Canadian Agency for Drugs and Technologies in Health

Agence canadienne des médicaments et des technologies de la santé

HTA

Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis

March 2009

Supporting Informed Decisions

Until April 2006, the Canadian Agency for Drugs and Technologies in Health (CADTH) was known as the Canadian Coordinating Office for Health Technology Assessment (CCOHTA).

Publications can be requested from:

CADTH
600-865 Carling Avenue
Ottawa ON Canada K1S 5S8
Tel.: 613-226-2553
Fax: 613-226-5392
Email: pubs@cadth.ca

or downloaded from CADTH's website:
http://www.cadth.ca

Cite as: Wells GA, Sultan SA, Chen L, Khan M, Coyle D. *Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis.* Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009.

Production of this report is made possible by financial contributions from Health Canada and the governments of Alberta, British Columbia, Manitoba, New Brunswick, Newfoundland and Labrador, Northwest Territories, Nova Scotia, Nunavut, Ontario, Prince Edward Island, Saskatchewan, and Yukon. The Canadian Agency for Drugs and Technologies in Health takes sole responsibility for the final form and content of this report. The views expressed herein do not necessarily represent the views of Health Canada or any provincial or territorial government.

CADTH is funded by Canadian federal, provincial, and territorial governments.

Legal Deposit – 2009
National Library of Canada
ISBN: 978-1-897465-88-2 (print)
ISBN: 987-1-897465-89-9 (online)
H0462 – March 2009

PUBLICATIONS MAIL AGREEMENT NO. 40026386
RETURN UNDELIVERABLE CANADIAN ADDRESSES TO
CANADIAN AGENCY FOR DRUGS AND TECHNOLOGIES IN HEALTH
600-865 CARLING AVENUE
OTTAWA ON K1S 5S8

Canadian Agency for Drugs and Technologies in Health

Indirect Evidence:
Indirect Treatment Comparisons in Meta-Analysis

George A Wells, MSc PhD[1,2,3]
Shagufta A. Sultan, MSc PhD[4]
Li Chen, MSc[2]
Maryam Khan, BHSc[2]
Doug Coyle, PhD[1,3]

March 2009

[1] University of Ottawa,, Ottawa, Ontario
[2] University of Ottawa Heart Institute, Ottawa, ON
[3] Ottawa Health Research Institute, Ottawa, ON
[4] Health Canada, Health Products and Food Branch, Ottawa, ON

Reviewers

These individuals kindly provided comments on this report:

External Reviewers

Gordon Guyatt, MD MSc
Professor
McMaster University
Hamilton, Ontario

Heiner C. Bucher, MD MPH
Professor,
Basel Institute for Clinical Epidemiology
Director,
Department of Internal Medicine
University Hospital Basel
Basel, Switzerland

Larry Stitt, MSc BSc
Assistant Director, Biostatistical Support Unit
Department of Epidemiology and Biostatistics
The University of Western Ontario
London, Ontario

Fujian Song, PhD
Reader in Research Synthesis
University of East Anglia
Norwich, UK

CADTH Peer Review Group Reviewers

Brian Hutton, MSc
Epidemiologist and Biostatistician
CADTH
Ottawa, Ontario

Authorship

George Wells was involved in the conception, design and implementation of the project, as well as all aspects of the development and writing of the original and revised reports and responses to internal and external reviewers.

Shagufta Sultan was involved in the writing of the original report, derivation of algebraic expressions, screening and analysis of the illustrative examples and the response to the reviewers.

Li Chen was involved in the design and conduct of the simulation and analyses associated with the illustrative examples.

Maryam Khan was involved in the writing of the report, the review of the methods, illustrative examples, and responses to the internal and external reviewers.

Doug Coyle assisted with the design of the project and concepts associated with the analysis and simulations.

All authors contributed to the revision of the report.

Conflicts of Interest

Gordon Guyatt is a co-author on one of the frequently cited papers.

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

ACE	angiotensin converting enzyme
AE	adverse event
AHA	American Heart Association
ANOVA	analysis of variance
APTC	Antiplatelet Trialists' Collaboration
ARB	angiotensin receptor blocker
BNZ	benzodiazepines
BPRS	Brief Psychiatric Rating Scale
CABG	coronary artery bypass graft surgery
CC	conventional care
CCB	calcium channel blocker
CHF	congestive heart failure
CI	confidence interval
CISCA	cisplatin and cisplatin, cyclophosphamide and adriamycin
CPM	confidence profile method
CPO	conditional predictive ordinate
CRT	cardiac resynchronization therapy
CV	coefficient of variation
EE	effect estimator
EF	ejection fraction
EPS	extra pyramidal side effects
ES	effect size
GC	gemcitabine cisplatin
GLM	general linear model
HR	hazard ratio
ICD	implantable cardioverter defibrillator
ICDF	inconsistency degrees of freedom
ITC	indirect treatment comparisons
LCL	lower confidence limit
LOR	log odds ratio
LV	left ventricular
MAOI	monoamine oxidase inhibitor
MCMC	Markov Chain Monte Carlo
MD	mean difference
MI	myocardial infarction
MSE	mean square error
MTC	mixed treatment comparison
MTM	mixed treatment meta-analysis

MVAC	methotrexate, vinblastine, and doxorubicin
NBNZ	nonbenzodiazepenes
NNRTI	non-nucleoside reverse transcriptase inhibitor
NRTI	nucleoside analogue reverse transcriptase inhibitors
NYHA	New York Heart Association
OMT	optimal medical therapy
OR	odds ratio
PANSS	Positive and Negative Syndrome Scale
PCTA	percutaneous transluminal coronary angioplasty
PI	protease inhibitor
RCT	randomized controlled trial
RD	risk difference
RR	relative risk
RRR	relative risk reduction
SE	standard error
SK	streptokinase
SOL	sleep onset latency
SQ	sleep quality
SRS	simple random sample
SSRI	selective serotonin reuptake inhibitor
TCA	tricyclic antidepressant
TIA	transient ischemic attack
t-PA	tissue type plasminogen activator
TST	total sleep time
UCL	upper confidence limit
WASO	wakefulness after sleep onset

1 INTRODUCTION

1.1 Description of the Problem

A direct assessment of two treatments, which are labeled A and C, is available if a comparative study of A and C has been conducted. Ideally, this comparative study is a randomized controlled trial (RCT) comparing A and C. However, many competing treatments have not been compared directly and/or such direct evidence is limited and insufficient. Indirect treatment comparisons (ITC) have been advocated when direct evidence is not available. These types of comparisons are becoming more commonplace.

In the simplest situation, we may have direct comparisons of A versus C and B versus C; indirect methods attempt to use the common comparator link C to yield an indirect comparison of A versus B. Various reasons may lead to the lack of a direct comparison and, in these situations, performing an indirect treatment comparison may be of interest. Firstly, often when the treatments of interest are drugs, due to commercial interests and the regulatory approval process, head-to-head comparison of two active treatments is often not available. Placebo-controlled trials are usually sufficient for acquiring regulatory approval of a new treatment, and there is no motivation from the commercial sector to compare the new treatment with existing active treatments. This is of particular importance for reviews where only placebo-controlled studies are directly available and indirect evidence assessing a new treatment with a standard treatment would be of interest in making appropriate assessments. In a setting in which a head-to-head comparison is not available, C could be a placebo comparator and the indirect comparison would provide evidence of a head-to-head comparison of A versus B.

Secondly, if there is strong evidence and belief that the current standard treatment is effective, then placebo-controlled trials may not be ethically conducted. New treatments are compared only with active treatments and there is an absence of comparison of a new treatment to placebo yielding the actual "true" effect of the drug.[1] To get an assessment of the true level of effect of the new drug, an indirect approach can provide some information. In this setting in which a new treatment cannot be compared to placebo, C could be the currently accepted standard of treatment and B is a placebo (for which the standard was compared to in the past); then the indirect comparison would yield evidence of a comparison of A with placebo.

Indirect treatment comparisons can also be useful when a meta-analysis groups together all treatments within the same class or for different doses of the same treatment. As such, when comparing to another treatment, class effects or the varying effects of different doses cannot be evaluated.[2] The lack of such an evaluation can result in erroneous conclusions if one class is recommended over another (even though each treatment within that class is not effective), or when any dose is recommended as being superior to placebo (when, in fact, only a specific dose or specific doses may be better than placebo).

More complex indirect evidence settings can arise. In the next simplest setting, we may have direct evidence from A versus C, B versus D, C versus D; with this evidence, we can attempt an indirect comparison of A versus B using, in particular, the direct evidence of C versus D. Even in the situation of A versus C, B versus C, D versus C, D versus F, the treatment F can be an

important contributor to the indirect comparison of A versus B. The network of direct and indirect evidence can be complex. Within this network of evidence, there is often a need to synthesize evidence from RCTs, and methods for deriving indirect treatment comparisons using meta-analysis are of prime interest.

Although there are situations in which ITCs may be performed and can provide useful information, these analyses present with several limitations. An ITC requires inference or extrapolation from known results to situations in which a study has not been done and, therefore, the validity of estimates obtained from ITCs may be questionable. Significant differences may exist between trials that compare one treatment to a control and trials that compare another treatment to the same control. For instance, the two sets of trials may be characterized by differences in patient characteristics, and such heterogeneity between patients may result in a different effect linking the treatment of interest. Differences in the length of follow-up, measurement of outcomes, and diagnostic criteria may also yield invalid results. Also, ITCs that include old trials may be based on data that does not represent current clinical practice and results generated from such analyses would not be observed in the present-day clinical setting.

To this end, the importance of direct evidence cannot be overemphasized. If a comparison between two treatments is of relevance and direct evidence does not exist, then investigators should plan a randomized controlled trial. In the event that an RCT cannot be conducted, investigators may resort to indirect comparisons and interpret, with caution, results based on such analyses. An understanding of the different methods and procedures available for making indirect treatment comparisons, of their limitations, and of the circumstances under which they may provide valid results is useful for health care decision makers who face the option of performing an ITC or who rely on information generated through indirect evidence.

1.2 Objectives

The first objective of this project was:

- to identify and review the different methods available for making indirect treatment comparisons (Chapter 2).

Additional objectives using the Bucher indirect treatment comparison approach were:

- to derive general methods and procedures for effect measures of discrete and continuous outcomes within complex networks of evidence. (Chapter 3)
- to assess the distributional properties of the indirect estimates using simulations (Chapter 3)
- to develop a user-friendly program for conducting indirect treatment comparisons for the methods and procedures derived. (Chapter 4)
- to illustrate the application of the empirically derived distributional properties of the indirect estimates and the program by applying it to examples selected from the literature. (Chapter 5).

The objective for the various methods identified was:

- to illustrate the application of the various methods for indirect treatment comparisons (Chapter 6).

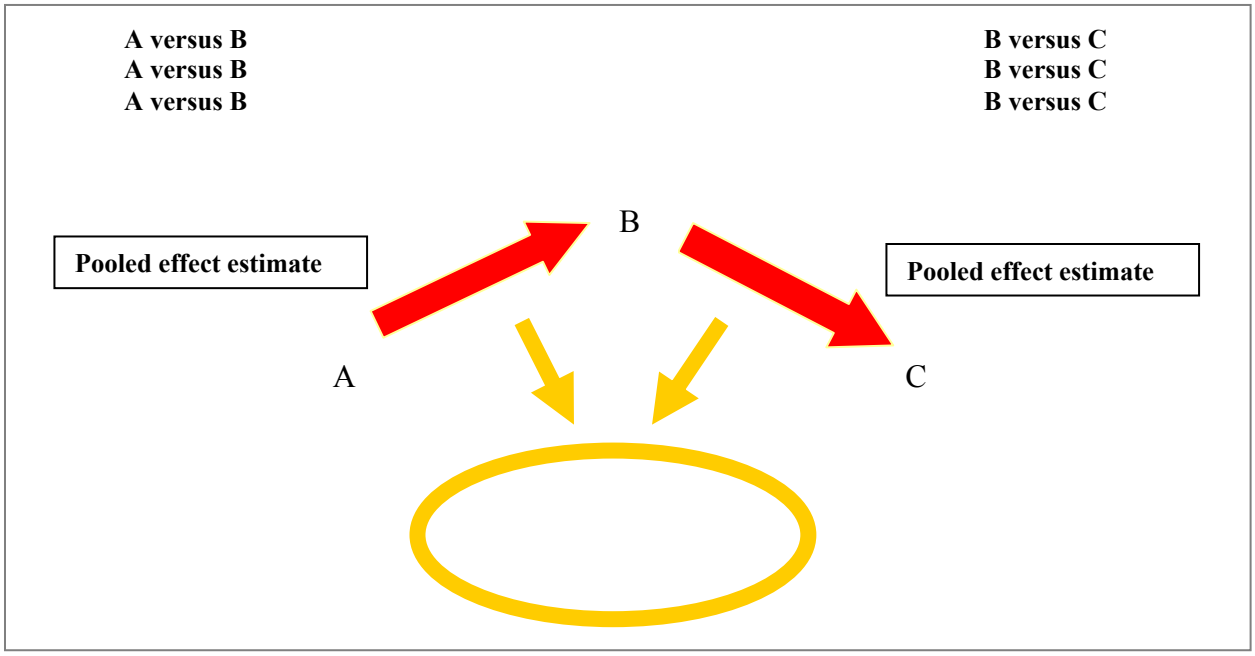
2 METHODS FOR MAKING INDIRECT COMPARISONS

The first objective was to identify and review the different methods available for making indirect treatment comparisons. To this end, a literature search was carried out to identify common methodologies that have been proposed for conducting indirect treatment comparisons. In theory, systematically applying selection criteria to information retrieved from a comprehensive literature search strategy (i.e., a systematic literature review) would provide the most complete set of references for articles that have discussed an ITC methodology. Due to a scarcity of appropriate indexing terms, a formal systematic search strategy to identify articles of all methods would be difficult to conduct; however, a comprehensive search was conducted and the more commonly reported methodologies were then selected for review.

2.1 Bucher Indirect Treatment Comparison

An indirect method proposed by Bucher et al. in 1997 has been a central article for considering indirect treatment comparisons in meta-analyses of RCTs for discrete data.[3] This model was developed with the odds ratio (OR) as the measure of treatment effect, and was specifically designed for the indirect comparison of A versus C when direct evidence of A versus B and B versus C was available. Using standard meta-analysis, the overall effect measure for A versus B is calculated as the usual weighted average of the individual effect measures of the included *g* studies, for example, and the association of B versus C is based on the usual weighted average of included *h* studies. The indirect estimate of A versus C is based on the paired comparisons of the direct estimates as follows: the sum of the logarithm of the summary effect measure for each direct estimate is taken as the summary effect measure for the indirect estimate; and the test of the association is based on the chi-squared value for the overall association of A versus B plus the chi-squared value for the overall association of B versus C, with $g + h$ degrees of freedom (Figure 1). Specific formulae for the effect estimates and derivation of the test statistics are given by the authors.

Figure 1: Adjusted Indirect Treatment Comparison Method



Similar to the case in which data is pooled during a direct comparison, Bucher et al. noted that in the absence of an interaction between covariates describing various subgroups of subjects included in the meta-analysis, the estimate would be unbiased in large samples; however, in the presence of an interaction, the data should not be combined. The authors tested their methodology for indirect treatment comparisons by performing an indirect analysis between two treatments and comparing the effect estimate to the estimate obtained through trials that directly compared the two treatments. The indirect result was more pronounced than the direct estimate, and the following reasons were provided for the observed discrepancy:

- *Difference between the weights given to studies included in the direct comparison versus those included in the indirect comparison:* The variance of the ln (OR) using the indirect method may be twice the magnitude of the variance for the direct method. In the indirect method, the variance is made up of two parts: the variance for the studies comparing A versus B and the variance for the trials comparing B versus C. If all the studies in the indirect comparison were given the same weight "w", the variance would be $(g + h)/w$. However, in the event that a direct comparison between A and C included as many patients as were included in the studies used to obtain an indirect effect estimate, each arm of the trial would have two times the number of patients per group in comparison to the indirect analysis. As such, the variance would be $(g + h)/2w$ and the direct comparison would be two times more efficient.
- *Methodological differences between studies that evaluated one treatment with the common comparator and studies that assessed the other treatment with the comparator:* In one set of studies, the design of trials may have been methodologically inferior and, as a result, may have exaggerated the treatment effects.
- *Differences in the measurement of the outcome:* Trials that compared one of the treatments to the comparator may have detected the outcome more efficiently than trials which compared the other treatment to the comparator.

- *Efficacy of treatments may vary among subgroups of patients:* If one subgroup of patients was over-represented in trials that included one of the treatments of interest in relation to trials for the other therapy, the difference in the magnitude of effectiveness between the two treatments could be more pronounced.

2.1.1 Assumptions

- The principal assumption of the model proposed by Bucher et al. is that the relative efficacy of a treatment is the same in all trials included in the indirect comparison. In the absence of a direct comparison between two treatments, A and C, an indirect estimate of the treatment effect can be obtained using results from trials that compared treatment A to treatment B and trials that compared treatment C to B, if the effect of treatment A observed in the trials comparing A to B would be the same had treatment A replaced treatment C in trials comparing B to C. As outlined by Song et al.,[4] if treatments A, B, and C were antibiotics, but A and C were compared in trials that included bacteria sensitive to both A and C, while B and C were compared in trials that included bacteria sensitive to B but resistant to C, then the method proposed by Bucher et al. could not be used to indirectly compare treatment A to treatment C because the results of one set of trials (studies comparing A versus C) is not generalizable to the other set of trials (studies comparing B versus C).
- This method assumes independence between pairwise comparisons, which is not found in three-arm trials.

2.1.2 Strengths

- As illustrated in Figure 1, the Bucher et al. method is an adjusted ITC in which the effect measure comparing two treatments within a RCT is used and not the individual results for each of the treatment groups. That is, the effect measure for each RCT comparing A versus B are combined and then compared to the corresponding combined effect measure based on the RCTs comparing B to C. This is in contrast to naïve ITC in which the randomization linking treatment groups is broken; treatment groups A are amalgamated, treatment groups B are amalgamated, and treatment groups C are amalgamated; and an effect measure is based on the amalgamated groups. An adjusted indirect comparison treatment method utilizes the magnitude of the effect measure reported in randomized controlled trials that separately compared each of two treatments to a common comparator. When an adjusted indirect comparison of two treatments is made, the strength of randomization is partially maintained and any differences between treatments that are observed through an indirect comparison are less likely due to patient differences unrelated to the treatment.

2.1.3 Limitations

- This model for indirect treatment comparisons can only be applied to data generated from two arm trials. The formulae provided by Bucher et al. are applicable only when there is no correlation between the pairwise comparisons. Since the estimates obtained from a three-arm trial are correlated, the Bucher method cannot be used.
- Only the effect measure OR was considered.
- Only the simple indirect treatment comparison involving three treatments (i.e., A versus B, B versus C) was considered.

2.1.4 Example

Coomarasamy et al.[5] used the Bucher et al. method in an indirect comparison of two tocolytics, nifedipine and atosiban, prescribed for the management of pre-term labour**.** The medications had been suggested by the Royal College of Obstetricians and Gynaecologists (UK) as alternatives to β-agonists because of their favourable adverse event profile, as well as their efficacy in the reduction of morbidity and mortality. Although evidence indicated that both nifedipine and atosiban had superior efficacy compared to β-agonists, Coomarasamy et al. performed an indirect comparison between nifedipine and atosiban because they had not been evaluated in a head-to-head trial.

For the outcome "reduction in neonatal respiratory distress syndrome," the authors determined that there was a significant difference between the two drugs (OR 0.55 [95% CI: 0.32, 0.97]) favouring nifedipine. Compared to atosiban, nifedipine was also associated with a greater rate of delay of delivery by 48 hours, but this difference was not significant. The authors did not have data for other outcomes of interest.

2.2 Lumley Network Meta-analysis for Indirect Treatment Comparisons

Lumley[6] has developed an indirect treatment comparison technique, known as network meta-analysis, to compare two treatments in the situation where an indirect comparison between two treatments of interest can be obtained through more than one common comparator or linking treatment. For instance, consider a setting where there is interest in performing an indirect comparison between treatment A and treatment B. If trials have separately compared treatment A to C, treatment B to C, treatment A to D, and treatment B to D, Lumley's method allows investigators to incorporate results from trials in which the common comparator was C, as well as trials in which the common comparator was D (i.e., more than once common treatment can be used to conduct an indirect comparison between two treatments). Network meta-analysis allows one to determine the amount of agreement between the results obtained when different linking treatments are used. Lumley has indicated that if the indirect comparison between two treatments yields the same result, regardless of which common comparator is used, there is a greater likelihood that the indirect treatment comparison represents the true relationship between the interventions; on the other hand, if there is a discrepancy in the results, "incoherence" exists, and Lumley has provided mechanisms to measure this incoherence.

Another situation in which the network meta-analytic approach may be of interest exists when an indirect comparison between two treatments can occur through "multiple paths", which require indirect comparisons within indirect comparisons. As an illustration, an assessment of the relative efficacy of A and E can occur through at least three paths (Figure 2, a to c). The first path (Figure 2, a) involves the construction of a model which:

- compares D to F through E; and
- compares D to G through the indirect treatment comparison between D and F and a direct treatment comparison between F and G.

In the second path (Figure 2, b), a model can be constructed to:

- compare D to B through C;
- compare B to G through A;
- compare D to G through the indirect comparison between D and B and the indirect comparison between B versus G.

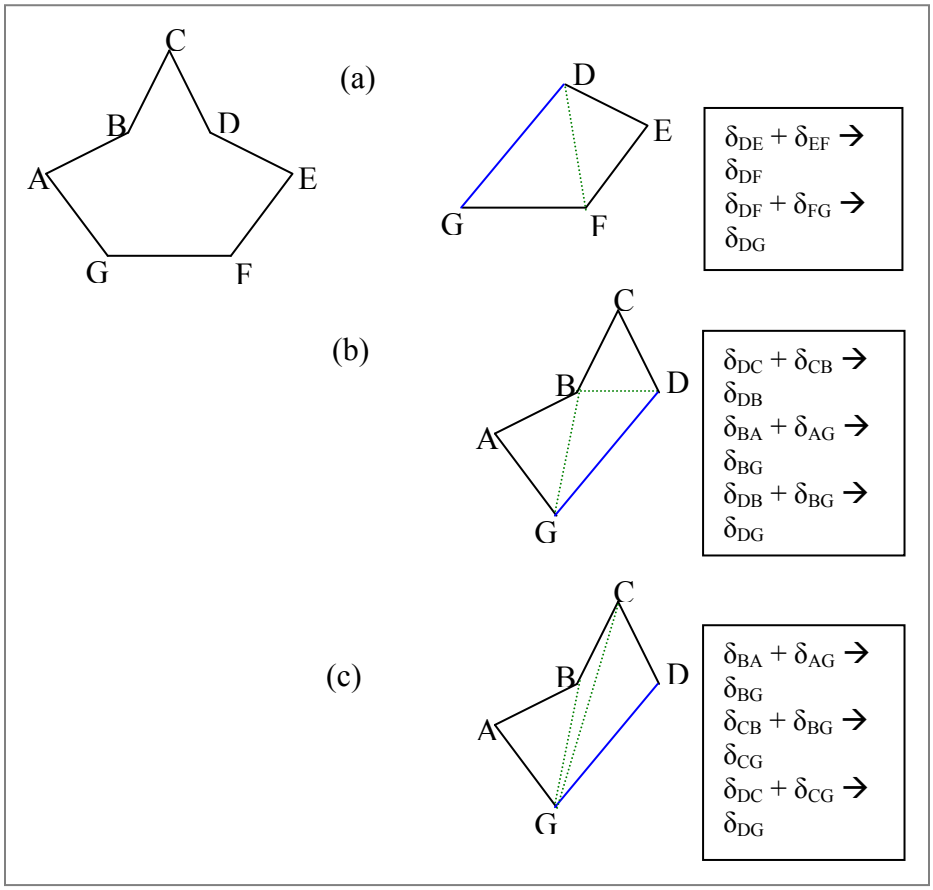
A third path (Figure 2, c) can be used to compare D versus G by considering treatment C. In this setting, an indirect comparison between D and G can be obtained through:

- the comparison of C to G through the indirect comparison between G versus B, as above, and a direct treatment comparison between B versus C;
- the comparison of D to G through the indirect comparison between G versus C and a direct comparison between C versus D.

In the network meta-analysis method, when more than one path can be used to compare two interventions, each path forms a component of a larger network and weights are assigned to different paths. The degree of agreement between the effect estimates obtained through different paths can be determined. The degree of agreement between paths is termed the "incoherence" of the network, and is incorporated in the calculation of the 95% confidence interval (CI) for the indirect estimate.

In the event that a limited amount of direct evidence is available for the comparison of two treatments, the network meta-analysis approach allows the direct evidence to be incorporated into the network model as well.

Figure 2: Closed Loop Network of Pathways for an Indirect Comparison in Network Meta-Analysis



The formal model[6] is

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

where:

Y_{ijk} is the treatment difference estimate from the kth RCT comparing treatment i and j;

σ_{ijk}^2 is the standard deviation error of Y_{ijk};

μ_i is the average effect of treatment i;

η_{ik} is a random effect with variance τ^2 representing the difference between the average effects of treatment i and j;

ξ_{ij} is a random effect with variance ω^2 representing a change in the effect of treatment i when it is compared to treatment j.

It should be noted that:

- η_{ik} random effects capture the heterogeneity of treatment effect;
- ξ_{ij} random effects capture the inconsistency of pairs of treatments. To combine different treatment comparisons, the effect of treatment i should be the same no matter what it is compared against (i.e. ξ_{ij} is close to 0), and ω^2 is called the incoherence of the network.

2.2.1 Assumptions

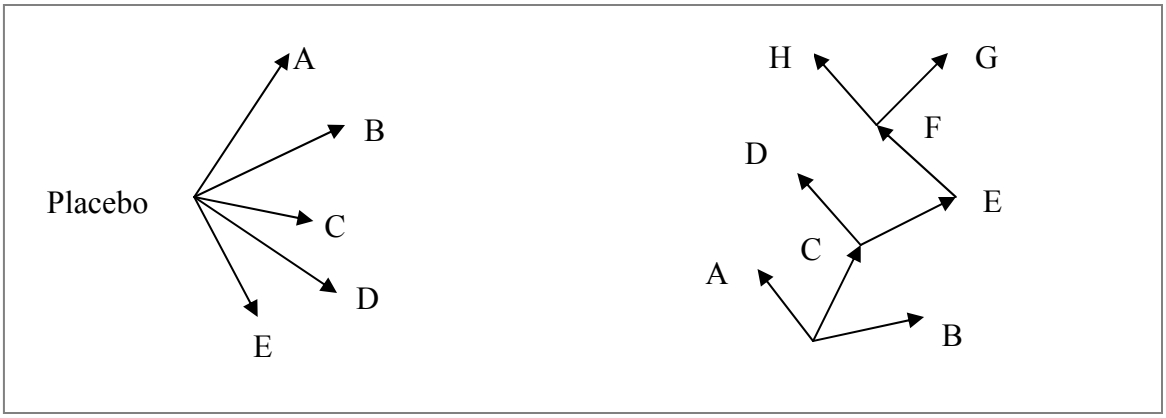
- The fundamental assumption underlying the network meta-analysis is that the comparison between two interventions, A and E (for example), will occur through a closed loop. The concept of a closed loop is best illustrated through the consideration of diagrams that depict the pathway of comparisons involved in the indirect comparison of two treatments (Figure 3). In a network diagram, a solid line is constructed between two treatments, A and F (for example), if they have been compared directly. The indirect comparison between the two treatments A and E follows a closed loop design if a solid line connecting all of the treatments between A and E can be drawn. A solid line can be replaced by a dotted line if the comparison between two interventions, A and B (for example), has been derived using indirect evidence and is then used in a pathway (network of comparisons) for the indirect comparison of two other treatments, such as A and E.

  A closed loop design is necessary for calculating the estimate of "incoherence," which is then used to construct a 95% confidence interval for the indirect estimate.

  Pathways that follow a star design or a ladder design cannot be used in the network meta-analysis (Figure 3). Such designs cannot quantify the amount of incoherence in a network of comparisons.

- Similar to the method proposed by Bucher et al., in order for the results of a network meta-analysis to be valid, the effect of any given treatment included in the model should be exchangeable across the other studies used to perform the network meta-analysis.

Figure 3: Open Loop Networks of Pathways for the Comparison of Competing Interventions



2.2.2 Strengths

- As for the method proposed by Bucher et al., the network meta-analysis is an adjusted indirect treatment comparisons model based on treatment effect measures observed in randomized controlled trials and, as such, partially preserves the randomization of study groups in the trials from which data are used.
- In the event that more than one comparator can be used to perform an indirect comparison, through the use of pathways, network meta-analysis can incorporate each comparator in a single model to arrive at an indirect estimate of treatment effect. The amount of agreement in the results obtained from the different paths for the indirect comparison can then be quantified.
- Although further investigation is needed, Lumley has indicated that the combination of both direct and indirect evidence may result in a narrower confidence interval than that which would be obtained if the relative efficacy of two treatments was based only on the limited direct evidence.[6]

2.2.3 Limitations

- Network meta-analysis does not involve a method to account for correlations that may exist between different effect estimates when they are obtained from a single multi-armed trial. Although a random effects model — in which the same random effect is applied to each treatment arm — can be used, this is not considered to be an optimal solution. Bayesian modelling has been proposed for situations in which trials with multiple treatment groups are included in the network meta-analytic model.[7] Nonetheless, the use of a Bayesian approach to appropriately model random effects in multi-armed trials has also been questioned due to its complexity and subsequent concern of the sensibility of conclusions drawn from such an approach.[8]
- Different paths may involve overlap, as illustrated by the presence of a comparison between A and B in both paths used to compare A with E (Figure 2). When the amount of incoherence is estimated, the network meta-analysis method does not provide a technique to account for any overlap that may exist. Because overlap cannot be accounted for, in situations where the

same comparison is performed in different paths, the estimated inconsistency will be less than the true amount. This is because the inconsistency between the result of a comparison, and the result of the same comparison in a different path, should be zero since the same data set is used each time.

2.2.4 Example

The network meta-analysis method was performed in an indirect comparison of the relative efficacy of various antihypertensive medications on the development of incident diabetes amongst patients with hypertension.[9,10] The authors cited several reasons for performing the indirect comparison. Specifically, although traditional meta-analyses have concluded that direct inhibitors of the renin-angiotensin system are effective in the prevention of incident diabetes, no comparison had been made between angiotensin-converting enzyme (ACE) inhibitors and angiotensin-receptor blockers (ARBs), both of which directly inhibit the renin-angiotensin system. Further to this, in traditional meta-analyses that assessed the effect of antihypertensive therapies in the prevention of incident diabetes, a comparison was made between a specific drug class and "any other treatment".[9,10] Because "any other treatment" amalgamates various different drug classes, such meta-analyses present with significant heterogeneity.

In order to compare ARBs with ACE inhibitors and to perform a comparison between only two drug classes, Elliot and Meyer performed a network meta-analysis.[9] In the analysis, various comparisons were made:

1. ARBs were compared to ACE inhibitors
2. each of the ACE inhibitors, ARBs, calcium channel blockers (CCBs), beta blockers, and a diuretic were compared to placebo
3. each of ACE inhibitors, ARBs, CCBs, beta blockers, and placebo were compared to diuretic.

The results were based on data from trials which consisted of a long-term follow-up, and which documented the number of cases of incident diabetes post-therapy. For the comparison of ACE inhibitors and ARBs, the network meta-analysis relied solely on indirect evidence. Because clinical trials had directly compared all of the ACE inhibitors, ARBs, CCBs, beta blockers, and a diuretic to placebo, for comparison (ii), the network meta-analysis included both direct and indirect evidence. Similarly, for comparison (iii) — in which the diuretic was the referent drug — the network meta- analysis included both direct and indirect evidence.

The results of the network meta-analysis indicated that there was no statistically significant difference between ACE inhibitors and ARBs in the risk of incident diabetes. For comparison (ii), diuretic (OR 1.34 [95% CI: 1.12, 1.60]) and beta-blocker therapy (OR 1.25 [95% CI: 1.05, 1.48]) were associated with significantly greater risk of incident diabetes than placebo. For comparison (iii), ARBs (OR 0.62 [95% CI: 0.51, 0.77]), ACE inhibitors (OR 0.67 [95% CI: 0.57, 0.79]), CCBs (OR 0.79 [95% CI: 0.67, 0.92]), and placebo (OR 0.75 [95% CI: 0.63, 0.89]) were associated with a lower risk of incident diabetes. The result between beta blocker therapy and diuretic was not statistically significant. The investigators also calculated the degree of incoherence in the model $\omega = 0.054$ indicating a reasonable level of coherence. Lumley in his paper identified an incoherence value of 0.001 to be small (i.e., good coherence) and 0.381 to be large (i.e., poor coherence).[6]

2.3 Models for Multi-parameter Synthesis and Consistency of Evidence

The confidence profile method (CPM) is a category of techniques used to conduct both direct and indirect treatment comparisons. The methodology was first proposed by Eddy et al.,[11] and later Ades[12] extended the CPM's methodology for indirect comparisons.

Analyses conducted in the CPM are based on Bayesian inference. In the Bayesian framework for the analysis of epidemiologic data, when a result for a parameter of interest is obtained, it presents itself in the form of a distribution, rather than a point estimate.[13] Additionally, before actual data is used to obtain information about a parameter, a mathematical model is constructed and includes a term to quantify prior knowledge about the parameter of interest. Prior knowledge about a parameter of interest is also presented in the form of a distribution and is called the "prior distribution."[13] In the event that there is no prior information known about the parameter of interest, the mathematical model consists of a "non-informative prior."[13] The model incorporates both the prior distribution and the actual data to generate an estimate for a given parameter. The result, presented as a distribution, is known as the "posterior distribution." The combining of evidence to generate an indirect estimate of treatment effect involves the addition of posterior distributions, rather than the addition of point estimates. Within the CPM structure, two techniques have been indicated for the generation of indirect evidence: i) intermediate outcomes; and ii) technology families. One or both of these mechanisms may be required to assess the relative efficacy of treatments.

Intermediate Outcomes: This method is used when interest lies in comparing the effect of two treatments on clinical endpoints or health outcomes, but the available data separately relates the effect of the treatments to intermediate endpoints or surrogate outcomes, and the effect of those intermediate endpoints on health outcomes. In this regard, the available data does not directly relate the interventions to clinical endpoints. Eddy et al.[11] developed formulas to combine the posterior distributions obtained from the two types of data. The combination of posterior distributions results in a parameter that represents the effect of the intervention on a clinical endpoint.

Technology Families: This method is used to compare two treatments, referred to as technologies, that have been compared to a common comparator, but not directly to each other. Eddy et al. have developed formulas to combine the posterior distributions that are generated when each of the treatments is compared to the control. The situation in which this methodology is used may be viewed as a Bayesian equivalent of the setting for which Bucher et al.[3] developed an indirect approach to combining evidence. Eddy et al. described influence diagrams to represent estimation problems in the CPM and explained their application to the pairwise comparisons that are performed within the technology family framework (Figure 4). In an influence diagram, information within a square represents experimental data, parameters of interest are enclosed in circles, and arrows specify the course of influence. $\theta_{K,i}$ denotes the value of the outcome parameter estimated for treatment K in trial i, where i=1,2,3…; $\varepsilon_{J-K,i}$ represents the value for the effect measure that estimates the difference between two treatments, J and K, in

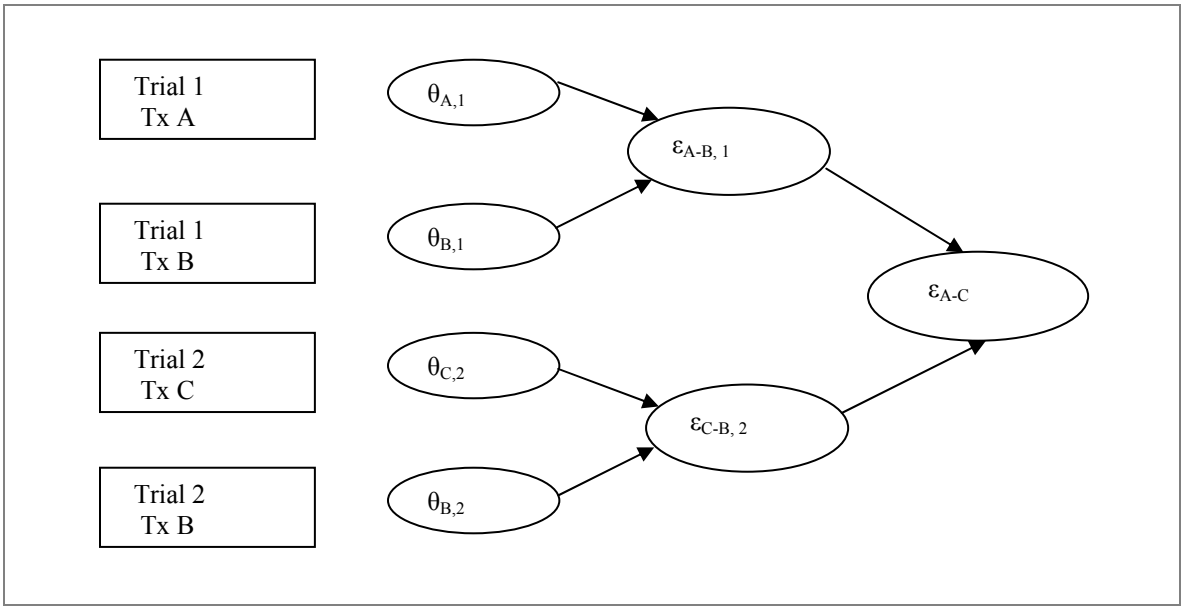
trial i. In an influence diagram, an arrow would point from $\theta_{K,i}$ to $\varepsilon_{J-K,i}$ because the value of $\varepsilon_{J-K,i}$ depends on $\theta_{K,i}$.

Ades extended the CPM model for indirect treatment comparisons and cited several reasons for its further development. Ades noted that the mathematical formulae outlined in the CPM framework assumed an equal study effects model in which the baseline probability of developing an intermediate outcome was the same across all trials. The CPM model also assumes a fixed treatment effect. Ades proposed a random study effects model in which, although the effect of treatment on an intermediate outcome is fixed, the baseline probability of the outcome is drawn from a normal distribution. Ades also extended both models in order that they may be used to combine direct and indirect evidence. Ades compared the two models by applying his model to a data set that was originally analyzed by Eddy et al. Three model-checking statistics were generated to determine the goodness of fit of each model, namely: the posterior mean deviance pD, which measures model fit for each parameter estimated in the model, and values greater than 1 indicate that the model fits the parameter poorly; posterior predictive value p(ext)%, which represents the probability of obtaining a more extreme result than that which is observed; and the conditional predictive ordinate (CPO), which indicates the probability of the observed result while considering the model and the rest of the data. Cross-validatory predictive checking was also performed and its purpose was to quantify whether or not different sources of data that may be chosen to perform an indirect comparison are similar enough to be validly combined. Checking the similarity of studies has been referred to as checking for "evidence consistency." Details of the model-checking statistics used by Ades and cross-validatory predictive checking can be found in the literature.[14]

2.3.1 Assumptions

- The method developed by Eddy et al. assumed a fixed study-effects and fixed treatment-effects model. Ades proposed a hierarchical model that assumes random study-effects and a fixed treatment-effects model.
- A primary assumption underlying both models is that it is valid to combine the different sources of data that have been selected for the indirect comparison. Validity is obtained when the studies are similar enough to each other such that the effect of a treatment is the same across all trials included in the comparison. This assumption also underlies the network meta-analysis approach and the methodology proposed by Bucher et al.
- For the analysis of intermediate outcomes, the method proposed by Eddy et al., and further extended by Ades, assumes that a clinical and causal relationship exists between the surrogate and clinical endpoints.
- The cross-validatory predictive checking method for evidence consistency requires the availability of direct evidence and can only determine whether or not the indirect sources of data can be validly combined with direct evidence. In the absence of direct evidence, cross-validatory predictive checking cannot be performed to determine whether or not the selected sources of data can be validly combined to perform an indirect comparison.

Figure 4: Influence Diagram for an Indirect Comparison Using the Confidence ProfileMethod



2.3.2 Strengths

- As noted by Ades, preserving the randomized nature of randomized controlled trial data was necessary to generate valid estimates of treatment efficacy.
- The model-checking statistics used by Ades are practical methods to determine whether or not direct and indirect sources can be combined to perform an indirect comparison.

2.3.3 Limitations

- The models described by Eddy et al. do not require that data from both arms of an RCT are used when the study is included in an indirect comparison. In the event that data from only one arm of an RCT is analyzed, the evidence from that trial equates to non-randomized evidence.
- Although appropriate use of both models requires that the validity of combining different sources of data be ascertained, this assumption was not explained or considered by Eddy et al. and was not always satisfied in his applications of the model.
- A lack of consistency between the direct and indirect estimates, as determined by cross-validatory predictive checking, does not necessarily mean that the selected sources of data for an indirect comparison cannot be combined. Rather, a discrepancy between direct and indirect results may indicate that the results of the trials used in the indirect comparison are not generalizable to those of the direct comparison, or vice versa. It is important for investigators to determine whether or not all sources of data are measuring the same elements and whether those elements are of clinical significance to the investigators.
- The complexity of the models may limit their use.

2.3.4 Example

Simple published examples in which investigators have used the methods outlined by Eddy et al.[11] and Ades[12] could not be found. For illustrative purposes, the aforementioned analysis conducted by Eddy et al. and later re-analyzed by Ades will be described.

Eddy et al.[11] used the CPM for intermediate outcomes, as well as technology families, in the comparison of tissue-type plasminogen activator (t-PA) and conventional care (CC) on one-year survival in patients with acute myocardial infarction. Three forms of evidence were used in the analysis. The first form of evidence included three RCTs which provided data on pairwise comparisons between t-PA, CC, and streptokinase (SK). The TIMI trial evaluated t-PA versus SK, the trial by Collen evaluated t-PA versus CC, and the Kennedy-R study evaluated SK versus CC. The trials evaluated an intermediate outcome, reperfusion. The second form of evidence consisted of data from the SK arm of the Kennedy-S RCT which evaluated the impact of reperfusion on a clinical endpoint, one-year survival. The third form of evidence was a meta-analysis of 20 trials that evaluated the relative efficacy between SK and CC on one-year survival.

Eddy et al. used the three sources of evidence to generate a posterior distribution for the relative efficacy of t-PA versus CC on the one-year survival outcome among patients with acute myocardial infarction. Using the equal study effects with fixed treatment effects model, Eddy et al. first performed an analysis of intermediate outcomes. Evidence from the TIMI study and the SK arm of the Kennedy-S study was combined to obtain an indirect estimate for the efficacy of t-PA versus SK with one-year survival as the primary outcome. An analysis of technology families was then carried out by combining the results of the intermediate outcomes analysis with data from meta-analysis of the 20 trials. A posterior distribution was generated for the comparison of t-PA versus CC for one-year survival.

Ades[12] re-analyzed the data set for a comparison between the effect of SK and CC on one-year survival. Ades applied the model used by Eddy et al., as well as the random study effects with fixed treatment effects model. Data from the Kennedy-R study was combined with data from the Kennedy-S study in an analysis of intermediate outcomes. The comparison provided an indirect estimate of the efficacy of SK versus CC on one-year survival. The models were also used to combine the direct evidence available through the 20 trials meta-analysis with the result of the intermediate outcomes analysis. Ades computed the three statistics for checking the model to determine the goodness of fit for each model and checked for evidence consistency. Using cross-validatory predictive checking, the direct evidence was excluded from the models and the indirect results were compared with the direct results of the meta-analysis of the 20 trials.

The model-checking statistics and cross-validatory predictive checking indicated that of the two models, the random study effects model fit the data with a greater degree of adequacy than the equal study effects model. Ades noted that the data set contained limited information on variance parameters, which were assigned weakly informative prior distributions. However, when the precision of the variance was increased, by assigning a more informative prior distribution to the variance parameter, the goodness of fit for the fixed effects model was greater than that for the random effects model. Ades indicated that for the dataset, assigning a greater amount of precision to the variance estimate also increases the amount of evidence consistency when the

indirect evidence is compared to the direct results obtained from the meta-analysis of the 20 trials.

In the fixed effects model, the SK-CC mean difference in survival was 0.038 (SE = 0.010). This value was the same as that obtained in the meta-analysis of the 20 trials. In the random effects model, the MD was 0.040 (SE = 0.010). Although the results of the analysis indicated that both models generate a similar estimate for the comparison of SK versus CC, for the comparison between SK and CC in an "average population" the random effects model presented a greater SE. The results of the cross-validatory predictive checking revealed that in the absence of the direct evidence, for the comparison between SK versus CC, the result of both the fixed and random effects models were similar and the 95% intervals contained the value observed in the 20-trials meta-analysis. For the comparison between SK and CC in an average population, the difference (0.827) observed in the 20-trials meta-analysis was just inside the lower bound of the 95% interval (difference 0.872 [95% interval: 0.825, 0.91]) for the fixed effects model.

2.4 Mixed Treatment Comparisons

Indirect comparisons may be performed through the mixed treatment comparison (MTC) method. The method is used in various situations, namely:

- To evaluate the relative efficacy between two treatments (A and B, for example) by combining both direct and indirect evidence. When the results of a direct comparison between two treatments are inconclusive, and an indirect comparison can be made, the direct and indirect sources of evidence are combined to strengthen the result of the direct comparison.
- When many competing interventions are available for the same medical condition, direct pair-wise comparisons are often only available for a subset of the treatments. MTC can be used to simultaneously perform indirect comparisons among the treatments for which results from direct comparisons do not exist. In this way, it is possible to obtain effect estimates for all possible pairwise comparisons and to rank the efficacy of these various competing treatments.

