available in the studies 7 - 10. The data structure is analogous to an

incomplete-blocks design, which has been investigated by several authors: Scheff'e (1959, page 161), Pocock (1989, page 121) and Hinkelmann and Kempthorne (1994, page 290). To define a data structure of multi-arm trials, we shall introduce an index set J_i comprising the treatments involved in the ith study. The data structure of multi-arm trials is represented as

$$\mathcal{D} = \{(r_{ij},\ n_{ij}) : i = 1,\ ...,\ M;\ j \in J_i\}\,. \tag{3.2}$$

3.3 Normal approximation model based on empirical logistic transform

According to the binomial distribution (3.1), the mean and variance of r_{ij} are $n_{ij}\pi_{ij}$ and $n_{ij}\pi_{ij}(1-\pi_{ij})$ respectively. An important property of the binomial distribution is that as the number of observation n_{ij} increases, the degree of asymmetry in the distribution decreases and also the binomial distribution becomes more closely approximated by the normal distribution (Collett, 1991, page 20). Let $\psi(x)$ be the function $\log(x/1-x)$ and let δ_{ij} be the parameter of interest, given by $\delta_{ij} = \psi(\pi_{ij})$. From Cox (1970, page 31) if n_{ij} is large and π_{ij} is not too near 0 or 1, we substitute π_{ij} by r_{ij}/n_{ij} in $\psi(\pi_{ij})$. Then the δ_{ij} is reasonably estimated by

$$X_{ij} = \psi(r_{ij}/n_{ij}) = \log\left(\frac{r_{ij}}{n_{ij} - r_{ij}}\right), \tag{3.3}$$

which is nearly normally distributed and we call X_{ij} the *empirical logistic transform of* (r_{ij}, n_{ij}). As n_{ij} approaches infinity, the asymptotic mean and variance are respectively

$$E(X_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) \qquad \text{and} \qquad Var(X_{ij}) = \frac{n_{ij}}{r_{ij}(n_{ij} - r_{ij})}.$$

Modifying the transformation, the empirical logistic transform X_{ij} and $Var(X_{ij})$ need modification only if $r_{ij} = 0$ or n_{ij} when the logistic transform in (3.3) is undefined. With extensive

data, occasional extreme values of r_{ij} are to be expected, even if on the whole the conditions for large-sample theory apply. Haldane and Smith (1948) and Anscombe (1956) proposed a transform defined by

$$X_{ij}(a) = \log\left(\frac{r_{ij}+a}{n_{ij}-r_{ij}+a}\right). \tag{3.4}$$

The idea is to choose the constant a so that the expected value of (3.4) is as nearly as possible $\delta_{ij} = \log(\pi_{ij}/(1-\pi_{ij}))$. As a result an appropriate choice of a is 1/2. We then have the empirical logistic transform as

$$X_{ij} = \log\left(\frac{r_{ij}+0.5}{n_{ij}-r_{ij}+0.5}\right). \tag{3.5}$$

The asymptotic mean and variance are respectively

$$E(X_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) \quad \text{and} \quad Var(X_{ij}) = \frac{n_{ij}+1}{(r_{ij}+0.5)(n_{ij}-r_{ij}+0.5)}. \tag{3.6}$$

3.4 Empirical log-odds model

In clinical trials without a control treatment, it is impossible to be sure that any response is due solely to the effect of the treatment and the importance of a new treatment can be over-stated. Thus the control treatment may be the standard treatment (a positive control treatment) or, if one does not exist, may be a negative control treatment, which can be a placebo (a treatment which looks and tastes like the new drug but which does not contain any active compound) (Petrie and Sabin, 2005, page 34). The control treatment corresponding to each study shall be called the '*baseline treatment*'. In a meta-analysis, more than one studies are combined so it is possible to have more than one baseline treatment in the meta-analysis. In comparing in multi-arm trials, we can have only one control treatment in a meta-analysis, thus we shall call the control treatment for a meta-analysis '*control group*'.

This section presents the model using the empirical logistic transform and based on the requirement of (r_{ij}, n_{ij}) that n_{ij} is large (larger than 20) and r_{ij} is not too small (near 0) and not too close to n_{ij}. We start with a special case of the model. Then the general model will be explained including the direct and indirect comparisons.

3.4.1 Meta-analysis of multi-arm trials

We first define a model for a special case in which each of the M studies includes all $K+1$ treatments. For this special case, the control group of meta-analysis and the baseline treatment for all studies are treatment '0'. There is a direct comparison only in this meta-analysis. Suppose that r_{i0} and r_{ij} have binomial distributions $Bin(n_{i0}, \pi_{i0})$ and $Bin(n_{ij}, \pi_{ij})$ respectively for $j = 1, \ldots, K$. The data structure is given in (3.2) where the set J_i for all M studies is $\{0, \ldots, K\}$. For the ith study, let X_{i0} and X_{ij} be the empirical logistic transforms (or *empirical log-odds*) for (r_{i0}, n_{i0}) and (r_{ij}, n_{ij}) respectively, as defined in (3.5). Based on the discussion in Section 3.3, normal approximation models for X_{i0} and X_{ij} on the log-odds scale can be defined by

$$X_{i0} = \alpha_i + \sigma_{i0}\epsilon_{i0}, \tag{3.7}$$

$$X_{ij} = \alpha_i + \delta_{i,0j} + \sigma_{ij}\epsilon_{ij}, \qquad j = 1, \ldots, K. \tag{3.8}$$

They are called an *empirical log-odds model*. The parameters σ_{i0}^2 and σ_{ij}^2 are the variances of X_{i0} and X_{ij} respectively, approximated from (3.6). The parameters ϵ_{i0} and ϵ_{ij} are independent, follow the standard normal distributions and correspond to the random sampling errors of the models X_{i0} and X_{ij} respectively. The random sampling errors ($\sigma_{i0}\epsilon_{i0}$ and $\sigma_{ij}\epsilon_{ij}$) are therefore independent and normally distributed as $N(0, \sigma_{i0}^2)$ and $N(0, \sigma_{ij}^2)$ respectively. The α_i in both models are the trial effects representing the difference across studies. The $\delta_{i,0j}$ is a parameter of interest, which is the treatment effect between the control group and

treatment j in the ith study. It is obtained from $\delta_{i,0j} = \delta_{ij} - \delta_{i0}$, called the log-odds ratio between treatment j and the control group.

Trial effect

Two assumptions are usually made about the trial effect α_i. The first one is that the trial effects are assumed to be study-level effects, which means that the α_is are different parameters and are treated as nuisance parameters in the model. We need to include M different unknown parameters in the model. The second one is that we may assume a model for the α_i's. A special case is to assume that the trial effect is a fixed effect, defined by $\alpha_i = \alpha_0$. Conversely, it may be assumed to be a random effect, given by $\alpha_i \sim N(\mu_{\alpha 0}, \tau^2_{\alpha 0})$, where $\mu_{\alpha 0}$ is the overall mean of the trial effect and $\tau_{\alpha 0}$ measures the magnitude of the variation between the studies. To capture skewness and heavy tails in the distribution of the trial effect, a mixture of normal distributions may be used, see Domenici et al. (1999). However, in practice the trial effects in most meta-analysis would not satisfy any model since different experiment designs and different data analysis models are used in different studies. Most of the existing methods therefore used the first assumption. However, the number of unknown parameters (for the trial effect) is the same as the number of studies if the first assumption of the trial effect is used. This will result in some theoretical and computational problems. The accuracy of the estimation depends on the sample size of each study not the overall sample size of the pool in the meta-analysis. The estimates of some parameters may not be consistent, see Lubin (1981). Due to the large number of parameters, the computation is usually unstable. We therefore propose the empirical log-odds ratio model in Section 3.5.

Treatment effect

The treatment effect can be assumed to be a fixed effect or a random effect. The fixed effect is defined as $\delta_{i,0j} = \mu_{0j}$, where μ_{0j} is a fixed treatment effect between the control group and treatment j for all studies. There are several different ways to deal with the random

effects, for example, see DerSimonian and Laird (1986). The treatment effect is assumed to be random and normally distributed as $\delta_{i,0j} \sim N(\mu_{0j}, \tau_{0j}^2)$

From the models (3.7) and (3.8), we shall assume that the trial effect follows the first assumption and the treatment effect is a random effect, i.e. all the α_is are different parameters and the treatment effect is a random effect $N(\mu_{0j}, \tau_{0j}^2)$. The treatment effects $\delta_{i,0j}$ and $\delta_{i,0k}$ for $j \neq k$ and $j, k \in \{1, \ldots, K\}$ may be dependent. This is because they involve δ_{i0} in the same way; thus the covariance between the treatment effects $\delta_{i,0j}$ and $\delta_{i,0k}$ are not equal to zero ($Cov(\delta_{i,0j}, \delta_{i,0k}) \neq 0$). Let ρ_{jk} be the correlation coefficient between each pair $(\delta_{i,0j}, \delta_{i,0k})$ for $j \neq k$ and $j, k \in \{1, \ldots, K\}$. The treatment effects $\delta_{i,0j}$, for $j = 1, \ldots, K$ in the ith study are therefore modelled by the following multivariate normal distribution,

$$\begin{pmatrix} \delta_{i,01} \\ \delta_{i,02} \\ \vdots \\ \delta_{i,0K} \end{pmatrix} \sim MVN \left(\begin{pmatrix} \mu_{01} \\ \mu_{02} \\ \vdots \\ \mu_{0K} \end{pmatrix}, \begin{pmatrix} \tau_{01}^2 & \rho_{12}\tau_{01}\tau_{02} & \cdots & \rho_{1K}\tau_{01}\tau_{0K} \\ \rho_{12}\tau_{01}\tau_{02} & \tau_{02}^2 & \cdots & \rho_{2K}\tau_{02}\tau_{0K} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1K}\tau_{01}\tau_{0K} & \rho_{2K}\tau_{02}\tau_{0K} & \cdots & \tau_{0K}^2 \end{pmatrix} \right). \quad (3.9)$$

The μ_{0k} is the overall mean effect between the control group and the treatment k. The τ_{0k}^2 is a measure of between-study heterogeneity of the treatment effect $\delta_{i,0k}$. The correlation coefficient ρ_{jk} measures the amount of linear association between the $\delta_{i,0j}$ and the $\delta_{i,0k}$. Also the $\rho_{jk}\tau_{0j}\tau_{0k}$ is the covariance between the treatment effects $\delta_{i,0j}$ and $\delta_{i,0k}$. From (3.9), the entries on the diagonal of the covariance matrix are often called the *heterogeneity parameters* of the treatment effects. The heterogeneity parameter measures the variation in the treatment effect between studies. If there is a very little variation between studies then a fixed effect may be appropriate for the treatment effect. The useful properties of the model parameterisation are the correlation structure of the covariance matrix:

1. An important special case is that the heterogeneity parameters of the treatment effects are assumed to be the same, called *homogeneity of variances*. The correlation coeffi-

cients between each pair $(\delta_{i,0j}, \delta_{i,0k})$, for $j \neq k$, and $j, k \in \{1, \ldots, K\}$ are equal and take the value 1/2 because the treatment effects $\delta_{i,0j}$ and $\delta_{i,0k}$ involve $\log(\pi_{i0}/1 - \pi_{i0})$ in the same way. The covariance matrix in (3.9) for this assumption is

$$\tau^2 \begin{pmatrix} 1 & \rho & \ldots & \rho \\ \rho & 1 & \ldots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \ldots & 1 \end{pmatrix}, \quad \text{where } \rho = 1/2.$$

2. The above assumption may not be reasonable in some applications. We thus allow the heterogeneity parameters of the treatment effects to be different for each treatment effect, called *heterogeneity of variances*. The covariance matrix will be in the standard form as shown in (3.9).

3.4.2 Meta-analysis of multi-arm trials with both direct and indirect comparisons

In some circumstances, a meta-analysis may contain different information to the special case. For example, some studies might compare fewer than $K + 1$ treatments, or some baseline treatments may be different, or both cases could occur simultaneously. We shall propose a general model adapted from the special case described in the previous section. Let $b(i)$ denote the baseline treatment corresponding to the ith study, which can be the control group or any other treatments. As mentioned earlier about indirect comparison, in a situation that the treatments in some studies can not be compared directly to the control group, we need to use evidence from the external studies. To make it clear, if $b(i) = 0$ then the direct comparison is involved in this study. Conversely, if $b(i) \neq 0$ then the study makes indirect comparison. Let $J_{(i)} = J_i \setminus \{b(i)\}$ represent the set of treatments involved in the ith study but excluding the baseline treatment $b(i)$. Let k_i and $k_i + 1$ denote the number of

treatments in the sets $J_{(i)}$ and J_i respectively. The $r_{ib(i)}$ and r_{ij} are binomially distributed as $Bin(n_{ib(i)}, \pi_{ib(i)})$ and $Bin(n_{ij}, \pi_{ij})$ for $j \in J_{(i)}$ respectively. The empirical log-odds models for the general case in the ith study are defined as

$$X_{ib(i)} = \alpha_i + \sigma_{ib(i)}\epsilon_{ib(i)}, \tag{3.10}$$

$$X_{ij} = \alpha_i + \delta_{i,b(i)j} + \sigma_{ij}\epsilon_{ij}, \qquad j \in J_{(i)}. \tag{3.11}$$

These models can be used for both comparisons. According to above discussion, let D and I be sets of studies that make the direct and indirect comparisons respectively. The assumptions of the trial effect and the treatment effect are similar to the special case (assumed to be different parameters and random effect respectively). The treatment effect $\delta_{i,b(i)j}$ in (3.11) can be direct treatment effect if $i \in D$ or indirect treatment effect if $i \in I$: they are defined as follows.

$$\delta_{i,b(i)j} = \begin{cases} \delta_{i,0j} \sim N(\mu_{0j}, \tau_{0j}^2) & \text{if } i \in D, \\ \delta_{i,0j} - \delta_{i,0b(i)} \sim N(\mu_{0j} - \mu_{0b(i)}, \tau_{0j}^2 + \tau_{0b(i)}^2 - 2\rho_{jb(i)}\tau_{0j}\tau_{0b(i)}) & \text{if } i \in I. \end{cases} \tag{3.12}$$

where $\rho_{jb(i)}$ is the correlation coefficient between $\delta_{i,0j}$ and $\delta_{i,0b(i)}$. For example, from the W1 data, suppose the treatment A, B, C represent aspirin plus dipyridamole, aspirin alone and control group respectively. The baseline treatment for the studies 7-10 is B thus the indirect treatment effect can be written as

$$\delta_{i,AB} = \delta_{i,AC} - \delta_{i,BC} \sim N(\mu_{AC} - \mu_{BC}, \tau_{AC}^2 + \tau_{BC}^2 - 2\rho_{AB}\tau_{AC}\tau_{BC}),\ i = 7, \ldots, 10.$$

Next, we shall consider the treatment effect in a matrix form, of which will be in the form of an index vector and the treatment effect model from the special case. From the treatment effect model (3.9), let $\boldsymbol{\delta}_{i,0}$ and $\boldsymbol{\mu}_0$ represent the vectors of $(\delta_{i,0j}, j = 1, \ldots, K)^t$ and $(\mu_{0j}, j = 1, \ldots, K)^t$ respectively where the superscript t stands for matrix transposition and let $\boldsymbol{\Omega}_0$

represent the $K \times K$ covariance matrix. The model (3.9) can be written as

$$\boldsymbol{\delta}_{i,0} \sim MVN(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0). \tag{3.13}$$

This is called the *basic model of random treatment effect*. Let $\mathbf{F}_{ij}$ be the index vector of length K consisting of elements 0 and 1 corresponding to $\delta_{i,b(i)j}$, given by

$$\mathbf{F}_{ij} = \begin{cases} (0, \ldots, \underbrace{0}_{b(i)th}, \ldots \underbrace{1}_{jth}, \ldots, 0) & \text{if } i \in D, \\ (0, \ldots, \underbrace{-1}_{b(i)th}, \ldots, \underbrace{1}_{jth}, \ldots, 0) & \text{if } i \in I. \end{cases} \tag{3.14}$$

Now, the random effect $\delta_{i,b(i)j}$ can be written in the form of (3.13) and (3.14):

$$\delta_{i,b(i)j} = \mathbf{F}_{ij}\boldsymbol{\delta}_{i,0} \sim N(\mathbf{F}_{ij}\boldsymbol{\mu}_0, \mathbf{F}_{ij}\boldsymbol{\Omega}_0\mathbf{F}_{ij}^t). \tag{3.15}$$

As before, the covariance between the treatment effects $\delta_{i,b(i)j}$ and $\delta_{i,b(i)k}$ for $j \neq k$ and $j, k \in J_{(i)}$ may be dependent. For the ith study, let $\mathbf{F}_i$ be the following $k_i \times K$ matrix

$$\mathbf{F}_i = (\mathbf{F}_{ij})_{k_i \times K}, \qquad \text{for } j \in J_{(i)}, \tag{3.16}$$

where $\mathbf{F}_{ij}$ is as defined in (3.14). Let $\boldsymbol{\delta}_i$ denote the vector $(\delta_{i,b(i)j}, j \in J_{(i)})^t$ then we have

$$\boldsymbol{\delta}_i = \mathbf{F}_i\boldsymbol{\delta}_{i,0} \sim MVN(\boldsymbol{\mu}_i, \boldsymbol{\Omega}_i), \tag{3.17}$$

where

$$\boldsymbol{\mu}_i = \mathbf{F}_i\boldsymbol{\mu}_0 \qquad \text{and} \qquad \boldsymbol{\Omega}_i = \mathbf{F}_i\boldsymbol{\Omega}_0\mathbf{F}_i^t. \tag{3.18}$$

Referring to the assumptions of covariance matrix $\boldsymbol{\Omega}_0$ in the previous subsection, the correlation structure of $\boldsymbol{\delta}_i$ can be considered accordingly. More discussion will be given in Section 3.8.

3.5 Empirical log-odds ratio model

To avoid the problem of many nuisance parameters and inconsistent estimate, the trial effects can be eliminated from the empirical log-odds models by using the empirical log-odds model on the log-odds ratio scale. Those models in Section 3.4.1 and 3.4.2 are considered here as following subsections.

3.5.1 Meta-analysis of multi-arm trials

Let $Y_{i,0j}$ be the empirical log-odds ratio between (r_{ij}, n_{ij}) for $j = 1, \ldots, K$ and (r_{i0}, n_{i0}). This can be written as $Y_{i,0j} = X_{ij} - X_{i0}$. According to the empirical log-odds models (3.7) and (3.8) in the special case, they can be defined on the log-odds ratio scale as

$$Y_{i,0j} = \delta_{i,0j} + \sigma_{i,0j}\epsilon_{i,0j}, \qquad j = 1 \ldots, K. \tag{3.19}$$

We shall call this an *empirical log-odds ratio model*. Notice that the trial effect is eliminated in the model. The $\delta_{i,0j}$ is a random treatment effect defined in (3.9). The variance $\sigma^2_{i,0j}$ is obtained from a summation of $\widehat{\sigma^2_{i0}}$ and $\widehat{\sigma^2_{ij}}$. For notational convenience, let $e_{i,0j}$ denote a random sampling error $\sigma_{i,0j}\epsilon_{i,0j}$ for the model $Y_{i,0j}$ and normally distributed as $N(0, \sigma^2_{i,0j})$. The model can be written as $Y_{i,0j} = \delta_{i,0j} + e_{i,0j}$. The $e_{i,0j}$ and $e_{i,0k}$ are not independent for $j \neq k$ and $j, k \in \{1, \ldots, K\}$, derived as

$$Cov(e_{i,0j}, e_{i,0k}) = Cov(X_{ij} - X_{i0}, X_{ik} - X_{i0}) = Var(X_{i0}) = \sigma^2_{i0}. \tag{3.20}$$

The random sampling errors $e_{i,0j}$ are distributed as a multivariate normal distribution, given by

$$\begin{pmatrix} e_{i,01} \\ e_{i,02} \\ \vdots \\ e_{i,0K} \end{pmatrix} \sim MVN \left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{i,01} & \sigma^2_{i0} & \dots & \sigma^2_{i0} \\ \sigma^2_{i0} & \sigma^2_{i,02} & \dots & \sigma^2_{i0} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{i0} & \sigma^2_{i0} & \dots & \sigma^2_{i,0K} \end{pmatrix} \right), \tag{3.21}$$

where $\sigma^2_{i,0j} = Var(Y_{i,0j}|\delta_{i,0j}) = \sigma^2_{i0} + \sigma^2_{ij}$. If we assume a random effect model for $\delta_{i,0j}$ as given in (3.9), the empirical log-odds ratio model for the ith study is the following multivariate normal distribution:

$$\begin{pmatrix} Y_{i,01} \\ Y_{i,02} \\ \vdots \\ Y_{i,0K} \end{pmatrix} \sim MVN \left(\begin{pmatrix} \mu_{01} \\ \mu_{02} \\ \vdots \\ \mu_{0K} \end{pmatrix}, \begin{pmatrix} \tau^2_{01} + \sigma^2_{01} & \rho_{12}\tau_{01}\tau_{02} + \sigma^2_{i0} & \dots & \rho_{1K}\tau_{01}\tau_{0K} + \sigma^2_{i0} \\ \rho_{12}\tau_{01}\tau_{02} + \sigma^2_{i0} & \tau^2_{02} + \sigma^2_{02} & \dots & \rho_{2K}\tau_{02}\tau_{0K} + \sigma^2_{i0} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1K}\tau_{01}\tau_{0K} + \sigma^2_{i0} & \rho_{2K}\tau_{02}\tau_{0K} + \sigma^2_{i0} & \dots & \tau^2_{0K} + \sigma^2_{0K} \end{pmatrix} \right). \tag{3.22}$$

The μ_{0k} is the overall mean effect between the control group and the treatment k obtaining from the mean of the treatment effect. The term $\tau^2_{0k} + \sigma^2_{0k}$ is the variance of $Y_{i,0k}$. The term $\rho_{jk}\tau_{0j}\tau_{0k} + \sigma^2_{i0}$ is the covariance between $Y_{i,0j}$ and $Y_{i,0k}$ where $j \neq k$ and $j, k \in \{1, \dots, K\}$.

3.5.2 Meta-analysis of multi-arm trials with both direct and indirect comparison

As in the previous subsection, let $Y_{i,b(i)j}$ be the empirical logistic transform between (r_{ij}, n_{ij}) and $(r_{ib(i)}, n_{ib(i)})$. The empirical log-odds models (3.10) and (3.11) can be defined on the log-odds ratio scale by

$$Y_{i,b(i)j} = \delta_{i,b(i)j} + \sigma_{i,b(i)j}\, \epsilon_{i,b(i)j}, \qquad j \in J_{(i)}. \tag{3.23}$$

The variance of $\sigma^2_{i,b(i)j}$ is approximated by $\widehat{\sigma^2_{ib(i)}} + \widehat{\sigma^2_{ij}}$. The random treatment effect model for $\delta_{i,b(i)j}$ is given in (3.15). As defined in the previous section, let $e_{i,b(i)j}$ represent the random sampling error $\sigma_{i,b(i)j}\epsilon_{i,b(i)j}$. The $e_{i,b(i)j}$ can be given in the form of the index vector and random sampling errors model. From (3.21), let $\mathbf{e}_{i,0}$ be the vector $(e_{i,0j}, j = 1, \ldots, K)^t$ and let $\boldsymbol{\Sigma}_{i,0}$ be the $K \times K$ covariance matrix. The model for random sampling errors given in (3.21) can then be rewritten as

$$\mathbf{e}_{i,0} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_{i,0}). \tag{3.24}$$

We call it as a *basic model of random sampling errors*. Using the index matrix defined in (3.14), the random sampling error $e_{i,b(i)j}$ is taken in the form of

$$e_{i,b(i)j} = \mathbf{F}_{ij}\mathbf{e}_{i,0} \sim N(\mathbf{0}, \mathbf{F}_{ij}\boldsymbol{\Sigma}_{i,0}\mathbf{F}^t_{ij}). \tag{3.25}$$

Let $\mathbf{e}_i$ be the vector $(e_{i,b(i)j}, j \in J_{(i)})^t$. From (3.21), we have

$$\mathbf{e}_i = \mathbf{F}_i\mathbf{e}_{i,0} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_i), \tag{3.26}$$

where $\mathbf{F}_i$ is given in (3.16) and $\boldsymbol{\Sigma}_i = \mathbf{F}_i\boldsymbol{\Sigma}_{i,0}\mathbf{F}^t_i$. Similarly, let $\mathbf{Y}_{i,0}$ be the vector $(Y_{i,0j}, j = 1, \ldots, K)^t$. The *basic model for empirical log-odds ratio model* (3.22) can be defined as

$$\mathbf{Y}_{i,0} \sim MVN(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0 + \boldsymbol{\Sigma}_{i,0}), \tag{3.27}$$

and the model (3.23) can be defined by

$$Y_{i,b(i)j} = \mathbf{F}_{ij}\mathbf{Y}_{i,0} \sim N(\boldsymbol{\mu}_i, \mathbf{F}_{ij}\boldsymbol{\Omega}_0\mathbf{F}^t_{ij} + \mathbf{F}_{ij}\boldsymbol{\Sigma}_{i,0}\mathbf{F}^t_{ij}). \tag{3.28}$$

Let $\mathbf{Y}_i$ be the vector $(Y_{i,b(i)j},\ j \in J_{(i)})^t$, which may be written as $\mathbf{Y}_i = \mathbf{F}_i \mathbf{Y}_{i,0}$. In matrix notation, the model (3.23) is

$$\mathbf{Y}_i = \boldsymbol{\delta}_i + \mathbf{e}_i \sim MVN(\boldsymbol{\mu}_i, \mathbf{V}_i), \tag{3.29}$$

where $\boldsymbol{\delta}_i$ and $\mathbf{e}_i$ are given in (3.17) and (3.26) respectively. The $\boldsymbol{\mu}_i$ is given in (3.18) and the covariance matrix $\mathbf{V}_i = \boldsymbol{\Omega}_i + \boldsymbol{\Sigma}_i$.

3.6 Maximum likelihood estimation

From model (3.29), $\mathbf{Y}_i$ is distributed as a multivariate normal distribution $MVN(\boldsymbol{\mu}_i, \mathbf{V}_i)$. The probability density function for $\mathbf{Y}_i$ is in the form

$$p(\mathbf{Y}_i) = \frac{1}{(2\pi)^{k_i/2} \left|\mathbf{V}_i\right|^{1/2}} e^{-(\mathbf{Y}_i - \boldsymbol{\mu}_i)' \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_i)/2}. \tag{3.30}$$

We aim to estimate the unknown parameters for the meta-analysis consisting of M studies. Let $\boldsymbol{\theta}$ be the collection of all unknown parameters of $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$. Suppose that $\boldsymbol{\theta}$ can take any value within an admissible range Θ. Let $\mathbf{Y}$ denote the collection $\mathbf{Y}_i$ for $i = 1, \ldots, M$. The likelihood function for the meta-analysis is defined as $L(\boldsymbol{\theta}|\mathbf{Y})$, taking the form

$$L(\boldsymbol{\theta}|\mathbf{Y}) = \frac{1}{\prod_{i=1}^{M} (2\pi)^{k_i^2/2} \left|\mathbf{V}_i\right|^{k_i/2}} e^{-\sum_{i=1}^{M} (\mathbf{Y}_i - \boldsymbol{\mu}_i)' \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_i)/2}. \tag{3.31}$$

The method of maximum likelihood (ML) is to find the value $\widehat{\boldsymbol{\theta}}$ within Θ which maximises the likelihood functions of $\boldsymbol{\theta}$. In other words

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta}|\mathbf{Y}).$$

This is the maximum likelihood estimator of $\boldsymbol{\theta}$. The likelihood function $L(\boldsymbol{\theta}|\mathbf{Y})$ represents the joint probability, or likelihood of observing data that has been collected in the meta-

analysis. The term *joint probability* means a probability that combines the contributions of all the studies in the meta-analysis. Let l_n stand for the log-likelihood function in the normal approximation model based on the empirical logistic transform. The MLE is usually determined by maximizing the log-likelihood function $l_n(\boldsymbol{\theta}|\mathbf{Y}) = \log L(\boldsymbol{\theta}|\mathbf{Y})$. Differentiating $l_n(\boldsymbol{\theta}|\mathbf{Y})$ with respect to $\boldsymbol{\theta}$, termed as *a score function*, gives

$$U(\boldsymbol{\theta}) = \frac{\partial l_n(\boldsymbol{\theta}|\mathbf{Y})}{\partial \boldsymbol{\theta}}.$$

By setting the score function to zero and solving for $\boldsymbol{\theta}$, the MLE $\widehat{\boldsymbol{\theta}}$ can be obtained.

3.7 Standard error of parameter estimation

Following the estimation of the unknown parameters in the empirical log-odds ratio model, suppose that m unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ are in the set $\boldsymbol{\theta}$ of a meta-analysis. The m derivatives of the log-likelihood function with respect to $\theta_1, \theta_2, \ldots,$ and θ_m are called the *efficient scores*, whose jth component is $\partial l_n(\boldsymbol{\theta}|\mathbf{Y})/\partial\theta_j$ for $j = 1, 2, \ldots, m$. Now let $\mathcal{H}(\boldsymbol{\theta})$ be the $m \times m$ matrix of second partial derivatives of $l_n(\boldsymbol{\theta}|\mathbf{Y})$,where the (j,k)th entry of $\mathcal{H}(\boldsymbol{\theta})$ is

$$\frac{\partial^2 l_n(\boldsymbol{\theta}|\mathbf{Y})}{\partial\theta_j \partial\theta_k},$$

for $j = 1, 2, \ldots, m$ and $k = 1, 2, \ldots, m$. The *observed Fisher information* (Palmgren, 1981) $\mathcal{I}(\boldsymbol{\theta})$ with (j,k)th entry is given by

$$(\mathcal{I}(\boldsymbol{\theta}))_{j,k} = -\left(\frac{\partial^2 l_n(\boldsymbol{\theta}|\mathbf{Y})}{\partial\theta_j \partial\theta_k}\right), \tag{3.32}$$

for $j = 1, 2, \ldots, m$ and $k = 1, 2, \ldots, m$. The observed Fisher information matrix $\mathcal{I}(\boldsymbol{\theta})$ plays a particularly important role in maximum likelihood estimation. The inverse of $\mathcal{I}(\boldsymbol{\theta})$, denoted by $\mathcal{I}(\boldsymbol{\theta})^{-1}$, is the *asymptotic variance-covariance matrix* of the maximum likelihood estimates of the unknown parameters. Additionally, standard errors for MLEs can be found

approximately by removing the dependence of $\mathcal{I}(\boldsymbol{\theta})$ on $\boldsymbol{\theta}$, i.e. $\mathcal{I}(\boldsymbol{\theta}) \approx \mathcal{I}(\widehat{\boldsymbol{\theta}})$. In other words, the asymptotic standard error $(s.e.)$ of $\widehat{\theta_j}$ is the square root of the jth diagonal entry of $\mathcal{I}(\widehat{\boldsymbol{\theta}})^{-1}$, given by

$$s.e.(\widehat{\theta_j}) \approx \sqrt{\mathcal{I}(\widehat{\boldsymbol{\theta}})^{jj}}, \tag{3.33}$$

for $j = 1, 2, \ldots, m$. We can also determine approximately the ellipsoidal confidence regions for $\boldsymbol{\theta}$ using

$$(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^t \mathcal{I}(\widehat{\boldsymbol{\theta}})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \sim \chi^2_m.$$

From the standard error of $\widehat{\theta_j}$, 100(1-γ)% confidence limits for the corresponding true value θ_j are

$$\widehat{\theta_j} \pm z_{\gamma/2} s.e.(\widehat{\theta_j}), \tag{3.34}$$

where $z_{\gamma/2}$ is the upper $\gamma/2$ point of the standard normal distribution. By the proposition of consistency (Bulmer, 1979), suppose that the estimator $\widehat{\boldsymbol{\theta}} = (\widehat{\theta_1}, \ldots, \widehat{\theta_m})$ is the MLE for $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$. Then the $\widehat{\theta_j}$ are consistent for θ_j where $j = 1, \ldots, m$. By the proposition of asymptotic normality (Bulmer, 1979), the estimator $\widehat{\boldsymbol{\theta}}$ is approximately distributed as $\widehat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, \mathcal{I}(\boldsymbol{\theta})^{-1})$.

3.8 Application to antiplatelet therapy data (W1)

In this section, we shall use the proposed model to the W1 data given in Table 2.1 of Chapter 2. According to this data, most of total number of patients are large (larger than 20), thus the normal approximation model can be applied.

3.8.1 The model

From the W1 data, there are 31 studies (or RCTs) in total, investigating the use of aspirin plus dipyridamole or aspirin alone in comparison with the control group. The studies compare three treatments: aspirin plus dipyridamole (A), aspirin alone (B) and control group

(C). Six studies compare A, B and C, four studies compare A and B, thirteen studies compare A and C and seven studies compare B and C. For convenience, we partition the dataset into four groups of studies,

$$G_1 = \{1, \ldots, 6\},\ G_2 = \{7, \ldots, 10\},\ G_3 = \{11, \ldots, 24\} \text{ and } G_4 = \{25, \ldots, 31\},$$

comparing treatment A versus B versus C, A versus B, A versus C and B versus C, respectively. Let r_{iA}, r_{iB} and r_{iC} be the numbers of patients who have reocclusions on treatments A, B and C respectively where the ith study is in $G_1 \cup G_2 \cup G_3$, $G_1 \cup G_2 \cup G_4$ and $G_1 \cup G_3 \cup G_4$, respectively. The total numbers of patients are n_{iA}, n_{iB} and n_{iC} respectively. Let π_{iA}, π_{iB} and π_{iC} be the probabilities of patients that have reocclusions on treatments A, B and C respectively in the ith study. The r_{iA}, r_{iB} and r_{iC} are thus binomially distributed as

$$\begin{aligned} r_{iA} &\sim Bin(\pi_{iA}, n_{iA}), && i \in G_1 \cup G_2 \cup G_3, \\ r_{iB} &\sim Bin(\pi_{iB}, n_{iB}), && i \in G_1 \cup G_2 \cup G_4, \\ r_{iC} &\sim Bin(\pi_{iC}, n_{iC}), && i \in G_1 \cup G_3 \cup G_4. \end{aligned}$$

Suppose that X_{iA}, X_{iB} and X_{iC} are the empirical logistic transforms for $(r_{iA}, n_{iA}), (r_{iB}, n_{iB})$ and (r_{iC}, n_{iC}) respectively and are formulated in (3.5). For example, the empirical logistic transform of X_{iA} is defined by $\log(r_{iA}+0.5)/(n_{iA}-r_{iA}+0.5)$. From the discussion in Section 3.3, the X_{iA}, X_{iB} and X_{iC} have approximate normal distributions with means and variances given in (3.6). For example, the X_{iA} has an approximate normal distribution with mean $\log(\pi_{iA}/(1-\pi_{iA}))$ and variance $\widehat{\sigma_{iA}^2} = (n_{iA}+1)/((r_{iA}+0.5)(n_{iA}-r_{iA}+0.5))$. The normal approximation models using the empirical logistic transforms can therefore be applied with the data.

