A COMPARISON OF BAYESIAN AND CLASSICAL META-ANALYSES

A SYSTEMATIC REVIEW AND META-ANALYSIS OF STUDIES OF PREOPERATIVE ASPIRIN ON BLEEDING AND CARDIOVASCULAR OUTCOMES OF PATIENTS UNDERGOING CORONARY ARTERY BYPASS SURGERY: A COMPARISON OF BAYESIAN AND CLASSICAL APPROACHES

By

JI CHENG, B.A., B.Sc.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Master of Science

McMaster University
© Copyright by Ji Cheng, April 2007

MASTER OF SCIENCE (2007) McMaster University

(Statistics) Hamilton, Ontario

TITLE: A Systematic Review and Meta-analysis of Studies of Preoperative Aspirin on Bleeding and Cardiovascular Outcomes of Patients Undergoing Coronary Artery Bypass Surgery: A Comparison of Bayesian and Classical Approaches

AUTHOR: Ji Cheng, B.A., B.Sc. (McMaster University)

SUPERVISOR: Dr. Lehana Thabane

NUMBER OF PAGES: vii, 48

Abstract

Meta-analysis is a statistical method to summarize the overall evidence of effects on intervention by systematically combining outcomes from available studies in the literature which are homogeneous in research methodology and research interest. The objective of this project is to evaluate the treatment effects of preoperative aspirin on bleeding and other cardiovascular outcomes from 11 randomized control trials (RCT) and 19 observational (non-RCT) studies. Both Bayesian meta-analysis and classical (frequentist) meta-analysis were applied to continuous and binary outcomes, and the results were compared.

The robustness of the Bayesian approach is assessed by examining the performances of different likelihood functions and priors. We also discuss strategies on dealing with zero-event studies for binary outcomes, and the implementation of multiple imputation (MI) technique to missing data for continuous outcomes.

Most results of primary analysis agree between the Bayesian and classical approaches. We suggest that the final conclusion of a meta-analysis should be based on the comparison of the results from both Bayesian and classical approaches.

Acknowledgements

First and foremost, I would like to extend my deepest gratitude to Dr. Lehana Thabane for giving me the opportunity to work on this project. Without his expert guidance and all-out support, this project would not have been possible.

Special thanks to my committee members, Dr. Peter Macdonald and Dr. Roman Viveros-Aguilera, who offered valuable advice.

Many thanks to Jack Sun and Richard Whitlock, who conducted the literature review and extracted the data. Also, thanks to Susan Tomlinson and Janine Arkinson for their proofreading.

Finally, I would like to thank all of my friends and family for their love and help, especially my husband, Xiaosong Nie, who endured this long process with me, always offering unwavering support and understanding.

Table of Contents

Chapter 1

Introduction

1.1 Background

Aspirin (ASA) is one of the common anti-platelet therapies used to treat patients with ischemic heart disease and clot-related strokes [1]. Coronary artery bypass grafting (CABG) is a type of heart surgery which re-routes, or "bypasses", blood around clogged arteries to improve blood flow and oxygen delivery to the heart [2]. Many studies have reported that preoperative aspirin increases patient blood loss, blood transfusion requirement, and incidence of re-operation due to bleeding [3]. However, other studies have shown a benefit of increased early and late vein graft patency in patients who received peri-operative aspirin therapy [4]. Therefore, risk-benefit assessment of aspirin therapy for CABG patients is controversial in the literature.

Despite the CABG guidelines recommending that aspirin should be discontinued seven days prior to coronary bypass surgery [5], many cardiac surgeons allow their patients to continue to take aspirin until the day of surgery, due to the aforementioned conflicting evidence [6]. Therefore, a systematic review is needed to summarize the results of the independent studies in the literature, and determine conclusive evidence on how aspirin affects the postoperative bleeding and the other associated risks around the time of CABG.

1.2 Objectives

The clinical objective of this study is to quantitatively summarize the findings of available literature on CABG surgery. To achieve this goal, a systematic review and related meta-analysis were conducted. The results from our meta-analysis will help to diminish the conflict surrounding the preoperative aspirin therapy and the risk of increased peri-operative bleeding and other adverse events. These conclusions will provide a more conclusive finding to assist surgeons in making an evidence-based decision.

The statistical objective of this thesis is to compare the performance of Bayesian meta-analysis to classical meta-analysis on different data settings. The robustness of the Bayesian model was assessed by sensitivity analysis on different likelihood functions and prior distributions. The advantages and

limitations of Bayesian approach are discussed based on the comparison to classical meta-analysis. We also propose strategies to deal with problematic data such as zero-event studies and missing data. We hope the experience gained from this project will assist others conducting meta-analyses.

1.3 Scope of Study

In the following chapters, we will discuss the methodological issues of a systematic review, the methods of statistical analysis, the results, the strategies used to deal with problematic data, and, finally, our conclusions.

Specifically, in Chapter 2, we describe the steps used to conduct our systematic review and the characteristics of the data which we used to perform our meta-analysis.

Chapter 3 discusses the statistical methods of meta-analysis, including the random effect model using both Bayesian and classical approaches and the sensitivity analysis of the Bayesian models.

Chapter 4 compares the results obtained from the Bayesian approach and the classical approach. The assessment of the robustness of the Bayesian model is also discussed in this chapter.

Subsequently, in Chapter 5, we discuss the strategies used to deal with zero-events in binary outcomes, and propose a sensitivity analysis to assess the impact of missing values for continuous outcomes.

Finally, in the last chapter, we discuss the advantages and limitations of Bayesian meta-analysis by comparing it to classical meta-analysis.

Chapter 2

Methods of Literature Review

2.1 Literature Search Strategy

To conduct a reproducible literature review with minimized bias, we used a systematic review to identify all papers in the relevant literature on our pre-defined research questions. The primary question is whether giving preoperative aspirin increases blood transfusion and bleeding in coronary artery bypass grafting (CABG) compared to giving no aspirin. In the following section, we discuss the steps of conducting the literature search.

2.2 Sources of Literature Search

MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials, the Cochrane database of systematic reviews, the Cochrane database of abstracts and effects, and the Web of Science database were each searched for relevant articles from January 1975 to February 2007. The key words

included *aspirin, coronary artery bypass, acetylsalicylic acid* and *coronary artery bypass surgery.* Other terms used in search included *graft occlusion, vascular patency, graft patency,* and *antifibrinolytic agents.* All abstracts of published papers and technical reports were scanned and relevant articles identified using predefined inclusion and exclusion criteria.

2.3 Study Selection and Data Extraction

The study selection was based on the following pre-defined inclusion and exclusion criteria. First, the primary research objective had to be the use of aspirin preoperatively in coronary artery bypass surgery. Second, the patients must have been undergoing their first cardiac surgery. Third, for randomized controlled trials (RCTs), the primary endpoints must have included postoperative blood transfusion requirement or bleeding. In the case of non-RCTs (retrospective studies), blood transfusion requirement or bleeding must have been part of the primary analysis or a clearly defined secondary analysis.

All study abstracts which satisfied the inclusion criteria were highlighted, and read to ensure compliance with exclusion criteria. Any studies which did not satisfy the following exclusion criteria were omitted from the final analysis. First, studies were excluded if the treatment group received aspirin concurrently with other drugs. Second, non-RCTs were excluded if studies did not have a control group. Additionally, studies which included patients

who had previously undergone thoracotomy surgery were not included. Finally, studies involving only off-pump coronary bypass were excluded.

After relevant studies were identified, data were extracted by two independent reviewers using a customized data collection form. Subsequently, abstracted data were compared, and any disagreements between the two reviewers were resolved by consensus.

The qualities of the RCT studies were assessed by the two reviewers independently using the Jadad scale [9]. The Jadad scores of each RCT study were compared, and any disagreements were resolved by consensus.

2.4 Clinical Endpoints and Study Characteristics

Following study selection and data extraction, the data from nine RCT studies [10-18] and 14 non-RCT studies [19-32] were available for a meta-analysis. Since RCTs and retrospective observational studies (non-RCT) have completely different design characteristics [8], data were analyzed separately.

The treatment group was defined as those patients consuming aspirin (ASA) until the day of operation. The control group included those patients who received no-aspirin treatment (noASA) at least seven days prior to the operation [5]. Of the nine RCTs, two studies had two treatment groups, and of

the 14 non-RCTs, two studies had three treatment groups and one study had two treatment groups. To properly incorporate multi-treatment-group studies into final analysis, the control groups in those studies were evenly divided according into two or three groups according to the number of treatment groups [33].

The primary endpoint in our study was homologous blood transfusion requirement (measured as packed red blood cells unit). The secondary endpoints were the amount of peri-operative bleeding (measured as millilitre (ml)), event of peri-operative myocardial infarction (MI), event of peri-operative mortality, and event of re-operation for bleeding. Of these endpoints, the requirement for blood transfusion and the amount of bleeding were continuous variables, while MI, mortality and re-operation are counts.

All studies included in our analysis were published between 1978 and 2005. Of the 30 studies analyzed, two were non-English studies.

Detailed information regarding these studies is found in Table 1, Appendix B.

Chapter 3

Statistical Methods

3.1 Methods of Analysis

In this study, both classical (Frequentist) meta-analysis and Bayesian meta-analysis were adopted to estimate aspirin's treatment effects on five endpoints of the blood transfusion requirement, amount of bleed, myocardial infarction (MI), mortality and re-operation respectively. Eleven RCT studies and 19 non-RCT studies are analyzed separately. In the case of continuous endpoints, i.e. blood transfusion requirement and the amount of bleeding, the treatment effects were reported as absolute mean difference, and for the other three binary outcomes, the treatment effects were reported as odds ratio (OR). Besides the point estimates, we also reported the associated 95% confidence interval for the classical approach and the associated 95% credible interval for the Bayesian approach.

Heterogeneity (between study variance) is a big obstacle in meta-analysis for pooling data from different studies and for interpreting the overall estimates obtained from pooled data. Our approach to deal with heterogeneity was that instead of exploring it by using subgroup analysis, we incorporated it by adopting the random effects model in both the classical and Bayesian approaches. The main concern of avoiding subgroup analysis is the relatively small number of studies. For example, in the case of blood transfusion requirement in RCTs, there were only six studies available to be pooled together. Any result from subgroup analysis obtained by dividing such a small dataset into two or even more subsets might be driven solely by chance. The results of the primary analysis of the above five endpoints from the classical meta-analysis and from the Bayesian meta-analysis are compared in Chapter 3.

To assess the robustness of the Bayesian models, we conducted sensitivity analyses to test the impacts of different likelihood functions for continuous outcomes, and the impacts of different priors for binary outcomes. The results are presented in Chapter 3.

The software used to perform the classical meta-analysis was STATA 9.2 with meta-analysis package [34]. WinBUGS 14.1 (Windows Version of Bayesian Inference Using Gibbs Sampling) was used to conduct the Bayesian meta-analysis [35].

3.2 Classical Meta-Analysis with Random Effects Model

In the random effects model of classical meta-analysis, the treatment effect of each individual study is assumed to be an independent random variable pulled from a population with normal distribution

$$\widehat{\theta}_i \sim N(\theta, w_i^{-1} + \hat{\tau}^2),$$

Where w_i^{-1} is the random variance in each study, the so-called within-study variance, and $\hat{\tau}^2$ is the between-study variance, the so-called heterogeneity [36]. For the continuous outcomes, $\widehat{\theta}_i$ is the absolute mean difference, and for the binary outcome, $\widehat{\theta}_i$ is the odds ratio on logarithmic scale.

The overall estimate of treatment effect is obtained by calculating the weighted average of the treatment effect from each individual study. The weight of each study is given by the reciprocal of total variance (the so-called inverse variance method) in individual study, i.e. within-study variance plus between-study variance. When all studies are homogeneous, the between-study variance equals zero, and the random effects model will reduce to the fixed effects model.

In classical meta-analysis, the statistical heterogeneity is assessed by Cochran's Q-test

$$Q = \sum_{i=1}^{k} W_i(\hat{\theta}_i - \bar{\theta})^2 ,$$

where W_i is the weight of each study, $\bar{\theta}$ is the average of $\hat{\theta}_i$ and k is the number of studies. The test statistic Q is assumed to have a chi-squared distribution with degrees of freedom *(df)* equal to the number of studies minus one. The problem with this test is that when the number of studies is large, the heterogeneity will become significant by adding up very small between-study variances from each study [37-38].

Instead of testing homogeneity, a better approach is quantifying heterogeneity. The formula of calculation the quantity of heterogeneity is I^2=100% × *(Q-df)/Q* [37-38], where Q is Cochran's heterogeneity statistics, and *df* is the degrees of freedom. Since this quantity deducts the effect of the number of studies in calculating heterogeneity, it measures between-study variance more accurately. When the value of I^2 *is* less than or equal to zero, there is no heterogeneity observed, and the larger the value of I^2 the severe the heterogeneity.

The results of the overall estimates of treatment effects obtained from classical meta-analysis are discussed in Section 4.1, including the mean difference for continuous outcomes and the odds ratio for binary outcomes,

and associated 95% confidence intervals. The forest plots of the pooled studies are presented in Appendix A.

3.3 Bayesian Meta-Analysis with Random Effects Model

Unlike the frequentist inference which is used in the classical meta-analysis, the Bayesian inference uses the observed data as new information to update a researcher's pre-belief or external information. In a Bayesian frame, the observed data are presented as likelihood functions, the pre-beliefs or external information are presented as prior distributions, and the updated results are presented as posterior distributions. Therefore, in Bayesian inference, all parameters are treated as random variables.

The mathematical link between the researcher's pre-belief or external information, the observed data and the updated result are expressed as

$$p(\psi, w, \theta, \tau^2, \sigma \mid y) \propto f(y \mid \psi, w)p(\psi \mid \theta, \tau^2)p(w \mid \sigma)p(\tau^2 \mid \sigma)p(\sigma)$$

The vector of observed treatment effect y has a joint density function $f(y \mid \psi, w)$ which is a proportion of the likelihood function, with the corresponding parameter of mean ψ and of within-study variability w. The parameter ψ has a prior distribution $p(\psi \mid \theta, \tau^2)$ with the parameter of mean

θ and of the between-study variability τ^2. The within-study variability has a prior distribution $p(w \mid \sigma)$, and the between-study variability has a prior distribution $p(\tau^2 \mid \sigma)$. The parameter σ has a prior distribution $p(\sigma)$.

In our analysis, we were interested in the estimates of the mean of the treatment effect and the between-study variability which were obtained from posterior distribution $p(\psi, w, \theta, \tau^2, \sigma \mid y)$ directly.

In the Bayesian random effects model for primary analysis, we specified the likelihood function and the prior of parameter mean as normal distributions. The parameter σ had a uniform distribution. In WinBUGS, the normal distribution is parameterized as mean and precision, thus the variance of a normal distribution is defined as $1/\sigma^2$.

To make the results of the Bayesian approach comparable to the results of the classical approach, non-informative priors were chosen which minimized the influence of the researcher's pre-belief or external information on the observed data [36, 39].

In our Bayesian model, the prior used for the mean was N (0, 1.0E-6) and the prior for the precision was uniform (0, 10). The total number of iterations to obtain the posterior distribution for each endpoint was 500,000, and the burned-in number was 10,000. To reduce the influence of the autocorrelation

from previous iteration, we thinned our samples by five, i.e., every set of five samples generated one valid sample.