Another term for MTC is mixed treatment meta-analysis (MTM).[15] Because of its similarity to the model proposed by Lumley, MTC has also been referred to as "network meta-analysis."[15,16]

2.4.1 Combination of direct and indirect evidence in mixed treatment comparisons

Lu and Ades have described the statistical methods for performing MTC in a Bayesian framework.[17] Earlier attempts to outline methodologies for mixed treatment comparisons can be found in the literature.[18,19]

This section provides an overview of Lu and Ades' method for MTC. In the case where k ($k \geq 2$) treatments have been compared, treatment one represents the reference treatment in a model for the meta-analysis of k treatment comparisons. When there is interest in comparing two treatments (let us say 1 and 2, by way of illustration), but there is limited evidence from RCTs for the comparison, then an MTC can be performed if there is sufficient evidence for the

comparison of each of treatment 1 and 2 to a third treatment, denoted 3. Treatment 3 acts as the reference treatment, and each of treatments 1 and 2 can be compared to treatment 3. d_{31} and d_{32} represent the relative efficacy of treatments 1 and 2 compared to reference treatment 3, respectively. Thereafter, the relative efficacy of treatment 2 versus 1 can be given by $d_{12} = d_{32} - d_{31}$, where d_{32} is the numerator treatment and d_{31} is the denominator treatment. A random treatment effects model simultaneously evaluates, on the logit scale, the relative efficacy of each treatment in comparison to the reference treatment 3.

The general formulation for the random effects models proposed by Lu and Ades[17] is the following:

$$\delta_{i1k} = \text{logit}(p_{ik}) - \text{logit}(p_{i1}),\ k = 2,\ldots,K$$

$$\delta_{i1k} \sim N(d_{k1}, \sum)$$

where i is the trial number, the index number 1 refers to the reference treatment, and k represents the treatment which is compared to the reference treatment 1. The δ_{i1k} are the relative treatment effects with respect to reference treatment 1, all being on the logit scale. δ_{i1k} is considered a sample from a bivariate normal distribution whose mean is $d_{k1,}$ and $\sum$ is the variance-covariance matrix that accounts for the correlation observed between the different groups in a multi-armed trial. In this way, the MTC model can be applied to two-arm and multi-arm trials. The structure of $\sum$ varies, depending on whether or not the model assumes homogeneous treatment variance, heterogeneous treatment variance, or a random effects covariance structure. It should be noted that d_{k1} are referred to as "basic parameters" and d_{23} is an example of a "functional parameter," which is determined through the combination of two basic parameters.[11] Because the model assumes random treatment effects, both the basic and functional parameters represent the mean of the relative random treatment effect for any comparison of interest. For the MTC model, Lu and Ades specified five variations to explore the effects of different assumptions. The models are, as follows:

- random multivariate treatment effects, unconstrained (fixed) baselines, homogeneous treatment variance
- random multivariate treatment effects, unconstrained baselines, heterogeneous treatment variance
- random multivariate treatment effects, random effects baselines, homogeneous treatment variance
- random multivariate treatment effects, random effects baselines, heterogeneous treatment variance
- random multivariate treatment effects, random effects baselines, random effects covariance structure.

When applied to a dataset, the MTC should be performed using Markov Chain Monte Carlo models in WinBUGS software.[17]

2.4.2 Evaluating evidence consistency

Lu and Ades extended various aspects of their initial methodology.[20] A model has been developed to estimate the inconsistency between direct and indirect evidence. The model has also been extended to specify more than one reference treatment when performing a MTC.

Salanti et al.[16] have provided a detailed discussion of this methodology. A summary of the information provided by Lu and Ades[20] and Salanti et al.[16] is subsequently described.

Consider three treatments A, B, C and the parameters μ_{AB}, μ_{AC}, and μ_{BC}, each of which has been estimated in a randomized controlled trial. These parameters can be estimated through a MTC, because they form a "consistency model". A consistency model is one in which a parameter, say μ_{BC}, can be estimated directly through a BC comparison, and indirectly through a comparison between AC and AB. In the indirect comparison, one of the treatments, in this case treatment A, is chosen as the reference to which both B and C are compared. An indirect comparison between B versus C can then be obtained through combining μ_{AB} and μ_{AC}. When μ_{BC} is measured indirectly, it is referred to as a "functional parameter" that is obtained through the combination of two "basic" parameters, μ_{AC} and μ_{AB}.

In a consistency model, μ_{BC} is expressed by the following relation:

$$\mu_{BC} = \mu_{AC} - \mu_{AB}$$

The assumption underlying a consistency model is that there is no discrepancy between the direct estimate of μ_{BC} and the estimate obtained by combining the basic parameters on the right side of the consistency relation. Because of this assumption, the consistency relation can be expressed in two other, equivalent ways:

$$\mu_{AB} = \mu_{AC} - \mu_{BC}, \quad \mu_{AC} = \mu_{AB} + \mu_{BC}$$

A property of a consistency model is that the pairwise comparisons in the consistency relation are part of a network that takes the shape of a closed loop.

As mentioned previously, in some situations, direct and indirect evidence for all pairwise comparisons, J_{max}, among T treatments of interest may not be available. A direct evaluation of only a subset of J_{max} comparisons, denoted J, may have been performed in RCTs. Furthermore, the J comparisons may not involve a common reference treatment. In this case, if there are direct estimates for T-1 basic parameters, the MTC method can be used to calculate all possible pairwise contrasts. However, effect estimates that are only based on indirect comparisons cannot be expressed in a consistency relation. A consistency relation implies that its parameters form a closed loop and can therefore be estimated through direct and indirect evidence. When $J \geq T$, at least one consistency model can be specified. The number of independent consistency relations is determined through the following formula: $J - (T - 1)$.

The MTC analysis proceeds by choosing a reference treatment for each trial included in the analysis. For instance, let us suppose there is interest in obtaining estimates of all possible ORs for treatments 1, 2, 3, and 4 and sets of studies have directly compared d_{12}, d_{13}, d_{14} d_{23}, d_{24} and d_{34}. A baseline is chosen for each trial, such that $G_{(1)}$, $G_{(2)}$, and $G_{(3)}$ represent those trials for which 1, 2, and 3 were chosen as baseline. The formal statistical model is thus,

$$\text{logit}(p_{it}) = \begin{cases} \mu_{ib} & t = b; \quad b = 1,2,3,4 \text{ for i} \in G_{(X)}, X = b(i) \\ \mu_{ib} + \delta_{itb} & t > b; \quad t = 1,2,3,4; t \in k_{(i)} \end{cases}$$

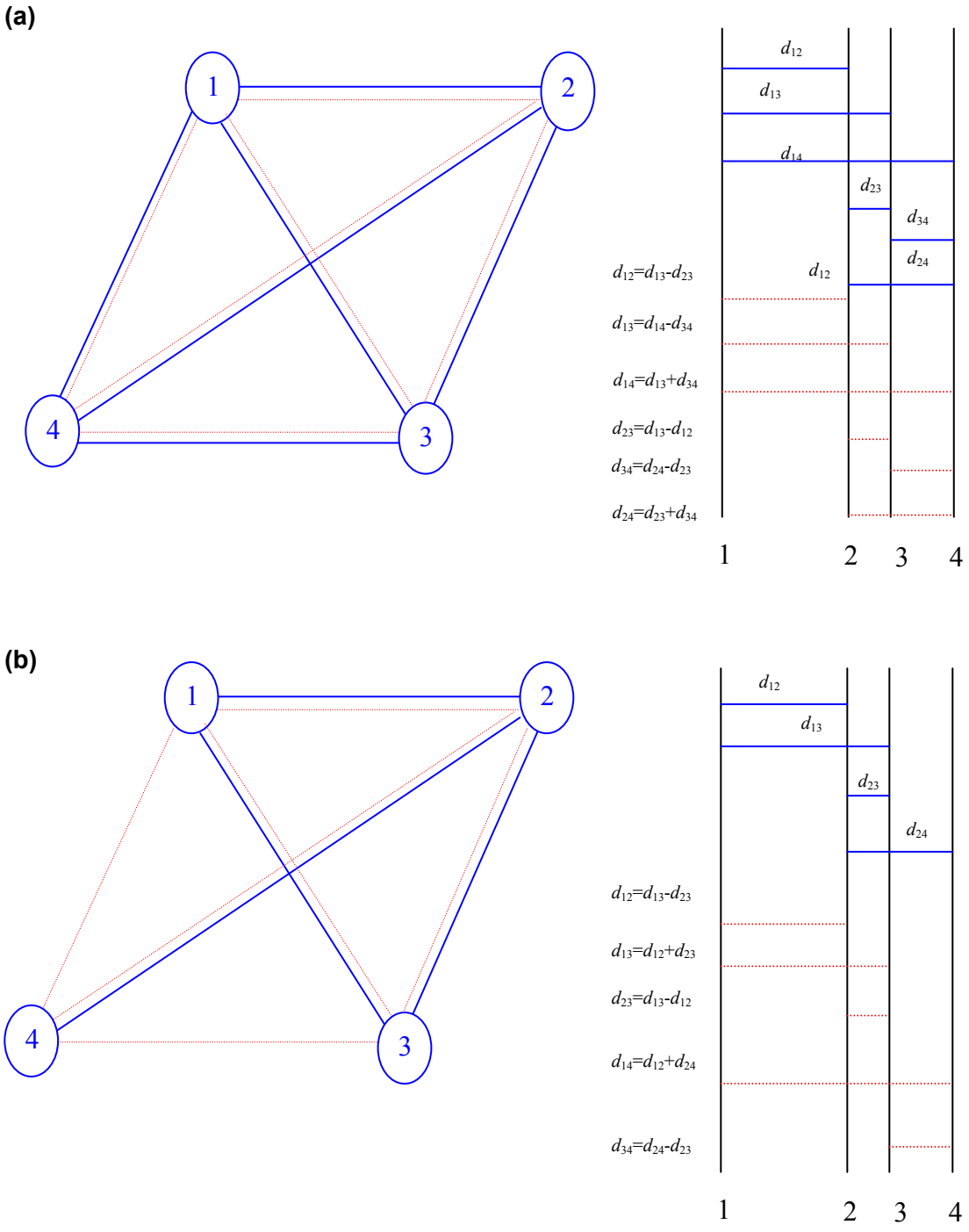
$$\delta_{ibt} \sim N(d_{bt}, \sigma_{bt}^2)$$

where:

- p_{it} is the probability of the event for treatment t in trial i;
- μ_{ib} is the log odds of the event for the reference treatment b in trial i;
- δ_{itb} is the trial specific log odds ratio of treatment t relative to the reference treatment b ($t > b$ signifies that t is numerically after b);
- $G_{(X)}$ is the set of trials for which X is chosen as the baseline;
- $_{(i)}$ represents the set of treatments evaluated in trial;
- d_{bt} represents the mean of the distribution;
- $\sigma^2 = 0$ corresponds to a fixed effects model and $\sigma_{bk}^2 = \sigma^2$ represents a random effects model that assumes homogeneity of between-trial variation.

Visual illustrations of the methodology for MTC have been described by various sources.[15,20,21] Figure 5 presents pairwise comparisons in an MTC framework for T=1,2,3,4.

Figure 5: Network of Trials for Mixed-Treatment Comparisons*



* Solid blue lines represent treatment contrasts that have been measured directly. Dashed red lines represent indirect comparisons. In figure 5a, all functional parameters can be represented, determined in alternate ways, as follows: $d_{12} = d_{14}-d_{24}$, $d_{13} = d_{12}+d_{23}$, $d_{14} = d_{12}+d_{24}$, $d_{24} = d_{14}-d_{12}$, $d_{23} = d_{24}-d_{34}$, $d_{34} = d_{14}-d_{13}$

For effect parameters that can form a consistency relation, Lu and Ades have proposed a method to determine whether or not the consistency assumption is valid.[20] An extra term, *w*, can be added to the consistency relation, as follows,

$$\mu_{BC} = \mu_{AC} - \mu_{AB} + w$$

$$w \sim N(0, \sigma_w^2)$$

where:

- w is an inconsistency factor that represents the amount of inconsistency between a direct estimate of μ_{BC} and an indirect estimate of μ_{BC}
- σ_w^2 represents the inconsistency variance.

Inconsistency degrees of freedom (ICDF), denoted L, correspond to the number of independent loops in an evidence network for T treatments being analyzed through MTC. L is determined through the same formula that calculates the number of independent consistency relations — $L = J - (T - 1)$ — and represents the number of potential inconsistencies. The measurement of inconsistency is included in the hierarchical models proposed by Lu and Ades.[20] A description of the statistical model is beyond the scope of this report. Readers are referred to the original references.[16,20] Bayesian measures for determining the goodness of fit of a model and model criticism techniques can also be used to assess inconsistency.[16,20]

2.4.3 Other contributions to the mixed treatment comparison framework

Salanti et al.[16] have discussed various alternative approaches to the parameterization of consistency models and the detection of inconsistency. The authors have also introduced the concept of network asymmetry. Asymmetry refers to the extent to which specific treatments or specific comparisons are represented more heavily than others in a network of treatment comparisons. Salanti et al. have proposed two methods to evaluate asymmetry. The first method determines whether some comparisons tend to occur more frequently than expected by chance alone. A second method determines whether some treatments occur more frequently than others in the network. Specific formulas to assess network asymmetry can be found in the original reference.

The results of the mixed treatment comparison methodology can be used to determine the probability that each treatment has the greatest efficacy for a specified outcome. Caldwell et al.[21] determined these probabilities in a mixed treatment comparison for the effect of various treatments in subjects with myocardial infarction. The probabilities are determined through Markov Chain Monte Carlo methods. Computational details can be found in Caldwell et al.'s example. In another example, Jansen et al.[22] used the mixed treatment comparison method to evaluate the relative efficacy of treatments for the management of chronic insomnia. For the different treatments considered, the probability that each was the best was determined.

2.4.4 Assumptions

- The true effect of a given treatment is the same in all trials included in the indirect comparison.
- The homogeneous variance models assume that the variance of the true treatment response rate, on the logit scale, is the same for all treatments evaluated in a single study.
- The models assume that the correlation terms in the variance-covariance matrix have the same value.
- All of the models assume that the event rates, on the logit scale, follow a multivariate normal distribution.

2.4.5 Strengths

- The MTC methodology can be used to perform an indirect comparison involving more than two interventions.
- Analyses are based on the pooling of effect estimates across trials rather than individual treatment groups.
- Results obtained through combining both direct and indirect evidence in the evaluation of two interventions may provide more precise estimates, as indicated by narrower confidence intervals, than results based on direct evidence alone.[4,17]
- MTC can be applied to data from trials with more than two treatment groups (i.e., multi-arm trial).

2.4.6 Limitations

- Because this method, like other Bayesian techniques, involves judgments in specifying prior distributions, these judgments may or may not be valid.
- The complexity of this method may limit its use.

2.4.7 Example

Caldwell et al.[21] consider the limitations of two standard pairwise meta-analyses. The data from these meta-analyses are used to perform a MTC to strengthen the results of the reviews.

The meta-analyses used to perform mixed comparisons evaluated the effects of different treatments in patients who had experienced a myocardial infarction. In one of the standard meta-analyses, Boland et al.[23] summarized the results of two- or three-arm studies performing different pairwise comparisons among six different thrombolytics: streptokinase, alteplase, accelerated alteplase, streptokinase+alteplase, reteplase, and tenecteplase. The review provided summary estimates of the pairwise comparisons for which direct evidence was available in the literature. Because a summary estimate for all possible pairwise comparisons could not be calculated, Caldwell et al. indicated that it was not possible to rank the treatments from best to worst. In the second meta-analysis, Keeley et al.[24] summarized the results of trials making various pairwise comparisons between primary percutaneous transluminal coronary angioplasty (PCTA) and streptokinase, alteplase, or accelerated alteplase. The investigators concluded that PCTA had a greater efficacy than thrombolytics; however, Caldwell et al. indicated that the thrombolytics were grouped together, and that it is not clear if a class effect for those medications can be assumed.

Using data from the two meta-analyses, Caldwell et al. performed a MTC analysis. The authors implemented a fixed-effects and random-effects logistic regression model to perform the analysis. The authors also calculated the probability that each treatment was best based on the premise of using mortality reduction as the outcome of interest. The MTC analysis generated a total of 21 odds ratios that represented the effect estimates for all possible pairwise comparisons. The results of the comparisons for which only direct evidence was available were in agreement with the results based on combining direct and indirect evidence. For example, for the comparison between accelerated alteplase and reteplase, results from each method indicated that there was no statistically significant difference between the treatments (OR_{direct} 1.02 [95%CI: 0.64, 1.02]; $OR_{MTC\ fixed\ effects}$ 1.05 [95% CI: 0.94, 1.17]; $OR_{MTC\ random\ effects}$ 1.04 [95% CI: 0.81,1.28]). The confidence intervals for results based on the MTC random effects model were generally wider than those for the MTC fixed effects model. The results of the MTC analysis illustrated an increase in precision when all available evidence was combined. For instance, for the comparison between accelerated alteplase and PCTA, direct evidence was inconclusive. The results from both fixed and random effects MTC analysis consisted of narrower confidence intervals. Based on all available evidence, Caldwell et al. concluded that there was a statistically significant difference between the two treatments. The probability that each treatment is most effective was determined. The results indicated that PCTA was associated with the highest probability for being the best treatment.

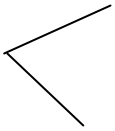
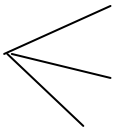
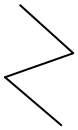
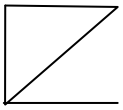
2.4.8 Software for conducting the mixed treatment comparison method

The WinBUGs software code for conducting the MTC methods for both fixed and random effects models are available on the website https://www.bris.ac.uk/cobm/research/mpes/mixed-treatment-comparisons.html.[25]

In particular, programs are available for performing a fixed effects model, a random effects model with no correlation structure for multi-arm trials, a random effects model with a correlation structure for three-arm trials, and a random effects model for multi-arm trials. Examples of applying these programs are given in Chapter 6.

2.5 Summary of the Methods

Table 1 indicates the various different networks of evidence that can be analyzed by the indirect comparison methods. These networks represent the star, ladder and closed and partially closed-loop designs. The MTC method can be used to obtain measures of effect for each of the indicated patterns. The network meta-analysis method proposed by Lumley[6] can compare treatments in a network geometry that contains at least one closed loop. The adjusted indirect comparison method proposed by Bucher et al.[3] can be used to evaluate the effect of treatments that form a simple star design. The Bucher method has been proposed to perform indirect comparisons when direct evidence is not available, and the method is not applicable to the closed loop pattern. For the other designs, the Bucher method can be used to determine the indirect evidence of the pair-wise contrasts that have not been directly compared in the star, ladder and network with one closed loop designs.

Table 1: Network Patterns that the Various Indirect Treatment Comparison MethodsCan Process				
Pattern Description	**Network Pattern**	**Indirect Comparison Method**		
		Bucher Method	**Network Meta-analysis**	**Mixed Treatment Comparison**
Simple star		✓	—	✓
Star		— (Pairwise contrasts)	—	✓
Ladder		✓	—	✓
Closed loop		—	✓	✓
Network with at least one closed loop		✓ (Pairwise contrasts)	✓	✓

When MTC or network meta-analysis is used to evaluate the evidence network depicted by the closed loop pattern, the methods can simultaneously combine the direct and indirect evidence and can evaluate the incoherence of the closed loop. The variance parameter w^2 from Lumley's model is equivalent to inconsistency variance σ_w^2 estimated in the MTC models.[20] However, the two methods will calculate different values for treatment effects because of differences in the way that inconsistency is evaluated. As indicated by Salanti et al.[16] and Lu and Ades,[20] in the network meta-analysis approach, the number of incoherence terms ξ_{ij} is equal to the number of different comparisons. In the MTC framework, the number of inconsistency terms is equal to the number of different independent closed loops.

The MTC model described by Lu and Ades measures the relative efficacy of treatments using the log OR effect measure. Various investigators have performed the MTC for other effect measures. For instance, Vandermeer et al. considered direct and indirect evidence to evaluate the relative efficacy between benzodiazepines and nonbenzodiazepenes[26] based on mean differences for five of their clinical outcomes and risk difference on the adverse event outcome. Jansen et al.[22] have outlined an MTC model to be applied to continuous outcomes.

Various approaches for indirect treatment comparisons have been reviewed. The mixed treatment comparison approaches by Lu and Ades are elegant, but require information that may not be available. The challenge of Lumley's network meta-analysis is that it needs a data-based assessment of trial consistency; therefore, it requires information from a large number of different treatment comparisons. When analyzing a network of comparisons, the inconsistency of the network needs to be considered, as well as between-trial heterogeneity and sampling error. Large inconsistencies rule out a meta-analysis, and small inconsistencies should add uncertainty to the results. The inconsistency of the network can only be assessed for a closed loop of treatments, with more loops allowing for better diagnosis of consistency. Estimating inconsistency will be reliable to the extent that the trials in these closed loops are similar to other trials. In addition, consistency cannot be assessed for a star design comparing everything to placebo, or for a ladder design where new treatments are always compared to current standard.

The attractiveness of the Bucher approach is that it has been designed to apply with minimal information to the common indirect treatment comparison involving a simple star design: using the direct comparisons X versus A and X versus B with the common comparator link "X," to yield an indirect comparison of A versus B. The Bucher approach has not been shown to work for the ladder design. That is, we have X versus C, C versus E, E versus F, F versus G and we want to use the comparator links "C", "E" and "F" to yield an indirect comparison of X versus G. In Chapter 3, we extend the Bucher approach to apply to the ladder design and, as well, extend the approach for the effect measures relative risk, risk difference, hazard ratio, and mean difference.

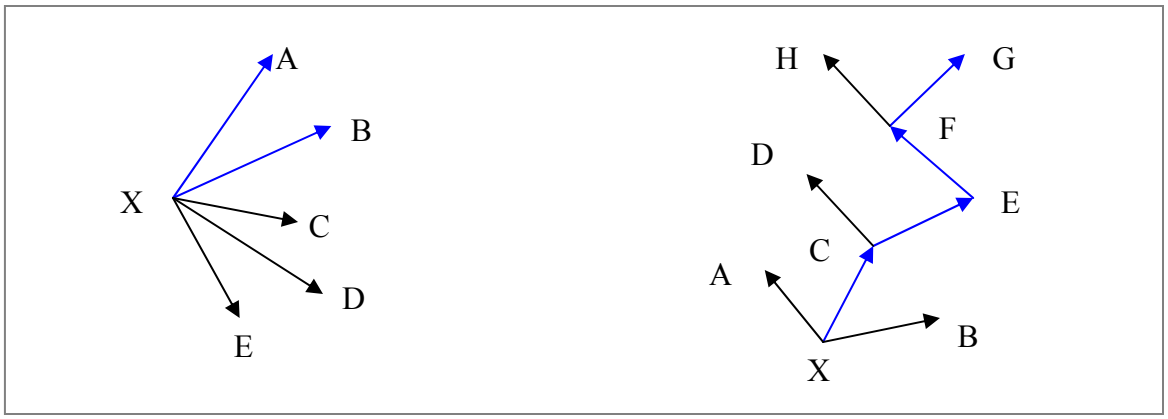
3 EXTENSION OF THE BUCHER APPROACH AND EMPIRICAL EVALUATION OF THE ESTIMATORS

As noted in Chapter 2, the attractiveness of the Bucher approach is that it has been designed to apply with minimal information to the common indirect treatment comparison involving a simple star design: using the direct comparisons X versus A and X versus B with the common comparator link "X" to yield an indirect comparison of A versus B (Figure 6).

The second objective was to derive general methods and procedures for effect measures of discrete and continuous outcomes within complex networks of evidence for the Bucher approach. In this chapter (Section 3.1), we extend the Bucher approach to another common comparator link involving a ladder design in which several direct comparisons can be linked by common comparators. That is, we have X versus C, C versus E, E versus F, F versus G and we want to use the comparator links "C", "E" and "F" to yield an indirect comparison of X versus G (Figure 6). Also, in this chapter, we extend the Bucher approach to different measures of association (i.e., relative risk, risk difference, hazard ratio, and mean difference).

The third objective was to assess the distributional properties of the indirect estimates using simulations for the Bucher approach. Later in this chapter (Section 3.2), simulations were conducted to determine the frequency distribution and the bias and mean square error for the various measures of association.

Figure 6: Open Loop Networks of Pathways — Star and Ladder Designs



3.1 Extension of the Bucher Approach

For discrete outcomes, the indirect odds ratio approach by Bucher et al.[3] was extended to more complex networks of evidence ("ladder" design) involving several direct comparisons. This generalized approach was also considered for the relative risk, hazard ratio, risk difference, and mean difference. Two fundamental propositions underlie the estimation and hypothesis testing procedures for these indirect measures, and details are provided in Appendix A. These propositions were applied to specific estimators of association, namely:

Indirect Comparisons of Odds Ratios (OR)
Indirect Comparisons of Relative Risks (RR)
Indirect Comparisons of Hazard Ratios (HR)
Indirect Comparisons of Risk Differences (RD)
Indirect Comparisons of Mean Differences (MD).

The summary of the indirect point and confidence interval estimators and test of association for these measures of association is provided in Table 2 when k treatments $A_1, A_2, \cdots, A_k$ are considered. The notation is such that for consecutive pairs of treatmen, the direct estimator of the measure of association of interest (RR, as an example) for treatment A_i and A_{i+1} is denoted by $RR_{A_iA_{i+1}}$. Full statements of the estimators and tests for each of these measures of association are provided in Appendix A.

Using empirical approaches, the extent and direction of biases associated with these indirect estimators were explored, with the goal of developing specific guidelines for using these methods (Section 3.2). A "reviewer-friendly" program was developed and made available to facilitate the evaluation of indirect evidence for reviewers (Chapter 4).

Table 2: Indirect Point and Confidence Interval Estimators and Test of Association			
Measure of Association	**Indirect Estimator**	**Indirect 100(1-α)% Confidence Interval Estimator**	
		In Terms of Variance	**In Terms of Confidence Limits***
Log odds ratio	$\sum_{i=1}^{k-1}\ln(OR_{A_iA_{i+1}})$	$\sum_{i=1}^{k-1}\ln(OR_{A_iA_{i+1}}) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(\ln(OR_{A_iA_{i+1}}))}$	$\sum_{i=1}^{k-1}\ln(OR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2}$
Odds ratio	$\prod_{i=1}^{k-1}OR_{A_iA_{i+1}}$	$\exp\left(\sum_{i=1}^{k-1}\ln(OR_{A_iA_{i+1}}) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(\ln(OR_{A_iA_{i+1}}))}\right)$	$\exp\left(\sum_{i=1}^{k-1}\ln(OR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2}\right)$
Log relative risk	$\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}})$	$\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(\ln(RR_{A_iA_{i+1}}))}$	$\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2}$
Relative risk	$\prod_{i=1}^{k-1}RR_{A_iA_{i+1}}$	$\exp\left(\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(\ln(RR_{A_iA_{i+1}}))}\right)$	$\exp\left(\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2}\right)$
Log hazard ratio	$\sum_{i=1}^{k-1}\ln(HR_{A_iA_{i+1}})$	$\sum_{i=1}^{k-1}\ln(HR_{A_iA_{i+1}}) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(\ln(HR_{A_iA_{i+1}}))}$	$\sum_{i=1}^{k-1}\ln(HR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2}$
Hazard ratio	$\prod_{i=1}^{k-1}HR_{A_iA_{i+1}}$	$\exp\left(\sum_{i=1}^{k-1}\ln(HR_{A_iA_{i+1}}) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(\ln(HR_{A_iA_{i+1}}))}\right)$	$\exp\left(\sum_{i=1}^{k-1}\ln(HR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2}\right)$
Risk difference	$\sum_{i=1}^{k-1}RD_{A_iA_{i+1}}$	$\sum_{i=1}^{k-1}RD_{A_iA_{i+1}} \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(RD_{A_iA_{i+1}})}$	$\sum_{i=1}^{k-1}RD_{A_iA_{i+1}} \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(ucl_{A_iA_{i+1}} - lcl_{A_iA_{i+1}})^2}$
Mean difference	$\sum_{i=1}^{k-1}MD_{A_iA_{i+1}}$	$\sum_{i=1}^{k-1}MD_{A_iA_{i+1}} \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(MD_{A_iA_{i+1}})}$	$\sum_{i=1}^{k-1}MD_{A_iA_{i+1}} \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(ucl_{A_iA_{i+1}} - lcl_{A_iA_{i+1}})^2}$
Test of Association			
For effect estimators (EE) ln(OR), ln(RR), ln(HR), RD, MD the test of association is: $\chi^2_{A_1A_k Indirect\ association} = \left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}}W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}}W_{A_{i'}A_{i'+1},j}\right)(EE_{A_iA_{i+1}} - EE_{A_{i'}A_{i'+1}})^2\right] \Big/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}}W_{A_iA_{i+1},j}\right)$		

HR=hazard ratio; MD=mean difference; OD=odds ratio; RD=risk difference
*ucl, lcl indicate upper confidence limit and lower confidence limit, respectively

3.2 Empirical Evaluation of the Estimators

A simulation was undertaken to determine the precision and accuracy of the indirect approach. For the case of k = 3 treatments, a simulated data set was created for each of three populations (A, B, C) and a simple random sample size of 100 was selected from each population to calculate direct estimates of the measure of association of interest relating A and B and for relating C and B, along with an indirect estimate of the measure of association for relating A and C. This process was repeated 1,000 times. The bias, variance, and mean square error (MSE) of the direct and indirect estimates were calculated and compared to determine the accuracy and precision of the indirect measures of association.

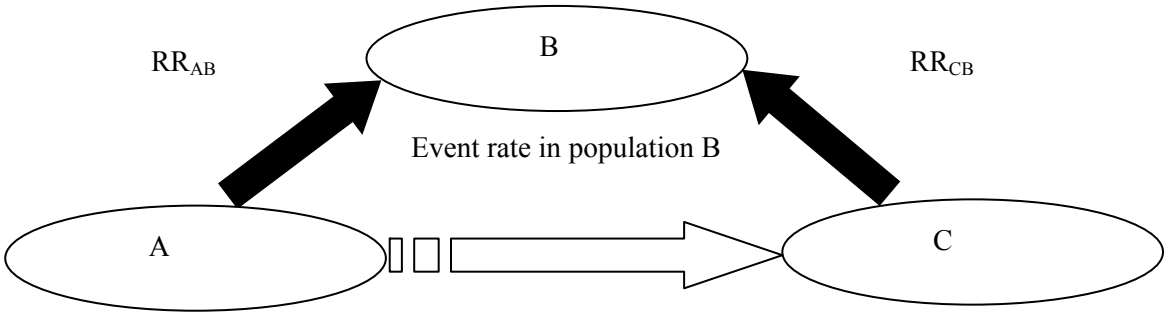
Bias is the expected difference between the estimator and the parameter to be estimated, and the MSE is the expected squared deviation between the estimator and this parameter. The MSE summarizes information about the bias and variance of the estimator under study.

This section provides a summary of the simulation results for each measure of association considered. Details of the simulation process and results can be found in Appendix B.

3.2.1 Simulation results for the relative risk

The simulation for RR followed a plan that would mimic the setting in which the indirect treatment comparison would be considered, as depicted in Figure 7. First, we need to have a common comparator (B); second, we would need to consider the RR relating treatment A and the common comparator B (RR_{AB}); and third, we would need to consider the RR relating treatment C and the common comparator B (RR_{CB}). The actual bias and MSE of the indirect treatment estimator depends on these values, as well as the likelihood of the event of interest in the common comparator group B (denoted by $P(E \mid B)$, where E denotes the event of interest).

Figure 7: Schematic of the Parameters Considered in the Simulation for Relative Risk

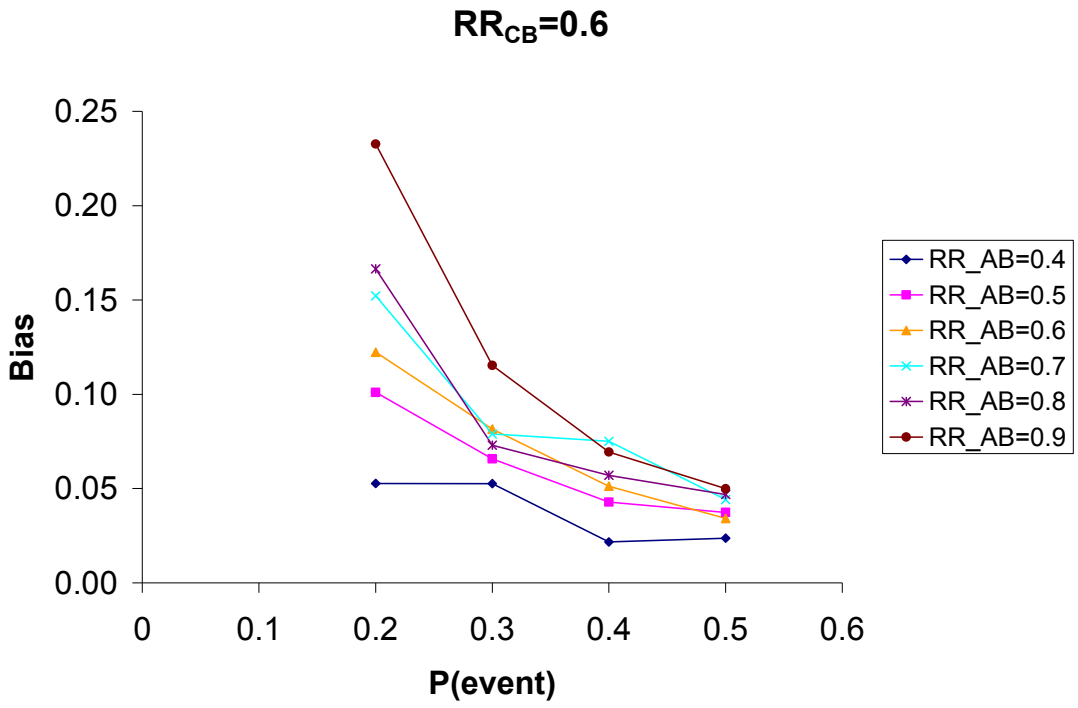


RR=relative risk

Details of the simulation results for the bias and MSE are given in Tables B.2.1 and Figures B.2.1 to B.2.4 in Appendix B. These tables and figures are instructive in providing information on the bias and MSE for the indirect treatment estimator of RR. In particular, Figures B.2.3 and B.2.4 provide graphs that can be used to quantify the degree of bias and MSE, respectively, when applying the indirect approach to a particular problem and will provide the user with an idea of

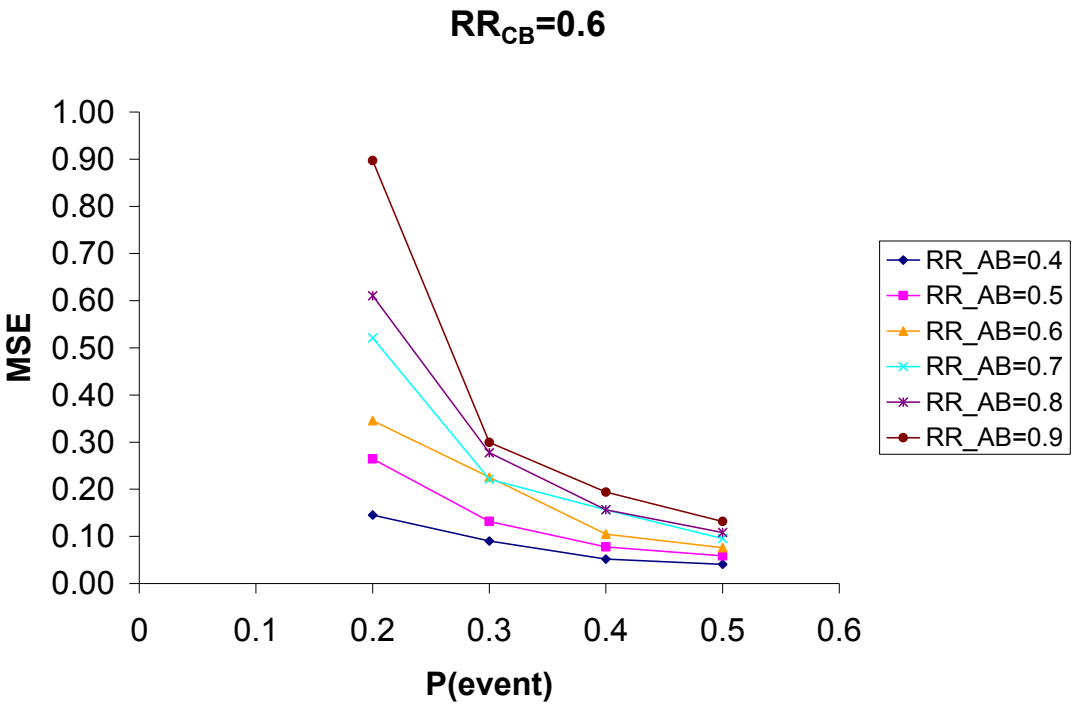
the degree of bias and precision of the resulting indirect treatment comparison. For example, consider Panel f of Figure B.2.3 and Panel f of Figure B.2.4 which are reproduced in Figures 8 and 9 respectively. If in a particular application $RR_{AB} = 0.7$, $RR_{CB} = 0.6$ and the likelihood of the event in the comparator population B was 0.2, then the bias of the indirect estimator is approximately 0.15 (see Figure 8) and MSE is 0.50 (see Figure 9).

Figure 8: Bias for Indirect Relative Risk Estimates



RR=relative risk

Figure 9: Mean Square Error for Indirect Relative Risk Estimates



MSE=mean square error; RR=relative risk

Several general observations regarding the bias and MSE for the direct and indirect estimators of the relative risk are provided in Table 3.

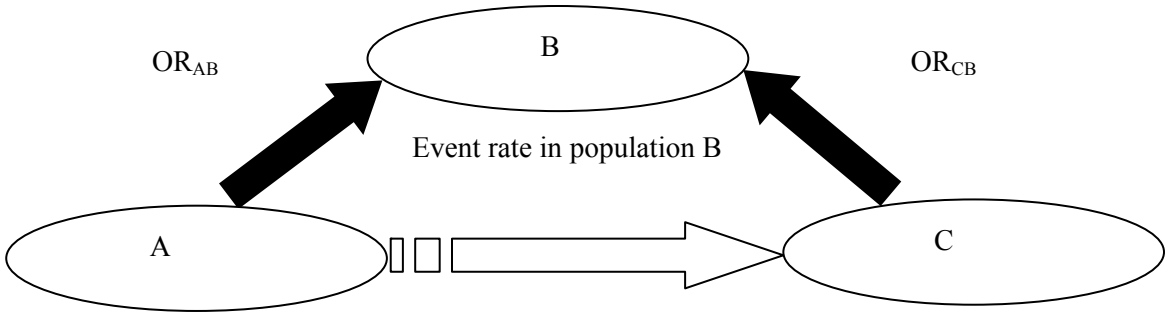
Table 3: Summary of Simulation Results for the Direct and Indirect Estimators of the Relative Risk			
Description of RR_{AB}	**Description of RR_{CB}**	**Description of Event Rate in Population B**	**General Observed Pattern for bias and Mean Square Error**
Bias			
Fixed	Fixed	Increasing to 0.5	Bias for both the direct and indirect estimators decrease
Fixed	Increasing to 1	Fixed	Bias for both the direct and indirect estimators decrease
Increasing to 1	Fixed	Fixed	Bias for both the direct and indirect estimators increase
Fixed	Fixed	Fixed	Bias for the indirect estimator is greater than for the direct estimator, particularly for event rates below 0.2
MSE			
Fixed	Fixed	Increasing to 0.5	MSE for both the direct and indirect estimators decrease
Fixed	Increasing to 1	Fixed	MSE for both the direct and indirect estimators decrease
Increasing to 1	Fixed	Fixed	MSE for both the direct and indirect estimators increase
Fixed	Fixed	Fixed	MSE for the indirect estimator is greater than for the direct estimator, particularly for event rates below 0.2

MSE=mean square error; RR=relative risk

3.2.2 Simulation results for the odds ratio

The simulation for OR followed a plan that would mimic the setting in which the indirect treatment comparison would be considered, as depicted in Figure 10. First, we need to have a common comparator (B); second we would need to consider the OR relating treatment A and the common comparator B (OR_{AB}); and third we would need to consider the OR relating treatment C and the common comparator B (OR_{CB}). The actual bias and MSE of the indirect treatment estimator depends on these values, as well as the likelihood of the event of interest in the common comparator group B (denoted by $P(E \mid B)$, where E denotes the event of interest).

Figure 10: Schematic of the Parameters Considered in the Simulation for Odds Ratio



OR=odds ratio

Details of the simulation results for the bias and MSE are given in Tables B.3.1 and Figures B.3.1 to B.3.4 in Appendix B. These tables and figures are instructive in providing information on the bias and MSE for the indirect treatment estimator of OR. Several general observations regarding the bias and MSE for the direct and indirect estimators of the odds ratio are provided in Table 4.

Table 4: Summary of Simulation Results for the Direct and Indirect Estimators of the Odds Ratio

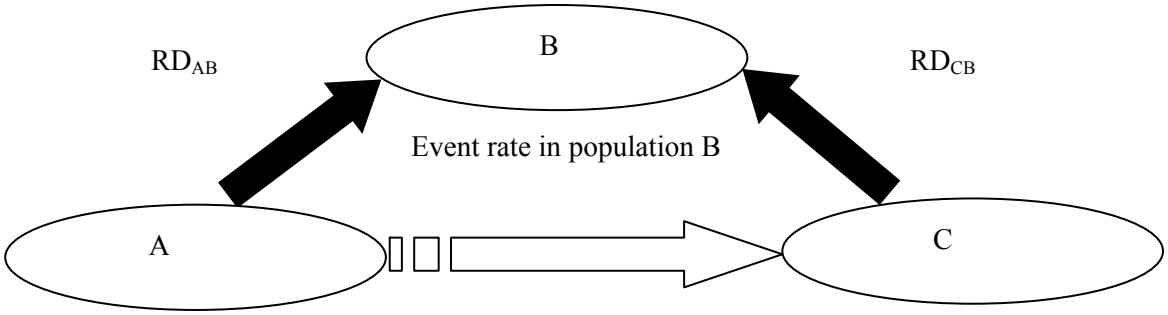
Description of OR_{AB}	Description of OR_{CB}	Description of Event Rate in Population B	General Observed Pattern for Bias and Mean Square Error (MSE)
Bias			
Fixed	Fixed	Increasing to 0.5	Bias for both the direct and indirect estimators decrease
Fixed	Increasing to 1	Fixed	Bias for both the direct and indirect estimators decrease
Increasing to 1	Fixed	Fixed	Bias for both the direct and indirect estimators increase
Fixed	Fixed	Fixed	Bias for the indirect estimator is greater than for the direct estimator, particularly for event rates below 0.2
MSE			
Fixed	Fixed	Increasing to 0.5	MSE for both the direct and indirect estimators decrease
Fixed	Increasing to 1	Fixed	MSE for both the direct and indirect estimators decrease
Increasing to 1	Fixed	Fixed	MSE for both the direct and indirect estimators increase
Fixed	Fixed	Fixed	MSE for the indirect estimator is greater than for the direct estimator, particularly for event rates below 0.2

OR=odds ratio; MSE=mean square error

3.2.3 Simulation results for the risk difference

The simulation for RD followed a plan that would mimic the setting in which the indirect treatment comparison would be considered, as depicted in Figure 11. First, there must be a common comparator (B); second, one would have to consider the RD relating treatment A and the common comparator B (RD_{AB}); and third, one would have to consider the RD relating treatment C and the common comparator B (RD_{CB}). The actual bias and MSE of the indirect treatment estimator depends on these values, as well as the likelihood of the event of interest in the common comparator group B (denoted by $P(E \mid B)$, where E denotes the event of interest.

Figure 11: Schematic of the Parameters Considered in the Simulation for the RiskDifference



RD=risk difference

Details of the simulation results for the bias and MSE are given in Tables B.4.1 and Figures B.4.1 to B.4.4 in Appendix B. These tables and figures are instructive in providing information on the bias and MSE for the indirect treatment estimator of RD. Several general observations regarding the bias, variance, and MSE for the direct and indirect estimators of the risk difference are provided in Table 5. Since the risk difference is an unbiased estimator, the simulation results for the bias are essentially zero and the variance and MSE results are the same.