The baseline treatment for G_1, G_3 and G_4 is the control group, can be written as $b(i) = \text{C}$ for $i \in G_1 \cup G_3 \cup G_4$. While the baseline treatment for G_2 is the treatment B, $b(i) = \text{B}$ for

$i \in G_2$. The meta-analysis involves the direct comparison in G_1, G_3 and G_4 and indirect comparison in G_2. The sets for both comparisons are $D = \{G_1, G_3, G_4\}$ and $I = \{G_2\}$ respectively. First, by using the models (3.7) and (3.8), the empirical log-odds models for each group can be given by

$$
\begin{aligned}
i \in G_1, \quad & \begin{cases} X_{iC} = \alpha_i + \sigma_{iC}\epsilon_{iC}, \\ X_{iA} = \alpha_i + \delta_{i,AC} + \sigma_{iA}\epsilon_{iA}, \\ X_{iB} = \alpha_i + \delta_{i,BC} + \sigma_{iB}\epsilon_{iB}, \end{cases} \\
i \in G_2, \quad & \begin{cases} X_{iB} = \alpha_i + \sigma_{iB}\epsilon_{iB}, \\ X_{iA} = \alpha_i + \delta_{i,AB} + \sigma_{iA}\epsilon_{iA}, \end{cases} \\
i \in G_3, \quad & \begin{cases} X_{iC} = \alpha_i + \sigma_{iC}\epsilon_{iC}, \\ X_{iA} = \alpha_i + \delta_{i,AC} + \sigma_{iA}\epsilon_{iA}, \end{cases} \\
i \in G_4, \quad & \begin{cases} X_{iC} = \alpha_i + \sigma_{iC}\epsilon_{iC}, \\ X_{iB} = \alpha_i + \delta_{i,BC} + \sigma_{iB}\epsilon_{iB}. \end{cases}
\end{aligned}
$$

The trial effects are assumed to be different and the treatment effects $\delta_{i,AC}$, $\delta_{i,BC}$ and $\delta_{i,AB}$ are assumed to be random as in (3.12). The ϵ_{iA}, ϵ_{iB} and ϵ_{iC} are independent, following the standard normal distributions and corresponding to the random sampling errors of X_{iA}, X_{iB} and X_{iC} respectively. All random sampling errors are therefore independent and normally distributed as $N(0, \sigma_{iA}^2)$, $N(0, \sigma_{iB}^2)$ and $N(0, \sigma_{iC}^2)$, respectively.

Next, we will determine the *basic model for the random treatment effect*. Let $\boldsymbol{\delta}_{i,0}$ and $\boldsymbol{\mu}_0$ represent the vectors $(\delta_{i,AC}, \delta_{i,BC})^t$ and $(\mu_{AC}, \mu_{BC})^t$ respectively and let $\boldsymbol{\Omega}_0$ denote the 2×2 covariance matrix corresponding to $\boldsymbol{\delta}_{i,0}$. Thus, the model $\boldsymbol{\delta}_{i,0}$ is distributed as $MVN(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$, i.e.

$$
\begin{pmatrix} \delta_{i,AC} \\ \delta_{i,BC} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \mu_{AC} \\ \mu_{BC} \end{pmatrix}, \begin{pmatrix} \tau_{AC}^2 & \rho\tau_{AC}\tau_{BC} \\ \rho\tau_{AC}\tau_{BC} & \tau_{BC}^2 \end{pmatrix} \right). \tag{3.35}
$$

The μ_{AC} and μ_{BC} are the overall mean effects between the control group C and treatments A and B, respectively. The τ^2_{AC} and τ^2_{BC} measure the between-study heterogeneities of the treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ respectively. For notational convenience, we let ρ be the correlation coefficient between $\delta_{i,AC}$ and $\delta_{i,BC}$. By using the index matrix given in (3.16), the index matrix $\mathbf{F}_i$ for G_1 is the 2×2 identity matrix; the $\mathbf{F}_i$ for G_2, G_3 and G_4 are $(1,-1)$, $(1,0)$ and $(0,1)$ respectively. The treatment effect for the ith study is defined as $\boldsymbol{\delta}_i = \mathbf{F}_i\boldsymbol{\delta}_{i,0} \sim MVN(\mathbf{F}_i\boldsymbol{\mu}_0, \mathbf{F}_i\boldsymbol{\Omega}_0\mathbf{F}_i^t)$. The model $\boldsymbol{\delta}_i$ for G_1 consisting of three arms which is modelling by the basic model (3.35). The $\boldsymbol{\delta}_i$ for G_2 is given by

$$\delta_{i,AB} = \mathbf{F}_i\boldsymbol{\delta}_{i,0} = \delta_{i,AC} - \delta_{i,BC} \sim N(\mu_{AC} - \mu_{BC}, \tau^2_{AC} + \tau^2_{BC} - 2\rho\tau_{AC}\tau_{BC}). \tag{3.36}$$

Similarly the $\delta_{i,AC}$ and $\delta_{i,BC}$ for G_3 and G_4 are normally distributed as $N(\mu_{AC}, \tau^2_{AC})$ and $N(\mu_{BC}, \tau^2_{BC})$ respectively. Now we have the treatmenet effect models for each group. As mentioned before, we have 31 nuisance parameters in the models. To overcome the problem of inconsistency, the empirical log-odds ratio models are suggested here in order to eliminate the trial effects.

To present the empirical log-odds ratio models for the data, we first need to identify the *basic model for random sampling errors and empirical log-odds ratio models*. Let $\mathbf{e}_{i,0}$ be the vector $(e_{i,AC}, e_{i,BC})^t$ and let $\boldsymbol{\Sigma}_{i,0}$ denote the 2×2 covariance matrix of $\mathbf{e}_{i,0}$. The basic model $\mathbf{e}_{i,0}$ is normally distributed as $MVN(\mathbf{0}, \boldsymbol{\Sigma}_{i,0})$, given by

$$\begin{pmatrix} e_{i,AC} \\ e_{i,BC} \end{pmatrix} \sim MVN\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{i,AC} & \sigma^2_{iC} \\ \sigma^2_{iC} & \sigma^2_{i,BC} \end{pmatrix}\right). \tag{3.37}$$

To obtain the basic model for empirical log-odds ratio models, let $\mathbf{Y}_{i,0}$ be the vector $(Y_{i,AC}, Y_{i,BC})^t$.

Then, $\mathbf{Y}_i$ is distributed as $MVN(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0 + \boldsymbol{\Sigma}_{i,0})$, i.e.

$$\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{AC} \\ \mu_{BC} \end{pmatrix}, \begin{pmatrix} \tau_{AC}^2 + \sigma_{i,AC}^2 & \rho\tau_{AC}\tau_{BC} + \sigma_{iC}^2 \\ \rho\tau_{AC}\tau_{BC} + \sigma_{iC}^2 & \tau_{BC}^2 + \sigma_{i,BC}^2 \end{pmatrix}\right). \tag{3.38}$$

By setting $\mathbf{Y}_i = \mathbf{F}_i\mathbf{Y}_{i,0}$, the $\mathbf{Y}_i$s for G_1, G_2, G_3 and G_4 are $(Y_{i,AC}, Y_{i,BC})^t$, $Y_{i,AB}$, $Y_{i,AC}$ and $Y_{i,BC}$ respectively. The $\boldsymbol{\delta}_i$ for each group is the same as defined in the empirical log-odds models. The random sampling error $\mathbf{e}_i$ for each group is $\mathbf{F}_i\mathbf{e}_{i,0} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_i)$; where $\boldsymbol{\Sigma}_i = \mathbf{F}_i\boldsymbol{\Sigma}_{i,0}\mathbf{F}_i^t$. Specially, the log-odds ratio models are

$$i \in G_1, \qquad \begin{cases} Y_{i,AC} = \delta_{i,AC} + e_{i,AC}, \\ Y_{i,BC} = \delta_{i,BC} + e_{i,BC}, \end{cases} \tag{3.39}$$

$$i \in G_2, \qquad Y_{i,AB} = \delta_{i,AB} + e_{i,AB}, \tag{3.40}$$

$$i \in G_3, \qquad Y_{i,AC} = \delta_{i,AC} + e_{i,AC}, \tag{3.41}$$

$$i \in G_4, \qquad Y_{i,BC} = \delta_{i,BC} + e_{i,BC}. \tag{3.42}$$

The trial effects are no longer in the models. The model (3.39) is normally distributed as shown in (3.38). Additionally the empirical log-odds ratio models (3.40)-(3.42) for G_2 - G_4 are normally distributed as $N(\mu_{AC}-\mu_{BC}, \tau_{AB}^2+\sigma_{i,AB}^2)$, $N(\mu_{AC}, \tau_{AC}^2+\sigma_{i,AC}^2)$ and $N(\mu_{BC}, \tau_{BC}^2+\sigma_{i,BC}^2)$ respectively, where $\tau_{AB}^2 = \tau_{AC}^2 + \tau_{BC}^2 - 2\rho\tau_{AC}\tau_{BC}$.

3.8.2 Maximum likelihood estimation

To make inferences, the maximum likelihood method is used to estimate the unknown parameters in the empirical log-odds ratio models (3.39) - (3.42). The aim is to estimate the unknown parameters for the meta-analysis consisting of 31 studies. The log-likelihood

function $l_n(\boldsymbol{\theta})$ for the empirical log-odds ratio models is performed as

$$\sum_{i\in G_1} \log p(Y_{i,AC}, Y_{i,BC}|\boldsymbol{\theta}) + \sum_{i\in G_2} \log p(Y_{i,AB}|\boldsymbol{\theta}) + \sum_{i\in G_3} \log p(Y_{i,AC}|\boldsymbol{\theta}) + \sum_{i\in G_4} \log p(Y_{i,BC}|\boldsymbol{\theta}).$$

The $l_n(\boldsymbol{\theta})$ is the summation of the log-likelihoods from G_1 to G_4 where G_1, G_3 and G_4 are in the set D and G_2 is in the set I. The $p(Y_{i,AC}, Y_{i,BC}|\boldsymbol{\theta})$, $p(Y_{i,AB}|\boldsymbol{\theta})$, $p(Y_{i,AC}|\boldsymbol{\theta})$ and $p(Y_{i,BC}|\boldsymbol{\theta})$ represent the joint probabilities of observing data that has been collected in G_1, G_2, G_3 and G_4 respectively. We used the function *nlme* in the software R to calculate the MLEs (R Development Core Team, 2007). As described in Section 3.4.1, there are two assumptions of heterogeneity parameters: homogeneity and heterogeneity variances. Since there are only 4 studies in G_2, in absence of additional information, we assume homogeneity of variance for the model. The heterogeneity parameters for the models (3.39) - (3.42) are assumed to be the same: $\tau_{AC} = \tau_{BC} = \tau_{AB} = \tau$ and the correlation coefficient between the δ_{AC} and δ_{BC} takes the value 1/2. The collection of unknown parameters is therefore $\boldsymbol{\theta} = \{\mu_{AC}, \mu_{BC}, \tau^2\}$. For convenience, let θ_1, θ_2 and θ_3 stand for μ_{AC}, μ_{BC} and τ^2 respectively. To estimate the standard error of maximum likelihood estimator, let l_{ij} stand for the related partition of the second derivatives of the log likelihood function in terms of θ_i and θ_j. Using the equation (3.32), the 3×3 observed Fisher information matrix $\mathcal{I}(\boldsymbol{\theta})$ is written as

$$\mathcal{I}(\boldsymbol{\theta}) = -\begin{pmatrix} l_{\theta_1\theta_1}(\boldsymbol{\theta}) & l_{\theta_1\theta_2}(\boldsymbol{\theta}) & l_{\theta_1\theta_3}(\boldsymbol{\theta}) \\ l_{\theta_2\theta_1}(\boldsymbol{\theta}) & l_{\theta_2\theta_2}(\boldsymbol{\theta}) & l_{\theta_2\theta_3}(\boldsymbol{\theta}) \\ l_{\theta_3\theta_1}(\boldsymbol{\theta}) & l_{\theta_3\theta_1}(\boldsymbol{\theta}) & l_{\theta_3\theta_3}(\boldsymbol{\theta}) \end{pmatrix}. \tag{3.43}$$

Standard errors can be calculated from the inverse matrix of $\mathcal{I}(\boldsymbol{\theta})$.

3.8.3 Numerical results

The estimates of the unknown parameters μ_{AC}, μ_{BC} and τ are shown in Table 3.1.

Table 3.1: The results for the empirical log-odds ratio models on the log-odds ratio (LOR) and odds ratio (LO) scales

	δ_{AB}		δ_{AC}		δ_{BC}	
	μ_{AB}	τ_{AB}	μ_{AC}	τ_{AC}	μ_{BC}	τ_{BC}
LOR	0.108146	0.275320	-0.568930	0.275320	-0.677076	0.275320
(SD)	(0.118645)	(0.136747)	(0.161554)	0.136747)	(0.150660)	(0.136747)
95%CI	(-0.12,0.34)	(0.007,0.54	(-0.88,-0.25)	(0.007,0.54)	(-0.97,-0.38)	(0.007,0.54)
OR	1.114210		0.566130		0.508100	
95%CI	(0.88,1.40)		(0.41,0.77)		(0.37,0.68)	

They are denoted by $\widehat{\mu_{AC}}$, $\widehat{\mu_{BC}}$ and $\widehat{\tau}$, respectively. Note that the estimate of μ_{AB} is obtained from $\widehat{\mu_{AC}} - \widehat{\mu_{BC}}$. The overall means of the treatment effects A versus B, A versus C and B versus C are 0.108146, -0.568930 and -0.677076 respectively and the variations between studies in those comparisons are the same, 0.275320. Taking the inverse of the observed Fisher information matrix (3.43), the asymptotic variance-covariance matrix of the unknown parameters for the models (3.39) - (3.42) is

$$\mathcal{I}(\widehat{\boldsymbol{\theta}})^{-1} = \begin{pmatrix} 0.026110 & 0.013530 & -0.0040 \\ 0.013530 & 0.022732 & -0.001220 \\ -0.0040 & -0.001220 & 0.018778 \end{pmatrix} \tag{3.44}$$

From this matrix, the asymptotic variances of $\widehat{\mu_{AC}}$, $\widehat{\mu_{BC}}$ and $\widehat{\tau}$ are the entries on the diagonal of the matrix, 0.26110, 0.022732 and 0.018778 respectively. As a result their asymptotic standard errors are 0.161554, 0.150660 and 0.136747 respectively. The variance of $\widehat{\mu_{AB}}$ is estimated from

$$Var(\widehat{\mu_{AB}}) = Var(\widehat{\mu_{AC}}) + Var(\widehat{\mu_{BC}}) - 2\rho\ se(\widehat{\mu_{AC}})se(\widehat{\mu_{BC}}).$$

The standard error of this estimate is 0.118645. Using these results, approximate 95%

confidence intervals on the log-odds-ratio scale for the estimators of μ_{AB}, μ_{AC},μ_{BC} and τ are $(-0.12, 0.34)$, $(-0.88, -0.25)$, $(-0.97, -0.38)$ and $(0.007, 0.54)$ respectively. All treatment effects are estimated on the LOR scale. The overall means of the treatment effects δ_{AB}, δ_{AC} and δ_{BC} are 1.114210, 0.566130 and 0.508100 on the OR scale. The results indicate that both treatment A and treatment B reduce the rates of reocclusion significantly by over 40% compared to the control group. However the difference between treatment A and treatment B is almost negligible although treatment B is slightly better than treatment A (improved by about 11%). The confidence intervals for the true values, μ_{AB}, μ_{AC} and μ_{BC} on the OR scale can be calculated for the related CI on the LOR scale, which are (0.88,1.40), (0.41,0.77) and (0.37,0.68) respectively.

3.9 Discussion

This chapter has demonstrated the normal approximation model based on the empirical logistic transform to multi-arm trials data. We first proposed the special case of empirical log-odds model with each of M studies comprising all $K+1$ treatments. The model did not cover all possible cases of multi-arm trials, e.g. if baseline treatments in some studies are different. Thus the general case of the empirical log-odds model was considered to model any multi-arm trial data set, including the direct and indirect comparisons. The treatment effect was defined in term of both comparisons using the *basic model of random treatment effect*. The mean and variance of the model that involves the indirect comparison cannot be estimated directly. Note that whenever there is no or insufficient evidence of direct comparison from RCTs, the indirect comparison may provide useful or supplementary information on the treatment effect. However the validity of the indirect comparisons depends on the internal validity and similarity of the included studies, see Song et al. (2003); Lu and Ades (2006). Additionally, we also described the assumptions of heterogeneity parameters – homogeneity and heterogeneity of variances - for the model. Generally, the assumption of variance homogeneity has been most used , see e.g. Higgins and Whitehead (1996); Lu and

Ades (2004, 2006).

In practice the trial effects in most meta-analysis would not satisfy any model (fixed effect or random effect) since different experiment designs and different data analysis models are used in different studies. Most of the existing methods assume that they are study-level effect. We also use this assumpiton in this thesis. Additionally, the treatment effects are assumed to be random because we do not believe that results from different studies and different designs can have the same treatment effect.

From the empirical log-odds model mentioned above, the trial effects are different, thus the number of unknown parameters (from the trial effect) are the same as the number of studies. The estimation may be unstable as many parameters are involved in the model, especially if the number of studies is large. The accuracy of estimation thus depends on the number of individual observations from each study, e.g. if this number is large enough then the estimate may be accurate. Also this may lead to a problem of inconsistent estimate. To avoid this problem, we suggested the empirical log-odds ratio model to eliminate the trial effects from the empirical log-odds model. There are at least three advantages of using the empirical log-odds ratio model over other methods:(1) the model excludes the trial effects and give a consistent estimate for treatment effect while the other methods (e.g. the empirical log-odds model) may give an inconsistent estimate in some circumstances;(2) the approximation is usually quite good if the number of individual observations is not too small (the number of samples in a single study should usually be larger than 20); (3) the computation is very efficient and very stable, it converges very fast for almost any starting point. It takes less than 2 seconds to get the results.

From the application to the W1 data, the studies $7-10$ (G_2) involve the indirect treatment effect $\delta_{i,AB}$, obtained from $\delta_{i,AC} - \delta_{i,BC}$. As mentioned before, the correlation coefficients

between the treatment effects under the assumption of variance heterogeneity are estimable if enough information is provided in the indirect comparison. Since there is not enough information in G_2 thus the correlation coefficient between treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ cannot be estimated. The assumption of variance heterogeneity is not valid for the model. Conversly, if the numbers of studies in G_1 and G_2 were 24 and 12 respectively, the correlation coefficient could be estimated by borrowing strength from indirect comparison (Higgins and Whitehead, 1996). Collaboration (1994a) concluded that antiplatelet therapy (aspirin plus dipyridamole (A) or aspirin alone (B)) produced a highly significant ($2p \leq 0.00001$) reduction in vascular occlusion in a wide range of patients. The odds of vascular graft or arterial occlusion were reduced by about 40% while treatment continued. Our numerical results in Table 3.1 are similar to those of Collaboration (1994a).

Even though the efficiency of computation for the empirical log-odds ratio model is good and the model gives a consistent estimate comparing to other methods, the model requires the large number of individual observations (larger than 20) and the probability of an unsuccessful outcome π_{ij} to be not too near zero or one. The MLEs of the model may not be accurate when compared to the model with the exact binomial distribution. We shall introduce the exact binomial model in the next chapter.

Chapter 4

Meta-analysis of multi-arm trials using binomial approach

4.1 Introduction

In the previous chapter, we proposed the normal approximation model using an empirical logistic transform. The model requires a large number of individual observations n_{ij} and the probability of an unsuccessful outcome π_{ij} to be not too near zero or one. If the number of individual observations is small, the model in Chapter 3 is not suitable. In this chapter, we introduce an exact binomial model to fit the binary multi-arm trials data. There are two alternative maximum likelihood approaches that can be used to make inferences for the unknown parameters in the logistic regression model. These are the *unconditional method* and *conditional method*. The logistic regression model has become increasingly popular with the easy availability of appropriate computer routines. Many authors have described maximum likelihood estimation procedures which turn out to be iterative, for example Cox (1970, page 61). Albert and Anderson (1984) dealt with the existence of maximum likelihood in logistic regression models and proved on existence theorem by considering the possible pattern of data points. The use of the conditional likelihood in logistic models is well established and

routines for fitting it are provided by major statistical software (Pendergast et al., 1996). Hirji (1994) proposed an efficient algorithm to generate the exact distribution for the bivariate logistic model with common and sub-unit-specific covariates and also presented the exact unconditional and conditional distribution of the model. Bellio and Sartori (2003) proposed the modified profile likelihood as an ideal extension of the conditional likelihood in generalized linear models for binary data with the generic link function, and also suggested that an important feature of the implementation was the standard outputs of routines for the generalized linear models. With an application in biology, Zhao and Aragaki (2000) investigated a conditional likelihood approach of candidate genes and showed analytically the consistency of this approach. There have been a large number of studies about unconditional and conditional methods, for example, see Cox (1972); Prentice (1976); Tritchler (1984) and Sartori (2003).

As mentioned in Chapter 1, most existing methods for meta-analysis of multi arm trials use the logistic regression model with the unconditional approach. Thompson and Sharp (1999) used the random effects logistic regression model with the unconditional method to explain heterogeneity in meta-analysis of serum cholesterol reduction. Lu and Ades (2004) introduced the Bayesian hierarchical model for multi-arm trials using the unconditional method to estimate unknown parameters. More examples can be found in Lu and Ades (2006); Lu et al. (2007). Using the unconditional maximum likelihood approach, note that if the number of studies is large and the number of individual observations is small then the estimate may be biased or misleading (Cox and Snell, 1989, page 103). For example, if the individual observations n_{i0} and n_{ij} are equal to 1 then for large M, the estimate of unconditional maximum likelihood $\widehat{\delta_{i,b(i)j}}$ is close to $2\delta_{i,b(i)j}$ (Cox and Snell, 1989, page 59). We thus introduce the logistic regression model using the conditional approach in this chapter.

The structure of this chapter is arranged as follows. We introduce the logistic regression

model for the direct and indirect comparisons in Section 4.2. Unconditional maximum likelihood approach for the model including the standard error of MLEs are described in Section 4.3. Similarly, conditional maximum likelihood approach for the model is presented in Section 4.4. In Section 4.5, we illustrate the logistic regression model with the unconditional and conditional approaches with the W2 data. We discuss the advantages and the limitations of the two approaches in the final section.

4.2 Fitting the logistic regression model

This section illustrates how to fit the logistic regression model to the binary data related to multi-arm trials including the direct and indirect comparisons. Logistic regression is a regression model for a binomially distributed response/dependent variable. It is useful for modelling the probability of an event occurring as a function of other factors. Logistic regression is part of a category of statistical models called *generalized linear models* and uses the logit as its link function. Logistic regression can be used only with two types of dependent variables: one is a categorical dependent variable that has exactly two categories (i.e. a binary or dichotomous variable). The other is a continuous dependent variable that has values in the range 0 to 1 representing the probability values or the proportions. The names for logistic regression used in various other application areas are *logistic model* or *logit model*. Logistic regression is similar to linear regression in that we are interested in the relationship of a group of independent variables with a response or dependent variable. In linear regression, the ultimate objective for the study may be either estimation of the coefficient values, or prediction of the response value. One significant difference between logistic and linear models is that the linear model has a continuous response variable and the logistic model uses a binary or dichotomous response.

All notations used in this chapter are the same as defined in Chapter 3 unless stated. Suppose that the $r_{ib(i)}$ and r_{ij} are binomially distributed, respectively as $Bin(n_{ib(i)}, \pi_{ib(i)})$ and

$Bin(n_{ij}, \pi_{ij})$ for $i = 1, \ldots, M$ and $j \in J_{(i)}$. Logistic regression models for the ith study can be defined by

$$\log\left(\frac{\pi_{ib(i)}}{1-\pi_{ib(i)}}\right) = \alpha_i, \tag{4.1}$$

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha_i + \delta_{i,b(i)j}, \qquad j \in J_{(i)}. \tag{4.2}$$

The assumptions of the trial effect and the treatment effect are the same as were assumed in the empirical log-odds models: the α_i's are assumed to be different and the $\delta_{i,b(i)j}$ are assumed to be random as presented in (3.17). The above models can be used for both treatment comparisons. From model (4.2), we call log $\pi_{ij}/(1-\pi_{ij})$ the *logistic transform of probability* π_{ij}, or alternatively *log odds* π_{ij} or *logit* π_{ij}. Having considered the properties of logit π_{ij}, the term $\pi_{ij}/(1-\pi_{ij})$ is the odds of an unsuccessful outcome from a patient treated with treatment j and so logit π_{ij} is the log odds of an unsuccessful outcome. It is easily seen that a value of π_{ij} in the range $(0, 1)$ corresponds to a value of logit π_{ij} in $(-\infty, \infty)$. As $\pi_{ij} \to 0$, logit $\pi_{ij} \to -\infty$; as $\pi_{ij} \to 1$, logit $\pi_{ij} \to \infty$ and for $\pi_{ij} = 0.5$, logit $\pi_{ij} = 0$. After some rearrangement, the logistic regression models (4.1) and (4.2) have equivalent formulations as

$$\pi_{ib(i)} = \left(\frac{e^{\alpha_i}}{1+e^{\alpha_i}}\right) \qquad \text{and} \qquad \pi_{ij} = \left(\frac{e^{\alpha_i+\delta_{i,b(i)j}}}{1+e^{\alpha_i+\delta_{i,b(i)j}}}\right). \tag{4.3}$$

There are two alternative ML approaches, the unconditional and conditional approaches, that can be used to estimate the unknown parameters in a logistic regression model. They will be performed in the following sections.

4.3 Unconditional maximum likelihood approach

Generally, unconditional ML estimation is prefered if the number of parameters in the model is small relative to the number of studies in a meta-analysis (Kleinbaum, 1994, page 106).

4.3.1 Probability functions

To demonstrate the unconditional ML estimation, let $p(r_{ib(i)}|\alpha_i)$ and $p(r_{ij}|\alpha_i, \delta_{i,b(i)j})$ denote the probability functions associated with the distributions of $r_{ib(i)}|\alpha_i$ and $r_{ij}|\alpha_i, \delta_{i,b(i)j}$ respectively for $i = 1, \ldots, M$ and $j \in J_{(i)}$, defined as follows.

For the baseline treatment,

$$p(r_{ib(i)}|\alpha_i) = \begin{pmatrix} n_{ib(i)} \\ r_{ib(i)} \end{pmatrix} \pi_{ib(i)}^{r_{ib(i)}} (1 - \pi_{ib(i)})^{n_{ib(i)} - r_{ib(i)}} = \begin{pmatrix} n_{ib(i)} \\ r_{ib(i)} \end{pmatrix} \frac{e^{\alpha_i r_{ib(i)}}}{(1 + e^{\alpha_i})^{n_{ib(i)}}}. \tag{4.4}$$

For the treatments j, $j \in J_{(i)}$

$$p(r_{ij}|\alpha_i, \delta_{i,b(i)j}) = \begin{pmatrix} n_{ij} \\ r_{ij} \end{pmatrix} \pi_{ij}^{r_{ij}} (1 - \pi_{ij})^{n_{ij} - r_{ij}} = \begin{pmatrix} n_{ij} \\ r_{ij} \end{pmatrix} \frac{e^{(\alpha_i + \delta_{i,b(i)j}) r_{ij}}}{(1 + e^{(\alpha_i + \delta_{i,b(i)j})})^{n_{ij}}}. \tag{4.5}$$

The combination in (4.4) represents the number of possible combinations of observations $n_{ib(i)}$ taken $r_{ib(i)}$ at a time. The $\pi_{ib(i)}$ in the middle term of (4.4) is substituted from (4.3) and $(1 - \pi_{ib(i)})$ becomes $1/1 + e^{\alpha_i}$. The combination in (4.5) can be considered in the same way.

4.3.2 The unconditional likelihood

From the probability functions (4.4) and (4.5), the trial effects α_i's are study-level effects. They are assumed to be different and also included in both probability functions. While the $\delta_{i,b(i)j}$ is a random effect, thus the $p(r_{ij}|\alpha_i, \delta_{i,b(i)j})$ involves the vector of random effects, $\boldsymbol{\delta}_i$, given in (3.17). The standard method of handling a probability function which involves random variables that have a fully specified probability is to integrate the probability function with respect to the distribution of those variables. To deal with the random effects $\boldsymbol{\delta}_i$, let $\mathbf{r}_{(i)}$ be the vector $(r_{ij}, j \in J_{(i)})^t$. We shall integrate the probability function $p(r_{ij}|\boldsymbol{\delta}_i)$ with respect to $\boldsymbol{\delta}_i$. The $p(\mathbf{r}_{(i)})$ contains k_i integrals, which is the number of treatments in the set

$J_{(i)}$, and is given by

$$p(\mathbf{r}_{(i)}) = \int_{\boldsymbol{\delta}_i} \prod_{j \in J_{(i)}} p(r_{ij}|\boldsymbol{\delta}_i)\phi(\boldsymbol{\delta}_i; \boldsymbol{\mu}_i, \boldsymbol{\Omega}_i)d\boldsymbol{\delta}_i, \tag{4.6}$$

where $\phi(\boldsymbol{\delta}_i; \boldsymbol{\mu}_i, \boldsymbol{\Omega}_i)$ is the probability density function of the normal distribution with mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Omega}_i$ defined in (3.17), given by

$$\phi(\boldsymbol{\delta}_i; \boldsymbol{\mu}_i, \boldsymbol{\Omega}_i) = \frac{1}{(2\pi)^{k_i/2}\left|\boldsymbol{\Omega}_i\right|^{1/2}} e^{-(\boldsymbol{\delta}_i - \boldsymbol{\mu}_i)' \boldsymbol{\Omega}_i^{-1}(\boldsymbol{\delta}_i - \boldsymbol{\mu}_i)/2}. \tag{4.7}$$

The integral (4.6) can be calculated numerically; one way to do it is to use the Gauss-Hermite method. To apply Gauss-Hermite approximation, the probability function $p(\mathbf{r}_{(i)})$ for the ith study can be estimated by

$$p(\mathbf{r}_{(i)}) \approx \pi^{-k_i/2} \sum_{n_1=1}^{l_1} w_{n_1}^{(1)} \ldots \sum_{n_{k_i}=1}^{l_{k_i}} w_{n_{k_i}}^{(k_i)} \left\{ \prod_{j \in J_{(i)}} \begin{pmatrix} n_{ij} \\ r_{ij} \end{pmatrix} \frac{e^{\left(\alpha_i + (\boldsymbol{\mu}_i + \sqrt{2}\boldsymbol{\Omega}_i^{1/2}\mathbf{d}_{i,n})\right) r_{ij}}}{\left(1 + e^{\alpha_i + (\boldsymbol{\mu}_i + \sqrt{2}\boldsymbol{\Omega}_i^{1/2}\mathbf{d}_{i,n})}\right)^{n_{ij}}} \right\}, \tag{4.8}$$

where the sampling nodes are at $\boldsymbol{\mu}_i + \sqrt{2}\boldsymbol{\Omega}_i^{1/2}\mathbf{d}_{i,n}$ and $\mathbf{d}_{i,n} = (x_{n_1}^{(1)}, \ldots, x_{n_{k_i}}^{(k_i)})$. The vector $\mathbf{d}_{i,n}$ depends on the number k_i, which is the number of treatments comprising in the ith study. The resulting function (4.8) does not depend on the $\boldsymbol{\delta}_i$. For most practical purposes, l_{k_i} need not be greater than 20, although some authors suggest using even smaller values (Collett, 1991, page 208). The assumptions of the heterogeneity parameters (variances for $\boldsymbol{\delta}_i$) are similar to those described in Section 3.4.2 of Chapter 3.

As before, let $\boldsymbol{\theta}$ be the collection of all unknown parameters for the meta-analysis including all trial effets $(\alpha_1, \ldots, \alpha_M)$, $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ and let $\mathbf{r}_i$ be the vector $(r_{ij}, j \in J_i)$. The likelihood function for the ith study can be written as

$$L(\boldsymbol{\theta}|\mathbf{r}_i) = \prod_{j \in J_i} p(r_{ij}) = p(r_{ib(i)}|\alpha_i).p(\mathbf{r}_{(i)}), \tag{4.9}$$

where $p(r_{ib(i)}|\alpha_i)$ and $p(\mathbf{r}_{(i)})$ are given in (4.4) and (4.8) respectively. Let $l_{u,i} = \log L(\boldsymbol{\theta}|r_i)$,

standing for the unconditional log-likelihood function of the logistic regression model for the ith study. The log-likelihood function of $\boldsymbol{\theta}$ for the models (4.1) and (4.2) is given by

$$l_u(\boldsymbol{\theta}) = \sum_{i=1}^{M} l_{u,i}. \tag{4.10}$$

Bear in mind that the number of α_i's is the same as the number of studies. The computation of MLEs may be quite unstable if the number of studies is large while the sample size of each study is small. As discussed at the beginning of this chapter, this may also result in a biased or misleading estimate. We thus suggest using a conditional approach to eliminate all nuisance parameters in Section 4.4.

4.3.3 Asymptotic variance-covariance matrix

In this section, we will show how to calculate the standard errors for the MLEs of the logistic regression model using the unconditional approach. Since there are random effects in the model, some integrals are involved in the likelihood function. The unconditional log-likelihood function (4.10) can be written as

$$\begin{aligned} l_u(\boldsymbol{\theta}) &= \sum_{i=1}^{M} \log p(r_{ib(i)}) + \sum_{i=1}^{M} \log p(\mathbf{r}_i), \\ &= \sum_{i=1}^{M} \log p(r_{ib(i)}) + \sum_{i=1}^{M} \log \int_{\boldsymbol{\delta}_i} \prod_{j \in J_{(i)}} p(r_{ij}|\boldsymbol{\delta}_i) \phi(\boldsymbol{\delta}_i; \boldsymbol{\mu}_i, \boldsymbol{\Omega}_i) d\boldsymbol{\delta}_i. \end{aligned} \tag{4.11}$$

We let l_1 and l_2 stand for the first and second terms of the above log-likelihood function, given by $l_u(\boldsymbol{\theta}) = l_1 + l_2$. Three types of unknown parameters are involved in $\boldsymbol{\theta}$; the trial effects, α_i's, the overall mean effects μ's (for $\boldsymbol{\mu}$), and the variances τ's and the correlation coefficients ρ's in the covariance matrix $\boldsymbol{\Omega}$. For convenience, we let τ represent a parameter (either τ or ρ) involved in $\boldsymbol{\Omega}$. There is no random effect involved in l_1.