The reporting of the results is in accordance with the recommendations by Sung et al. [39].

The results from the Bayesian meta-analysis are presented in Section 4.1, which include estimates of the mean difference for continuous outcomes and the estimates of the odds ratio for binary outcomes, and the associated 95% credible intervals. To evaluate the convergence of the Markov Chain [39], plots of entire posterior distributions including dynamic trace plots, times series plot, density plots and autocorrelation plots are provided in Appendix C. The codes for running Bayesian models on WinBUGS along with the initials can be found in Appendix D.

3.4 Sensitivity Analysis of Likelihood Functions and Priors

Sometimes we might suspect that the observed data have heavy tails. In this case, the likelihood function of the observed data was specified as student's t distribution [40]. For our continuous outcomes, i.e. the blood transfusion requirement and the amount of bleeding, to test the impact of likelihood functions, we changed the normal distribution used in the primary

analysis to the student's t distributions with different degrees of freedom (df) equal to 4, 8, 16, and 32.

In the primary analysis, we followed the common practice by specifying the priors of variance as uniform distribution with the upper bound parameter equal to 10 for the binary data, i.e. MI, mortality and re-operation. However, the robustness of this specification needs to be assessed [41]. We evaluated the impact of different priors by giving the upper bound parameter of uniform distribution different values used as 1, 5, 10, 25, 50 and 100.

The comparisons of the results obtained from different likelihood functions and priors are discussed the Section 4.2 and Section 4.3.

Chapter 4

Results

4.1 Results of Primary Analysis

The overall estimates obtained from the classical approach and the Bayesian approach from 11 RCT studies were similar except for re-operation. The results from both approaches reported that no significant differences on treatment effect were detected between the treatment group (ASA group) and the control group (noASA group) for blood transfusion requirement, the amount of bleeding, MI and mortality. The only disagreement between the classical approach and the Bayesian approach happened on re-operation. The classical approach reported that the re-operation rate of the ASA group was significantly higher than the re-operation rate in the noASA group as OR = 2.52 (1.18, 5.38). However, the Bayesian approach reported that there was no significant difference on treatment effect of re-operation as OR = 2.78

(0.96, 8.86). In our case, the result from Bayesian meta-analysis was more meaningful by comparing it to the difference of re-operation rate in non-RCT studies. In non-RCTs, the results from both the classical approach and the Bayesian approach agreed that there was no significant difference in the re-operation rate between the treatment group and the control group (classical: OR = 1.12 (0.69, 1.83); Bayesian: OR = 1.20 (0.63, 2.39)). Since re-operation is a very objective measure which is little influenced by confounders and other bias factors, the result from RCT studies should be similar to the result from non-RCT studies.

		Classical Method		Bayesian Method	
RCT	(Continuous Outcome)	WMD	C.I.	Mean	Cr.I.
	Blood Transfusion	0.36	(-0.37, 1.08)	0.45	(-1.11, 1.97)
	Bleeding	89.75	(-17.85, 179.43)	99.00	(-46.37, 227.60)
	(Binary Outcome)	OR	C.I.	OR	Cr.I.
	MI	1.04	(0.35, 3.07)	1.04	(0.05, 24.05)
	Mortality	1.21	(0.31, 4.71)	1.23	(0.18, 8.18)
	Re-operation	2.52	(1.18, 5.38)	2.78	(0.96, 8.86)
Non-RCT	(Continuous Outcome)	WMD	C.I.	Mean	Cr.I.
	Blood Transfusion	0.34	(0.12. 0.56)	0.36	(0.06, 0.66)
	Bleeding	131.10	(61.72, 200.47)	134.00	(7.26, 240.60)
	(Binary Outcome)	OR	C.I.	OR	
	*MI	1.29	(0.02, 68.92)	NA	
	Mortality	0.59	(0.34, 1.02)	0.91	(0.33, 3.59)
	Mortality without Bybee's	1.39	(0.55, 3.53)	1.53	(0.43, 6.21)
	Re-operation	1.12	(0.69, 1.83)	1.20	(0.63, 2.39)

Table 2: Overall estimates of primary analysis from classical meta-analysis and from Bayesian meta-analysis (*only one study available)

We also noticed that the credible interval of MI was unrealistically wide (OR = 1.04 (0.05, 24.05)) in the Bayesian approach. This problem is discussed in Section 4.2 as part of the results of the sensitivity analysis

The overall estimates of treatment effect on five endpoints in 19 non-RCT studies obtained from the classical approach and the Bayesian approach were similar. The significant difference between the treatment group (ASA group) and the control group (noASA group) were detected on the blood transfusion requirement and the amount of bleeding. For mortality and re-operation (the Bayesian approach on MI could not be conducted due to insufficient data), no significant differences of treatment effect were found.

In the overall estimates of non-RCT studies, we noticed that the point estimate of mortality from the classical approach and the Bayesian approach were quite different in magnitude ((classical: OR = 0.59 (0.34, 1.02)); Bayesian: OR = 0.91 (0.33, 3.59)). From the forest plot, we found that the data from Bybee's study [32] might be an outlier which behaved very differently from other studies. By excluding Bybee's studies, the estimates of treatment effect from the classical and the Bayesian approaches became similar (classical: OR = 1.39 (0.55, 3.53): Bayesian: OR = 1.53 (0.43, 6.21)).

Overall, the point estimates from both the classical and the Bayesian approaches were more similar than the interval estimates. The credible

intervals from the Bayesian approach were wider than the confidence intervals from the classical approach, and the smaller the number of studies the wider the interval. The main reason for the difference was that the Bayesian approach incorporated all kinds of variability from all parameters. The difference and similarity between classical meta-analysis and Bayesian meta-analysis are discussed in the last chapter.

In the results of the classical meta-analysis, significant heterogeneities were found in the blood transfusion requirement (in RCTs: $p = 0.01$, $I^2 = 66.9\%$; in non-RCTs: $p < 0.001$, $I^2 = 75.0\%$) and the amount of bleeding (in RCTs: $p = 0.006$, $I^2 = 61.2\%$; in non-RCT: $p = 0.001$, $I^2 = 68.4\%$). The reason is that these two endpoints are more subjective so that the variation among independent studies easily happens.

4.2 Impact of Likelihood Functions

To examine the influence of likelihood functions in Bayesian meta-analysis, we also specified the likelihood functions as the Student's t distribution which has heavier tails than the normal distribution, with the degrees of freedom equal to 4, 8, 16, and 32 respectively for continuous endpoints, i.e. blood transfusion requirement and bleeding.

The results of the blood transfusion requirement in RCTs showed that the estimates of odds ratios and credible intervals from the Student's t likelihood

function converged with those estimates from the normal likelihood function gradually while the degrees of freedom of the student's t increases. For example, the overall estimates from the student's t likelihood function of the treatment effect on the blood transfusion requirement in RCT studies were OR = 0.38 (-1.06, 1.88) (df = 4), OR = 0.42 (-1.09, 1.95) (df = 8), OR = 0.44 (-1.08, 1.96) (df = 16), and OR = 0.44 (-1.10, 1.97) (df = 32). When the degrees of freedom reached 32, the result was very similar to the result from the normal likelihood distribution which had OR = 0.45 (-1.11, 1.97). The detailed information for this comparison can be found in Appendix B, Table 3, and a visual plot of this comparison can be found in Appendix A, Figure 2.

The sensitivity analysis about different likelihood functions proved that our specification of the likelihood function as a normal distribution was proper. However, when the observed data belong to the student's t distributions, the choice of the degrees of freedom will influence the results.

4.3 Impact of Priors of Variance Parameter

In Bayesian meta-analysis, the uniform prior distribution is a preferred choice of non-informative prior for variance parameters [41]. It works properly for most cases, and generates estimates similar to those from the classical approach. However, when the number of studies is small, this choice will lead to an extremely large variance and an unrealistically broad credible

interval [41]. In this situation, a restriction of the upper bound parameter for uniform distribution has to be applied.

To examine the relationship between the number of studies and the uniform prior, we changed the upper bound parameter of the uniform prior from small to large by setting it equal to 1, 5, 25, 50, and 100 respectively. These five different priors were applied on binary outcomes in which the number of studies for those outcomes varied from 3 to 7 to 11. The results from the above priors were compared to the results from uniform (0, 10), which is a common choice for binary endpoints and which was used in our primary analysis.

The results of the above sensitivity analysis showed that when the number of studies is small, three in our example, the length of the credible interval increased dramatically while the upper bound parameter of the uniform prior increased. When the number of studies is moderate, seven in our example, the credible interval became stable when the upper bound parameter of the uniform prior was equal to or greater than 10. When the number of studies is large, eleven in our example, all choices of the upper bound parameter for the uniform priors produced similar results. Therefore, uniform (0, 10) is a proper choice as a non-informative prior for most binary endpoints with number of studies from moderate to large. But when the number of studies is small, according to Gelman's suggestion in his paper of the simulation study, we

have to use the empirical prior, i.e. the prior needs to be estimated from observed data [41].

The results of this comparison are presented in Appendix B, Table 4, and a visual plot of this comparison can be found in Appendix A, Figure 3.

Chapter 5

Dealing with Problematic Data

5.1 Problems of Missing Data and Zero-events in Meta-analysis

Meta-analysis uses summary statistics extracted from limited studies as input data to calculate the overall estimate. If summarized data from several valid studies cannot be incorporated into the meta-analysis for some reason, we not only face the problem of low statistical power, but also lose the integrity of the evidence in the literature.

There are two kinds of problematic data which are difficult or impossible to incorporate into meta-analysis. One is missing data and the other is zero-event studies. Though the standard software for meta-analysis does not fully support these problems, other statistical techniques can be borrowed to either

solve the problems or evaluate the impact of these problems. In the following two sections, we discuss some common approaches in the literature and new attempts used to deal with these problems. The results from those approaches are compared to the results from the primary analysis to evaluate the influence of importing these techniques.

5.2 Dealing with Zero-events Studies

Mortality or occurrence of rare events is very often used as an outcome to monitor the treatment effect in clinical research. Sometimes the event rate is so low that we do not have enough observations, which leads to lack of power to detect the treatment effect. The worst case scenario is that we may not observe any event in a single study. Therefore meta-analysis can be used as an efficient tool to increase the statistical power by combining the single insufficient studies. To increase the power, we should pool all available data into the analysis, but it has been argued whether including those zero-event studies into meta-analysis is meaningful [42]. For those either-arm zero-event studies, i.e., the zero-event happened either in the treatment group or in the control group, the consensus of opinion in the literature is that those studies should be included. However, there is no common method for incorporating these data [42-43]. For both-arm zero-event studies, i.e. no observation in both the treatment and the control groups, the majority of researchers think

that those studies should be excluded from the analysis. The reason for excluding both-arm zero-event studies is that those studies do not make any contribution to the overall estimates. One paper published in 2007 has suggested that including both-arm zero events would reduce the risk of over-estimated treatment effect and narrow the confidence interval for classical meta-analysis [44].

In our study, we compared the results of excluding all kinds of zero-event studies, including only either-arm zero-event studies, and including all zero-event studies by using classical meta-analysis and Bayesian meta-analysis. In classical meta-analysis, we used continuity correction (correction factor 0.5) with the inverse variance (IV) method and the Peto method [42-43] to incorporate either-arm zero-event studies. To incorporate two-arm zero-event studies, we adopted continuity correction into IV method by manually adding 0.5 to each cell in both-arm zero-event studies. This is the approach we used in our primary analysis to include all kinds of zero-event studies into the final data analysis.

The endpoint we chose to illustrate the different approaches for dealing with zero-event studies was re-operation in RCT studies. Of the seven available studies, three were either-arm zero-event studies, and one was both-arm zero-event study.

The results showed that by including all kinds of zero-event studies, both the confidence intervals from classical meta-analysis and the credible intervals from Bayesian meta-analysis were narrowed, particularly for the Bayesian approach. The results of the classical approach showed a substantial change before and after including zero-event studies. The result of excluding any kind of zero-event studies reported that no significant treatment effect was detected with OR = 2.33 (0.99, 5.48). However, after including either-arm zero-event studies in analysis, the result showed that there was a significant difference between treatment group and control group with OR = 2.68 (1.34, 6.10) from the IV method and with OR = 2.90 (1.44, 5.82) from the Peto method. The reason for this change is that including zero-event studies increased sample size and thus increased the statistical power to detect the difference between treatment group and control group. We also found that, compared to the results of including only either-arm zero-event studies, adding both-arm zero-event studies in the analysis can pull the estimates of treatment effect closer to the null hypothesis to produce a more conservative estimate.

Another benefit gained from including all zero-event studies into the analysis was reduced publication bias. Publication bias is caused by lack of small studies [45], and zero-events often happen in studies having a small number of patients.

The numerical summary of this comparison is presented in Appendix B, Table 5, and a visual plot of this comparison can be found in Appendix A, Figure 4.

5.3 Dealing with Missing Data

Most of the data used in meta-analysis are extracted from separate and independent papers in the literature. Since the process of conducting the analyses and the format of the reporting of results varies from study to study, we very often cannot find all the needed data from all the studies. For example, in our database, every endpoint had different levels of missing data.

In the literature, almost all meta-analyses automatically deleted the studies with missing data from the final analysis without any further exploration. This strategy seems to be the only choice, but ignoring missing data leads to smaller sample size and the loss of integrity of evidence. Firstly, after following a strict selection procedure, the number of valid studies is very limited. Any missing data might make the total sample size too small to detect the statistical significance of the treatment effect. Thus, one of the major advantages of meta-analysis is lost. Secondly, even if the sample size is large enough to detect the significance, the overall conclusion based on partial literature is still susceptible. Readers might doubt what will happen if all missed studies do not favour this conclusion. Furthermore, an important data

handling principle in RCTs is the intention-to-treat analysis (ITT) [8]. This principle states that all data of patients who have been randomized into the trials should be included in the final analysis. Otherwise the result will be biased in favour of the treatment group.

In this study, the proportions of missing data were very large. For example, in the primary endpoint, blood transfusion requirement, five out of 11 RCT studies and five out of 19 non-RCT studies could not be pooled into the final analysis due to missing data. To evaluate the impact of missing data on the final result for this endpoint, we imputed ten datasets simultaneously by using the multiple imputation method (MI) [46]. The patient demographics and clinically related information were used to fill in the missing data in the blood transfusion requirement. They were study design type, number of patient, patient average age, proportion of males, average number of graft, aspirin dose, pumping time of bleeding, and the amount of bleeding. The seed of the imputation was 84446.

After combining the results for the ten imputed datasets, the overall estimates of the difference of blood transfusion requirement in RCT studies (0.50 (0.02, 0.97)) was significant. However, the result from our primary analysis reported that there was no significant difference between the treatment group and the control group (0.36 (-0.37, 1.08)). Comparing the dramatically increased sample size (58%) by applying imputation, the original

sample size might have been too small to have enough power to detect the difference. Therefore, we need to interpret the results of the primary analysis with caution. The results based on imputation for non-RCT studies (0.47 (0.16, 0.77)) was similar to the results from primary analysis (0.34 (0.12, 0.56)), thus from the statistical point of view, the results of non-RCT studies from the primary analysis were more reliable than the result of RCT studies.

The detailed comparison of the overall estimate of blood transfusion before and after the imputation can be found in Appendix B, Table 6.