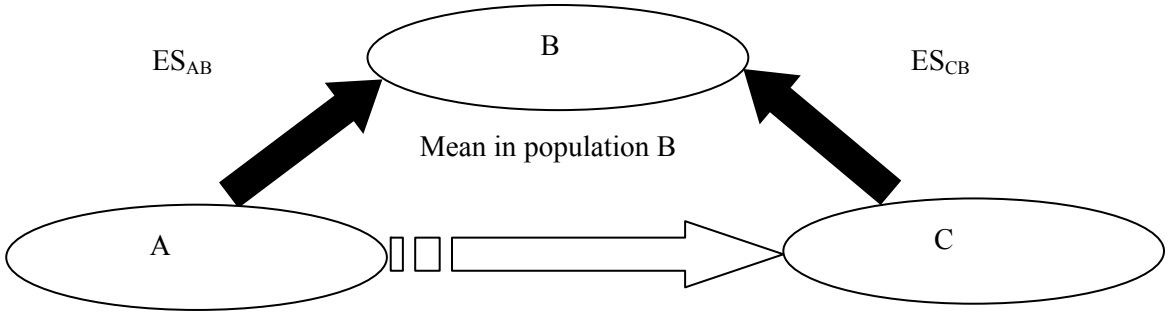
Table 5: Summary of Simulation Results for the Direct and Indirect Estimators of the Risk Difference			
Description of RD_{AB}	**Description of RD_{CB}**	**Description of Event Rate in Population B**	**Observed Pattern for Bias and Mean Square Error or Variance**
Bias			
Fixed	Fixed	Increasing to 1	Bias essentially zero
Fixed	Decreasing to -6	Fixed	Bias essentially zero
Decreasing to -6	Fixed	Fixed	Bias essentially zero
Fixed	Fixed	Fixed	Bias (absolute) for the indirect estimator is greater than for the direct estimator
MSE			
Fixed	Fixed	Increasing	MSE for the direct estimators increase; MSE for the indirect estimators have a fairly similar value regardless of the event rate
Fixed	Increasing	Fixed	No effect on the MSE for the direct estimators; MSE for the indirect estimators increase
Increasing	Fixed	Fixed	No effect on the MSE for the direct estimators; MSE for the indirect estimators increase
Fixed	Fixed	Fixed	MSE for the indirect estimators is consistently larger than the MSE for the direct estimators

MSE=mean square error; RD=risk difference

3.2.4 Summary of the simulation results for mean difference

The simulation for RD followed a plan that would mimic the setting in which the indirect treatment comparison would be considered, as depicted in Figure 12. First, there must be a common comparator (B); second, one would have to consider the MD relating treatment A and the common comparator B (MD_{AB}), which is expressed in terms of effect size $ES_{AB} = MD_{AB}/SD_{AB}$ where SD_{AB} is the standard deviation of MD_{AB}; and third, one would need to consider the MD relating treatment C and the common comparator B (MD_{CB}) expressed in terms of the effect size $ES_{CB} = MD_{CB}/SD_{CB}$ where SD_{CB} is the standard deviation of MD_{CB}. The actual bias and MSE of the indirect treatment estimator depends on these values, as well as the mean of the outcome of interest in the comparator group (M_B) and the coefficient of variation $CV_B = SD_B/M_B$ where SD_B is the standard deviation of the outcome in the comparator group B.

Figure 12: Schematic of the Parameters Considered in the Simulation for Mean Difference



Details of the simulation results for the bias and MSE are given in Tables B.5.1 and Figures B.5.1 to B.5.4 in Appendix B. These tables and figures are instructive in providing information on the bias and MSE for the indirect treatment estimator of MD. Several general observations regarding the bias, variance, and MSE for the direct and indirect estimators of the risk difference are provided in Table 6. Since the mean difference is an unbiased estimator, the simulation results for the bias are essentially zero and the variance and MSE results are the same.

Table 6: Summary of Simulation Results for the Direct and Indirect Estimators of the Mean Difference

Description of ES_{AB}	Description of ES_{CB}	Description of CV_B	Description of Mean in Population B	General Observed Pattern for Bias and Mean Square Error or Variance
Bias				
Fixed	Fixed	Fixed	Increasing	For both the direct and indirect estimators, the absolute value of the bias increases
Fixed	Increasing	Fixed	Fixed	For both the direct and indirect estimators, a similar value for the bias is observed, regardless of a change in ES_{CB}
Increasing	Fixed	Fixed	Fixed	For both the direct and indirect estimators, a similar value for the bias is observed, regardless of a change in ES_{AB}
Fixed	Fixed	Increasing to 0.5	Fixed	Bias increases for both the direct and indirect estimates
Fixed	Fixed	Fixed	Fixed	No general pattern observed
MSE				
Fixed	Fixed	Fixed	Increasing	For both the direct and indirect estimates, the MSE increases
Fixed	Increasing	Fixed	Fixed	For both the direct and indirect estimators, a similar value for the MSE is observed, regardless of a change in ES_{CB}

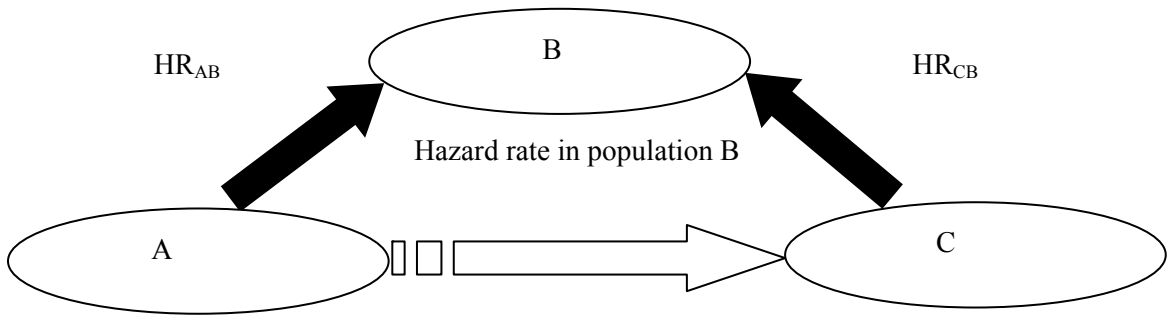
Table 6: Summary of Simulation Results for the Direct and Indirect Estimators of the Mean Difference				
Description of ES_{AB}	**Description of ES_{CB}**	**Description of CV_B**	**Description of Mean in Population B**	**General Observed Pattern for Bias and Mean Square Error or Variance**
Increasing	Fixed	Fixed	Fixed	For both the direct and indirect estimators, a similar value for the MSE is observed, regardless of a change in ES_{AB}
Fixed	Fixed	Increasing to 0.5	Fixed	MSE increases for both the direct and indirect estimators
Fixed	Fixed	Fixed	Fixed	The MSE of the indirect estimator is consistently larger than that for the direct estimator

MD=mean difference; MSE=mean square error

3.2.5 Simulation results for the hazard ratio

The simulation for HR followed a plan that would mimic the setting in which the indirect treatment comparison would be considered, as depicted in Figure 13. First, we need to have a common comparator (B); second we would need to consider the HR relating treatment A and the common comparator B (HR_{AB}); and third we would need to consider the HR relating treatment C and the common comparator B (HR_{CB}). The actual bias and MSE of the indirect treatment estimator depends on these values, as well as the hazard rate in the common comparator group B.

Figure 13: Schematic of the Parameters Considered in the Simulation for Hazard Ratio



HR=hazard ratio

Details of the simulation results for the bias and MSE are given in Tables B.6.1 and Figures B.6.1 to B.6.4 in Appendix B. These tables and figures are instructive in providing information on the bias and MSE for the indirect treatment estimator of HR. Several general observations regarding the bias and MSE for the direct and indirect estimators of the hazard ratio are provided in Table 7.

Table 7: Summary of Simulation Results for the Direct and Indirect Estimators of the Hazard Ratio			
Description of HR_{AB}	**Description of HR_{CB}**	**Description of Hazard Rate in Population B**	**Observed Pattern for Bias and Mean Square Error**
Bias			
Fixed	Fixed	Increasing	Bias for the direct estimator is generally zero regardless of the event rate; Bias for the indirect estimate, no pattern observed
Fixed	Increasing	Fixed	Bias for the direct estimator is generally zero; Bias for the indirect estimate, no pattern observed
Increasing	Fixed	Fixed	Bias for the direct estimator is generally zero; Bias for the indirect estimator, no pattern observed
Fixed	Fixed	Fixed	For any combination of the parameters, the bias of both the direct and indirect estimators is small
MSE			
Fixed	Fixed	Increasing	For both the direct and indirect estimators, generally, the MSE is constant regardless of the event rate
Fixed	Increasing	Fixed	For both the direct and indirect estimators, for any given HR_{AB}, the larger HR_{CB} then the smaller the MSE for both direct and indirect estimators
Increasing	Fixed	Fixed	For both the direct and indirect estimators, for any given HR_{CB}, the larger HR_{AB} then the larger the MSE for both direct and indirect estimators
Fixed	Fixed	Fixed	For any combination of the parameters, the MSE of the indirect estimator is consistently larger than that for the direct estimator

HR=hazard ratio; MSE=mean square error

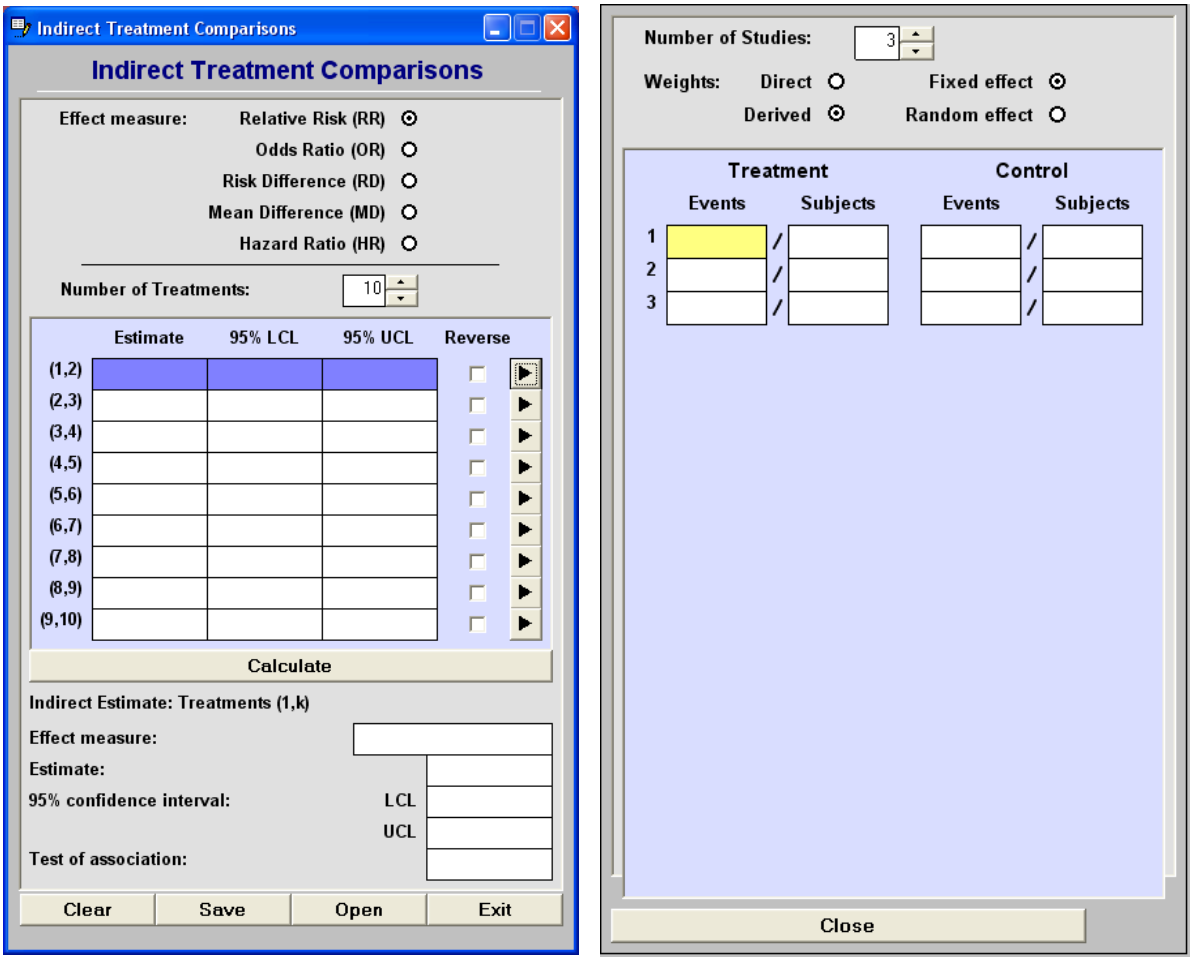
4 ITC PROGRAM

To meet the fourth objective of developing a user-friendly program for conducting indirect treatment comparisons for the methods and procedures derived for the Bucher approach, a program is described in this chapter for making the needed calculations.

4.1 Introduction

The ITC (Indirect Treatment Comparison) program has been developed in Visual Basic to assist with the various calculations associated with indirect comparisons that are described in Chapter 3. It consists of two screens, which are described in detail in Section 4.2. On the first screen, the effect measure of interest is identified, and information for each consecutive pair of treatments of interest is requested regarding the point estimate and 95% CI of the effect measure for each direct comparison involved in the indirect comparison. The resulting indirect comparison estimates for the effect measure and the 95% CI, as well as the p-value for the test of association corresponding to this effect measure, are provided. On the second screen, the weights needed for a specific direct comparison is requested in order to calculate the test statistic for the test of association. There are various formats in which the information to calculate these weights can be provided and are identified through the weight selections (direct versus derived; fixed versus random), and the specific information for each study involved in the direct comparison is then identified and requested.

SCREEN 1 **SCREEN 2**



4.2 Instructions

Input (Screen 1)

1. Check circle indicating effect estimate of interest:

Effect measure:	**Relative Risk (RR)**	O
	Odds Ratio (OR)	O
	Risk Difference (RD)	O
	Mean Difference (MD)	O
	Hazard Ratio (HR)	O

2. Select the number of treatments k (maximum 10):
   Number of Treatments: _____

3. For each consecutive pair of treatments, provide the direct estimates of the measure of association and the 95% lower and upper confidence limits. The order of entry of the treatment pairs must follow the exact sequence indicated with the bridging comparison groups linking the treatment pairs.

Treatment	**Estimate**	**95%LCL**	**95%UCL**
(1,2)	_____	_____	_____
(2,3)	_____	_____	_____
(3,4)	_____	_____	_____
.			
.			
.			
(k-1,k)	_____	_____	_____

4. For each treatment comparison, an option to reverse the order of a treatment comparison is provided. For example, if 1 = Treatment A, 2 = Treatment B, and 3 = Treatment C, then (1,2) is (A,B), and for (2,3) we can enter (C,B) and use the reverse option to switch it to (B,C). This option can be useful when B is placebo and results are given as the active treatment versus placebo; i.e., A versus Placebo (B) and C versus Placebo (B).

   Reverse

Input (Screen 2)

If the test of association is needed, then the weights used for the calculation of each weighted average estimate from Screen 1 are required for its calculation. This is the purpose of Screen 2.

1. For a direct treatment comparison (i, i+1) from Screen 1, press the arrow bar which will activate Screen 2 for inputting the weights used for the effect measure estimate for that direct treatment comparison.

2. For (i, i+1), indicate the number of studies upon which the estimate is based:
   Number of Studies: _____

3. The option is available to enter the weights directly. In particular, if the effect estimate is based on one study, then a single weight of 1 can be inputted.

4. The weights can also be computed from first principles, based on the frequencies (for RR, OR, and RD) or the standard errors (for MD), using either the fixed or random effects model.

Output (Screen 1):

Indirect Estimate: Treatments (1,k)
Effect measure: _____
Estimate: _____
95% confidence interval: **LCL** _____
UCL _____
Test of Association (p-value): _____

4.3 Worked Example

Osteoporosis is associated with important medical, social, and financial implications and its incidence is expected to increase significantly as the Canadian population ages. Many of the consequences of osteoporosis are potentially lessened through use of a number of non-pharmacological and pharmacological interventions. The oral bisphosphonate drugs— etidronate, alendronate and risedronate — have been introduced as pharmacological options for the primary and secondary prevention of osteoporotic fractures. We have conducted a systematic review assessing the clinical effectiveness of etidronate, alendronate, and risedronate in the primary and secondary prevention of osteoporotic fractures in postmenopausal women receiving these agents compared with untreated women over a follow-up period of at least one year.[27] A systematic literature search of the evidence from randomized placebo-controlled trials of each of the three drugs was conducted using a standardized Cochrane Collaboration approach to literature search, article selection, data extraction, and quality assessment. Clinical data analysis was conducted according to the methodology of The Cochrane Collaboration for systematic reviews and meta-analyses.

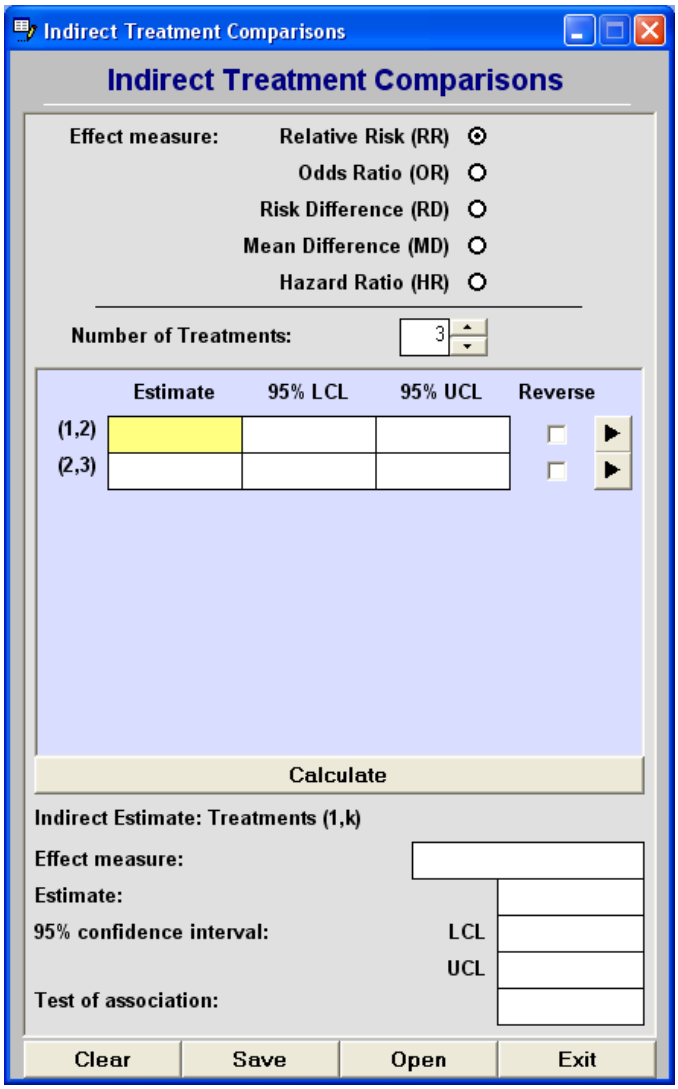
Considering the data available for the longest treatment duration in the trials and using the follow-up denominators for the number of patients in the trial, a detailed worked example will be considered. For this detailed worked example, the weighted relative risk effect estimates of fracture after treatment with the bisphosphonates alendronate and etidronate compared to placebo will be used to derive an indirect estimate. The indirect treatment comparison method will be used to evaluate the head-to-head comparison of alendroante to etidronate using the placebo as the bridging group in the one-step comparison (i.e., $k = 3$).

On the first screen, the effect measure of interest is identified and information for each consecutive pair of treatments of interest is requested regarding the point estimate and 95% CI of the effect measure for each direct comparison involved in the indirect comparison.

The effect measure of interest is the relative risk and this was selected by checking the circle corresponding to "Relative Risk (RR)".

There are three treatments involved in this indirect comparison (i.e., alendronate, etidronate, and placebo) and, therefore, the "Number of Treatments" box is selected to be "3".

For each consecutive pair of treatments, the direct estimates of the measure of association and the 95% lower and upper confidence limits need to be provided. The order of entry of the treatment pairs must follow the exact sequence indicated with the bridging comparison groups linking the treatment pairs. The interest here is to compare alendronate to placebo and then placebo to etidronate, and so use placebo as the bridging comparison group.

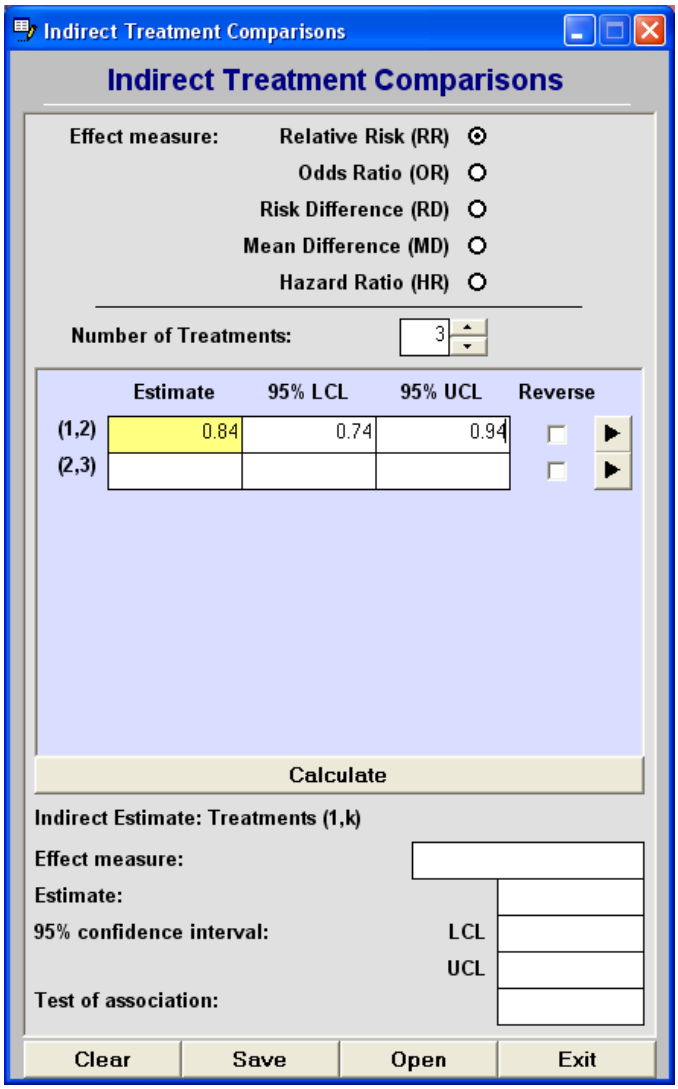


A systematic review was conducted for trials that compared alendronate to placebo for primary or secondary prevention. Non-vertebral fractures were reported in eight trials. One trial did not report fractures separately by treatment groups and one trial reported that no fractures occurred in either treatment group.

The pooled estimate of the RR of non-vertebral fractures from the five trials that could be analyzed demonstrated a significant reduction (16%) in non-vertebral fractures (RR 0.84 [95% CI: 0.74, 0.94]).

The direct estimate from this meta-analysis for the relative risk (0.84) and the 95% lower confidence limit (0.74) and 95% upper confidence limit (0.94) are entered on the first direct comparison line (1 = alendronate, 2 = placebo).

The arrow at the end of the "(1,2)" line is pressed to activate Screen 2. If the test of association is not of interest, then Screen 2 need not be activated and the next step can be skipped.



On Screen 2, the weights used for calculating the weighted effect measure for the direct comparison on Screen 1 are requested. These weights are needed for calculating the test statistic for the test of association. There are various formats in which the information to calculate these weights can be provided and these formats are identified through the weight selections (direct versus derived; fixed versus random); the specific information for each study involved in the direct comparison is then identified and requested.

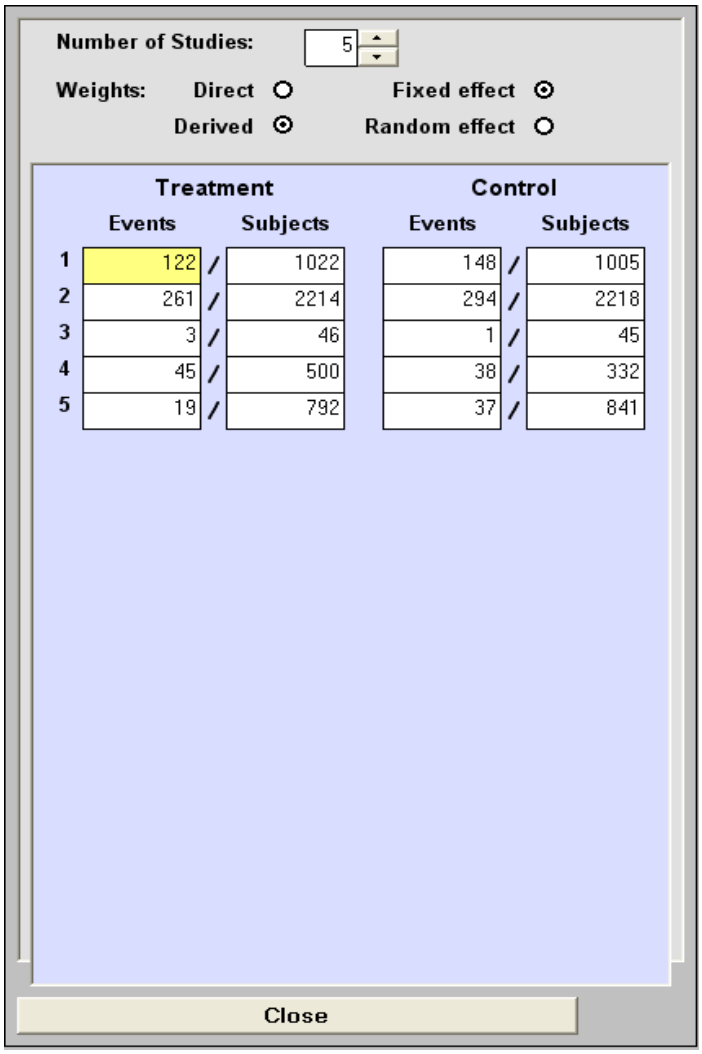
The pooled estimate of the RR of non-vertebral fractures (RR 0.84 [95% CI: 0.74, 0.94]) was based on five trials. To calculate the weights that were used for this weighted relative risk, the "Derived" circle was checked and the "Fixed effect" model weights were selected since heterogeneity was not an issue.

Rates for the treatment and control groups for each study are requested in the form of numerator (number of events): denominator (number of subjects). Rates from the systematic review are shown here:

Treatment (n/N)	Control (n/N)
122/1022	148/1005
261/2214	294/2218
3/46	1/45
45/500	38/332
19/792	37/841

These results are then entered in the corresponding lines provided.

The "Close" bar is pressed to save the entries.

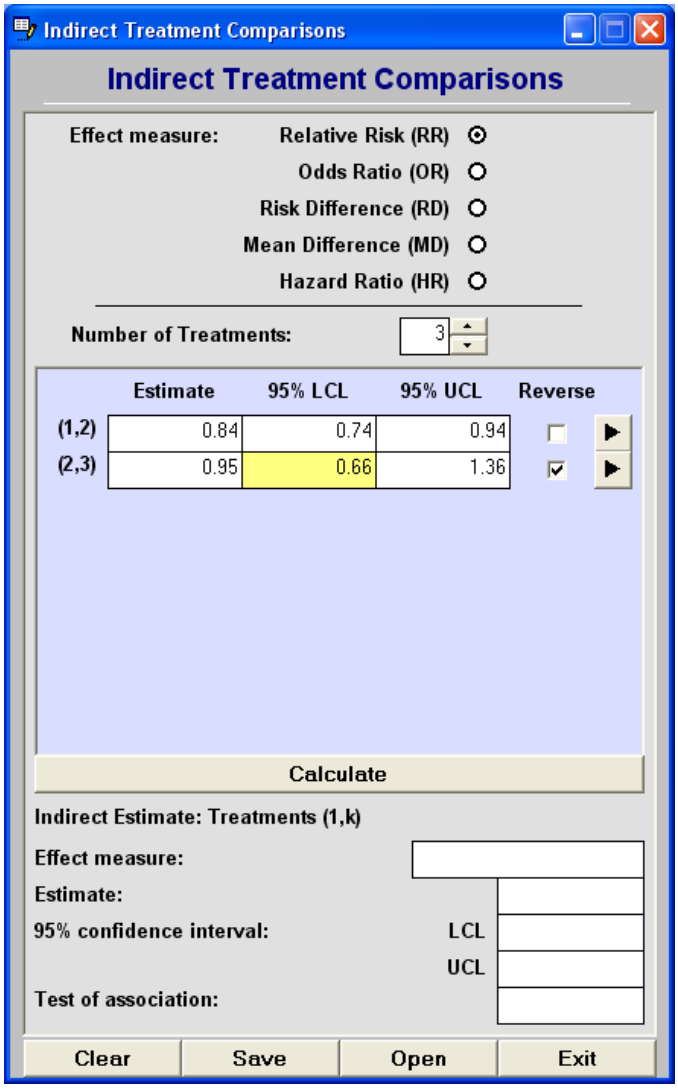


A systematic review was conducted for trials that compared etidronate to placebo for primary or secondary prevention. Non-vertebral fractures were reported in seven trials.

The pooled estimate of the RR of non-vertebral fractures from the seven trials indicated a lack of effect of etidronate on non-vertebral fractures. The 95% CI around the RR estimate for all non-vertebral fractures was wide with a relative risk reduction (RRR) of approximately 32% and a RRR increase of 42% (RR 0.95 [95% CI: 0.66, 1.36]). Results were consistent across the seven trials.

The direct estimate from this meta-analysis for the relative risk (0.95) and the 95% lower confidence limit (0.66) and 95% upper confidence limit (1.36) are entered on the second direct comparison. The results entered compare etidronate to placebo, however, the "Reverse" box was checked and so these results will be reversed so that placebo is compared to etidronate and the (2,3) will correspond to (2 = placebo, 3 = etidronate).

The arrow at the end of the (2,3) line is pressed to activate Screen 2. If the test of association is not of interest, then Screen 2 need not be activated and the next step can be skipped.



On Screen 2, the weights used for calculating the weighted effect measure for the direct comparison on Screen 1 are requested. These weights are needed for calculating the test statistic for the test of association. There are various formats in which the information to calculate these weights can be provided and these formats are identified through the weight selections (direct versus derived; fixed versus random), and the specific information for each study involved in the direct comparison is then identified and requested.

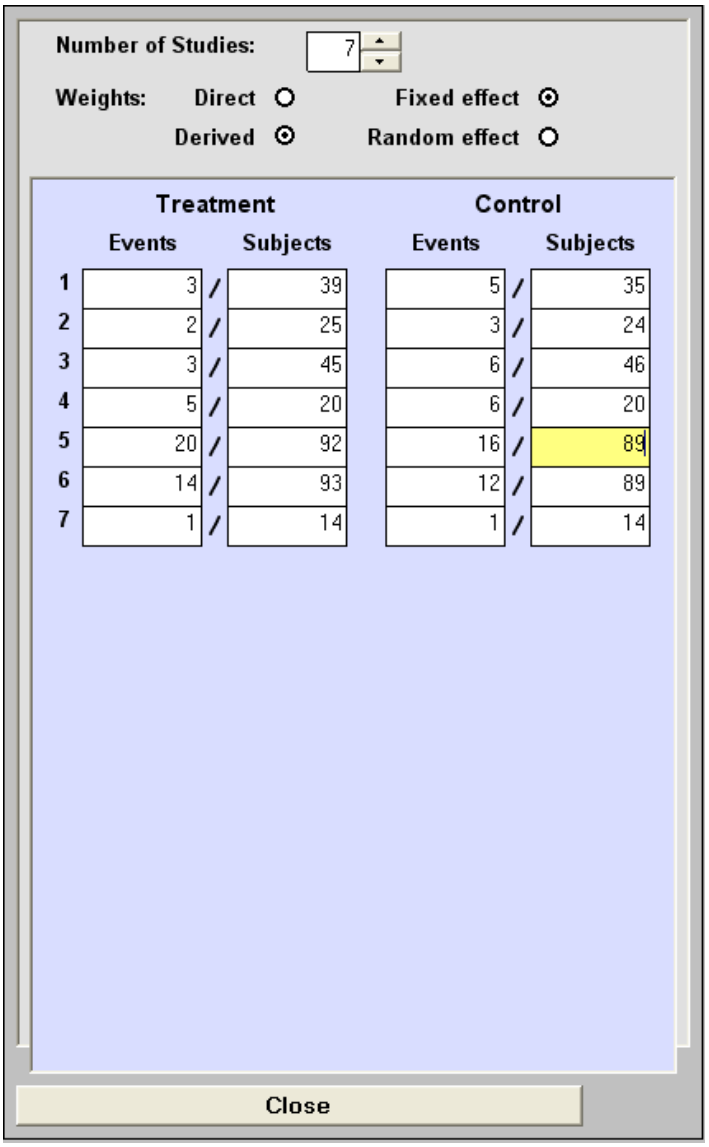
The pooled estimate of the RR of non-vertebral fractures (RR 0.95 [95% CI: 0.66, 1.36]) was based on seven trials. To calculate the weights that were used for this weighted relative risk, the "Derived" circle was checked and the "Fixed effect" model weights were selected since heterogeneity was not an issue.

Rates for the treatment and control groups for each study are requested in the form of numerator (number of events): denominator (number of subjects). Rates from the systematic review are shown here:

Treatment (n/N)	Control (n/N)
3/39	5/35
2/25	3/24
3/45	6/46
5/20	6/20
20/92	16/89
14/91	12/89
1/14	1/14

These results are then entered in the corresponding lines provided.

The "Close" bar is pressed to save the entries.

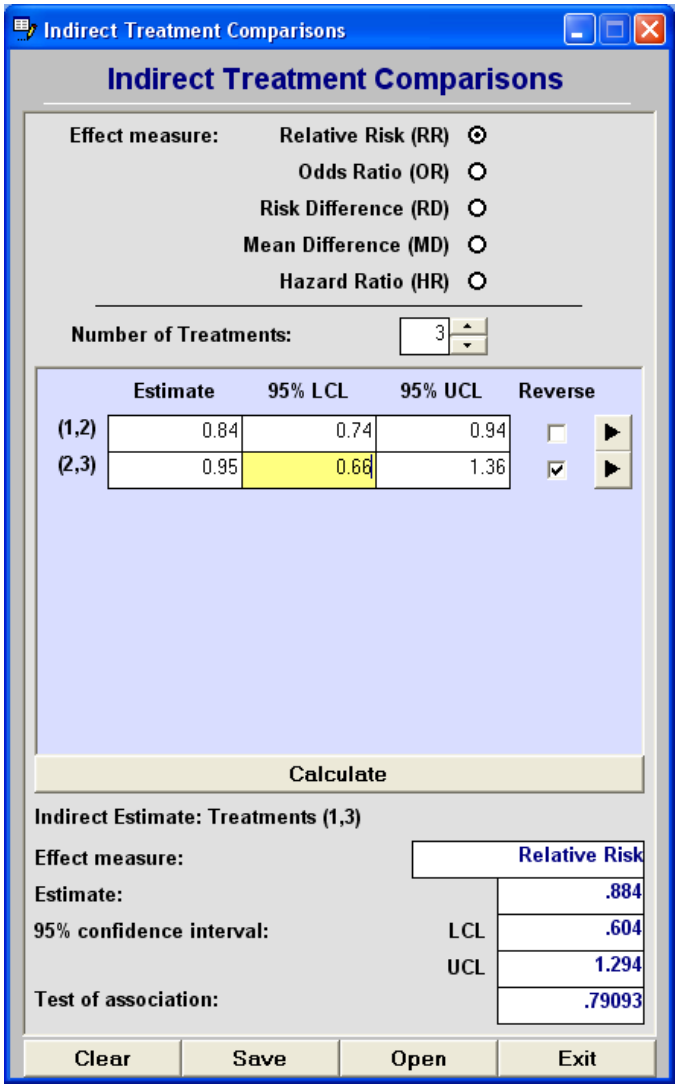


Once all the data are entered, the resulting indirect comparison estimates for the effect measure and the 95% CI, as well as the P value for the test of association corresponding to this effect measure, are provided on Screen 1 for the comparison of treatments (1,3) using treatment 2 as the bridging comparison.

All the data was entered and the "Calculate" bar was pressed.

The indirect treatment effect estimate for the relative risk of alendronate compared to etidronate was 0.88 with the 95% CI (0.60, 1.29). The result indicates that alendronate and etidronate are not significantly different. This is confirmed with the P value for the test of association of 0.79.

"Save" can be pressed to save the results.



5 APPLICATION OF THE ITC PROGRAM: ILLUSTRATIONS FROM THE LITERATURE

The fifth objective was to illustrate the application of the empirically derived distributional properties of the indirect estimates and the program by applying it to examples selected from the literature. In this chapter, the ITC program is applied using examples from the literature in which indirect treatment comparisons were used. Five examples were considered from the literature illustrating the indirect treatment comparisons for RR, OR, RD, MD, HD.

For each example, the background and analysis from the article are provided. The results based on the ITC program are then provided and compared to the published results. Finally, assessments of the bias and MSE associated with the indirect treatment comparison based on the simulated results are provided.

5.1 Relative Risk

Reference: Lim et al.[28]

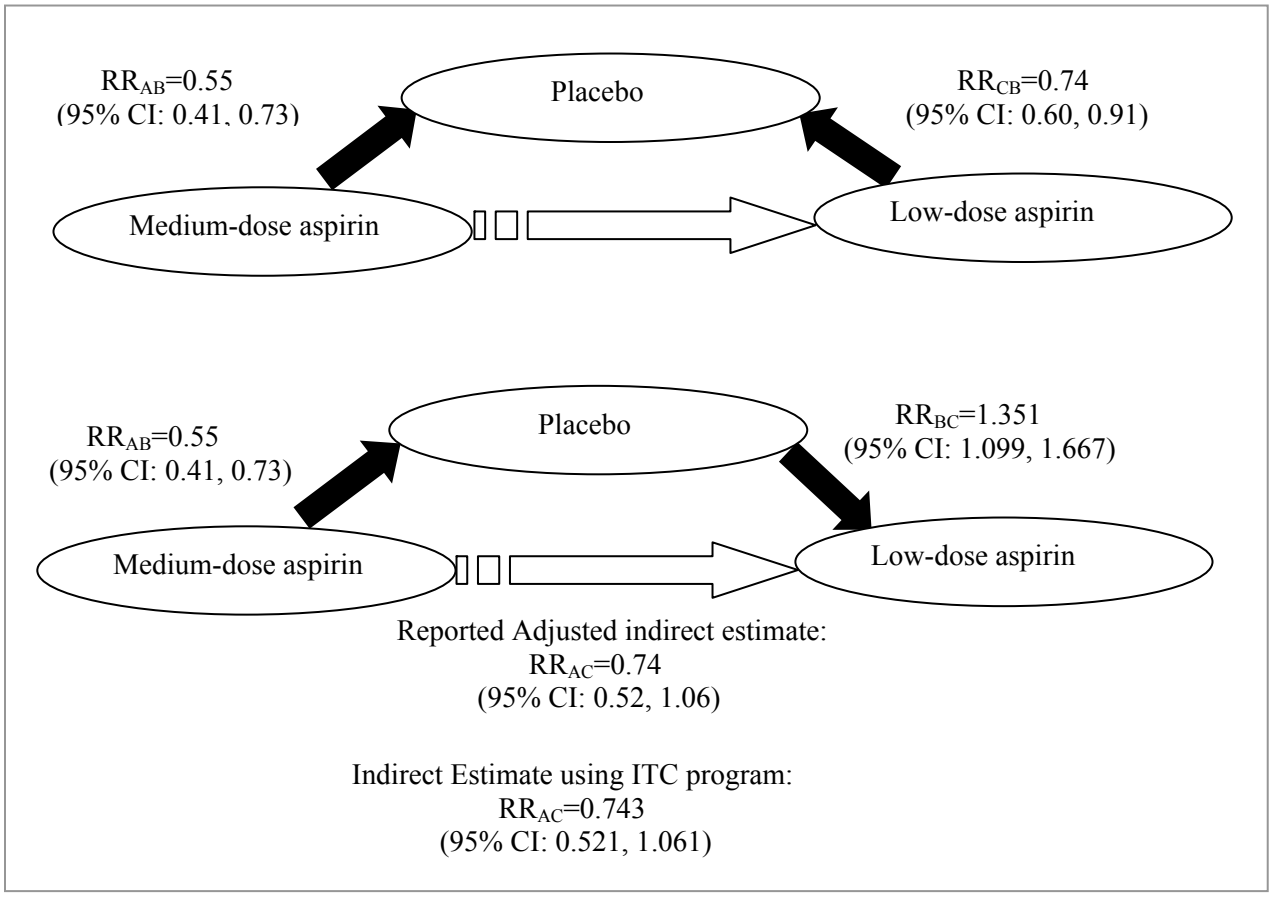
Background: Coronary artery bypass graft (CABG) surgery is a frequently performed procedure; in the United Kingdom, the estimated number of procedures exceeds 25,000 annually. The saphenous vein graft is commonly used in the procedure and its occlusion rate is approximately 15% to 30% in the first year after surgery. Antiplatelet therapy is an important intervention that is prescribed post-operatively after CABG surgery to inhibit the platelet aggregation that results from the physiological stress experienced during surgery.

Aspirin has been shown to be an effective antiplatelet therapy for graft patency. Although three meta-analyses, based on trials between 1979 to 1993, have each illustrated the efficacy of aspirin, none of these meta-analyses determined whether the observed benefits of aspirin were consistent across different doses of the medication. The range of aspirin doses analyzed in the meta-analyses varied between 75 mg and 325 mg. Low-dose aspirin (75 mg to 100 mg) is often prescribed, even though its relative efficacy compared to medium-dose aspirin has not been established.

Analysis: Lim et al.[28] (2003) used the adjusted indirect comparison method to evaluate the efficacy of low dose aspirin in their study (50 mg to150 mg) compared to medium dose (300 mg to 325 mg) therapy on graft patency after coronary artery surgery. The investigators used placebo as the common comparator and their analysis included trials that compared either low dose aspirin or medium dose aspirin to placebo therapy. The primary endpoint was graft patency and was reported as graft occlusion and event rate. Graft occlusion was defined as a distal anastamosis that could not be visually detected through angiography. An event rate was defined as one or more occlusions of the saphaneous vein graft.

Comparison of indirect treatment comparisons program with reported indirect estimate: We used the ITC program to recalculate the adjusted indirect estimate for graft occlusion. Figure 14 shows the RR and its 95% CI calculated by Lim et al., as well as the ITC program. The result obtained from the ITC program is in agreement with the estimate calculated by Lim et al.

Figure 14: Relative Risk for the Indirect Comparison Between Low- and Medium-Dose Aspirin

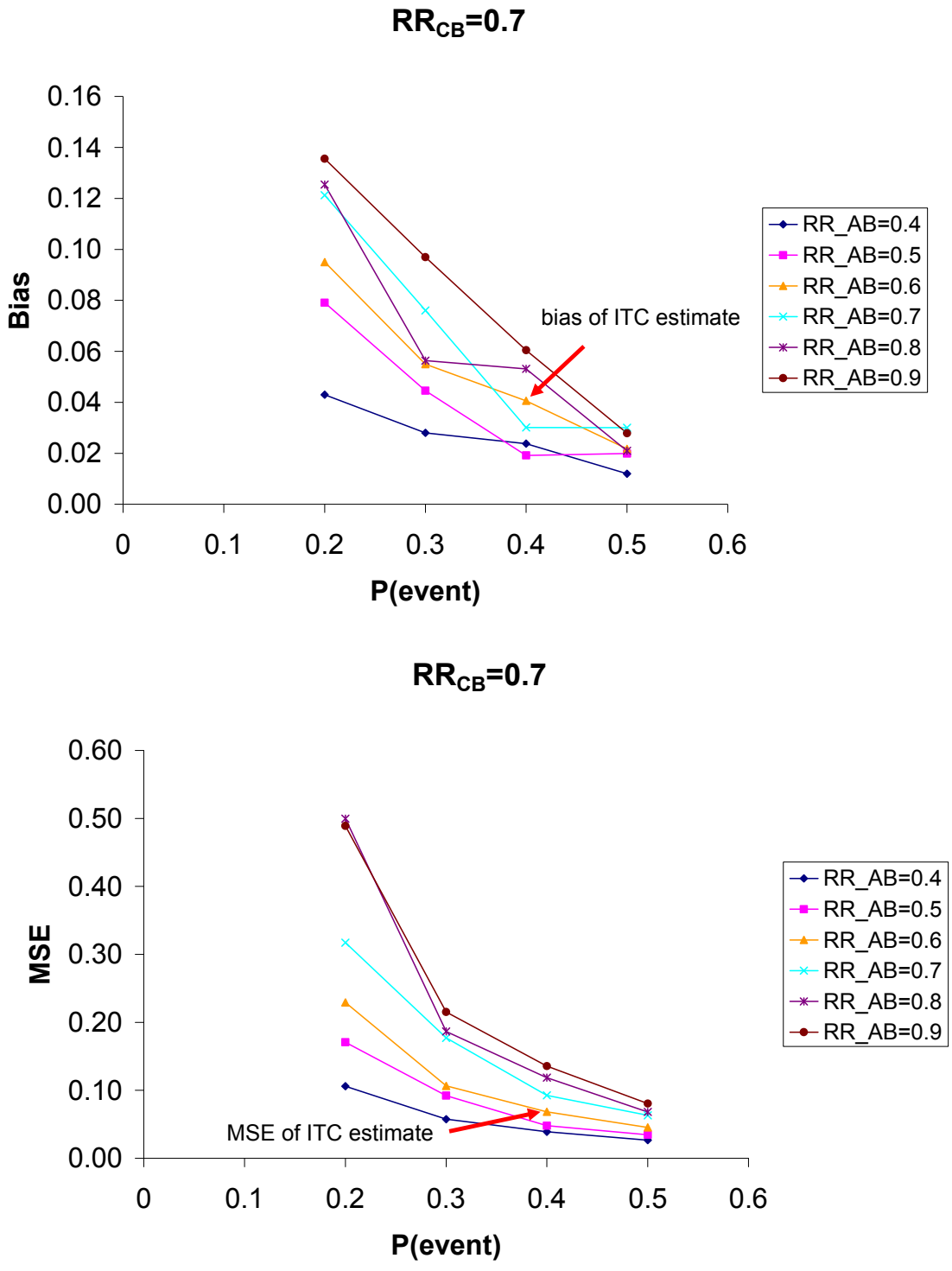


CI=confidence interval; RR=relative risk

Results from a direct comparison: Not available.