First, the second-order partial derivative $\partial^2 l_1 / \partial \alpha_i^2$ can be calculated in the usual way; while

the other terms are

$$\frac{\partial^2 l_1}{\partial \alpha_i \alpha_j} = \frac{\partial^2 l_1}{\partial \alpha_i \mu} = \frac{\partial^2 l_1}{\partial \alpha_i \tau} = 0, \qquad i \neq j \ \text{ and } \ i, j \in \{1, ..., M\}$$

Next, let us consider the second term of (4.11), for notational convinence, let $P_i(\boldsymbol{\delta}_i)$ represent the function $\prod_{j \in J_{(i)}} p(r_{ij}|\boldsymbol{\delta}_i)$ in l_2. Now the term l_2 takes the form

$$l_2(\alpha_1, ..., \alpha_M, \boldsymbol{\mu}, \boldsymbol{\Omega}) = \sum_{i=1}^{M} \log \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i)\phi(\boldsymbol{\delta}_i) d\boldsymbol{\delta}_i = \sum_{i=1}^{M} l_{2i},$$

where $\phi(\boldsymbol{\delta}_i)$ is the density of the multivariate normal distribution with mean $\boldsymbol{\mu}_i$ and variance matrix $\boldsymbol{\Omega}_i$, and l_{2i} is a summand of the log-likelihood involving the integrals ($\log \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i)\phi(\boldsymbol{\delta}_i) d\boldsymbol{\delta}_i$). The first-order partial derivatives relating to l_2 are shown as follows

$$\begin{aligned}
\frac{\partial l_2}{\partial \alpha_i} &= \sum_{i=1}^{M} e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} \frac{\partial P_i(\boldsymbol{\delta}_i)}{\partial \alpha_i} \phi(\boldsymbol{\delta}_i) d_{\boldsymbol{\delta}_i}, \\
\frac{\partial l_2}{\partial \mu} &= \sum_{i=1}^{M} e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i) \frac{\partial \phi(\boldsymbol{\delta}_i)}{\partial \mu} d_{\boldsymbol{\delta}_i}, \\
\frac{\partial l_2}{\partial \tau} &= \sum_{i=1}^{M} e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i) \frac{\partial \phi(\boldsymbol{\delta}_i)}{\partial \tau} d_{\boldsymbol{\delta}_i}.
\end{aligned}$$

Similarly, the second-order partial derivatives are

$$\frac{\partial^2 l_2}{\partial \alpha_i^2} = \sum_{i=1}^{M} \left(e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} \frac{\partial^2 P_i(\boldsymbol{\delta}_i)}{\partial \alpha_i^2} \phi(\boldsymbol{\delta}_i) d_{\boldsymbol{\delta}_i} - \left(e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} \frac{\partial P_i(\boldsymbol{\delta}_i)}{\partial \alpha_i} \phi(\boldsymbol{\delta}_i) d_{\boldsymbol{\delta}_i} \right)^2 \right), \quad (4.12)$$

$$\frac{\partial^2 l_2}{\partial \mu^2} = \sum_{i=1}^{M} \left(e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i) \frac{\partial^2 \phi(\boldsymbol{\delta}_i)}{\partial \mu^2} d_{\boldsymbol{\delta}_i} - \left(e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i) \frac{\partial \phi(\boldsymbol{\delta}_i)}{\partial \mu} d_{\boldsymbol{\delta}_i} \right)^2 \right), \quad (4.13)$$

$$\frac{\partial^2 l_2}{\partial \tau^2} = \sum_{i=1}^{M} \left(e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i) \frac{\partial^2 \phi(\boldsymbol{\delta}_i)}{\partial \tau^2} d_{\boldsymbol{\delta}_i} - \left(e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i) \frac{\partial \phi(\boldsymbol{\delta}_i)}{\partial \tau} d_{\boldsymbol{\delta}_i} \right)^2 \right), \quad (4.14)$$

$$\frac{\partial^2 l_2}{\partial \mu \partial \tau} = \sum_{i=1}^{M} \left(e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i) \frac{\partial^2 \phi(\boldsymbol{\delta}_i)}{\partial \mu \partial \tau} d_{\boldsymbol{\delta}_i} - \left(e^{-l_{2i}} \int_{\boldsymbol{\delta}_i} P_i(\boldsymbol{\delta}_i) \frac{\partial \phi(\boldsymbol{\delta}_i)}{\partial \mu} d_{\boldsymbol{\delta}_i} \right) \right)^2. \quad (4.15)$$

Note that the second-order partial derivative $\partial^2 l_2 / \partial \alpha_i \partial \alpha_j$ is equal to zero. The second-order

partial derivatives of $\partial^2 l_2/\partial\mu_i\partial\mu_j$ and $\partial^2 l_2/\partial\tau_i\partial\tau_j$ can be expressed in similar equations to (4.13) and (4.14). The integrals in the first-order and second-order partial derivatives can be approximated by Gaussian quadrature.

From the log-likelihood (4.11), the second-order partial derivatives for the observed Fisher information matrix can be calculated as

$$\frac{\partial^2 l_u}{\partial\alpha_i^2} = \frac{\partial^2 l_1}{\partial\alpha_i^2} + \frac{\partial^2 l_2}{\partial\alpha_i^2}, \qquad \frac{\partial^2 l_u}{\partial\mu^2} = \frac{\partial^2 l_2}{\partial\mu^2}, \qquad \frac{\partial^2 l_u}{\partial\mu_i\partial\mu_j} = \frac{\partial^2 l_2}{\partial\mu_i\partial\mu_j},$$

$$\frac{\partial^2 l_u}{\partial\tau^2} = \frac{\partial^2 l_2}{\partial\tau^2}, \qquad \frac{\partial^2 l_u}{\partial\tau_i\partial\tau_j} = \frac{\partial^2 l_2}{\partial\tau_i\partial\tau_j}, \qquad \frac{\partial^2 l_u}{\partial\mu\partial\tau} = \frac{\partial^2 l_2}{\partial\mu\partial\tau}.$$

As set earlier, the second partial derivatives of $\partial^2 l_u/\partial\rho^2$ and $\partial^2 l_u/\partial\mu\partial\rho$ (and $\partial^2 l_u/\partial\tau\partial\rho$) can be calculated in similar equations to $\partial^2 l_u/\partial\tau^2$ and $\partial^2 l_u/\partial\mu\partial\tau$ respectively. Notice that the second-order derivative of l_1 is only related in $\partial^2 l_u/\partial\alpha_i^2$. We can partition the matrix of second partial derivatives into a block matrix with null matrices in the off diagonals:

$$\mathcal{H}(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{H}_\alpha(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \mathcal{H}_{\mu\tau\rho}(\boldsymbol{\theta}) \end{pmatrix},$$

where $\mathcal{H}_\alpha(\boldsymbol{\theta})$ and $\mathcal{H}_{\mu\tau\rho}(\boldsymbol{\theta})$ are the second-order partial derivatives about α_i, and μ, τ and ρ respectively. By multiplying $\mathcal{H}(\boldsymbol{\theta})$ by -1, the observed Fisher information matrix $\mathcal{I}(\theta)$ is obtained. The inverse of $\mathcal{I}(\theta)$ is the asymptotic variance-covariance matrix of MLEs and their standard errors are the square roots of the diagonal of $\mathcal{I}(\theta)^{-1}$.

4.4 Conditional maximum likelihood approach

Conditional likelihood is widely used in logistic regression models with binary data. In particular, this leads to accurate inferences for the parameters of interest and eliminates all nuisance parameters (Kleinbaum, 1994). We shall define the conditional likelihood and

describe the maximum likelihood estimation in this section.

4.4.1 Conditional likelihood

From the logistic regression models (4.1) and (4.2), the conditional likelihood $\mathbf{r}_i$ given that $C_i = \sum_{j \in J_i} r_{ij} = c_i$ for the ith study, is given by

$$f(\mathbf{r}_i | C_i = c_i; \boldsymbol{\delta}_i) = f(\mathbf{r}_i | \sum_{j \in J_i} r_{ij} = c_i; \boldsymbol{\delta}_i) = \frac{f(\mathbf{r}_i | \boldsymbol{\delta}_i)}{f(\sum_{j \in J_i} r_{ij} = c_i | \boldsymbol{\delta}_i)}. \tag{4.16}$$

The conditional likelihood reflects the probability of the observed data configuration relative to the probability of all possible configurations of the given data. The numerator $f(\mathbf{r}_i|\boldsymbol{\delta}_i)$ is exactly the same as the unconditional likelihood obtained from (4.4) and (4.5) . The denominator is what makes the conditional likelihood different from the unconditional likelihood; it sums the joint probability for all possible configurations. To derive the equation (4.16), the conditional likelihood $\mathbf{r}_i$ given C_i can be simplified as

$$f(\mathbf{r}_i | C_i = c_i; \boldsymbol{\delta}_i) = \frac{\prod_{j \in J_i} \begin{pmatrix} n_{ij} \\ r_{ij} \end{pmatrix} e^{(\boldsymbol{\delta}_i r_{ij})}}{\sum_{\mathbf{u}_i \in \mathcal{U}_i} \begin{pmatrix} n_{i0} \\ c_i - \sum_{j \in J_{(i)}} u_{ij} \end{pmatrix} \prod_{j \in J_{(i)}} \begin{pmatrix} n_{ij} \\ u_{ij} \end{pmatrix} e^{(\boldsymbol{\delta}_i u_{ij})}}, \tag{4.17}$$

where $\mathbf{u}_i = (u_{ij},\ j \in J_{(i)})^t$ and

$$\mathcal{U}_i = \left\{ \mathbf{u}_i : 0 \leq u_{ij} \leq n_{ij}, j \in J_{(i)} \text{and } c_i - n_{i0} \leq \sum_{j \in J_{(i)}} u_{ij} \leq c_i \right\}.$$

Notice that this likelihood function does not involve any nuisance parameters α_i's and is a function of $\boldsymbol{\delta}_i$ alone. The removal of the trial effects from the conditional likelihood is important because it means that when the conditional likelihood is used, estimates are obtained only for the parameters of interest in the model and not for the α_i's.

4.4.2 Estimation

The conditional likelihood (4.17) has k_i random effects so the likelihood $f(\mathbf{r}_i|\sum_{j\in J_i} r_{ij} = c_i)$ involves k_i integrations:

$$f(\mathbf{r}_i|\sum_{j\in J_i} r_{ij} = c_i) = \int_{\boldsymbol{\delta}_i} f(\mathbf{r}_i|\sum_{j\in J_i} r_{ij} = c_i;\boldsymbol{\delta}_i)\phi(\boldsymbol{\delta}_i;\boldsymbol{\mu}_i,\boldsymbol{\Omega}_i)d\boldsymbol{\delta}_i, \tag{4.18}$$

where $\phi(\boldsymbol{\delta}_i;\boldsymbol{\mu}_i,\boldsymbol{\Omega}_i)$ is the probability density function of multivariate normal distribution with mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Omega}_i$, given in (4.7). Similar to the discussion in the previous section, we apply Gauss-Hermite approximation to (4.18) and obtain:

$$f(\mathbf{r}_i|\sum_{j\in J_i} r_{ij} = c_i) \approx \pi^{-k_i/2}\sum_{n_1=1}^{l_1} w_{n_1}^{(1)} \ldots \sum_{n_{k_i}=1}^{l_{k_i}} w_{n_{k_i}}^{(k_i)} f(\mathbf{r}_i|\sum_{j\in J_i} r_{ij} = c_i;\boldsymbol{\delta}_{i,n}), \tag{4.19}$$

where $f(\mathbf{r}_i|\sum_{j\in J_i} r_{ij} = c_i;\boldsymbol{\delta}_{i,n})$ is obtained from (4.17) where the sampling nodes is $\boldsymbol{\delta}_{i,n} = \boldsymbol{\mu}_i + \sqrt{2}\boldsymbol{\Omega}_i^{1/2}\mathbf{d}_{i,n}$ and $\mathbf{d}_{i,n} = (d_{n_1}^{(1)},\ldots,d_{n_{k_i}}^{(k_i)})$. Again, let $\boldsymbol{\theta}$ be the collection of all unknown parameters for the meta-analysis. The likelihood for the ith study $L(\boldsymbol{\theta}|\mathbf{r}_i)$ can be written as

$$L(\boldsymbol{\theta}|\mathbf{r}_i) = f(\mathbf{r}_i|\sum_{j\in J_i} r_{ij} = c_i;\boldsymbol{\delta}_i).$$

The log-likelihood function of the logistic regression models using the conditional approach is

$$l_c(\boldsymbol{\theta}) \;=\; \log L(\boldsymbol{\theta}|\mathbf{r}) = \sum_{i=1}^{M}\log L(\boldsymbol{\theta}|\mathbf{r}_i), \tag{4.20}$$

By maximising the conditional likelihood function over $\boldsymbol{\theta}$ we obtain an exact parameter estimate for $\boldsymbol{\theta}$, called the *conditional maximum likelihood estimate*. To calculate the standard

error of their MLEs, the log-likelihood function (4.20) can be written as

$$\begin{aligned} l_c(\boldsymbol{\theta}) &= \sum_{i=1}^{M} \log f(\mathbf{r}_i | \sum_{j \in J_i} r_{ij} = c_i), \\ &= \sum_{i=1}^{M} \log \int_{\boldsymbol{\delta}_i} f(\mathbf{r}_i | \sum_{j \in J_i} r_{ij} = c_i, \boldsymbol{\delta}_i) \phi(\boldsymbol{\delta}_i; \boldsymbol{\mu}_i, \boldsymbol{\Omega}_i) d\boldsymbol{\delta}_i. \end{aligned} \tag{4.21}$$

Let $P_i(\boldsymbol{\delta}_i)$ represent $f(\mathbf{r}_i | \sum_{j \in J_i} r_{ij} = c_i, \boldsymbol{\delta}_i)$ in the above equation. The second-order partial derivatives of $\partial^2 l_c / \partial \mu^2$, $\partial^2 l_c / \partial \tau^2$ and $\partial^2 l_c / \partial \mu \partial \tau$ are similar to the equations (4.13) - (4.15) respectively. In a similar way to the previous section, the standard errors for the MLEs are obtained.

4.5 Application to antiplatelet therapy data (W2)

From the W2 data given in Table 2.2 of Chapter 3, the number of individual observations is small thus the empirical log-odds model is not appropriate. In this section, we shall apply the logistic regression model using the unconditional and conditional approaches with the W2 data.

4.5.1 Unconditional inference

From the W2 data, there are 27 studies investigating the use of aspirin plus dipyridamole or aspirin alone in comparison with the control group. The studies compare three treatments: aspirin plus dipyridamole (A), aspirin alone (B) and the control treatment (C). Seven studies compare A, B and C, ten studies compare A and C and ten studies compare B and C. There is no indirect comparison for this dataset, so the set D is $\{1, \ldots, 27\}$. The baseline treatment for all studies is the control group ($b(i) = 0$).

The indices $i = 1, \ldots, 27$ and $j = 0, 1, 2$ stand for the studies and the treatments C, A and B, respectively. The data is partitioned into three groups: $G_1 = \{1, \ldots, 7\}, G_2 = \{8, \ldots, 17\}$

and $G_3 = \{18, \ldots, 27\}$. The sets J_i and $J_{(i)}$ are given by

$$\begin{cases} J_i = \{0,1,2\}, & J_{(i)} = \{1,2\} & \text{for } i \in G_1, \\ J_i = \{0,1\}, & J_{(i)} = \{1\} & \text{for } i \in G_2, \\ J_i = \{0,2\}, & J_{(i)} = \{2\} & \text{for } i \in G_3. \end{cases} \tag{4.22}$$

Let r_{i0}, r_{i1} and r_{i2} be the numbers of patients who suffered reocclusions on treatments C, A and B respectively, where the ith study is in $G_1 \cup G_2 \cup G_3$, $G_1 \cup G_2$ and $G_1 \cup G_3$, respectively. The total numbers of patients are n_{i0}, n_{i1} and n_{i2}. Let π_{i0}, π_{i1} and π_{i2} be the probabilities that patients have reocclusions on treatments C, A and B respectively in the ith study. The r_{i0}, r_{i1} and r_{i2} are binomially distributed as

$$\begin{aligned} r_{i0} &\sim Bin(\pi_{i0}, n_{i0}), \quad i \in G_1 \cup G_2 \cup G_3, \\ r_{i1} &\sim Bin(\pi_{i1}, n_{i1}), \quad i \in G_1 \cup G_2, \\ r_{i2} &\sim Bin(\pi_{i2}, n_{i2}), \quad i \in G_1 \cup G_3. \end{aligned}$$

The treatment effect models can be obtained in the same way to that described in Section 3.8 of Chapter 3. For example, the treatment effect $\boldsymbol{\delta}_i$ for G_1 is defined as

$$\begin{pmatrix} \delta_{i,01} \\ \delta_{i,02} \end{pmatrix} \sim MVN \left(\begin{pmatrix} \mu_{01} \\ \mu_{02} \end{pmatrix}, \begin{pmatrix} \tau_{01}^2 & \rho\tau_{01}\tau_{02} \\ \rho\tau_{01}\tau_{02} & \tau_{02}^2 \end{pmatrix} \right). \tag{4.23}$$

Logistic regression models for the data can be fitted using the equations (4.1) and (4.2) where $b(i) = 0$ and $J_{(i)}$ is given in (4.22). Note that the trial effects are assumed to be different in each study. To define the unconditional likelihood function, let $\mathbf{r}_{(i)}$ represent the vector (r_{i1}, r_{i2}). The probability functions $p(r_{i0})$ and $p(\mathbf{r}_{(i)})$ are formulated from (4.4) and (4.8) respectively.

From $\boldsymbol{\delta}_i$ for G_1, the correlation coefficient ρ between $\delta_{i,01}$ and $\delta_{i,02}$ is in the form $\tau_0^2/\tau_{01}\tau_{02}$, where is obtained from $\delta_{i0} \sim N(\mu_0, \tau_0^2)$. Note that μ_0 and τ_0^2 are not estimable unless

some other information is used. We shall consider the assumption of homogeneity variance here. Suppose that all heterogeneity parameters are the same: $\tau_{01} = \tau_{02} = \tau$ and the correlation coefficient takes the value 1/2. The unknown parameter $\boldsymbol{\theta}$ for the models is $\{\alpha_1, \alpha_2, \ldots, \alpha_{27}, \mu_{01}, \mu_{02}, \tau^2\}$. The log-likelihood function $l_u(\boldsymbol{\theta})$ is obtained from (4.10). By maximizing the log-likelihood function, the MLEs can be estimated. Also we calculate their standard errors from the observed Fisher information matrix given in Section 4.3.3.

The results for the treatment effects δ_{01} and δ_{02} are given in Table 4.1. The trial effects are presented in Table 4.2. The overall means on the LOR scale for δ_{01} and δ_{02} are -1.17849 (SD 0.08499) and -0.63700 (SD 0.03728), and the heterogeneity parameter is 0.0372 (SD 0.04752). On the OR scale, the means are 0.30774 and 0.52800 respectively. Their confidence interval can be calculated from the related CI on the LOR scale. We conclude that treatments aspirin plus dipyridamole and aspirin only in antiplatelet therapy reduce deep venous thrombosis by over 70% and 45% respectively. The average of both treatments reduce deep venous thrombosis by over 55 %.

Table 4.1: The results of the treatment effects for the model using the unconditional method

	δ_{01}		δ_{02}	
	μ_{01}	τ_{01}	μ_{02}	τ_{02}
LOR scale	-1.17849	0.00372	-0.63700	0.00372
(SD)	(0.08499)	(0.04752)	(0.03728)	(0.04752)
95%CI	(-1.33,-1.00)	(-0.08,0.09)	(-0.64,-0.62)	(-0.08,0.09)
OR scale	0.30774		0.52800	
95%CI	(0.26,0.36)		(0.52,0.53)	

Table 4.2: The trial effects of the model using the unconditional method

Study	1-5	-0.72387	1.20619	-0.54688	-3.01061	0.55283
(SD)		(0.01021)	(0.00934)	(0.00993)	(0.0117)	(0.00926)
95%CI		(-0.74,-0.70)	(1.18, 1.22)	(-0.56,-0.52)	(-3.03, -2.98)	(0.53,0.57)
Study	6-10	-0.85773	-1.69947	-0.34480	-0.65231	-1.29308
(SD)		(0.01206)	(0.01264)	(0.00735)	(0.00770)	(0.01087)
95%CI		(-0.88,-0.83)	(-1.72,-1.67)	(-0.35,-0.33)	(-0.66, -0.63)	(-1.31,-1.27)
Study	11-15	-2.18147	1.68130	1.17724	0.68567	-0.14132
(SD)		(0.01231)	(0.00636)	(0.00811)	(0.00869)	(0.01102)
95%CI		(-2.20,-2.15)	(1.66,1.69)	(1.16,1.19)	(0.66,0.70)	(-0.16,-0.11)
Study	16-20	-1.53114	-0.57320	-0.33486	-4.24972	-1.05748
(SD)		(0.01214)	(0.00941)	(0.00661)	(0.01367)	(0.01199)
95%CI		(-1.55,-1.50)	(-0.59, -0.55)	(-0.34,-0.32)	(-4.27,-4.22)	(-1.08,-1.03)
Study	21-25	-3.01727	-0.11773	-2.23388	0.24853	0.04007
(SD)		(0.01184)	(0.00977)	(0.01252)	(0.00802)	(0.00987)
95%CI		(-3.04,-2.99)	(-0.13,-0.09)	(-2.25,-2.20)	(0.23,0.26)	(0.02,0.05)
Study	26-27	-0.37573	-0.76995			
(SD)		(0.01277)	(0.01098)			
95%CI		(-0.40,-0.35)	(-0.79,-0.74)			

4.5.2 Conditional inference

The models and other parameters are similar to those defined in the unconditional method. The function C_i for the data can be defined by

$$\begin{cases} C_i = r_{i0} + r_{i1} + r_{i2} & \text{for } i \in G_1, \\ C_i = r_{i0} + r_{i1} & \text{for } i \in G_2, \\ C_i = r_{i0} + r_{i2} & \text{for } i \in G_3. \end{cases} \tag{4.24}$$

Let $\mathbf{r}_i$ denote the vector (r_{i0}, r_{i1}, r_{i2}). The conditional likelihood $f(\mathbf{r}_i|C_i)$ for the ith study is given in (4.18). To handle the random treatment effect $\boldsymbol{\delta}_i$, the likelihood function is approx-

imated by Gaussian-Hermite approximation as defined in (4.19). The unknown parameter $\boldsymbol{\theta}$ for the models is $\{\mu_{01}, \mu_{02}, \tau\}$. By using the log-likelihood function (4.20), the results of the models are given in Table 4.3. On the LOR scale, the overall mean effects for both treatment effects are -0.87516 (SD 0.04340) and -0.39000 (SD 0.31160) while their variation between studies is 0.37000 (SD 0.03900). Those means on the OR scale are 0.41679 and 0.67434. As before their confidence intervals are obtained from the related CI on the LOR scale. The results indicate that treatments aspirin plus dipyridamole and aspirin only produce a reduction in deep venous thrombosis by over 55% and 30% respectively. The average of both treatments in antiplatelet therapy reduces deep venous thrombosis by over 40 %.

As seen from Tables 4.1 and 4.3, the results from using the unconditional likelihood (on the LOR scale) are smaller than from using conditional likelihood. Note that those results are negative. That is to say that estimation with unconditional likelihood may cause underestimation or bias. Collaboration (1994b) summarized that antiplatelet therapy produced a highly significant ($2p \leq 0.00001$) reduction in deep venous thrombosis of about 40%. The results from the model using the conditional likelihood support this.

Table 4.3: The results of the treatment effects for the model using the conditional method

	δ_{01}		δ_{02}	
	μ_{01}	τ_{01}	μ_{02}	τ_{02}
LOR scale	-0.87516	0.37000	-0.39000	0.37000
(SD)	(0.04340)	(0.03900)	(0.31160)	(0.03900)
95%CI	(-0.96,-0.79)	(0.29,0.44)	(-1.00,0.22)	
OR scale	0.41679		0.67434	
95%CI	(0.38,0.45)		(0.36,1.24)	

4.6 Discussion

In Chapter 3, we presented the normal approximation model for a large number of individual observations. In this chapter, we have introduced the logistic regression model for the exact binomial distribution. Two types of comparisons, direct and indirect, can be used with the model. Two alternative approaches for making inferences were presented. The unconditional likelihood involves nuisance parameters (from the trial effects). If the number of studies (M) is large, it may lead to inconsistent estimate. Cox and Snell (1989, page 103) concluded for the unconditional likelihood that if the number of studies (M) is large and the number of individual observations (n_{ij}) is small then it makes estimation inaccurate and inconsistent. Thus we introduced the conditional maximum likelihood approach for the model to eliminate all nuisance parameters. In making a choice between the two approaches, we need to consider the number of studies and the number of individual observations. However, the use of this method can be expensive in term of the cost of computer running time, especially if the number of individual observations is large. Simulation studies will be conducted in the next chapter to compare these two approaches. Some other methods can be used in the logistic regression model, for example, using a pseudo-loglikelihood, see Severini (1998); or the modified profile likelihood, see Bellio and Sartori (2003).

Gaussian-Hermite quadrature was used to calculate the integral forms of the probabilities including random effects in the likelihood functions for both approaches. The approximation is reasonably effective for low-oder integrations (Crouch and Spiegelman, 1990). Implementing Gaussian-Hermite approximation, we used the function 'gauss.quad' in the software R to estimate MLEs for the model. The number of integrands depends on the number of treatments involved in those studies. If this number is large then it makes the dimensionality of the integral large and so it cannot be approximated accurately. Other approximations such as Laplace approximation or Monte Carlo method can be used, see Ripatti and Palmgren (2000); Shi and Copas (2002). Laplace approximation could make the calculation of

second-order derivatives for the observed Fisher information matrix easier than using Gaussian approximation since there is no weight term in the approximation (Liu and Pierce, 1994).

Chapter 5

Simulation study

5.1 Introduction

We saw the normal approximation model used with an empirical logistic transform with the W1 data in Chapter 3. Computation of the model is efficient and converges very fast for almost any starting point. The approximation of the model is quite good if the number of individual observations is large (Chootrakool and Shi, 2008). In Chapter 4, the logistic regression model was introduced for the exact binomial distribution including the unconditional and conditional approaches to making inferences. We applied the logistic regression model with the W2 data because some of the numbers of individual observations were not large enough (less than 20) to use the normal approximation model. By comparing the results from both approaches in Chapter 4, we concluded that the results from the unconditional approach may be inconsistent. This bias can be eliminated by considering the conditional approach to the logistic regression model (Prentice and Breslow, 1978; Lubin, 1981). Thus, the conditional maximum likelihood estimate may be more accurate in a certain situation. The theory of exact conditional logistic regression analysis (or exact inference) was first proposed by Cox (1970) (McCarthy, 2007). The unconditional approach (or asymptotic inference) is suitable for the small number of parameters in the model (Kleinbaum, 1994).

In this chapter, we examine the performance of various inference methods from the normal approximation model and the logistic regression model using the unconditional and conditional methods. The main aim is to compare the unconditional and conditional methods of the logistic regression model in different cases. We demonstrate the procedure for generating data in Section 5.2. The models that are used to make inferences in the simulation are presented in Section 5.3. We discuss and compare some models in Section 5.4. Simulation details and the results are given in Sections 5.5 and 5.6 respectively. Finally, Section 5.7 concludes and gives some discussions about the chapter.

5.2 Simulated data

In this section, we aim to generate the data set which will be used in the simulation study. The basic data structure is the same as the W2 data. The baseline treatment for all studies is the control group, which means there is only direct comparison here. The data is generated from binomial distribution (3.1) with logistic regression models given in (4.1) and (4.2) where $b(i) = 0$. The indices $i = 1, \ldots, M$ and $j \in J_{(i)}$ represent the studies and the treatments, respectively. The general scheme of generating the data is given as follows:

1. Give the numbers of individual observations n_{i0} and n_{ij};
2. Generate the trial effect α_i;
3. Generate the treatment effect $\delta_{i,0j}$;
4. Calculate the probabilities π_{i0} and π_{ij}: substituting the generated trial effect and the generated treatment effect into models (4.1) and (4.2), and the probabilities are obtained;
5. Generate the r_{i0} and r_{ij} from binomial distribution (3.1) ;
6. Repeat steps 1-5 until the data is generated for all M studies.

Similar to the W2 data, we consider three treatments in the simulation study. Two scenarios are employed here.

- S1: The values of α_i's are different. The treatment effects are $\delta_{i,01} \sim N(-1.0, 0.2^2)$ and $\delta_{i,02} \sim N(-0.30, 0.05^2)$.
- S2: The values of α_i's are generated from a distribution $N(-0.92, 0.2^2)$. The treatment effects are $\delta_{i,01} \sim N(-1.0, 0.2^2)$ and $\delta_{i,02} \sim N(-0.30, 0.05^2)$.

Note that there is no association between the treatment effects $\delta_{i,01}$ and $\delta_{i,02}$. In S1, the values of α_i's are quite different (this is the case we usually encounter in practice). In S2, we assume α_i's come from a normal distribution.

5.3 The models

Eight different models related to the normal approximation model and the logistic regression model will be considered. For convenience, let 'M1', 'M2' and 'M3' represent the empirical log-odds ratio model and the logistic regression model using the unconditional and conditional methods, respectively. The correlation coefficients between the treatment effects in this section are assumed to be zero. This may be written as $\rho_{jk} = 0$ where $j \neq k$ and $j, k \in J_{(i)}$. Therefore, the covariance between the treatment effects is $Cov(\delta_{ij}, \delta_{ik}) = 0$. Let 'F' and 'R' denote the fixed-treatment effect model and the random-treatment effect model, respectively. For the logistic regression model, let 'd' and 'N' represent the different-trial effect and the random-trial effect model, respectively.

5.3.1 The empirical log-odds ratio model

We compare the empirical log-odds ratio model with the fixed-treatment effect and the random-treatment effect. The empirical log-odds ratio model is given in (3.29), and the treatment effects are assumed to be random; this is model '**M1-R**'. By setting the variances

of all the treatment effects in the 'M1-R' model equal to zero, then the model '**M1-F**' is obtained.

5.3.2 The logistic regression model

The logistic regression models given in (4.1) and (4.2) are applied here with different choices of treatment effect ('R' or 'F') and trial effect ('d' or 'N'). The following logistic regression models are used to make inferences with the unconditional maximum likelihood approach as described in Chapter 4.

- **M2-F-d**: the treatment effects are fixed: $\delta_{i,0j} = \mu_{0j}$ and the trial effects are different parameters.

- **M2-R-d**: the treatment effects are assumed to be random and normally distributed as $\delta_{i,0j} \sim N(\mu_{0j}, \tau_{0j}^2)$ and the trial effects are different parameters.

- **M2-F-N**: the treatment effects are fixed as above and the trial effect is assumed to be random as $N(\mu_{\alpha 0}, \tau_{\alpha 0}^2)$. Hence, the probability function for the baseline treatment $p(r_{i0}|\alpha_i)$ has a random effect α_i. By integrating $p(r_{i0}|\alpha_i)$ with respect to α_i, we obtain

$$p(r_{i0}) = \int p(r_{i0}|\alpha_i)\phi(\alpha_i; \mu_{\alpha 0}, \tau_{\alpha 0}^2)d\alpha_i, \tag{5.1}$$

  where $p(r_{i0}|\alpha_i)$ is defined in (4.4) where the trial effect is normally distributed. The $\phi(\alpha_i; \mu_{\alpha 0}, \tau_{\alpha 0}^2)$ is the normal distribution with mean $\mu_{\alpha 0}$ and variance $\tau_{\alpha 0}^2$. As discussed in Chapter 4, the integral (5.1) can be calculated numerically by a Gauss-Hermite approximation, taking the form

$$p(r_{i0}) \approx \pi^{-1/2} \begin{pmatrix} n_{i0} \\ r_{i0} \end{pmatrix} \sum_{n=1}^{l} w_n \left\{ \frac{e^{\left(\mu_{\alpha 0} + \sqrt{2}\tau_{\alpha 0} d_n\right) r_{i0}}}{\left(1 + e^{(\mu_{\alpha 0} + \sqrt{2}\tau_{\alpha 0} d_n)}\right)^{n_{i0}}} \right\}, \tag{5.2}$$

  where the sampling nodes are at $\mu_{\alpha 0} + \sqrt{2}\tau_{\alpha 0} d_n$ for $n = 1, \ldots, l$.

- **M2-R-N**: the treatment effects and the trial effect are assumed to be random as defined above. The probability function $p(r_{i0})$ here is the same as (5.2). The probability function $p(r_{ij}|\alpha_i, \delta_{i,0j})$ involves two random effects of the trial effect α_i and the treatment effect $\delta_{i,0j}$. The probability $p(r_{ij})$ is given by

$$p(r_{ij}) = \int\int p(r_{ij}|\alpha_i, \delta_{i,0j})\phi(\alpha_i; \mu_{\alpha 0}, \tau^2_{\alpha 0})\phi(\delta_{i,0j}; \mu_{0j}, \tau^2_{0j})d\alpha_i d\delta_{i,0j}, \tag{5.3}$$

  where $\phi(\alpha_i; \mu_{\alpha 0}, \tau^2_{\alpha 0})$ and $\phi(\delta_{i,0j}; \mu_{0j}, \tau^2_{0j})$ are the probability density functions of normal distributions for α_i and $\delta_{i,0j}$ respectively. As before, the $p(r_{ij})$ can be approximated by

$$p(r_{ij}) \approx \pi^{-1} \begin{pmatrix} n_{ij} \\ r_{ij} \end{pmatrix} \sum_{n_1=1}^{l_1} w^{(1)}_{n_1} \sum_{n_2=1}^{l_2} w^{(2)}_{n_2} \left\{ \frac{e^{\left((\mu_{\alpha 0}+\sqrt{2}\tau_{\alpha 0}d_{n_1})+(\mu_{0j}+\sqrt{2}\tau_{0j}d_{n_2})\right)r_{ij}}}{\left(1+e^{(\mu_{\alpha 0}+\sqrt{2}\tau_{\alpha 0}d_{n_1})+(\mu_{0j}+\sqrt{2}\tau_{0j}d_{n_2})}\right)^{n_{ij}}} \right\}, \tag{5.4}$$

  where the sampling nodes are at $(\mu_{\alpha 0} + \sqrt{2}\tau_{\alpha 0}d_{n_1}) + (\mu_{0j} + \sqrt{2}\tau_{0j}d_{n_2})$ for $n_1 = 1, \ldots, l_1$ and $n_2 = 1, \ldots, l_2$.

For the conditional maximum likelihood approach, we consider the fixed-effect model (for treatment effect) denoted by '**M3-F**', and the random-effect model denoted by '**M3-R**'. We do not need to consider the trial effects since they are eliminated.

5.4 Comparison of models

We shall compare three models, M1-R, M2-R-d and M3-R, in terms of limitations, computation and drawbacks. Those are the mostly used models in practice. Assuming that the multi-arm trials data is similar to the special case given in Section 3.4.1. The treatment effects for all models are assumed to be random and the trial effects are assumed to be different parameters for the logistic regression model. The brief conclusions are summarized in Table 5.1.

Table 5.1: Conclusions of the models

Model	Limitations	Computation	Drawbacks
Empirical log-odds ratio (M1-R)	n_{i0} and n_{ij} are large π_{i0} and π_{ij} are not near 0 or 1	fast	not accurate if sample size is small
Logistic regression with unconditional method (M2-R-d)	-	medium	biased estimate and unstable computation if n_{i0} and n_{ij} are small and M is large
Logistic regression with conditional method (M3-R)	-	slow	time consuming if n_{i0},n_{ij} and K are large

Regarding the limitations of each model, if the number of individual observations n_{i0} and n_{ij} is reasonable large (larger than 20) and the probability π_{ij} is not near 0 or 1 then the empirical log-odds ratio model is appropriate (Shi and Copas, 2002; Chootrakool and Shi, 2008). According to the discussion in Section 3.3 of Chapter 3, with the opposite conditions, the empirical log-odds ratio model is not valid because the empirical logistic transforms for (r_{i0}, n_{i0}) and (r_{ij}, n_{ij}) are not approximately normally distributed. While the logistic regression model can be used for the exact binomial distribution without any limitations. The unconditional or conditional maximum likelihood approaches can be employed with the logistic regression model for making inferences.