Chapter 6

Discussion and Conclusions

6.1 Difference between Classical and Bayesian Meta-analysis

Meta-analysis is an effective tool for quantitatively synthesizing overall evidence by pooling information from all qualified independent studies in the literature. Though classical meta-analysis and Bayesian meta-analysis have the same goal, they have different philosophies on statistical inference.

Classical meta-analysis belongs to frequentist inference, in which point estimates, confidence intervals and hypothesis tests are based on the assumption of infinite but identical repetitions on the fixed but unknown parameters [36]. This process is more deductive, and estimates of parameters are summarized from the observations directly.

For Bayesian inference, all unknown parameters are treated as random variables. The estimates of treatment effect and between-study variance which we are interested in are directly obtained from their posterior distributions. The Bayesian approach works more inductively, in the way that a researcher's pre-beliefs (or external information) are expressed as prior distribution functions; these beliefs then are updated by incorporating the likelihood function of observed data. The updated beliefs form the posterior distribution functions, and the point estimates and credible intervals are directly estimated from the density of posterior distributions.

Despite the above philosophical differences between the frequentist inference and the Bayesian inference, in practice, by using non-informative priors in the Bayesian model, these two different methods often lead to similar numerical results.

6.2 Advantages of Bayesian Meta-analysis

The availability of well-developed statistical software on Bayesian inference such as WinBUGS, and powerful computers has greatly increased the use of the Bayesian approach for meta-analysis. However, the debate regarding the advantages and limitations of Bayesian meta-analysis versus classical meta-analysis continues [53]. In our study, we compared the process

and the results between Bayesian meta-analysis and classical meta-analysis, and found that Bayesian analysis has many advantages.

One of the major advantages of Bayesian approach is the ability to incorporate all kinds of variability. In general, this feature leads to credible intervals wider than the confidence intervals from the classical approach, thus providing readers with more conservative evidence.

Another attractive point of the Bayesian approach is that it allows researchers to make probability statements on the results. For example, the probability of the correctness of a 95% credible interval is 0.95.

Furthermore, the Bayesian approach provides researchers the flexibility to choose different priors and likelihood functions. The prior distribution can express a researcher's pre-beliefs more intuitively and also allows a researcher to construct the prior distribution by borrowing the evidence from external sources. Unlike the common used method for classical meta-analysis where the data are always assumed to have a normal distribution, the Bayesian approach allows a researcher to specify the likelihood function of observed data as other distributions, e.g. a student's t distribution when the data were suspected having heavy tails. When the number of observations is large, the data will approximate to a normal distribution; when number of observation is small, the Bayesian model is more flexible to generate more accurate estimates.

In this study, we found an advantage of the Bayesian model which is ignored in most research papers. Compared to the classical approach, the Bayesian model is relatively robust to outliers. Since the weight of the individual study is given by the reciprocal of the total variance in classical meta-analysis, when a large study happens to have an outlier, the estimates of overall effect will be driven by this heavily weighted study. For example, Bybee's study contributed 65.1% weight, and it was the only study which claimed that the control group had higher mortality (OR = 0.37 (0.19, 0.73)). When we pooled this study with other non-RCT studies by performing classical meta-analysis, its large weight might have been the reason which caused the estimates of odds ratio of mortality in non-RCT studies (OR = 0.59 (0.34, 1.02)) to differ in direction and magnitude from the results of RCT studies (OR = 1.21 (0.31, 4.71)). However, when we compared the estimates of odds ratios of mortality in RCT studies and non-RCT studies generated by the Bayesian approach, the difference (RCTs: OR = 1.23 (0.18, 8.18); non-RCTs: OR = 0.91 (0.33, 3.59)) between the two study designs was much smaller than the difference in the classical approach. Since mortality is a very objective measurement, the event rate in RCT studies and non-RCT studies should be similar, thus the result from the Bayesian meta-analysis is more reasonable. The reason for the Bayesian approach's outlier-free (relatively) character is that the Bayesian model captures total variability

from all parameters, so the influence of sample size is reduced. Actually, when we excluded Bybee's study from our analysis, the estimates of odds ratio of mortality in RCTs from the classical approach (1.39 (0.55, 3.53)) and the results from the Bayesian approach (1.53 (0.43, 6.21)) became similar.

6.3 Limitations of Bayesian Meta-analysis

The ability to capture all variability and interpret the results on an exact probability makes Bayesian meta-analysis attractive to many researchers [47]. However, the Bayesian approach has its own limitations.

First, the results from the Bayesian approach are somehow dependent on pre-defined priors, which express a researcher's subjective beliefs (or external information). Though the influence of priors can be maximally diminished by using non-informative priors, non-informative priors do not work properly when the number of studies is small, particularly for the prior of variance. When this situation happens, the priors have to be estimated from observed data, thus the advantage of the full Bayesian model is partially lost.

Probability distributions have to be specified for all parameters and the likelihood functions for observed data. When the number of studies is small, the misspecification of priors or likelihood functions has some impact on the posteriors. For example, for the blood transfusion requirement in RCT studies, the student's t likelihood function with a small degree of freedom yielded

different point estimates compared to the normal likelihood function. Though the significance of results did not change in this case, the different magnitude of point estimation might influence the clinical interpretation.

In the Bayesian frame, the posterior distribution of one parameter is requires integration over all other parameters. When we have many parameters in the model, the integral on a high dimension parameter-space might cause problems. Gibbs sampling within the Markov Chain Monte Carlo (MCMC) simulation method used in WinBUGS simplifies the high dimensional integral to sampling the posterior conditional distributions over a sufficiently long time period [35]. However, for some complex models, the convergence of Markov Chain is sensitive to the specification of priors and initial values [48].

Another unavoidable disadvantage is the computation burden for the Bayesian model. If one uses state-of-the-art computers, this is not always a problem, but when the number of parameter or the number of studies is large, using the Bayesian method is still a time-consuming process. For example, the Bayesian model used for binary endpoints in our studies needed at least twice the time of than the model for continuous endpoints.

6.4 Choice of Analysis Methods for Small Datasets

The data used to perform meta-analysis were extracted from separate studies, and the studies were skimmed using strictly defined inclusion and exclusion criteria. Very often, the number of valid studies is small. When we faced a small dataset, we had to carefully choose the proper method to carry out the analysis, and interpret the results with caution, i.e. verifying the results by finding the coincidence across methods or finding the clinical rationales.

Firstly, when number of studies was small, the better choice for dealing with heterogeneity rather than explaining it by using subgroup analysis was to incorporate it by using a random effects model. Further dividing the dataset into smaller subsets might cause the loss of statistical power to detect statistical significance. Even if the result showed a significant difference between groups, the results might be caused only by random chance.

Secondly, for small datasets, the classical approach and the Bayesian approach might lead to different results. The results from the classical meta-analysis can be overly optimistic by assuming that the observed data can represent the whole population. On the other hand, the unrealistically broad credible interval from the Bayesian approach provides little information for

decision making. When such disagreement rises, we should not make any quick conclusions based on the result from either method.

In Chapter 4 we discussed the strategies of increasing sample size by including zero-event studies for binary outcomes. We also proposed a sensitivity analysis for continuous outcomes by using the multiple imputation method to evaluate the impact of the decreased sample size due to missing values.

6.5 Conclusions

In general, our results from the Bayesian approach agreed with those from the classical meta-analysis. The only exception was the odds ratio in RCT studies.

For the blood transfusion requirement and the amount of bleeding, the disagreement between RCT studies and non-RCT studies made it difficult for us to make the final conclusion. The results of the RCT studies showed that there was no significant difference between the aspirin group and the no-aspirin group. Usually the results from RCT studies are more reliable than the results from non-RCT studies since RCT design is considered relatively bias free. However, the small number of RCT studies, e.g. six studies of blood transfusion requirement, was questionable. We doubted that we had enough statistical power to detect the significance of the treatment effect, and the

results of the sensitivity analysis using multiple imputation methods provided further evidence to confirm this. On the other hand, non-RCT designs, known as confounders, and bias-associated design, were likely to provide over-estimated results favouring the treatment group [7]. Therefore, we could not make our final conclusion based on the results of the non-RCT studies.

In the case of mortality, we were more confident in concluding that there was significant difference between the aspirin group and the no-aspirin group.

For myocardial infarction (MI), all available results showed that the treatment effect of aspirin was not significant. However, considering how small the sample sizes were and how much wider the confidence or credible intervals were, we could not make any conclusions based on this study without caution.

For re-operation, most results showed that the re-operation rate of the aspirin patients was not higher than the no-aspirin group, which went against the results in RCTs from the classical meta-analysis. However, if we notice that the lower limit of the credible interval from Bayesian meta-analysis in RCT studies was just across 1 (OR = 2.78 (0.96, 8.86)), the certainty of these results is questionable.

Incorporating all available evidence, we suggest that discontinuing aspirin seven days before coronary artery bypass grafting surgery is a safer choice.

The comparisons between Bayesian meta-analysis and classical analysis cannot prove that one method is superior to another. Our conclusion from a statistical aspect is that to conduct a proper meta-analysis, we should apply both the Bayesian approach and the classical approach. Any final conclusions made should be based on the comparison of the results from these two approaches.

Bibliography

1. Willard JE, Lange RA, Hillis LD. The use of aspirin in ischemic heart disease. *New England Journal of Medicine* 1992; **327**:175-181.
2. America Heart Association Website http://www.americanheart.org/, accessed 02-17-2007.
3. Dacey LJ, Munoz JJ, Johnson ER, et al. Effect of Preoperative Aspirin Use on Mortality in Coronary Artery Bypass Grafting Patients. *The Annals of Thoracic Surgery* 2000: **70**:1986-1990.
4. Chesebro JH, Clements IP, Fuster V, et al. A platelet-inhibitor-drug trial in coronary-artery bypass operations: benefit of perioperative dipyridamole and aspirin therapy on early postoperative vein-graft patency. *New England Journal of Medicine* 1982; **307**:73-78.
5. Eagle KA, Guyton RA, Davidoff R, et al. ACC/AHA Guidelines for Coronary Artery Bypass Graft Surgery: Executive Summary and Recommendations. *Circulation* 1999; **100**:1461-1480.
6. Sun JC, Crowther MA, Warkentin TE, et al. Should Aspirin Be Discontinued Before Artery Bypass Surgery? *Circulation* 2005; **112**:e85-e90.

7. Guyatt GH, Rennie D. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Chicago, Illinois: American Medical Association, 2002.
8. Haynes RB, Sackett DL, Guyatt GH, Tugwell P. Clinical Epidemiology: How to do Clinical Practice Research, 3rd Ed. New York NY: Lippincott Williams & Wilkins, 2005.
9. Moher D, Jones A, Lepage L. Use of the CONSORT Statement and Quality of Reports of Randomized Trials. *Journal of the American Medical Association* 2001; **285**:1992-1995.
10. Fuller JK, Copeland JG. Does Short-Term Pre-Operative Aspirin in Coronary Bypass Patients Increase Post-Operative Bleeding? *Vascular Surgery* 1985; **19**:174-178.
11. Karwander SV, Weksler BB, Gay WA, et al. Effect of Preoperative Antiplatelet Drugs on Vascular Prostacyclin Synthesis. *The Annals of Thoracic Surgery* 1987; **43**:318-322.
12. Ferraris VA, Ferraris SP, Lough FC, et al. Preoperative Aspirin Ingestion Increases Operative Blood Loss after Coronary Artery Bypass Grafting. *The Annals of Thoracic Surgery* 1988; **45**:71-74.
13. Goldman S, Copeland J, Moritz T, et al. Starting Aspirin Therapy After Operation–Effects on Early Graft Patency. *Circulation* 1991; **84**:250-256.

14. Hockings BE, Ireland MA, Goth-Martin KF, et al. Placebo-Controlled Trial of Enteric Coated Aspirin in Coronary Bypass Graft Patients. *The Medical Journal of Australia* 1993; **159**:376-378.

15. Kallis P, Tooze JA, Talbot S, et al. Pre-operative Aspirin Decreases Platelet Aggregation and Increases Post-operative Blood Loss – A Prospective, Randomized, Placebo Controlled, Double-Blind Clinical Trial in 100 Patients with Chronic Stable Angina. *European Journal of Cardio-Thoracic Surgery* 1994; **8**:404-409.

16. Matsuzaki K, Okabe H, Kajihara N, et al. A Prospective Study on the Timing of Discontinuation of Aspirin before Coronary Artery Bypass Grafting. *Journal of Japanese Annals of Thoracic Surgery* 1997; **45**:1710-1714.

17. Klein M, Keith PR, Dauben HP, et al. Aprotinin Counterbalances an Increased Risk of Peri-operative Hemorrhage in CABG Patients Pre-treated with Aspirin. *European Journal of Cardio-thoracic Surgery* 1998; **14**:360-366.

18. Morawski W, Sanak M, Cisowski M, et al. Prediction of the Excessive Perioperative Bleeding in Patients Undergoing Coronary Artery Bypass Grafting: Role of Aspirin and Platelet Glycoprotein IIIa Polymorphism. *Surgery for Acquired Cardiovascular Disease* 2005; **130**:791-796.

19. Michelson EL, Morganroth J, Torosian M, el at. Relation of Preoperative Use of Aspirin to Increased Mediastinal Blood Loss after Coronary Artery Bypass Graft Surgery. *The Journal of Thoracic and Cardiovascular Surgery* 1978; **76**:694-697.

20. Torosian M, Michelson EL, Morganroth J, et al. Aspirin and Coumadin Related Bleeding After Coronary Artery Bypass Graft Surgery. *Annals of Internal Medicine* 1978; **89**:325-328.

21. Taggart DP, Siddiqui A, Wheatley DJ, et al. Low-Dose Preoperative Aspirin Therapy, Postoperative Blood Loss, and Transfusion Requirements. *The Annals of Thoracic Surgery* 1990; **50**:425-428.

22. Rawitscher RE, Jones JW, McCoy TA, et al. A Prospective Study of Aspirin's Effect on Red Blood Cell Loss in Cardiac Surgery. *The Journal of Cardiovascular Surgery* 1991; **32**:1-7.

23. Reich DV, Patel GC, Vela-Cantos F, et al. Aspirin Does Not Increase Homologous Blood Requirements in Elective Coronary Bypass Surgery. *Cardiovascular Anesthesia* 1994; **79**:4-8.

24. Vuylsteke A, Oduro A, Cardan E, et al. Effect of Aspirin in Coronary Artery Bypass Crafting. *Journal of Cardiothoracic and Vascular Anesthesia* 1997; **11**: 831-834.

25. Jakics J, Lee J, Ikeda S. Preoperative Aspirin and Heparin Therapy Does Not Increase Perioperative Blood Loss and Blood Product Requirements in

Coronary Artery Bypass Graft Surgery. *Journal of Anesthesia* 1999; **13**:8-13.