Estimated bias and mean square error associated with the indirect estimate: The following parameter settings were chosen from the simulation data in order to determine the bias and mean square error (MSE) of the indirect estimate: $RR_{AB} = 0.6$ ($RR_{AB(reported)} = 0.55$), $RR_{CB} = 0.7$ ($RR_{CB(reported)}=0.74$) and $P(E|B) = 0.4$. The event rate in the placebo group was based on the calculation of the average event rate across the placebo groups in all trials included in the indirect comparison. Based on these parameter settings, the bias was 0.041 and the MSE was 0.068 (Figure 15).

Figure 15: Expected Bias and Mean Square Error of the Relative Risk for theIndirect Comparison Between Low- and Medium-Dose Aspirin for GraftOcclusion and Event Rate



ITC=indirect treatment comparison; MSE=mean square error; RR=relative risk

5.2 Odds Ratio

Reference: Fisher et al.[1]

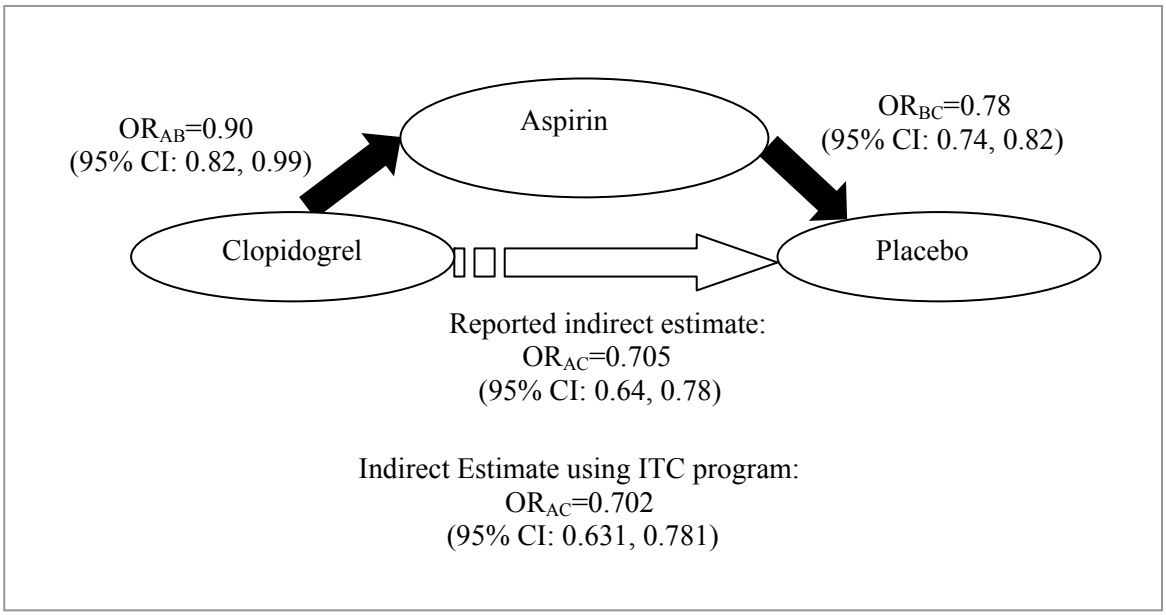
Background: Atherothrombotic disease involves platelet activation that can subsequently result in an increased risk of ischemic stroke, myocardial infarction, and vascular death. Antiplatelet therapy is recommended for platelet inhibition in order to reduce the occurrence of such events in patients with atherothrombotic disease. The CAPRIE trial,[29] a randomized controlled trial which followed patients between 1992 to 1996, established the relative efficacy of clopidogrel — a new antiplatelet therapy — over aspirin in reducing the risk of a composite outcome that included ischemic stroke, myocardial infarction, and vascular death. Clopidogrel was compared to aspirin rather than placebo because aspirin was considered standard treatment for patients at high risk of developing negative outcomes associated with platelet activation. The efficacy of aspirin over placebo was established in a large number of clinical trials that were also meta-analyzed by the Antiplatelet Trialists' Collaboration (APTC) systematic review.[30] The review assessed the effect of various antiplatelet therapies on the risk of developing the composite outcome.

Although the CAPRIE trial provided evidence of the benefit associated with clopidrogrel in relation to aspirin, data from that trial cannot provide an estimate of the total effect of the drug. To calculate this estimate, clopidogrel would have had to be compared with placebo. The ethics and usefulness of comparing the new therapy against placebo may be questioned in the presence of an established active control; however, for regulatory purposes, there is interest in knowing the true effect of the drug. In some countries, in order for regulatory agencies to approve a drug for use, its superiority against placebo needs to be illustrated; such a comparison demonstrates the total effect of the drug. Although "add-on" trials (in which a new therapy administered in combination with standard treatment is compared to standard therapy alone) may be performed in order to compare a new drug with placebo, such trials do not provide any information about the total effect of a new therapy alone.

Analysis: Fisher et al. used the adjusted indirect comparison methodology to evaluate clopidogrel versus placebo. The investigators used aspirin as the common comparator and their analysis included the CAPRIE trial that assessed clopidogrel versus aspirin and the APTC meta-analysis that compared aspirin to placebo. Fisher et al. calculated the adjusted indirect estimate of the OR and its 95% CI for various outcomes. In particular, a composite outcome cluster consisting of any stroke, myocardial infarction, and vascular death (including hemorrhage) was analyzed.

Comparison of indirect treatment comparisons program with reported indirect estimate: We used the ITC program to recalculate the indirect estimate of the OR for the composite outcome consisting of any stroke, myocardial infarction, and vascular death. Figure 16 shows the OR and its 95% CI that was calculated by Fisher et al., as well as the estimate derived from the ITC program. The result obtained from the ITC program is in agreement with the estimate calculated by Fisher et al.

Figure 16: Odds ratio for the indirect comparison between clopidogrel andplacebo

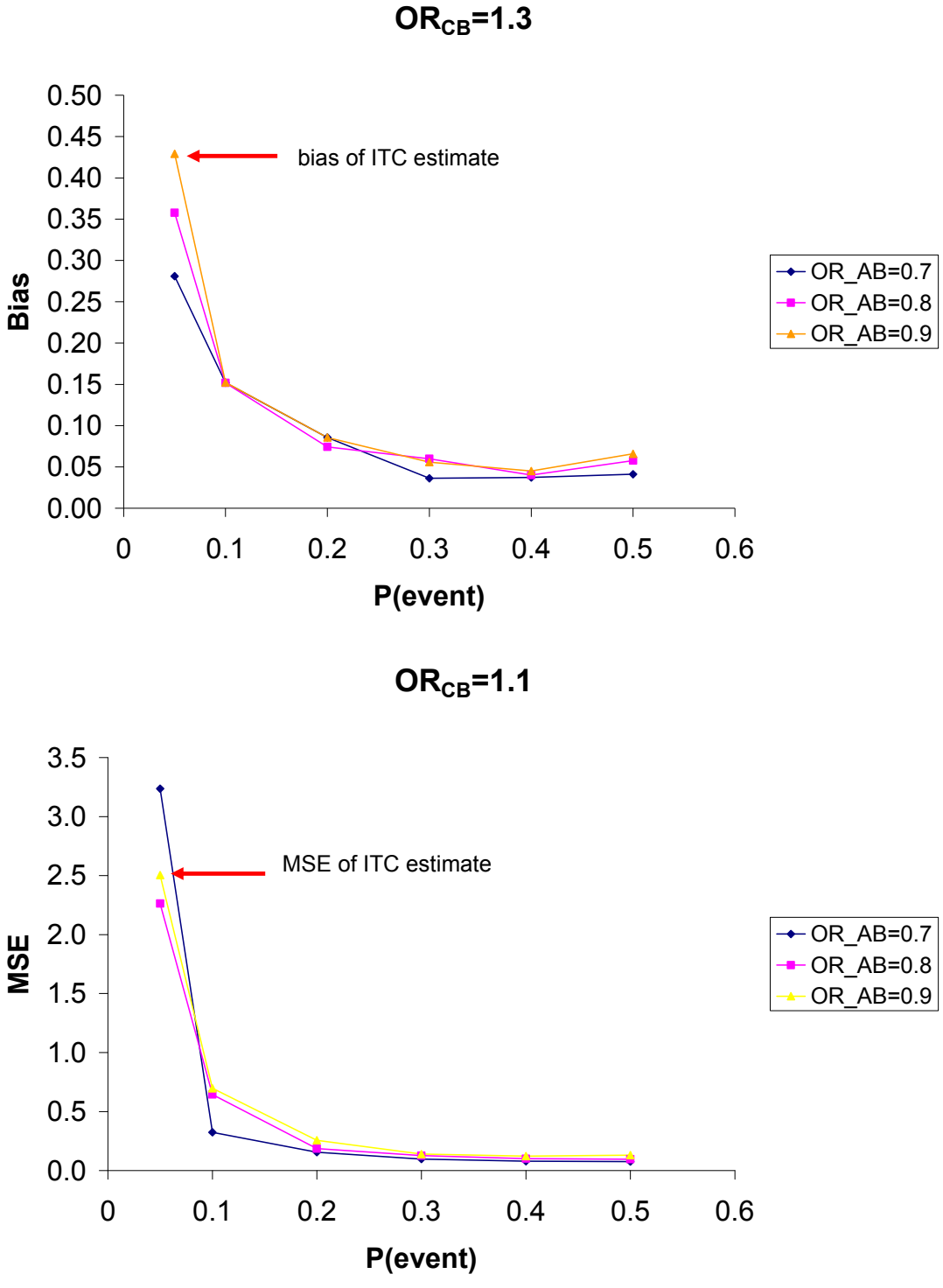


ITC=indirect treatment comparison; OR=odds ratio

Results from a direct comparison*:* Not available.

Estimated bias and mean square error associated with the indirect estimate: The following parameter settings were chosen from the simulation data in order to determine the bias and MSE of the indirect estimate: OR_{AB}=0.9 ($OR_{AB(reported)}$=0.9), OR_{CB}=1.3 ($OR_{CB(reported)}$=1.28) and P(E|B)=0.05 ($P(E|B)_{reported}$=0.0583). Based on these parameter values, the bias and MSE associated with the indirect estimate is 0.43 and 2.5, respectively (Figure 17). The bias and MSE are quite large and care must be exercised in interpreting the indirect estimate.

Figure 17: Expected Bias and Mean Square Error of the Odds Ratio for the Indirect Comparison Between Clopidogrel and Placebo for the Composite Outcome Consisting of Stroke, Myocardial Infarction, or Vascular Death



ITC=indirect treatment comparison; MSE=mean square error; OR=odds ratio

5.3 Risk Difference

A literature search did not result in a published example of an indirect comparison that used both the Bucher methodology and reported treatment effect on the risk difference scale. However, a report in which a general linear model (GLM) approach was utilized to perform indirect comparisons, and in which efficacy was expressed as risk difference (RD), had been published.[31] Although the method used for the indirect treatment comparisons performed in the published example varies from the Bucher method (and a discussion of the GLM approach is not included in this report), this example has been included to demonstrate that the results obtained through the use of the ITC program are similar to the result obtained through the GLM approach.

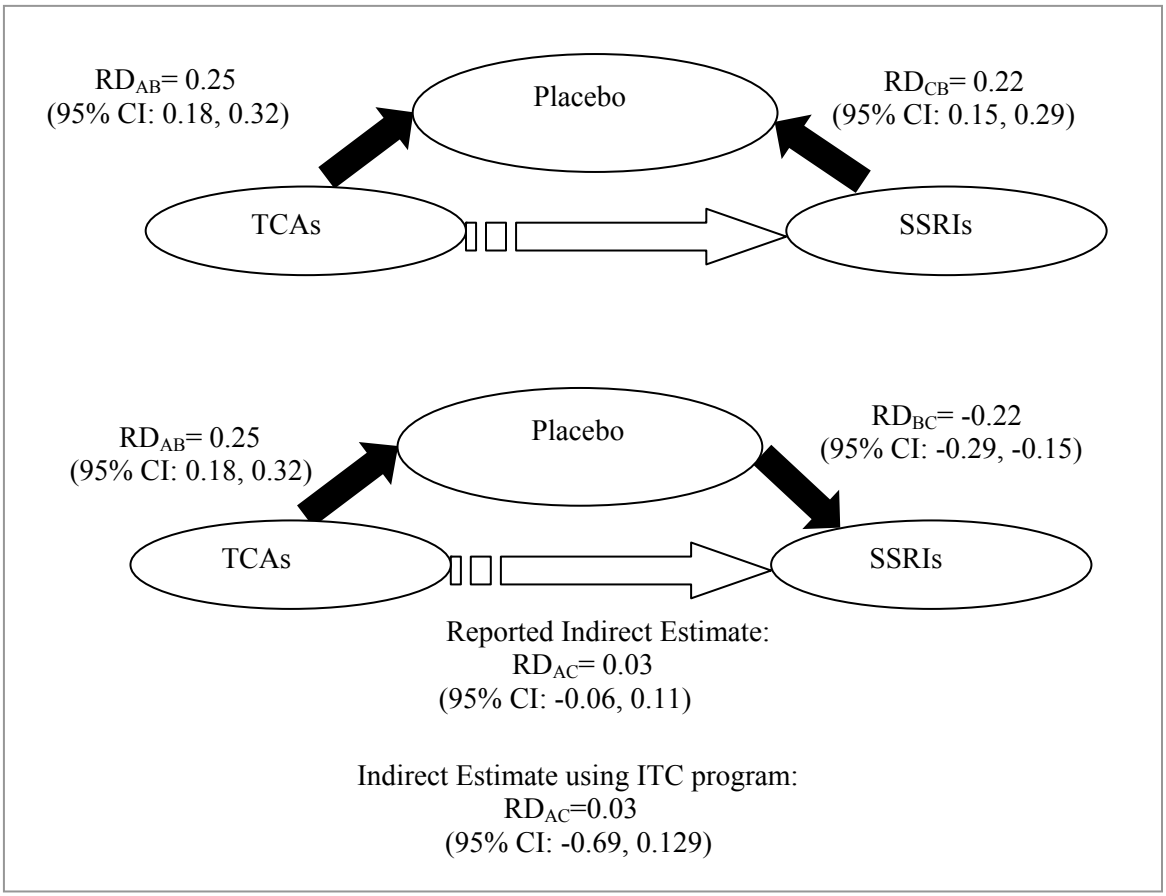
Reference: Ballesteros[31]

Background: Medications from various drug classes have been prescribed for the treatment of dysthymia and placebo-controlled trials have evaluated the relative efficacy of tricyclic antidepressants (TCAs), selective serotonin reuptake inhibitors (SSRIs), as well as monoamine oxidase inhibitors (MAOIs). From a regulatory perspective, placebo-controlled trials have been useful in evaluating the effectiveness of each of the three psychotropic drugs. Clinically, however, evidence about the relative efficacy between these classes of antidepressant medications provides useful information to physicians who need to consider various different treatment options for their patients. Ballesteros has indicated that, although direct comparisons between TCAs versus SSRIs have been performed, there is limited direct evidence for the relative efficacy of each of TCAs and SSRIs to MAOIs.

Analysis: To illustrate the use of the GLM approach, Ballesteros performed an indirect comparison to evaluate the efficacy of TCAs versus MAOIs, SSRIs versus MAOIs, and TCAs versus SSRIs. The investigators used placebo as the common comparator, and the analysis included trials in which each of TCAs, SSRIs, and MAOIs had been compared to placebo. Ballesteros calculated the indirect estimate of the risk difference for the outcome defined as a 50% decrease from baseline, in depressive symptoms, or other similar criteria.

Comparison of indirect treatment comparisons program with reported indirect estimate: We used the ITC program to recalculate the indirect estimate of the RD for the aforementioned outcome. Figure 18 shows the RD and its 95% CI that was calculated by Ballesteros, as well as the estimate derived from the ITC program. The result obtained from the ITC program is in agreement with the estimate calculated by Ballesteros.

Figure 18: Risk Difference for the Indirect Comparison Between Tricyclic Antidepressants and Selective Serotonin Reuptake Inhibitors

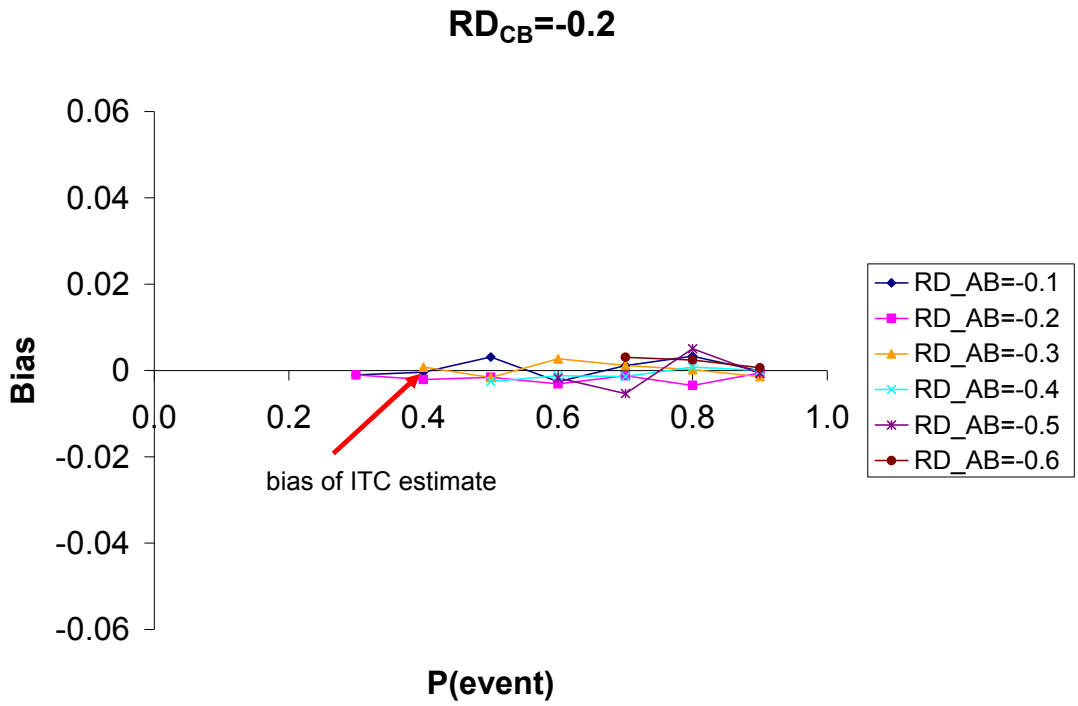


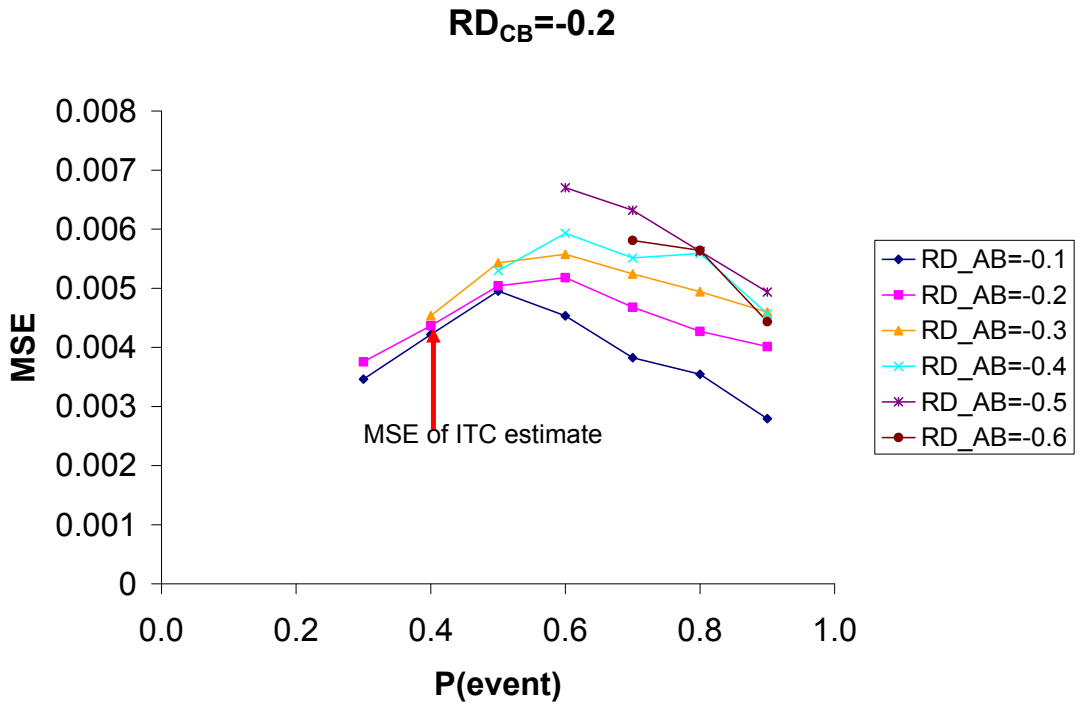
ITC=indirect treatment comparison; RD=risk difference; SSRIs=selective serotonin reuptake inhibitors; TCAs=tricyclic antidepressants

Results from a direct comparison: Ballesteros provided the results of a study which directly compared SSRIs versus TCAs. In the direct comparison, there was no statistically significant difference between TCAs and SSRIs (RD -0.05 [95% CI: -0.07, 0.17]).

Estimated bias and MSE associated with the indirect estimate: The following parameter settings were chosen from the simulation data to determine the bias and MSE of the indirect estimate. RD_{AB}=-0.3 ($RD_{AB(reported)}$=0.25), RD_{CB}=-0.2 ($RD_{CB(reported)}$=0.22) and P(E|B)=0.4. It should be noted that in the Ballesteros comparison, the probability of an event referred to the success rate; however, the event rate in the simulation data represents the rate of a failed response. Therefore, in order to apply the results of the simulation, we changed the direction of the RD_{AB} and RD_{CB} provided by the authors. The event rate was calculated as an average of the placebo group event rates reported in all the trials included in the ITC. Based on the parameter values, the bias and MSE associated with the indirect estimate is 0.0008 and 0.0045, respectively (Figure 19).

Figure 19: Expected Bias and Mean Square Error of the Risk Difference for the IndirectComparison Between Tricyclic Antidepressants and Selective Serotonin Reuptake Inhibitors, for a 50% Decrease, Since Baseline, in Depressive Symptoms or Other Similar Criteria





ITC=indirect treatment comparison; RD=risk difference; SSRIs=selective serotonin reuptake inhibitors; TCAs=tricyclic antidepressant;s

5.4 Mean Difference

Reference: Sauriol et al.[32]

Background: Conventional antipsychotic medications for the treatment of schizophrenia have been effective in the alleviation of positive symptoms; however, their efficacy is limited for the reduction of negative symptoms and these drugs present with significant side effects. Second generation therapies have been developed to provide improved effectiveness and fewer side effects. Second generation therapies, olanzapine and respiridone, have each been compared to haloperidol, a first generation medication that is considered the standard therapy in treatment of schizophrenia. Although a number of published clinical trial reports have compared each of olanzapine and respiridone to haloperidol, clinical trial data comparing olanzapine to respiridone is very limited. Only one study, conducted by Tran et al.[33] (1997), has evaluated these two medications in a head-to-head trial.

Analysis: Sauriol et al.[32] (2001) performed an adjusted indirect comparison to evaluate the relative efficacy of olanzapine versus respiridone, and compared their results with the result of the single head-to-head trial. The investigators used haloperidol as the common comparator, and the analysis included trials of olanzapine versus haloperidol and respiridone versus haloperidol. Each of the studies utilized for the indirect comparison presented the mean difference between olanzapine versus haloperidol or respiridone versus haloperidol for various efficacy and safety outcomes. The following safety outcomes were analyzed, change from baseline in:

- Brief Psychiatric Rating Scale (BPRS)
- Positive and Negative Syndrome Scale (PANSS), negative subscale score.

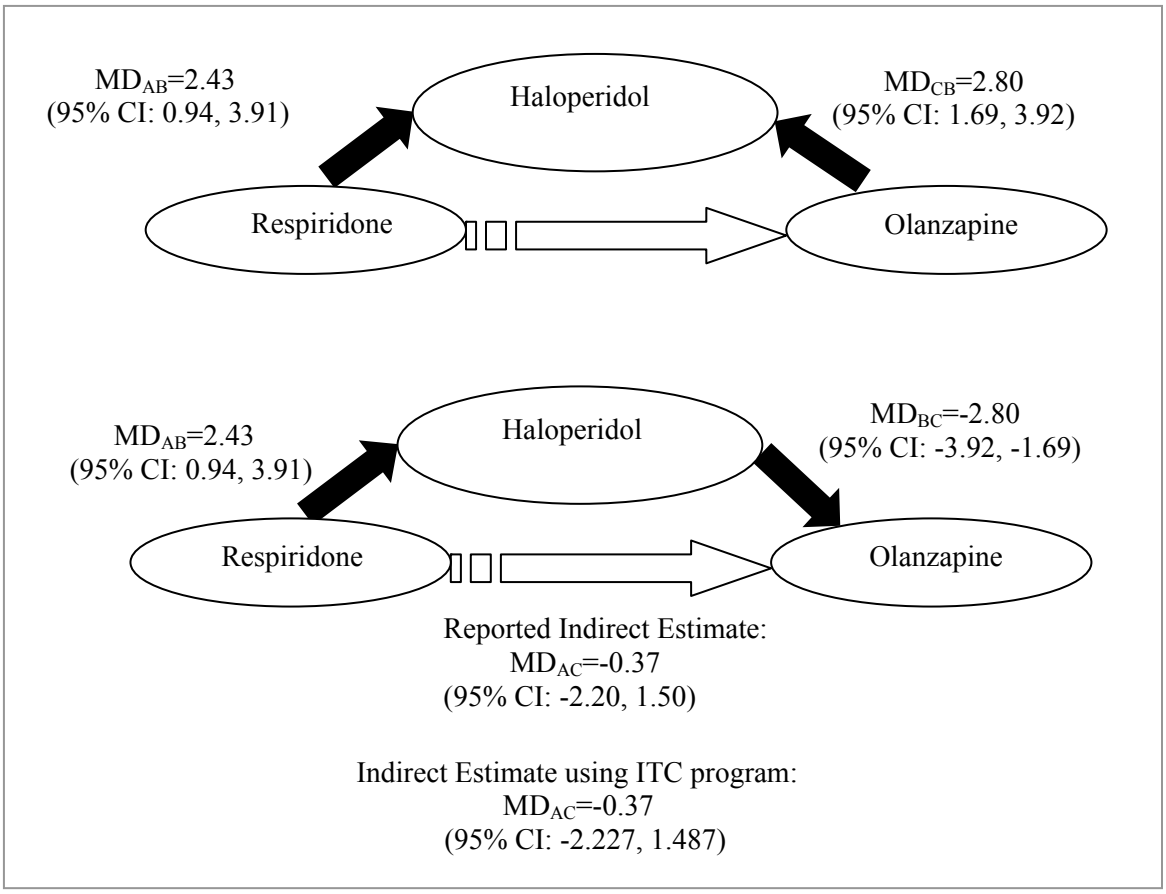
The following tolerability outcomes were assessed, difference, from baseline, in the percentage of patients who:

- used anticholinergic drugs
- dropped out due to side effects
- dropped out due to lack of efficacy
- dropped out due to any cause.

Through use of the results for the aforementioned outcomes, Sauriol et al. indirectly calculated the MD between olanzapine and respiridone for the same outcomes.

Comparison of ITC program with reported indirect estimate: We used the ITC program to recalculate the indirect estimates of the MDs for all of the efficacy and safety parameters. For the outcome BPRS total score change, Figure 20 provides the indirect MD and its 95% CI that was calculated by Sauriol et al., as well as the ITC program. The value for the MD obtained using the ITC program is in agreement with Sauriol et al.'s result. For all other safety and efficacy outcomes, Table 8 provides the adjusted indirect estimates reported by the investigators and those obtained from the ITC program. All of the indirect estimates obtained from the ITC program are in agreement with those provided by Sauriol et al.

Figure 20: Mean Difference for the Indirect Comparison Between Respiridone andOlanzapine



CI=confidence interval; MD=mean difference

Results from a direct comparison: Sauriol et al. provided the results of a study which directly compared olanzapine versus respiridone. In the direct comparison, for the outcome BPRS total score change, there was no statistically significant difference between olanzapine and respiridone (MD 1.80 [95% CI: -1.40, 5.00]). For all other safety and efficacy outcomes, Table 8 provides the results of the direct comparison.

Estimated bias and mean square error associated with the indirect estimate: For the outcome BPRS total score change, the following simulation parameter setting for effect size (ES) and coefficient of variation (CV) were used to determine the bias and MSE for the indirect MD reported in Sauriol et al.'s example: ES_{AB}=0.2, ES_{CB}=0.2, CV_B=0.5, $Mean_B$=10 ($Mean_{B(reported)}$=9.38). To obtain values for ES_{AB} and ES_{CB}, we divided MD_{AB} and MD_{CB} by their standard deviations. CV_B=0.5 was selected to represent the high level of variation observed in the haloperidol group. Based on the values of the various parameters, the expected bias for the indirect estimate is -0.366 and the expected MSE is 1.101 (Figure 21).

Table 8: Direct and Adjusted Indirect Results for Various Outcomes in the Comparison of Olanzapine and Risperidone for the Treatment of Schizophrenia					
Variables	**Two Treatments Comparison MD (95% CI)**		**Direct Method MD (95% CI)**	**Adjusted Indirect Method MD (95% CI)**	
	[1]Risperidone (A) vs. Haloperidol (B)	**[1,2]Haloperidol (B) vs. Olanzapine (C)**	**Risperidone (A) vs. Olanzapine (C)**	**Risperidone (A) vs. Olanzapine (C)**	
	Reported	**Reported**	**Reported**	**Reported**	**Using ITC Program**
Efficacy					
BPRS total score change	2.43 (0.94, 3.91)	-2.80 (-3.92, -1.69)	-1.80 (-5.00, 1.40)	-0.37 (-2.20, 1.50)	-0.37 (-2.227, 1.487)
PANSS negative score change	0.81 (-0.07, 1.69)	-1.35 (-1.89, -0.81)	-1.10 (-2.60, 0.44)	-0.54 (-1.60, 0.49)	-0.54 (-1.572, 0.492)
Tolerability (difference in % of patients)					
Anticholinergic drug use	13.4 (8.0, 18.9)	-33.0 (-36.8, -29.1)	-13.1 (-22.5, -3.9)	-19.5 (-26.2, -12.8)	-19.6 (-26.273, -12.927)
Dropped out due to side effects	1.2 (-2.3, 4.7)	-3.4 (-5.5, -1.2)	-0.3 (-6.7, 6.1)	-2.2 (-6.7, 2.4)	-2.2 (-6.308, 1.908)
Dropped out due to lack of efficacy	4.0 (-0.1, 8.0)	-9.7 (-13.4, -6.0)	-2.8 (-10.5, 4.9)	-5.7 (-11.2, -0.2)	-5.7 (-11.186, -0.214)
Dropped out due to any cause	7.1 (1.9, 12.2)	-17.0 (-21.2, -12.9)	-10.3 (-20.8, 0.3)	-10.0 (-16.6, -3.3)	-9.9 (-16.514, -3.28)

BPRS= Brief Psychiatric Rating Scale; CI=confidence interval; MD=mean difference; PANSS= Positive and Negative Syndrome Scale; vs=versus

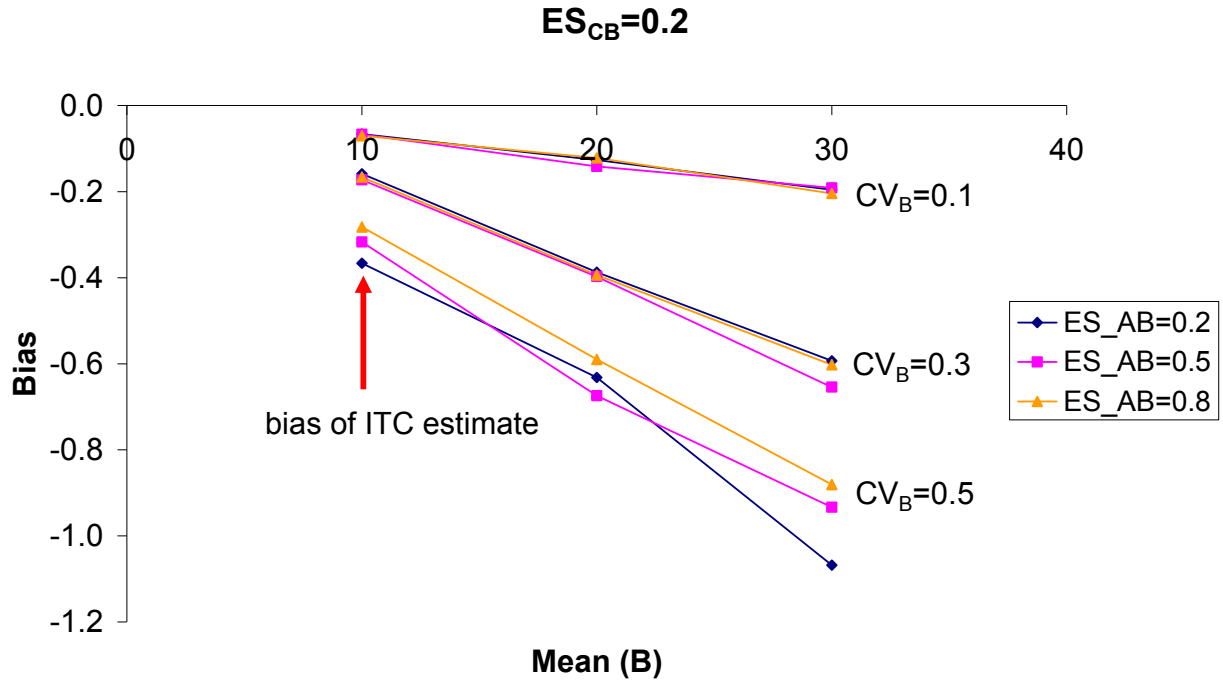
[1]Difference in variations from baseline to end point;

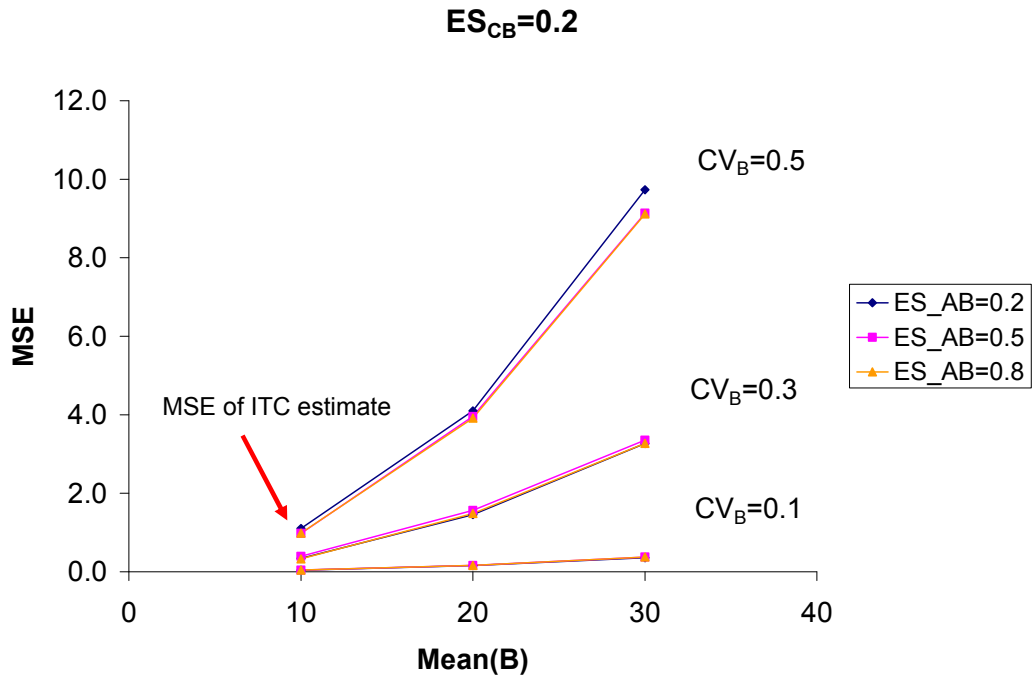
[2] The B versus C risk difference, MD_{BC}, and 95% CI (lcl_{BC}, ucl_{BC}) are obtained from the corresponding values for MD_{CB} by using the relations: $MD_{BC} = -MD_{CB}$, $lcl_{BC} = -ucl_{CB}$ and $ucl_{BC} = -lcl_{CB}$.

Table 8: Direct and Adjusted Indirect Results for Various Outcomes in the Comparison of Olanzapine and Risperidone for the Treatment of Schizophrenia (cont'd)			
Outcome measure	**Differences from head-to-head RCT of Risperadone versus Olanzapine** **MD (95% CI)**	**Adjusted Indirect Comparison of Risperidone versus Olanzapine** **MD (95% CI)**	
	Reported by Tran et al[33]	**Reported by Sauriol et al[32]**	**Using ITC program**
BPRS total score change	-1.80 (-5.00, 1.40)	-0.37 (-2.20, 1.50)	-0.37 (-2.227, 1.487)
PANSS negative score change	-1.10 (-2.60, 0.44)	-0.54 (-1.60, 0.49)	-0.54 (-1.572, 0.492)
Anticholinergic drug use (%)	-13.1 (-22.5, -3.9)	-19.5 (-26.2, -12.8)	-19.6 (-26.273, -12.927)
Dropped out due to side effects (%)	-0.3 (-6.7, 6.1)	-2.2 (-6.7, 2.4)	-2.2 (-6.308, 1.908)
Dropped out due to lack of efficacy (%)	-2.8 (-10.5, 4.9)	-5.7 (-11.2, -0.2)	-5.7 (-11.186, -0.214)
Dropped out due to any cause (%)	-10.3 (-20.8, 0.3)	-10.0 (-16.6, -3.3)	-9.9 (-16.514, -3.28)

BPRS= Brief Psychiatric Rating Scale; CI=confidence interval; MD=mean difference; PANSS= Positive and Negative Syndrome Scale; vs=versus

Figure 21: Expected Bias and Mean Square Error of the Mean Difference for the IndirectComparison Between Respiridone and Olanzapine for the Outcome BPRSTotal Score Change





BPRS= Brief Psychiatric Rating Scale; ITC=indirect treatment comparisons; MSE=mean square error

5.5 Hazard Ratio

Reference: von der Maase et al.[34]

Background: Various combination therapies have been established for the treatment of bladder cancer, most of which are based on the anti-tumor medication, cisplatin. Although never compared to placebo, until at least 2000, methotrexate, vinblastine, doxorubicin and cisplatin (MVAC) was considered the standard of care for the treatment of bladder cancer.[34] The efficacy of the multi-modal MVAC therapy was established through trials[35,36] in which it was compared separately to two strategies, cisplatin alone or cisplatin, cyclophosphamide and adriamycin (CISCA), both of which are considered ineffective. Despite its efficacy, MVAC presents with a significant toxicity profile. As such, there was a need for the development of therapy that had comparable or superior efficacy to MVAC and an improved toxicity profile. Phase 2 studies have shown gemcitabine/cisplatin (GC) to have comparable efficacy to MVAC, but significantly less toxic side effects. Von der Maase et al. conducted a phase 3 clinical trial to compare GC versus MVAC. The results of the trial did not show a significant difference in efficacy between GC and MVAC; however, GC was associated with fewer side effects. Based on the results of the trial, von der Maase et al. asserted that GC should be considered the standard therapy for treatment of bladder cancer.

In a letter to the editor, Cohen and Rothmann[37] (2001) criticised the abovementioned suggestion and disagreed with the authors. Cohen and Rothmann stated that the lack of a statistically significant difference between the two combination therapies does not equate to the non-inferiority of GC. Also, they stated that in order for the non inferiority of GC to be established, it was necessary to first determine the efficacy of MVAC on the outcome survival and then determine the amount of survival benefit from MVAC that is maintained when GC is used instead of MVAC. To determine the true effect of MVAC on survival, a comparison between MVAC and placebo was required. Since no placebo-controlled trials had been performed to evaluate MVAC, its effect on survival was based on the results of the trials in which it was compared separately to CISCA and cisplatin. Cohen and Rothmann performed several calculations. Of interest, they pooled the results of the MVAC versus CISCA trial and the MVAC versus cisplatin trial to determine the hazard ratio for the survival benefit of MVAC versus "control therapy", which included cisplatin and CISCA. Once the effect of MVAC was determined, Cohen and Rothmann utilized various methodologies to determine the amount of survival benefit that could potentially be maintained through the use of GC instead of MVAC. Their results indicated that the survival benefit observed when MVAC is administered in comparison to CISCA or cisplatin would be lost if GC were administered instead. As such, Cohen and Rothmann concluded that MVAC should remain the standard of care.

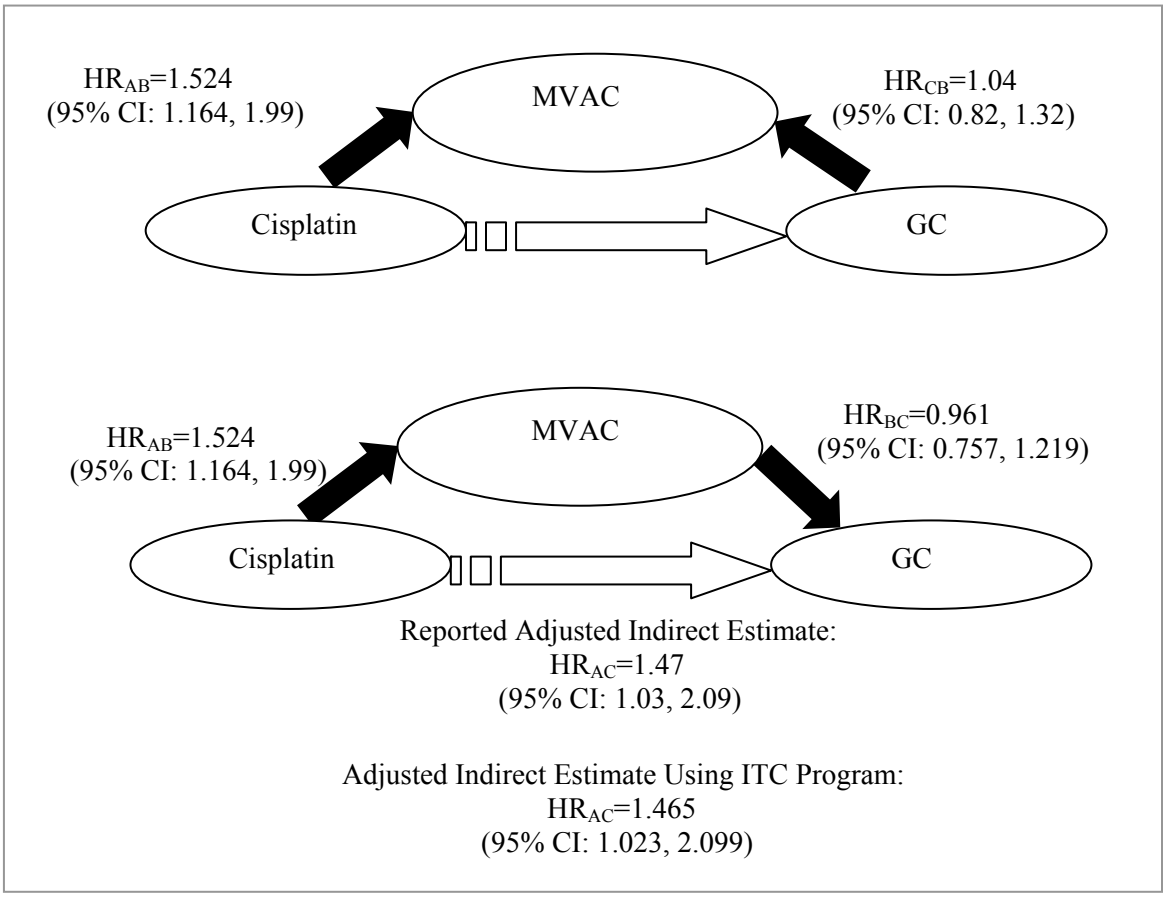
In response to Cohen and Rothmann's letter, von der Maase et al. opposed several aspects of the methodologies that Cohen and Rothman used to compare GC with MVAC. Additionally, von der Maase discussed reasons why it was inappropriate to use CISCA as a control therapy in order to quantify MVAC's efficacy.

Analysis: In order to illustrate that GC was able to maintain the observed efficacy of MVAC in comparison to cisplatin, von der Maase et al. first performed an indirect comparison between GC

and cisplatin and then used a Bayesian methodology to illustrate the amount MVAC's survival benefit is maintained when GC is used instead of MVAC in the treatment of bladder cancer. To indirectly compare GC with cisplatin, von der Maase et al. chose MVAC as the common comparator; their analysis included the trial in which MVAC was compared to cisplatin, as well as their own trial in which GC was compared to MVAC. The investigators calculated the HR for the outcome survival.

Comparison of indirect treatment comparisons program with reported indirect estimate: We used the ITC program to recalculate the adjusted indirect estimate of the hazard ratio for survival. Figure 22 provides the indirect hazard ratio and its 95% CI for GC versus cisplatin that was reported by von der Maase et al., as well as the estimate obtained from the ITC program. The value for the HR obtained from the ITC program is in agreement with the result obtained by von der Maase et al.