In term of computation, the empirical log-odds model is distributed as a multivariate normal distribution. Its likelihood function is straightforward, as shown in (3.31). If the numbers of studies (M) and/or treatments (K) are large, it will not affect the computation much compared to the other models. Therefore, the computation of MLEs for the empirical log-odds model is fast. For the logistic regression model, we use the Gaussian-Hermite approximation to deal with random variables for both inference methods. This is one of the reasons to make

the calculation for the logistic regression model take more time than the empirical log-odds model. By comparing both inference methods, as described in Chapter 4, the numerator of the conditional likelihood is exactly the same as the likelihood for the unconditional likelihood but the denominator of the conditional likelihood requires summing $\mathbf{u}_i$ terms, where the $\mathbf{u}_i$ are defined by

$$\mathcal{U}_i = \left\{ \mathbf{u}_i : 0 \leq u_{ij} \leq n_{ij}, \text{ and } c_i - n_{i0} \leq \sum_j u_{ij} \leq c_i; j = 1, \ldots, K \right\}.$$

This is often computationally prohibitive. The computation is tedious and slow, particularly if n_{i0}, n_{ij} and K are large (Lubin, 1981; Prentice and Breslow, 1978). Consequently the conditional maximum likelihood estimation is slower than the unconditional maximum likelihood estimation.

To conclude the drawbacks, the estimate from the empirical log-odds model is not as accurate as from the model using exact binomial distribution unless the sample size for each studies is sufficiently large. The logistic regression model using the unconditional method includes nuisance parameters; the model should be used with a small number of studies. The estimate may be biased if the number of observations n_{i0} and n_{ij} are small and the number of studies M is large (Lubin, 1981; Cox and Snell, 1989). As mentioned above, if n_{i0}, n_{ij} and K are large for the logistic regression model using the conditional method, it can be time consuming. The main advantage of the conditional likelihood approach is that the likelihood depends only on the parameter of interest.

5.5 Simulation details

If the sample size of each individual study is large, the empirical logistic transform model is always the best choice. Here, we just compare the different models in two different cases: small number of individual observations with a medium number of studies, and very small

number of individual observations with a large number of studies. By using the scheme of generating data in Section 5.2, we use the numbers of individual observations n_{i0} and n_{ij} from the two original data sets as follows:

- 27 studies with small number of individual observations (n_{i0} and n_{ij} are the same as for the W2 data but r_{i0} and r_{ij} are generated as discussed in Section 5.2);
- 54 studies with very small number of individual observations. To construct this data set, we double the data set from 27 to 54 studies but the number of each individual study is halved from the W2 data.

For notational convenience, let '$M = 27$' and '$M = 54$' represent two types of the simulated data sets respectively. Following the steps given in Section 5.2, the trial effect and the treatment effect are generated from the models S1 and S2. For each generated data set, the eight models discussed in Section 5.3 are used. The estimates of treatment effects and other parameters are calculated for each model.

We compute 1000 replications in our simulation study. The root mean squared error ($r.m.s.e.$) is used to measure the performance for different models. Suppose that θ_0 is the true value and $\widehat{\theta_i}$ is the value of estimation obtained in the ith replication. The $r.m.s.e$ for θ is defined as

$$r.m.s.e(\widehat{\theta}) = \left(\frac{1}{n_r} \sum_{i=1}^{n_r} (\widehat{\theta_i} - \theta_0)^2 \right)^{1/2}.$$

where n_r is the number of replications, and $n_r = 1000$ in our simulation study. The value of $r.m.s.e$ and the sample means of $\widehat{\theta_i}$'s are calculated. The results of the simulation study for S1 and S2 are shown in Tables 5.2 and 5.3, and 5.4 and 5.5 (see the end of this chapter).

5.6 Results

5.6.1 Scenario 1

(i). **S1 with M = 27**

Table 5.2 gives the simulation study results based on the data generated from S1 with $M = 27$. The sample size for each individual study is quite small. Note that simulation model S1 is the logistic regression model with the different-trial effects and random-treatment effects thus the models with random-treatment effects may give good estimates. The true values from S1 are $\mu_{01} = -1.0$, $\tau_{01} = 0.2$, $\mu_{02} = -0.3$ and $\tau_{02} = 0.05$. The trial effects in S1 are different. We compare eight different models, and the sample means and r.m.s.e.'s are reported in Table 5.2. The following conclusions are our findings.

(a) Overall, the sample means from model M3-R (the logistic regression model with random-treatment effects by using the conditional likelihood) are the ones most close to the true values. That is to say that the model gives the least bias. Also the values of r.m.s.e for this model give the best performance.

(b) Since the sample size for some studies is very small, as expected, the accuracy of the estimates (sample means) from the empirical log-odds ratio models (M1) are not good as the logistic regression models (M2 and M3) except the models with random-trial effect by using the unconditional likelihood (M2-F-N and M2-R-N).

(c) By comparing the unconditional and conditional methods for the logistic regression models with random-treatment effect, the estimates and the values of r.m.s.e from M3-R give respectively the better results and performance than M2-R-d.

(d) We shall use the quantile-quantile plot (or Q-Q plot) to test the normality of the trial effect assumption we used in M2-F-N and M2-R-N. The Q-Q plot for α_i's is shown in Figure 5.1. Some plotted values fall off on a straight-line. This means that the trial effects do not follow the normal distribution. The normality

assumption for the trial effect fails in models M2-F-N and M2-R-N. The simulation study results given in Table 5.2 indicate that those two models perform badly comparing to other models.

(ii). **S1 with M = 54**

The sample means from the models in Table 5.3 are based on the data generated from S1 with $M = 54$. The sample size in each individual studies is very small but the number of studies is large. The true values for μ_{01}, τ_{01}, μ_{02} and τ_{02} are the same as S1 with $M = 27$. In addition, we shall compare the sample means and performance from the models in Tables 5.2 (27 studies) and 5.3 (54 studies). The data generated from S1 with $M = 27$ has small sample size of each individual study and medium number of studies. While the data from S1 with $M = 54$ has very small sample size of each individual study and large number of studies. We expect these results from the comparison: (1) the empirical log-odds ratio models ($M1$) from $M = 54$ should perform even worse than those from $M = 27$, because the sample size of individual studies in $M = 54$ is even smaller; (2) the logistic regression model with unconditional method ($M2$) from $M = 54$ may give inconsistent or biased estimates due to very small sample size of individual studies and the large number of study in meta-analysis. The results from our simulation study are summarized as follows.

(a) As expected, the model M3-R gives the best estimates.

(b) Similar to S1 with $M = 27$, the sample means from models M1 are least accurate. In comparison models M1 from $M = 27$ and $M = 54$, the empirical model with $M = 54$ has larger bias and less accuracy, this is because the normal approximation is deteriorated for smaller sample size.

(c) The estimates from M3-R give the better estimates than M2-R-d. By comparing models M2 and M3 from $M = 27$ and $M = 54$, as expected, models M2 from $M = 27$ give the less bias than from $M = 54$. The performance of models M2

from $M = 27$ is better than those from $M = 54$ except models M2-F-N and M2-R-N.

(d) Models M2-F-N and M2-R-N assume normality for the trial effect wrongly, the models fail for the data. The Q-Q plot of the trial effects (from 54 studies) is given in Figure 5.2. The plot does not support the normality of trial effects either. Similar to S1 with $M = 27$, the trial effect cannot be assumed to be normally distributed in models M2-F-N and M2-R-N.

5.6.2 Scenario 2

(i). **S2 with M= 27**

In Table 5.4, the means and r.m.s.e.'s of the models are obtained from simulation model S2 with 27 studies; this is the logistic regression model with the random-trial effect and random-treatment effects. The true values for S2 are $\mu_{\alpha 0} = -0.92$, $\tau_{\alpha 0} = 0.2$, $\mu_{01} = -1.0$, $\tau_{01} = 0.2$, $\mu_{02} = -0.3$ and $\tau_{02} = 0.05$. Notice that this simulation model assume a normal distribution for the trial effect. We expect the same results as described in S1 with $M = 27$ but models M2-F-N and M2-R-N would perform better. The simulation results are summarized as follows.

(a) Again, the estimates from the empirical log-odds ratio model are least accurate; because the sample size of each individual study is small as used in S1 with $M = 27$.

(b) By comparing the unconditional and conditional methods for the logistic regression models, the estimates and the values of r.m.s.e from M2-R-N give the better results and performance than M3-R because the model M2-R-N is exactly the same model as S2 although the performance of conditional likelihood method for M3 is still very good.

(ii). **S2 with M= 54**

The estimates from the models are obtained from the data generated from S2 with $M = 54$. The results from the simulation study are the same as S2 with $M = 27$. For example, model M2-R-N gives the best estimates; means of models M1 give the most bias. The summaries of comparisons between S1 with $M = 27$ and S1 with $M = 54$ are similar from as described in (ii) of Section 5.6.1.

5.7 Discussion

The simulation provides opportunities to analyse the data that are not available when using the real data set alone. Generally, the results from the simulation give more robust and dependable solutions. The empirical log-odds ratio model was proposed for a certain situation in Chapter 3. In Chapter 4, we introduced the exact binomial model (logistic regression model) for binary multi-arm trials data. We also expected that the conditional maximum likelihood estimation of the model would be more accurate than the unconditional maximum likelihood estimation because there were no nuisance parameters involved. In this chapter we have examined the performance of estimation in those models in different situations. Additionally we made some general conclusions on the comparisons of mostly used models.

The sample means and r.m.s.e.'s from the empirical log-odds ratio models (M1) between $M = 27$ and $M = 54$ suggest that the models are suitable for large individual observations only (Cox and Snell, 1989; Shi and Copas, 2002). The individual observations from 27 studies are larger than from 54 studies thus their MLEs are close to the true values than from 54 studies.

For the logistic regression models using the unconditional method (M2), there are nuisance parameters involved in the model. The accuracy of estimates depends on the number of individual observations and nuisance parameters. The estimates from M2 from the simulation with 54 studies confirm that the use of the unconditional method leads to biased estimates

if the number of individual observations is small and the number of studies is large (Cox and Snell, 1989; Hirji et al., 1987), although their standard errors are very small (Lubin, 1981).

The logistic regression models using the conditional method (M3) perform well in almost all the cases. However, as described in Section 5.4, one obstacle of the conditional method is the computational complexity. From simulation study results of $M = 27$ (from S1 and S2), the large number of individual studies makes the estimation of models M3 difficult to compute, see e.g. Prentice and Breslow (1978); Hirji et al. (1987). The number of individual observations in $M = 54$ is small and the number of studies is large. The computation is not heavy as for $M = 27$.

Overall, we have the following conclusions for meta-analysis of multi-arm trials. If the sample size in each individual study is large enough (larger than 20), see Chootrakool and Shi (2008), we shall use an empirical logistic transform model; otherwise we should use an exact logistic regression model with conditional likelihood. However, in the case that the number of studies is not very large but the sample size in each individual study is not very small, the performance of conditional and unconditional likelihood approaches are quite similar (Cox and Snell, 1989, page 103), we can use the unconditional likelihood approach to reduce the computation burden.

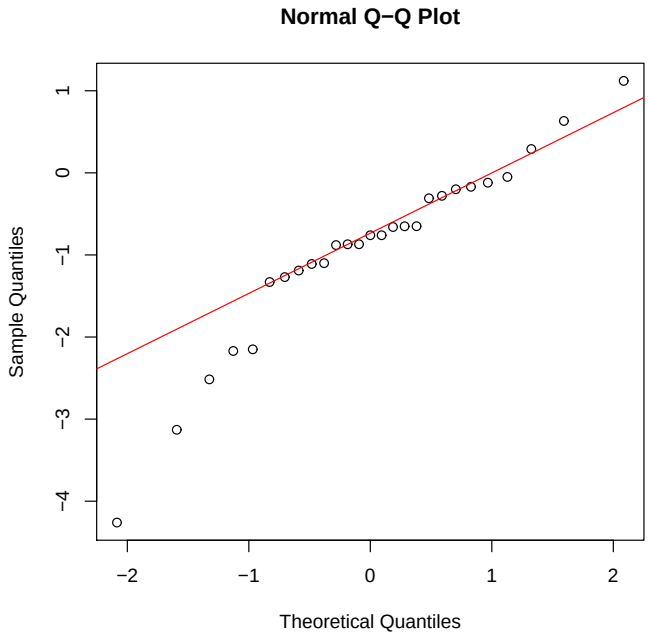


Figure 5.1: The Q-Q plot: the trial effects for M =27

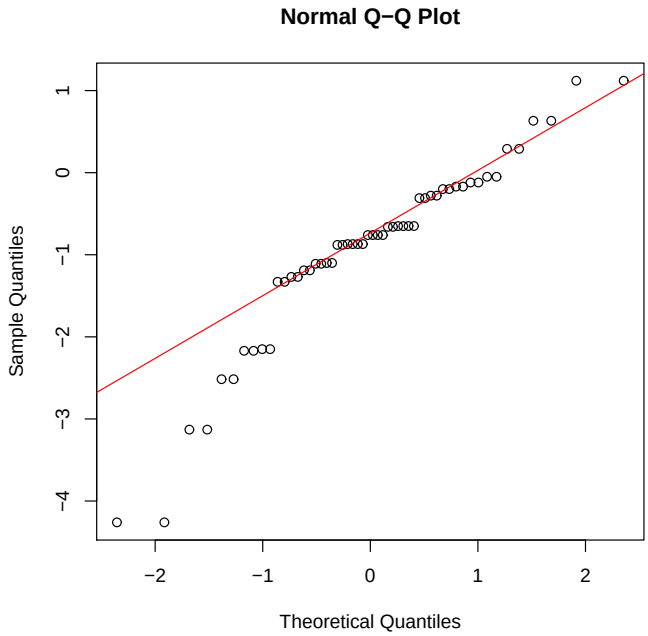


Figure 5.2: The Q-Q plot: the trial effects for M =54

Table 5.2: Simulation study results based on the data generated from S1 with M = 27

Model	parameters (True value)	α		δ_{01}		δ_{02}	
		$\mu_{\alpha 0}$	$\tau_{\alpha 0}$	μ_{01} -1.0	τ_{01} 0.2	μ_{02} -0.3	τ_{02} 0.05
1) M1-F							
	mean			-0.90123		-0.28788	
	r.m.s.e			0.17720		0.10696	
2) M1-R							
	mean			-0.90360	0.08654	-0.28828	0.04605
	r.m.s.e			0.17578	0.18135	0.10718	0.09132
3) M2-F-d							
	mean			-1.00566		-0.30219	
	r.m.s.e			0.16261		0.11259	
4)M2-R-d							
	mean			-1.00931	0.02977	-0.31490	0.03803
	r.m.s.e			0.16640	0.20307	0.27665	0.16084
5) M2-F-N							
	mean	-0.67698	1.00675	-1.01345		-0.73103	
	r.m.s.e			0.20104		0.50908	
6) M2-R-N							
	mean	-0.92330	0.96800	-0.99082	0.04267	0.08951	0.03030
	r.m.s.e			0.20110	0.29985	0.53850	0.53932
7) M3-F							
	mean			-0.99176		-0.29990	
	r.m.s.e			0.16034		0.11168	
8) M3-R							
	mean			-1.00000	0.17030	-0.30141	0.03026
	r.m.s.e			0.16245	0.20459	0.11229	0.14836

Table 5.3: Simulation study results based on the data generated from S1 with M = 54

Model	parameters (True value)	α		δ_{01}		δ_{02}	
		$\mu_{\alpha 0}$	$\tau_{\alpha 0}$	μ_{01} -1.0	τ_{01} 0.2	μ_{02} -0.3	τ_{02} 0.05
1) M1-F							
	mean			-0.78602		-0.26570	
	r.m.s.e			0.25923		0.13718	
2) M1-R							
	mean			-0.78706	0.02606	-0.26569	0.01536
	r.m.s.e			0.25834	0.19311	0.13737	0.06792
3) M2-F-d							
	mean			-1.02375		-0.31619	
	r.m.s.e			0.19243		0.16091	
4) M2-R-d							
	mean			-1.02501	0.02784	-0.30247	0.03933
	r.m.s.e			0.19728	0.21311	0.25730	0.17855
5) M2-F-N							
	mean	-0.92121	0.96196	-1.00059		-0.08400	
	r.m.s.e			0.19461		0.31555	
6) M2-R-N							
	mean	-0.92330	0.96800	-0.99082	0.04267	-0.08951	0.03038
	r.m.s.e			0.21176	0.36846	0.31422	0.35829
7) M3-F							
	mean			-0.98839		-0.30544	
	r.m.s.e			0.18460		0.15488	
8) M3-R							
	mean			-1.00021	0.25955	-0.30686	0.09526
	r.m.s.e			0.18862	0.28027	0.15710	0.24840

Table 5.4: Simulation study results based on the data generated from S2 with M = 27

Model	parameters	α		δ_{01}		δ_{02}	
		$\mu_{\alpha 0}$	$\tau_{\alpha 0}$	μ_{01}	τ_{01}	μ_{02}	τ_{02}
	(True value)	-0.92	0.2	-1.0	0.2	-0.3	0.05
1) M1-F							
	mean			-0.91355		-0.30170	
	r.m.s.e			0.15194		0.07769	
2) M1-R							
	mean			-0.91863	0.11031	-0.30098	0.02966
	r.m.s.e			0.15035	0.07223	0.19608	0.06211
3) M2-F-d							
	mean			-1.03935		-0.30879	
	r.m.s.e			0.16839		0.08244	
4) M2-R-d							
	mean			-1.04805	0.05279	-0.33293	0.04529
	r.m.s.e			0.17288	0.23431	0.15533	0.09380
5) M2-F-N							
	mean	-0.93680	0.19283	-1.02629		-0.36060	
	r.m.s.e	0.09379	0.10139	0.15523		0.16195	
6) M2-R-N							
	mean	-0.93793	0.11769	-1.04984	0.06497	-0.35672	0.00826
	r.m.s.e	0.09579	0.17453	0.17591	0.38942	0.15960	0.17326
7) M3-F							
	mean			-0.97811		-0.31870	
	r.m.s.e			0.16480		0.27833	
8) M3-R							
	mean			-0.91198	0.19711	-0.31237	0.11499
	r.m.s.e			0.34845	0.30415	0.22113	0.27201

Table 5.5: Simulation study results based on the data generated from S2 with M = 54

Model	parameters	α		δ_{01}		δ_{02}	
		$\mu_{\alpha 0}$	$\tau_{\alpha 0}$	μ_{01}	τ_{01}	μ_{02}	τ_{02}
	(True value)	-0.92	0.2	-1.0	0.2	-0.3	0.05
1) M1-F							
	mean			-0.76590		-0.25500	
	r.m.s.e			0.27700		0.13429	
2)M1-R							
	mean			-0.76642	0.01049	-0.25519	0.01110
	r.m.s.e			0.27647	0.19779	0.13428	0.06617
3)M2-F-d							
	mean			-1.01269		-0.31020	
	r.m.s.e			0.19552		0.15237	
4)M2-R-d							
	mean			-1.01784	0.03457	-0.31176	0.02672
	r.m.s.e			0.20172	0.22038	0.27105	0.13303
5)M2-F-N							
	mean	-0.94138	0.15866	-0.97917		-0.30087	
	r.m.s.e	0.11654	0.16723	0.19263		0.23158	
6)M2-R-N							
	mean	-0.94216	0.15912	-1.00124	0.07353	-0.30892	0.01922
	r.m.s.e	0.11721	0.17700	0.20349	0.33496	0.23695	0.29187
7) M3-F							
	mean			-0.97882		-0.27944	
	r.m.s.e			0.18996		0.14711	
8) M3-R							
	mean			-1.00012	0.25748	-0.30244	0.11425
	r.m.s.e			0.19574	0.28954	0.14917	0.26059

Chapter 6

Sensitivity analysis to bivariate normal approximation model

In Chapter 3, we used the empirical log-odds ratio model for the W1 data without considering selection bias. In fact, we do not know how the studies in the W1 data were selected in the meta-analysis. As explained in Chapter 1, various tools to detect selection bias can be used in meta-analysis and in this thesis, we use the funnel plot. If studies with positive results were mostly selected in the meta-analysis then it could make the meta-analysis positively biased. Conversely, if more studies with negative results were selected then the meta-analysis would be negatively biased. In either case, the results may give us incorrect results. To solve this problem, we will use a selection model to investigate the mechanism of selection process. The empirical log-odds ratio model will be used as a standard meta-analysis model in this chapter. The exact logistic regression model will be discussed in the next chapter.

The funnel plot has been widely used to detect selection bias in medical research. Egger et al. (1997) concluded from the investigation of the funnel plot with 37 meta-analyses that the funnel plots provided a useful test for the likely presence of bias in meta-analyses, but

the capacity to detect bias will be limited to a small number of studies in meta-analysis. Copas and Shi (2000) used the funnel plot (the relative risk against the standard deviation) to reanalyse the 37 published epidemiological studies of passive smoking and lung cancer data and proposed a sensitivity analysis method to address the problem of selection bias. Song et al. (2002) examined a funnel plot along with three other statistical methods: rank correlation, regression analysis and Trim and Fill, to 28 meta-analyses from the Database of Abstracts of Reviews of Effectiveness (DARE).

There are various approaches that a researcher confronting the problem of selection bias may take. One is to apply a selection model for bias using a weight function to represent the process of selection. Several classes of selection model have been proposed. Iyengar and Greenhouse (1988) employed the selection model, or weighted distributions, to deal with bias and corrected the results. They also suggested using families of weight functions to model plausible biasing mechanisms to study the sensitivity analysis of inferences about the treatment effects. A similar idea was studied in the area of education, see Hedges (1984). Alternatively, the weight function of the selection model can be defined depending on the treatment effect estimate and its standard error (Copas, 1999; Copas and Shi, 2001, 2002); because some parameters are inestimable and a sensitivity analysis has to be conducted. We will use the similar idea to address the problem of selection bias in meta-analysis for multi-arm trials.

The chapter is outlined as follows. Section 6.1 describes how to detect selection bias in the multi-arm trials model. Section 6.2 illustrates selection bias including the population and selection models, and some mathematical consequences are also given. Section 6.3 presents the likelihood of combined models between the empirical log-odds ratio models and the selection models. Section 6.4 shows a goodness-of-fit test for the funnel plots of combined models. The details of the procedure for sensitivity analysis are described in Section 6.5.

Section 6.6 examines the use of sensitivity analysis with the simulated data. Some related theorems and derivations applying to this chapter are proved in Section 6.7. Finally, some comments are made in Section 6.8. Throughout the chapter, the W1 data will be used to illustrate the idea and the model. There is no difficulty to extend to other data sets.

6.1 Identifying selection bias in multi-arm trials

The basic idea of funnel plot is to plot the estimated treatment effects from individual studies (e.g. empirical log-odds ratios) against their standard errors. If a set of studies is a good sample of a meta-analysis, the funnel plot will be symmetrical between the negative and positive on the treatment effect estimate axis. Asymmetry is a sign of selection bias (see detailed discussion in Rothstein et al., chapter 4, 2005). In multi-arm trials data, there are multiple-pairwise comparisons in RCTs. We thus need to consider the funnel plot in each pairwise-comparison involved in those studies. By using the empirical log-odds ratio model for the W1 data in Chapter 3, it would be convenient and reasonable to use the empirical log-odds ratio and its standard error on the axes because these quantities are already available in the data set.

Recall that the W1 data is partitioned into four groups of studies: $G_1 = \{1, \ldots, 6\}, G_2 = \{7, \ldots, 10\}, G_3 = \{11, \ldots, 24\}$ and $G_4 = \{25, \ldots, 31\}$ where the studies in G_1, G_2, G_3 and G_4 compare treatments A versus B versus C, A versus B, A versus C, B versus C respectively. Let $Y_{i,AC}$, $Y_{i,BC}$ and $Y_{i,AB}$ be the empirical log-odds ratios between the treatments A versus C, B versus C and A versus B, and let $s_{i,AC}$, $s_{i,BC}$ and $s_{i,AB}$ be their respective standard errors. To detect selection bias, we shall apply the funnel plot to the individual studies in each group of multi-arm trials with the empirical log-odds ratio on the vertical axis and the standard error on the horizontal axis. Note that the studies in G_2 are not considered here since there are only indirect comparisons in G_2.

For the W1 data, we consider the number of 'event' that the patients in whom reocclusion on treatments (A, B and C) was detected. Thus, the negative value of, for example, $Y_{i,AC}$ means positive effect. For convenience, we multiplied the value -1 to all the empirical log-odds ratios in our analysis. Thus, the larger positive value of $Y_{i,AC}$ means the more positive effect of treatment A comparing to the control group C. The two funnel plots for G_1: $Y_{i,AC}$ against $s_{i,AC}$ and $Y_{i,BC}$ against $s_{i,BC}$, are displayed in Figures 6.1(a) and 6.1(b) respectively. The funnel plots corresponding to G_3 and G_4 are given in Figures 6.1(c) and 6.1(d) respectively. There are strong tendencies in the funnel plots displayed in Figures 6.1(a) and 6.1(b). Also, signs of selection bias can be seen in the top right-hand corner of both funnel plots – smaller studies (larger standard errors) give more positive results than larger studies (smaller standard errors). Figure 6.1(c) shows a set of studies in G_3 with no evidence of selection bias. Plot 6.1(d) is asymmetrical with a suggestive lack of studies in the bottom right-hand corner. It shows that small studies with negative results are missing in G_4.

From Figures 6.1(a), 6.1(b) and 6.1(d), the problem of selection bias has arisen in G_1 and G_4. As a result, we suspect that there might be other small studies, comparing treatments A, B and C and treatments B and C, respectively which have been carried out or published, but which have not been selected in our meta-analysis.

6.2 Selection bias

The empirical log-odds ratio models for the W1 data are defined in (3.39) - (3.42) of Chapter 3. All treatment effects are assumed to be random. The estimated treatment effects (empirical log-odds ratios) $Y_{i,AC}$, $Y_{i,BC}$, and $Y_{i,AB}$, and their standard errors $s_{i,AC}$, $s_{i,BC}$, and $s_{i,AB}$ are known from the meta-analysis.

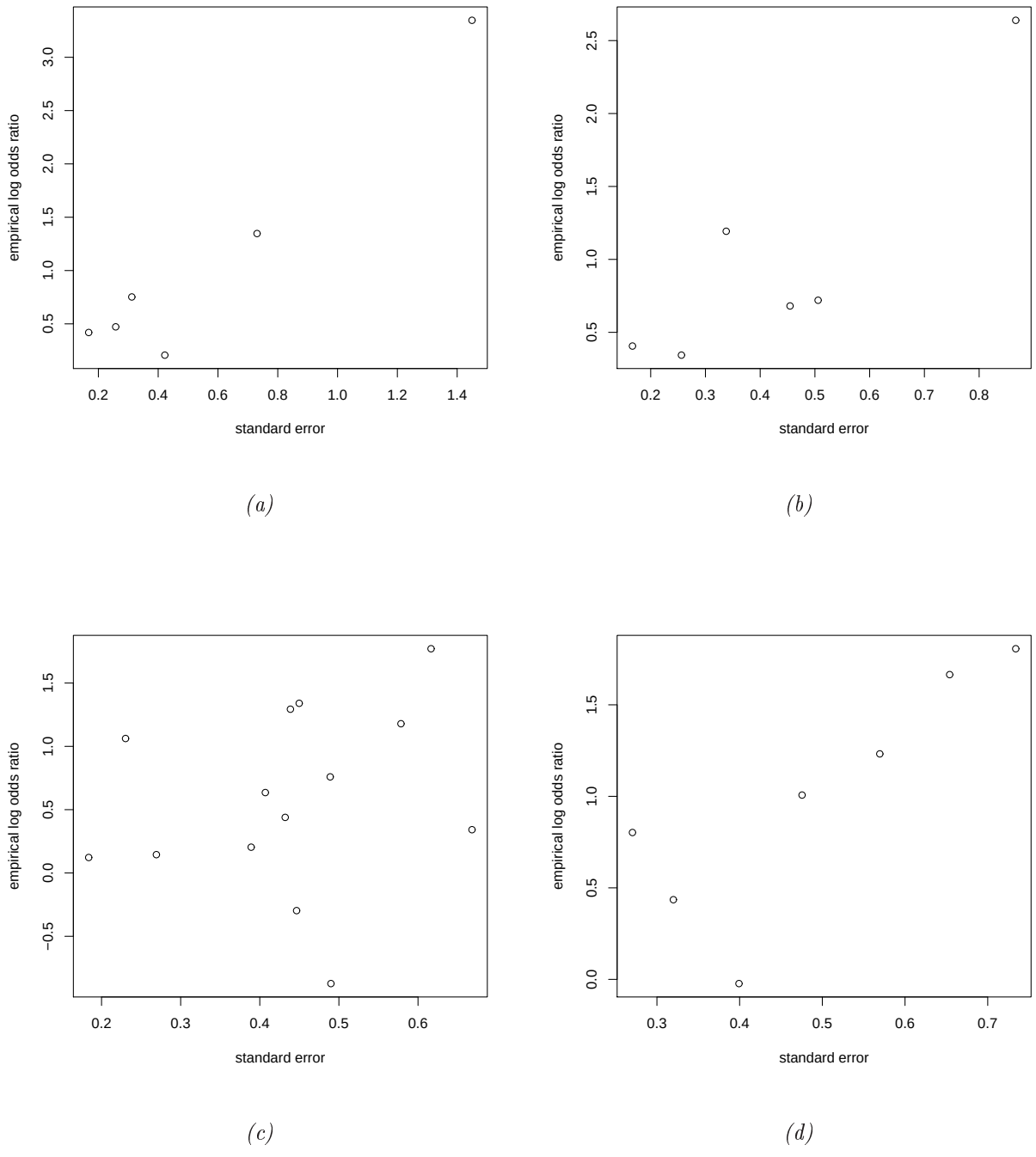


Figure 6.1: The funnel plots:(a) $Y_{i,AC}$ against $s_{i,AC}$ for G_1; (b) $Y_{i,BC}$ against $s_{i,BC}$ for G_1;(c) $Y_{i,AC}$ against $s_{i,AC}$ for G_3;(d) $Y_{i,BC}$ against $s_{i,BC}$ for G_4.

6.2.1 Assumption for population model

From the previous section, there is a presence of selection bias in G_1 and G_4. Recall the empirical log-odds ratios models for G_1 and G_4:

$$\begin{cases} Y_{i,AC} = \delta_{i,AC} + \sigma_{i,AC}\epsilon_{i,AC}, \\ Y_{i,BC} = \delta_{i,BC} + \sigma_{i,BC}\epsilon_{i,BC}, \end{cases} \tag{6.1}$$

$$Y_{i,BC} = \delta_{i,BC} + \sigma_{i,BC}\epsilon_{i,BC}. \tag{6.2}$$

We will make some assumptions for both models to allow us to explore the selection process. We assume that the models (6.1) and (6.2) represent the population of studies, comparing treatment A versus B versus C and treatment B versus C respectively, that have been or could be carried out. In theory, the empirical log-odds ratios are not dependent on their standard errors (Copas and Shi, 2002). For example, $Y_{i,AC}$ and $s_{i,AC}$ are independent to each other. From now on, the models (6.1) and (6.2) are our population models.

6.2.2 Selection model

We first define a selection model for studies in G_1 via a latent variable Z_{i1}. The ith study is selected when Z_{i1} is greater than zero. The latent variable Z_{i1} is defined by

$$Z_{i1} = a_1 + \frac{b_1}{\varphi_i} + \xi_i, \tag{6.3}$$

where ξ_i is a standard normal distribution $N(0,1)$. By adding this selection model to the population model (6.1), the random residuals $(\epsilon_{i,AC}, \xi_i)$ and $(\epsilon_{i,BC}, \xi_i)$ are bivariate normal distributions with both means equal to zero and both variances equal to one. Also their correlations ϱ_1 and ϱ_2 are respectively as

$$corr(\epsilon_{i,AC}, \xi_i) = \varrho_1 \qquad \text{and} \qquad corr(\epsilon_{i,BC}, \xi_i) = \varrho_2.$$

The latent variable Z_{i1} in (6.3) can be interpreted as the inclination for the selection. The quantity φ_i is the average of the standard errors involved in the ith study, given by $(s_{i,AC} + s_{i,BC})/2$. Thus, larger study will have smaller value of φ_i. The parameter a_1 controls the overall proportion of the studies selected; parameter b_1 controls how fast the probability of selection increases as φ_i decreases. In practice, the parameters b_1, ϱ_1 and ϱ_2 are expected to be positive. We will explain this later.

As mentioned earlier, the outcome $(Y_{i,AC}, Y_{i,BC})$ in the population model (6.1) will be selected only if Z_{i1} is greater than 0. In other words, a study comparing treatments A, B and C will be selected in the meta-analysis if and only if the value of the random quantity Z_{i1} is positive. Therefore, the available data from G_1 (the 6 studies from Table 1 of Chapter 2) can be written as $(Y_{i,AC}, Y_{i,BC})|Z_{i1} > 0$ and the related density function for those observations is $p(Y_{i,AC}, Y_{i,BC}|Z_{i1} > 0)$.

If the population and the selection models are independent then the correlations ϱ_1 and ϱ_2 are zero. This will be the ordinary bivariate normal distribution of $(Y_{i,AC}, Y_{i,BC})$. Also, it indicates that the set of studies from the original model is a well-selected sample of the meta-analysis (no selection bias in the model). If $\varrho_1 > 0$ or $\varrho_2 > 0$ then the selected studies will have $Z_{i1} > 0$, and are more likely to have positive ξ_i and positive $\epsilon_{i,AC}$ or $\epsilon_{i,BC}$, leading to a positive bias value of $(Y_{i,AC}, Y_{i,BC})$.

We can define a selection model for studies in G_4 similarly. Let Z_{i2} be a latent variable, the selection model is defined by

$$Z_{i2} = a_2 + \frac{b_2}{s_{i,BC}} + \xi_i, \tag{6.4}$$

where ξ_i is normally distributed as $N(0, 1)$. The ith study in G_4 is selected when Z_{i2} is greater than zero. The random residual $(\epsilon_{i,BC}, \xi_i)$ is a bivariate normal distribution with both means equal to zero and both variances equal to one. The correlation between $\epsilon_{i,BC}$

and ξ_i is ϱ_3. Notice that the denominator of b_2 is the standard error of $s_{i,BC}$ because there are only two arms in G_4.

6.2.3 Relating mathematical consequences

Some related mathematical consequences for the population model (6.1) and the selection model (6.3) for Z_{i1} are given here. All proofs are given in Section 6.7. The equations below can be derived in a similar way to the population model (6.2) and the selection model (6.4).

(i). **The probability of selection**

From the selection model (6.3), the marginal probability of selection can be calculated as

$$p(Z_{i1} > 0|\varphi_i) = \Phi\left(a_1 + \frac{b_1}{\varphi_i}\right), \tag{6.5}$$

where Φ is a standard normal cumulative distribution (see the proof of (6.5) in Theorem 6.7.1). If the parameters a_1 and b_1 are fixed in the selection model then the probability will depend only on the function φ_i. For example, if φ_i is small (small $s_{i,AC}$ and $s_{i,BC}$, from a large study) then the probability of selection is close to 1. In contrast, it makes the selection probability less than 1 for large φ_i (i.e. the small study). Thus, if b_1 is positive then large studies (small φ_i) are more likely to be selected than small studies.