26. Chavanon O, Durand M, Romain-Sorin V, et al. L Does the Time Between Preoperative Interruption of Aspirin Intake and Operation Influence Postoperative Blood Loss and Transfusion Requirement in Coronary Artery Bypass Graft? *Annales françaises d'anesthèsie et de réanimation* 2002; **21**:458-463.
27. Weightman WM, Gibbs NM, Weidmann CR, et al. The Effect of Preoperative Aspirin-Free Interval on Red Blood Cell Transfusion Requirements in Cardiac Surgical Patients. *Journal of Cardiothoracic and Vascular Anesthesia* 2002; **16**:54-58.
28. Gerrah R, Izhar U. Beneficial Effect of Aspirin on Renal Function Post-Cardiopulmonary Bypass. *Asian Cardiovascular and Thoracic Annals* 2003; **11**:304-308.
29. Ray JG, Deniz S, Olivieri A, et al. Increased Blood Product Use among Coronary Artery Bypass Patients Prescribed Preoperative Aspirin and Clopidogrel. *BMC Cardiovascular Disorder* 2003; **3**:1-6.
30. Gerrah R, Ehrlich S, Tshori S, et al. Beneficial Effect of Aspirin on Renal Function in Patients with Renal Insufficiency Postcardiac Surgery. *The Journal of Cardiovascular Surgery* 2004; **45**: 545-50.

31. Gerrah Babin, Elami A, Stamler A, et al. Preoperative Aspirin Administration Improves Oxygenation in Patients Undergoing Coronary Artery Bypass Grafting. *CHEST* 2005; **127**:1622-1626.

32. Bybee KA, Powell BD, Valeti U, et al. Preoperative Aspirin Therapy Is Associated with Improved Postoperative Outcomes in Patients Undergoing Coronary Artery Bypass Grafting. *Circulation* 2005; **112**:286-292.

33. http://www.cochrane-net.org/openlearning/HTML/modA2-5.htm, accessed 2007-02-17.

34. Sterne JA, Bradburn MJ, Egger M. Systematic Reviews in Health Care: Meta-Analysis in Context, 2nd Edition: Chapter 18 Meta-analysis in Stata. London, England: British Medical Journal Publishing Group, 2001.

35. http://www.mrcbsu.cam.ac.uk/bugs/winbugs/contents.shtml, accessed 2006-09-16.

36. Whitehead A. Meta-Analysis of Controlled Clinical Trials. West Sussex, England: John Wiley and Sons, 2002.

37. Higgins JP, Thompson SG. Quantifying Heterogeneity in a Meta-Analysis. *Statistics in Medicine* 2002; **21**:1539-1558.

38. Higgins JO, Thompson SG, Deeks JJ, et al. Measuring Inconsistency in Meta-Analyses. *British Medical Journal* 2003; **327**:557-560.

39. Sung L, Hayden J, Greenberg ML, et al. Seven Items Were Identified for Inclusion When Reporting a Bayesian Analysis of a Clinical study. *Journal of Clinical Epidemiology* 2005; **58**: 261-268.

40. Angers JF. Use of Student-t Prior for the Estimation of Normal Means: A Computational Approach. *Bayesian Statistics* 1992; 4:567-575.

41. Gelman A. Prior Distributions for Variance Parameters in Hierarchical Models. *Bayesian Analysis* 2006; **1**: 515-533.

42. Sweeting MJ, Sutton AJ, Lambert SP. What to Add to Nothing? Use and Avoidance of Continuity Correction in Meta-analysis of Sparse Data. *Statistics in Medicine.* 2004; **23**:1351-1375.

43. Bradburn MJ, Deeks JJ, Belin JS, et al. Much Ado about Nothing: A Comparison of the Performance of Meta-Analytical Methods with Rare Events. *Statistics in Medicine.* 2007; **26**:53-77.

44. Friedrich JO, Adhikari NK, Beyene J. Inclusion of Zero Total Event Trials in Meta-Analyses Maintains Analytic Consistency and Incorporates All Available Data. *BMC Medical Research Methodology* 2007; **7**:5-10.

45. Sterne JA, Gavaghan D, Egger M. Publication and Related Bias in Meta-Analysis: Power of Statistical Tests and Prevalence in the Literature. *Journal of Clinical Epidemiology* 2000; **53**:1119-1129.

46. Yuan YC. Multiple Imputation for Missing Data, Concepts and New Development: SAS Documentation Experimental MI. http://support.sas.com/rnd/app/da/new/dami.html, accessed 2006-12-15

47. Bloom BS, de Pouvourville N, Libert S. Classical or Bayesian Research Design and Analysis: Does it Makes a Difference? *International Journal of Technology Assessment in Health Care* 2002; **18**: 120-126.

48. Sutton AJ, Abrams KR. Bayesian Methods in Meta-Analysis and Evidence Synthesis. *Statistical Methods in Medical Research* 2001; **10**:277-303.

Appendix A

Forest Plots of the Primary Analysis and Comparison Plots of Sensitivity Analysis

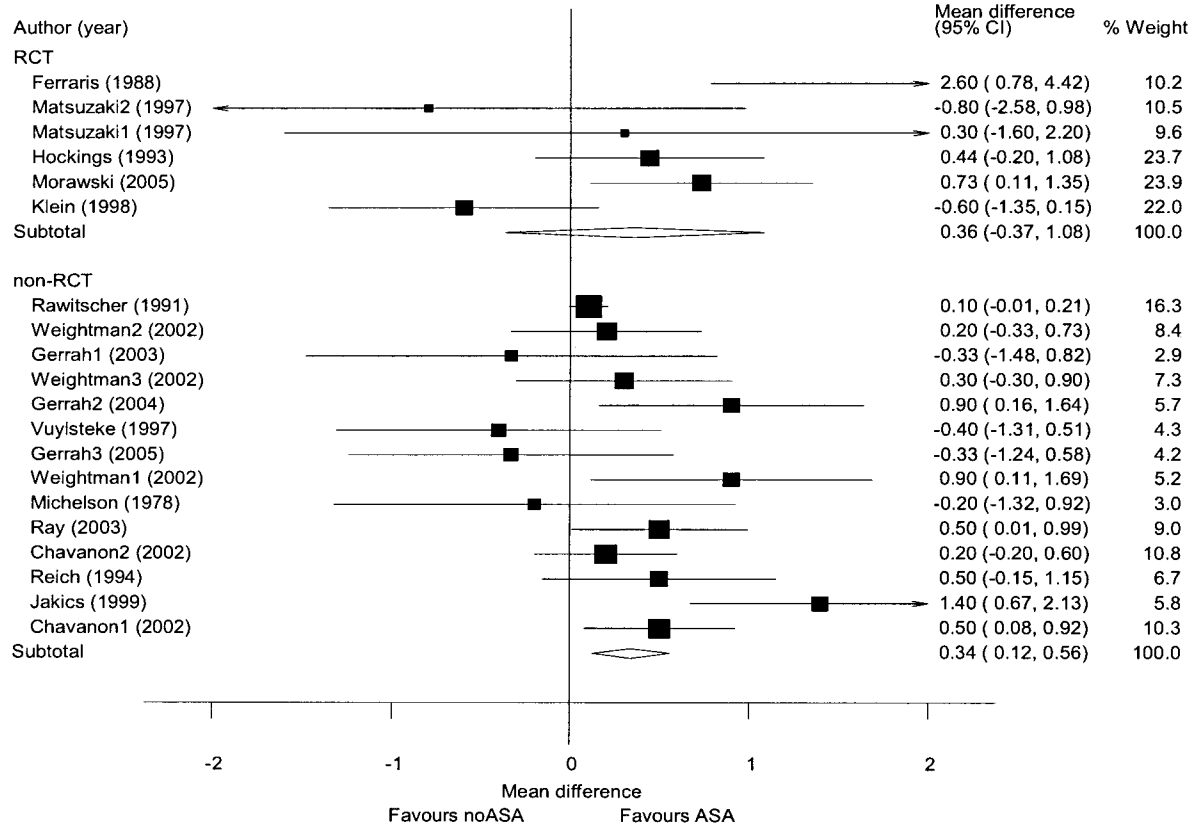


Figure 1a Classical Random Effects Model of Pooled Blood Transfusion

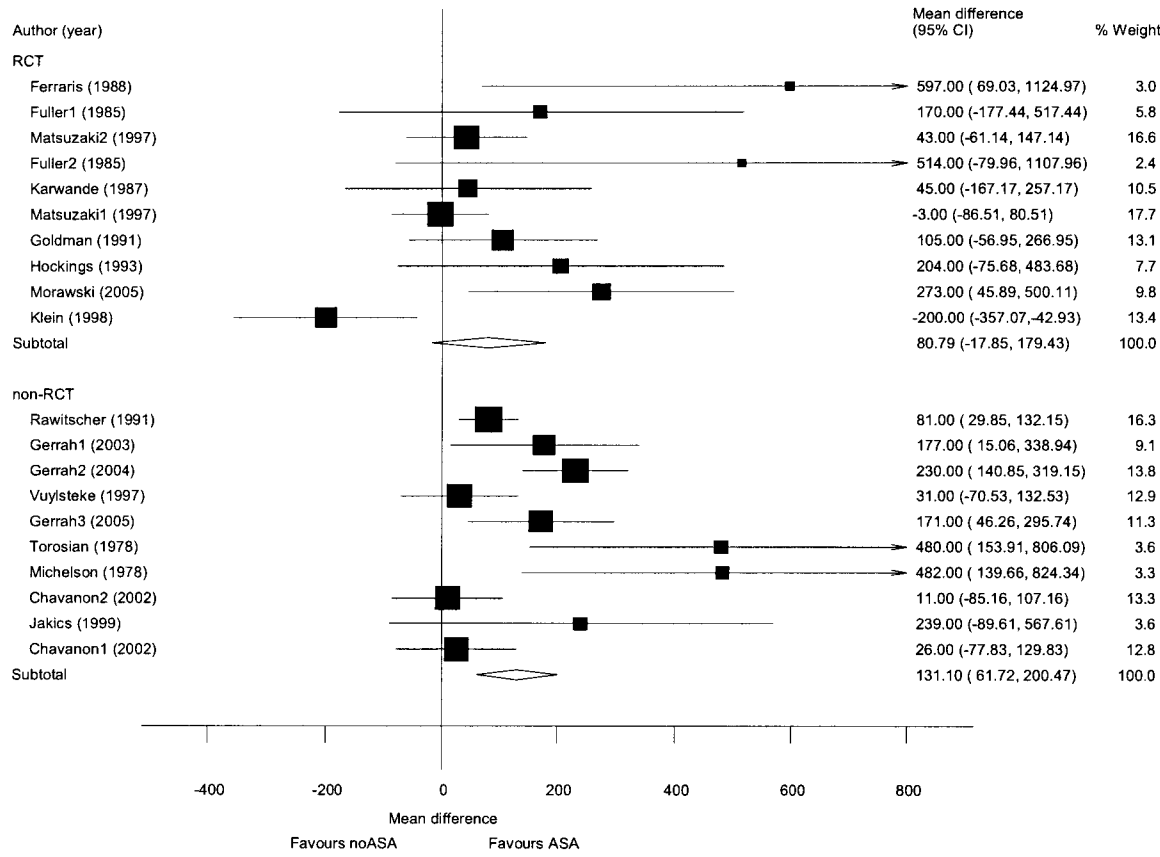


Figure 1b Classical Random Effects Model of Pooled Amount of Bleeding

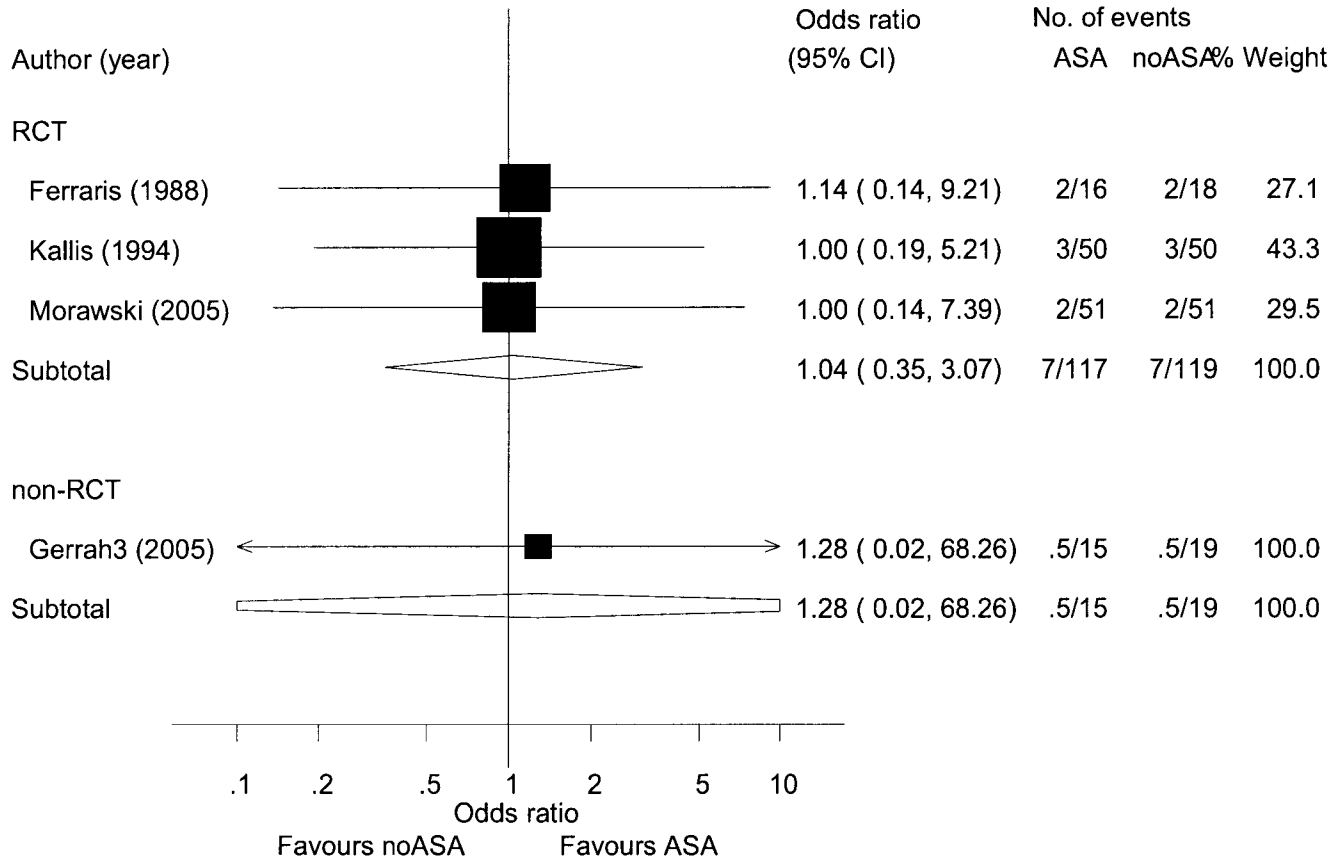


Figure 1c Classical Random Effects Model of Pooled Myocardial Infarction (MI)