Figure 22: Hazard Ratio for the Indirect Comparison Betweeen Gemcitabin and Cisplatinand Methotrexate, Vinblastine, Doxorubicin, and Cisplatin for the OutcomeSurvival

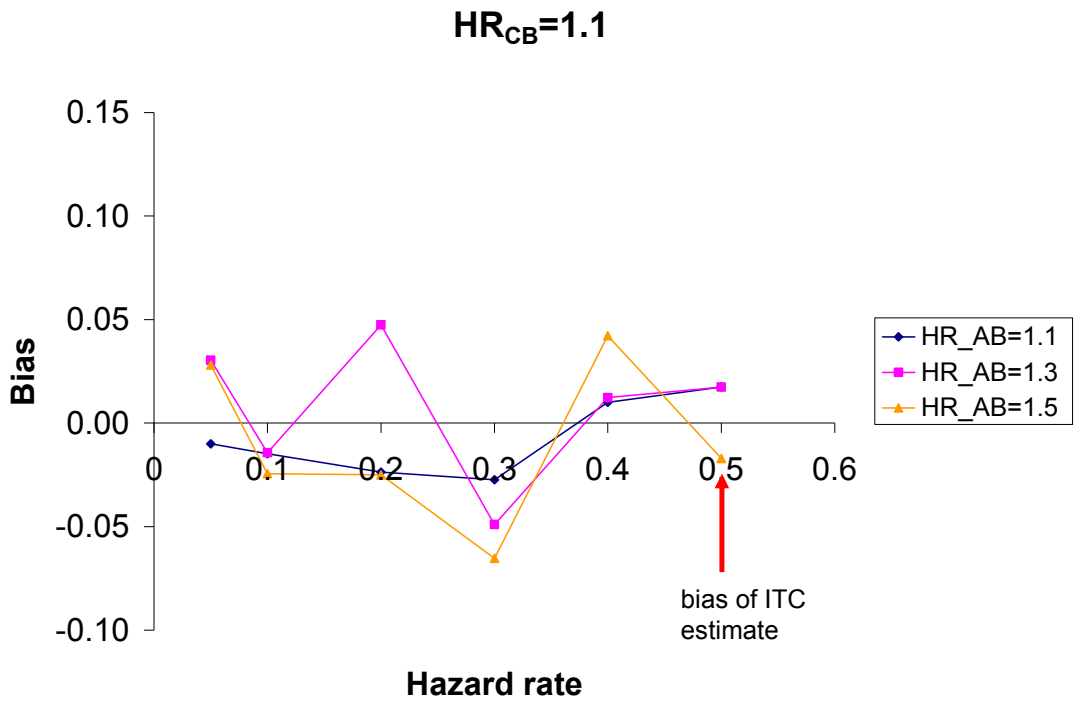


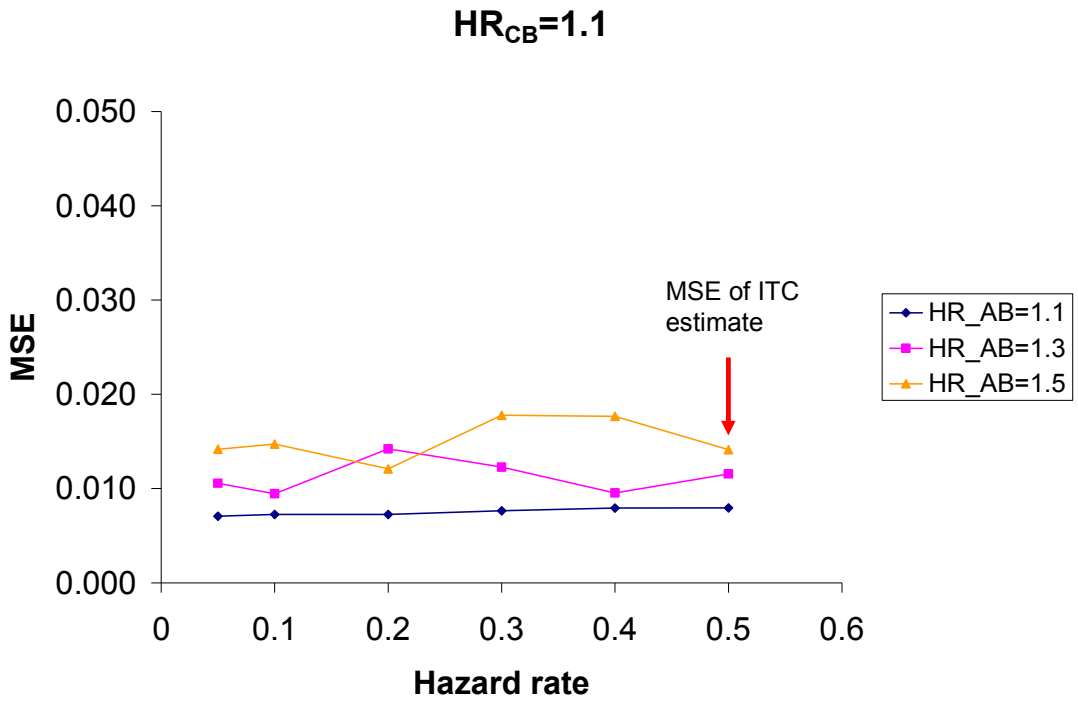
CI=confidence interval; GC=gemcitabi and cisplatin; HR=hazard ratio; ITC=indirect treatment comparisons; MVAC= methotrexate, vinblastine, doxorubicin, and cisplatin

Results of a direct comparison: Not available

Estimated bias and mean square error associated with the indirect estimate:
The parameter settings chosen from the simulation data to estimate the bias and MSE associated with the indirect HR were, as follows: HR_{AB}=1.5 ($HR_{AB(reported)}$=1.524), HR_{CB}=1.1 ($HR_{CB(reported)}$=1.04), baseline hazard rate in the MVAC group = 0.5. To calculate the hazard rate in the MVAC group, we used data from the trial in which MVAC was compared to cisplatin. An average was calculated for the hazard rate at three different time points in the survival curve for patients treated with MVAC. Like von der Maase et al., we assumed a constant survival hazard because of limited information in the study. Based on these parameter settings, the bias associated with the indirect estimate is -0.017 and the value for the MSE is 0.014 (Figure 23).

Figure 23: Expected Bias and Mean Square Error of the Hazard Ratio for the Indirect Comparison Betweeen Gemcitabin and Cisplatin and Methotrexate, Vinblastine, Doxorubicin, and Cisplatin for the Outcome Survival





GC=gemcitabi and cisplatin; HR=hazard ratio; ITC=indirect treatment comparisons; MVAC= methotrexate, vinblastine, doxorubicin, and cisplatin; MSE=mean square error

5.6 Interpretation of Indirect Comparisons From the Illustrative Examples

In addition to considering the magnitude of the bias and MSE associated with an indirect estimate, the external and internal validity of the indirect estimate should be assessed.[4] When considering the external validity, sources of discrepancies between the direct and indirect results should be considered. Specifically, it should be determined whether or not there exists a lack of comparability between the linking treatment, patient/clinical characteristics, methodological quality, and study design, as well as date of publication in the set of trials that estimate measure of association Y_{CB} or Y_{AB} and any trials that estimate Y_{AC} directly. Any other sources of heterogeneity should also be examined. When considering internal validity, comparability between trials for Y_{AB} and those for Y_{CB} should be evaluated to determine whether the trials were similar enough to be combined.

5.6.1 Bias and Mean Square Error

For each of the indirect estimates of Y_{AC} in the illustrative examples, the magnitude of the bias and MSE has been provided.

5.6.2 External Validity

Whether the indirect estimates discussed in the illustrative examples are significantly different from the results of a direct comparison could only be determined for the comparison performed by Sauriol et al.[32] and that performed by Ballesteros.[31] For the other examples, a direct estimate was not available.

For the Sauriol et al.[32] comparison, where there were differences between the direct and indirect results, the authors attributed them to the inclusion of older trials in the meta-analysis for the indirect comparison. Specifically, the average dose at which respiridone was administered in the trials included in the indirect comparison was 8.13 mg, but some studies included doses as high as 20 mg. The recommended dosage for respiridone has decreased over time — the single RCT included respiridone doses that ranged from 4 mg to 12 mg and at the time of Sauriol et al.'s publication, the recommended dosage for respiridone was 6 mg. Higher doses of respiridone may be associated with higher levels of extra pyramidal side effects (EPS) and, consequently, higher use of anticholinergic drug use. If the indirect analysis was based on doses of respiridone prescribed today, then the drug may have been associated with lower rates of EPS and, therefore, lower rates of anticholinergic drug use. As such, the results may not represent the clinical side effect profile that would be observed in the current clinical setting.

In Ballesteros'[31] study, for the comparison between tricyclic antidepressants (TCAs) and SSRIs, the result of the indirect comparison were similar to the direct estimate. Although not focused on in this report, for the comparison between monoamine oxidase inhibitors (MAOIs) versus TCAs, the direct and indirect approach generated different magnitudes for the point estimate and the range of values within the 95% CI; however, both approaches indicate that there is no statistically significant difference in efficacy between MAOIs and TCAs. Ballesteros has not provided reasons for the observed discrepancy between the direct and indirect risk difference. As such, there is a need for clinical investigators to determine whether or not there exists a lack of comparability between the linking treatment, patient/clinical characteristics, methodological quality, and study design, as well as date of publication in the set of trials that estimated the risk difference for MAOIs versus placebo or TCAs versus placebo and the trials that directly compared MAOIs with TCAs.

5.6.3 Internal Validity

Two studies considered comparability between the set of trials that estimated Y_{AB} and those that estimated Y_{CB}.[31,32]

In Sauriol et al.'s[32] indirect comparison, the authors noted the presence of hetereogeneity/clinical differences between the olanzapine versus haloperidol trials and the respiridone versus haloperidol trials. For instance, the baseline BPRS scores for the studies included in the respiridone versus haloperidol analysis were higher than those for the studies included in the olanzapine versus haloperidol studies.

Further to this, the follow-up time and the rate of anticholinergic drug use was higher in the respiridone versus haloperidol than in the olanzapine versus haloperidol studies. The authors indicated that the greater use of anticholinergic drugs in the respiridone versus haloperidol

studies perhaps occurred due to a longer follow-up period. Also, extra pyramidal side effects are more likely to occur over a longer treatment time and anticholinergic medication is used in the treatment of such side effects.

There were no differences in the linking treatment, haloperidol, in the two sets of studies included in the meta-analysis; however, due to the aforementioned heterogeneity in patient characteristics, the comparator may not have had the same effect in each of the two sets of trials and, therefore, may have influenced the observed variability between the direct and indirect estimates.

Although not mentioned by the authors, there may be methodological differences between the olanzapine versus haloperidol trials and the respiridone versus haloperidol trials that may also account for the observed differences between the direct and indirect results. For instance, the indirect comparison of respiridone and olanzapine seems to include trials that have three treatment arms. Because the method proposed by Bucher et al. does not provide a variance estimate for correlated comparison groups, trials that contain more than two comparison groups perhaps should not be analyzed through Bucher's indirect treatment comparison technique. As such, inclusion of the three-arm trials may have resulted in biased estimates for the relative efficacy of the two therapies.

In the Lim et al.[28] study, the authors indicated the presence of clinical heterogeneity between the two sets of trials included in the indirect comparison between low-dose aspirin and medium-dose aspirin. Specifically, the follow-up period was shorter in the clinical trial reports for the comparison of low-dose aspirin to placebo. In the largest trial comparing low-dose aspirin to placebo, the mean time to angiography for the low-dose aspirin group and the placebo group was 10 and 11 days, respectively. In the two other trials included in this comparison, the mean time to angiography, in days, was 131/129 (aspirin/placebo) and 180/180 (aspirin/placebo). In contrast, the two trials included in the comparison of medium-dose aspirin to placebo consisted of a follow-up period, in days, of 363/363 (aspirin/placebo) and 367/367 (aspirin/placebo). Lim et al. stated that graft occlusion at 10 and 11 days post-operatively is usually related to surgical technique rather than antiplatelet therapy. Because the trial with the shortest follow-up period was the largest of the trials included in the low-dose versus placebo comparison, its results had the greatest impact on the estimate of the relative risk. As such, the observed efficacy of the low-dose aspirin may, in actuality, not be due to the therapeutic effects of the medication, but, rather, due to a surgical technique that may have caused graft occlusions and may have been used more frequently in the placebo group. The result of this would be an overestimate of the beneficial effects of low-dose aspirin in comparison to placebo. The authors do not mention any additional differences in patient or clinical characteristics between the two sets of trials and have not commented on the methodological quality of all trials included in the indirect comparison. Any additional differences related to patient heterogeneity and differences in methodological quality that may exist between the trials included in the low-dose aspirin versus placebo comparison and the medium-dose aspirin versus placebo comparison could affect the validity of the reported results.

Importantly, the trials included in the indirect meta-analysis were published in the time period 1979 to 1993 and may not reflect the current clinical practice related to the administration of aspirin after coronary artery bypass graft surgery.

6 ILLUSTRATIONS OF THE VARIOUS METHODS: BUCHER, LUMLEY, MIXED TREATMENT COMPARISON

The sixth objective of this technical report was to illustrate the application of the various methods for indirect treatment comparisons. In this chapter, examples are presented in which the Bucher-adjusted indirect comparison method, Lumley's network meta-analysis approach, and mixed treatment comparisons (as described by Lu and Ades) methods are applied.

In their article, Vandermeer et al[26] considered different methods for evaluating the relative efficacy of benzodiazepines (BNZ) and nonbenzodiazepenes (NBNZ) in the management of chronic insomnia. The methods considered were:

- *Frequentist direct* — the inverse variance random effects meta-analysis method
- *Frequentist indirect* — the Bucher-adjusted indirect comparison method
- *Frequentist modified indirect* — the Bucher-adjusted indirect comparison method, which excluded 3-arm trials (comparing BNZ, NBMZ, placebo), since these trials appeared in the direct estimate, and the authors did not want to double-count the data from these three-arm trials when calculating the "frequentist combined estimate"
- *Frequentist combined estimate* — the frequentist direct and modified indirect estimates combined according to a normal random effects, meta-analysis model with weights determined using the inverse variance concept
- *Bayesian direct estimate* — a random effects, Bayesian meta-analysis
- *Bayesian combined estimate* — the mixed treatment comparisons method proposed by Lu and Ades.

These methods were used to compare the BNZ and NBNZ on six outcomes: sleep-onset latency, wakefulness after sleep onset, sleep efficiency, total sleep time, sleep quality, and adverse events. For each of the indirect, direct and combined methods, the meta-analytic results for each outcome are summarized in Table 9.

Table 9: Comparison of the Results for the Different Methods for Evaluating the Relative Efficacy of Benzodiazepines and Nonbenzodiazepenes in the Management of Chronic Insomnia[26]

	Frequentist				Bayesian	
Outcome	**Direct**	**Indirect**	**Modified Indirect**	**Combined**	**Direct**	**Combined**
Sleep onset latency (minutes: mean difference)						
Significant	No	No	No	No	No	No
Favours	NBNZ	NBNZ	NBNZ	NBNZ	NBNZ	NBNZ
Width of CI	9.92	11.96	13.68	8.03	11.69	9.95
Number of studies	11	62	54	65	11	65
Wakefulness after sleep onset (minutes: mean difference)						
Significant	No	No	No	No	No	No
Favours	BNZ	BNZ	BNZ	BNZ	NBNZ	BNZ
Width of CI	38.85	32.62	34.6	25.84	53.26	29.48
Number of studies	3	17	16	19	3	19
Sleep efficiency (% points: mean difference)						
Significant	Yes	No	No	Yes	Yes	No
Favours	BNZ	BNZ	NBNZ	BNZ	BNZ	BNZ
Width of CI	4.2	5.44	7.92	3.71	6.03	5.21
Number of studies	3	16	13	16	3	16
Total sleep time (minutes: mean difference)						
Significant	Yes	No	No	Yes	No	No
Favours	NBNZ	NBNZ	NBNZ	NBNZ	NBNZ	NBNZ
Width of CI	23.48	27.28	33.86	19.29	24.88	20.79
Number of studies	8	37	31	39	8	39
Sleep quality (standardized mean difference)						
Significant	No	Yes	Yes	Yes	No	Yes
Favours	NBNZ	NBNZ	NBNZ	NBNZ	NBNZ	NBNZ
Width of CI	0.31	0.36	0.44	0.26	0.34	0.29
Number of studies	11	45	38	49	11	49
Adverse events (risk difference)						
Significant	Yes	Yes	Yes	Yes	No	Yes
Favours	NBNZ	NBNZ	NBNZ	NBNZ	NBNZ	NBNZ
Width of CI	0.14	0.12	0.16	0.11	0.21	0.13
Number of studies	17	59	52	69	17	69

BNZ=benzodiazepines; CI=confidence interval; NBNZ=nonbenzodiazepines

Adapted from Vandermeer BW, Buscemi N, Liang Y, Witmans M. Comparison of meta-analytic results of indirect, direct, and combined comparisons of drugs for chronic insomnia in adults: a case study. *Med Care* 2007;45(10 Supl 2):S166-S172, with permission from Wolters Kluwer Health.

The following observations were noted regarding the summary estimates calculated:

- *Number of studies:* A greater number of studies were involved in the indirect comparisons than in the direct comparisons.
- *Comparison of direct and indirect evidence:* For each outcome, the following specific comparisons were made: frequentist direct versus frequentist modified indirect, frequentist direct versus frequentist combined, and Bayesian direct versus Bayesian combined. In these comparisons, differences between methods were small when compared to the effect size for the outcomes sleep onset latency, wakefulness after sleep onset, total sleep time, and adverse events. For the outcomes Sleep Efficiency and Sleep Quality, there was a moderate difference between methods. It was noted that, for some outcomes, some methods resulted in statistically significant differences between the two drugs, while others did not. When the direct comparisons resulted in a statistically significant difference, but the indirect methods did not, the authors stated that the difference could be attributable to the wider CIs of the indirect methods compared to the direct, assuming the same sample size. When the direct estimates were not significant but the indirect estimates were significant, then the difference in samples size resulting from the inclusion of a larger number of trials may be the reason.
- *Comparison of frequentist and Bayesian methods:* The frequentist direct and combined estimates have narrower CIs than Bayesian direct and Bayesian combined estimates, respectively. The CIs for frequentist methods were narrower than Bayesian methods since the frequentist approach assumes a constant known between-study variance and the Bayesian method was based on a varying between-study variance parameter.
- Assuming all studies have equal standard errors and variances, approximately four times as many studies are required in the indirect comparison relative to the direct comparison in order for the variance of the direct and indirect comparison to be the same.
- The Bayesian methods were insensitive to the prior chosen for the mean, but were sensitive to the prior chosen for the between study variance; between study prior sensitivity was directly related to the number of studies in the analysis.

Vandermeer et al. concluded that: (1) the indirect evidence was not substantially different from the direct evidence and can at least be used in sensitivity analyses; and (2) that frequentist and Bayesian indirect comparisons should be considered when conducting meta-analysis.

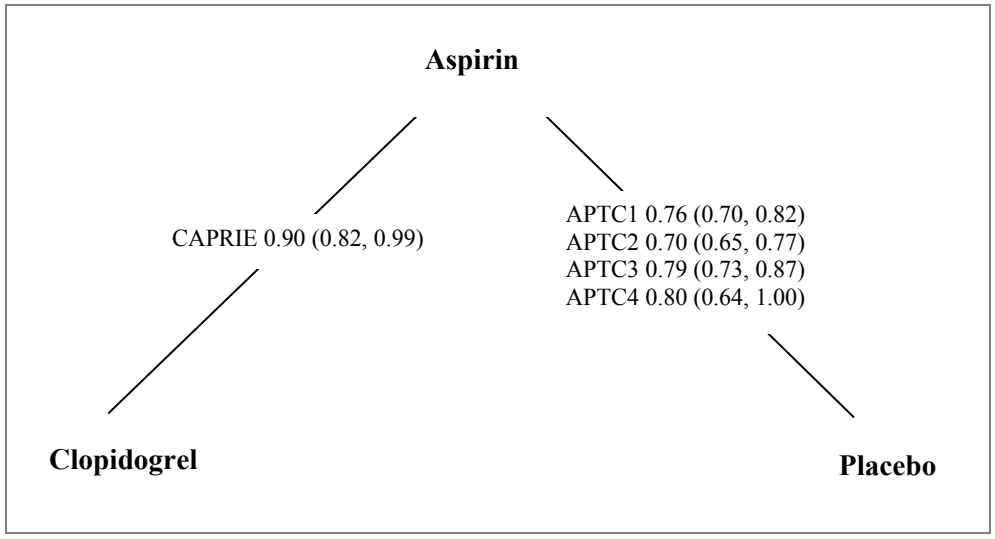
Three examples are considered in this section to explore the differences in the Bucher-adjusted indirect comparison, the Lumley network meta-analysis and the mixed treatment comparison methods.

6.1 Example 1: Clopidrogrel Versus Placebo in the Development of the Composite Outcome of Stroke, Myocardial Infraction or Vascular Death

Fisher et al.[1] (2001) used the adjusted indirect comparison methodology of Bucher to evaluate clopidogrel versus placebo. The investigators used aspirin as the common comparator, and their analysis included the CAPRIE trial,[29] which assessed clopidogrel versus aspirin, and the Antiplatelet Trialists' Collaboration (APTC) meta-analysis,[30] which compared aspirin to placebo. Fisher et al. calculated the adjusted indirect estimate of the OR and its 95% CI for various outcomes. In particular, a composite outcome consisting of stroke, myocardial infarction, and

vascular death (including hemorrhage) was analyzed. The APTC considered a number of meta-analyses for various patient subgroups, and in making this comparison of aspirin to placebo, the meta-analysis studies categorized under prior myocardial infarction (11 studies), acute myocardial infarction (eight studies), prior stroke/transient ischemic attack (15 studies), and intermittent claudication (22 studies) were homogeneous with respect to the odds ratio for the composite outcome. So, for this example, the meta-analytic result for each of these subgroups was considered as a separate "study" in the analysis. This example has been described in Section 5.2, and the estimate calculated using Bucher's adjusted indirect comparison method comparing clopidrogrel to placebo was OR = 0.71 (95%CI: 0.64, 0.78) (Figure 24). In this section, the mixed treatment comparisons method is used to analyze the data.

Figure 24: Evidence Network of the Reported Clinical Trials of Clopidogrel and Aspirin for the Occurrence of the Composite Outcome of Stroke, Myocardial Infarction, Or Vascular Death Considered in Fisher's Analysis Using the Bucher Adjusted Indirect Comparison Method*



*The summary odds ratio and associated 95% confidence interval appear below trial names in which drugs were compared directly; the arrowhead is directed to the drug for which there is a lower risk of the occurrence of the composite outcome of stroke, myocardial infarction or vascular death.

To apply the mixed treatment comparisons method, the frequency data of the number of events (i.e., composite outcome of any stroke, myocardial infarction, or vascular death) and the number of patients were needed (Table 10). The code and data used for applying the mixed treatment comparisons method is given in Box 1.

Table 10: Frequency Data Needed for Applying the Mixed Treatment Comparison Method to Clopidogrel, Aspirin, and Placebo for the Occurrence of the Composite Outcome of Stroke, Myocardial Infarction, or Vascular Death

Study	Placebo (Events/n)*	Aspirin (Events/n)	Clopidogrel (Events/n)
CAPRIE		1021/17519	939/17636
APTC meta-analysis (prior MI)	1693/9914	1331/9877	
APTC meta-analysis (acute MI)	1348/9385	992/9388	
APTC meta-analysis (prior stroke/TIA)	1301/5870	1076/5837	
APTC meta-analysis (intermittent claudication)	195/1649	160/1646	

APTC=Antiplatelet Trialists' Collaboration; MI=myocardial infarction; TIA=transient ischemic attack
* Event refers to the occurrence of the composite outcome of any stroke, myocardial infarction, or vascular death and n is the number of patients

Box 1: WinBUGS Code and Data for Applying the Mixed Treatment Comparison Method to Clopidogrel, Aspirin, and Placebo for the Occurrence of the Composite Outcome of Stroke, Myocardial infarction, or Vascular Death[20,21,25]

Random effects model: no correlation structure in multi-arm trials

```
model{

for(i in 1:N)  {   logit(p[i])<-mu[s[i]]+ delta[i]  * (1-equals(t[i],b[i]))          # model
                   r[i]~dbin(p[i],n[i])                                          # binomial likelihood
                   delta[i] ~ dnorm(md[i],tau)                                   # trial-specific LOR distributions
                   md[i] <- d[t[i]] - d[b[i]]   }                                # means of LOR distributions

for(j in 1:NS){ mu[j]~dnorm(0,.0001) }                                           # vague priors for 5 trial baselines

d[1]<-0
for (k in 2:NT)  {d[k] ~ dnorm(0,.0001) }                                        # vague priors for basic parameters

sd~dunif(0,2)                                                                    # vague prior for random effects standard deviation
tau<-1/pow(sd,2)


# Absolute treatment effects on mean response on Aspirin over 5 trials
mA ~ dnorm(0,1)
for (k in 1:NT)   { logit(T[k])<- mA  +d[k] }

# Ranking and probability {treatment k is best}
for (k in 1:NT) { rk[k]<- NT+1 - rank(T[],k)
                  best[k]<-equals(rk[k],1)}

# Pairwise odds ratios
for (c in 1:(NT-1))
        {  for (k in (c+1):NT)
             {  lor[c,k] <- d[k] - d[c]
                log(or[c,k]) <- lor[c,k]
             }
         }
}
```

Data:

```
#  treatment definitions: 1=aspirin, 2=clopidogrel, 3=placebo

list(N=10, NS=5, NT=3)

s[] t[]     r[]       n[]   b[]
1   1    1021  17519   1
1   2     939  17636   1
2   1    1331   9877   1
2   3    1693   9914   1
3   1     992   9388   1
3   3    1348   9385   1
4   1    1076   5837   1
4   3    1301   5870   1
5   1     160   1646   1
5   3     195   1649   1
END

#initial
list(
d=c(NA,0,0),
sd=1,
mu=c(0,0,0,0,  0,0,0,0,  0,0,0,0,  0,0,0,0,  0,0,0,0)
)
```

Source: https://www.bris.ac.uk/cobm/research/mpes/mixed-treatment-comparisons.html, reproduced with permission.

Applying the MTC method to these data yielded similar results to those obtained using the Bucher method (Table 11), with a slightly smaller point estimate for OR (0.67 versus 0.71) and a wider credible interval (0.32 versus 0.14). Recall that the estimated bias and MSE associated with the Bucher indirect estimate were quite large, and care must be exercised in interpreting this indirect estimate. Given the similarity of the results, the same caution must be noted for the mixed treatment comparisons method.

Table 11: Comparison of Clopidogrel and Placebo Using the Bucher-Adjusted Indirect Comparison Method and the Mixed Treatment Comparison Method for the Occurrence of the Composite Outcome of Stroke, Myocardial Infarction, or Vascular Death

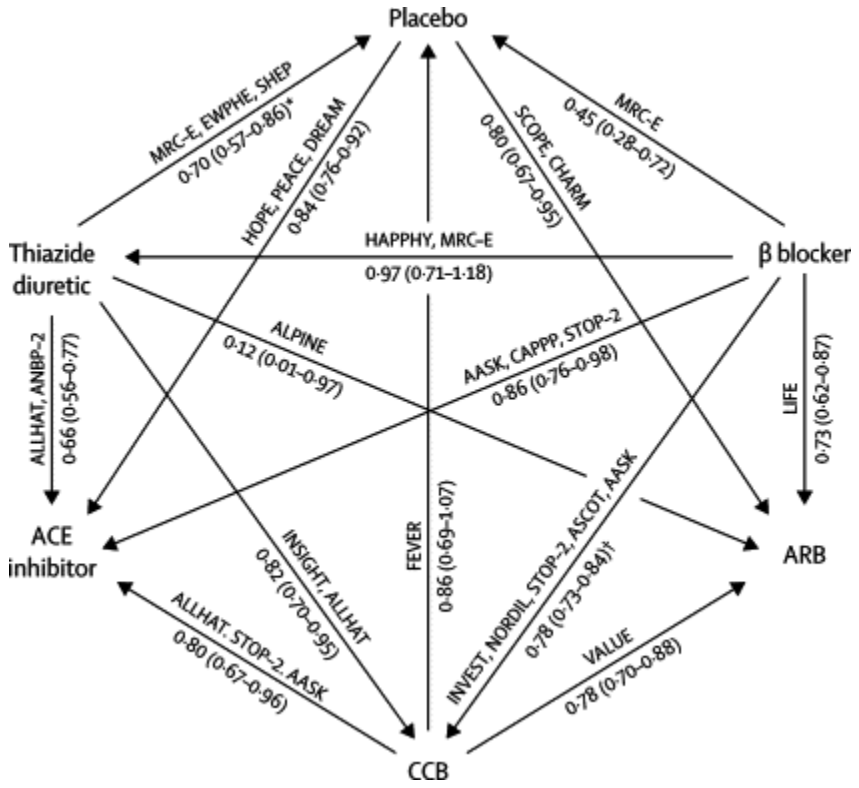
Comparison	Bucher OR (95% CI)	MTC OR (95% CrI)
Clopidogrel versus Placebo	0.71 (0.64,0.78)	0.67 (0.54,0.86)

OR=odds ration

6.2 Example 2: Antihypertensive Drugs and Development of Incident Diabetes

A complex network of clinical trials of antihypertensive drugs was reported by Elliott and Meyer[9,10] in assessing the development of incident diabetes (Figure 25). The Lumley network meta-analysis method was used to analyze the data. In this section, the mixed treatment comparisons method is used to analyze the data.

Figure 25: Evidence Network of the Reported Clinical Trials of Antihypertensive Drugs for the Development of Incident Diabetes Considered in Elliot and Meyer's Analysis Using the Lumley Network Meta-analysis Method*[9]



* The summary odds ratio and associated 95% confidence interval appear below the trial names in which drug classes were compared directly; the arrowhead is directed to the drug class for which there is a lower risk of developing incident diabetes

Reproduced from *The Lancet*, 369 (9557), 201-7, Elliott WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis[9] Copyright © (2007), with permission from Elsevier.

To apply the mixed treatment comparisons method, the frequency data of the number of events (i.e., incident diabetes) and the number of patients are needed (Table 12). The code and data used for applying the mixed treatment comparisons method is given in Box 2.

Table 12: Frequency Data needed for Applying the Mixed Treatment Comparison Method to Antihypertensive Drugs for the Development of Incident Diabetes

Study	Diuretic (Events/n)*	B Blocker (Events/n)	Placebo (Events/n)	CCB (events/n)	ACE (Events/n)	ARB (Events/n)
HAPPHY	75/3272	86/3297				
MRC-E	43/1081	37/1102	34/2213			
EWPHE	29/416		20/424			
SHEP	140/1631		118/1578			
INSIGHT	176/2511			136/2508		
ALLHAT	302/6766			154/3954	119/4096	
ANBP-2	200/2826				138/2800	
ALPINE	8/196					1/196
INVEST		665/8078		569/8098		
NORDIL		251/5059		216/5095		
ASCOT		799/7040		567/7072		
AASK		70/405		32/202	45/410	
STOP-2		97/1960		95/1965	93/1970	
CAPPP		380/5230			337/5183	
LIFE		320/3979				242/4020
FEVER			154/4870	177/4841		
HOPE			155/2883		102/2837	
PEACE			399/3472		335/3432	
DREAM			489/2646		449/2623	
SCOPE			115/2175			93/2167
CHARM			202/2721			163/2715
VALUE				845/5074		690/5087

* Event refers to the development of incident diabetes and n is the number of patients

Box 2: WinBUGS Code and Data for Applying the Mixed Treatment Comparison Method to Antihypertensive Drugs for the Development of Incident Diabetes[20,21,25]

Random effects model: includes correlation structure for 3-arm trials

```
model{
sw[1] <- 0
for(i in 1:N)  {
      logit(p[i])<-mu[s[i]]+ delta[i]  * (1-equals(t[i],b[i]))                    #  model
      r[i]~dbin(p[i],n[i])                                                       #  binomial likelihood
       delta[i] ~ dnorm(md[i],taud[i])                                           # trial-specific LOR distributions
       taud[i] <- tau * (1 + equals(m[i],3) /3)                                  # precisions of LOR distributions
      md[i] <- d[t[i]] - d[b[i]]  +  equals(m[i],3) * sw[i]                      # means of LOR distributions
      }
for (i in 2:N)  {   sw[i] <- (delta[i-1] -  d[t[i-1]] + d[b[i-1]])/2}           # adjustment for 3-arm trials

for(j in 1:NS){ mu[j]~dnorm(0,.0001) }                                          # vague priors for 22 trial baselines
```

```
d[1]<-0
for (k in 2:NT)  {d[k] ~ dnorm(0,.0001) }                    # vague priors for 6 basic parameters
 sd~dunif(0,2)                                    # vague prior for random effects standard deviation
tau<-1/pow(sd,2)

# Absolute treatment effects based on mean response on Diuretics over 8 trials
mA ~ dnorm(0,1)
for (k in 1:NT)   { logit(T[k])<- mA  +d[k] }

# Ranking and probability {treatment k is best}
for (k in 1:NT) { rk[k]<- NT+1 - rank(T[],k)
                  best[k]<-equals(rk[k],1)}

# All pairwise odds ratios
for (c in 1:(NT-1))
        {  for (k in (c+1):NT)
             {  lor[c,k] <- d[k] - d[c]
               log(or[c,k]) <- lor[c,k]
             }
         }
}
```

Data:

treatment definitions: 1=diuretic, 2=beta blocker, 3=CCB, 4=placebo, 5=ACE inhibitor, 6=ARB

list(N=48, NS=22, NT=6)

s[]	t[]	r[]	n[]	b[]	m[]
1	1	75	3272	1	1
1	2	86	3297	1	2
2	1	43	1081	1	1
2	2	37	1102	1	2
2	3	34	2213	1	3
3	1	29	416	1	1
3	3	20	424	1	2
4	1	140	1631	1	1
4	3	118	1578	1	2
5	1	176	2511	1	1
5	4	136	2508	1	2
6	1	302	6766	1	1
6	4	154	3954	1	2
6	5	119	4096	1	3
7	1	200	2826	1	1
7	5	138	2800	1	2
8	1	8	196	1	1
8	6	1	196	1	2
9	2	665	8078	2	1
9	4	569	8098	2	2
10	2	251	5059	2	1
10	4	216	5095	2	2
11	2	799	7040	2	1
11	4	567	7072	2	2
12	2	70	405	2	1
12	4	32	202	2	2
12	5	45	410	2	3

```
13  2   97  1960   2  1
13  4   95  1965   2  2
13  5   93  1970   2  3
14  2  380  5230   2  1
14  5  337  5183   2  2
15  2  320  3979   2  1
15  6  242  4020   2  2
16  3  154  4870   3  1
16  4  177  4841   3  2
17  3  155  2883   3  1
17  5  102  2837   3  2
18  3  399  3472   3  1
18  5  335  3432   3  2
19  3  489  2646   3  1
19  5  449  2623   3  2
20  3  115  2175   3  1
20  6   93  2167   3  2
21  3  202  2721   3  1
21  6  163  2715   3  2
22  4  845  5074   4  1
22  6  690  5087   4  2
END

#initial
list(
d=c(NA,0,0,0,0,0),
sd=1,
mu=c(0,0,0,0,0,0,0,  0,0,0,0,0,0,0,  0,0,0,0,0,0,0,  0,0,0,0,0,0,0,  0,0,0,0,0,0,0)
)
```

Source: https://www.bris.ac.uk/cobm/research/mpes/mixed-treatment-comparisons.html, reproduced with permission.

Applying the mixed treatment comparisons method to these data yielded similar results to those obtained using the Lumley network meta-analysis method (Table 13). The point estimates for odds ratios were sometimes larger or smaller using the mixed treatment comparisons method compared to those using Lumley's method, but the confidence intervals resulting from the mixed treatment comparisons method were always narrower.

Table 13: Comparisons of Various Antihypertensive Drugs and Diuretics Using the Lumley Network Meta-analysis Method and the Mixed Treatment Comparison Method for the Development of Incident Diabetes*

Comparison	MTC OR (95% CrI)	Lumley OR (95% CI)+
β blocker vs. Diuretic	0.93 (0.80,1.08)	0.93 (0.78,1.11)
CCB vs. Diuretic	0.74 (0.63,0.86)	0.79 (0.67,0.92)
Placebo vs. Diuretic	0.78 (0.68,0.90)	0.75 (0.63,0.89)
ACE inhibitor vs. Diuretic	0.66 (0.57,0.76)	0.67 (0.57,0.79)
ARB vs. Diuretic	0.62 (0.51,0.73)	0.62 (0.51,0.77)

CI=confidence interval; MTC=mixed treatment comparisons method; OR=odds ratio;vs.=versus
* Diuretic was used as the standard for comparisons as recommended.
+ Incoherence: 0.054

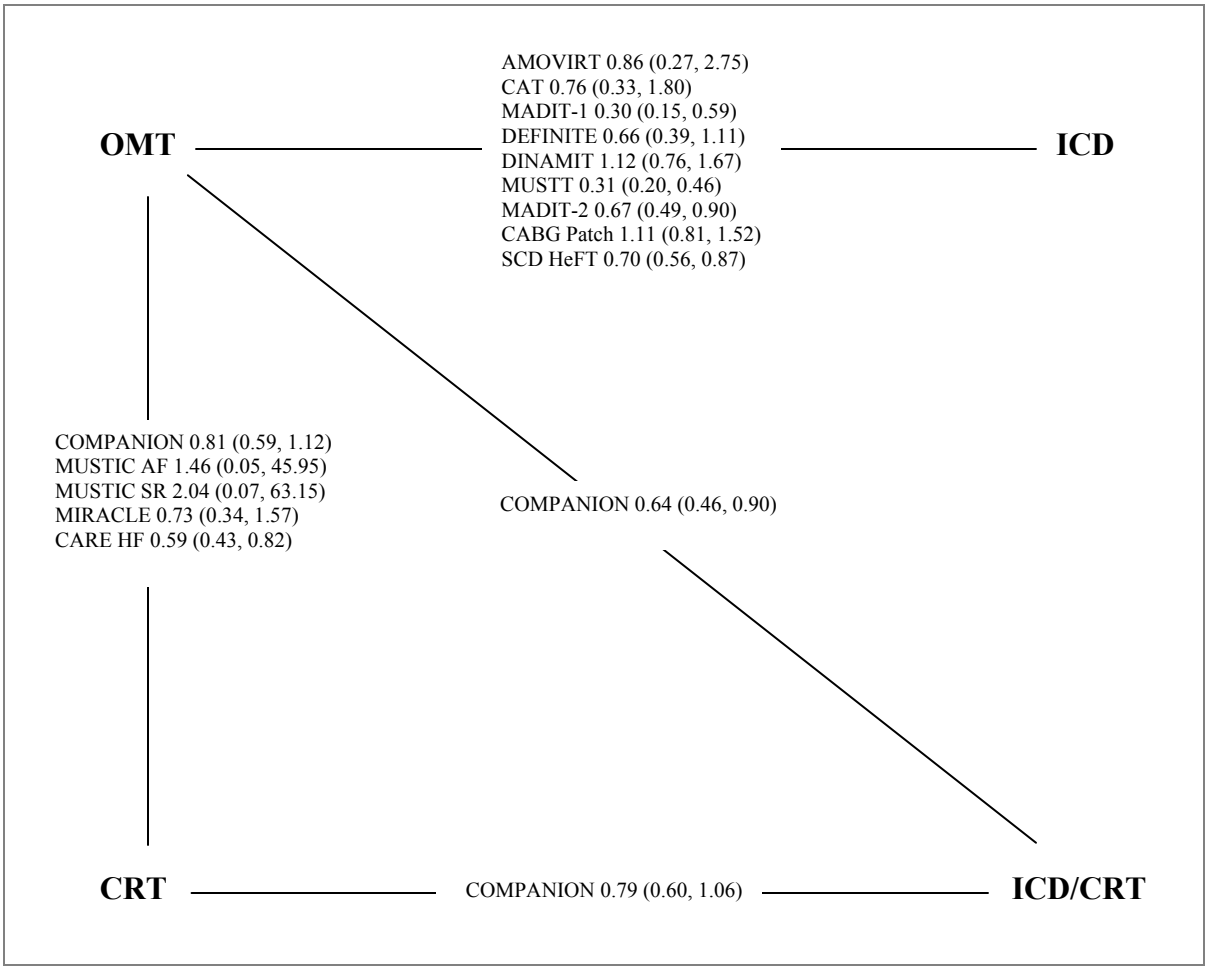
6.3 Example 3: Cardiac Devices ICD/CRT versus ICD Alone and Total Mortality in Heart Failure Patients

An international, multi-centre, randomized controlled trial known as the RAFT study (Resynchronization/Defibrillation for Ambulatory Heart Failure Trial)[38] is currently being conducted. The hypothesis being evaluated in this cardiac device trial is that, in patients with left ventricular (LV) dysfunction (ejection fraction ≤ 30%) and QRS duration ≥ 120 ms with moderate to severe congestive heart failure symptoms, the addition of cardiac resynchronization therapy (CRT) to Implantable Cardioverter Defibrillator (ICD) and optimal medical therapy (OMT) reduces the combined end point of mortality and CHF hospitalization. In the study, patients are randomized in a 1:1 proportion to: ICD (single or dual chamber) or ICD/CRT. The randomization is stratified for centre and single/dual ICD indication. Patients and heart failure care personnel are blinded, and only device care personnel are unblinded. The primary outcome measure is a composite of total mortality or hospitalization for CHF, where hospitalization for CHF is defined as an admission to hospital with a diagnosis of worsening CHF for > 24 hours. For this illustration, only total mortality is considered.

Emerging information from two recent device trials put into question recruitment of New York Heart Association (NYHA) Class III patients into the RAFT study. The COMPANION Trial was a three-arm, randomized controlled trial comparing ICD/CRT, CRT, and OMT, while the CARE-HF trial was a two-arm randomized controlled trial comparing ICD and OMT. Taking the results of these trials into consideration, the American Heart Association recommended ICD therapy for NYHA Class III patients. Although ICD/CRT has never been compared to ICD alone in a well-designed, randomized controlled trial, the decision by the American Heart Association has effectively prevented such a study from being conducted in NYHA Class III patients. Although the RAFT study continues to recruit NYHA Class II patients, Class III patients cannot be enrolled in the study, and the results for these patients will never be known.

A network meta-analysis of clinical trials comparing OMT, ICD, CRT, and ICD/CRT has been conducted assessing total mortality in heart failure patients (Figure 26). To compare ICD/CRT versus ICD alone, the Bucher-adjusted indirect comparison method, the Lumley network meta-analysis method, and the mixed treatment comparison method were used.

Figure 26: Evidence Network of the Reported Clinical Trials of Optimal Medical Therapy, Implantable Cardioverter Defibrillator (ICD), Cardiac Resynchronization Therapy (CRT) and ICD/CRT for Total Mortality in Heart Failure Patients*



CRT=cardiac resynchronization therapy; ICD= Implantable Cardioverter Defibrillator; OMT=optimal medical therapy;
*The summary odds ratio and associated 95% confidence interval appear below trial names in which treatments were compared directly; the arrowhead is directed to the treatment for which there is a lower risk of mortality.

To apply the optimal medical therapy, the frequency data of the number of events (i.e., all cause mortality) and the number of patients are needed (Table 14). For illustrative purposes, various branches of the overall network are considered in order to provide different network patterns, namely: simple star, star, ladder, and at least one closed loop. The code and data used to identify the optimal medical therapy for these various network patterns are given in Box 3.