From (6.5), the selection probability is determined by both parameters a_1 and b_1. As mentioned in the previous section, the value of a_1 controls the overall level of selection probability while b_1 controls how the chance of selection depends on the study size. In practice, we need to restrict that b_1 is greater than zero because the results of large studies are usually required to report, no matter that the finding is positive or negative (i.e have large selection probability). While the small studies with negative results are easy to be ignored (either rejected by journals or constrained by researchers themselves). It is important to note that a_1 and b_1 cannot be estimated

from the available data because the unselected studies in the population model (6.1) are unknown.

(ii). **The probability of selection for a typical study**

The probability of a study with the same observations as ith study being selected can be calculated by

$$p(Z_{i1} > 0|(Y_{i,AC}, Y_{i,BC})) = \Phi\left(\frac{\mu_{2i1}}{\sigma_{2i1}}\right), \tag{6.6}$$

where μ_{2i1} and σ^2_{2i1} are given by

$$\mu_{2i1} = \left(a_1 + \frac{b_1}{\varphi_i}\right) + \mathbf{w}_{12}\mathbf{w}_{22}^{-1}\begin{pmatrix} Y_{i,AC} - \mu_{AC} \\ Y_{i,BC} - \mu_{BC} \end{pmatrix}, \tag{6.7}$$

$$\sigma^2_{2i1} = \mathbf{w}_{11} - \mathbf{w}_{12}\mathbf{w}_{22}^{-1}\mathbf{w}_{21}, \tag{6.8}$$

where $\mathbf{w}_{11} = (1)$, $\mathbf{w}_{12} = \left(\varrho_1\sqrt{v_{1i}}, \varrho_2\sqrt{v_{2i}}\right)$, $\mathbf{w}_{21} = (\varrho_1\sqrt{v_{1i}}, \varrho_2\sqrt{v_{2i}})^t$ and $\mathbf{w}_{22}$ is

$$\begin{pmatrix} v_{1i} & v_{12i} \\ v_{21i} & v_{2i} \end{pmatrix} = \begin{pmatrix} \tau^2_{AC} + \sigma^2_{i,AC} & \rho\tau_{AC}\tau_{BC} + \sigma^2_{iC} \\ \rho\tau_{AC}\tau_{BC} + \sigma^2_{iC} & \tau^2_{BC} + \sigma^2_{i,BC} \end{pmatrix}.$$

The proof of (6.6) is given in Theorem 6.7.2. The probability (6.6) is a measure to determine that how much chance the outcome of ith study will be selected. The equation (6.6) will give a larger selection probability for larger values of $Y_{i,AC}$ or $Y_{i,BC}$.

(iii). **The means for selected studies**

The means of the log-odds ratios for selected studies are

$$E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i\right) = \begin{pmatrix} \mu_{AC} \\ \mu_{BC} \end{pmatrix} + \begin{pmatrix} \varrho_1\sigma_{iAC} \\ \varrho_2\sigma_{iBC} \end{pmatrix} \lambda\left(a_1 + \frac{b_1}{\varphi_i}\right). \tag{6.9}$$

The $\lambda(\cdot)$ is Mill's ratio $\phi(\cdot)/\Phi(\cdot)$, where ϕ and Φ are the density and distribution functions respectively, of the standard normal distribution (see the proof of (6.9) in

Theorem 6.7.3). The equation (6.9) gives the average of log-odds ratios for selected studies from the population model, allowing different amounts of selection bias. This average depends on the pair (a_1, b_1) in the selection model (6.3). It is also an increasing function of $(\sigma_{i,AC}, \sigma_{i,BC})$ and a decreasing function of φ_i. Since the second term in (6.9) is larger than zero, the selected studies has a larger mean than overall mean (μ_{AC}, μ_{BC}). For smaller studies of $(Y_{i,AC}, Y_{i,BC})$, it has even larger mean since Mill's ratio $\lambda(a_1 + b_1/\varphi_i)$ has larger values. Thus, the model can be used to model the data shown in Figure 6.1.

(iv). **Variance of selected studies**

From equation (6.9), the variance of the selected outcomes can be defined as

$$Var\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i\right) = \begin{pmatrix} \sigma^2_{i,AC}(1 + d^2_{i1}\varrho^2_1) & \sigma^2_{iC} \\ \sigma^2_{iC} & \sigma^2_{i,BC}(1 + d^2_{i1}\varrho^2_2) \end{pmatrix}, \quad (6.10)$$

where

$$d^2_{i1} = \lambda\left(a_1 + \frac{b_1}{\varphi_i}\right)\left(a_1 + \frac{b_1}{\varphi_i} + \lambda\left(a_1 + \frac{b_1}{\varphi_i}\right)\right).$$

The proof of (6.10) is given in Theorem 6.7.4. Now, we need to distinguish between $(\sigma^2_{i,AC}, \sigma^2_{i,BC})$ and $(s^2_{i,AC}, s^2_{i,BC})$. The $\sigma^2_{i,AC}$ and $\sigma^2_{i,BC}$ are the variances of the population models and may be written as

$$\sigma^2_{i,AC} = Var(Y_{i,AC}|\delta_{i,AC}) \qquad \text{and} \qquad \sigma^2_{i,BC} = Var(Y_{i,BC}|\delta_{i,BC}).$$

The parameters $s^2_{i,AC}$ and $s^2_{i,BC}$ are the variances of our meta-analysis, estimating from

$$s^2_{i,AC} = Var(Y_{i,AC}|Z_{i1} > 0) \qquad \text{and} \qquad s^2_{i,BC} = Var(Y_{i,BC}|Z_{i1} > 0).$$

For example, $\sigma^2_{i,AC}$, $\sigma^2_{i,BC}$ and σ^2_{iC} in (6.10) may be written, respectively, as

$$\sigma^2_{i,AC} = \frac{s^2_{i,AC}}{(1 + d^2_{i1}\varrho^2_1)}, \qquad \sigma^2_{i,BC} = \frac{s^2_{i,BC}}{(1 + d^2_{i1}\varrho^2_2)} \qquad \text{and} \qquad \sigma^2_{iC} = s^2_{iC}. \tag{6.11}$$

Note that the $\sigma^2_{i,AC}$ and $\sigma^2_{i,BC}$ in equation (6.9) are replaced by (6.11).

6.3 Likelihood

In this section, we will calculate MLEs of all unknown parameters by assuming that the values of (a_1, b_1) and (a_2, b_2) are given. As the previous section, we still use the W1 data as our illustrative example. The population models (6.1) and (6.2) are combined with the selection models (6.3) and (6.4) for G_1 and G_4 respectively; the empirical log odds ratio models for G_2 and G_3 are the same (without selectivity) as defined in Chapter 3 (models (3.40) and (3.41)). The log-likelihood function for the empirical log odds models with selection models (with selectivity) can be written as

$$\begin{aligned} l(\boldsymbol{\theta}) \;=\; & \sum_{i \in G_1} \log p((Y_{i,AC}, Y_{i,BC})|Z_{i1} > 0) + \sum_{i \in G_2} \log p(Y_{i,AB}|\boldsymbol{\theta}) \\ & + \sum_{i \in G_3} \log p(Y_{i,AC}|\boldsymbol{\theta}) + \sum_{i \in G_4} \log p(Y_{i,BC}|Z_{i2} > 0). \end{aligned} \tag{6.12}$$

As discussed in Chapter 3, the heterogeneity parameters are assumed to be the same: $\tau^2_{AC} = \tau^2_{BC} = \tau^2_{AB} = \tau^2$ and the correlation coefficient between the treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ takes the value 1/2. The collection of all unknown parameters is

$$\boldsymbol{\theta} = \left\{\mu_{AC}, \mu_{BC}, \tau^2, \varrho_1, \varrho_2, \varrho_3\right\}. \tag{6.13}$$

The likelihoods $p(Y_{i,AB}|\boldsymbol{\theta})$ and $p(Y_{i,AC}|\boldsymbol{\theta})$ are the same as given in Chapter 3. The log-

likelihood function for G_1 (the first term of (6.12)) can be written as

$$\begin{aligned} l_{G_1} &= \sum_{i \in G_1} \log p((Y_{i,AC}, Y_{i,BC}) | Z_{i1} > 0) \\ &= \sum_{i \in G_1} \left(\log p(Z_{i1} > 0 | (Y_{i,AC}, Y_{i,BC})) + \log p(Y_{i,AC}, Y_{i,BC}) - p(Z_{i1} > 0 | \varphi_i) \right). \end{aligned}$$

The formulae of $p(Z_{i1} > 0|(Y_{i,AC}, Y_{i,BC}))$ and $p(Z_{i1} > 0|\varphi_i)$ are given by Theorems 6.7.2 and 6.7.1 respectively. Note that $(Y_{i,AC}, Y_{i,BC})$ has a bivariate normal distribution shown in (3.38), we therefore have

$$\begin{aligned} l_{G_1} &= \sum_{i=1}^{6} \left(-\frac{1}{2} \log(\tau^2 + \sigma^2_{i,AC})(\tau^2 + \sigma^2_{i,BC})(1 - R_i^2) + \log \Phi(\mu_{2i1}/\sigma_{2i1}) - \log \Phi(a_1 + b_1/\varphi_i) \right) \\ &\quad - \sum_{i=1}^{6} \frac{1}{2(1 - R_i^2)} \left(\frac{(Y_{i,AC} - \mu_{AC})^2}{\tau^2 + \sigma^2_{i,AC}} - \frac{2R_i (Y_{i,AC} - \mu_{AC})(Y_{i,BC} - \mu_{BC})}{\sqrt{\tau^2 + \sigma^2_{i,AC}} \sqrt{\tau^2 + \sigma^2_{i,BC}}} + \frac{(Y_{i,BC} - \mu_{BC})^2}{\tau^2 + \sigma^2_{i,BC}} \right), \end{aligned}$$

where

$$R_i = \frac{\rho\tau^2 + \sigma^2_{iC}}{\sqrt{\tau^2 + \sigma^2_{i,AC}} \sqrt{\tau^2 + \sigma^2_{i,BC}}}.$$

The $\Phi(\mu_{2i1}/\sigma_{2i1})$ is obtained from the function $p(Z_{i1} > 0|(Y_{i,AC}, Y_{i,BC}))$ where μ_{2i1} and σ^2_{2i1} are given in (6.7) and (6.8) respectively. The $\Phi(a_1 + b_1/\varphi_i)$ is derived from $p(Z_{i1} > 0|\varphi_i)$. The parameter R_i is the correlation between $Y_{i,AC}$ and $Y_{i,BC}$. Its numerator $\rho\tau^2 + \sigma^2_{i,C}$ is the covariance of $Y_{i,AC}$ and $Y_{i,BC}$; the first term $(\rho\tau^2)$ is the covariance of $\delta_{i,AC}$ and $\delta_{i,BC}$, and the second (σ^2_{iC}) is the covariance of the random sampling errors from both models. In the denominator of R_i, the $\tau^2 + \sigma^2_{i,AC}$ and $\tau^2 + \sigma^2_{i,BC}$ are the variances of the models $Y_{i,AC}$ and $Y_{i,BC}$ respectively. The $\sigma^2_{i,AC}$, $\sigma^2_{i,BC}$ and σ^2_{iC} are replaced by (6.11).

Similarly, we can calculate the log-likelihood for studies in G_4, which is

$$\begin{aligned} l_{G_4} &= \sum_{i \in G_4} \log p(Y_{i,BC} | Z_{i2} > 0), \\ &= \sum_{i \in G_4} \left(\log p(Z_{i2} > 0 | Y_{i,BC}) + \log p(Y_{i,BC}) - p(Z_{i2} > 0 | s_{i,BC}) \right). \end{aligned}$$

The formulae $p(Y_{i,BC}|Z_{i2} > 0)$ and $p(Z_{i2} > 0|Y_{i,BC})$ can be expressed in the same way as Theorems 6.7.2 and 6.7.1 respectively. The $p(Y_{i,BC})$ is a density function of normal distribution. Therefore, the log-likelihood l_{G_4} is

$$l_{G_4} = \sum_{i \in G_4} \left(\frac{1}{2} \log(\tau^2 + \sigma^2_{i,BC}) - \frac{(Y_{i,BC} - \mu_{BC})^2}{2(\tau^2 + \sigma^2_{i,BC})} - \log \Phi(a_2 + b_2/s_{i,BC}) + \log(\mu_{2i2}/\sigma_{2i2}) \right),$$

where μ_{2i2} and σ_{2i2} are $E(Z_{i2}|Y_{i,BC})$ and $Var(Z_{i2}|Y_{i,BC})$ respectively, given by

$$\begin{aligned} \mu_{2i2} &= a_2 + \frac{b_2}{s_{i,BC}} + \varrho_3 \frac{(Y_{i,BC} - \mu_{BC})}{(\tau^2 + \sigma^2_{i,BC})^{1/2}}, \\ \sigma_{2i2} &= (1 - \varrho_3^2)^{1/2}. \end{aligned}$$

Also, the $\sigma^2_{i,BC}$ in l_{G_4} is replaced by $\sigma^2_{i,BC} = s^2_{i,BC}(1 + d^2_{i2}\varrho^2_3)$ where

$$d^2_{i2} = \lambda \left(a_2 + \frac{b_2}{s_{i,BC}} \right) \left(a_2 + \frac{b_2}{s_{i,BC}} + \lambda \left(a_2 + \frac{b_2}{s_{i,BC}} \right) \right).$$

From the log-likelihood functions l_{G_1} and l_{G_4}, the parameters (a_1, b_1) and (a_2, b_2) are not estimable because we do not know how many unpublished studies, comparing treatments A, B and C and treatments B and C, may have been carried out. Therefore, these parameters will be treated as free parameters in the sensitivity analysis. If the pairs (a_1, b_1) and (a_2, b_2) are given then the MLE for $\boldsymbol{\theta}$ can be estimated by maximizing the log-likelihood function directly.

It would be of interest to test the overall means of the treatment effect. For example, to test

the hypothesis $H_0 : \mu_{AC} = 0$ and $H_1 : \mu_{AC} \neq 0$, we can use the following likelihood ratio statistic:

$$2\left(l(\widehat{\boldsymbol{\theta}}) - l(\widehat{\boldsymbol{\theta}}_{\mu_{AC}=0})\right) \sim \chi^2_1 \text{ under } H_0, \tag{6.14}$$

where $\widehat{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$ and $\widehat{\boldsymbol{\theta}}_{\mu_{AC}=0}$ is the MLE of $\boldsymbol{\theta}$ with restriction $\mu_{AC} = 0$. The hypothesis test for $H_0 : \mu_{BC} = 0$ can be considered in the same way.

6.4 Goodness of fit

In this section, we suppose that the pairs (a_1, b_1) and (a_2, b_2) are given in the selection models (6.3) and (6.4) or the log-likelihood function (6.12). We will explain how to infer these pairs in the next section. From the profile of the log-likelihood function in the previous section, if a set of specific parameters (a_1, b_1, a_2, b_2) is a possible set for the selection models (6.3) and (6.4), then we need to check that the resulting models (combined models) from these selection models give reasonable fits to the data in funnel plots. For a study in G_1, if a selection model with a specific pair (a_1, b_1) is used, the mean of selected studies is given by (6.9). If another pair (a_1^*, b_1^*) is used, the difference of the means by selection model with these two pairs is given by

$$\begin{aligned} & E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i, a_1^*, b_1^*\right) - E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i, a_1, b_1\right) \\ \approx \;& c^* + \begin{pmatrix} \varrho_1 \\ \varrho_2 \end{pmatrix} (\lambda(a_1^*) - \lambda(a_1)) \begin{pmatrix} s_{i,AC} \\ s_{i,BC} \end{pmatrix}, \end{aligned} \tag{6.15}$$

where c^* is constant (see the proof of (6.15) in Theorem 6.7.5) when (a_1^*, b_1^*) is close to (a_1, b_1). The equation (6.15) is a linear equation in terms of $s_{i,AC}$ and $s_{i,BC}$. This suggests that local departures of the model in terms of (a_1, b_1) will be similar to adding the linear term β_1 in $s_{i,AC}$ and β_2 in $s_{i,BC}$ to the expected value of $(Y_{i,AC}, Y_{i,BC})$. Therefore, testing that if there

is another pair (a_1^*, b_1^*) better than (a_1, b_1) is equivalent to test $H_0 : \beta_1 = 0$ and $\beta_2 = 0$ in the following models.

$$\begin{aligned} Y_{i,AC} &= \delta_{i,AC} + \beta_1 s_{i,AC} + \sigma_{i,AC}\epsilon_{i,AC}, \\ Y_{i,BC} &= \delta_{i,BC} + \beta_2 s_{i,BC} + \sigma_{i,BC}\epsilon_{i,BC}, \\ Z_{i1} &= a_1 + \frac{b_1}{\varphi_i} + \xi_i. \end{aligned}$$

In a similar way to the selection model (6.4), the difference of the means with these two pairs (a_2, b_2) and (a_2^*, b_2^*) is

$$\begin{aligned} & E(Y_{i,BC}|Z_{i2} > 0, s_{i,BC}, a_2^*, b_2^*) - E(Y_{i,BC}|Z_{i2} > 0, s_{i,BC}, a_2, b_2) \\ \approx\ & c^* + \varrho_3(\lambda(a_2^*) - \lambda(a_2))s_{i,BC}. \end{aligned} \tag{6.16}$$

The proof of (6.16) can be obtained in a similar way as Theorem 6.7.5. Similar idea to (6.15), we add the term β_3 in $s_{i,BC}$ to the expected value of $Y_{i,BC}$. The refitted population model (6.2) and its selection model can be written as

$$\begin{aligned} Y_{i,BC} &= \delta_{i,BC} + \beta_3 s_{i,BC} + \sigma_{i,BC}\epsilon_{i,BC}, \\ Z_{i2} &= a_2 + \frac{b_2}{s_{i,BC}} + \xi_i. \end{aligned}$$

The purpose here is to consider the fit tests for the combined models $(Y_{i,AC}, Y_{i,BC}|Z_{i1} > 0)$ and $(Y_{i,BC}|Z_{i2} > 0)$ at the same time in the meta-analysis. To test whether or not the set $(a_1^*, b_1^*, a_2^*, b_2^*)$ is better than (a_1, b_1, a_2, b_2), we use the likelihood ratio test

$$H_0 : \boldsymbol{\beta} = \mathbf{0} \ \text{ versus } \ H_1 : \boldsymbol{\beta} \neq \mathbf{0},$$

where $\boldsymbol{\beta}$ is the vector $(\beta_1, \beta_2, \beta_3)$. If the null hypothesis is accepted it means that the selection models (6.3) and (6.4) have satisfactorily explained the linear relationships between

$(Y_{i,AC}, Y_{i,BC})$ and $(s_{i,AC}, s_{i,BC})$, and between $Y_{i,BC}$ and $s_{i,BC}$. The other meaning is that the set (a_1, b_1, a_2, b_2) makes the funnel plots of combined models fit well.

To test a goodness-of-fit for any given (a_1, b_1, a_2, b_2), the log-likelihood function (6.12) can be extended by adding the term $\beta_1 s_{i,AC}$ to μ_{AC} and the term $\beta_2 s_{i,BC}$ to μ_{BC} in the log-likelihood function l_{G_1}, and adding the term $\beta_3 s_{i,BC}$ to μ_{BC} in the log-likelihood function l_{G_4}. Thus

$$l^*(\boldsymbol{\theta}, \boldsymbol{\beta}) = l^*_{G_1}(\boldsymbol{\theta}, \boldsymbol{\beta}) + l_{G_2}(\boldsymbol{\theta}) + l_{G_3}(\boldsymbol{\theta}) + l^*_{G_4}(\boldsymbol{\theta}, \boldsymbol{\beta}), \tag{6.17}$$

where $l_{G_2}(\boldsymbol{\theta})$ and $l_{G_3}(\boldsymbol{\theta})$ are the same as given in Chapter 3. The log-likelihood $l^*_{G_1}(\boldsymbol{\theta}, \boldsymbol{\beta})$ is given by

$$\begin{aligned}
&\sum_{i=1}^{6}\left(-\frac{1}{2}\log((\tau^2+\sigma^2_{i,AC})(\tau^2+\sigma^2_{i,BC})(1-R_i^2)+\log\Phi(\mu^*_{2i1}/\sigma^*_{2i1})-\log\Phi(a_1+b_1/\varphi_i)\right)\\
&-\sum_{i=1}^{6}\frac{1}{2(1-R_i^2)}\left(\frac{(Y_{i,AC}-\mu_{AC}-\beta_1 s_{i,AC})^2}{\tau^2+\sigma^2_{i,AC}}-\frac{2R_i(Y_{i,AC}-\mu_{AC}-\beta_1 s_{i,AC})(Y_{i,BC}-\mu_{BC}-\beta_2 s_{i,BC})}{\sqrt{\tau^2+\sigma^2_{i,AC}}\sqrt{\tau^2+\sigma^2_{i,BC}}}\right)\\
&+\sum_{i=1}^{6}\frac{1}{2(1-R_i^2)}\left(\frac{(Y_{i,BC}-\mu_{BC}-\beta_2 s_{i,AC})^2}{\tau^2+\sigma^2_{i,BC}}\right).
\end{aligned}$$

Note that μ^*_{2i1} and σ^*_{2i1} are similar as defined before but they are added the term $\beta_1 s_{i,AC}$ to μ_{AC} and the term $\beta_2 s_{i,BC}$ to μ_{BC}. The log-likelihood $l^*_{G_4}(\boldsymbol{\theta}, \boldsymbol{\beta})$ can be calculated similarly. Then, the likelihood ratio statistic for the null hypothesis $H_0 : \boldsymbol{\beta} = \mathbf{0}$ is

$$2\left(l^*(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\beta}}) - l^*(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\beta}=0}, \boldsymbol{\beta} = \mathbf{0})\right) \sim \chi^2_3 \text{ under } H_0, \tag{6.18}$$

where $(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\beta}})$ is MLEs by maximizing (6.17) while $(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\beta}=0}, \boldsymbol{\beta} = \mathbf{0})$ is the MLEs from (6.17) with restriction $\boldsymbol{\beta} = 0$.

6.5 Sensitivity analysis

The idea of a sensitivity analysis is to use the selection models (6.3) and (6.4) to the population models (6.1) and (6.2) respectively by allowing different amounts of selection probability in the combined models $(Y_{i,AC}, Y_{i,BC})|Z_{i1} > 0$ and $(Y_{i,BC})|Z_{i2} > 0$, and investigate how sensitive the main interest parameters are changed when compared to the results of the standard model (without selectivity). The main parameters of interest in our meta-analysis are μ_{AC} and μ_{BC}, which are the overall mean effects from the treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ respectively. The procedure of sensitivity analysis is given as follows

- **Step 1**

  Determine the possible ranges of (a_1, b_1) and (a_2, b_2) for the selection models (6.3) and (6.4) by using the marginal selection probabilities $p(Z_{i1} > 0|\varphi_i)$ and $p(Z_{i2} > 0|s_{i,BC})$ respectively.

- **Step 2**

  For each combination of (a_1, b_1, a_2, b_2), we estimate $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ by maximizing (6.17) and use the goodness-of-fit test to test how the meta-analysis model with selection models fit in funnel plots. P-value will be calculated for each test.

- **Step 3**

  We conduct a sensitivity analysis based on p-value of the goodness-of-fit test given in step 2 and other quantities. For example, the overall estimates $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ obtaining from the combined models with p-value < 0.05 should be discarded. We will discuss the details for each step in the following subsections.

6.5.1 The possible range of (a_1, b_1) and (a_2, b_2) (Step 1)

As mentioned earlier, the parameters (a_1, b_1) and (a_2, b_2) cannot be estimated in the usual way; they need to be given in the log-likelihood function. In this section, we shall identify ranges of (a_1, b_1) and (a_2, b_2) which cover all reasonable possibilities for the selection models

(6.3) and (6.4) respectively. We use the selection model for G_1 in the W1 data to demonstrate how to choose such a range. Since the selection probability

$$p(Z_{i1} > 0|\varphi_i, a_1, b_1) = \Phi(a_1 + b_1/\varphi_i)$$

is a decreasing function of φ_i, we obtain

$$P_{min}(\text{selection}) = p(Z_{i1} > 0|\varphi_{max}, a_1, b_1) \qquad \text{and} \qquad P_{max}(\text{selection}) = p(Z_{i1} > 0|\varphi_{min}, a_1, b_1), \tag{6.19}$$

where φ_{max} and φ_{min} are the maximum and minimum values of $\{\varphi_i, i = 1, \ldots, 6\}$. Thus, the selection probability $p(Z_{i1} > 0|\varphi_i, a_1, b_1)$ can be written as

$$P_{min}(\text{selection}) \leq p(Z_{i1} > 0|\varphi_i, a_1, b_1) \leq P_{max}(\text{selection}). \tag{6.20}$$

If we take a grid in the following area

$$0.01 \leq P_{min}(\text{selection}) \leq P_{max}(\text{selection}) \leq 0.99. \tag{6.21}$$

This should cover all reasonable possibilities of selection. Each pair of (P_{min}, P_{max}) is corresponding to a pair of (a_1, b_1). For example, if $(P_{min}, P_{max}) = (0.7, 0.8)$, we have

$$P_{min} = \Phi(a_1 + b_1/\varphi_{max}) = 0.7 \qquad \text{and} \qquad P_{max} = \Phi(a_1 + b_1/\varphi_{min}) = 0.8.$$

For the W1 data, the smallest and largest values of φ are 0.16718 and 0.97771 respectively thus the pair (a_1, b_1) is (0.4589681, 0.06397397)(as shown on row 2 of Table 6.1).

So the first step of sensitivity analysis is to take a grid in the area (6.21) and then transfer them to a set of pairs (a_1, b_1). The range of (a_2, b_2) can be chosen similarly.

In the following sections, we will report the detailed results of the following six typical pairs in the area (6.21):

(0.8,0.7), (0.8,0.5),(0.7,0.4),(0.6,0.3),(0.4,0.1) and (0.2,0.01).

For these pairs, the related (a_1, b_1) for G_1 and (a_2, b_2) for G_4 in the W1 data are given in Tables (6.1) and (6.2) respectively. The first row in the table is corresponding to the model without assuming selection bias.

Table 6.1: The pairs (a_1, b_1) for the selection model (6.3)

Selection probability pairs	a_1	b_1
(1.0,1.0)	6.0	0.0
(0.80,0.70)	0.4589681	0.06397397
(0.80,0.50)	-0.1735993	0.16972995
(0.70,0.40)	-0.4137713	0.15684854
(0.60,0.30)	-0.6848247	0.15684854
(0.40,0.10)	-1.4936369	0.20735823
(0.20,0.01)	-2.6325990	0.29942516

Table 6.2: The pair of (a_2, b_2) for the selection model (6.4)

Selection probability pairs	a_2	b_2
(1.0,1.0)	6.0	0.0
(0.80,0.70)	0.33930785	0.1358371
(0.80,0.50)	-0.49107105	0.3603908
(0.70,0.40)	-0.70714903	0.3330395
(0.60,0.30)	-0.97820244	0.3330395
(0.40,0.10)	-1.88149062	0.4402876
(0.20,0.01)	-3.19265952	0.6357751

From Table 6.1, we can interpret a selection from the population model (6.1). The pair $(a_1, b_1) = (0.4589681, 0.06397397)$ (row 2 of Table 6.1) is calculated from the selection probability pair $(0.80, 0.70)$. This means that the marginal selection probability take 80% and 70% for the largest observed studies (smallest standard errors) and the smallest observed studies (largest standard errors) respectively in the population model (6.1). Also 80% of the largest studies will be selected but 70% of the smallest studies will be selected. Other pairs can be interpreted in the same way.

6.5.2 Estimation and goodness-of-fit test (Step 2)

In the sensitivity analysis, we consider the use of the pairs (a_1, b_1) and (a_2, b_2) together to select the studies from the population models (6.1) and (6.2) respectively. Each combination of (a_1, b_1, a_2, b_2) is corresponding to a particular selection model. The second step in our sensitivity analysis is to calculate the relative statistical quantities (e.g. the p-value of goodness-of-fit test) to judge if the underlying model is a reasonable choice. To do so, the following quantities are calculated for each combination of (a_1, b_1, a_2, b_2) for the W1 data.

1. $\widehat{\mu_{AC}}$;
2. p-value for testing $H_0 : \mu_{AC} = 0$;
3. lower limit of the 95% confidence interval for μ_{AC};
4. upper limit of the 95% confidence interval for μ_{AC};
5. P_{max}(selection) for the selection model (6.3);
6. P_{min}(selection) for the selection model (6.3);
7. estimated number of selected and unselected studies given for G_1 by $\sum_i \{p(Z_{i1} > 0|\varphi_i)\}$;
8. $\widehat{\mu_{BC}}$;
9. p-value for testing $H_0 : \mu_{BC} = 0$;
10. lower limit of the 95% confidence interval for μ_{BC};
11. upper limit of the 95% confidence interval for μ_{BC};
12. P_{max}(selection) for the selection model (6.4);
13. P_{min}(selection) for the selection model (6.4);
14. estimated number of selected and unselected studies given for G_4 by $\sum_i \{p(Z_{i2} > 0|s_{i,BC})\}$;

15. p-value for the fit for the funnel plot corresponding to the null hypothesis $H_0 : \boldsymbol{\beta} = 0$.

The 7th and 14th quantities present the overall severities of the selection models (6.3) and (6.4) respectively. The 15th quantity gives the p-value of goodness-of-fit test discussed in Section 6.4. For the W1 data, we listed the detailed results for seven typical combinations in Table 6.3. The quantities in each row of the table are calculated from the combination (a_1, b_1, a_2, b_2) corresponding to the same row in Tables 6.1 and 6.2. The first row represents the empirical log-odds ratio model without assuming selection bias. The conclusions for Table 6.3 are as follows.

Table 6.3: The W1 data with selection model: summary of outputs

	[, 1]	[, 2]	[, 3]	[, 4]	[, 5]	[, 6]	[, 7]	[, 8]
[1,]	0.5689296	1.7901e-06	0.2386970	0.8991622	1.0	1.00	6	0.6770754
[2,]	0.5438695	9.0884e-06	0.5311785	0.5565605	0.8	0.70	8	0.5802389
[3,]	0.5203228	1.8689e-05	0.4642864	0.5763592	0.8	0.50	9	0.5270842
[4,]	0.5029442	3.5987e-05	0.4417334	0.5641550	0.7	0.40	12	0.4703732
[5,]	0.4840191	6.3934e-05	0.3825303	0.5855079	0.6	0.30	15	0.4085832
[6,]	0.4446604	2.1316e-04	0.4196116	0.4697092	0.4	0.10	37	0.2726371
[7,]	0.4134910	3.9836e-04	0.3018102	0.5251718	0.2	0.01	264	0.1496556
	[, 9]	[, 10]	[, 11]	[, 12]	[, 13]	[, 14]	[, 15]	
[1,]	7.8409e-07	0.294571394	1.0595794	1.0	1.00	7	0.02812794	
[2,]	4.5603e-05	0.477201709	0.6832761	0.8	0.70	9	0.07719169	
[3,]	1.9897e-04	0.472615769	0.5815526	0.8	0.50	11	0.19807872	
[4,]	8.4598e-04	0.276823211	0.6639232	0.7	0.40	13	0.32012129	
[5,]	3.4631e-03	0.318677969	0.4984884	0.6	0.30	17	0.54505213	
[6,]	3.3151e-02	0.183261054	0.3620131	0.4	0.10	42	0.93498257	
[7,]	3.0103e-01	-0.004106355	0.3034176	0.2	0.01	292	0.55372482	

(i). The estimates of $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ corresponding to different amounts of seleciton bias are presented in columns 1 and 8 respectively. By using the asymptotic variance-covariance matrix in Chapter 3, their standard errors from each row of $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ are

$$SD(\widehat{\mu_{AC}}) = \{0.16848, 0.00647, 0.02859, 0.03123, 0.05178, 0.01278, 0.05698\},$$

$$SD(\widehat{\mu_{BC}}) = \{0.19515, 0.05257, 0.02779, 0.09875, 0.04587, 0.0456, 0.07845\}.$$

The lower and upper limits of the 95% confidence intervals for $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ are given in columns 3 and 4, and columns 10 and 11 respectively.

(ii). The inferences of standard model are presented in the first row of the table. The $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ of the standard models are 0.5689296 and 0.677075 respectively. The test for the presence of selection bias is obtained from the likelihood test for the null hypothesis $H_0 : \boldsymbol{\beta} = \mathbf{0}$. The p-value 0.02812794 in the first row confirms that there is strong evidence to reject H_0, i.e, there is selection bias in G_1 and G_4. The p-value of the goodness-of-fit test for the second row (0.077191) is close to 0.05 . This is the evidence that $\widehat{\mu_{AC}} = 0.5689296$ and 0.5438695, and that $\widehat{\mu_{BC}} = 0.677075$ and 0.5802389 are overestimated.

(iii). To consider the number of unselected studies for G_1 and G_4 (columns 7 and 14), rows 2 - 7 show that the numbers of unselected studies (or the study populations of treatments A vs B vs C and B vs C) increase while the estimates of $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ (columns 1 and 8) decrease gradually when reading downwards. However, the extreme number given in row 7 indicates that the underline model is not acceptable.

(iv). The p-values of the null hypothesis $H_0 : \widehat{\mu_{AC}} = 0$ (column 2) are significant on all rows while the p-value at row 7 (column 9) of the null hypothesis $H_0 : \widehat{\mu_{BC}} = 0$ is not significant.

(v). To analyze a goodness-of-fit test for the meta-analysis, we consider the p-value of $H_0 : \boldsymbol{\beta} = 0$ (column 15). The p-values from rows 3 - 7 give good fits for the funnel plots for the models $(Y_{i,AC}, Y_{i,BC}|Z_{i1} > 0)$ and $(Y_{i,BC}|Z_{i2} > 0)$ while the others including the standard estimates are overestimates.

(vi). Overall, the models corresponding to row 1,2 and 7 are not acceptable. The others are plausible. The fits of funnel plot for G_1 given by different values of (a_1, b_1, a_2, b_2)

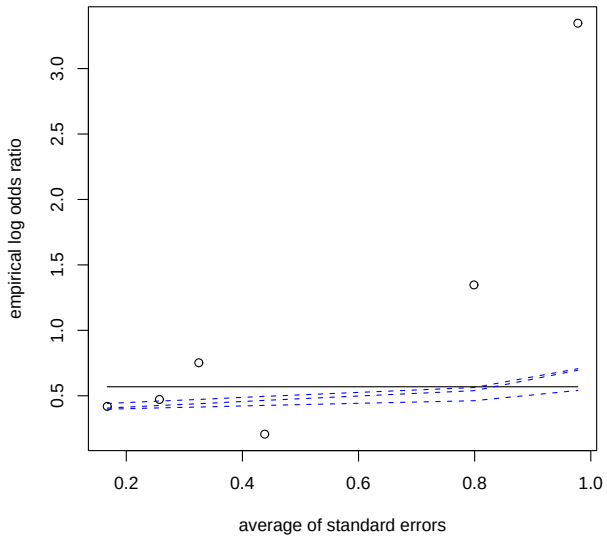


Figure 6.2: Funnel plot: $Y_{i,AC}$ against φ_i for G_1- the solid line represents the estimate without selectivity $\widehat{\mu_{AC}} = 0.5689296$; the dashed lines represent the fitted values for given (a_1, b_1, a_2, b_2) which $(a_1, b_1, a_2, b_2, \widehat{\mu_{AC}})$ are equal to (0.458,0.063,0.339,0.135, 0.54),(-0.17,0.16,-0.49,0.36,0.52) and (-0.41,0.15,-0.70,0.33,0.50).