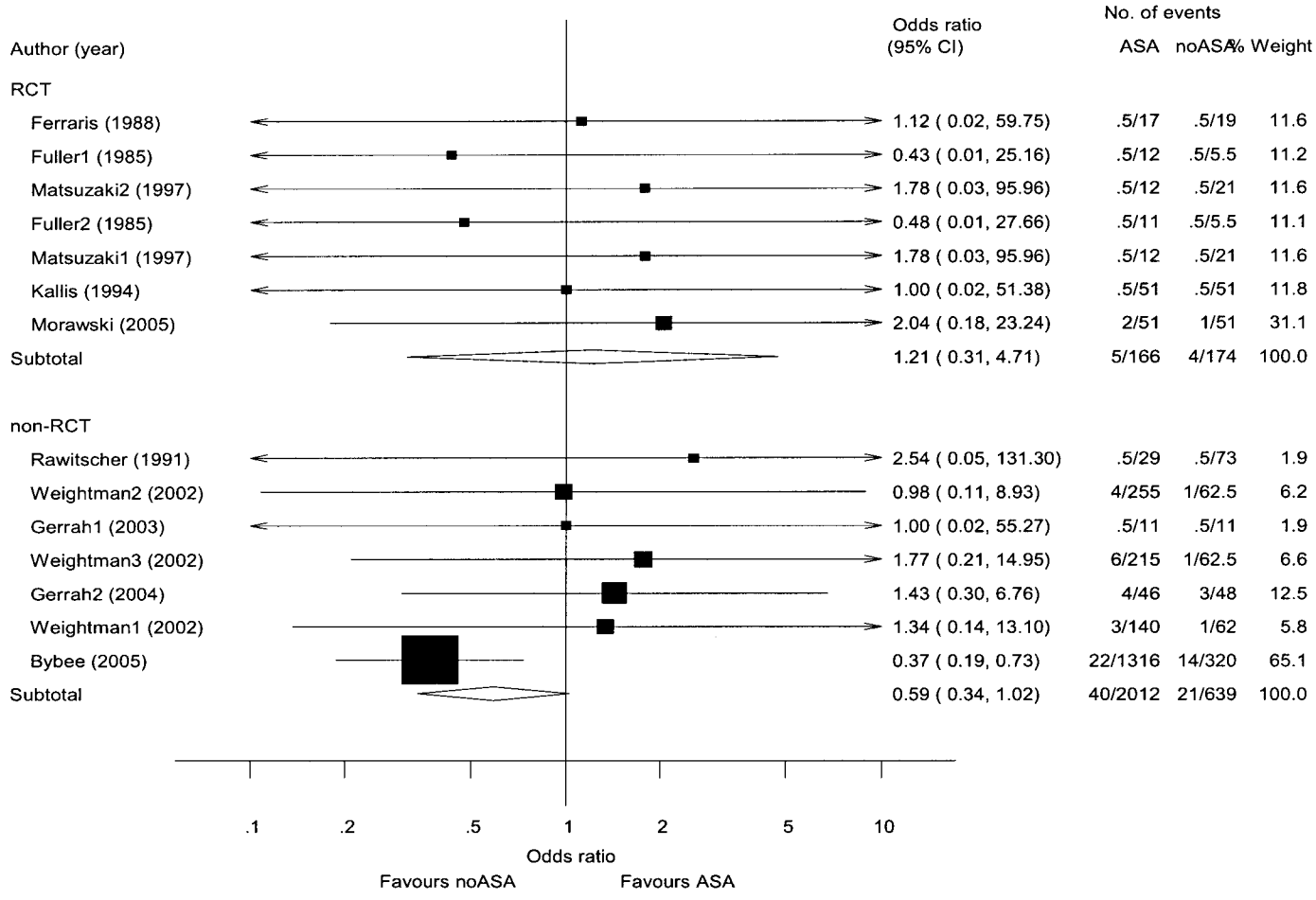


Figure 1d Classical Random Effects Model of Pooled Mortality

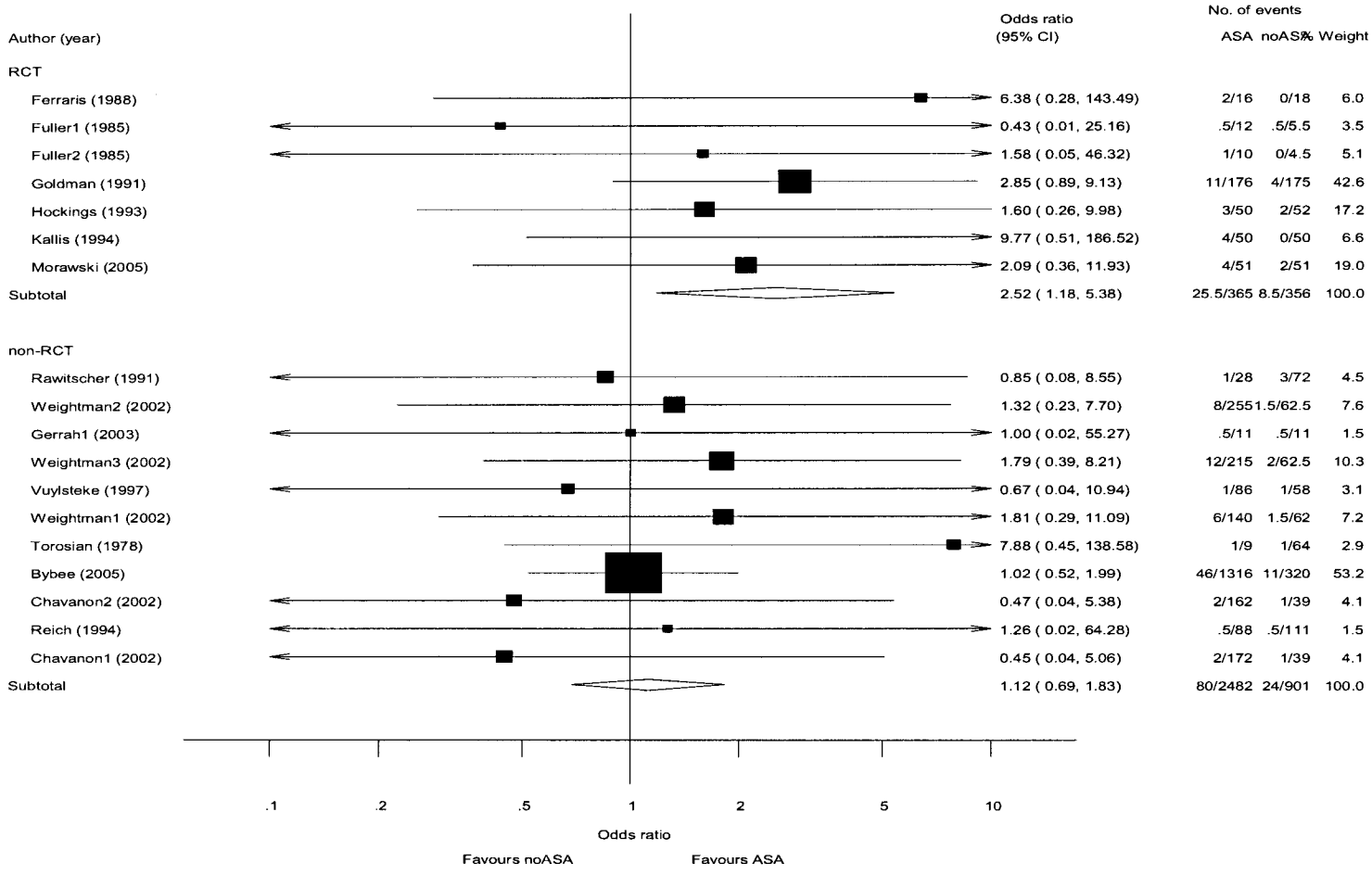


Figure 1e Classical Random Effects Model of Pooled Re-operation

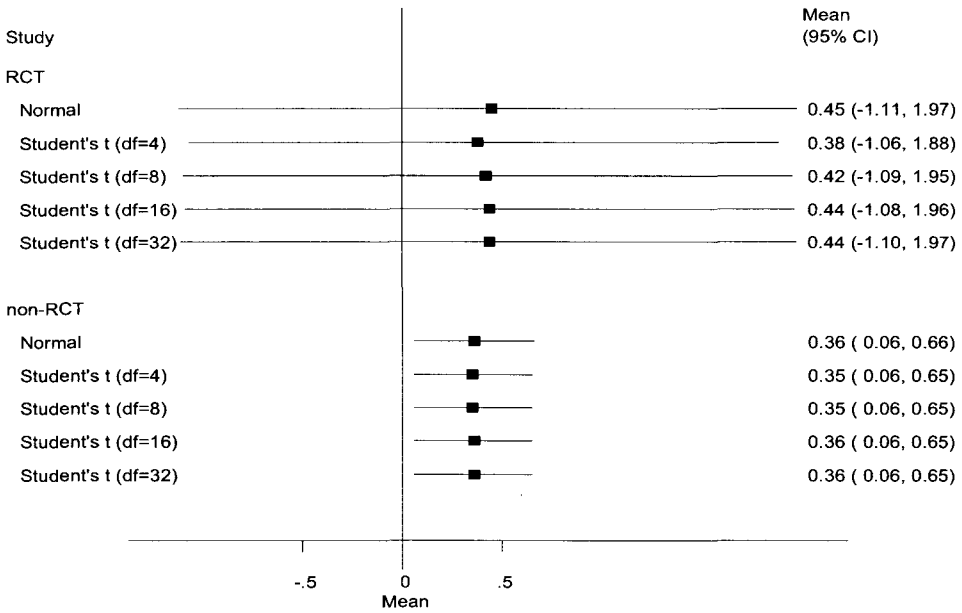


Figure 2 Impacts of Different Likelihood Functions on Overall Estimate of Blood Transfusion

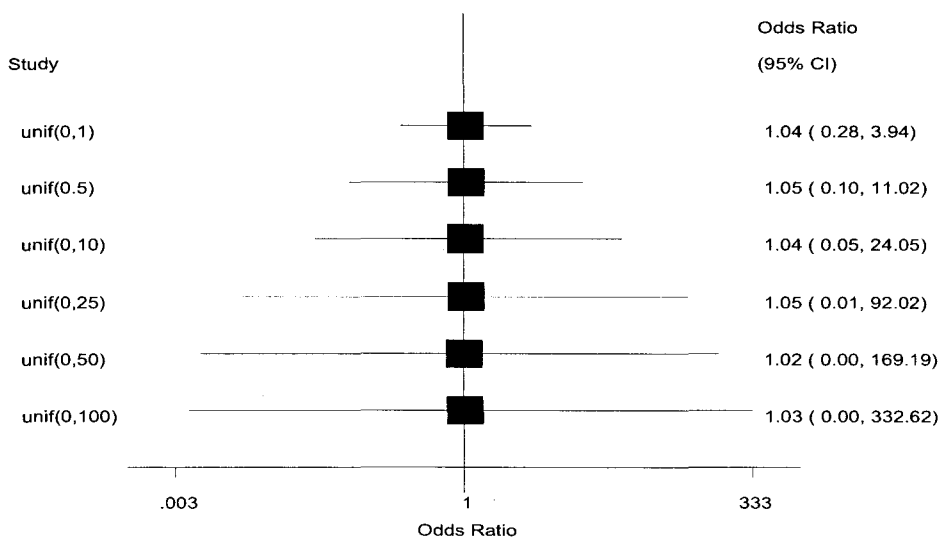


Figure 3 Impacts of Different Priors on Overall Estimate of Myocardial Infarction (MI)

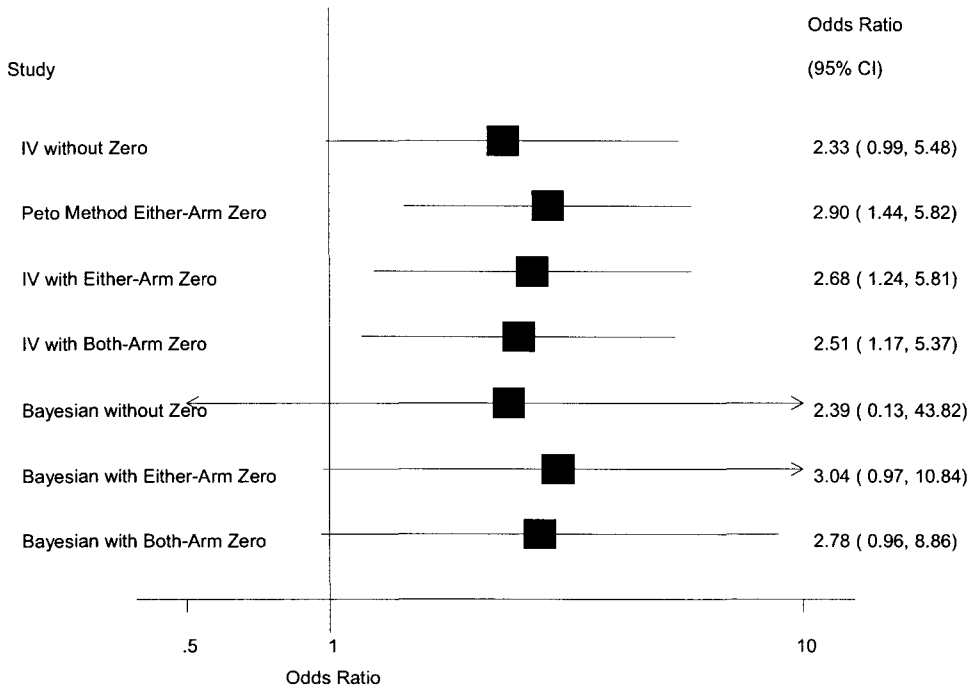


Figure 4 Comparison of Excluding Zero-events, Including Either-arm Zero-events and Including All Zero-events

Appendix B

Tables of Extracted Data, Summaries of the Primary Analysis and Summaries of the Sensitivity Analysis

					Aspirin Group				No Aspirin Group		
Reference Number	Author	Year	Design	Centre(s)	No. of Patient	Mean of Patient Age	Percentage of Male Patient	Aspirin Dose	No. of Patient	Mean of Patient Age	Percentage of Male Patient
10	Fuller1	1985	RAT	Single	11	53.0	.	325	5	59.0	.
10	Fuller2	1985	RAT	Single	10	60.0	.	2,600	5	59.0	.
11	Kewanee	1987	RAT	Single	14	58.0	75.0	80	10	70.0	70.0
12	Ferraris	1988	RAT	Single	16	64.3	87.5	325	18	60.7	88.9
13	Goldman	1991	RAT	Multiple	176	60.0	100.0	325	175	60.0	100.0
14	Hockings	1993	RAT	Single	50	60.0	94.0	100	52	60.0	92.3
15	Challis	1994	RAT	Single	50	62.0	82.0	300	50	62.0	80.0
16	Matsuzaki1	1997	RAT	Single	11	62.0	63.6	239	20	63.0	70.0
16	Matsuzaki2	1997	RAT	Single	11	64.0	72.7	262	20	63.0	70.0
17	Klein	1998	RAT	Multiple	37	.	.	100	36	.	.
18	Mora ski	2005	RAT	Single	51	61.3	94.1	300	51	61.1	82.4
19	Michelson	1978	non-RAT	Single	9	53.4	100.0	600	16	54.6	100.0
20	Torsion	1978	non-RAT	Single	9	53.4	100.0	1,500	64	53.2	89.1
21	Taggart1	1990	non-RAT	Single	28	58.0	89.3	150	33	56.0	86.9
21	Taggart2	1990	non-RAT	Single	29	58.0	82.8	300	33	56.0	86.9
21	Taggart3	1990	non-RAT	Single	44	56.0	75.0	75	33	56.0	86.9
22	Ravisher	1991	non-RAT	Single	28	62.8	82.1	85	72	60.7	73.6
23	Reich	1994	non-RAT	Single	87	67.7	65.5	.	110	68.2	69.1
24	Vuylsteke	1997	non-RAT	Single	86	61.0	81.4	325	58	62.0	81.0
25	Jakics	1999	non-RAT	Single	51	64.5	74.5	325	49	61.7	73.5
26	Chavanon1	2002	non-RAT	Single	172	.	.	.	39	.	.
26	Chavanon2	2002	non-RAT	Single	162	.	.	.	39	.	.
27	Weightman1	2002	non-RAT	Single	140	62.1	82.9	.	62	62.0	80.6
27	Weightman2	2002	non-RAT	Single	255	61.4	83.1	.	63	62.0	80.0
27	Weightman3	2002	non-RAT	Single	215	61.6	86.0	.	63	62.0	80.0
28	Gerrah1	2003	non-RAT	Single	10	64.0	70.0	100	10	62.0	80.0
29	Ray	2003	non-RAT	Single	105	66.7	67.6	.	497.00	64.9	78.9
30	Gerrah2	2004	non-RAT	Single	46	67.0	73.9	100	48.00	70.0	81.3
31	Gerrah3	2005	non-RAT	Single	14	64.0	85.7	100	18.00	62.0	66.7
32	Bybee	2005	non-RAT	Single	1,316	69.0	76.2	.	320.00	69.0	72.5