Table 14: Frequency Data Needed for Applying the Mixed Treatment Comparison Method to Implantable Cardioverter Defibrillator (ICD), Cardiac Resynchronization Therapy, (CRT), and ICD/CRT Cardiac Devices for Total Mortality

Study	OMT (Events/n)*	ICD (Events/n)	CRT (Events/n)	ICD/CRT (Events/n)
AMOVIRT	7/52	6/51		
CAT	17/54	13/50		
MADIT 1	39/101	15/95		
DEFINITE	40/229	28/229		
DINAMIT	58/342	62/332		
MUSTT	255/537	35/161		
MADIT 2	97/490	105/742		
CABG Patch	95/454	101/446		
SCD HeFT	244/847	182/829		
MUSTIC SR	0/29		1/29	
MUSTIC AF	0/18		1/25	
MIRACLE	16/225		12/228	
CARE HF	120/404		82/409	
COMPANION	77/308		131/617	105/595

CRT=Cardiac Resynchronization Therapy; ICD=Implantable Cardioverter Defibrillator; OMT=optimal medical therapy
* Event refers to any cause mortality and n is the number of heart failure patients

Box 3: WinBUGS Code and Data for Applying the Mixed Treatment Comparison Method to Implantable Cardioverter Defibrillator (ICD), Cardiac Resynchronization Therapy, (CRT), and ICD/CRT Cardiac Devices for Total Mortality[20,21,25]

Random effects model: no correlation structure in multi-arm trials

```
model{

for(i in 1:N)  {   logit(p[i])<-mu[s[i]]+ delta[i]  * (1-equals(t[i],b[i]))        # model
                   r[i]~dbin(p[i],n[i])                                          # binomial likelihood
                   delta[i] ~ dnorm(md[i],tau)                                   # random effects: trial-specific LORs
                   md[i] <- d[t[i]] - d[b[i]]   }                                # means of trials-specific LORs

for(j in 1:NS){ mu[j]~dnorm(0,.0001) }                                           # vague priors for trial baselines

d[1]<-0
for (k in 2:NT)  {d[k] ~ dnorm(0,.0001) }                                        # vague priors for basic parameters

sd~dunif(0,2)                                                                    # vague prior for random effects standard deviation
tau<-1/pow(sd,2)


# Absolute treatment effects on mean response on OMT over trials involved
mA ~ dnorm(0,1)
for (k in 1:NT)   { logit(T[k])<- mA  +d[k] }
```

```
# Ranking and probability {treatment k is best}
for (k in 1:NT) { rk[k]<- NT+1 - rank(T[],k)
                  best[k]<-equals(rk[k],1)}

# Pairwise odds ratios
for (c in 1:(NT-1))
        {  for (k in (c+1):NT)
              {  lor[c,k] <- d[k] - d[c]
                 log(or[c,k]) <- lor[c,k]
              }
         }
}
```

Data:

```
#2 arm star
list(N=20, NS=10, NT=3)

s[]  t[]  r[]      n[]   b[]
1    1     7      52    1
1    2     6      51    1
2    1    17      54    1
2    2    13      50    1
3    1    39     101    1
3    2    15      95    1
4    1    40     229    1
4    2    28     229    1
5    1    58     342    1
5    2    62     332    1
6    1   255     537    1
6    2    35     161    1
7    1    97     490    1
7    2   105     742    1
8    1    95     454    1
8    2   101     446    1
9    1   244     847    1
9    2   182     829    1
10   1    77     308    1
10   3   105     595    1
END

#initial
list(
d=c(NA,0,0),
sd=1,
mu=c(0,0,0,0,  0,0,0,0,  0,0,0,0,  0,0,0,0,  0,0,0,0)
)

#3 arm star
list(N=30, NS=15, NT=4)

s[]  t[]   r[]      n[]   b[]
1    1     7      52    1
1    2     6      51    1
2    1    17      54    1
2    2    13      50    1
```

```
3   1   39   101   1
3   2   15    95   1
4   1   40   229   1
4   2   28   229   1
5   1   58   342   1
5   2   62   332   1
6   1  255   537   1
6   2   35   161   1
7   1   97   490   1
7   2  105   742   1
8   1   95   454   1
8   2  101   446   1
9   1  244   847   1
9   2  182   829   1
10   1    77   308   1
10   3   105   595   1
11  1     0    18   1
11  3     1    25   1
12  1    16   225   1
12  3    12   228   1
13  1   120   404   1
13  3    82   409   1
14  1    77   308   1
14  3   131   617   1
15  1    77   308   1
15  4   105   595   1
END

#initial
list(
d=c(NA,0,0,0),
sd=1,
mu=c(0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0)
)

#ladder design
list(N=30, NS=15, NT=4)

s[]  t[]   r[]     n[]    b[]
1    1      7      52    1
1    2      6      51    1
2    1     17      54    1
2    2     13      50    1
3    1     39     101    1
3    2     15      95    1
4    1     40     229    1
4    2     28     229    1
5    1     58     342    1
5    2     62     332    1
6    1    255     537    1
6    2     35     161    1
7    1     97     490    1
7    2    105     742    1
8    1     95     454    1
8    2    101     446    1
9    1    244     847    1
```

```
9   2   182   829   1
10   1     0    29   1
10   3     1    29   1
11   1     0    18   1
11   3     1    25   1
12   1    16   225   1
12   3    12   228   1
13   1   120   404   1
13   3    82   409   1
14   1    77   308   1
14   3   131   617   1
15   3   131   617   3
15   4   105   595   3
END

#initial
list(
d=c(NA,0,0,0),
sd=1,
mu=c(0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0)
)

#abd design
list(N=32, NS=16, NT=4)

s[]  t[]   r[]     n[]   b[]
1    1     7      52    1
1    2     6      51    1
2    1    17      54    1
2    2    13      50    1
3    1    39     101    1
3    2    15      95    1
4    1    40     229    1
4    2    28     229    2
5    1    58     342    1
5    2    62     332    1
6    1   255     537    1
6    2    35     161    1
7    1    97     490    1
7    2   105     742    1
8    1    95     454    1
8    2   101     446    1
9    1   244     847    1
9    2   182     829    1
10   1     0      29    1
10   3     1      29    1
11   1     0      18    1
11   3     1      25    1
12   1    16     225    1
12   3    12     228    1
13   1   120     404    1
13   3    82     409    1
14   1    77     308    1
14   3   131     617    1
15   1    77     308    1
15   4   105     595    1
```

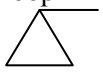
```
16  3    131   617    3
16  4    105   595    3
END

#initial
list(
d=c(NA,0,0,0),
sd=1,
mu=c(0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0,  0,0,0,0,0)
)
```

Source: https://www.bris.ac.uk/cobm/research/mpes/mixed-treatment-comparisons.html, reproduced with permission.

The results of applying the Bucher, Lumley, and mixed treatment comparison methods to these data are summarized in Table 15. The mixed treatment comparison method could be used for all four patterns, the Bucher method could be used directly for the simple star and ladder patterns, and the Lumley method could only be used for the closed loop. Although never reaching statistical significance, the point estimates for odds ratio always exceeded 1 when the mixed treatment comparison method was used, and were less than 1 for the Bucher and Lumley methods. The confidence intervals were always narrower for the Bucher method. For the one pattern in which the Lumley method could be used, the incoherence was large (0.38). As a result, the 95% CI was (-0.87, 2.75) when adding the incoherence (0.38) and standard deviation (0.84) of the odd ratio (0.94). As the incoherence was high,we got a negative lower confident limit. The 95% CI is presented as (0, 2.75) in Table 15, but it indicates that the odds ratios and 95% CIs from the Lumley method are not valid when the incoherence is high.

Table 15: Comparison of ICD/CRT versus ICD Cardiac Devices Using the Bucher-Adjusted Indirect Method, the Lumley Network Analysis Method, and the Mixed Treatment Comparison Method for Total Mortality

Pattern	Branches of Network Used in the Estimation: OR (95% CI)	ICD/CRT versus ICD		
		Bucher OR (95% CI)	Lumley OR (95% CI)	MTC OR (95% CI)
Simple star 	ICD vs OMT: 0.66(0.51,0.86) ICD/CRT vs OMT: 0.64(0.46,0.90)	0.97 (0.62,1.53)	NA	1.20 (0.37,2.61)
Star 	CRT vs OMT: 0.70(0.57,0.88) ICD vs OMT: 0.66(0.51,0.86) ICD/CRT vs OMT: 0.64(0.46,0.90)	NA*	NA	1.16 (0.38,2.49)
Ladder 	ICD/CRT vs CRT: 0.79(0.60,1.06) CRT vs OMT: 0.70(0.57,0.88) OMT vs ICD: 1.52(1.16,1.96)	0.84 (0.53,1.35)	NA	1.16 (0.32,2.59)
At least one closed loop 	ICD/CRT vs CRT: 0.79(0.60,1.06) CRT vs OMT: 0.70(0.57,0.88) ICD/CRT vs OMT: 0.64(0.46,0.90) OMT vs ICD: 1.52(1.16,1.96)	NA	0.94 (0.00,2.75)	1.06 (0.49,1.95)

CI=confidence interval; CRT=Cardiac Resynchronization Therapy; ICD=Implantable Cardioverter Defibrillator; OMT=optimal medical therapy; vs=versus
*NA=Not Applicable

6.4 Summary

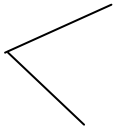
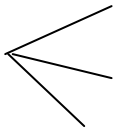
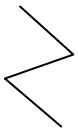
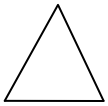
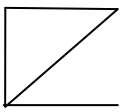
The Bucher-adjusted indirect comparison, the Lumley network meta-analysis, and the mixed treatment comparison methods can lead to different estimates of the odds ratio effect measure. Based on the examples, the following observations are made:

- Although point estimates were not the same, all methods agreed regarding the statistical significance of the effect measure.
- Point estimates could differ by being on opposite sides of the neutral point OR=1.
- Confidence intervals were the widest for the Lumley method, and the confidence intervals for the mixed treatment comparison method were wider than for the Bucher method.
- The software is readily available for computing estimates for odds ratio using all three methods, but not for the other effect measures except for the Bucher method.
- The mixed treatment comparison method could be used for all network patterns.
- The Bucher method required the least amount of information for computation.

7 SUMMARY

The review of the various approaches for indirect treatment comparisons has identified three general approaches: the Bucher-adjusted indirect comparison method, the Lumley network analysis method, and the mixed treatment comparison method. Table 16 indicates the various different networks of evidence that can be analyzed by the indirect comparison methods. These networks represent the star, ladder, and closed and partially closed-loop designs. The mixed treatment comparison method can be used to obtain measures of effect for each of the indicated patterns. The network meta-analysis method proposed by Lumley can compare treatments in a network geometry that contains at least one closed loop. The adjusted indirect comparison method proposed by Bucher can be used to evaluate the effect of treatments that form a simple star design. The Bucher method has been proposed to perform indirect comparisons when direct evidence is not available, and the method is not applicable to the closed loop pattern. For the other designs, the Bucher method can be used to determine the indirect evidence of the pairwise contrasts that have not been directly compared in the star, ladder, and network with one closed loop designs.

Table 16: Network Patterns that the Various Indirect Treatment Comparison Methods Can Process

Pattern Description	Network Pattern	Indirect Comparison Method		
		Bucher Method	Network Meta-analysis	Mixed Treatment Comparison
Simple star		✓	—	✓
Star		— (Pairwise contrasts)	—	✓
Ladder		✓	—	✓
Closed loop		—	✓	✓
Network with at least one closed loop		✓ (Pairwise contrasts)	✓	✓

When mixed treatment comparison or network meta-analysis is used to evaluate the evidence network depicted by the closed loop pattern, the methods can simultaneously combine direct and indirect evidence, and can evaluate the incoherence of the closed loop. The variance parameter w^2 from Lumley's model is equivalent to inconsistency variance σ_w^2 estimated in the mixed treatment comparison models.[20] However, the two methods will calculate different values for treatment effects because of differences in the way that inconsistency is evaluated. As indicated by Salanti et al.[16] and Lu and Ades,[20] in the network meta-analysis approach, the number of incoherence terms ξ_{ij} is equal to the number of different comparisons. In the mixed treatment comparison framework, the number of inconsistency terms is equal to the number of different independent closed loops.

The mixed treatment comparison model described by Lu and Ades measures the relative efficacy of treatments using the log odds ratio effect measure. Various investigators have performed the mixed treatment comparison for other effect measures. For instance, Vandermeer et al. considered direct and indirect evidence to evaluate the relative efficacy between benzodiazepines and nonbenzodiazepenes[26] based on mean differences for five of their clinical outcomes and risk difference on the adverse event outcome. Jansen et al.[22] have outlined a mixed treatment comparison model to be applied to continuous outcomes.

Various approaches for indirect treatment comparisons have been reviewed. The mixed treatment comparison approaches by Lu and Ades are elegant, but require information that may not be available. The challenge of Lumley's network meta-analysis is that it needs a data-based assessment of trial consistency; therefore, it requires information from a large number of different treatment comparisons. When analyzing a network of comparisons, the inconsistency of the network needs to be considered, as well as between-trial heterogeneity and sampling error. Large inconsistencies rule out a meta-analysis, small inconsistencies should add uncertainty to the results. The inconsistency of the network can only be assessed for a closed loop of treatments, with more loops allowing for better diagnosis of consistency. Estimating inconsistency will be reliable to the extent that the trials in these closed loops are similar to other trials. In addition, consistency cannot be assessed for a star design comparing everything to placebo, or for a ladder design where new treatments are always compared to current standard.

The attractiveness of the Bucher approach is that it has been designed for application with minimal information to the common indirect treatment comparison involving a simple star design: using the direct comparisons X versus A and X versus B with the common comparator link "X" to yield an indirect comparison of A versus B. The Bucher approach has not been shown to work for the ladder design. That is, we have X versus C, C versus E, E versus F, F versus G, and we want to use the comparator links "C," "E," and "F" to yield an indirect comparison of X versus G. In Chapter 3, we extended the Bucher approach to apply to the ladder design and, as well, extended the approach for the effect measures relative risk, risk difference, hazard ratio, and mean difference.

We extended the Bucher approach to different measures of association and to the "ladder" design in which several direct comparisons can be linked by common comparators. A methodology for indirect evidence for both discrete and continuous outcomes has been developed by expanding the indirect odds ratio approach by Bucher et al. (1997) involving k direct comparisons. This

generalized approach was then considered for the relative risk, hazard ratio, risk difference, and mean difference. The indirect point and confidence interval estimators and test of association for these different effect parameters were derived.

The distributional and statistical properties of the Bucher-adjusted indirect comparison estimators have been explored using Monte Carlo simulations for the case of k=3 treatments. In particular:

Frequency distribution for the indirect estimators ln(relative risk), ln(odds ratio), ln(hazard ratio), risk difference, and mean difference all are mound shape and symmetric.

For relative risk and odds ratio, the bias, variance, and mean square error:

- for the direct and indirect estimators decrease as the event rate approaches 0.5
- for the indirect estimator are larger than the direct estimator; in particular, for small event rates
- for both direct and indirect estimators increase as the effect measure being estimated increases.

For risk difference and mean difference:

- for any combination of the parameters, the bias of both the direct and indirect estimators is small (theoretically zero) and of similar magnitude
- for any combination of the parameters, the variance of the indirect estimator is consistently larger than that for the direct estimator
- similarly for any combination of the parameters, the mean square error of the indirect estimator is consistently larger than that for the direct estimator.

Although, in theory, there is no limit to the number of treatments that can be included in the indirect comparisons, in practice, the number of treatments should be limited. The confidence intervals continue to increase in width as the number of treatments increase and become impractically large. Further, the point estimates for the risk difference and mean difference continue to increase.

Our conclusions about the degree of bias associated with direct versus indirect effect estimates are not consistent with results of a study by Song et al.[39] For the evaluation of the relative efficacy between new and conventional drugs, Song et al. examined the discrepancy between treatment effects based on direct comparisons versus indirect comparisons using the approach of Bucher. The results showed that the effect sizes of direct estimates were greater than those of indirect estimates. For this reason, the authors concluded that the results of direct comparisons may be associated with a greater amount of bias than results of indirect comparisons. More specifically, these investigators performed meta-analyses for three pairwise comparisons. The authors evaluated buproprion versus nicotine replacement therapy for smoking cessation, rispiridone versus haloperidol for schizophrenia, and fluoxetine versus imipramine for depressive disorders. For each comparison, a meta-analysis was performed by pooling together the results of trials in which the drugs were compared directly. Another meta-analysis was conducted for each comparison by pooling the results of trials in which each drug was compared to placebo, and then performing an indirect comparison. Treatment differences were measured on the odds ratio

scale. The results indicated that the effect size of the odds ratio from the meta-analyses, based on direct evidence, was larger than the effect size of the indirect odds ratio. Their conclusions need to be verified by further investigations. Interestingly, the authors performed a simulation study which indicated that when placebo-controlled trials of new drugs are associated with less bias than placebo-controlled trials of conventional drugs, the adjusted indirect comparison will underestimate the true treatment effect. Song et al. also showed that, for any given level of bias in placebo-controlled trials of new drugs, as the level of bias for placebo-controlled trials of conventional drugs increases, the bias of the adjusted indirect estimate decreases. Based on the results of their simulation study, the authors concluded that although indirect estimates could be biased, the magnitude of this bias may still be less than the bias of estimates based on direct evidence. However, their simulation study did not specifically investigate the discrepancy between direct and indirect estimates, and such a simulation study should be conducted.

A "reviewer-friendly" program was developed and made available to facilitate the evaluation of indirect evidence for reviewers. The Indirect Treatment Comparison program has been developed in Visual Basic to assist with the various calculations associated with indirect comparisons. It consists of two screens. On the first screen, the effect measure of interest is identified and information for each consecutive pair of treatments of interest is requested for the point estimate and 95% CI of the effect measure for each direct comparison involved in the indirect comparison. The resulting indirect comparison estimates for the effect measure and the 95% CI, as well as the P value for the test of association corresponding to this effect measure, are provided. On the second screen, the weights needed for a specific direct comparison is requested in order to calculate the test statistic for the test of association. There are various formats in which the information to calculate these weights can be provided and are identified through the weight selections (direct versus derived; fixed versus random), and the specific information for each study involved in the direct comparison is then identified and requested.

This report has expanded on a previously published report on indirect comparisons.[39] In a health technology assessment report, Glenny et al.[40] performed a comprehensive survey of the literature for published examples of indirect comparisons. The authors reported the frequency of indirect comparisons in the published literature and described the methods being used at the time of their review to obtain indirect treatment effects. The report also consisted of empirical investigations to compare effect parameters based on direct versus indirect evidence, and to analyze the discrepancy between results from each of these sources of evidence.

The objectives of the investigation have been met. We have: identified and reviewed the different popular methods available for making indirect treatment comparisons; derived general methods and procedures for effect measures of discrete and continuous outcomes within complex webs of evidence following a ladder design; determined the distributional properties of the indirect estimates using simulations and derived bias and mean square error tables and charts providing guidance on the indirect treatment comparison results; and developed a user-friendly program for conducting indirect treatment comparisons for the methods and procedures derived. We have replicated the indirect treatment comparison results in several examples from the literature using the indirect treatment comparisons program, and have used the bias and mean squared error tables and charts to assess the goodness of the adjusted indirect comparison

treatment estimates. Further, for the various methods identified, we have illustrated the application of these methods for indirect treatment comparisons.

Note of caution: In the absence of previously performed randomized controlled trials in which two interventions of interest have been compared, indirect methods may be used. However, indirect treatment comparisons should be restricted to those situations in which it is not possible to perform a direct head-to-head trial. Furthermore, if there is interest in comparing treatment A to treatment C and trials have compared each of A and C to treatment B separately, it is important to be certain that the fundamental assumption underlying this method for indirect comparisons is fulfilled: the effect of A, observed in the A versus B trials, is expected to have been constant had it been administered instead of C in the C versus B trials. Likewise, the observed effect of C should be expected to be constant in the A versus B trials.

Whether an indirect treatment comparison provides a valid estimate of the relative efficacy for an intervention of interest significantly depends on the fulfillment of this primary assumption. To determine whether or not this assumption is met, trials included in the indirect comparison can be assessed according to three criteria:

- comparability of the linking treatment;
- comparability of patients/heterogeneity;
- methodological comparability of included trials.

If there are significant differences in the aforementioned criteria between the two sets of trials, the effect estimate will not represent the true value. Differences in the linking treatment may occur when, for instance, an active control was administered at different doses in each set of studies, or when a placebo is not truly equivalent in both sets of studies. This latter situation occurs when there is heterogeneity between patient groups. Heterogeneity may arise when there are clinical differences in the two sets of studies related to diagnostic criteria, disease severity, follow-up time, trial setting, assessment of outcomes, chosen outcome measures, age, and sex. Heterogeneity in patients may not only affect the comparability of the linking group but, also, the consistency of the observed effect of the intervention. Further to this, if one set of trials — the A versus B trials, for instance — was of relatively weaker methodological quality than the other set of trials (C versus B), the effect of A may be exaggerated and could not be expected to be reproducible if A were administered in place of C in the C versus B trials.

The validity of the indirect comparison also depends on the following additional criteria:

- the inclusion of non randomized studies
- date of publication.

Because non-randomized studies are associated with unmeasureable biases that can only be accounted for through the conduct of randomized controlled trials, these types of studies should not be included in an indirect comparison. Additionally, even in the case that the two sets of trials included in the indirect comparison are the same, if the trials used in the indirect comparison were published in a time period that does not reflect the current day clinical practice, any estimate derived from such comparisons may be insensible. This is especially important because indirect comparisons are often based on older trials. For instance, Yazdanpanah et al.[41] conducted an indirect treatment comparison to evaluate the relative efficacy of non-nucleoside

reverse transcriptase inhibitor (NNRTI) based triple therapy versus protease inhibitor-based triple therapy in preventing the primary outcome of an AIDS-defining disease or death among patients with advanced immunodeficiency. Each treatment had been compared to a two-drug regimen consisting of nucleoside analogue reverse transcriptase inhibitors (NRTIs). Based on the results of the comparison, Yazdanpanah et al. concluded that protease inhibitor-based triple therapy was more effective than NNRTI-based triple therapy (OR 0.54 [95% CI: 0.40, 0.73]).[41] However, a review article by Lundgren and Phillips[42] stated that Yazdanpanah's conclusions were based on protease inhibitors that are obsolete, due to their limited potency. Yazdanpanah et al.'s[41] results were not replicated in two recent randomized controlled trials, which directly compared NNRTI-based triple therapy to protease inhibitor-based triple therapy.

7.1 Future Work

- The methods and applications presented in this report apply to data from randomized controlled trials. Using data from non-randomized studies to perform indirect comparisons requires further methodological developments. Such methods would require the incorporation of techniques to adjust for biases associated with data from non-randomized studies. Using appropriate statistical methods, adjusted confidence interval estimates could be derived and used in the Bucher-adjusted indirect comparison method, but the interpretation of the bias in light of different levels of confounding should be investigated. For the Lumley network meta-analysis and mixed treatment comparison methods, the statistical models used will need to be generalized to incorporate potential confounding variables.
- A comprehensive simulation to formally evaluate the bias and mean square error properties of Lumley and mixed treatment comparison methods should be undertaken.
- This report extended the adjusted indirect comparison method for applicability to various measures of treatment effect. The use of the other indirect comparison methods to a range of different effect scales is in need of additional elaboration. The mixed treatment comparison model is formally based on the odds ratio. Further work is required to determine the validity of the results when the mixed treatment comparison model is applied to effect measures.
- This report has presented an indirect treatment comparison program to perform the Bucher-adjusted indirect comparison method for consecutive pairs of k treatments. The Lumley network meta-analysis method can be performed using statistical software programs such as S-plus or R, which are used for routine statistical analyses and programming.[1] The mixed treatment comparison method can be implemented in WinBUGs, a platform for performing advanced statistical procedures in a Bayesian framework.[2] Extensive statistical and computational knowledge is needed to properly use S-plus, R, and WinBUGS. The development of a software program that simplifies the process for performing network meta-analysis and mixed treatment comparison— such as the indirect treatment comparisons program described in this report — would be a valuable contribution.
- For treatment differences that are not statistically significant, a power analysis can be conducted to determine the power of the study to reject the null hypothesis. Considerations related to the power of studies are part of the general body of knowledge for statistical analyses. For this reason, a section on power has not been included in this report. However, a formal evaluation of power should be reviewed and, in addition, as indicated by Salanti et al.,[16] future research may indicate procedures through which results of meta-analyses based on indirect comparisons can be used to determine power when planning a new study.

8 REFERENCES

1. Fisher LD, Gent M, Büller HR. Active-control trials: how would a new agent compare with placebo? A method illustrated with clopidogrel, aspirin, and placebo. *Am Heart J* 2001;141(1):26-32.
2. McAlister FA, Laupacis A, Wells GA, Sackett DL. Users' guides to the medical literature: XIX. Applying clinical trial results B. Guidelines for determining whether a drug is exerting (more than) a class effect. *JAMA* 1999;282(14):1371-7.
3. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50(6):683-91.
4. Song F, Glenny AM, Altman DG. Indirect comparison in evaluating relative efficacy illustrated by antimicrobial prophylaxis in colorectal surgery. *Control Clin Trials* 2000;21(5):488-97.
5. Coomarasamy A, Knox EM, Gee H, Song F, Khan KS. Effectiveness of nifedipine versus atosiban for tocolysis in preterm labour: a meta-analysis with an indirect comparison of randomised trials. *BJOG* 2003;110(12):1045-9.
6. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21(16):2313-24.
7. Lam SK, Owen A. Incident diabetes in clinical trials of antihypertensive drugs. *Lancet* 2007;369(9572):1513-4.
8. Pocock SJ. Safety of drug-eluting stents: demystifying network meta-analysis. *Lancet* 2007;370(9605):2099-100.
9. Elliott WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis. *Lancet* 2007;369(9557):201-7.
10. Elliot WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis [erratum]. *Lancet* 2007;369(9572):1518.
11. Eddy DM, Hasselblad V, Shachter R. *Meta-analysis by the confidence profile method: the statistical synthesis of evidence* [Statistical modeling and decision science]. Boston: Academic Press; 1992.
12. Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Stat Med* 2003;22(19):2995-3016.
13. Moye LA. *Elementary Bayesian biostatistics*. Boca Roton: Chapman & Hall; 2008.
14. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian data analysis*. London: Chapman & Hall; 1995.
15. Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions*. Version 5.0.1 [updated September 2008]. Oxford: The Cochrane Collaboration; 2008. Available: http://www.cochrane-handbook.org (accessed 2009 Jan 23).
16. Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17(3):279-301.
17. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23(20):3105-24.
18. Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med* 1996;15(24):2733-49.
19. Smith TC, Spiegelhalter DJ, Thomas A. Bayesian approaches to random-effects meta-analysis: a comparative study. *Stat Med* 1995;14(24):2685-99.
20. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2006;101(474):447-59.
21. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331(7521):897-900.

22. Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value Health* 2008;11(5):956-64.

23. Boland A, Dundar Y, Bagust A, Haycox A, Hill R, Mujica MR, et al. Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation. *Health Technol Assess* 2003;7(15):1-136.

24. Keeley EC, Boura JA, Grines CL. Primary angioplasty versus intravenous thrombolytic therapy for acute myocardial infarction: a quantitative review of 23 randomised trials. *Lancet* 2003;361(9351):13-20.

25. Mixed treatment comparisons. In: *Community based medicine*. Bristol (UK): University of Bristol, Department of Community Based Medicine; 2009. Available: https://www.bris.ac.uk/cobm/research/mpes/mixed-treatment-comparisons.html (accessed 2009 Jan 23).

26. Vandermeer BW, Buscemi N, Liang Y, Witmans M. Comparison of meta-analytic results of indirect, direct, and combined comparisons of drugs for chronic insomnia in adults: a case study. *Med Care* 2007;45(10 Supl 2):S166-S172.

27. Wells GA, Cranney A, Boucher M, Peterson J, Shea B, Robinson V, et al. *Bisphosphonates for the primary and secondary prevention of osteoporotic fractures in postmenopausal women: a meta-analysis* [Technology report]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.

28. Lim E, Ali Z, Ali A, Routledge T, Edmonds L, Altman DG, et al. Indirect comparison meta-analysis of aspirin therapy after coronary surgery. *BMJ* 2003;327(7427):1309.

29. CAPRIE Steering Committee. A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). *Lancet* 1996;348(9038):1329-39.

30. Collaborative overview of randomised trials of antiplatelet therapy--I: Prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients. Antiplatelet Trialists' Collaboration. *BMJ* 1994;308(6921):81-106.

31. Ballesteros J. Orphan comparisons and indirect meta-analysis: a case study on antidepressant efficacy in dysthymia comparing tricyclic antidepressants, selective serotonin reuptake inhibitors, and monoamine oxidase inhibitors by using general linear models. *J Clin Psychopharmacol* 2005;25(2):127-31.

32. Sauriol L, Laporta M, Edwardes MD, Deslandes M, Ricard N, Suissa S. Meta-analysis comparing newer antipsychotic drugs for the treatment of schizophrenia: evaluating the indirect approach. *Clin Ther* 2001;23(6):942-56.

33. Tran PV, Hamilton SH, Kuntz AJ, Potvin JH, Andersen SW, Beasley C, et al. Double-blind comparison of olanzapine versus risperidone in the treatment of schizophrenia and other psychotic disorders. *J Clin Psychopharmacol* 1997;17(5):407-18.

34. von der Maase H, Hansen SW, Roberts JT, Dogliotti L, Oliver T, Moore MJ, et al. Gemcitabine and cisplatin versus methotrexate, vinblastine, doxorubicin, and cisplatin in advanced or metastatic bladder cancer: results of a large, randomized, multinational, multicenter, phase III study. *J Clin Oncol* 2000;18(17):3068-77.

35. Logothetis CJ, Dexeus FH, Finn L, Sella A, Amato RJ, Ayala AG, et al. A prospective randomized trial comparing MVAC and CISCA chemotherapy for patients with metastatic urothelial tumors. *J Clin Oncol* 1990;8(6):1050-5.

36. Loehrer PJ, Sr., Einhorn LH, Elson PJ, Crawford ED, Kuebler P, Tannock I, et al. A randomized comparison of cisplatin alone or in combination with methotrexate, vinblastine, and doxorubicin in patients with metastatic urothelial carcinoma: a cooperative group study. *J Clin Oncol* 1992;10(7):1066-73.

37. Cohen MH, Rothmann M. Gemcitabine and cisplatin for advanced, metastatic bladder cancer. *J Clin Oncol* 2001;19(4):1229-31.

38. Tang AS, Wells GA, Arnold M, Connolly S, Hohnloser S, Nichol G, et al. Resynchronization/defibrillation for ambulatory heart failure trial: rationale and trial design. *Curr Opin Cardiol* 2009;24(1):1-8.

39. Song F, Harvey I, Lilford R. Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol* 2008;61(5):455-63.

40. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005;9(26):1-134.

41. Yazdanpanah Y, Sissoko D, Egger M, Mouton Y, Zwahlen M, Chene G. Clinical efficacy of antiretroviral combination therapy based on protease inhibitors or non-nucleoside analogue reverse transcriptase inhibitors: indirect comparison of controlled trials. *BMJ* 2004;328(7434):249.

42. Lundgren JD, Phillips AN. Indirect comparisons: a novel approach to assessing the effect of anti-HIV drugs. *BMJ* 2004;328(7434):253.

APPENDIX A: GENERAL THEORY FOR EFFECT SIZE ESTIMATOR AND TEST STATISTIC OF ASSOCIATION

Two fundamental propositions underlie the estimation and hypothesis testing procedures for the indirect measures.

Consider k treatments $A_1, A_2, \cdots, A_k$. If for consecutive pairs of treatment the direct estimator of the measure of association (Y) for treatment A_i and A_{i+1} is $\hat{Y}_{A_iA_{i+1}} (i = 1,2,...,k-1)$ then the effect size estimator and test statistic for evaluating the indirect association between treatments A_1 and A_k can be expressed in terms of the direct estimators as summarized in Figure A.1.

Figure A.1: Indirect effect size estimator and test statistic for association

Result	Treatment Comparison	Estimator of Association	Number of Studies	Weights*
1	A_1 and A_2	$\hat{Y}_{A_1A_2}$	$h_{A_1A_2}$	$W_{A_1A_2,j}$
2	A_2 and A_3	$\hat{Y}_{A_2A_3}$	$h_{A_2A_3}$	$W_{A_2A_3,j}$
⋮	⋮	⋮	⋮	⋮
k-1	A_{k-1} and A_k	$\hat{Y}_{A_{k-1}A_k}$	$h_{A_{k-1}A_k}$	$W_{A_{k-1}A_k,j}$

* $W_{A_iA_{i+1},j} = 1/\text{var}\,iance(Y_{A_iA_{i+1},j})$ is the weight assigned to the measure of association $Y_{A_iA_{i+1},j}$ for the jth study evaluating treatments A_i and A_{i+1} $(i = 1,2,...,k-1; j = 1,2,...,h_{A_iA_{i+1}})$

- Indirect estimator of Y for treatments A_1 and A_k: $\hat{Y}_{A_1A_kIndirect} = \sum_{i=1}^{k-1} \hat{Y}_{A_iA_{i+1}}$

  provided the direct estimator has the general functional form $\hat{Y}_{A_iA_{i+1}} = f(A_i) - f(A_{i+1})$, $(i = 1,2,...,k-1)$

- Indirect 100(1-α)% confidence interval estimator of Y for treatments A_1 and A_k:

  $$\sum_{i=1}^{k-1} \hat{Y}_{A_iA_{i+1}} \pm Z_{\alpha/2} \sqrt{\sum_{i=1}^{k-1} Var(\hat{Y}_{A_iA_{i+1}})}$$

  where $Z_{\alpha/2}$ is the 100(1-α) percentile of the standard normal distribution

- Test statistic for evaluating the indirect association between treatments A_1 and A_k:

$$\chi^2_{A_1A_k Indirect\ association} = \left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)(\hat{Y}_{A_iA_{i+1}} - \hat{Y}_{A_{i'}A_{i'+1}})^2\right] \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)$$

The formal statement of the two theorems that underlie the estimation and hypothesis testing procedures are:

A.1 Effect Size Estimator

Proposition:
Consider k treatments $A_1, A_2, \cdots, A_k$. If for consecutive pairs of treatments the direct estimator of the measure of association (Y) for treatment A_i and A_{i+1} is $\hat{Y}_{A_iA_{i+1}}$ and this has the functional form $\hat{Y}_{A_iA_{i+1}} = f(A_i) - f(A_{i+1}),\ (i = 1,2,...,k-1)$ then the indirect estimator of Y for treatments A_1 and A_k is

$$\hat{Y}_{A_1A_k Indirect} = \sum_{i=1}^{k-1} \hat{Y}_{A_iA_{i+1}} \qquad \text{(A.1.1)}$$

The indirect 100(1-α)% confidence interval estimator of Y is given by

$$\sum_{i=1}^{k-1} \hat{Y}_{A_iA_{i+1}} \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1} Var(\hat{Y}_{A_iA_{i+1}})} \qquad \text{(A.1.2)}$$

where $Z_{\alpha/2}$ is the upper 100(1-α/2) percentile of the standard normal distribution.

Proof:

The proof of the relationship will be done by mathematical induction. First, for the case of an indirect estimator of n=1 steps (i.e. k=3). Consider k=3 treatments, the direct estimator of Y for treatments A_1 and A_3 is

$$\hat{Y}_{A_1A_3} = f(A_1) - f(A_3) \qquad \text{(A.1.3)}$$

Adding and subtracting $f(A_2)$ in (A.1.3) yields

$$\big(f(A_1) - f(A_2)\big) + \big(f(A_2) - f(A_3)\big) = \hat{Y}_{A_1A_2} + \hat{Y}_{A_2A_3}$$

which is the indirect estimator of Y. That is

$$\hat{Y}_{A_1A_3 Indirect} = \hat{Y}_{A_1A_2} + \hat{Y}_{A_2A_3}$$

implying that the relationship is true for a 1 step indirect estimator.

Assume that the relationship is true for an indirect estimator of n-3 (n>5) steps (i.e. k-1 treatments). That is, for k-1 treatments it is assumed that the indirect estimator of the measure of association between treatments A_1 and A_{k-1} is given by

$$\hat{Y}_{A_1A_{k-1}Indirect} = \sum_{i=1}^{k-2} \hat{Y}_{A_iA_{i+1}} \quad \text{where} \quad \hat{Y}_{A_iA_{i+1}} = f(A_i) - f(A_{i+1}) \qquad (i = 1,2,....,k-2).$$

The direct estimator of the measure of association between treatments A_1 and A_k is

$$\hat{Y}_{A_1A_k} = f(A_1) - f(A_k) \qquad \text{(A.1.4)}$$

Adding and subtracting $f(A_i), i = 2,...,k-1$ in (A.1.4) and simplifying yields

$$\sum_{i=1}^{k-1}(f(A_i) - f(A_{i+1})) = \sum_{i=1}^{k-2} \hat{Y}_{A_iA_{i+1}} + \hat{Y}_{A_{k-1}A_k} = \hat{Y}_{A_1A_{k-1}Indirect} + \hat{Y}_{A_{k-1}A_k}$$

which is the indirect estimator of the measure of association for k treatments. That is

$$\hat{Y}_{A_1A_kIndirect} = \sum_{i=1}^{k-1} \hat{Y}_{A_iA_{i+1}}$$

implying that the relationship is true for a n-3 step indirect estimator (i.e. k treatments). By the principle of mathematical induction this relationship holds for all values of k treatments (k>2) or equivalently n steps (n>3, n=k-2).

The indirect 100(1-α)% confidence interval estimator of Y is given by

$$\hat{Y}_{A_1A_kIndirect} \pm Z_{\alpha/2}\sqrt{Var(\hat{Y}_{A_1A_kIndirect})} \qquad \text{(A.1.5)}$$

Since the $\hat{Y}_{A_iA_{i+1}} (i = 1,2,...,k-1)$ are estimated from different studies, they are statistically independent and hence

$$Var(\hat{Y}_{A_1A_kIndirect}) = \sum_{i=1}^{k-1} Var(\hat{Y}_{A_iA_{i+1}}) \qquad \text{(A.1.6)}$$

Substituting (A.1.6) into (A.1.5) yields the indirect 100(1-α)% confidence interval estimator of Y as given in (A.1.2). This completes the proof.

Remark:

Given the 100(1-α)% CI estimator $(lcl_{A_iA_{i+1}}, ucl_{A_iA_{i+1}})$ for $\hat{Y}_{A_iA_{i+1}}$ $(i = 1,2,...,k-1)$, the standard error of the indirect estimator $\hat{Y}_{A_1A_kIndirect}$ can be expressed as

$$\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \Big/ 2Z_{\alpha/2}$$

and the 100(1-α)% confidence interval estimator for Y is

$$\sum_{i=1}^{k-1}\hat{Y}_{A_iA_{i+1}} \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(ucl_{A_iA_{i+1}} - lcl_{A_iA_{i+1}})^2}\ . \qquad \text{(A.1.7)}$$

A.2 Test Statistic of Association

In combining measures of association Y_j $(j = 1,2,\cdots,h)$ across several (h) studies, the overall measure of association $\hat{Y}$ is a weighted average of the measures with weights W_j $(j = 1,2,\cdots,h)$ being the inverse of the variance of the measures for each study. Further, as indicated in Figure A.2.1, under the null hypothesis of no association in any of the studies, various sums of squares of the measures have approximate chi-square distributions that can be used to assess the degree of association and the homogeneity of the measures across the studies.

Figure A.2.1: Combining measures of association

Study	Measure of Association*	Standard Error	Weight $(1/s_i^2)$
1	Y_1	s_1	W_1
2	Y_2	s_2	W_2
⋮	⋮	⋮	⋮
h	Y_h	s_h	W_h
* no association is given by $Y_i = 0$			

$$\chi^2_{total} = \chi^2_{association} + \chi^2_{heterogeneity}$$

where

$$\chi^2_{total} = \sum_{j=1}^{h} W_j Y_j^2$$

$$\chi^2_{association} = \left(\sum_{j=1}^{h} W_j Y_j \right)^2 \Bigg/ \sum_{j=1}^{h} W_j$$

$$\chi^2_{heterogeneity} = \sum_{j=1}^{h} W_j (Y_j - \hat{Y})^2 \qquad \textit{where} \quad \hat{Y} = \sum_{j=1}^{h} W_j Y_j \Bigg/ \sum_{j=1}^{h} W_j$$

and under the null hypothesis of no association

χ^2_{total} *is distributed as* χ^2_h

$\chi^2_{association}$ *is distributed as* χ^2_1

$\chi^2_{heterogeneity}$ *is distributed as* χ^2_{h-1}

For the direct comparison of A_i and A_{i+1} involving $h_{A_iA_{i+1}}$ studies, we have

$$\chi^2_{A_iA_{i+1}total} = \chi^2_{A_iA_{i+1}association} + \chi^2_{A_iA_{i+1}heterogeneity}$$

where $\chi^2_{A_iA_{i+1}total} = \sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} Y^2_{A_iA_{i+1},j}$

$$\chi^2_{A_iA_{i+1}association} = \left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} Y_{A_iA_{i+1},j} \right)^2 \Bigg/ \sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}$$

$$\chi^2_{A_iA_{i+1}heterogeneity} = \sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} (Y_{A_iA_{i+1},j} - \hat{Y}_{A_iA_{i+1}})^2 \qquad \text{(A.2.1)}$$

$$\text{where} \quad \hat{Y}_{A_iA_{i+1}} = \sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} Y_{A_iA_{i+1},j} \Bigg/ \sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} \qquad \text{(A.2.2)}$$

and under the null hypothesis of no association

$\chi^2_{A_iA_{i+1}total}$ is distributed as $\chi^2_{h_{A_iA_{i+1}}}$

$\chi^2_{A_iA_{i+1}association}$ is distributed as χ^2_1

$\chi^2_{A_iA_{i+1}heterogeneity}$ is distributed as $\chi^2_{h_{A_iA_{i+1}}-1}$

Combining all $\sum_{i=1}^{k-1} h_{A_iA_{i+1}}$ studies, we have

$$\chi^2_{A_1A_2\cdots A_k total} = \chi^2_{A_1A_2\cdots A_k association} + \chi^2_{A_1A_2\cdots A_k heterogeneity}$$

where $\chi^2_{A_1A_2\cdots A_k total} = \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} Y^2_{A_iA_{i+1},j} = \sum_{i=1}^{k-1} \chi^2_{A_iA_{i+1} total}$

$$\chi^2_{A_1A_2\cdots A_k association} = \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} Y_{A_iA_{i+1},j}\right)^2 \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_1A_2}} W_{A_iA_{i+1},j}\right)$$

$$\chi^2_{A_1A_2\cdots A_k heterogeneity} = \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}(Y_{A_iA_{i+1},j} - \hat{Y}_{A_1A_2\cdots A_k})^2 \qquad \text{(A.2.3)}$$

$$\text{where} \quad \hat{Y}_{A_1A_2\cdots A_k} = \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} Y_{A_iA_{i+1},j}\right) \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \qquad \text{(A.2.4)}$$

and under the null hypothesis of no association

$\chi^2_{A_1A_2\cdots A_k total}$ is distributed as $\chi^2_{\sum_{i=1}^{k-1} h_{A_iA_{i+1}}}$

$\chi^2_{A_1A_2\cdots A_k association}$ is distributed as χ^2_1

$\chi^2_{A_1A_2\cdots A_k heterogeneity}$ is distributed as $\chi^2_{\sum_{i=1}^{k-1} h_{A_iA_{i+1}} - 1}$

In considering a test statistic $\chi^2_{A_1A_k Indirect\ association}$ for evaluating the association between treatments A_1 and A_k when only an indirect comparison of the treatments A_1 and A_k is available, based on direct estimators for consecutive pairs of treatment A_i and A_{i+1} (i=1,2,…,k-1), the following is proposed

$$\chi^2_{A_1A_k Indirect\ association} = \chi^2_{A_1A_2\cdots A_k heterogeneity} - \sum_{i=1}^{k-1} \chi^2_{A_iA_{i+1} heterogeneity} \qquad \text{(A.2.5)}$$

which is approximately distributed as χ^2_1. A rationale for this formulation is provided after the proof of the following theorem.

Proposition:
Consider k treatments $A_1, A_2, \cdots, A_k$. If for consecutive pairs of treatments the direct estimator of the measure of association (Y) for treatment A_i and A_{i+1} is $\hat{Y}_{A_iA_{i+1}} (i = 1,2,...,k-1)$ then the test statistic for evaluating the indirect association between treatments A_1 and A_k can be expressed as

$$\chi^2_{A_1A_k Indirect\ association} = \left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)(\hat{Y}_{A_iA_{i+1}} - \hat{Y}_{A_{i'}A_{i'+1}})^2\right] \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)$$

(A.2.6)

where $W_{A_iA_{i+1},j} = 1/\text{var}\,iance(Y_{A_iA_{i+1},j})$ is the weight assigned to the measure of association $Y_{A_iA_{i+1},j}$ for the jth study evaluating treatments A_i and A_{i+1} $(i = 1,2,...,k-1;\ j = 1,2,...,h_{A_iA_{i+1}})$ and $h_{A_iA_{i+1}}$ are the number of studies involved in the direct comparison of A_i and A_{i+1}.

Proof:

Substituting (A.2.1) and (A.2.3) into (A.2.5) yields

$$\chi^2_{A_1A_2\cdots A_k heterogeneity} - \sum_{i=1}^{k-1}\chi^2_{A_iA_{i+1} heterogeneity}$$

$$= \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}(Y_{A_iA_{i+1},j} - \hat{Y}_{A_1A_2\cdots A_k})^2 - \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}(Y_{A_iA_{i+1},j} - \hat{Y}_{A_iA_{i+1}})^2$$

$$where\ \ \hat{Y}_{A_1A_2\cdots A_k} = \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}Y_{A_iA_{i+1},j}\right) \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)$$

$$and\ \ \ \ \hat{Y}_{A_iA_{i+1}} = \sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}Y_{A_iA_{i+1},j} \Bigg/ \sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}$$

$$= \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}(Y^2_{A_iA_{i+1},j} - 2Y_{A_iA_{i+1},j}\hat{Y}_{A_1A_2\cdots A_k} + \hat{Y}^2_{A_1A_2\cdots A_k}) - \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}(Y^2_{A_iA_{i+1},j} - 2Y_{A_iA_{i+1},j}\hat{Y}_{A_iA_{i+1}} + \hat{Y}^2_{A_iA_{i+1}})$$

$$= \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}(Y^2_{A_iA_{i+1},j} - 2Y_{A_iA_{i+1},j}\hat{Y}_{A_1A_2\cdots A_k} + \hat{Y}^2_{A_1A_2\cdots A_k} - Y^2_{A_iA_{i+1},j} + 2Y_{A_iA_{i+1},j}\hat{Y}_{A_iA_{i+1}} - \hat{Y}^2_{A_iA_{i+1}})$$

$$= -2\hat{Y}_{A_1A_2\cdots A_k}\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}Y_{A_iA_{i+1},j} + \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_1A_2\cdots A_k} + 2\sum_{i=1}^{k-1}\hat{Y}_{A_iA_{i+1}}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}Y_{A_iA_{i+1},j}$$
$$- \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_iA_{i+1}}$$

$$= -2\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_1A_2\cdots A_k} + \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_1A_2\cdots A_k} + 2\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_iA_{i+1}} - \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_iA_{i+1}}$$

since from (A.2.2) and (A.2.4)

$$\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} Y_{A_iA_{i+1},j} = \hat{Y}_{A_1A_2\cdots A_k}\left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \quad \textit{and} \quad \sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} Y_{A_iA_{i+1},j} = \hat{Y}_{A_iA_{i+1}}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)$$

respectively

$$= \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_iA_{i+1}} - \sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_1A_2\cdots A_k}$$

$$= \left[\left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_iA_{i+1}}\right) - \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}_{A_iA_{i+1}}\right)^2\right] \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \tag{A.2.7}$$

since substituting (A.2.2) into (A.2.4) and simplifying yields

$$\hat{Y}_{A_1A_2\cdots A_k} = \left(\sum_{i=1}^{k-1}\hat{Y}_{A_iA_{i+1}}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\right) \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \quad \textit{and}$$

$$\hat{Y}_{A_1A_2\cdots A_k}\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j} = \sum_{i=1}^{k-1}\hat{Y}_{A_iA_{i+1}}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)$$

The following terms in (A.2.7) can be expressed as:

$$\left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}_{A_iA_{i+1}}\right)^2 = \sum_{i=1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)^2 \hat{Y}^2_{A_iA_{i+1}} + 2\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)\hat{Y}_{A_iA_{i+1}}\hat{Y}_{A_iA_{i'+1}} \tag{A.2.8}$$

and

$$\left(\sum_{i'=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_{i'}A_{i'+1},j}\right)\left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\hat{Y}^2_{A_iA_{i+1}}\right)$$

$$= \sum_{i'=1}^{k-1}\sum_{i=1}^{k-1}\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\hat{Y}^2_{A_iA_{i+1}} + \sum_{i=1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)^2 \hat{Y}^2_{A_iA_{i+1}} \tag{A.2.9}$$

Substituting (A.2.8) and (A.2.9) in (A.2.7) yields

$$\left[\sum_{i'=1}^{k-1}\sum_{i=1}^{k-1}\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\hat{Y}^2_{A_iA_{i+1}} - 2\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)\hat{Y}_{A_iA_{i+1}}\hat{Y}_{A_{i'}A_{i'+1}}\right]$$
$$\Big/\left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)$$

$$=\left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)(\hat{Y}_{A_iA_{i+1}} - \hat{Y}_{A_{i'}A_{i'+1}})^2\right]\Big/\left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)$$

This completes the proof.