(rows 2, 3 and 4 from Tables 6.1 and 6.2) are presented by the dashed lines in Figures 6.2 and 6.3 respectively. These curves are calculated from the equation (6.9) (mean for selected studies). As described in Section 6.2.3, note that the smaller number of studies of population model (6.1) (column 7) gives larger means as shown in columns 7. Two values of (a_1, b_1, a_2, b_2) obtained from the selection probability pairs $(0.80, 0.50)$ (row 3 in Tables 6.1 and 6.2) and $(0.70, 0.40)$ (row 4 in Tables 6.1 and 6.2) give good fits while the first one (the first dashed line, near the solid line) is unacceptable.

Similarly, the fit of funnel plot for G_4 given in Figure 6.4 is evaluated from $\widehat{\mu_{BC}} + \varrho_3 \sigma_{i,BC} \lambda(a_2 + b_2/s_{i,BC})$. The fit for G_4 in Figure 6.4 gives similar results as for G_1.

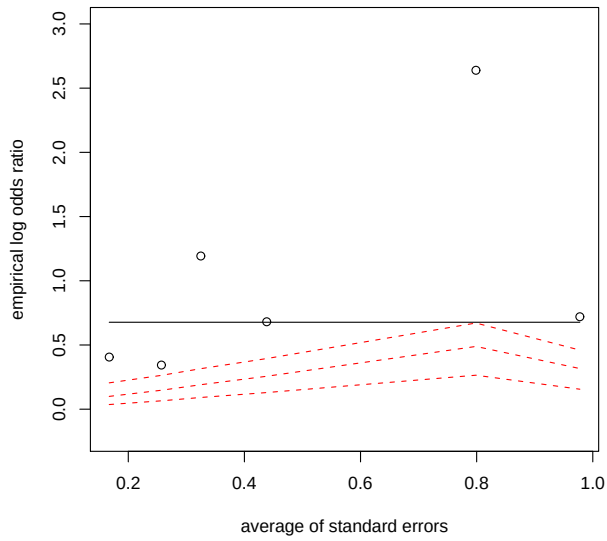


Figure 6.3: Funnel plot: $Y_{i,AC}$ against φ_i for G_1- the solid line represents the estimate without selectivity $\widehat{\mu_{BC}} = 0.6770754$; the dashed lines represent the fitted values for given (a_1, b_1, a_2, b_2), which $(a_1, b_1, a_2, b_2, \widehat{\mu_{AC}})$ are equal to (0.458,0.063,0.339,0.135, 0.58),(-0.17,0.16,-0.49,0.36,0.52) and (-0.41,0.15,-0.70,0.33,0.47).

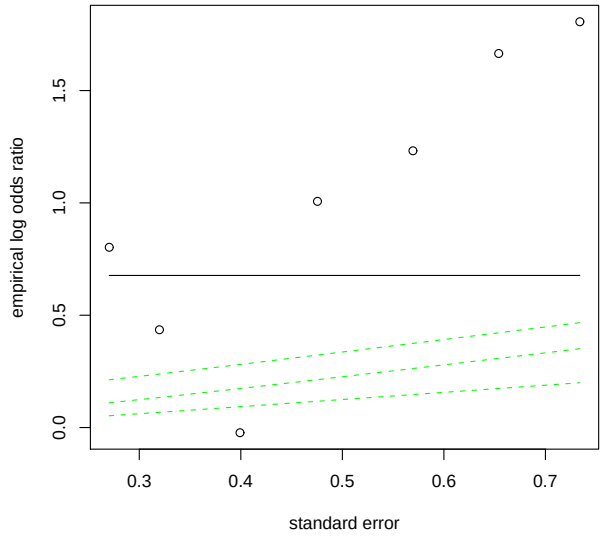


Figure 6.4: Funnel plot: $Y_{i,BC}$ against $s_{i,BC}$ for G_4- the solid line represents the estimate without selectivity $\widehat{\mu_{BC}} = 0.6770754$; the dashed lines represent the fitted values for given (a_1, b_1, a_2, b_2), which $(a_1, b_1, a_2, b_2, \widehat{\mu_{BC}})$ are equal to (0.458,0.063,0.339,0.135, 0.58),(-0.17,0.16,-0.49,0.36,0.52) and (-0.41,0.15,-0.70,0.33,0.47).

6.5.3 Sensitivity analysis (Step 3)

The idea of sensitivity analysis is to calculate all the statistical quantities for any combinations of (a_1, b_1, a_2, b_2) transformed from a grid in (6.21) for G_1 and a similar grid for G_4. Then, we can plot the estimates for example $\widehat{\mu_{AC}}$ against the p-value of the goodness-of-fit test. All the estimates with p-value less than a significant level (say 0.05) can be discarded. The estimates corresponding to model with p-value around 0.5 can be treated as the most plausible estimates. Some other quantities can also be used to find plausible estimates.

For the W1 data and all combinations of selection probability pairs presented in Tables 6.1 and 6.2, the plots of $\widehat{\mu_{AC}}$ against p-value of $H_0 : \boldsymbol{\beta} = 0$ and $\widehat{\mu_{BC}}$ against p-value of $H_0 : \boldsymbol{\beta} = 0$ are given in Figures 6.5 and 6.6 respectively. The plots indicate that the $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ can be anything less than 0.55 and 0.60 respectively. The overall $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ should come from the models with p-value greater than 0.05 and plausible overall estimates should be the ones from the models with p-value around 0.5. Therefore the plausible estimates for $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ should be around 0.47 and 0.40 respectively.

Bear in mind that we put a negative sign for all the empirical log-odds ratios in this chapter. That means that the overall log-odds ratio having reocclusion for treatment A comparing to the control group should be around -0.47 (OR = 0.625, i.e reduced the rates of reocclusion 37%). The estimate from the standard model (row 1 of Table 6.3) is -0.5689296 (OR = 0.566, reduced the rate of reocclusion by over 40%), which is overestimated.

Comparing treatment B and the control group, the overall log-odds ratio is around -0.40 (OR = 0.67, reduced the rate of reocclusion 33%), while the model without assuming selection bias gives the estimate of μ_{BC} -0.677075 (OR = 0.50, reduced the rate of reocclusion 50%), which is clearly overestimated.

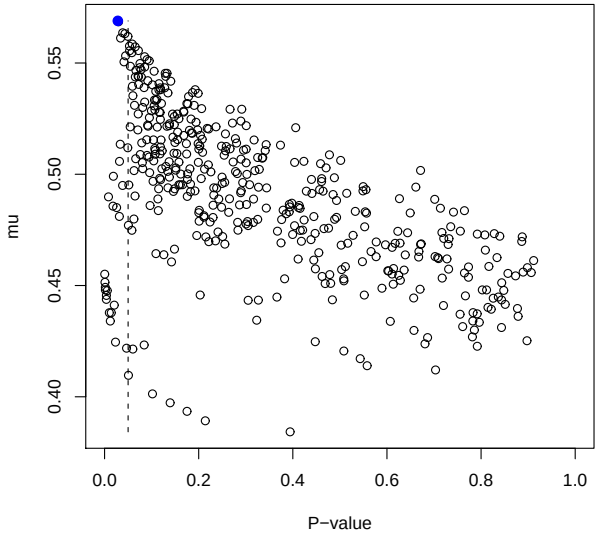


Figure 6.5: The W1 data: $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = 0$.

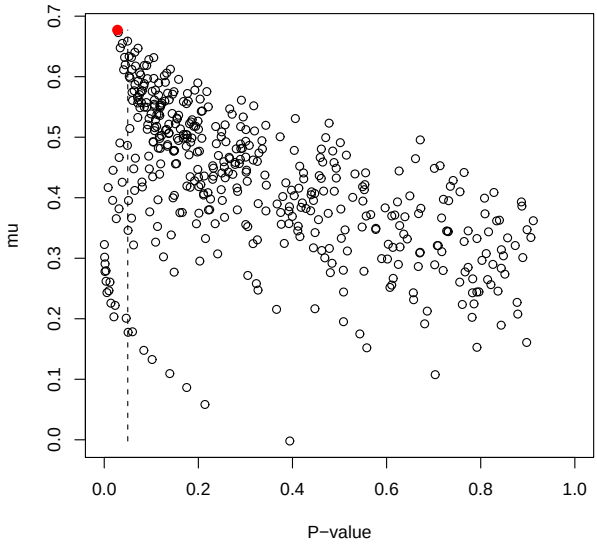


Figure 6.6: The W1 data: $\widehat{\mu_{BC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = 0$.

6.6 Simulation study

This section aims to examine a sensitivity analysis of the bias from the generated three-arm data. The steps of the sensitivity analysis with the generated data are as follows.

1. The population data

   We generate the three-arm data with 24 studies to represent the population data of treatment A versus B versus C. Note that the simulation model is from the different-trial effects and the treatment effects $\delta_{i,AC} \sim N(0.90, 0.10^2)$ and $\delta_{i,BC} \sim N(0.60, 0.10^2)$. The correlation coefficient between both treatment effects is assumed to be zero. This implies that the covariance between both treatment effects is zero. The main parameters μ_{AC} and μ_{BC} obtained from the generated data are 0.96 and 0.62 respectively.

2. Make the selection bias

   We shall select each study by the selection probability for a typical study given in (6.6). The parameters ϱ_1 and ϱ_2 are selected from the pair $(0.80, 0.80)$. We choose the values of $(P_{max}(\text{selection}), P_{min}(\text{selection}))$ as (0.90,0.10), (0.80,0.20) and (0.60,0.30) then determine the values of (a_1, b_1) for the selection probability (6.6).

   The requirement of selection for a study from the population model is that larger studies are likely to be selected than smaller studies. Let P_i be a probability of the population data being selected. The probability of selection for the ith study is

   $$P_i = p(Z_{i1} > 0 | (Y_{i,AC}, Y_{i,BC})) = \Phi\left(\frac{\mu_{2i1}}{\sigma_{2i1}}\right), \tag{6.22}$$

   where μ_{2i1} and σ^2_{2i1} are the same as (6.7) and (6.8). Let U_i be a random number generated from an uniform distribution $U(0, 1)$. The ith study from the population data in (i) will be selected if P_i is greater than U_i.

It is clear that the above steps would generate a set of studies with selection bias. We first calculated the MLEs for μ_{AC} and μ_{BC} by using the model without assuming selection bias. The results are given in Table 6.4. The p-value in the table is the one for goodness-of-fit test with $H_0 : \beta_1 = \beta_2 = 0$. The p-value 0.06705 of the model (the last row of Table 6.4) shows a good fit for funnel plots while the other p-values are significant at significance level 0.5. The funnel plots for $Y_{i,AC}$ against $s_{i,AC}$ and $Y_{i,BC}$ against $s_{i,BC}$ corresponding to the first two models in Table 6.4 are given in Figures 6.7(a)-(b) and 6.8 (a)-(b). All funnel plots show signs of selection bias, i.e. some studies may be unselected. Thus, we shall use the sensitivity for the first two models in Table 6.4.

The procedure of sensitivity analysis is as discussed in the previous section. To save space, we present only the scatter plots of μ_{AC} and μ_{BC} against their p-values of the goodness-of-fit test $H_0 : \beta_1 = \beta_2 = 0$, given in Figures 6.7(c)-(d) and 6.8 (c)-(d). The dashed line in the plots represents the true mean effect of the standard model (from the simulated data).

Table 6.4: The simulated three-arm data: summary of outputs

ϱ_1	ϱ_2	P_{max}(selection)	P_{min}(selection)	$\widehat{\mu_{AC}}$ (0.96)	$\widehat{\mu_{BC}}$ (0.62)	p-value	number of selected studies
0.8	0.8	0.90	0.10	1.34348	1.02130	0.03032	11
		0.80	0.20	1.34701	0.96602	0.04459	9
		0.60	0.30	1.39359	1.06687	0.06705	11

(i). The estimates for μ_{AC} and μ_{BC} from the models without assuming selection bias (see Table 6.4) are overestimated comparing to their true mean effects and presented in the blue and red solid circles in (c) and (d) of all figures respectively.

(ii). By using the sensitivity analysis to those models, we can conclude as follows

(a) The estimates for μ_{AC} and μ_{BC} with p-value less than 0.05 can be discarded, thus

$\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ corresponding to the first two rows of Table 6.4 can be anything less than 1.33 and 0.98 , and 1.33 and 0.93 respectively.

(b) As described in Section 6.5, the most plausible estimates should come from the model with p-value around 0.5. Therefore, the plausible estimates for μ_{AC} and μ_{BC} should be around 1.26 and 0.62, and 1.0 and 0.64 respectively. Notice these estimates are quite close to the true mean effects of μ_{AC} and μ_{BC} (0.96 and 0.62).

Based on the simulation study, sensitivity analysis approach used in this thesis can be used to adjust the over-estimates which the standard model usually give when there is selection bias.

6.7 Some theorems of mathematical consequences

In this section, we will prove the statistical theorems presented in Section 6.2.3.

Theorem 6.7.1 (The probability of selection). *Suppose that there is selection bias in G_1 and the empirical log-odds ratio model (6.1) is assumed to be population model. The selection model is defined as $Z_{i1} = a_1 + b_1/\varphi_i$ where a_1 and b_1 control the marginal probability and the φ_i is the average of the standard errors involved in the ith study. Then the probability of being selected for the ith study is*

$$p(Z_{i1} > 0|\varphi_i) = \Phi(g_{i1}), \qquad \textit{where} \qquad g_{i1} = a_1 + \frac{b_1}{\varphi_i}.$$

Proof. The selection model Z_{i1} is normally distributed with mean g_{i1} and variance 1: $Z_{i1} \sim N(g_{i1}, 1)$ where $g_{i1} = a_1 + b_1/\varphi_i$. The marginal probability of the selection model can be written as

$$p(Z_{i1} > 0|\varphi_i) = p(Z_{i1} - g_{i1} > -g_{i1}) = \Phi(g_{i1}).$$ □

Theorem 6.7.2 (The probability of selection for a typical study $(Y_{i,AC}, Y_{i,BC})$). *From the population model (6.1) and the selection model (6.3), the probability of being selected for*

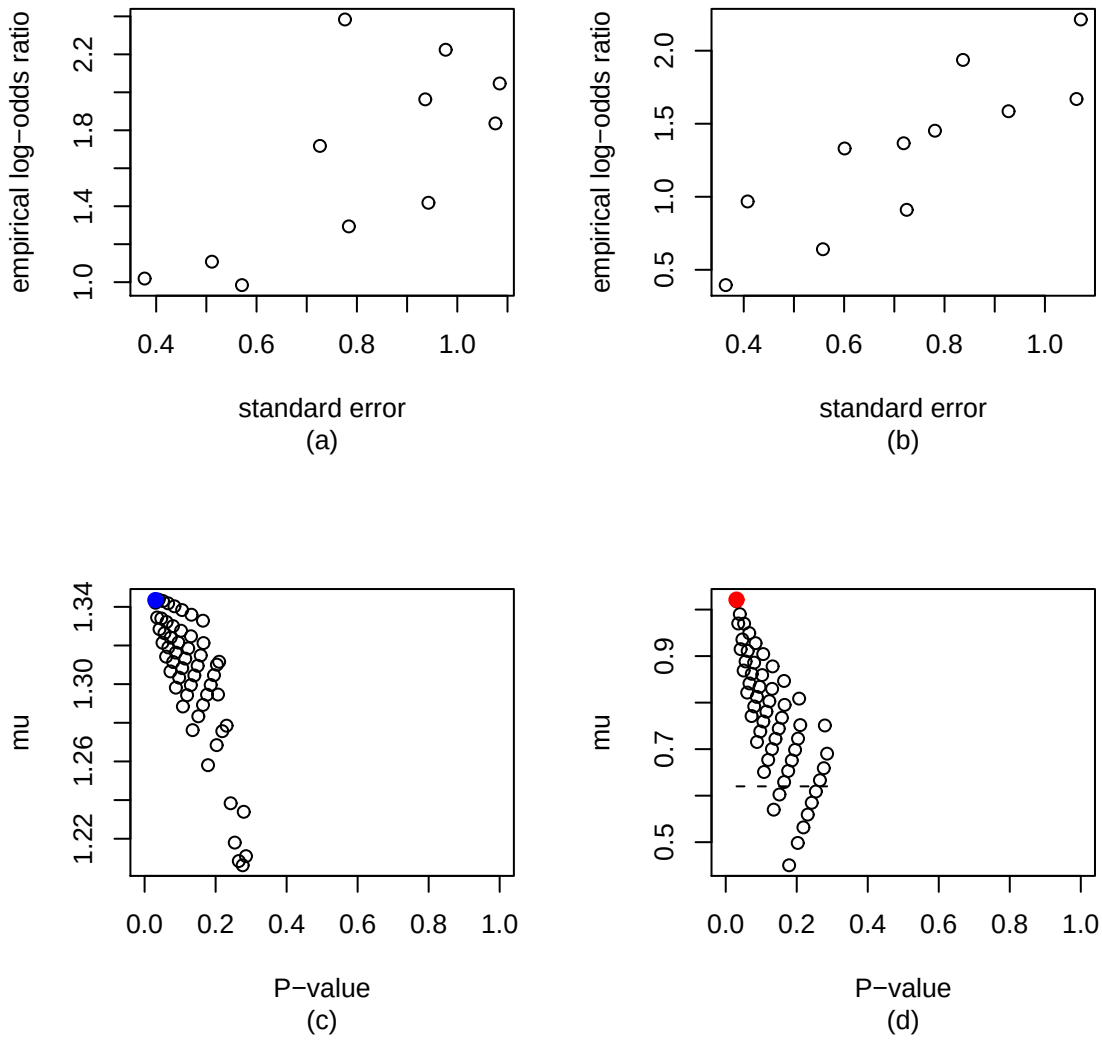


Figure 6.7: The generated data with bias where $\varrho_1 = \varrho_2 = 0.8$ corresponds to $(P_{min}(\text{selection}), P_{max}(\text{selection}) = (0.90, 0.10)$: (a) funnel plot of $Y_{i,AC}$ against $s_{i,AC}$; (b) funnel plot of $Y_{i,BC}$ against $s_{i,BC}$; (c) $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \beta_1 = \beta_2 = 0$; (d)$\widehat{\mu_{BC}}$ against the p-value of $H_0 : \beta_1 = \beta_2 = 0$

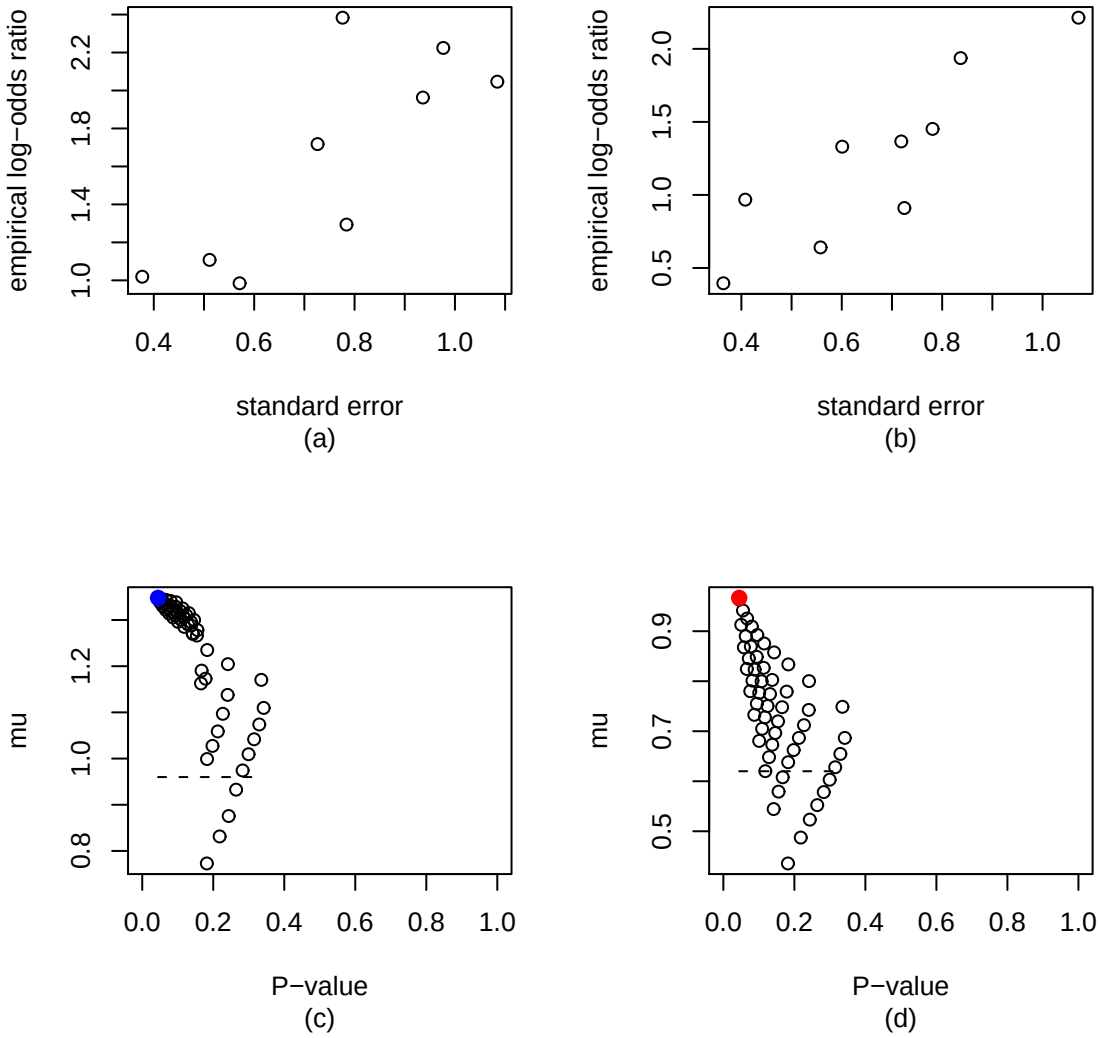


Figure 6.8: The generated data with bias where $\varrho_1 = \varrho_2 = 0.8$ corresponds to $(P_{min}(\text{selection}), P_{max}(\text{selection}) = (0.80, 0.20)$: (a) funnel plot of $Y_{i,AC}$ against $s_{i,AC}$; (b) funnel plot of $Y_{i,BC}$ against $s_{i,BC}$; (c) $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \beta_1 = \beta_2 = 0$; (d)$\widehat{\mu_{BC}}$ against the p-value of $H_0 : \beta_1 = \beta_2 = 0$

a typical study $(Y_{i,AC}, Y_{i,BC})$ *is*

$$p(Z_{i1} > 0|(Y_{i,AC}, Y_{i,BC})) = \Phi\left(\frac{\mu_{2i1}}{\sigma_{2i1}}\right),$$

where $\mu_{2i1} = E(Z_{i1}|(Y_{i,AC}, Y_{i,BC}))$ *and* $\sigma^2_{2i1} = Var(Z_{i1}|(Y_{i,AC}, Y_{i,BC}))$ *are given in (6.7) and (6.8) respectively.*

Proof. The selection model Z_{i1} is normally distributed as $N(g_{i1}, 1)$ where $g_{i1} = a_1 + b_1/\varphi_i$. The outcome $(Y_{i,AC}, Y_{i,BC})$ is normally distributed as presented in (3.38) of Chapter 3. The variance-covariance matrix of (3.38) is

$$\begin{pmatrix} v_{1i} & v_{12i} \\ v_{21i} & v_{2i} \end{pmatrix} = \begin{pmatrix} \tau^2_{AC} + \sigma^2_{i,AC} & \rho\tau_{AC}\tau_{BC} + \sigma^2_{iC} \\ \rho\tau_{AC}\tau_{BC} + \sigma^2_{iC} & \tau^2_{BC} + \sigma^2_{i,BC} \end{pmatrix}. \tag{6.23}$$

The variance-covariance matrix between the selection model Z_{i1} and the distribution $(Y_{i,AC}, Y_{i,BC})$ is

$$Cov(Z_{i1}, (Y_{i,AC}, Y_{i,BC})) = \begin{pmatrix} \mathbf{w}_{11} & \mathbf{w}_{12} \\ \mathbf{w}_{21} & \mathbf{w}_{22} \end{pmatrix},$$

where $\mathbf{w}_{11} = (1)$, $\mathbf{w}_{12} = \left(\varrho_1\sqrt{v_{1i}}, \varrho_2\sqrt{v_{2i}}\right)$, $\mathbf{w}_{21} = (\varrho_1\sqrt{v_{1i}}, \varrho_2\sqrt{v_{2i}})^t$ and $\mathbf{w}_{22}$ is given in (6.23). From the property of conditional distribution, the conditional distribution of the selection model (6.3) given $(Y_{i,AC}, Y_{i,BC})$ is a multivariate normal distribution with mean $E(Z_{i1}|(Y_{i,AC}, Y_{i,BC})$ and variance $Var(Z_{i1}|(Y_{i,AC}, Y_{i,BC}))$. The $E(Z_{i1}|(Y_{i,AC}, Y_{i,BC})$ can be calculated as

$$E(Z_{i1}|(Y_{i,AC}, Y_{i,BC})) \;=\; g_{i1} + \mathbf{w}_{12}\mathbf{w}^{-1}_{22}\begin{pmatrix} Y_{i,AC} - \mu_{AC} \\ Y_{i,BC} - \mu_{BC} \end{pmatrix} = \mu_{2i1}. \tag{6.24}$$

Likewise, $Var(Z_{i1}|(Y_{i,AC}, Y_{i,BC}))$ is formulated as

$$Var(Z_{i1}|(Y_{i,AC}, Y_{i,BC})) \;=\; \mathbf{w}_{11} - \mathbf{w}_{12}\mathbf{w}^{-1}_{22}\mathbf{w}_{21} = \sigma^2_{2i1}. \tag{6.25}$$

Hence, the probability of being selected for a typical study $(Y_{i,AC}, Y_{i,BC})$ may be written as

$$p(Z_{i1} > 0, \varphi_i | (Y_{i,AC}, Y_{i,BC})) = p\left(\frac{Z_{i1}|(Y_{i,AC},Y_{i,BC}) - E(Z_{i1}|(Y_{i,AC},Y_{i,BC})}{\sqrt{Var(Z_{i1}|(Y_{i,AC},Y_{i,BC}))}} > -\frac{E(Z_{i1}|(Y_{i,AC},Y_{i,BC}))}{\sqrt{Var(Z_{i1}|(Y_{i,AC},Y_{i,BC}))}} \right),$$

$$= \Phi\left(\frac{E(Z_{i1}|(Y_{i,AC},Y_{i,BC}))}{\sqrt{Var(Z_{i1}|(Y_{i,AC},Y_{i,BC}))}} \right),$$

where $E(Z_{i1}|(Y_{i,AC}, Y_{i,BC}))$ and $Var(Z_{i1}|(Y_{i,AC}, Y_{i,BC}))$ are given in (6.24) and (6.25) respectively. □

Theorem 6.7.3 (The means for selected studies). *From the population model (6.1) and the selection model (6.3), the means for selected studies are*

$$E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i \right) = \begin{pmatrix} \mu_{AC} \\ \mu_{BC} \end{pmatrix} + \begin{pmatrix} \varrho_1 \sigma_{i,AC} \\ \varrho_2 \sigma_{i,BC} \end{pmatrix} \lambda\left(a_1 + \frac{b_1}{\varphi_i} \right).$$

Proof. The expected value of the conditional distribution $(Y_{i,AC}, Y_{i,BC})$ given Z_{i1} is estimated by

$$E(Z_{i1}|(Y_{i,AC}, Y_{i,BC})) = \begin{pmatrix} \mu_{AC} \\ \mu_{BC} \end{pmatrix} + \begin{pmatrix} \varrho_1 \sigma_{iAC} \\ \varrho_2 \sigma_{iBC} \end{pmatrix} (Z_{i1} - g_{i1}). \tag{6.26}$$

The expected value of the conditional distribution $(Y_{i,AC}, Y_{i,BC})$ given $Z_{i1} > 0$ is

$$
\begin{aligned}
& E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i\right) \\
&= \int_{-\infty}^{\infty} p(Y_{i,AC}, Y_{i,BC}) p((Y_{i,AC}, Y_{i,BC}) | Z_{i1} > 0, \varphi_i) d(Y_{i,AC}, Y_{i,BC}), \\
&= \frac{\int_0^{\infty} \int_{-\infty}^{\infty} p(Y_{i,AC}, Y_{i,BC}) p((Y_{i,AC}, Y_{i,BC}), Z_{i1}) d(Y_{i,AC}, Y_{i,BC}) dZ_{i1}}{\int_0^{\infty} p(Z_{i1}) dZ_{i1}}, \\
&= \frac{\int_0^{\infty} \int_{-\infty}^{\infty} p(Y_{i,AC}, Y_{i,BC}) p((Y_{i,AC}, Y_{i,BC}) | Z_{i1}) p(Z_{i1}) d(Y_{i,AC}, Y_{i,BC}) dZ_{i1}}{\int_0^{\infty} p(Z_{i1}) dZ_{i1}}, \\
&= \frac{\int_0^{\infty} p(Z_{i1}) \int_{-\infty}^{\infty} p(Y_{i,AC}, Y_{i,BC}) p((Y_{i,AC}, Y_{i,BC}) | Z_{i1}) d(Y_{i,AC}, Y_{i,BC}) dZ_{i1}}{\int_0^{\infty} p(Z_{i1}) dZ_{i1}}, \\
&= \frac{\int_0^{\infty} p(Z_{i1}) E((Y_{i,AC}, Y_{i,BC}) | Z_{i1}) dZ_{i1}}{\int_0^{\infty} p(Z_{i1}) dZ_{i1}}, \qquad (6.27) \\
&= \begin{pmatrix} \mu_{AC} \\ \mu_{BC} \end{pmatrix} + \begin{pmatrix} \varrho_1 \sigma_{i,AC} \\ \varrho_2 \sigma_{i,BC} \end{pmatrix} \lambda\left(a_1 + \frac{b_1}{\varphi_i}\right).
\end{aligned}
$$

Inserting equation (6.26) in equation (6.27), results are obtained in the above formula. □

Theorem 6.7.4 (The variance of selected studies). *From the population model (6.1) and the selection model (6.3), the variance of selected studies is*

$$
Var\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i\right) = \begin{pmatrix} \sigma_{i,AC}^2 (1 + d_{i1}^2 \varrho_1^2) & \sigma_{iC}^2 \\ \sigma_{iC}^2 & \sigma_{i,BC}^2 (1 + d_{i1}^2 \varrho_2^2) \end{pmatrix},
$$

where $d_{i1}^2 = \lambda(g_{i1})(g_{i1} + \lambda(g_{i1}))$. *and* $g_{i1} = a_1 + b_1/\varphi_i$.

Proof. The variance from the above equation can be written as

$$\begin{pmatrix} Var(Y_{i,AC}|Z_{i1}>0,\varphi_i) & Cov(Y_{i,AC}|Z_{i1}>0,Y_{i,BC}|Z_{i1}>0) \\ Cov(Y_{i,BC}|Z_{i1}>0,Y_{i,AC}|Z_{i1}>0) & Var(Y_{i,BC}|Z_{i1}>0,\varphi_i) \end{pmatrix}. \tag{6.28}$$

We shall prove the entries on the diagonal first.

1. The entry on the diagonal $Var(Y_{i,AC}|Z_{i1}>0,\varphi_i)$ can be written in the form

$$Var(Y_{i,AC}|Z_{i1}>0,\varphi_i) = E(Y_{i,AC}^2|Z_{i1}>0,\varphi_i) - \left(E(Y_{i,AC}|Z_{i1}>0,\varphi_{i1})\right)^2. \tag{6.29}$$

The last term of (6.29) is calculated as

$$\begin{aligned} \left(E\left(Y_{i,AC}|Z_{i1}>0,\varphi_i\right)\right)^2 &= \left(\mu_{AC}+\varrho_1\sigma_{i,AC}\lambda(g_{i1})\right)^2, \\ &= \mu_{AC}^2 + 2\mu_{AC}\varrho\sigma_{i,AC}\lambda(g_{i1}) + \varrho_1^2\sigma_{i,AC}^2(\lambda(g_{i1}))^2. \end{aligned} \tag{6.30}$$

The first term of equation (6.29) can calculate

$$\begin{aligned} E\left(Y_{i,AC}^2|Z_{i1},\varphi_i\right) &= E\left(Y_{i,AC}|Z_{i1},\varphi_i\right) + \left(E(Y_{i,AC}|Z_{i1},\varphi_i)\right)^2, \\ &= \sigma_{i,AC}^2(1-\varrho_1^2) + (\mu_{AC}+\varrho_1\sigma_{i,AC}(Z_{i1}-\mu_{BC}))^2. \end{aligned} \tag{6.31}$$

The first term of (6.29) is an integral

$$E(Y_{i,AC}^2|Z_{i1}>0,\varphi_i)$$

$$=\int_{-\infty}^{\infty} Y_{i,AC}^2 p(Y_{i,AC}|Z_{i1}>0,\varphi_i)dY_{i,AC},$$

$$=\frac{\int_{-\infty}^{\infty} Y_{i,AC}^2 \int_0^{\infty} p(Y_{i,AC},Z_{i1})dZ_{i1}dY_{i,AC}}{\int_0^{\infty} p(Z_{i1},\varphi_i)dZ_{i1}},$$

$$=\frac{\int_{-\infty}^{\infty} Y_{i,AC}^2 \int_0^{\infty} p(Y_{i,AC}|Z_{i1})p(Z_{i1})dZ_{i1}dY_{i,AC}}{\int_0^{\infty} p(Z_{i1},\varphi_i)dZ_{i1}},$$

$$=\frac{\int_0^{\infty} p(Z_{i1}) \int_{-\infty}^{\infty} Y_{i,AC}^2 p(Y_{i,AC}|Z_{i1})dY_{i,AC}dZ_{i1}}{\int_0^{\infty} p(Z_{i1},\varphi_i)dZ_{i1}},$$

$$=\frac{\int_0^{\infty} p(Z_{i1})E(Y_{i,AC}^2|Z_{i1})dZ_{i1}}{\int_0^{\infty} p(Z_{i1},\varphi_i)dZ_{i1}}, \tag{6.32}$$

$$=\frac{\int_0^{\infty} p(Z_{i1})(\sigma_{i,AC}^2(1-\varrho_1^2)+(\mu_{AC}+\varrho_1\sigma_{i,AC}(Z_{i1}-\mu_{BC}))^2)dZ_{i1}}{\int_0^{\infty} p(Z_{i1},\varphi_i)dZ_{i1}},$$

$$=\sigma_{i,AC}^2-\varrho_1^2+\mu_{AC}^2+2\mu_{AC}\varrho_1\sigma_{i,AC}\lambda(\mu_{BC})+\varrho_1^2\sigma_{i,AC}^2(\mu_{BC}\lambda(\mu_{BC})+1). \tag{6.33}$$

Using the results from (6.30) in (6.32), the equation (6.33) is obtained. Substituting equations (6.30) and (6.33) into the first term and the second term of (6.29) respectively

gives

$$Var(Y_{i,AC}|Z_{i1} > 0, \varphi_i) = \sigma^2_{i,AC} + \varrho_1^2\sigma^2_{i,AC}\mu_{BC}\lambda(\mu_{BC}) + \varrho_1^2\sigma^2_{i,AC}(\lambda(\mu_{BC}))^2,$$
$$= \sigma^2_{i,AC}(1 + d^2_{i1}\varrho_1^2),$$

where $d^2_{i1} = \lambda(g_{i1})(g_{i1} + \lambda(g_{i1}))$. In similar way, we have $Var(Y_{i,BC}|Z_{i1} > 0, \varphi_i) = \sigma^2_{i,BC}(1 + d^2_{i1}\varrho_2^2)$.