Table 1 Study Characteristics and Patient Demographics
(reference number: corresponding number in bibliography)

	RCT				non-RCT			
	Blood Transfusion (n=6)		Bleeding (n=14)		Blood Transfusion (n=10)		Bleeding (n=10)	
Likelihood functions	MD	Cr.I.	MD	Cr.I.	MD	Cr.I.	MD	Cr.I.
Normal	0.45	(-1.11, 1.97)	99.00	(-46.37, 227.60)	0.36	(0.06, 0.66)	134	(7.26, 240.60)
Student's *t* (df=4)	0.38	(-1.06, 1.88)	95.87	(-43.32, 222.90)	0.35	(0.06, 0.65)	130.1	(9.52, 233.50)
Student's *t* (df=8)	0.42	(-1.09, 1.95)	97.23	(-46.00, 225.30)	0.35	(0.06, 0.65)	132.5	(8.55, 237.10)
Student's *t* (df=16)	0.44	(-1.08, 1.96)	98.16	(-45.49, 226.00)	0.36	(0.06, 0.65)	133.9	(7.091, 239.9)
Student's *t* (df=32)	0.44	(-1.10, 1.97)	98.16	(-46.32, 277.70)	0.36	(0.06, 0.65)	134.2	(8.44, 239.20)

Table 3 Comparison of the Impact of Different Likelihood Functions on Continuous Outcomes for the Bayesian Random Effects Model
(MD: mean difference; Cr.I.: credible interval; df: degrees of freedom; n: number of studies)

Priors	RCT						non-RCT			
	MI (n=3)		Mortality (n=7)		Re-operation (n=7)		Mortality (n=7)		Re-operation (n=11)	
	OR	Cr.I.	MD	Cr.I.	MD	Cr.I.	MD	Cr.I.	MD	Cr.I.
unif(0,1)	1.04	(0.28, 3.94)	1.26	(0.27, 6.07)	2.74	(1.15, 6.91)	0.83	(0.37, 2.17)	1.19	(0.65, 2.26)
unif(0,5)	1.05	(0.10, 11.02)	1.23	(0.19, 7.88)	2.77	(0.95, 8.71)	0.91	(0.33, 3.60)	1.20	(0.63, 2.39)
unif(0,10)	1.04	(0.05, 24.05)	1.23	(0.18, 8.18)	2.78	(0.96, 8.86)	0.91	(0.33, 3.59)	1.20	(0.63, 2.39)
unif(0,25)	1.05	(0.01, 92.02)	1.22	(0.18, 8.22)	2.77	(0.94, 8.87)	0.92	(0.34, 3.62)	1.20	(0.63, 2.36)
unif(0,50)	1.02	(0.005, 169.19)	1.22	(0.18. 8.25)	2.77	(0.94, 8.87)	0.92	(0.33, 3.60)	1.20	(0.63, 2.36)
unif(0,100)	1.03	(0.004, 332.62)	1.22	(0.18, 8.25)	2.77	(0.94, 8.87)	0.92	(0.33, 3.60)	1.20	(0.63, 2.36)

Table 4 Comparison of the Impact of Different Priors on Binary Outcomes for Bayesian Random Effects Model
(unif: uniform distribution; OR: odds ratio; Cr.I.: credible interval; n: number of studies)

	OR	C.I	Length of C.I. / Cr.I.
IV without Zero Event Studies	2.33	(0.99, 5.48)	4.49
Peto Method	2.90	(1.44, 5.82)	4.38
IV with Either Arm Zero Studies	2.68	(1.24, 5.81)	4.57
IV with Both Arm Zero Studies	2.51	(1.17, 5.37)	4.20
Bayesian without Any Zeros	2.39	(0.13, 43.82)	43.69
Bayesian with Either Arm Zeros	3.04	(0.97, 10.84)	9.87
Bayesian with Both Arm Zeros	2.78	(0.96, 8.86)	7.90

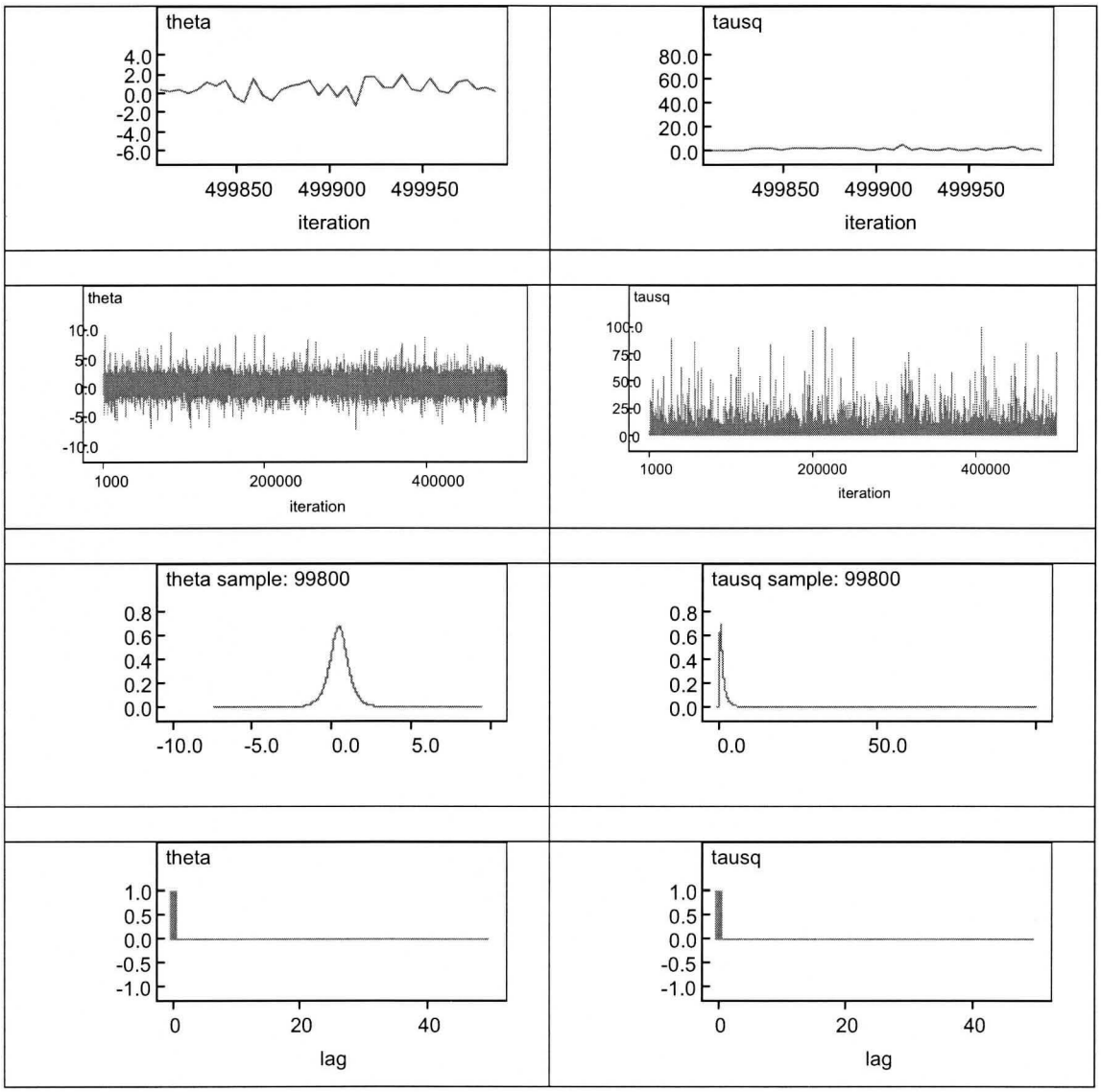
Table 5 Comparison of Excluding and Including Zero-event Studies on Re-operation
(number of studies = 7, either-arm zero-event studies = 3, two-arm zero-event studies = 1)

	Imputation	Estimates	L.L.	U.L.
RCT Blood Transfusion	1	0.489	0.059	0.919
	2	0.400	-0.19	0.991
	3	0.493	0.063	0.923
	4	0.667	0.175	1.158
	5	0.566	0.108	1.023
	6	0.461	-0.009	0.931
	7	0.463	-0.004	0.93
	8	0.538	0.131	0.944
	9	0.442	-0.07	0.952
	10	0.461	-0.029	0.952
	Average	0.50	0.02	0.97
	Original	0.36	-0.37	1.08
Non-RCT Blood Transfusion	1	0.554	0.099	1.01
	2	0.520	0.26	0.779
	3	0.606	0.278	0.933
	4	0.534	0.268	0.801
	5	0.488	0.251	0.73
	6	0.343	0.11	0.576
	7	0.446	0.203	0.688
	8	0.565	0.152	0.977
	9	0.275	-0.011	0.561
	10	0.324	0.03	0.62
	Average	0.47	0.16	0.77
	Original	0.34	0.12	0.56

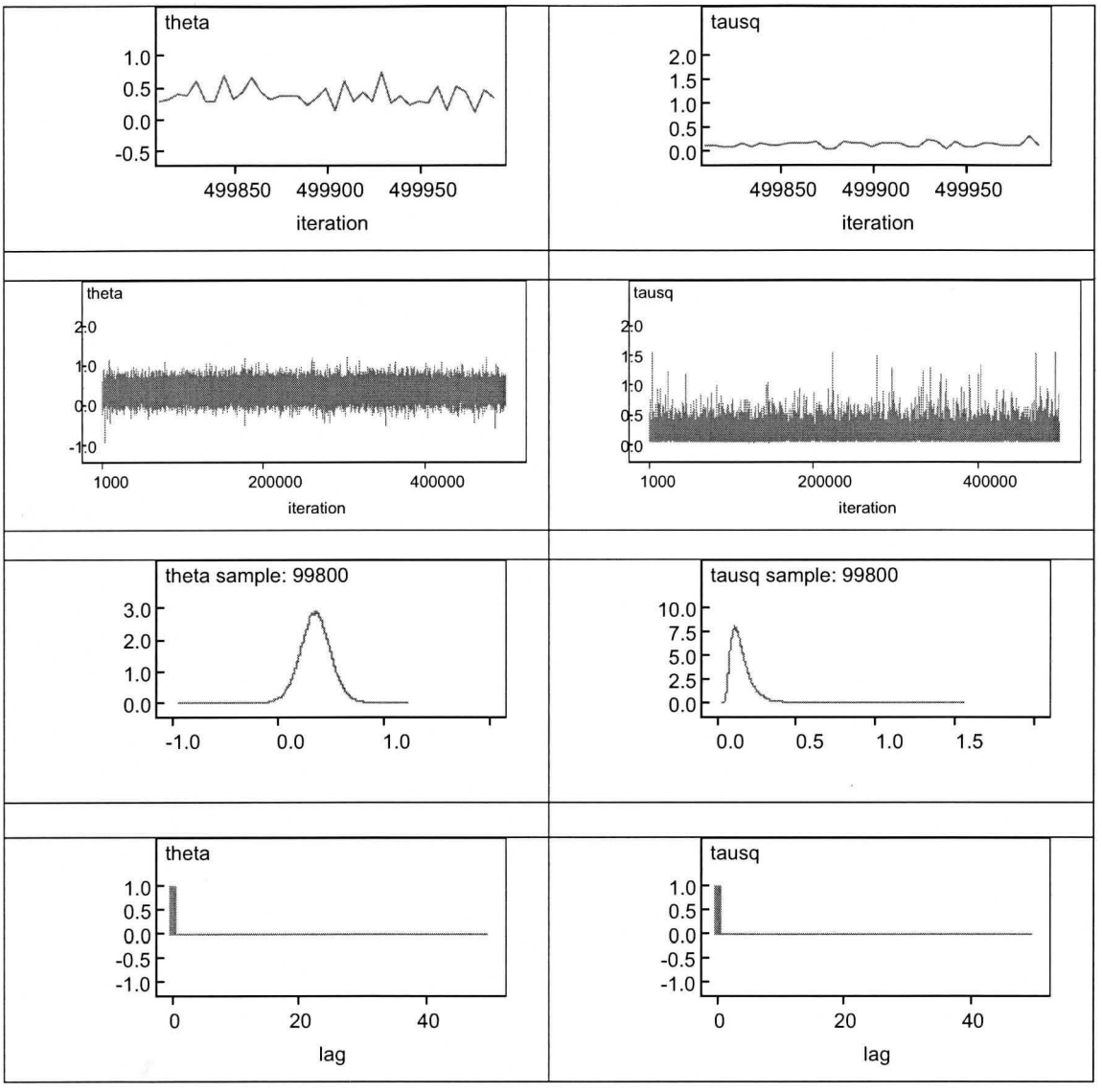
Table 6 Comparison of the Overall Estimates of Blood Transfusion After-Before Imputation (missing rate in RCTs = 0.45; missing rate in non-RCT = 0.26)