Rationale for equation (A.2.5)

For analysis of variance (ANOVA) with one factor having k levels, the total sum of squares (SST) is

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \overline{X}_{..})^2 = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \overline{X}_{.j} + \overline{X}_{.j} - \overline{X}_{..})^2 \qquad \text{(A.2.10)}$$

where

n_j is the number of observations in j th level $j = 1, \cdots, k$

$X_{ij} = i$ th observation for j th level

$\overline{X}_{..} = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} X_{ij} = \frac{1}{n}\sum_{j=1}^{k} n_j \overline{X}_{.j}$ and $n = \sum_{j=1}^{k} n_j$, $\overline{X}_{.j} = \sum_{i=1}^{n_j}\frac{X_{ij}}{n_j}$

Expanding the square on right hand side (A.2.10), yields

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \overline{X}_{.j} + \overline{X}_{.j} - \overline{X}_{..})^2$$
$$= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \overline{X}_{.j})^2 + 2\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \overline{X}_{.j})(\overline{X}_{.j} - \overline{X}_{..}) + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(\overline{X}_{.j} - \overline{X}_{..})^2$$

(A.2.11)

The cross-product term in (A.2.11) is zero, that is

$$2\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X}_{.j})(\overline{X}_{.j}-\overline{X}_{..}) = 2\sum_{j=1}^{k}(\overline{X}_{.j}-\overline{X}_{..})\sum_{i=1}^{n_j}(X_{ij}-\overline{X}_{.j})$$

$$= 2\sum_{j=1}^{k}(\overline{X}_{.j}-\overline{X}_{..})\left[\sum_{i=1}^{n_j}X_{ij}-n_j\overline{X}_{.j}\right]$$

$$= 2\sum_{j=1}^{k}(\overline{X}_{.j}-\overline{X}_{..})(n_j\overline{X}_{.j}-n_j\overline{X}_{.j})$$

$$= 0$$

Thus (A.2.11) reduces to

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X}_{..})^2 = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X}_{.j})^2 + \sum_{j=1}^{k}n_j(\overline{X}_{.j}-\overline{X}_{..})^2$$

and

$$\sum_{j=1}^{k}n_j(\overline{X}_{.j}-\overline{X}_{..})^2 = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X}_{..})^2 - \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X}_{.j})^2 \qquad \text{(A.2.12)}$$

For k=2, (A.2.12) reduces to

$$\sum_{j=1}^{2}n_j(\overline{X}_{.j}-\overline{X}_{..})^2 = \sum_{j=1}^{2}\sum_{i=1}^{n_j}(X_{ij}-\overline{X}_{..})^2 - \sum_{i=1}^{n_1}(X_{i1}-\overline{X}_{.1})^2 - \sum_{i=1}^{n_2}(X_{i2}-\overline{X}_{.2})^2 \qquad \text{(A.2.13)}$$

which can be expressed as SSR = SST - SSW_1 - SSW_2 where

SST represents the overall heterogeneity, that is the sum of squares of differences of all observations from overall mean; this corresponds to $\chi^2_{heterogeneity}$.

SSW_1 represents the heterogeneity due to the A_1A_2 comparison, that is sum of squares of differences of observations in A_1A_2 comparison group from their group mean; this corresponds to $\chi^2_{A_1A_2heterogeneity}$.

SSW_2 represents the heterogeneity due to A_2A_3 comparison, that is sum of squares of differences of observations in A_2A_3 comparison group from their group mean; this corresponds to $\chi^2_{A_2A_3heterogeneity}$.

SSR is the residual sum of squares and represents the statistic $\chi^2_{A_1A_3heterogeneity}$.

Similarly for k treatments $A_1, A_2, \cdots, A_k$, for the indirect comparison between A_1 and A_k using k-1 pairwise direct comparisons between A_i and A_{i+1}, $i = 1, \cdots, k-1$, equation (A.2.12) can be expressed as an extension of (A.2.13). That is

$$\sum_{j=1}^{k} n_j (\overline{X}_{.j} - \overline{X}_{..})^2 = \sum_{j=1}^{k-1} \sum_{i=1}^{n_j} (X_{ij} - \overline{X}_{..})^2 - \sum_{j=1}^{k-1} \sum_{i=1}^{n_j} (X_{ij} - \overline{X}_{.j})^2$$

which can be expressed as $SSR = SST - \sum_{j=1}^{k-1} SSW_j$

Thus

$$\chi^2_{A_1 A_k Indirect\ association} = \chi^2_{A_1 A_2 \cdots A_k heterogeneity} - \sum_{i=1}^{k-1} \chi^2_{A_i A_{i+1} heterogeneity}$$

This completes the rationale.

A.3 Model for Indirect Comparisons of Odds Ratios (OR)

Consider k treatments $A_1, A_2, \cdots, A_k$. For consecutive pairs of treatments, the direct estimator of the logarithm odds ratio (ln(OR)) for treatment A_i and A_{i+1} is $\ln(OR_{A_i A_{i+1}}) = \ln(P_i/(1-P_i)) - \ln(P_{i+1}/(1-P_{i+1}))$ where $P_i = P(A_i)$ denotes the outcome rate for patients on treatment A_i (i=1,2,…,k-1). This has the functional form $f(A_i) - f(A_{i+1})$ where $f(A_i) = \ln(P_i/(1-P_i))$. Applying the Effect Size Estimator Proposition to the special case of estimating the logarithm odds ratio, the indirect estimator of the ln(OR) for treatments A_1 and A_k is

$$\ln(OR_{A_1 A_k Indirect}) = \sum_{i=1}^{k-1} \ln(OR_{A_i A_{i+1}}) \qquad \text{(A.3.1)}$$

and the indirect 100(1-α)% confidence interval estimator of ln(OR) is given by

$$\sum_{i=1}^{k-1} \ln(OR_{A_i A_{i+1}}) \pm Z_{\alpha/2} \sqrt{\sum_{i=1}^{k-1} Var(\ln(OR_{A_i A_{i+1}}))} \qquad \text{(A.3.2)}$$

where $Z_{\alpha/2}$ is the upper 100(1-α/2) percentile of the standard normal distribution. On the arithmetic scale the corresponding indirect estimators for the odds ratio are

$$OR_{A_1 A_k Indirect} = \prod_{i=1}^{k-1} OR_{A_i A_{i+1}} \qquad \text{(A.3.3)}$$

and

$$\exp\left(\sum_{i=1}^{k-1} \ln(OR_{A_i A_{i+1}}) \pm Z_{\alpha/2} \sqrt{\sum_{i=1}^{k-1} Var(\ln(OR_{A_i A_{i+1}}))} \right) \qquad \text{(A.3.4)}$$

respectively.

Remark:

Given the 100(1-α)% CI estimator $(lcl_{A_iA_{i+1}}, ucl_{A_iA_{i+1}})$ for $OR_{A_iA_{i+1}}$ $(i = 1,2,...,k-1)$, the standard error of the indirect estimator $\ln(OR_{A_1A_kIndirect})$ can be expressed as

$$\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \Big/ 2Z_{\alpha/2}$$

and the 100(1-α)% confidence interval estimator of ln(OR) and OR are

$$\sum_{i=1}^{k-1}\ln(OR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \qquad \text{(A.3.5)}$$

and

$$\exp\left(\sum_{i=1}^{k-1}\ln(OR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2}\right) \qquad \text{(A.3.6)}$$

respectively.

The test statistic for evaluating the indirect association OR between treatments A_1 and A_k can be expressed as:

$$\chi^2_{A_1A_kIndirect\ association} = \left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)(\ln(OR_{A_iA_{i+1}}) - \ln(OR_{A_{i'}A_{i'+1}}))^2\right] \Big/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \qquad \text{(A.3.7)}$$

This is a special case of Test Statistic for Association Proposition applied to the estimation of the odds ratio.

A.4 Model for Indirect Comparisons of Relative Risk (RR)

Consider k treatments $A_1, A_2, \cdots, A_k$. For consecutive pairs of treatments, the direct estimator of the logarithm relative risk (ln(RR)) for treatment A_i and A_{i+1} is $\ln(RR_{A_iA_{i+1}}) = \ln(P_i) - \ln(P_{i+1})$ where $P_i = P(A_i)$ denotes the outcome rate for patients on treatment A_i (i=1,2,…,k-1). This has the functional form $f(A_i) - f(A_{i+1})$ where $f(A_i) = \ln(P_i)$. Applying the Effect Size Estimator Proposition to the special case of estimating the logarithm relative risk, the indirect estimator of the ln(RR) for treatments A_1 and A_k is

$$\ln(RR_{A_1A_kIndirect}) = \sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \qquad \text{(A.4.1)}$$

and the indirect 100(1-α)% confidence interval estimator of ln(RR) is given by

$$\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(\ln(RR_{A_iA_{i+1}}))} \qquad \text{(A.4.2)}$$

where $Z_{\alpha/2}$ is the upper 100(1-α/2) percentile of the standard normal distribution. On the arithmetic scale the corresponding indirect estimators for the relative risk are

$$RR_{A_1A_k Indirect} = \prod_{i=1}^{k-1} RR_{A_iA_{i+1}} \qquad \text{(A.4.3)}$$

and

$$\exp\left(\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1}Var(\ln(RR_{A_iA_{i+1}}))}\right) \qquad \text{(A.4.4)}$$

respectively.

Remark:

Given the 100(1-α)% CI estimator $(lcl_{A_iA_{i+1}}, ucl_{A_iA_{i+1}})$ for $RR_{A_iA_{i+1}}$ $(i = 1,2,...,k-1)$, the standard error of the indirect estimator $\ln(RR_{A_1A_k Indirect})$ can be expressed as

$$\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \Big/ 2Z_{\alpha/2}$$

and the 100(1-α)% confidence interval estimator of ln(RR) and RR are

$$\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \qquad \text{(A.4.5)}$$

and

$$\exp\left(\sum_{i=1}^{k-1}\ln(RR_{A_iA_{i+1}}) \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2}\right) \qquad \text{(A.4.6)}$$

respectively.

The test statistic for evaluating the indirect association RR between treatments A_1 and A_k can be expressed as:

$$\chi^2_{A_1A_k Indirect\ association} = \left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)(\ln(RR_{A_iA_{i+1}}) - \ln(RR_{A_{i'}A_{i'+1}}))^2\right] \Big/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \qquad \text{(A.4.7)}$$

This is a special case of Test Statistic for Association Proposition applied to the estimation of the relative risk.

A.5 Model for Indirect Comparisons of Hazard Ratio (HR)

Consider k treatments $A_1, A_2, \cdots, A_k$. For consecutive pairs of treatments, the direct estimator of the logarithm hazard ratio (ln(HR)) for treatment A_i and A_{i+1} is $\ln(HR_{A_iA_{i+1}}) = \ln(h_i(t)) - \ln(h_{i+1}(t))$ where $h_i(t)$ denotes the hazard of the outcome for patients on treatment A_i at time t (i=1,2,…,k-1). This has the functional form $f(A_i) - f(A_{i+1})$ where $f(A_i) = \ln(h_i(t))$. Applying the Effect Size Estimator Proposition to the special case of estimating the logarithm hazard ratio, the indirect estimator of the ln(HR) for treatments A_1 and A_k is

$$\ln(HR_{A_1A_k\,Indirect}) = \sum_{i=1}^{k-1} \ln(HR_{A_iA_{i+1}}) \qquad \text{(A.5.1)}$$

and the indirect 100(1-α)% confidence interval estimator of ln(HR) is given by

$$\sum_{i=1}^{k-1} \ln(HR_{A_iA_{i+1}}) \pm Z_{\alpha/2} \sqrt{\sum_{i=1}^{k-1} Var(\ln(HR_{A_iA_{i+1}}))} \qquad \text{(A.5.2)}$$

where $Z_{\alpha/2}$ is the upper 100(1-α/2) percentile of the standard normal distribution. On the arithmetic scale the corresponding indirect estimators for the hazard ratio are

$$HR_{A_1A_k\,Indirect} = \prod_{i=1}^{k-1} HR_{A_iA_{i+1}} \qquad \text{(A.5.3)}$$

and

$$\exp\left(\sum_{i=1}^{k-1} \ln(HR_{A_iA_{i+1}}) \pm Z_{\alpha/2} \sqrt{\sum_{i=1}^{k-1} Var(\ln(HR_{A_iA_{i+1}}))} \right) \qquad \text{(A.5.4)}$$

respectively.

Remark:

Given the 100(1-α)% CI estimator $(lcl_{A_iA_{i+1}}, ucl_{A_iA_{i+1}})$ for $HR_{A_iA_{i+1}}$ $(i = 1,2,...,k-1)$, the standard error of the indirect estimator $\ln(HR_{A_1A_k\,Indirect})$ can be expressed as

$$\sqrt{\sum_{i=1}^{k-1} (\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \Big/ 2Z_{\alpha/2}$$

and the 100(1-α)% confidence interval estimator of ln(HR) and HR are

$$\sum_{i=1}^{k-1} \ln(HR_{A_iA_{i+1}}) \pm \frac{1}{2} \sqrt{\sum_{i=1}^{k-1} (\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \qquad \text{(A.5.5)}$$

and

$$\exp\left(\sum_{i=1}^{k-1} \ln(HR_{A_iA_{i+1}}) \pm \frac{1}{2} \sqrt{\sum_{i=1}^{k-1} (\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \right) \qquad \text{(A.5.6)}$$

respectively.

The test statistic for evaluating the indirect association HR between treatments A_1 and A_k can be expressed as:

$$\chi^2_{A_1A_k Indirect\ association} = \left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)(\ln(HR_{A_iA_{i+1}}) - \ln(HR_{A_{i'}A_{i'+1}}))^2\right] \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \quad \text{(A.5.7)}$$

This is a special case of Test Statistic for Association Proposition applied to the estimation of the relative risk.

A.6 Model for Indirect Comparisons of Risk Difference (RD)

Consider k treatments $A_1, A_2, \cdots, A_k$. For consecutive pairs of treatments, the direct estimator of the risk difference for treatment A_i and A_{i+1} is $RD_{A_iA_{i+1}} = P_i - P_{i+1}$ where $P_i = P(A_i)$ denotes the the outcome rate for patients on treatment A_i (i=1,2,…,k-1). This has the functional form $f(A_i) - f(A_{i+1})$ where $f(A_i) = P(A_i)$. Applying the Effect Size Estimator Proposition to the special case of estimating the risk difference, the indirect estimator of the RD for treatments A_1 and A_k is

$$RD_{A_1A_k Indirect} = \sum_{i=1}^{k-1} RD_{A_iA_{i+1}} \quad \text{(A.6.1)}$$

and the indirect 100(1-α)% confidence interval estimator of RD is given by

$$\sum_{i=1}^{k-1} RD_{A_iA_{i+1}} \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1} Var(RD_{A_iA_{i+1}})} \quad \text{(A.6.2)}$$

where $Z_{\alpha/2}$ is the upper 100(1-α/2) percentile of the standard normal distribution.

Remark:

Given the 100(1-α)% CI estimator $(lcl_{A_iA_{i+1}}, ucl_{A_iA_{i+1}})$ for $RD_{A_iA_{i+1}}$ $(i = 1,2,...,k-1)$, the standard error of the indirect estimator $RD_{A_1A_k Indirect}$ can be expressed as

$$\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \Big/ 2Z_{\alpha/2}$$

and the 100(1-α)% confidence interval estimator for RD is

$$\sum_{i=1}^{k-1} RD_{A_iA_{i+1}} \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(ucl_{A_iA_{i+1}} - lcl_{A_iA_{i+1}})^2}\ . \quad \text{(A.6.3)}$$

The test statistic for evaluating the indirect association between treatments A_1 and A_k can be expressed as:

$$\chi^2_{A_1A_k Indirect\ association} = \left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)(RD_{A_iA_{i+1}} - RD_{A_{i'}A_{i'+1}})^2\right] \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \quad \text{(A.6.4)}$$

This is a special case of Test Statistic for Association Proposition applied to the estimation of the risk difference.

A.7 Model for Indirect Comparisons of Mean Difference (MD)

Consider k treatments $A_1, A_2, \cdots, A_k$. For consecutive pairs of treatments, the direct estimator of the mean difference for treatment A_i and A_{i+1} is $MD_{A_iA_{i+1}} = M_i - M_{i+1}$ where $M_i = M(A_i)$ denotes the mean of the outcome for patients on treatment A_i (i=1,2,…,k-1). This has the functional form $f(A_i) - f(A_{i+1})$ where $f(A_i) = M(A_i)$. Applying the Effect Size Estimator Proposition to the special case of estimating the mean difference, the indirect estimator of the MD for treatments A_1 and A_k is

$$MD_{A_1A_k Indirect} = \sum_{i=1}^{k-1} MD_{A_iA_{i+1}} \quad \text{(A.7.1)}$$

and the indirect 100(1-α)% confidence interval estimator of MD is given by

$$\sum_{i=1}^{k-1} MD_{A_iA_{i+1}} \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k-1} Var(MD_{A_iA_{i+1}})} \quad \text{(A.7.2)}$$

where $Z_{\alpha/2}$ is the upper 100(1-α/2) percentile of the standard normal distribution.

Remark:

Given the 100(1-α)% CI estimator $(lcl_{A_iA_{i+1}}, ucl_{A_iA_{i+1}})$ for $MD_{A_iA_{i+1}}$ $(i = 1,2,...,k-1)$, the standard error of the indirect estimator $MD_{A_1A_k Indirect}$ can be expressed as

$$\sqrt{\sum_{i=1}^{k-1}(\ln(ucl_{A_iA_{i+1}}) - \ln(lcl_{A_iA_{i+1}}))^2} \Big/ 2Z_{\alpha/2}$$

and the 100(1-α)% confidence interval estimator for MD is

$$\sum_{i=1}^{k-1} MD_{A_iA_{i+1}} \pm \frac{1}{2}\sqrt{\sum_{i=1}^{k-1}(ucl_{A_iA_{i+1}} - lcl_{A_iA_{i+1}})^2}\ . \quad \text{(A.7.3)}$$

The mean difference (MD) for treatment A_i and A_{i+1} is $MD_{A_iA_{i+1}}(i=1,2,...,k-1)$ then the test statistic for evaluating the indirect association between treatments A_1 and A_k can be expressed as:

$$\chi^2_{A_1A_k Indirect\ association} = \left[\sum_{i=1}^{k-2}\sum_{i'=i+1}^{k-1}\left(\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right)\left(\sum_{j=1}^{h_{A_{i'}A_{i'+1}}} W_{A_{i'}A_{i'+1},j}\right)(MD_{A_iA_{i+1}} - MD_{A_{i'}A_{i'+1}})^2\right] \Bigg/ \left(\sum_{i=1}^{k-1}\sum_{j=1}^{h_{A_iA_{i+1}}} W_{A_iA_{i+1},j}\right) \quad \text{(A.7.4)}$$

This is a special case of Test Statistic for Association Proposition applied to the estimation of the mean difference.

APPENDIX B: SIMULATION- DESCRIPTION, TABLE AND FIGURES

B.1 Empirical Evaluation of the Estimators

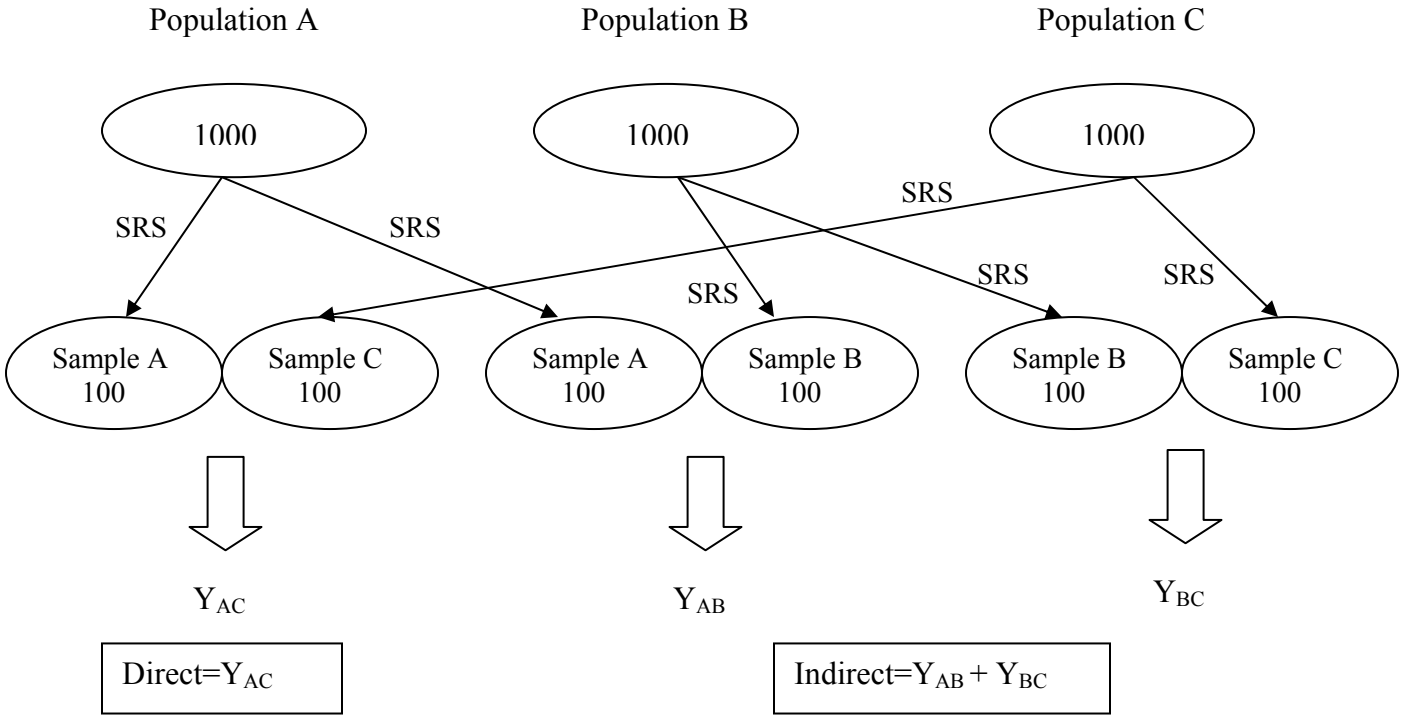
Bias and Mean Square Error

The bias is the expected difference between the estimator and the parameter to be estimated and the mean square error (MSE) is the expected squared deviation between the estimator and this parameter. The MSE summarizes information about the bias and variance of the estimator under study. The purpose of the analyzing the bias, variance and MSE was evaluate the accuracy and precision of the various measures of association, Y (i.e., OR, RR, RD, MD and HR) for both the direct and indirect approaches.

Description of the Monte Carlo Simulation Process

In order to determine the precision and accuracy of the indirect approach, a Monte Carlo simulation analysis was undertaken. The initial step was to generate the population according to the specified outcome risk level. For the case of k=3 treatments, a data set was created for each of the three populations (A, B and C) according to the specified outcome risk level. The second step was to take a simple random sample (SRS) from each of the three populations and calculate the direct and indirect estimates being considered. This step would be repeated 1000 times. The third step was to calculate the bias, variance and MSE for the 1000 direct and indirect estimates and to compare the accuracy and precision of direct and indirect estimators in order to assess whether the indirect method performed adequately. A schematic of the Monte Carlo simulation for the case of the relative risk is provided in Figure B.1.

Figure B.1: Monte Carlo simulation process



Note: SRS-Simple Random Sampling

Step 1: Simulating the populations
A data set with 1000 observations was generated for each of the three populations (A, B and C) according to the specified outcome risk levels (P(E | A, B or C)) determined through the indicated parameters (Y_{AB}, Y_{CB}, and P(E | B)). For time to event data, the uniform random number generator was used to generate a simulated data set of 5000 observations according to the specified outcome risk level in each population based on Cox proportional hazards models assuming exponential distributions for survival time.

Step 2: Simple random sampling
In order to calculate the direct and indirect effect estimates, a simple random sampling without replacement procedure was used to select 100 observations from each of the three populations. For time to event data, 500 observations were selected from each of the three populations. In this procedure, each observation in the population has an equal chance of being selected, once selected it cannot be chosen again. This procedure was repeated 1000 times. Zero event in simulated samples was corrected by adding 0.5.

Step 3: Statistical analysis
The direct and indirect estimates of Y_{AC} were computed for each of the 1000 simulations. The bias, variance and MSE of the direct and indirect estimates based on 1000 simulations were calculated and evaluated in order to assess the precision and accuracy of indirect estimation

approach. The sampling errors in the generation of three populations were adjusted in the calculation of bias and MSE.

It should be noted that the simulation does not take into account differences that could exist related to methods of patient selection and ascertainment of outcome.

For each measure of association, steps 1-3 were repeated by choosing different settings for the indicated parameters: Y_{AB}, Y_{CB}, and $P(E \mid B)$. These settings and simulation results are provided in section B.2 to B.6 for RR, OR, RD, MD and HR, respectively.

B.2 Simulation Results for the Relative Risk

For each of the three populations (A, B and C) in the Monte Carlo simulation for the relative risk, the outcome risk level for each population was selected according to a specific combination of values for the parameters RR_{AB}, RR_{CB}, and the probability of the event in population B $P(E \mid B)$) as follows:

RR_{AB}: 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
RR_{CB}: 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
$P(E \mid B)$): 0.05, 0.1, 0.2, 0.3, 0.4, 0.5

For the various combinations of these parameters, the results of the simulation for the variance, bias and MSE for the direct and indirect estimators of the RR are provided in Table B.2.1.

As an illustration of the frequency distribution of the estimators, the frequency distributions for the parameter settings RR_{AB}=0.6, RR_{CB}=0.8 and $P(E \mid B)$=0.05, 0.3, 0.5 for the direct and indirect estimators are presented graphically in Figure B.2.1 on the logarithmic scale. It is apparent from these figures that both estimators have a mound shape, symmetric distribution (on the logarithmic scale). The indirect estimator has a larger variance and bias.

Table B.2.1: Bias, variance and mean square error (MSE) of direct and indirect relative risk (RR) estimators for different settings of the indicated parameters (k=3 treatments)

Relative Risk Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
RR_{AB}=0.9	**0.05**	2.364	0.392	2.518	7.349	0.750	7.912
RR_{CB}=0.9	**0.1**	0.356	0.124	0.371	0.797	0.232	0.851
	0.2	0.108	0.055	0.111	0.216	0.060	0.219
	0.3	0.057	0.034	0.058	0.112	0.054	0.115
	0.4	0.031	0.009	0.031	0.068	0.026	0.069
	0.5	0.023	0.005	0.023	0.041	0.016	0.041
RR_{AB}=0.9	**0.05**	1.917	0.379	2.060	6.304	0.733	6.841
RR_{CB}=0.8	**0.1**	0.563	0.138	0.582	1.290	0.259	1.357
	0.2	0.153	0.055	0.156	0.323	0.095	0.332
	0.3	0.076	0.038	0.077	0.164	0.067	0.168
	0.4	0.050	0.027	0.050	0.093	0.046	0.095
	0.5	0.033	0.020	0.033	0.059	0.020	0.059
RR_{AB}=0.9	**0.05**	3.601	0.616	3.980	51.588	1.197	53.021
RR_{CB}=0.7	**0.1**	1.222	0.230	1.275	2.158	0.450	2.360
	0.2	0.241	0.083	0.248	0.471	0.136	0.489
	0.3	0.115	0.054	0.118	0.206	0.097	0.215
	0.4	0.081	0.040	0.083	0.132	0.060	0.136
	0.5	0.052	0.038	0.053	0.080	0.028	0.081
RR_{AB}=0.9	**0.05**	4.624	0.648	5.043	11.509	1.122	12.769
RR_{CB}=0.6	**0.1**	1.080	0.262	1.149	2.547	0.423	2.726
	0.2	0.402	0.141	0.422	0.843	0.233	0.897
	0.3	0.208	0.080	0.214	0.286	0.115	0.299
	0.4	0.111	0.037	0.113	0.189	0.069	0.194
	0.5	0.081	0.041	0.083	0.129	0.050	0.132
RR_{AB}=0.9	**0.05**	7.252	0.957	8.169	20.774	1.472	22.941
RR_{CB}=0.5	**0.1**	3.923	0.574	4.253	9.451	0.823	10.127
	0.2	0.802	0.209	0.846	1.201	0.270	1.274
	0.3	0.367	0.126	0.383	0.645	0.151	0.668
	0.4	0.192	0.055	0.195	0.327	0.087	0.334
	0.5	0.143	0.037	0.144	0.214	0.102	0.225
RR_{AB}=0.9	**0.05**	9.614	1.188	11.026	24.417	1.696	27.294
RR_{CB}=0.4	**0.1**	10.159	0.996	11.150	19.845	1.286	21.500
	0.2	2.472	0.325	2.578	2.275	0.495	2.520
	0.3	0.676	0.129	0.693	1.347	0.267	1.419
	0.4	0.479	0.118	0.493	0.617	0.154	0.641
	0.5	0.263	0.068	0.268	0.420	0.080	0.427
RR_{AB}=0.8	**0.05**	1.436	0.305	1.529	9.253	0.521	9.524
RR_{CB}=0.9	**0.1**	0.265	0.110	0.277	0.641	0.225	0.692
	0.2	0.084	0.039	0.085	0.186	0.079	0.192

Relative Risk Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.3	0.048	0.018	0.048	0.085	0.029	0.085
	0.4	0.031	0.013	0.031	0.050	0.030	0.051
	0.5	0.021	0.014	0.021	0.034	0.009	0.034
RR_{AB}=0.8	**0.05**	1.698	0.323	1.802	8.455	0.787	9.075
RR_{CB}=0.8	**0.1**	0.803	0.169	0.832	0.811	0.239	0.868
	0.2	0.118	0.042	0.120	0.222	0.086	0.230
	0.3	0.078	0.052	0.080	0.117	0.043	0.119
	0.4	0.045	0.026	0.045	0.075	0.027	0.076
	0.5	0.026	0.010	0.026	0.047	0.029	0.048
RR_{AB}=0.8	**0.05**	3.136	0.569	3.460	8.460	0.896	9.262
RR_{CB}=0.7	**0.1**	1.230	0.248	1.291	1.500	0.361	1.630
	0.2	0.222	0.077	0.228	0.484	0.125	0.500
	0.3	0.102	0.039	0.104	0.183	0.056	0.187
	0.4	0.064	0.023	0.064	0.116	0.053	0.119
	0.5	0.044	0.020	0.045	0.068	0.021	0.068
RR_{AB}=0.8	**0.05**	4.126	0.647	4.545	9.855	0.955	10.766
RR_{CB}=0.6	**0.1**	1.638	0.298	1.727	4.389	0.547	4.688
	0.2	0.331	0.131	0.348	0.582	0.167	0.610
	0.3	0.179	0.052	0.182	0.272	0.073	0.278
	0.4	0.095	0.037	0.097	0.153	0.057	0.157
	0.5	0.077	0.015	0.078	0.106	0.047	0.108
RR_{AB}=0.8	**0.05**	6.143	0.914	6.978	24.769	1.571	27.237
RR_{CB}=0.5	**0.1**	3.965	0.474	4.190	4.975	0.653	5.401
	0.2	0.974	0.166	1.002	1.044	0.219	1.092
	0.3	0.288	0.106	0.299	0.412	0.115	0.425
	0.4	0.196	0.071	0.201	0.287	0.127	0.303
	0.5	0.125	0.036	0.126	0.195	0.063	0.199
RR_{AB}=0.8	**0.05**	7.209	0.883	7.990	55.287	1.985	59.228
RR_{CB}=0.4	**0.1**	7.029	0.827	7.713	13.207	1.058	14.325
	0.2	1.478	0.324	1.583	1.925	0.417	2.099
	0.3	0.667	0.178	0.698	1.035	0.218	1.083
	0.4	0.417	0.115	0.431	0.571	0.136	0.590
	0.5	0.230	0.068	0.234	0.358	0.105	0.369
RR_{AB}=0.7	**0.05**	1.497	0.316	1.597	3.972	0.533	4.257
RR_{CB}=0.9	**0.1**	0.190	0.078	0.196	0.451	0.168	0.479
	0.2	0.076	0.045	0.078	0.144	0.062	0.148
	0.3	0.042	0.031	0.043	0.084	0.057	0.088
	0.4	0.026	0.012	0.026	0.052	0.036	0.054
	0.5	0.018	0.006	0.019	0.032	0.026	0.033
RR_{AB}=0.7	**0.05**	1.500	0.326	1.606	6.093	0.709	6.595
RR_{CB}=0.8	**0.1**	0.458	0.127	0.474	0.774	0.225	0.825
	0.2	0.103	0.049	0.106	0.197	0.076	0.203

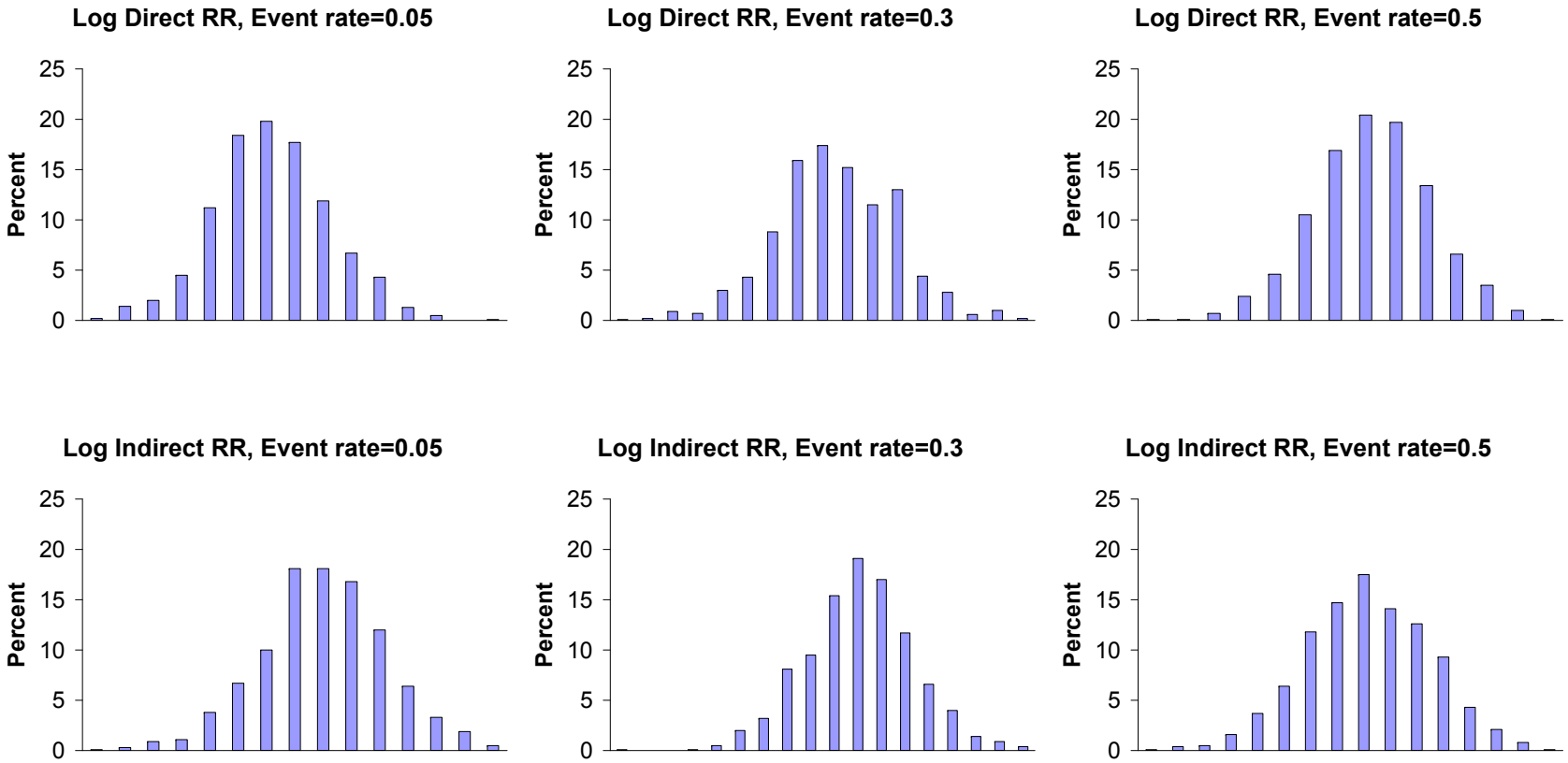
Relative Risk Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.3	0.060	0.034	0.061	0.113	0.058	0.117
	0.4	0.038	0.014	0.038	0.061	0.032	0.062
	0.5	0.025	0.003	0.025	0.041	0.005	0.041
RR_{AB}=0.7	**0.05**	2.833	0.449	3.034	5.689	0.689	6.163
RR_{CB}=0.7	**0.1**	0.785	0.164	0.812	2.680	0.336	2.793
	0.2	0.141	0.051	0.144	0.303	0.121	0.317
	0.3	0.079	0.032	0.080	0.171	0.076	0.177
	0.4	0.051	0.023	0.051	0.092	0.030	0.093
	0.5	0.036	0.020	0.036	0.062	0.030	0.063
RR_{AB}=0.7	**0.05**	3.136	0.623	3.524	8.963	0.799	9.601
RR_{CB}=0.6	**0.1**	1.265	0.287	1.348	2.598	0.429	2.781
	0.2	0.250	0.077	0.256	0.498	0.152	0.521
	0.3	0.135	0.049	0.137	0.215	0.079	0.221
	0.4	0.082	0.036	0.083	0.151	0.075	0.157
	0.5	0.064	0.032	0.066	0.093	0.044	0.095
RR_{AB}=0.7	**0.05**	4.103	0.650	4.526	18.552	1.315	20.282
RR_{CB}=0.5	**0.1**	1.731	0.344	1.849	5.299	0.670	5.747
	0.2	0.459	0.146	0.481	0.679	0.212	0.723
	0.3	0.210	0.075	0.215	0.341	0.121	0.355
	0.4	0.157	0.064	0.161	0.221	0.086	0.228
	0.5	0.098	0.047	0.100	0.149	0.074	0.155
RR_{AB}=0.7	**0.05**	6.884	0.874	7.647	65.695	1.984	69.633
RR_{CB}=0.4	**0.1**	6.421	0.806	7.070	9.906	1.033	10.974
	0.2	2.410	0.275	2.485	4.262	0.396	4.419
	0.3	0.507	0.161	0.533	0.595	0.148	0.617
	0.4	0.324	0.112	0.337	0.483	0.097	0.493
	0.5	0.213	0.072	0.218	0.346	0.104	0.356
RR_{AB}=0.6	**0.05**	1.170	0.244	1.230	2.696	0.534	2.981
RR_{CB}=0.9	**0.1**	0.180	0.086	0.188	0.403	0.175	0.434
	0.2	0.062	0.019	0.063	0.128	0.062	0.131
	0.3	0.037	0.030	0.038	0.055	0.024	0.056
	0.4	0.022	0.017	0.022	0.037	0.015	0.037
	0.5	0.014	0.009	0.014	0.027	0.024	0.027
RR_{AB}=0.6	**0.05**	1.343	0.282	1.423	3.777	0.639	4.186
RR_{CB}=0.8	**0.1**	0.284	0.110	0.296	0.542	0.172	0.572
	0.2	0.075	0.025	0.075	0.159	0.075	0.165
	0.3	0.051	0.023	0.052	0.081	0.042	0.082
	0.4	0.035	0.020	0.035	0.050	0.030	0.051
	0.5	0.022	0.020	0.022	0.034	0.023	0.034
RR_{AB}=0.6	**0.05**	1.924	0.366	2.058	9.423	0.680	9.885
RR_{CB}=0.7	**0.1**	0.423	0.159	0.448	1.063	0.305	1.156
	0.2	0.127	0.059	0.131	0.220	0.095	0.229

Relative Risk Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.3	0.071	0.025	0.072	0.103	0.055	0.107
	0.4	0.043	0.019	0.043	0.067	0.041	0.068
	0.5	0.031	0.019	0.032	0.045	0.022	0.045
RR_{AB}=0.6	**0.05**	2.714	0.510	2.974	9.318	0.858	10.054
RR_{CB}=0.6	**0.1**	0.764	0.204	0.806	1.548	0.334	1.659
	0.2	0.178	0.057	0.181	0.330	0.122	0.345
	0.3	0.107	0.054	0.110	0.219	0.082	0.226
	0.4	0.068	0.031	0.069	0.102	0.051	0.105
	0.5	0.047	0.007	0.047	0.075	0.034	0.076
RR_{AB}=0.6	**0.05**	3.215	0.636	3.619	16.220	1.205	17.671
RR_{CB}=0.5	**0.1**	2.224	0.367	2.359	3.945	0.485	4.181
	0.2	0.438	0.158	0.464	0.710	0.165	0.737
	0.3	0.176	0.062	0.180	0.257	0.084	0.264
	0.4	0.122	0.026	0.123	0.171	0.075	0.176
	0.5	0.079	0.025	0.080	0.109	0.050	0.111
RR_{AB}=0.6	**0.05**	4.275	0.772	4.871	52.812	1.656	55.555
RR_{CB}=0.4	**0.1**	3.407	0.526	3.683	5.668	0.668	6.115
	0.2	0.866	0.228	0.917	1.937	0.265	2.007
	0.3	0.360	0.104	0.371	0.720	0.216	0.766
	0.4	0.359	0.117	0.373	0.312	0.089	0.320
	0.5	0.164	0.060	0.168	0.219	0.066	0.223
RR_{AB}=0.5	**0.05**	0.582	0.202	0.623	3.801	0.424	3.981
RR_{CB}=0.9	**0.1**	0.148	0.069	0.153	0.321	0.134	0.339
	0.2	0.047	0.021	0.048	0.090	0.049	0.092
	0.3	0.026	0.020	0.026	0.042	0.026	0.042
	0.4	0.018	0.009	0.018	0.028	0.026	0.028
	0.5	0.011	0.004	0.011	0.019	0.021	0.020
RR_{AB}=0.5	**0.05**	0.849	0.243	0.909	4.274	0.551	4.578
RR_{CB}=0.8	**0.1**	0.245	0.107	0.256	0.359	0.162	0.385
	0.2	0.070	0.040	0.072	0.128	0.069	0.132
	0.3	0.033	0.017	0.033	0.057	0.031	0.058
	0.4	0.025	0.018	0.026	0.039	0.013	0.039
	0.5	0.020	0.016	0.020	0.026	0.020	0.027
RR_{AB}=0.5	**0.05**	1.155	0.301	1.246	3.746	0.533	4.030
RR_{CB}=0.7	**0.1**	0.383	0.106	0.394	0.922	0.234	0.977
	0.2	0.096	0.053	0.099	0.164	0.079	0.171
	0.3	0.061	0.037	0.062	0.090	0.045	0.092
	0.4	0.039	0.026	0.039	0.048	0.019	0.048
	0.5	0.026	0.006	0.026	0.034	0.020	0.035
RR_{AB}=0.5	**0.05**	1.972	0.391	2.126	7.379	0.736	7.921
RR_{CB}=0.6	**0.1**	0.712	0.187	0.747	1.329	0.261	1.397
	0.2	0.140	0.061	0.144	0.254	0.101	0.264

Relative Risk Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.3	0.074	0.027	0.075	0.128	0.066	0.132
	0.4	0.054	0.025	0.055	0.076	0.043	0.078
	0.5	0.039	0.025	0.040	0.058	0.037	0.059
RR_{AB}=0.5	**0.05**	2.317	0.500	2.567	7.246	0.872	8.007
RR_{CB}=0.5	**0.1**	1.128	0.292	1.213	1.917	0.404	2.080
	0.2	0.277	0.098	0.287	0.471	0.161	0.497
	0.3	0.143	0.054	0.146	0.212	0.083	0.219
	0.4	0.089	0.047	0.091	0.118	0.063	0.122
	0.5	0.070	0.025	0.071	0.089	0.038	0.091
RR_{AB}=0.5	**0.05**	3.467	0.659	3.901	19.141	1.216	20.620
RR_{CB}=0.4	**0.1**	3.334	0.548	3.635	3.668	0.665	4.111
	0.2	0.532	0.164	0.559	1.029	0.239	1.086
	0.3	0.300	0.118	0.314	0.393	0.141	0.412
	0.4	0.162	0.055	0.165	0.238	0.062	0.242
	0.5	0.128	0.041	0.130	0.164	0.075	0.170
RR_{AB}=0.4	**0.05**	0.451	0.138	0.470	1.269	0.308	1.364
RR_{CB}=0.9	**0.1**	0.101	0.063	0.105	0.237	0.098	0.246
	0.2	0.040	0.023	0.041	0.057	0.039	0.058
	0.3	0.020	0.011	0.020	0.031	0.021	0.032
	0.4	0.016	0.009	0.016	0.021	0.007	0.021
	0.5	0.009	0.002	0.009	0.015	0.019	0.015
RR_{AB}=0.4	**0.05**	0.933	0.227	0.984	5.017	0.431	5.203
RR_{CB}=0.8	**0.1**	0.180	0.049	0.182	0.262	0.123	0.277
	0.2	0.045	0.021	0.045	0.078	0.050	0.081
	0.3	0.029	0.015	0.029	0.042	0.037	0.043
	0.4	0.019	0.008	0.019	0.025	0.015	0.025
	0.5	0.014	0.007	0.014	0.020	0.016	0.020
RR_{AB}=0.4	**0.05**	1.071	0.284	1.151	1.879	0.403	2.041
RR_{CB}=0.7	**0.1**	0.235	0.109	0.247	0.469	0.171	0.498
	0.2	0.072	0.049	0.074	0.104	0.043	0.106
	0.3	0.043	0.034	0.044	0.057	0.028	0.057
	0.4	0.027	0.015	0.027	0.039	0.024	0.039
	0.5	0.019	0.008	0.019	0.027	0.012	0.027
RR_{AB}=0.4	**0.05**	1.226	0.340	1.341	9.732	0.776	10.334
RR_{CB}=0.6	**0.1**	0.752	0.216	0.799	1.734	0.275	1.809
	0.2	0.109	0.051	0.112	0.142	0.053	0.145
	0.3	0.059	0.036	0.060	0.087	0.053	0.090
	0.4	0.040	0.024	0.041	0.051	0.022	0.052
RR_{AB}=0.4	**0.05**	0.030	0.009	0.031	0.040	0.024	0.041
RR_{CB}=0.5	**0.1**	1.594	0.410	1.762	5.997	0.802	6.641
	0.2	0.854	0.261	0.922	1.619	0.357	1.747
	0.3	0.219	0.069	0.224	0.283	0.129	0.299

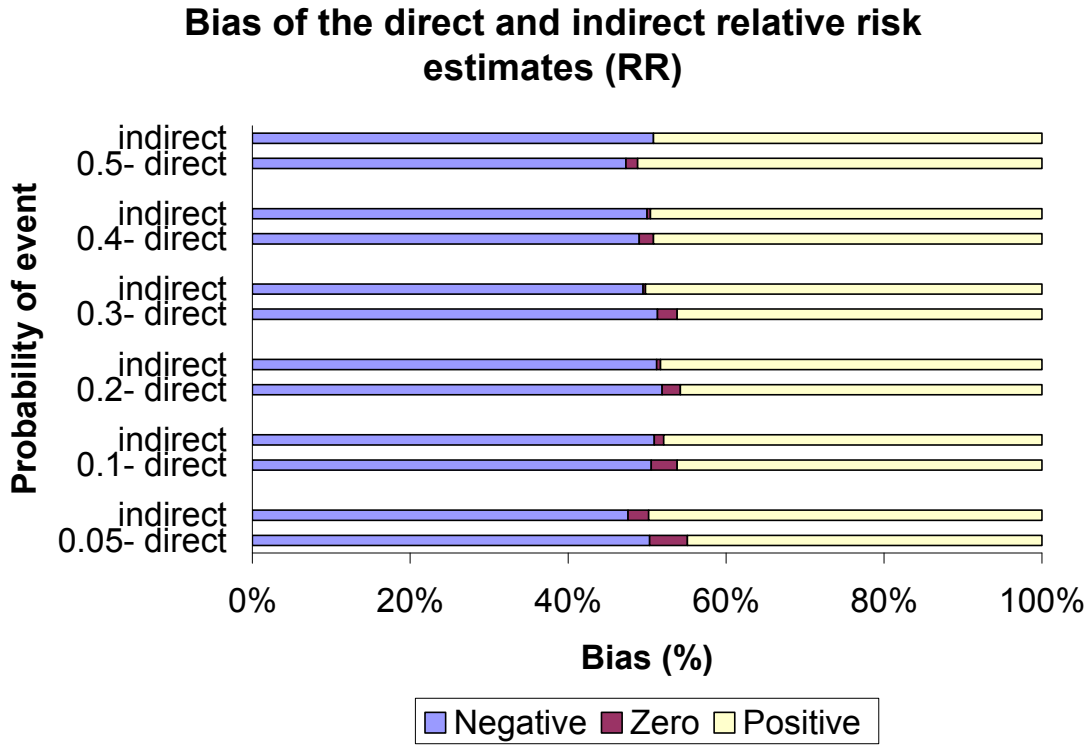
Relative Risk Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.4	0.108	0.044	0.110	0.149	0.064	0.153
	0.5	0.068	0.048	0.070	0.104	0.041	0.105
RR_{AB}=0.4	**0.05**	0.049	0.024	0.049	0.064	0.028	0.065
RR_{CB}=0.4	**0.1**	2.404	0.512	2.667	6.567	0.775	7.166
	0.2	2.313	0.398	2.472	4.269	0.653	4.696
	0.3	0.471	0.162	0.497	0.713	0.178	0.745
	0.4	0.184	0.080	0.190	0.267	0.101	0.277
	0.5	0.126	0.041	0.128	0.151	0.049	0.154
		0.081	0.014	0.081	0.111	0.030	0.112

Figure B.2.1: Frequency distribution of direct and indirect relative risk (RR) estimators for the parameter settings RR_{AB}=0.6, RR_{CB}=0.8, P(E | B)=0.05, 0.3, 0.5 (k=3 treatments)



For the parameter settings RR_{AB}=0.6, RR_{CB}=0.8, the direction of the biases over the 1000 samples are illustrated in Figure B.2.2 for the direct and indirect approaches under the different settings of the event rate P(E | B): 0.05, 0.1, 0.2, 0.3, 0.4, 0.5. For both the direct and indirect estimators, generally, as the event rate increases, the percentage of estimators which are not associated with any bias decreases.