2. Considering the covariance of $Var\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} |Z_{i1} > 0, \varphi_{i1}\right)$, the G_1 has the treatment C as the baseline treatment thus

$$Cov(Y_{i,AC}|Z_{i1} > 0, Y_{i,BC}|Z_{i1} > 0) = Cov(Y_{i,BC}|Z_{i1} > 0, Y_{i,AC}|Z_{i1} > 0) = \sigma^2_{iC}. \quad \square$$

Theorem 6.7.5 (The difference of means). *From the population model (6.1) and the selection model (6.3), we assume that there is another pair (a_1^*, b_1^*) for the selection model which is better than (a_1, b_1). The difference of the means by the selection model (6.3) with the two pairs (a_1, b_1) and (a_1^*, b_1^*) is*

$$E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} |Z_{i1} > 0, \varphi_i, a_1^*, b_1^*\right) - E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} |Z_{i1} > 0, \varphi_i, a_1, b_1\right)$$
$$\approx \quad c^* + \begin{pmatrix} \varrho_1 \\ \varrho_2 \end{pmatrix} (\lambda(a_1^*) - \lambda(a_1)) \begin{pmatrix} s_{i,AC} \\ s_{i,BC} \end{pmatrix}.$$

where c^* is constant where (a_1^*, b_1^*) is close to (a_1, b_1).

Proof. The above equation can be written as

$$
\begin{aligned}
& E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i, a_1^*, b_1^*\right) - E\left(\begin{pmatrix} Y_{i,AC} \\ Y_{i,BC} \end{pmatrix} | Z_{i1} > 0, \varphi_i, a_1, b_1\right) \\
= & \begin{pmatrix} \mu_{AC} \\ \mu_{BC} \end{pmatrix} + \begin{pmatrix} \varrho_1\sigma_{iAC} \\ \varrho_2\sigma_{iBC} \end{pmatrix} \lambda(a_1^* + b_1^*/\varphi_i) - \begin{pmatrix} \mu_{AC} \\ \mu_{BC} \end{pmatrix} - \begin{pmatrix} \varrho_1\sigma_{iAC} \\ \varrho_2\sigma_{iBC} \end{pmatrix} \lambda(a_1 + b_1/\varphi_i), \\
\approx & c^* + \begin{pmatrix} \varrho_1 s_{iAC} \\ \varrho_2 s_{iBC} \end{pmatrix} \lambda(a_1^* + b_1^*/\varphi_i) - \lambda(a_1 + b_1/\varphi_i). \qquad (6.34)
\end{aligned}
$$

By using Theorem 6.7.4, we obtain $\sigma_{i,AC}^2 = s_{i,AC}^2/(1 + d_{i1}^2\varrho_1^2)$ and $\sigma_{i,BC}^2 = s_{i,BC}^2/(1 + d_{i1}^2\varrho_2^2)$. We substitute $\sigma_{i,AC}^2$ and $\sigma_{i,BC}^2$ by $s_{i,AC}^2$ and $s_{i,BC}^2$ in the above equation. From Taylor series

$$f(x + \Delta) = f(x) + \Delta x f^{'}(x) + \frac{\Delta x^2}{2} f^{''}(x) + \ldots$$

By using the Teylor series, the functions $\lambda(a_1^* + b_1^*/\varphi_i)$ and $\lambda(a_1 + b_1/\varphi_i)$ in (6.34) are given by

$$
\begin{aligned}
\lambda(a_1^* + b_1^*/\varphi_i) &= \lambda(a_1^*) + \frac{b^*}{\varphi_i} + \left(\frac{b^*}{\varphi_i}\right)^2 \frac{\lambda^{''}(a_1^*)}{2} + \ldots \\
\lambda(a_1 + b_1/\varphi_i) &= \lambda(a_1) + \frac{b}{\varphi_i} + \left(\frac{b}{\varphi_i}\right)^2 \frac{\lambda^{''}(a_1)}{2} + \ldots .
\end{aligned}
$$

Hence, the equation (6.34) is approximated by

$$c^* + \begin{pmatrix} \varrho_1 \\ \varrho_2 \end{pmatrix} \lambda(a_1^*) - \lambda(a_1) \begin{pmatrix} s_{iAC} \\ s_{iBC} \end{pmatrix}. \qquad \square$$

6.8 Discussion

An important role of meta-analysis is to combine information from different studies to summarize an overall estimate of a treatment effect. Studies with a greater effect may be more likely to be selected or published than studies with a less statistically significant effect. Chapter 3 presented the empirical log-odds ratio model for the W1 data without considering the problem of selection in the meta-analysis. In this chapter, we employed the sensitivity analysis using the selection model to examine the selection bias and corrected the results under the controlled assumptions for the model. We regard the selection model as a tool of sensitivity analysis.

The funnel plot was used to test a selection bias in this thesis. For studies with the binary outcomes, the standard error is the best measure of study size, while risk ratios or odds ratios should be used for the measure of treatment effect. We plotted the empirical log-odds ratios against their standard errors for the funnel plot. From the funnel plot 6.1(d), the studies missed at the bottom right-hand corner can be treated as non-ignorable missing data in meta-analysis, see e.g Little and Rubin (2002). Note that a funnel plot is a simple graphical tool for the investigation of selection bias in meta-analysis. It cannot be claimed that visually interpreted asymmetry of a plot always reflects selection bias. For example, studies of lower quality may exaggerate the estimate of the treatment effects. Selection bias is only one of a number of possible causes of funnel plot asymmetry. Other sources of asymmetry in funnel plots may be true heterogeneity, data irregularities, artefact and chance (Egger et al., 1987). These may give the low power of tests for the funnel plot asymmetry.

The basic idea of the selection model is that the probability of selection depends on both the empirical log-odds ratio and its standard error. Also the model is made under the requirement that larger studies are more likely to be selected than smaller studies. When the number of studies is small, two problems arise for the selection model. There may be

numerical problems in obtaining stable estimates of the parameters. More importantly, the standard errors of estimates will be large, perhaps so large as to make any specific inferences impossible or meaningless. In this case, we need to use an exact logistic regression model as discussed in Chapter 4. The related selection model will be discussed in the next chapter.

In addition, the pairs (a_1, b_1) and (a_2, b_2) for the selection models Z_{i1} and Z_{i2} cannot be estimated from the log-likelihood function in the usual way, because we do not know that how many unselected studies are there in the population of treatment comparisons A vs B vs C and B and C. Thus, we calculate those pairs from the given probabilities obtaining from the largest studies and smallest studies in meta-analysis. These probabilities represent the different amounts of selection bias for the models assuming the selection bias. The funnel plot examines whether or not there is selection bias in meta-analysis but cannot tell that how many of unselected studies are. Therefore, the sensitivity analysis is needed.

As discussed in Chapter 3, the assumption of variance homogeneity applies to all the treatment effects, and the correlation coefficients between treatment effects are 1/2. If both direct and indirect comparisons are in meta analysis and the number of indirect comparison studies is sufficiently large then the correlation coefficient between those treatment effects is estimable (Chootrakool and Shi, 2008). However it would make the model more complicated in the sensitivity analysis.

If more than three treatments are compared in the meta-analysis, the sensitivity analysis can be applied in the same way but each group of treatment comparisons should have enough information (studies) if we would like to add the selection model in those studies (for bias-suspected model).

Chapter 7

Sensitivity analysis to logistic regression model

7.1 Introduction

We have described how to inspect selection bias by a funnel plot and how to address the selection bias by using of a sensitivity analysis for normal approximation model in Chapter 6. However, if the sample size for each study is very small, an exact binomial model should be used. The multi-arm trials model based on the binomial approach for the binary data was presented in Chapter 4. In this chapter, we extend the sensitivity analysis to the exact logistic regression model when there is selection bias, i.e. studies with statistically significant results might have been selected more predominantly. Regarding to discussion in Chapter 5, the conditional likelihood estimates for the logistic regression model usually gives a better result, therefore a conditional method will be applied for the logistic regression model in this chapter. We will use a simulated data to perform the sensitivity analysis in this chapter.

The simulated data is given in the first section. The rest of this chapter is arranged as follows. The multi-arm trials model for the exact conditional distribution is given in Section 7.3. We

present an inspection of selection bias for the data in Section 7.4. Section 7.5 performs the conditional probability with selection using some formulae from Chapter 6. The log-likelihood function of the model with selection is produced in Section 7.6. Section 7.7 and 7.8 illustrate goodness of fit and sensitivity analysis respectively. Finally, the conclusion and some comments are given in Section 7.9.

7.2 Simulated data

In this chapter, we will employ the following simulated data to illustrate how sensitivity analysis is used to address the problem of selection bias in meta-analysis with the logistic regression model using conditional method. Essentially, the following steps of generating the data and making the selection bias are similar to the steps in Section 6.6 from the previous chapter. Those steps are

1. The population data

   To generate the population data of treatment A versus B versus C, we generate three-arm data for 14 studies. We assume the different-trial effects, and the treatment effects $\delta_{i,AC} \sim N(0.40, 0.10^2)$ and $\delta_{i,BC} \sim N(0.60, 0.10^2)$. Similar to Section 6.6, the covariance between both treatment effects is assumed to be zero.

2. Make the selection bias

   The parameters ϱ_1 and ϱ_2, and the selection probabilities $(P_{max}(\text{selection}), P_{min}(\text{selection}))$ are 0.5,0.5 and (0.90,0.30) respectively. We use these parameters in the selection model Z_{i1} to determine the values of (a_1, b_1). Following the step 2 of Section 6.6, we will obtain the selected studies.

A group of selected studies obtaining from the above steps is supposed to be biased and the number of studies in the meta-analysis is now 9. From here, nine selected studies are used in our meta-analysis and the treatment C is the control group. We shall present the

exact conditional distribution of logistic regression model for this meta-analysis in the next section.

7.3 Multi-arm trials with the conditional probability

According to the selected studies of three-arm comparisons in the previous section, the r_{iC}, r_{iA} and r_{iB} are binomially distributed as $Bin(\pi_{iA}, n_{iA})$, $Bin(\pi_{iB}, n_{iB})$ and $Bin(\pi_{iC}, n_{iC})$ respectively for $i = 1, \ldots, 9$. If n_{iA}, n_{iB} and n_{iC} are large and r_{iA}, r_{iB} or r_{iC} are not equal to n_{iA}, n_{iB} or n_{iC} or zero. From the discussion in Section 3.3 of Chapter 3, we can define normal approximation model (see discussion in Chapter 3). For example, the empirical log-odds ratios between (r_{iA}, n_{iA}) and (r_{iC}, n_{iC}), and (r_{iB}, n_{iB}) and (r_{iC}, n_{iC}) are respectively

$$Y_{i,AC} = \log\left(\frac{r_{iA}+0.5}{n_{iA}-r_{iA}+0.5}\frac{n_{iC}-r_{iC}+0.5}{r_{iC}+0.5}\right), \tag{7.1}$$

$$Y_{i,BC} = \log\left(\frac{r_{iB}+0.5}{n_{iB}-r_{iB}+0.5}\frac{n_{iC}-r_{iC}+0.5}{r_{iC}+0.5}\right). \tag{7.2}$$

The logistic regression models for our meta-analysis can be defined as

$$\log\left(\frac{\pi_{iC}}{1-\pi_{iC}}\right) = \alpha_i, \tag{7.3}$$

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \alpha_i + \delta_{i,Cj}, \qquad j \in J_{(i)}, \tag{7.4}$$

where $J_{(i)} = \{A, B\}$. We allow the heterogeneity in the model. Both treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ are thus assumed to be random. We assume that there is no association between two treatment effects then the covariance between them is zero. Let $\mathbf{r}_i$ be the vector (r_{iA}, r_{iB}) and the function C_i represent $r_{iA}+r_{iB}+r_{iC} = c_i$. By using the conditional method as illustrated in Chapter 4, the conditional probability $\mathbf{r}_i$ given C_i for our meta-analysis is

given by

$$f(\mathbf{r}_i|\delta_{i,AC},\delta_{i,BC}) = f(\mathbf{r}_i|r_{iA}+r_{iB}+r_{iC}=c_i,\delta_{i,AC},\delta_{i,BC}),$$

$$= \frac{\binom{n_{iA}}{r_{iA}}\binom{n_{iB}}{r_{iB}}\binom{n_{iC}}{r_{iC}} e^{(\delta_{i,AC}r_{iA}+\delta_{i,BC}r_{iB})}}{\sum_{\mathbf{u}_i}\binom{n_{iC}}{c_i-u_{i1}-u_{i2}}\binom{n_{iA}}{r_{iA}}\binom{n_{iB}}{r_{iB}} e^{(\delta_{i,AC}u_{i1}+\delta_{i,BC}u_{i2})}}, \quad (7.5)$$

where $\mathbf{u}_i$ is the vector (u_{i1},u_{i2}) and is in the boundary of

$$\max(0,c_i-n_{iC}) \le u_{i1} \le \min(c_i,n_{iA}) \text{ and } \max(0,c_i-n_{iB}) \le u_{i2} \le \min(c_i,n_{iB}). \quad (7.6)$$

We use the homogeneity of variance for the model. Thus the heterogeneity parameters for the treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ are the same: τ^2_{AC}= $\tau^2_{BC}=\tau^2$. As described in Chapter 4, we integrate the conditional probability function $f(\mathbf{r}_i|\delta_{i,AC},\delta_{i,BC})$ with respect to $\delta_{i,AC}$ and $\delta_{i,BC}$ respectively. The probability $f(\mathbf{r}_i)$ now involves two integrals and is given by

$$f(\mathbf{r}_i)=\int\int f(\mathbf{r}_i|\delta_{i,AC},\delta_{i,BC})\phi(\delta_{i,AC};\mu_{AC},\tau)\phi(\delta_{i,BC};\mu_{BC},\tau)d\delta_{i,AC}d\delta_{i,BC},$$

where $\phi(\delta_{i,AC};\mu_{AC},\tau)$ and $\phi(\delta_{i,BC};\mu_{BC},\tau)$ are the probability density functions of normal distribution for $\delta_{i,AC}$ and $\delta_{i,BC}$ respectively. By applying Gaussian-Hermite approximation, the above probability is approximated as

$$f(\mathbf{r}_i) \approx \pi^{-1}\sum_{n_1=1}^{l_1} w^{(1)}_{n_1}\sum_{n_2=1}^{l_2} w^{(2)}_{n_2} f(\mathbf{r}_i|\delta_{i,AC},\delta_{i,BC}), \quad (7.7)$$

where $f(\mathbf{r}_i|\delta_{i,AC},\delta_{i,BC})$ is given in (7.5) and the sampling nodes are at $\delta_{i,AC}=\mu_{AC}+\sqrt{2}\tau d_{n_1}$ and $\delta_{i,BC}=\mu_{BC}+\sqrt{2}\tau d_{n_2}$ for $n_1=1,\ldots,l_1$ and $n_2=1,\ldots,l_2$.

7.4 Detecting the selection bias

As illustrated in Section 6.1 of Chapter 6, the funnel plot was used to detect selection bias for the W1 data. The plot of the empirical log-odds ratios against their standard errors from each pairwise-comparison is considered for those groups (G_1, G_3 and G_4). Since this chapter aims to use the sensitivity analysis for the conditional probability model, the values for the sample size axis (standard errors) and means of the model cannot be calculated in the usual way as used in Chapter 6.

To detect selection bias in this chapter, we will plot the empirical log-odds ratios on the vertical axis and the estimated conditional standard errors on the horizontal axis, and use the conditional mean instead of the conventional mean. As before, we consider the funnel plot in each pairwise-comparison of meta-analysis, e.g by considering our three-arm simulated data, the funnel plots are for treatment A versus C and B versus C . Here we need to estimate the *conditional variance* and the *conditional mean* for the funnel plot. In probability theory, the conditional variance is the variance of a conditional probability distribution. While the conditional mean (also known as *conditional expected value* or *conditional expectation*) is the expected value of a real random variable with respect to a conditional probability distribution.

7.4.1 Conditional variance

To calculate the conditional variance, let $\nu^2_{i,AC}$ and $\nu^2_{i,BC}$ be the conditional variances of $Y_{i,AC}$ given c_i, and $Y_{i,BC}$ given c_i respectively corresponding to the ith study, may be written as

$$\nu^2_{i,AC} = Var(Y_{i,AC}|c_i) \text{ and } \nu^2_{i,BC} = Var(Y_{i,BC}|c_i),$$

Note that $Y_{i,AC}$ and $Y_{i,BC}$ are empirical log-odds ratios for treatments A versus C and B versus C and defined in (7.1) and (7.2) respectively. The above conditional variances can be

estimated respectively by

$$\nu^2_{i,AC} = E(Y^2_{i,AC}|c_i) - (E(Y_{i,AC}|c_i))^2 \text{ and } \nu^2_{i,BC} = E(Y^2_{i,BC}|c_i) - (E(Y_{i,BC}|c_i))^2, \quad (7.8)$$

where E represents the expectation operator.

7.4.2 Conditional mean

From (7.8), the $E(Y_{i,AC}|c_i)$ and $E(Y_{i,BC}|c_i)$ are the conditional means of $Y_{i,AC}$ given c_i and $Y_{i,AC}$ given c_i, respectively. They can be calculated as

$$E(Y_{i,AC}|c_i) = \sum_{r_{iA}} (Y_{i,AC}.f(r_{iA}|\delta_{i,AC})) \text{ and } E(Y_{i,BC}|c_i) = \sum_{r_{iB}} (Y_{i,BC}.f(r_{iB}|\delta_{i,BC})). \quad (7.9)$$

The r_{iA} and r_{iB} are treated as discrete random variables and play the important role for $E(Y_{i,AC}|c_i)$ and $E(Y_{i,BC}|c_i)$ respectivley. The conditional probability functions $f(r_{iA}|\delta_{i,AC})$ and $f(r_{iB}|\delta_{i,BC})$ can be obtained from (7.5) and estimated in the same way as (7.7). For example, by using (7.5) , $f(r_{iA}|\delta_{i,AC})$ is given by

$$f(r_{iA}|\delta_{i,AC}) = \frac{\binom{n_{iA}}{r_{iA}}\binom{n_{iC}}{r_{iC}} e^{\delta_{i,AC} r_{iA}}}{\sum_{u_{i1}} \binom{n_{iC}}{c_i - u_{i1}}\binom{n_{iA}}{r_{iA}} e^{\delta_{i,AC} u_{i1}}}, \quad (7.10)$$

where u_{i1} is given in (7.6). The above equation is approximated by

$$f(r_{iA}) \approx \pi^{-1/2} \sum_{n=1}^{l} w_n f(r_{iA}|\delta_{i,AC}), \quad (7.11)$$

where the sampling nodes are at $\delta_{i,AC} = \mu_{AC} + \sqrt{2}\tau d_n$ for $n = 1, \ldots, l$. Notice that the values of conditional means $E(Y_{i,AC}|c_i)$ and $E(Y_{i,BC}|c_i)$ depend on the ith study and are

conditioned on the function c_i; this will give the rough function of their funnel plots. From (7.8), the conditional means of $Y^2_{i,AC}$ given c_i and $Y^2_{i,BC}$ given c_i can be evaluated from

$$E(Y^2_{i,AC}|c_i) = \sum_{r_{iA}} (Y^2_{i,AC}.f(r_{iA}|\delta_{i,AC})) \text{ and } E(Y^2_{i,BC}|c_i) = \sum_{r_{iB}} (Y^2_{i,BC}.f(r_{iB}|\delta_{i,BC})). \quad (7.12)$$

7.4.3 Funnel plot

From our meta-analysis (9 studies), the funnel plots $Y_{i,AC}$ against $\nu_{i,AC}$ and $Y_{i,BC}$ against $\nu_{i,BC}$ are shown in Figures 7.1 and 7.2 respectively. The conditional means $E(Y_{i,AC}|c_i)$ and $E(Y_{i,BC}|c_i)$ are represented by the dashed line in both figures. As mentioned earlier, notice that the conditional mean in both figures are not smooth functions when plotted against the conditional variance. Plot 7.1 indicates that smaller studies (larger $\nu_{i,AC}$) give more positive results than larger studies (smaller $\nu_{i,BC}$) and this plot has a trend. Funnel plot 7.2 shows a similar sign of selection bias to Figure 7.1. The problem of selection bias has arisen in the meta-analysis. Therefore, we would assume here that there might be other small studies comparing the treatments A, B and C, which have been carried out but which have not been selected in the meta-analysis.

7.4.4 Standard error

The standard errors of $Y_{i,AC}$ and $Y_{i,BC}$ in (7.8) for logistic regression model depend on value of treatment effects. Now we shall calculate the standard error for the model without treatment effect for the use later. Let n_{1i} and n_{2i} represent the summations $n_{iA} + n_{iC}$ and $n_{iB} + n_{iC}$ respectively. From the empirical log-odds ratios $Y_{i,AC}$ and $Y_{i,BC}$, and the conditional probabilities $f(r_{iA}|\delta_{i,AC})$ and $f(r_{iB}|\delta_{i,BC})$, we obtain the following standard errors (see the

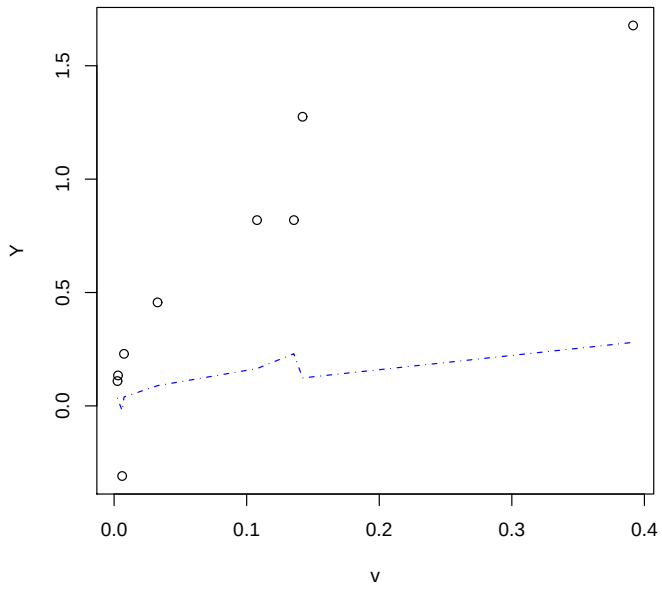


Figure 7.1: The funnel plot:$Y_{i,AC}$ against $v_{i,AC}$-the dashed lines represent the conditional mean of $Y_{i,AC}$ given c_i.

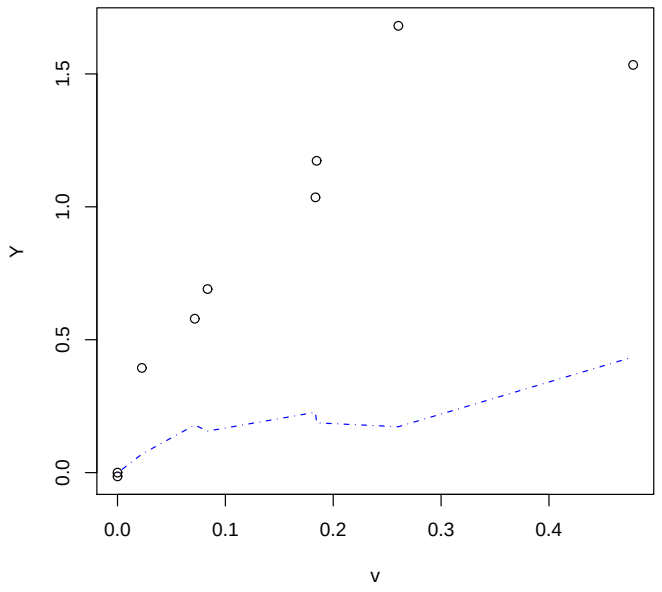


Figure 7.2: The funnel plot:$Y_{i,BC}$ against $v_{i,BC}$-the dashed lines represent the conditional mean of $Y_{i,BC}$ given c_i.

details in Shi and Copas, 2002)

$$s^*_{i,AC} = \sqrt{Var(Y_{i,AC}|\delta_{i,AC}=0,c_i)} = \left(\frac{n^3_{1i}}{c_i(n_{1i}-c_i)n_{iA}n_{iC}}\right)^{1/2}, \quad (7.13)$$

$$s^*_{i,BC} = \sqrt{Var(Y_{i,BC}|\delta_{i,BC}=0,c_i)} = \left(\frac{n^3_{2i}}{c_i(n_{2i}-c_i)n_{iB}n_{iC}}\right)^{1/2}. \quad (7.14)$$

7.5 Selection bias

As seen from the preceding section, there is selection bias in our meta-analysis. We apply the idea of the use of selection model from Chapter 6 in this section. We shall demonstrate how we assume a population model and how a study from the model will be selected. They are described as follows

1. **Population model**

   We shall assume the logistic regression model with conditional probability (7.5) to be a population model for treatment A versus B versus C.

2. **Selection event**

   A selection of studies from the population model can be chosen to represent our meta-analysis. To illustrate this selection, let $\mathcal{S}_1$ be the event that a study from the population model will be selected. This is under the expectation that larger studies are more likely to be selected than those smaller studies. Supposing that the event $\mathcal{S}_1$ has occurred then the population model with assuming $\mathcal{S}_1$ happened can be written as $p(\mathbf{r}_i|\mathcal{S}_1,c_i)$ (or called the combined model).

The combined model $p(\mathbf{r}_i|\mathcal{S}_1,c_i)$ can be derived as

$$p(\mathbf{r}_i|\mathcal{S}_1,c_i) = \frac{p(\mathcal{S}_1,\mathbf{r}_i|c_i)}{p(\mathcal{S}_1|c_i)},$$

$$= \frac{p(\mathbf{r}_i|c_i)p(\mathcal{S}_1|\mathbf{r}_i,c_i)}{p(\mathcal{S}_1|c_i)}, \quad (7.15)$$

where $p(\mathbf{r}_i|c_i)$ is the population model and given in (7.5) while $p(\mathcal{S}_1|\mathbf{r}_i, c_i)$ and $p(\mathcal{S}_1|c_i)$ are the probability of selection ($\mathcal{S}_1$ happened) for a typical study ($\mathbf{r}_i$) and the probability of selection ($\mathcal{S}_1$ happened). We need to define a selection model and calculate these probabilities. Note that the random treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ are included in those probabilities. The following details are for $p(\mathcal{S}_1|\mathbf{r}_i, c_i)$ and $p(\mathcal{S}_1|c_i)$ respectively.

(i). **The probability of selection event happened for a typical study**

Let q be the function of probability of a typical study as the ith study being selected, defined by

$$q(\mathbf{r}_i|\delta_{i,AC}, \delta_{i,BC}) = p(\mathcal{S}_1|\mathbf{r}_i, c_i, \delta_{i,AC}, \delta_{i,BC}). \tag{7.16}$$

We need to define the above selection probability. Now let us revise the selection model we used in Chapter 6 for normal approximation model. The normal approximaton model and the selection model Z_{i1} are given by

$$\begin{cases} Y_{i,AC} = \delta_{i,AC} + s^*_{i,AC}\epsilon^*_{i,AC}, \\ Y_{i,BC} = \delta_{i,BC} + s^*_{i,BC}\epsilon^*_{i,BC}, \end{cases} \tag{7.17}$$

$$Z_{i1} = a_1 + \frac{b_1}{\varphi^*_i} + \xi^*_i, \tag{7.18}$$

where $s^*_{i,AC}$ and $s^*_{i,BC}$ are the standard errors of $Y_{i,AC}$ and $Y_{i,BC}$ respectively. The function φ^*_i is the average of standard errors in the ith study, can be written as $(s^*_{i,AC} + s^*_{i,BC})/2$. The random residuals $(\epsilon^*_{i,AC}, \xi^*_i)$ and $(\epsilon^*_{i,BC}, \xi^*_i)$ are bivariate normal distributions with both means equal to zero and both variances equal to one. Their correlations are

$$corr(\epsilon^*_{i,AC}, \xi^*_i) = \varrho^*_1 \qquad \text{and} \qquad corr(\epsilon^*_{i,BC}, \xi^*_i) = \varrho^*_2.$$

If ϱ^*_1 and ϱ^*_2 are zero then it shows that the r_{iA}, r_{iB} and r_{iC} from the meta-analysis (or the outcome $(Y_{i,AC}, Y_{i,BC})$) have no effect on whether the study is selected or not.

This will be the model without assuming selection bias. On the other hand, if $\varrho_1^* > 0$ and $\varrho_2^* > 0$ then the selected studies are biased by the large values of $Y_{i,AC}$ and $Y_{i,BC}$.

Following the discussion given in Chapter 6, we have the following formula:

$$q(\mathbf{r}_i|\delta_{i,AC}, \delta_{i,BC}) = \Phi\left(\frac{\mu_{2i1}^*}{\sigma_{2i1}^*}\right), \tag{7.19}$$

where Φ is the standard normal cumulative distribution, and $\mu_{2i1}^* = E(Z_{i1}|(Y_{i,AC}, Y_{i,BC}))$ and $\sigma_{2i1}^{*2} = Var(Z_{i1}|(Y_{i,AC}, Y_{i,BC}))$ given in (6.7) and (6.8) respectively. For the logistic regression models (7.3) and (7.4) with conditional approach (7.5), we will still adopt the selection probability but $s_{i,AC}^*$ and $s_{i,BC}^*$ here are replaced by (7.13) and (7.14). For simplifying the computation, as assumed earlier, there is no association between the treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$.

The selection model defined as above is reasonable. Actually, the only requirement for selection probability is that it can model the phenomena shown in Figures 7.1 and 7.2, i.e. the large studies and the studies with positive results would tend to have larger selection probabilities than others. Selection probability (7.19) would satisfy the requirement. As discussed in Chapter 6, the parameters a_1 and b_1 are inestimable and whether the meta-analysis model with a selection model fit to the data will be checked by goodness-of-fit test and other statistical quantities in a sensitivity analysis.

(ii). **The marginal selection probability**

To estimate the probability of selection $p(\mathcal{S}_1|c_i)$, let Q_{i1} be the marginal selection probability given $\delta_{i,AC}$ and $\delta_{i,BC}$ and derived as

$$\begin{aligned} Q_{i1}(\delta_{i,AC}, \delta_{i,BC}) &= p(\mathcal{S}_1|c_i, \delta_{i,AC}, \delta_{i,BC}), \\ &= \sum_{\mathbf{u}_i} p(\mathcal{S}_1|\mathbf{r}_i = \mathbf{u}_i, c_i, \delta_{i,AC}, \delta_{i,BC}) p(\mathbf{r}_i = \mathbf{u}_i|c_i, \delta_{i,AC}, \delta_{i,BC}), \end{aligned} \tag{7.20}$$

where $p(\mathcal{S}_1|\mathbf{r}_i = \mathbf{u}_i, c_i, \delta_{i,AC}, \delta_{i,BC})$ is the probability of selection for a study including three arms and given in (7.19), and $p(\mathbf{r}_i = \mathbf{u}_i|c_i, \delta_{i,AC}, \delta_{i,BC})$ is the conditional probability model of $\mathbf{r}_i$ given c_i and given in (7.7). Note that the vector $\mathbf{u}_i$ is given in (7.6). Thus, we have

$$Q_{i1}(\delta_{i,AC}, \delta_{i,BC}) = \sum_{\mathbf{u}_i} q(\mathbf{u}_i|\delta_{i,AC}, \delta_{i,BC}) f(\mathbf{u}_i|\delta_{i,AC}, \delta_{i,BC}). \tag{7.21}$$

Equation (7.21) includes two random treatment effects. We shall integrate the marginal selection probability $Q_{i1}(\delta_{i,AC}, \delta_{i,BC})$ with respect to treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ respectively. The overall marginal selection probability is

$$\begin{aligned} Q_{i1} &= \int\int Q_{i1}(\delta_{i,AC}, \delta_{i,BC})\phi(\delta_{i,AC}; \mu_{AC}, \tau^2)\phi(\delta_{i,BC}; \mu_{BC}, \tau^2) d\delta_{i,AC} d\delta_{i,BC}, \\ &= \int\int \sum_{\mathbf{u}_i} f(\mathbf{u}_i|\delta_{i,AC}, \delta_{i,BC}) q(\mathbf{u}_i|\delta_{i,AC}, \delta_{i,BC})\phi(\delta_{i,AC})\phi(\delta_{i,BC}) d\delta_{i,AC} d\delta_{i,BC}, \end{aligned} \tag{7.22}$$

where $\phi(\delta_{i,AC})$ and $\phi(\delta_{i,BC})$ are the probability density functions of the normal distributions $N(\mu_{AC}, \tau^2)$ and $N(\mu_{BC}, \tau^2)$ respectively. Notice that the function $f(\mathbf{u}_i|\delta_{i,AC}, \delta_{i,BC})$ involves the random treatment effects so we need Guassian-Hermnite approximation to estimate in the usual way. After integrating, the Q_{i1} is an unconditional probability and does not depend on the $\delta_{i,AC}$ and $\delta_{i,BC}$. Note that the estimate from marginal selection probabily Q_{i1} is close to $\Phi(a_1 + b_1/\varphi_i)$ (obtained from equations (7.17) and (7.18)) (see the discussion from Shi and Copas, 2002).

7.6 Likelihood

The log-likelihood function of the conditional probability model with assuming selection event happened can be written as

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{9} \log p(\mathbf{r}_i|\mathcal{S}_1, c_i) = \sum_{i=1}^{9} \log \left(\frac{p(\mathbf{r}_i, \mathcal{S}_1|c_i)}{p(\mathcal{S}_1|c_i)} \right). \tag{7.23}$$

The right-hand side of above equation is obtained from the probability property. The collection of unknown parameters is

$$\boldsymbol{\theta} = \{\mu_{AC}, \mu_{BC}, \tau, \varrho_1^*, \varrho_2^*\}\,. \tag{7.24}$$

We need to handle with two random treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$ in the log-likelihood function $l(\boldsymbol{\theta})$. The probability $p(\mathbf{r}_i, \mathcal{S}_1|c_i)$ thus comprises two integrations which are with respect to both treatment effects. While $p(\mathcal{S}_1|c_i)$ is marginal selection probability given in (7.21) and involved integrations as givne in (7.22). Then, the right- hand side of $l(\boldsymbol{\theta})$ can be derived as

$$\begin{aligned} & \sum_{i=1}^{9} \left(\log p(\mathbf{r}_i, \mathcal{S}_1|c_i) - \log p(\mathcal{S}_1|c_i)\right) \\ = & \sum_{i=1}^{9} \left(\log \int\int p(\mathbf{r}_i, \mathcal{S}_1|c_i, \delta_{i,AC}, \delta_{i,BC})\phi(\delta_{i,AC})\phi(\delta_{i,BC})d\delta_{i,AC}d\delta_{i,BC} - \log(Q_{i1})\right). \end{aligned}$$

By using equation (7.15) in the term $p(\mathbf{r}_i, \mathcal{S}_1|c_i, \delta_{i,AC}, \delta_{i,BC})$, the log-likelihood function $l(\boldsymbol{\theta})$ is

$$\sum_{i=1}^{9} \left(\log \int\int f(\mathbf{r}_i|\delta_{i,AC}, \delta_{i,BC})q(\mathbf{r}_i|\delta_{i,AC}, \delta_{i,BC})\phi(\delta_{i,AC})\phi(\delta_{i,BC})d\delta_{i,AC}d\delta_{i,BC} - \log(Q_{i1})\right) \tag{7.25}$$

where $f(\mathbf{r}_i|\delta_{i,AC}, \delta_{i,BC})$ and $q(\mathbf{r}_i|\delta_{i,AC}, \delta_{i,BC})$ are given in the equations (7.5) and (7.19) respectively. The Q_{i1} in the last term is given in (7.22).