Appendix C

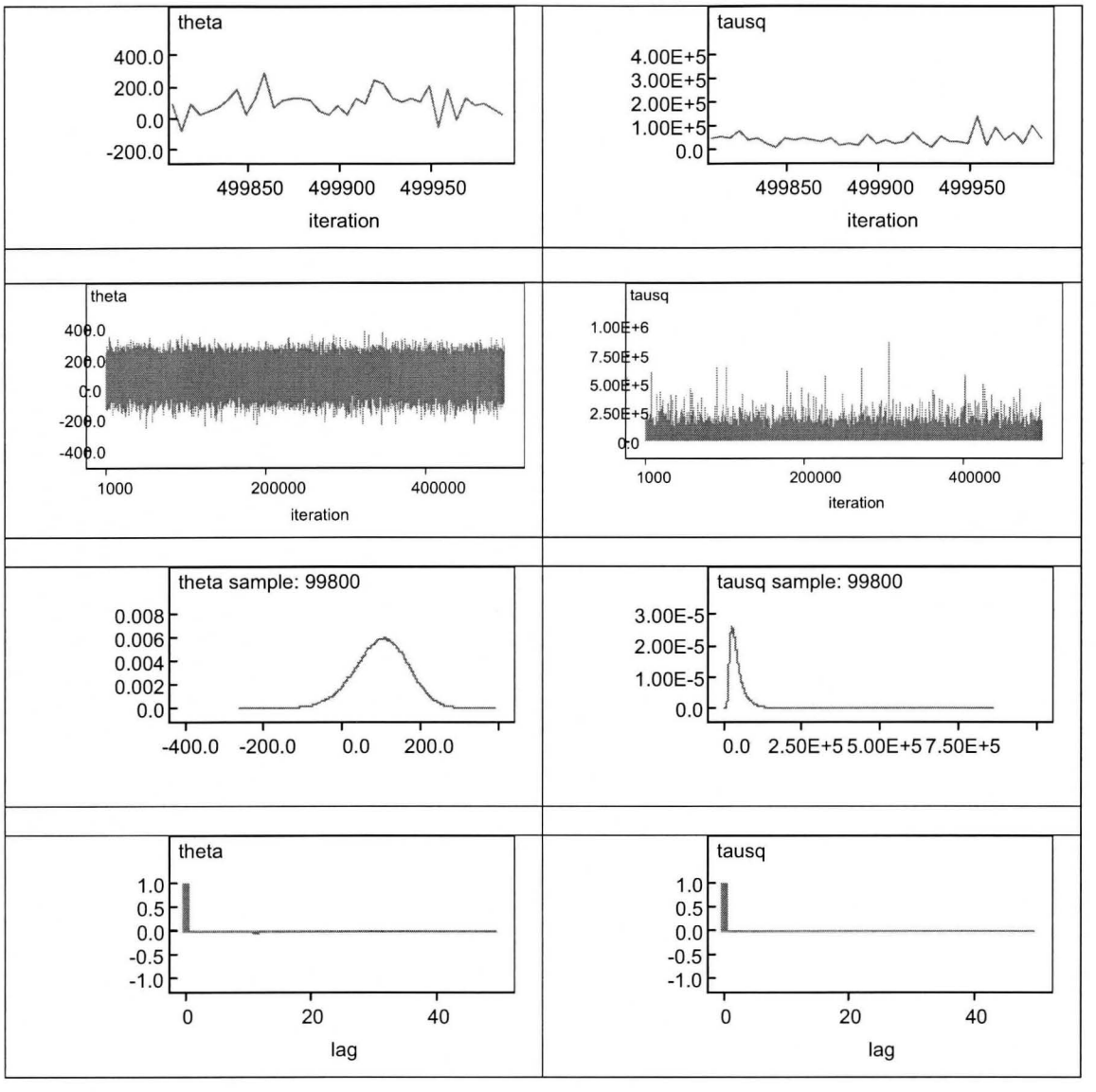
Plots of Dynamic Trace, Kernel Density, Time Series and Autocorrelation of Bayesian Random Effects Models



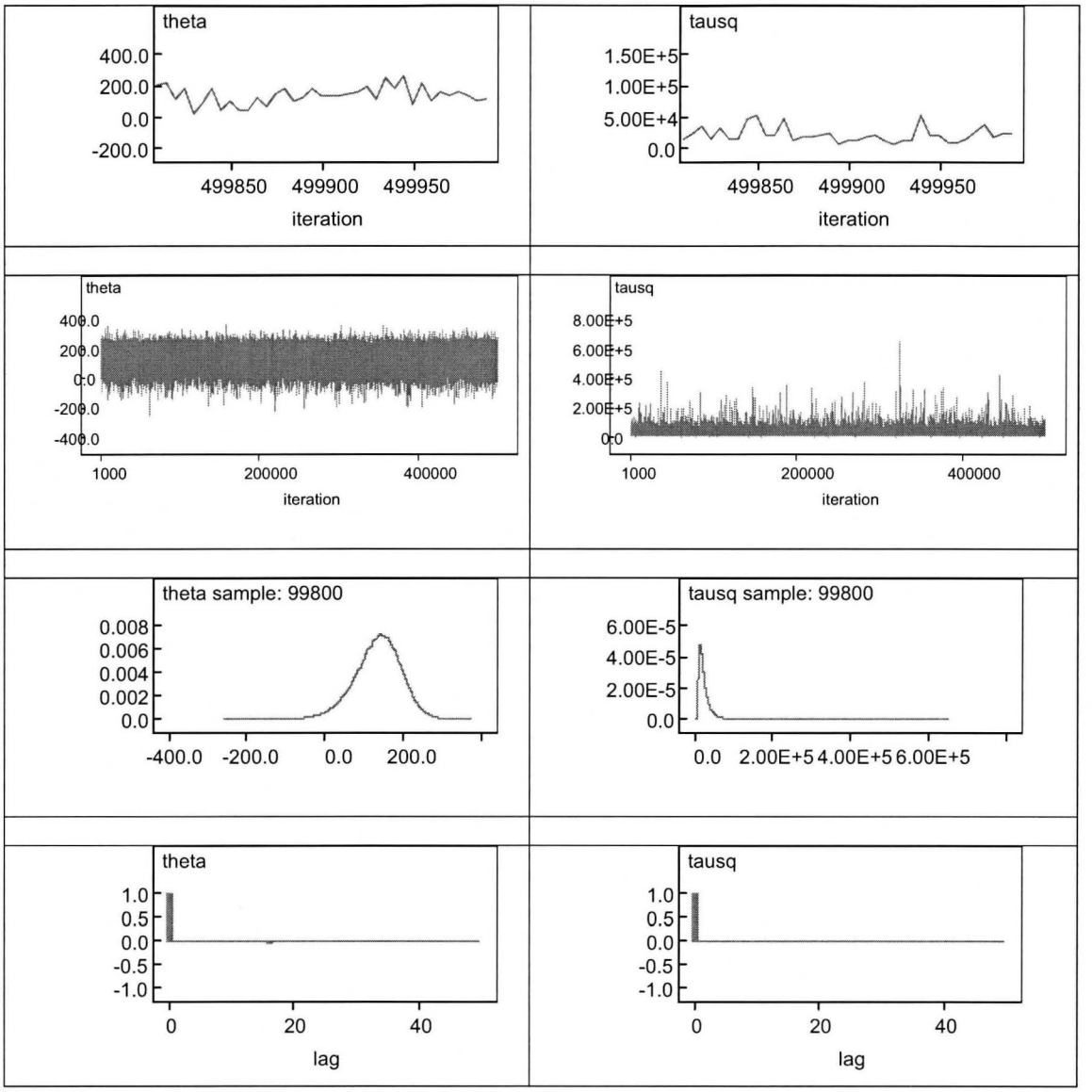
Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of Mean and Between-Variance of Blood Transfusion in RCT Studies



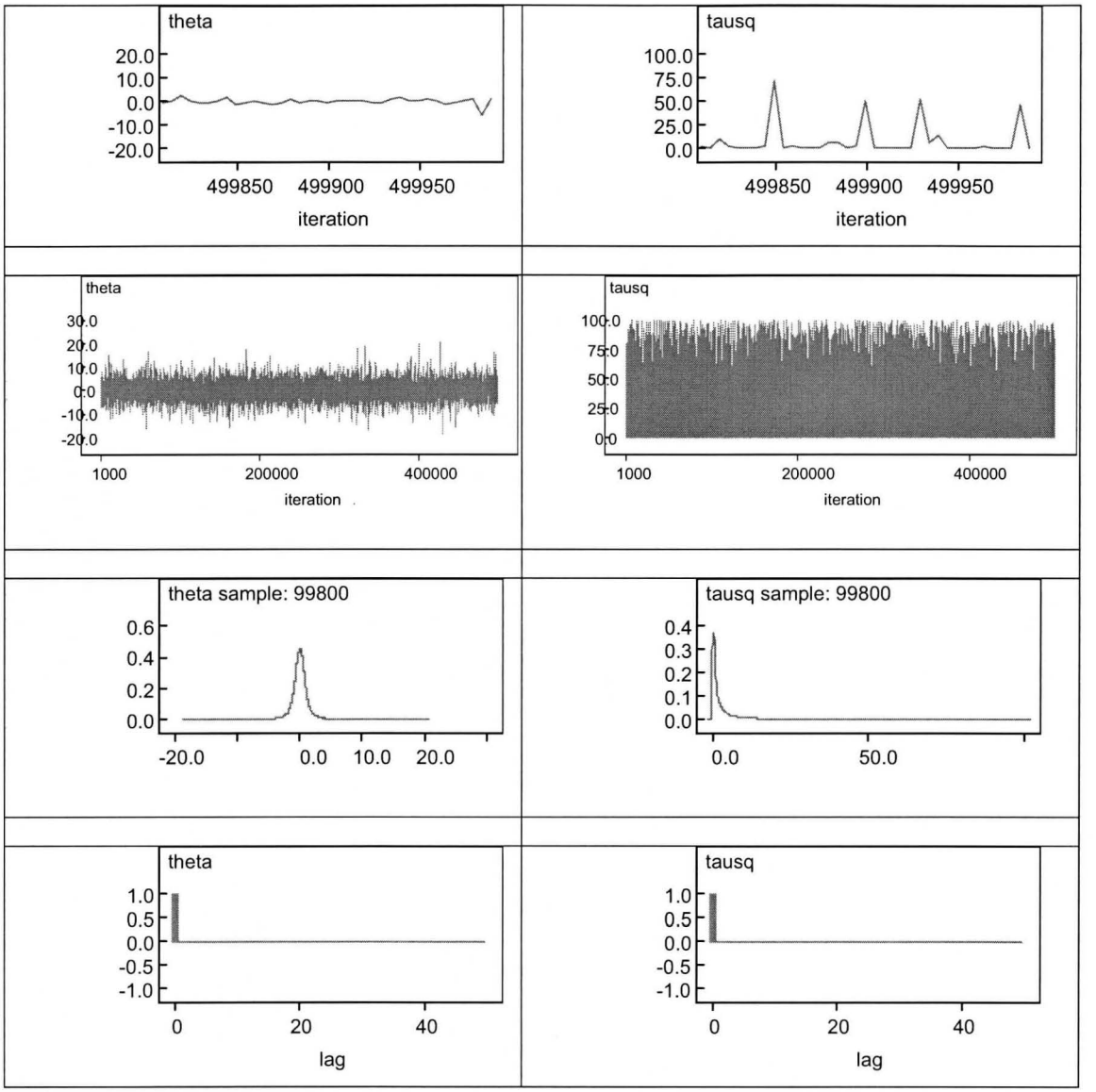
Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of
Mean and Between-Variance of Blood Transfusion in non-RCT Studies



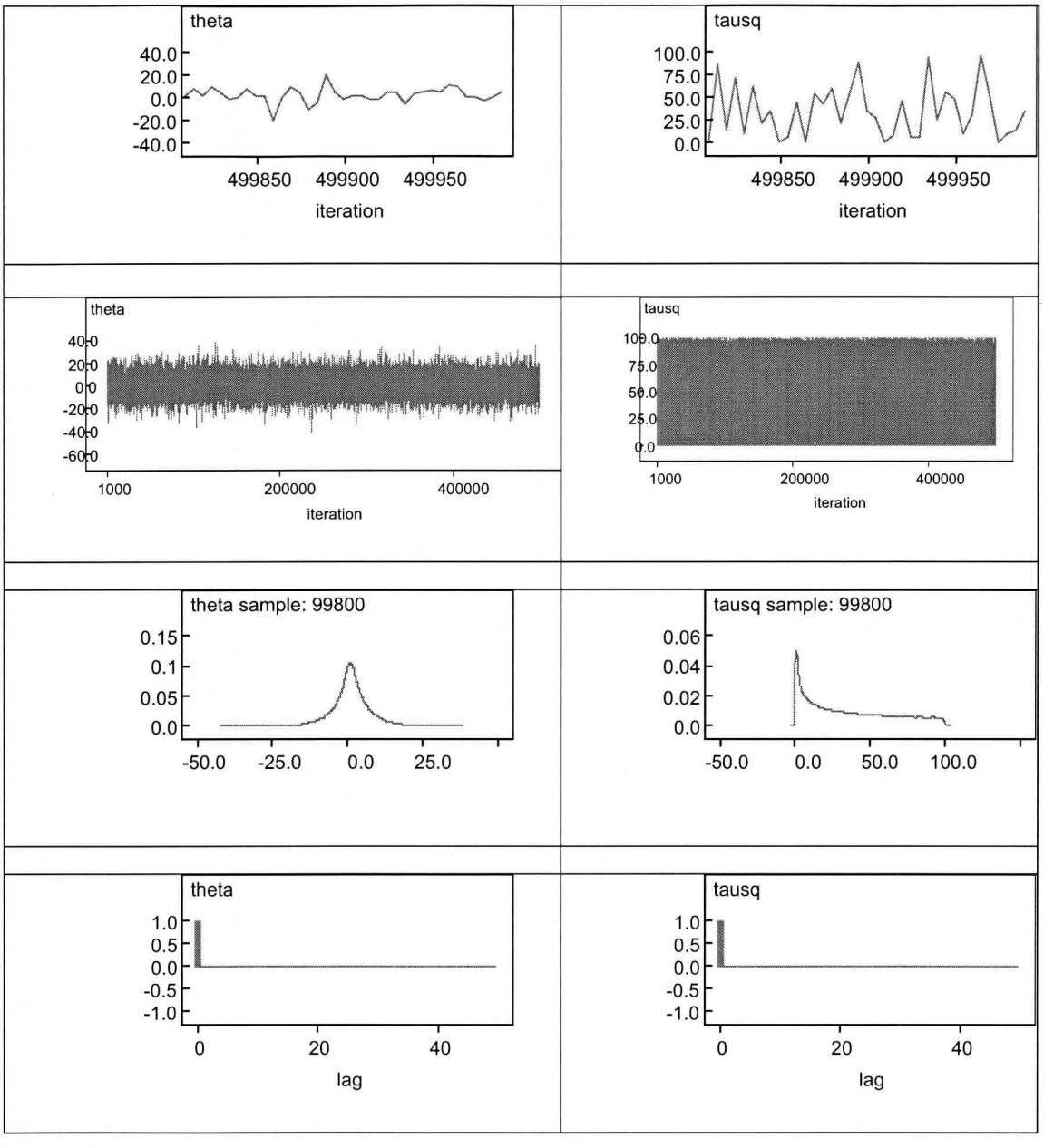
Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of
Mean and Between-Variance of Bleeding in RCT Studies



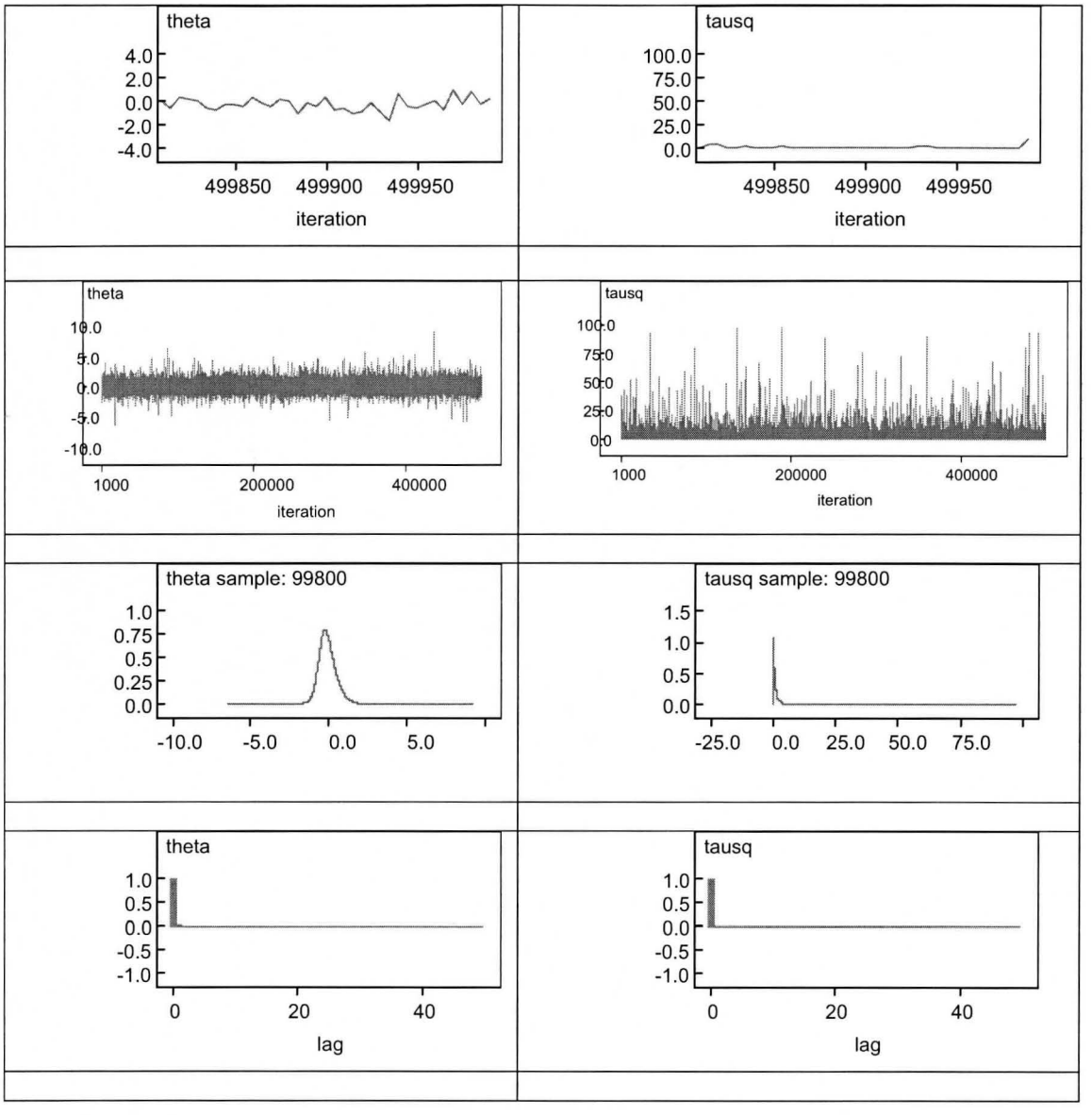
Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of Mean and Between-Variance of Bleeding in non-RCT Studies



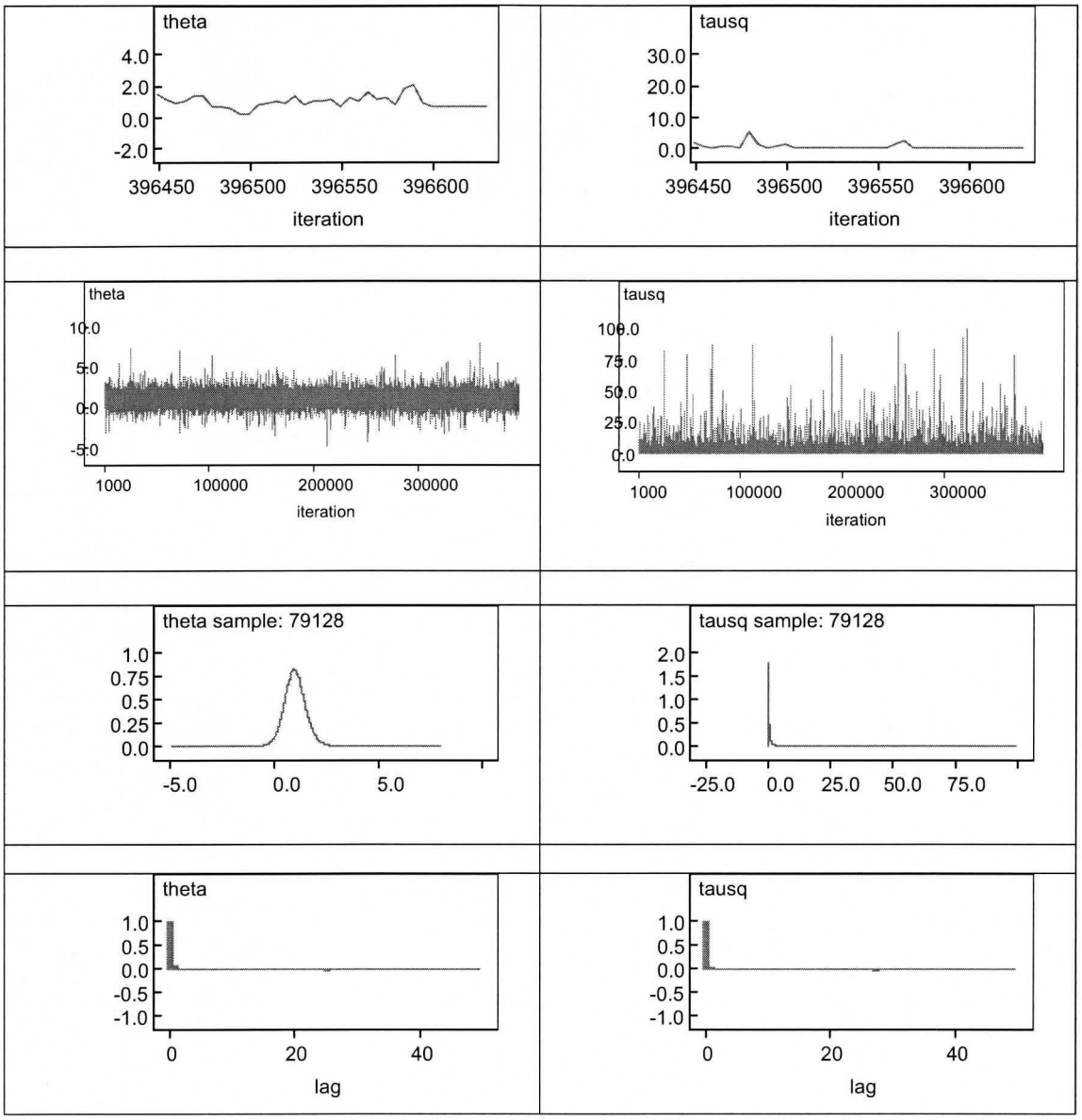
Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of
Mean and Between-Variance of Myocardial Infarction (MI) RCT Studies



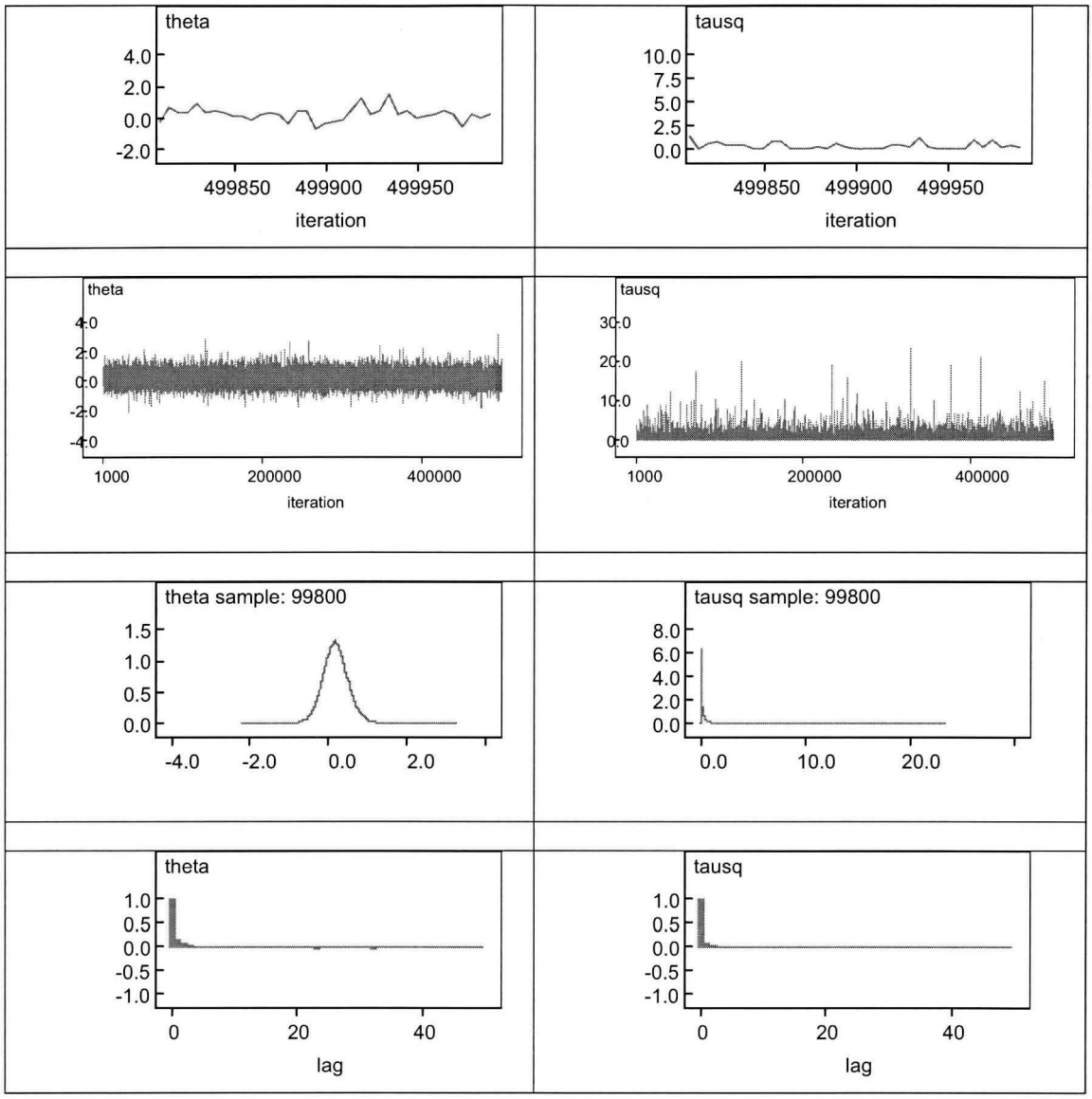
Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of
Mean and Between-Variance of Mortality RCT Studies



Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of
Mean and Between-Variance of Mortality non-RCT Studies



Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of Mean and Between-Variance of Re-operation RCT Studies



Plots (Dynamic Trace, Time Series, Density and Autocorrelation) of Mean and Between-Variance of Re-operation non-RCT Studies

Appendix D

Codes for Bayesian Random Effects Model with Input Data, Initial Values and Seeds

```
model # Random effect model for continuous outcomes
{ #k<-200 # k=4, 8, 16, 32 #(df of Student't distribution)
      for (i in 1:r)
      {
         y[i]~dnorm(psi[i],w[i]) #normal likelihood
            # y[i]~dt(psi[i],w[i],k) # student's likelihood
            psi[i]~dnorm(theta,t)
            w[i]<-1/(sigma*sigma)

      }
      theta~dnorm(0,1.0E-6)
      sigma~dunif(0,10)
      t<-1/(sigma*sigma)
      tausq<-1/t  #between-study variance of normal likelihood
      #tausq<-k/(t*(k-2))  #between-study variance of student's t
      #likelihood
                  }


#blood transfusion in RCT
list(y=c(2.6,0.44,-0.6,0.3,-0.8,0.73),r=6)

#initial
list(theta=0,sigma=1,psi=c(0,0,0,0,0,0))



#blood transfusion in non-RCT
list(y=c(0.5,0.2,-0.33,0.9,-0.33,1.4,0.2,0.1,0.5,0.5,0.4,0.9,0.2,0.3),
r=14)

#initial
list(theta=0,sigma=1,psi=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0))



#bleeding in RCT
list(y=c(597,170,514,105,204,45,-200,-3,43,273),r=10)

# initial
list(theta=0,sigma=1,psi=c(0,0,0,0,0,0,0,0,0,0))



#bleeding in non-RCT
list(y=c(26,11,177,230,171,239,482,81,480,31), r=10)

#initial
list(theta=0,sigma=1,psi=c(0,0,0,0,0,0,0,0,0,0))


#seed = 14721869
```

```
model # Random effect model for binary outcomes
      {
                for (i in 1:r) {
                  rc[i] ~ dbin(pc[i],nc[i])
                  rt[i] ~ dbin(pt[i],nt[i])
                  logit(pc[i])<-mu[i]
                  logit(pt[i])<-mu[i]+delta[i]
                  mu[i]~dnorm(0,1.0E-5)
                  delta[i]~dnorm(theta,t)
                  }
             theta~dnorm(0, 1.0E-6)
             t<-1/(sigma*sigma)
             sigma~dunif(0,10)
             #sigma~dunif(0,1) #prior for sensitivity analysis
             #sigma~dunif(0,5) #prior for sensitivity analysis
             #sigma~dunif(0,25) #prior for sensitivity analysis
             #sigma~dunif(0,50) #prior for sensitivity analysis
             #sigma~dunif(0,100) #prior for sensitivity analysis
             #tausq<-1/t  #between-study variance

      }


#MI in RCT
list(nt=c(16,50,51),rt=c(2,3,2),nc=c(18,50,51),rc=c(2,3,2),r=3)

#initial
list(theta = 0, sigma=1, mu = c( 0, 0, 0), delta = c(0, 0, 0) )



#Mortality in RCT
list(nt=c(17,12,11,51,12,12,51),rt=c(0.5,0.5,0.5,0.5,0.5,0.5,2),
      nc=c(19,5.5,5.5,51,21,21,51),rc=c(0.5,0.5,0.5,0.5,0.5,0.5,1),r=7)

#initial
list(theta = 0, sigma=1, mu = c( 0,0,0,0,0,0,0), delta =
c(0,0,0,0,0,0,0) )



#Mortality in non-RCT
list(nt=c(1316,11,29,46,140,255,215),rt=c(22,0.5,0.5,4,3,4,6),
      nc=c(320,11,73,48,62.5,62.5,62),rc=c(14,0.5,0.5,3,1,1,1),r=7)

#initial
list(theta = 0, sigma=1, mu = c(0,0,0,0,0,0,0),  delta =
c(0,0,0,0,0,0,0) )
```

```
#Re-operation in RCT
list(nt=c(51,12,10,16,176,50,50),rt=c(4,0.5,1,2,11,3,4),
      nc=c(51,5.5,5.5,19,175,52,51),rc=c(2,0.5,0.5,0.5,4,2,0.5),r=7)

#initial
list(theta = 0, sigma=1, mu = c( 0,0,0,0,0,0,0), delta =
c(0,0,0,0,0,0,0) )



#Re-operation in no-RCT
list(nt=c(1316,172,162,11,28,87,9,86,140,255,215),
   rt=c(46,2,2,0.5,1,0.5,1,1,6,8,12),
   nc=c(320,39,39,11,72,111,64,58,62,62.5,62.5),rc=c(11,1,1,0.5,3,0.5,1,1
   ,2,1.5,1.5),r=11)

#initial
list(theta= 0, sigma=1, mu = c( 0,0,0,0,0,0,0,0,0,0,0), delta =
c(0,0,0,0,0,0,0,0,0,0,0) )


#seed = 14721869
```