7.7 Goodness of fit

Suppose that the pair (a_1, b_1) is used in the selection process. To test whether the set (a_1, b_1) is a possible pair in the conditional probability model $p(\mathbf{r}_i|\mathcal{S}_1, c_1)$ or not, we adopt the test based on the goodness-of-fit test in Chapter 6. The null hypothesis for the test is $H_0 : \boldsymbol{\beta} = 0$ where $\boldsymbol{\beta}$ is the vector (β_1, β_2). We shall add the term $\beta_1 s_{i,AC}^*$ to μ_{AC} and $\beta_2 s_{i,BC}^*$ to μ_{BC} for

the treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$. This can be set to

$$\delta^*_{i,AC} \sim N(\mu_{AC} + \beta_1 s^*_{i,AC}, \tau^2) \qquad \text{and} \qquad \delta^*_{i,BC} \sim N(\mu_{BC} + \beta_2 s^*_{i,BC}, \tau^2), \tag{7.26}$$

where $s^*_{i,AC}$ and $s^*_{i,BC}$ are estimated from (7.13) and (7.14) respectively. After that, the treatment effects in (7.26) are applied to the log-likelihood function $l^*(\boldsymbol{\theta}, \boldsymbol{\beta})$, given by

$$\sum_{i=1}^{9} \left(\log \int \int f(\mathbf{r}_i | \delta^*_{i,AC}, \delta^*_{i,BC}) q(\mathbf{r}_i | \delta^*_{i,AC}, \delta^*_{i,BC}) \phi(\delta_{i,AC}) \phi(\delta_{i,BC}) d\delta_{i,AC} d\delta_{i,BC} - \log(Q^*_{i1}) \right), \tag{7.27}$$

where

$$Q^*_{i1} \quad = \quad \int \int \sum_{\mathbf{u}_i} f(\mathbf{u}_i | \delta^*_{i,AC}, \delta^*_{i,BC}) q(\mathbf{u}_i | \delta^*_{i,AC}, \delta^*_{i,BC}) \phi(\delta^*_{i,AC}) \phi(\delta^*_{i,BC}) d\delta_{i,AC} d\delta_{i,BC}. \tag{7.28}$$

The likelihood ratio statistics for $H_0 : \boldsymbol{\beta} = 0$ is

$$2 \left(l^*(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\beta}}) - l^*(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\beta}=0}, \boldsymbol{\beta} = \mathbf{0}) \right) \sim \chi^2_2 \text{ under } H_0, \tag{7.29}$$

where $(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\beta}})$ is MLEs by maximizing the log-likelihood function (7.27) while $(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\beta}=0}, \boldsymbol{\beta} = \mathbf{0})$ is the MLEs from (7.27) with restriction $\boldsymbol{\beta} = 0$. The interpretation of test is similar as explained in Chapter 6. If the null hypothesis is accepted, it means that the pair (a_1, b_1) is a plausible choice of the model $p(\mathbf{r}_i | \mathcal{S}_1, c_1)$ and makes the funnel plots fit well.

7.8 Sensitivity analysis

We use the similar idea in Section 6.5 to conduct a sensitivity analysis here. We allow the conditional probability model $p(\mathbf{r}_i | \mathcal{S}_1, c_i)$ to have different amounts of selection bias depending on the pair (a_1, b_1) in the selection model $p(\mathcal{S}_1 | \mathbf{r}_i, c_i)$ or $p(\mathcal{S}_1 | c_i)$. The steps of sensitivity analysis are given as follows.

- **Step 1: Determine the range of** (a_1, b_1)

We use three typical pairs: (0.99,0.80), (0.80,0.50) and (0.60,0.30), in the area of

$$0.01 \leq P_{min}(\text{selection}) \leq P_{max}(\text{selection}) \leq 0.99,$$

where $P_{min}(\text{selection})$ and $P_{max}(\text{selection})$ are given in (6.19). By using three typical pairs to identify the pair (a_1, b_1), the pairs relating to the selection probability pair are given in Table 7.1. The model without assuming $\mathcal{S}_1$ happened (standard model) is obtained by using the first pair of Table 7.1 in the model $p(\mathbf{r}_i|\mathcal{S}_1, c_i)$.

Table 7.1: The pairs of (a_1, b_1) for the selection model Z_{i1}

Selection probability pairs	a_1	b_1
(1.0,1.0)	6.0	0.0
(0.99,0.80)	0.3793294	0.4492381
(0.80,0.50)	-0.8844004	0.8594276
(0.60,0.30)	-1.3416807	0.7942026

- **Step 2: Estimation and goodness-of-fit test**

  We will use each combination (a_1, b_1) in Table 7.1 to calculate each of the following quantities.

  1. $\widehat{\mu_{AC}}$;
  2. p-value for testing $H_0 : \mu_{AC} = 0$;
  3. standard error of $\widehat{\mu_{AC}}$;
  4. lower limit of the 95% confidence interval for μ_{AC};
  5. upper limit of the 95% confidence interval for μ_{AC};
  6. $\widehat{\mu_{BC}}$;
  7. p-value for testing $H_0 : \mu_{BC} = 0$;
  8. standard error of $\widehat{\mu_{AC}}$;
  9. lower limit of the 95% confidence interval for μ_{BC};

10. upper limit of the 95% confidence interval for μ_{BC};
11. P_{max}(selection) for the selection model Z_{i1};
12. P_{min}(selection) for the selection model Z_{i1};
13. estimated number of selected and unselected studies given by $\sum_i \{p(Z_{i1} > 0|\varphi_i)\}$;
14. p-value for the fit for the funnel plot corresponding to the null hypothesis H_0 : $\boldsymbol{\beta} = 0$.

Table 7.2: The bias-simulated data with selection: summary of outputs

[, 1]	[, 2]	[, 3]	[, 4]	[, 5]	[, 6]	[, 7]
0.5054664	0.0124615	0.014580	0.4768896	0.5340432	0.6369195	0.0014782
0.5050151	0.0124782	0.038970	0.4286339	0.5813963	0.6364665	0.0021454
0.3380378	0.0430073	0.135470	0.0725166	0.6035590	0.4668118	0.0062834
0.1314767	0.3789370	0.481100	-0.8114793	1.0744327	0.2589700	0.0879419
[, 8]	[, 9]	[, 10]	[, 11]	[, 12]	[, 13]	[, 14]
0.254810	0.1374919	1.1363471	1.00	1.00	9	0.0999526
0.021450	0.5948775	0.6789615	0.99	0.80	10	0.1446150
0.654800	-0.8165962	1.7502198	0.80	0.50	13	0.4337069
0.258900	-0.2484740	0.7664140	0.60	0.30	20	0.6969515

The MLEs for μ_{AC} and μ_{BC} are presented in columns 1 and 6 respectively. By calculating the asymptotic variance-covariance matrix, described in Chapter 4, their standard errors are shown in columns 3 and 8 respectively. Columns 4 and 5, and columns 9 and 10 are the lower and upper limits of the 95% confidence intervals for $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ respectively. Note that the significance level in this section is 0.10.

(i). The first row represents the results for the standard conditional probability model without assuming selection event. The estimates for μ_{AC} and μ_{BC} are 0.5054664 and 0.6369195 respectively. The presence of selection bias can be detected from the p-value of $H_0 : \boldsymbol{\beta} = \mathbf{0}$ (column 14). The p-value 0.0999526 shows that the model is slightly biased.

(ii). Considering the selection bias, the $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ decrease gradually while the estimated number of population studies (column 13) increases.

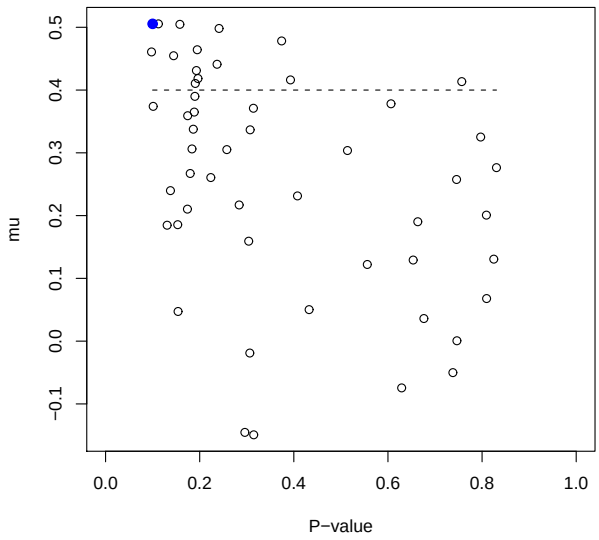


Figure 7.3: $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = \mathbf{0}$

(iii). P-values of $H_0 : \widehat{\mu_{AC}} = 0$ (column 2) and $H_0 : \widehat{\mu_{BC}} = 0$ (column 7) are significant in all rows.

(iv). Using the goodness-of-fit test, the p-value of $H_0 : \boldsymbol{\beta} = \mathbf{0}$ indicates that the model with assuming the selection event has improved from reading downward.

- **Step 3: Sensitivity analysis**

  The plots of $\widehat{\mu_{AC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = \mathbf{0}$ and $\widehat{\mu_{BC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = \mathbf{0}$ are shown in Figures 7.3 and 7.4 respectively. The estimates for μ_{AC} and μ_{BC} from our meta-analysis (9 studies) are presented in the blue and red solid dots in Figures 7.3 and 7.4 respectively. By using our sensitivity analysis, the plots show that $\widehat{\mu_{AC}}$ and $\widehat{\mu_{BC}}$ can be anything less than 0.45 and 0.50 respectively. Also their plausible estimates with p-value 0.5 should be around 0.30 and 0.40 respectively.

To conclude, we can see the plausible estimates are acceptable comparing to the true values (0.4 and 0.6) of μ_{AC} and μ_{BC} in the population data of Section 7.2.

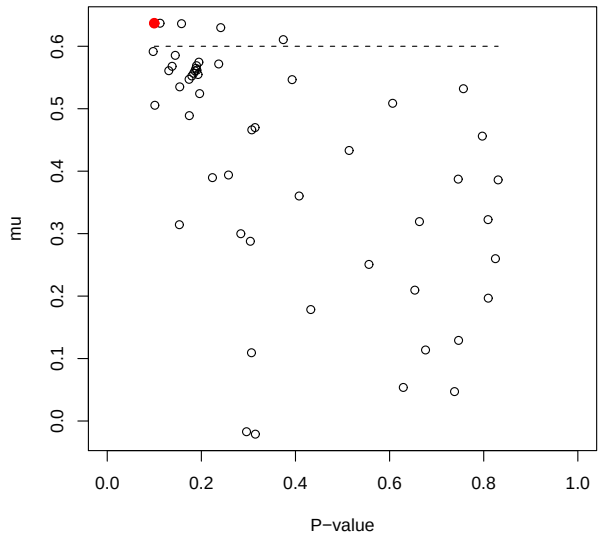


Figure 7.4: $\widehat{\mu_{BC}}$ against the p-value of $H_0 : \boldsymbol{\beta} = \mathbf{0}$

7.9 Discussion

We first used the sensitivity analysis to the W2 data but there was no evidence of selection bias. Consequently, we generated the three-arm data to be the population and made this data bias from selection. In this chapter, we assume that there is no association between both treatment effects $\delta_{i,AC}$ and $\delta_{i,BC}$.

In general, we extend the sensitivity analysis from the previous chapter to the conditional probability model. We use the exact distribution of the data with the conditional method to represent the population model and apply the formulae of the normal approximation model with selection, expressed in Chapter 6, for the selection of the event $\mathcal{S}_1$. Thus, the probability of being selected for a typical study $q(\mathbf{r}_i|\delta_{i,AC}, \delta_{i,BC})$ is obtained from $p(Z_{i1} > 0|Y_{i,AC}, Y_{i,BC})$. As we have discussed in Section 7.5, this selection probability model is still relevant for modelling selection bias such as appeared in Figures 7.1 and 7.2, and it can be used in a

sensitivity analysis.

The test for the pair (a_1, b_1) of the selection model is similar to the goodness-of-fit in Chapter 6. Since there are two random effects involved in the likelihood function, the estimation for the log-likelihood function is complicated and takes long time. As before, Gaussian quadrature has been used for integral estimation. Alternatively, we can use the other methods, mentioned in Chapter 4 to estimate the integral. Shi and Copas (2002) used a Markov chain Mote Carlo EM algorithm to estimate MLEs for the meta-analysis of 2×2 tables using exact conditional distributions.

In this chapter, we discuss a model with three treatments. More treatment comparisons can be applied to the sensitivity analysis here but the complexity of conditional model would make the calculation difficult, particularly its denominator. In addition, we will have more free parameters in the likelihood if multiple-selection models are exploited.

Chapter 8

Conclusions and further development

8.1 Conclusions

Meta-analysis is a statistical tool that summarizes evidence from multiple studies of a particular topic and attempts to provide an estimate of true effect. The aims of meta-analysis of multi-arm trials are to combine evidence from all possible similar studies and draw inferences about the effectiveness of multiple compared-treatments. Throughout the thesis, we have used two meta-analyses of multi-arm trials data (W1 and W2) to different model strategies. If the number of individual studies (n_{ij}) is large enough (larger than 20) and r_{ij} is not too small and not too close to n_{ij}, for example from the W1 data then the normal approximation model is appropriate. For the empirical log-odds model, the trial effects in meta-analysis would not satisfy any model (fixed effect or random effect) because they are pooled from different design models. Thus, the trial effects were assumed to be different. This makes the logistic regression model include M (the number of studies in meta-analysis) unknown parameters in the likelihood function and may cause the problem of many nuisance parameters and inconsistent estimate. To avoid these problems, the empirical log-odds ratios model can be proposed. We compare the small and large numbers of n_{ij} for empirical log-odds ratio model in simulation study of Chapter 5. The results show that the model are suitable for large individual studies. However, if M is not too large; the empirical log-odds

and empirical log-odds ratio models may give the similar results.

The logistic regression model can be employed to any multi-arm trials data. Two approaches, unconditional and conditional are used to make inferences. The logistic regression models are applied to the W2 data due to the small number of n_{ij}. The logistic regression model using the unconditional method includes nuisance parameters. The model should be used with a small number of studies. The unconditional maximum likelihood estimate may be biased if n_{ij} is small and M is large (Lubin, 1981; Cox and Snell, 1989). The main advantage of the conditional likelihood approach is that the likelihood depends only on the parameters of interest. This gives a consistent estimates and the computation is stable. The results from simulation study of Chapter 5 support our conclusions for the normal approximation model and the logistic regression model using unconditional and conditional methods.

The empirical log-odds ratio models have been used for the W1 data in Chapter 3. However we found that studies with positive results were more likely to be selected, it could therefore lead to selection bias (positive bias). A sensitivity analysis by using a selection model has been employed to examine the selection bias and corrected the results under the controlled assumptions for the model. The selection model is regarded as a tool of sensitivity analysis. The missing studies in funnel plot can be treated as non-ignorable missing data in meta-analysis. Similarly, the sensitivity analysis is extended to the logistic regression model.

8.2 Further devolopment

We proposed unconditional and conditional likelihood for meta-analysis with the logistic regression model in Chapters 4 and 5. Although conditional approach shows good performance in theory and in our simulation studies, it is of interest to compare the method with some other methods, for example, restricted maximum likelihood estimation (REML), penalized quasi-likelihood (PQL) estimation.

Gauss-Hermite quadrature approximation has been used to approximate the integral form of probabilities including random effects in the likelihood function for the logistic regression model. By using different number of nodes for approximation, the results from the model were not much different. As mentioned in Chapter 4, the approximation is reasonably effective for low-order integrations depending on the number of treatments involved in those studies. If this number is large then it makes the dimensionality of the integral large and the approximation cannot give an accurate approximation. If there are more than three treatments (two pairwise-comparisons) in multi-arm trials, we may need to use some other methods, for example, Laplace approximation method or Monte Carlo EM algorithm, see Ripatti and Palmgren (2000); Shi and Copas (2002).

In Chapter 5, we focus on comparing three methods used in this thesis with a special case that there is no association between the treatment effects ($\rho = 0$) and the direct comparisons are only involved. The parameter ρ is of interest. It is estimable if enough information is provided for indirect comparison. It is worth a further study on this parameter, by a comprehensive simulation study and analysis of more real data. From Chapter 7, the estimation of log-likelihood function for the model with selection models is complicated and takes long time. Alternatively, we can use a Markov chain Monte Carlo EM algorithm to estimate MLEs. We used the method to the simulated data. Further, more real data can be applied to the method.

Bibliography

Abramowitz, M. and Stegun, I. (1972). *Handbook of mathematical Functions*. Dover.

Ades, A. (2003). A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Statistics in Medicine*, 22:2995–3016.

Ades, A. and Cliffe, S. (2002). Markov chain monte carlo estimation of a multi-parameter decision model:consistency of evidence and the accurate assesment of uncertainty. *Medical Decision Making*, 22:359–371.

Albert, A. and Anderson, J. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika*, 71(1):1–10.

Anscombe, F. (1956). On estimating binomial response relations. *Biometrica*, 43:461–464.

Bailey, K. (1987). Inter-study differences-how should they influence the interpretation and analysis of results. *Statistics in Medicine*, 6:351–60.

Begg, C. and Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50:1088–101.

Bellio, R. and Sartori, N. (2003). Extending conditional likelihood in models for stratified binary data. *SMA*, 12:121–132.

Boissel, J., Blanchard, J., Panak, E., Peyrieux, J., and Sacks, H. (1989). Considerations for the meta-analysis of randomized clinical trials:summary of a panel discussion. *Controlled Clin. Trials*, 10:254–81.

Bonchek, L., Boerboom, L., and Olinger, G. (1982). Prevention of lipid accumulation in experimental vein bypass grafts by antiplatelet therapy. *Circulation*, 66:338–41.

Bucher, H., Guyatt, G., Griffith, L., and Walter, S. (1997). The result of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J. Clin Epidemiol*, 50:683–691.

Bulmer, M. (1979). *Principle of statistics, 2nd edition.* Dover Publication, London.

Bussuyt, P., Reitsma, J., Bruns, D., Gatsonsis, C.A.and Glasziou, P., and Irwig, L. f. t. S. g. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy:the stard initiative. *Clinical Chemistry*, 49:1–6.

Campeau, L., Enjalbert, M., and Lesperance, J. (1984). The relation of risk factors to the development of atherosclerosis in saphenous vein bypass grafts. *N Engl J Med*, 311:1329–32.

Chalmers, T., Matta, R.J.and Smith, H., and Kunzler, A. (1977). Evidence favouring the use of anticoagulants in the hospital phase of acute myocardial infarction. *New Engl. J. Med.*, 297:1091–6.

Chootrakool, H. and Shi, J. (2008). Meta-analysis of multi-arm trials using empirical logistic transform. *The Open Medical Informatics Journal*, 2:105–109.

Collaboration, A. (1994a). Collaborative overview of randomised trials of antiplatelet therapy- 2:maintenance of vascular graft or arterial patency by antiplatelet therapy. *BMJ*, 308:159–68.

Collaboration, A. (1994b). Collaborative overview of randomised trials of antiplatelet therapy- 3:reduction in venous thrombosis and pulmonary embolism by antiplatelet prophylaxis among surgical and medical patients. *BMJ*, 308:235–46.

Collett, D. (1991). *Modelling binary data.* Chapman and Hall, London.

Collins, R., Scrimgeour, A., Yusuf, S., and Peto, R. (1988). Reduction in fatal pulmonary embolism and venous thrombosis by perioperative administration of subcutaneous heparin. *N Engl J Med*, 318:1162–73.

Copas, J. (1999). What works?:selectivity models and meta-analysis. *J.R. Statist. Soc. A*, 162:95–109.

Copas, J. and Shi, J. (2000). Reanalysis of epidemilogical evidence on lung cancer and passive smoking. *BMJ*, 320:417–18.

Copas, J. and Shi, J. (2001). A sensitivity analysis for publication bias in systematic reviews. *Statistics in Medicine*, 10:251–265.

Copas, J. and Shi, J. (2002). Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics*, 1:247–262.

Coursol, A. and Wagner, E. (1986). Effect of positive findings on submission and acceptance rates:a note on meta-analysis bias. *Professional Psychol*, 17:136–7.

Cox, D. (1970). *Analysis of Binary data*. Chapman and Hall, London.

Cox, D. (1972). Regression models and life table (with discussion). *J.R.Statist.Soc. B*, 34:187–220.

Cox, D. and Snell, E. (1989). *Analysis of Binary data, 2nd ed.* Chapman and Hall, London.

Crouch, E. A. and Spiegelman, D. (1990). The evaluation of integrals of the form $\int_{-\infty}^{\infty} f(t)\exp(-t^2)dt$:application to logistic-normal models. *J. Am. Statist. Assoc.*, 85:464–9.

Dalen, J., Paraskos, J., Ockene, I., Alpert, J., and Hirsh, J. (1986). Venous thromboembolism: scope of the problem. *Chest*, 89:370–3S.

DerSimonian, R. and Laird, N. (1986). Meta-analysis in clinical trials. *Controlled clin Trials*, 7:177–88.

Domenici, F., Parmigiani, G.and Wolpert, R., and Hasselblad, V. (1999). Meta-analysis of migraine headache treatments: combining information from heterogeneous designs. *Journal of the American Statistical Association*, 94:16–28.

Duval, S. and Tweedie, R. (2000a). A non-parametric 'trim and fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95:89–98.

Duval, S. and Tweedie, R. (2000b). Trim and fill:a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56:455–463.

Eddy, D., Hasselblad, V., and Shachter, R. (1992). *Meta-Analysis by the Confidence Profile Method*. Academic Press.

Egger, M. and Davey Smith, G. (1995). Misleading meta-analysis. *British Medical Journal*, 310:752–754.

Egger, M., Davey Smith, G., Schneider, M., and Minder, C. (1987). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315:629–634.

Egger, M., Smith, G., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Br. Med. J.*, 315:629–34.

Emmons, P., Harrison, M., and A.F., H. (1965). Effect of a pyrimido-pyrimidine derivative on thrombus formation in the rabbit. *Nature*, 208:255–7.

Fuster, V. and Chesebro, J. (1986). Role of platelets and platelet inhibitors in aorto-coronary vein graft disease. *Circulation*, 73:227–32.

Galbraith, R. (1988). A note on graphical presentation of estimated odds ratio from several clinical trials. *Stat. Med.*, 7:889–94.

Gent M., R. R. (1986). A meta-analysis of the studies of dihydroergotamine plus heparin in the prophylaxis of deep vein thrombosis. *Chest*, 89:396–400S.

Gillum, R. (1987). Coronary artery bypass surgery and coronary angiography in the united states,1979-1985. *Am Heart J*, 113:1255–60.

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

Glasziou, P. and Sanders, S. (2002). Investigating causes of heterogeneity in systematic reviews. *Statistics in Medicine*, 21:1503–1511.

Glenny, A., Altman, D.G.and Song, F. S., Deeks, J., and D'Amico, R, e. a. (2005). Indirect comparisons of completing interventions. *Health Technol Assess*, 9(26):1–148.

Golub, G. and Welsch, J. (1969). Calulation of gauss quadrature rules. *Mathematics of Computation*, 23(106):221–230.

Green, S., Benedetti, J., and Crowlet, J. (1997). *Clinical trials in oncology*. Chapman & Hall.

Greenwald, A. (1975). Consequences of prejudice against the null pypothesis. *Psychol Bull*, 82:1–20.

Haldane, J. and Smith, C. (1948). A simple exact test for birth-order effect. *Ann. Eugenics*, 14:117–124.

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

Hasselblad, V. (1998). Meta-analysis of multi-treatment studies. *Medical Decision Making*, 18:37–43.

Hedges, L. (1984). Estimation of effect size under nonrandom sampling: The effects of censoring studies yielding statistically insignificant mean differences. *Journal of Educational Statistics*, 9:61–85.

Hedges, L. and Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic press, Inc, London.

Hedges, L. and Pigott, T. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, 6:203–217.

Higgins, J. and Whitehead, A. (1996). Borrowing strength from external trials in a meta-analysis. *Statistics in Medicine*, 15:2733–2749.

Higgins, J., Whitehead, A., Turner, R., Omar, R., and Thompson, S. (2001). Meta-analysis of continuous outcome data from individual patients. *Statistics in Medicine*, 20:2219–2241.

Hill, A. (1990). Memories of the British Streptomycin trial in tuberculosis. *Controlled Clinical Trials*, 11:77–9.

Hinkelmann, K. and Kempthorne, O. (1994). *Design and Analysis of Experimental Design*. Wiley.

Hirji, K. (1994). Exact analysis for paired binary data. *Biometrika*, 50(4):964–974.

Hirji, K., Mehta, C., and Patel, N. (1987). Computing distribution for exact logistic regression. *Journal of the American Statistical Association*, 82(400):1110–1117.

Iyengar, S. and Greenhouse, J. (1988). Selection models and the file drawer problem. *statistical Science*, 3:109–135.

Jackson, D. (2006). The power of the statistical test in meta-analysis. *Statistics in Medicine*, 25:2688–2699.

Jadad, A. (1998). *Randomized Controlled Trials*. BMJ publishing Group.

Jenicek, M. (1989). Meta-analysis in medicine:where we are and where we want to go. *J. Clin. Epidemiol*, 42:35–44.

Jennison, C. and Turnbull, B. (1990). Statistical approaches to interim monitoring of medical trials:a review and commentary. *Statistical Science*, 5:299–317.

Kakkar, V. (1981). Prevention of venous thromboembolism. *Clin Haematol*, 10:543–82.

Kleinbaum, D. (1994). *Logistic regression.* Springer-Verlag.

Lewis, S. and Clarke, M. (2001). Forest plot:trying to see the wood and the trees. *BMJ*, 322:1479–1480.

Light, R. and Pillemer, D. (1984). *Summing up: The science of reviewing research.* Harvard University press.

Little, R. and Rubin, D. (2002). *Statistical Analysis with missing data, 2nd.* John Wiley & Sons,Inc.

Liu, Q. and Pierce, D. (1994). A note on Gauss-Hermite quadrature. *Biometrika*, 81:624–9.

Lu, G. and Ades, A. (2004). Combination of direct and indirect evidence in mixed treatment comparison. *Statistics in Medicine*, 23:3105–3124.

Lu, G. and Ades, A. (2006). Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association*, 101(474):447–459.

Lu, G., Ades, A., Sutton, A.J.and Cooper, N. B. A., and Caldwell, D. (2007). Meta-analysis of mixed treatment comparisons at multiple follow-up times. *Statistics in Medicine*, 26:3681–3699.

Lubin, J. (1981). An empirical evaluation of the use of conditional and unconditional likelihoods for case-control data. *Biometrika*, 68(2):567–71.

Lumley, T. (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine*, 21:2313–2324.

Matthews, J. (2005). *Randomized Controlled Clinical Trials*. Academic Press.

McCarthy, W. (2007). The existence of maximum likelihood estimates for the binary response logistic regression model. *COBRA Preprint Series*, Article 26.

Moayyedi, P. (2004). Meta-analysis:can we mix apples and oranges. *American Journal of Gastroenterology*, pages 2297–2301.

Mother, D., Cook, D., Eastwood, S., Olikin, I., Rennie, D., and Stroup, D. (1999). Improving the quality of reporting of meta-analysis of randomised controlled trials: the quorum statement. *The Lancet*, 354:1896–1900.

Mother, D., Schulz, K., and Altman, D. f. t. C. g. (2001). The consort statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet*, 357:1191–1194.

Oxman, A., Clarke, M., and Stewart, L. (1995). From science to practice. meta-analyses using individual patient data are needed. *Journal of the American Medical Association*, 274:845–6.

Pagliaro, L., D'Amico, G., Sőrensen, T., Lebrec, D.and Burroughts, A. M. A. T. F., Politi, F., and Traina, M. (1992). Prevention of first bleeding in cirrhosis:a meta-analysis of randomzed trials of nonsurgical treatment. *Annals of Internal Medicine*, 117:59–70.

Palmgren, J. (1981). The Fisher information matrix for log linear models arguing conditionally on observed explanatory variables. *Biometrika*, 68(2):563–566.

Party, B., Lumley, T., Furberg, C.D.and Schellenbaum, G., Pahor, M., Alderman, M., and Weiss, N. (2003). Health outcomes associates with various antihypertensive therapies used as first-line agents. *Journal of the American Medical Association*, 289:2534–2544.

Pearson, K. (1904). Report on certain enteric fever inoculation statistics. *Br.Med.J.*, 3:1243–6.

Pendergast et al., J. (1996). A survey of methods for analyzing clustered binary response data. *Internat.Statist. Rev.*, 64:89–118.

Petrie, A. and Sabin, C. (2005). *Medical Statistics at a glance, 2nd ed.* Blackwell.

Petticrew, M. (2001). Systematics reviews from astronomy to zoology:myth and misconceptions. *British Medical Journal*, 322:98–101.

Pinello, D. (1999). Linking party to judicial ideology in american courts: A meta-analysis. *The justice System Journal*, 20:219–254.

Pirk, J., Ruzbarsky, V., and Konig, J. (1990). The effect of antiaggregating drugs on the patency of grafts in the arterial system. *World J Surg*, 4:615–20.

Pocock, S. (1989). *Clinical Trials: A practical approach.* Wiley, Chichester.

Prentice, R. (1976). Use of the logistic regression model in retrospective studies. *Biometrics*, 32:599–606.

Prentice, R. and Breslow, N. (1978). Restrospective studies and failure time models. *Biometrika*, 65:167–80.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna,Austria. ISBN3-900051-07-0.

Ripatti, S. and Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, 56:1016–1022.

Rothstein, H., Sutton, A., and Borenstein, M. (2005). *Publication bias in Meta-analysis: prevention, assesment and adjustments.* Wiley, Chichester.

Sartori, N. (2003). Modifiled profile likelihoods in models with stratum nuisance parameters. *Biometrika*, 90:533–549.

Scheff'e, H. (1959). *Analysis of Variace*. Wiley, New York.

Scheid, F. (1988). *Numerical Analysis 2/ed*. McGraw-Hill, New york.

Schmidt, F. (1988). Validity generalization and the future of criterion-related validity. *Test validity*, pages 173–189.

Schmidt, F. and Hunter, J. (1981). Employment testing:ole theories and new research findings. *American Psychologist*, 36:1128–1137.

Schmidt, F. and Hunter, J. (1998). The validity and utidity of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124:1128–1137.

Severini, T. A. (1998). Likelihood functions for inference in the presence of a nuisance parameter. *Biometrika*, 85:507–522.

Shadish, W. and Haddock, C. (1987). Combining estimates of effect size. in: Cooper, h. hedges,l.v., editors. *Statistics in Medicine*, 6:351–60.

Shi, J. and Copas, J. (2002). Publication bias and meta-analysis for 2× 2. tables: an average markov chain monte carlo em algorithm. *J. R. Statist Soc. B*, 64:221–236.

Shi, J. and Copas, J. (2004). Meta-analysis for trend estimation. *Statistics in Medicine*, 23:3–19.

Simmonds, M., Higgins, J., Stewart, L., Tierney, J., and Clark, M. T. S. (2005). Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clinical Trials*, 2:209–217.

Smith, T., Spiegelhalter, D., and Thomas, S. (1995). Bayesian approaches to random-effects meta-analysis:a comparative study. *Statistics in Medicine*, 14:2685–2699.

Sommer, B. (1987). The file drawer effect and publication rates in menstrual cycle research. *Psychol. of Women Quart.*, 11:233–42.

Song, F., Altman, D., Glenny, A., and Deeks, J. (2003). Validity of indirect comparison for estimating efficacy of competing interventions:empirical evidence from published meta-analyses. *BMJ*, 326:472–6.

Song, F., Khan, K., Dinnes, J., and Sutton, A. (2002). Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *International Epidemiological Association*, 31:88–95.

Song, F., Sheldon, T.A.and Sutton, A., Abrams, K., and Jones, D. (2001). Method for exploring heterogeniety in meta-analysis. *Evaluation & the Health Professions*, 24(2):126–151.

Stanley, T. (1998). New wine in old bottles: A meta-analysis of ricardian equivalence. *Southern Economic Journal*, 64:713–727.

Stanley, T. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, 15:131–150.

Stanley, T. and Jarrel, S. (1989). Meta-regression analysis: A quantitative method of literature surveys. *Journal of Economics Surveys*, 3:161–169.

Stanley, T. and Jarrel, S. (1998). Gender wage discrimination. *Journal of Human Resources*, 33:947–973.

Sterne, J. and Egger, M. (2001). Funnel plots for detecting bias in meta-analysis. *Journal of Clinical Epidemiology*, 54:1046–1055.

Stewart, L. and Parmar, M. (1993). Meta-analysis of the literature or of individual patient data:is there a difference? *Lancet*, 341:845–6.

Sutton, A., Abrams, K., Jones, D., Sheldon, T., and Song, F. (2000). *Methods for Meta-Analysis in Medical Research*. Wiley.

Thompson, S. and Sharp, S. (1999). Explaining heterogeneity in meta-analysis:a comparison of methods. *Statistics in medicine*, 18:2693–2708.

Tritchler, D. (1984). An algorithm for exact logistic regression. *Journal of the American Statistical Association*, 79:709–711.

Van Houwelingen, H. (1997). The future of biostatitics:expecting the unexpected. *Statistics in Medicine*, 16:2773–2784.

Von Elm, E. and Egger, M. (2004). The scandal of poor epidemilogical research. *British Medical Journal*, 329:868–869.

Welton, N. and Ades, A. (2005). A toxoplasmosis incidence in the uk:evidence synthesis and consistency of evidence. *Applied Statistics*, 54:385–404.

Whitehead, A. (2002). *Meta-Analysis of controlled clinical trials*. Wiley.

Yates, F. and Cochran, W. (1938). The analysis of groups of experiments. *J.Agricultural Sci*, 28:556–80.

Yazdanpanah, Y., Sissoko, D., Egger, M., Mouton, Y., and Zwahlen, M. Chene, G. (2004). Clinical efficacy of antiretroviral combination therapy based on protease inhibitors or non-nucleoside analogue reverse transcriptase inhibitors:indirect comparisonof controlled trials. *BMJ*, 328:249–53.

Yusuf, S., Peto, R., Lewis, J., Collins, R., and Sleight, P. (1985). Beta blockade during and after myocardial infarction:an overview of the randomised trials. *Progress in Cardiovascular Diseases*, 27:335–71.

Zhao, L. L. P. and Aragaki, C. (2000). A note on a conditional-likelihood approach for family-based assciation studies of candidate genes. *Hum Hered*, 50:194–200.

Ziegler, S., Koch, A., and Victor, N. (2001). Deficits and remedy of the standard ranodm effects methods in meta-analysis. *Methods of Information in Medicine*, 40:148–155.