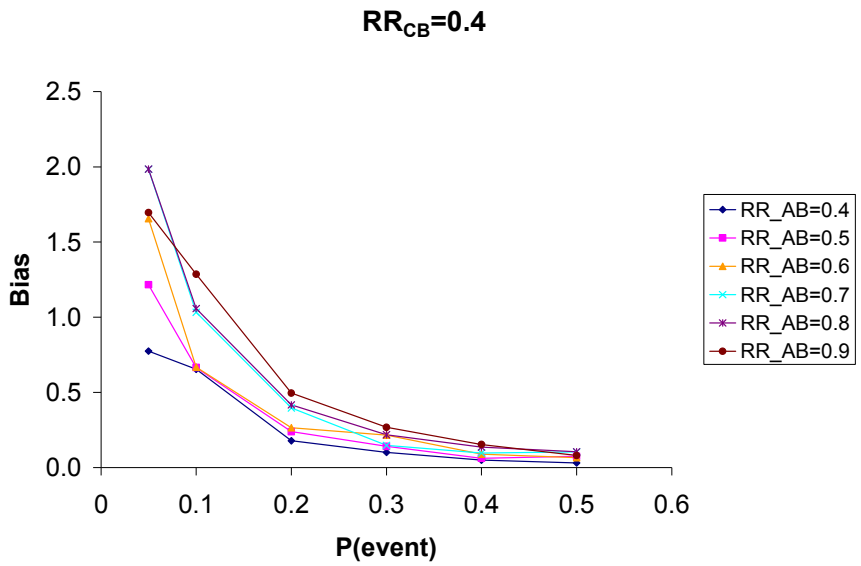
Figure B.2.2: Direction of the bias of the direct and indirect relative risk (RR) estimates for the parameter settings RR_{AB}=0.6, RR_{CB}=0.8, P(E | B) =0.05, 0.1, 0.2, 0.3, 0.4, 0.5 (k=3 treatments)



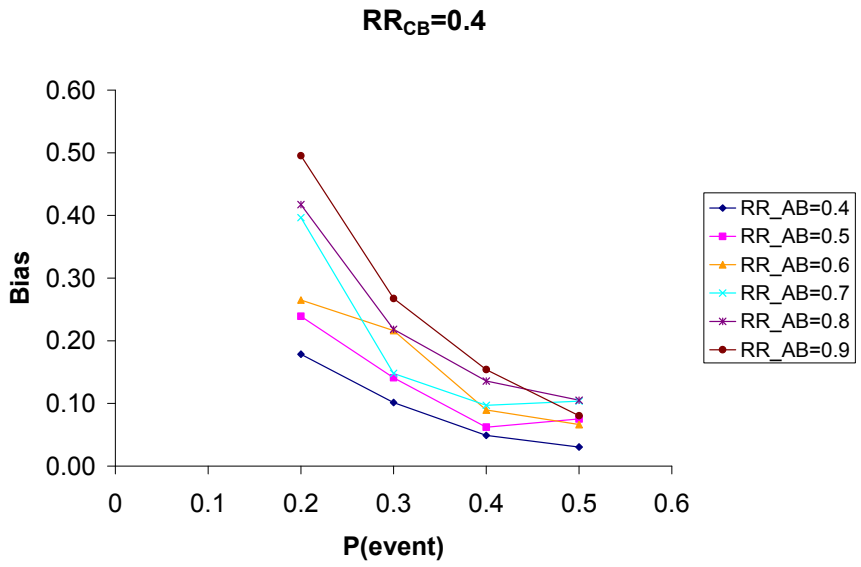
Patterns of bias and MSE for the indirect relative risk (RR) estimates are displayed for different settings of RR_{AB}, RR_{CB}, and the event rates (Figure B.2.3 and Figure B.2.4). The patterns are shown for the event rate 0.05 to 0.5 and the results are symmetric about 0.5. As such, patterns for event rates ranging from 0.5 to 0.95 are not shown. For each of the settings, the figures are displayed for event rates that start at 0.05. The figures are also displayed for event rates that start at 0.2 in order to improve the resolution of the graphs.

Figure B.2.3: Bias for indirect relative risk (RR) estimates for various parameter settings.

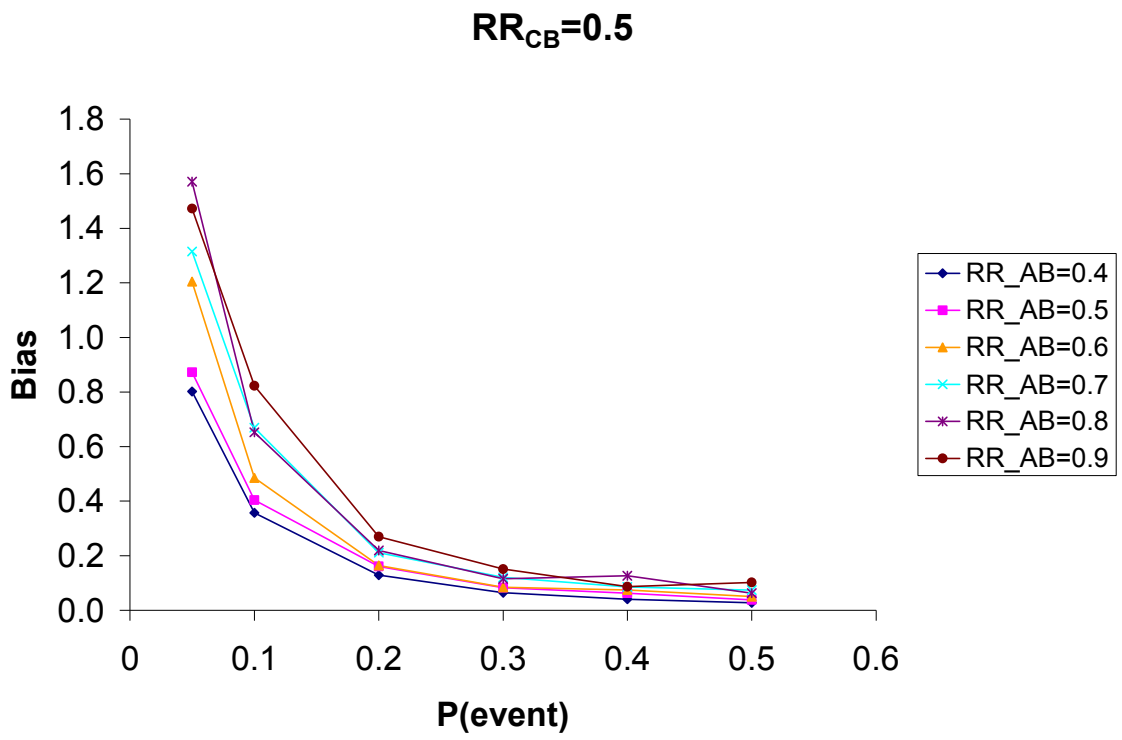
a)



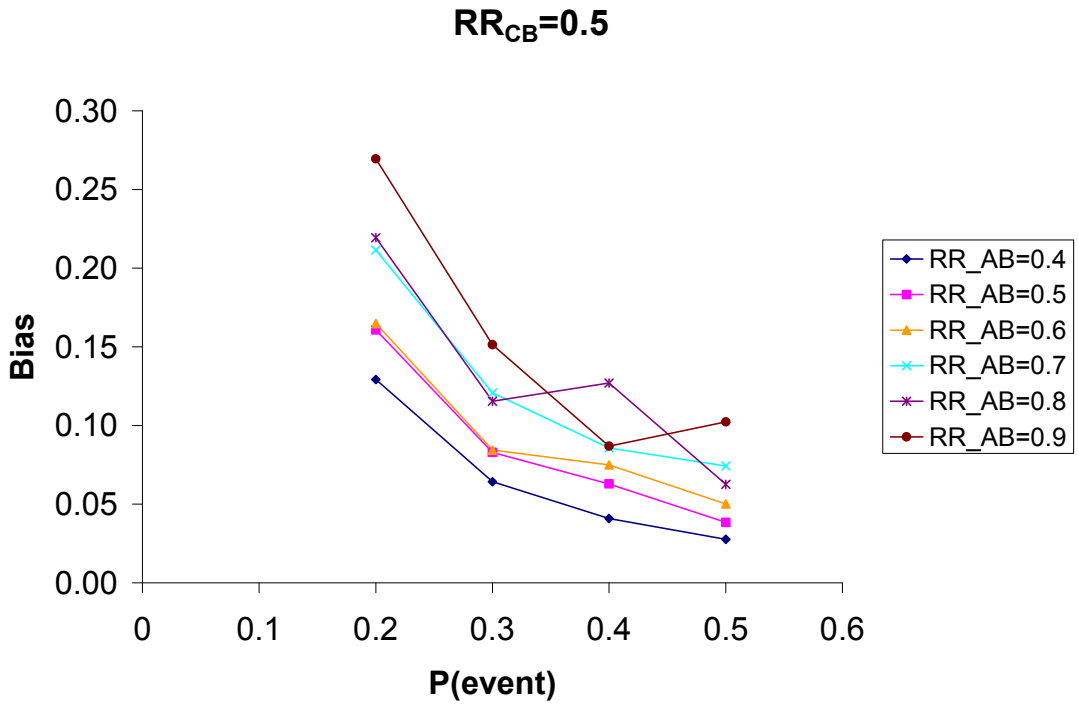
b)



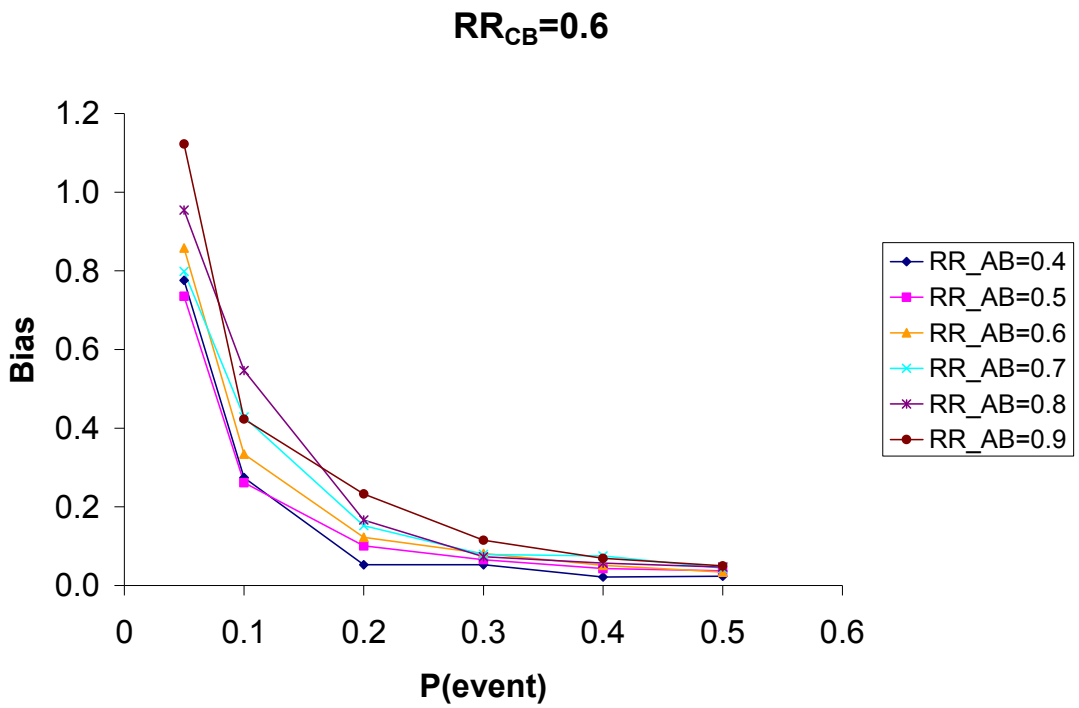
c)


RR_CB=0.5
Bias
P(event)
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

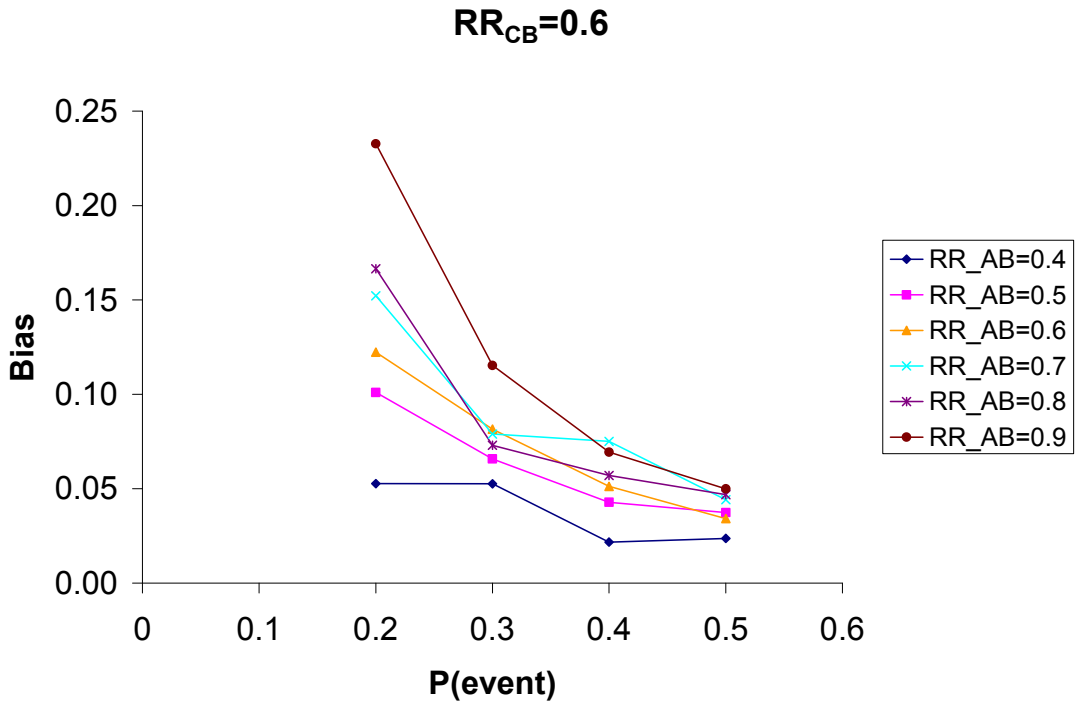
d)


RR_CB=0.5
Bias
P(event)
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

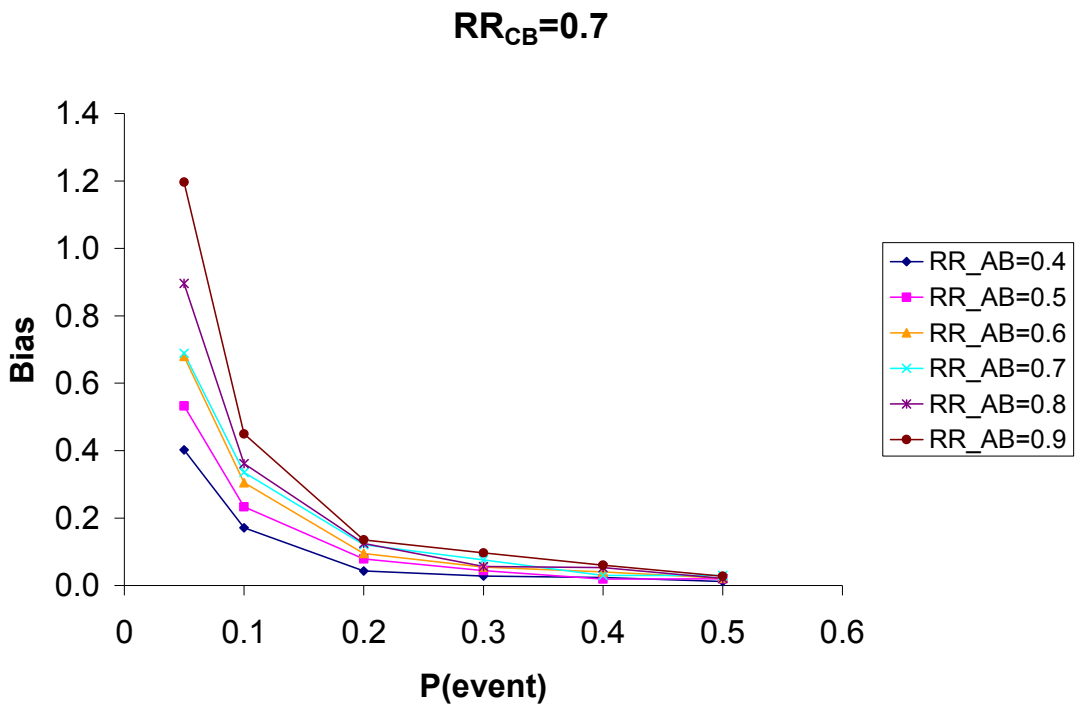
e)


RR_CB=0.6
Bias
P(event)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

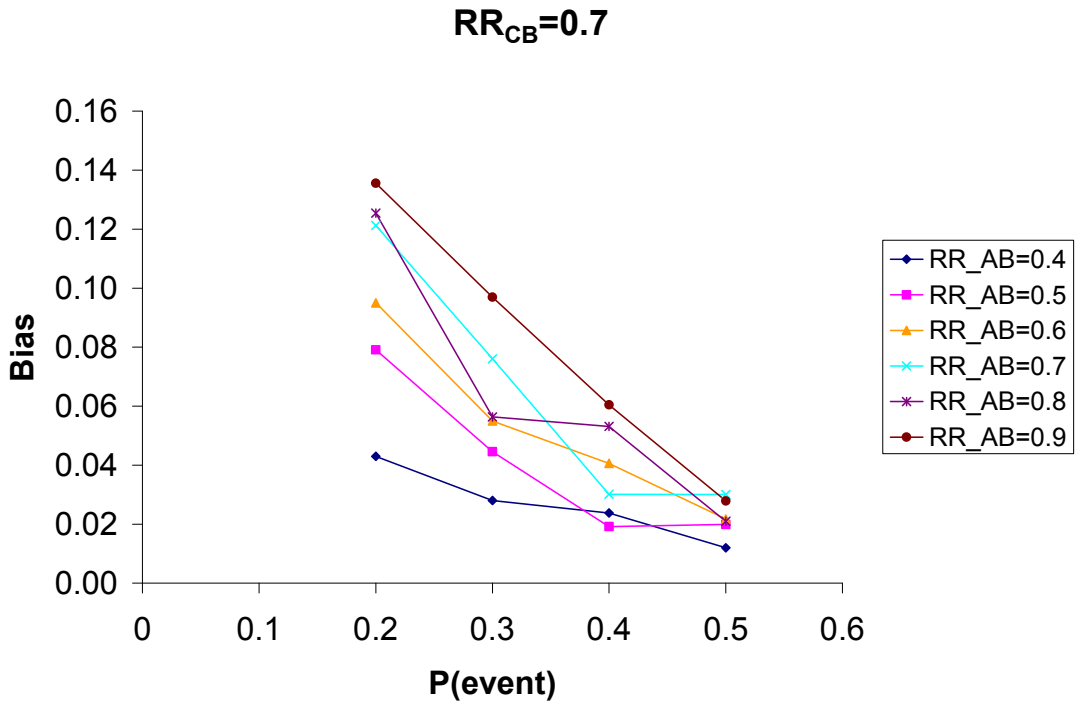
f)


RR_CB=0.6
Bias
P(event)
0.25
0.20
0.15
0.10
0.05
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

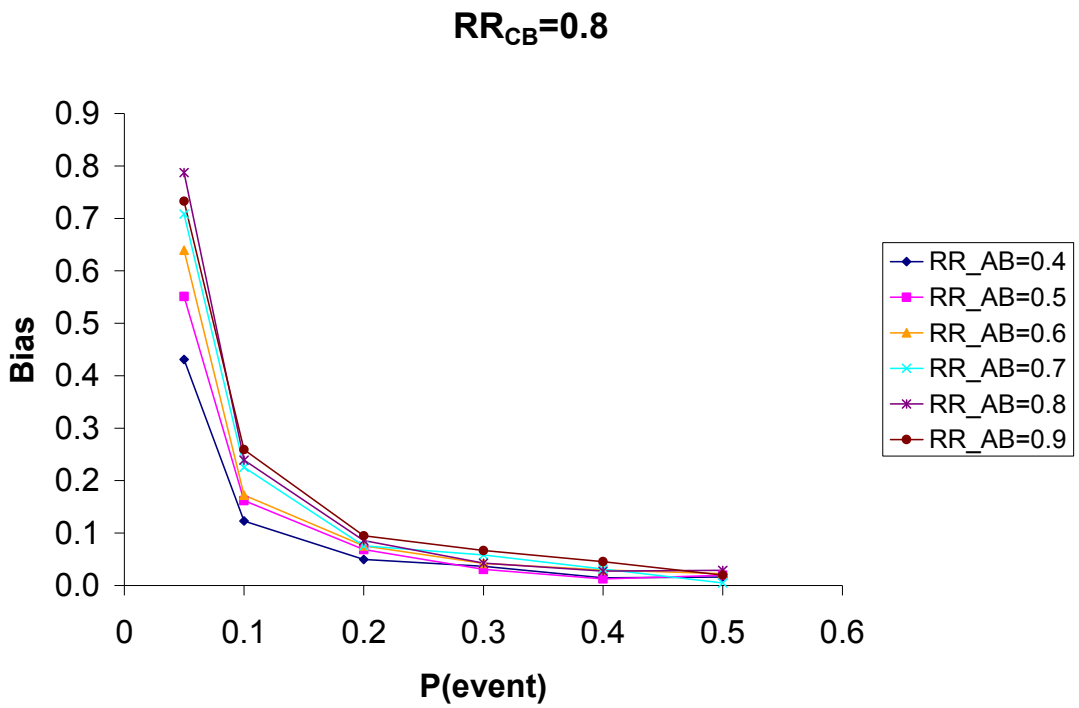
g)


RR_CB=0.7
Bias
P(event)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

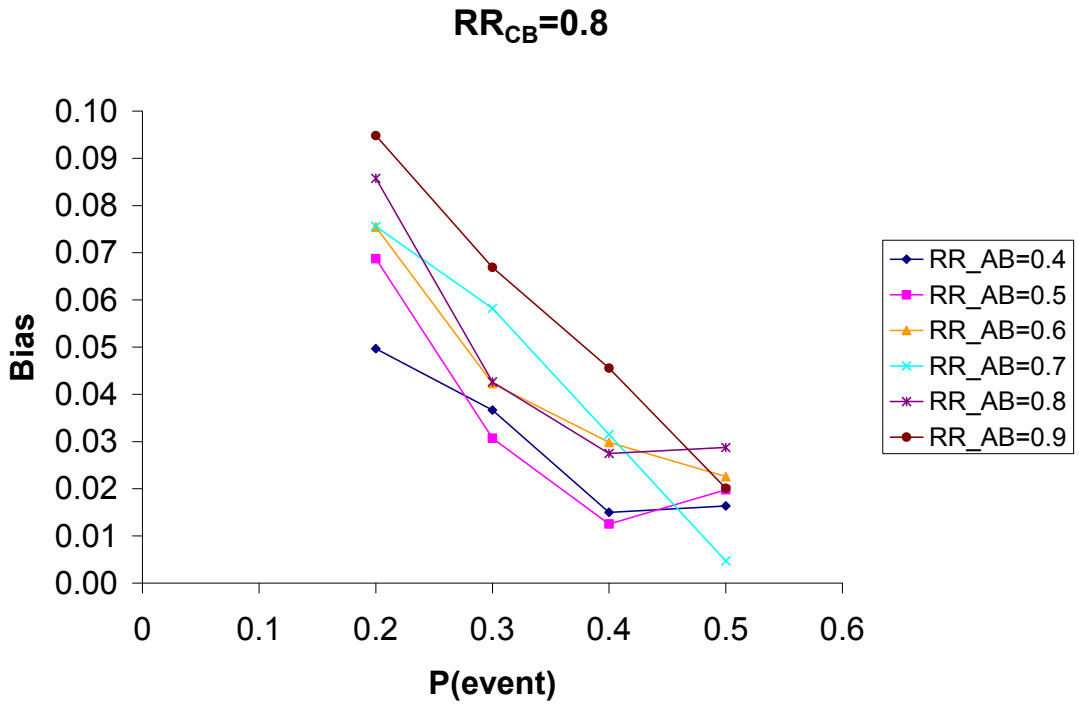
h)


RR_CB=0.7
Bias
P(event)
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

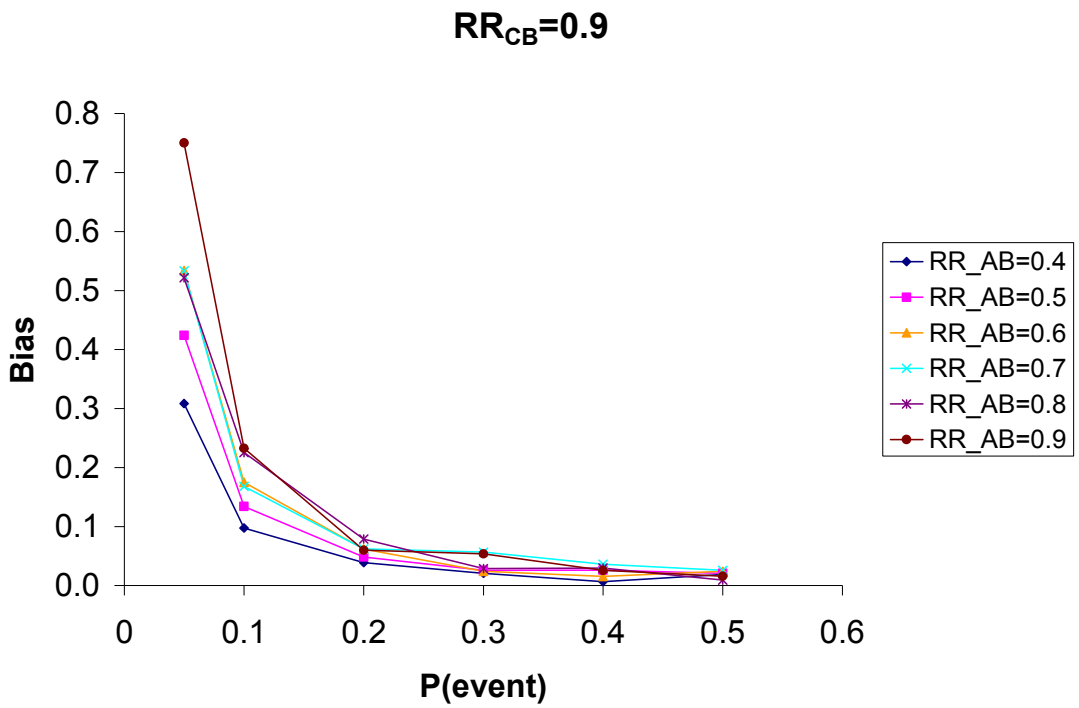
i)


RR_CB=0.8
Bias
P(event)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

j)


RR_CB=0.8
Bias
P(event)
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.10
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

k)


RR_CB=0.9
Bias
P(event)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

l)

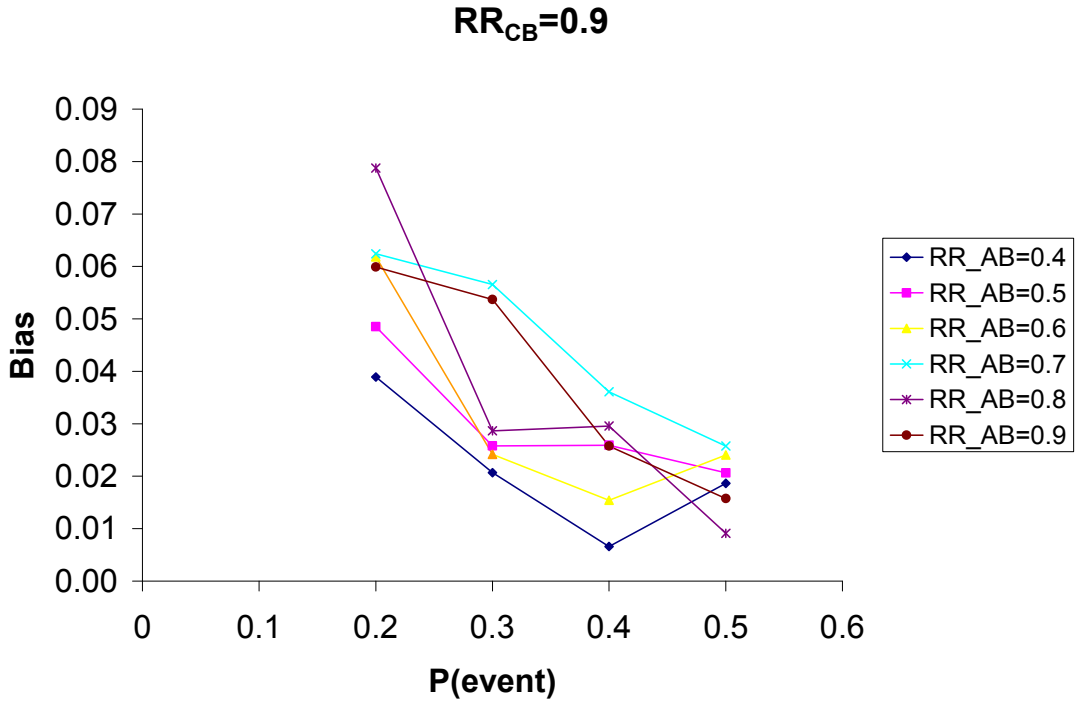
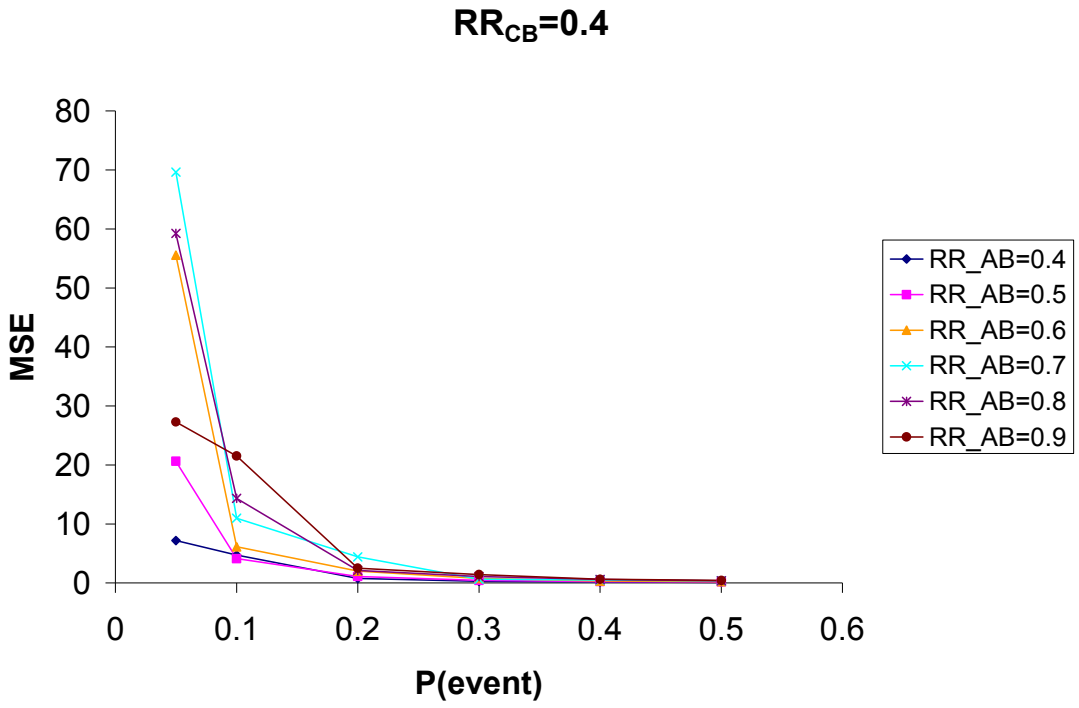
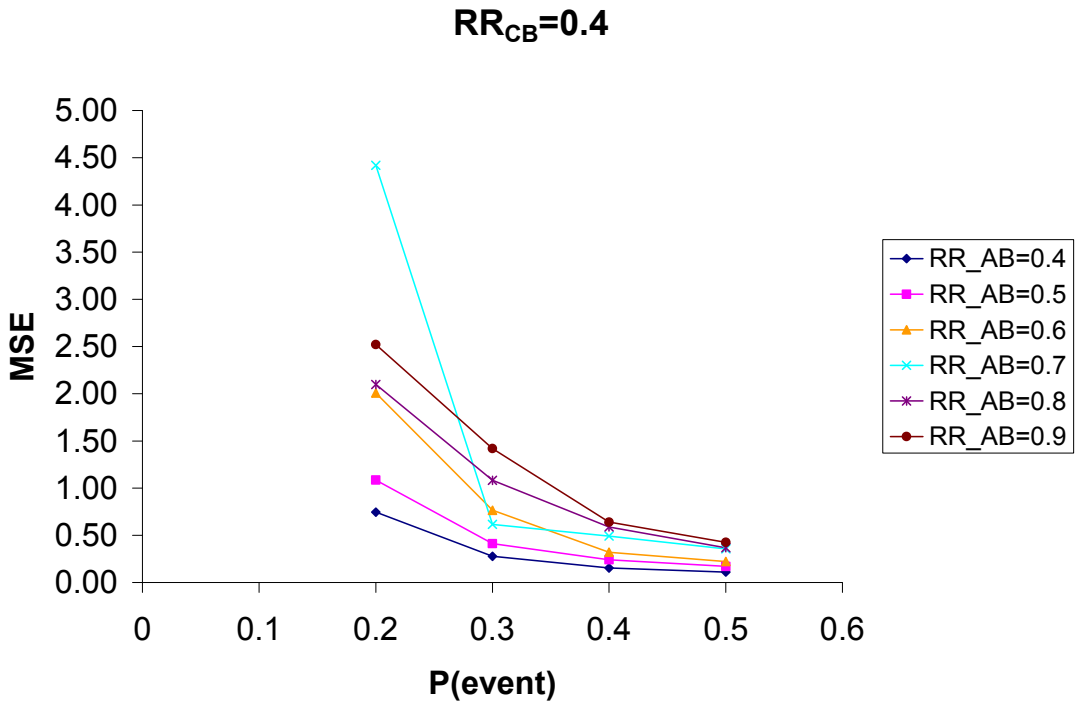

RR_CB=0.9
Bias
P(event)
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

Figure B.2.4: Mean square error (MSE) for indirect relative risk (RR) estimates for various parameter settings.

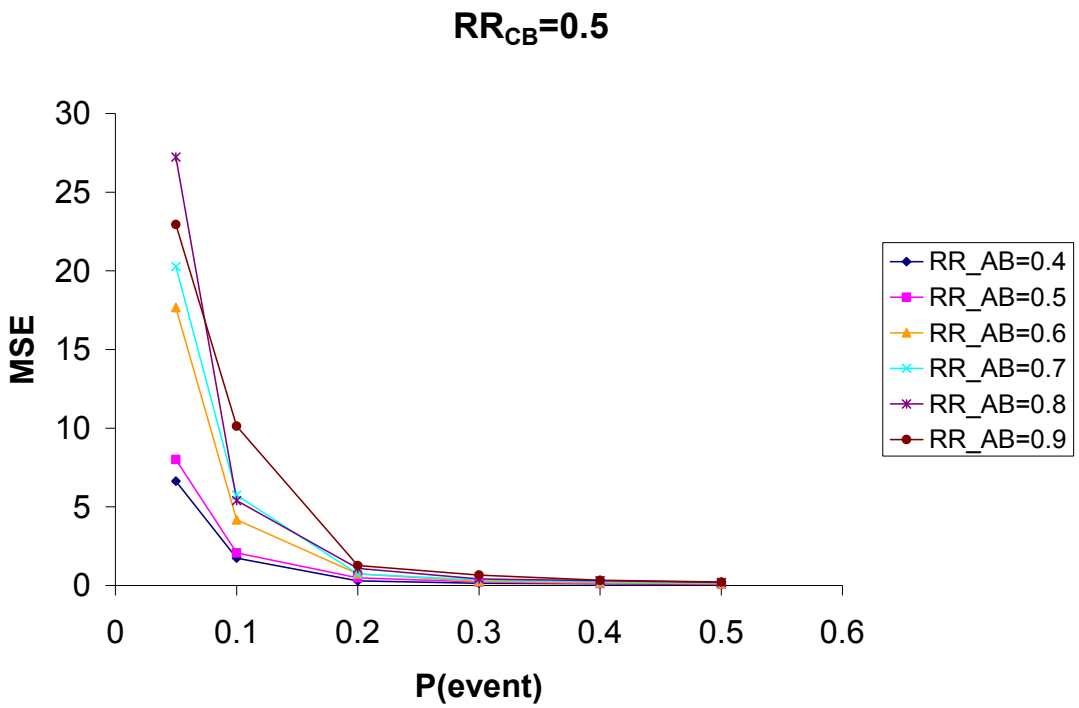
a)



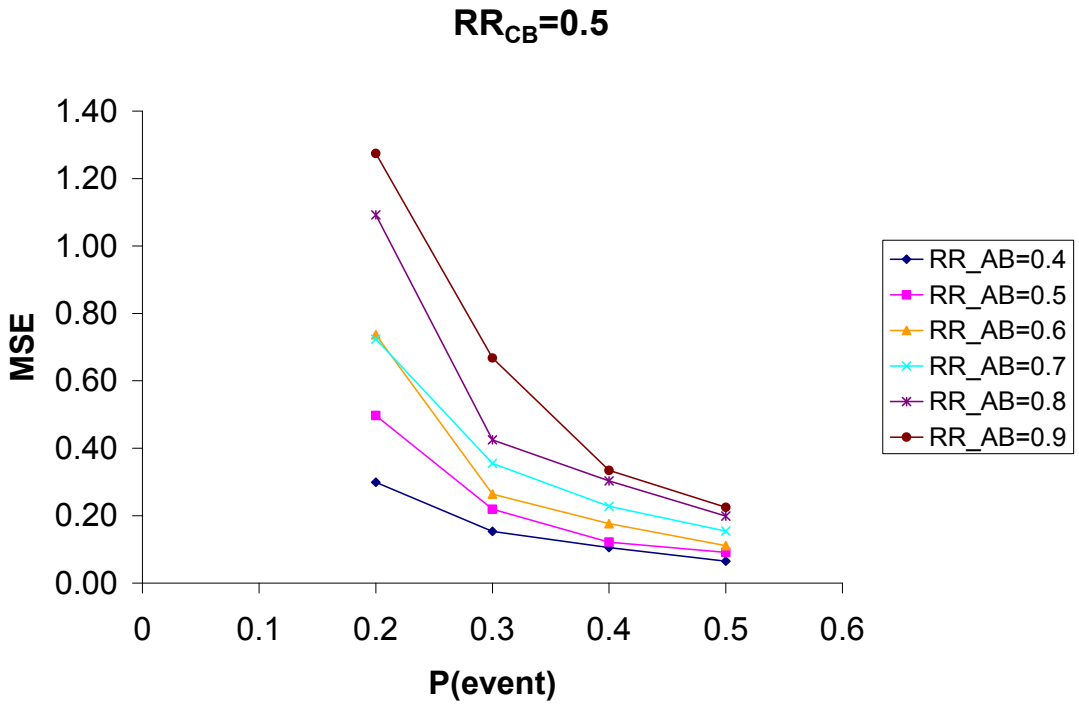
b)



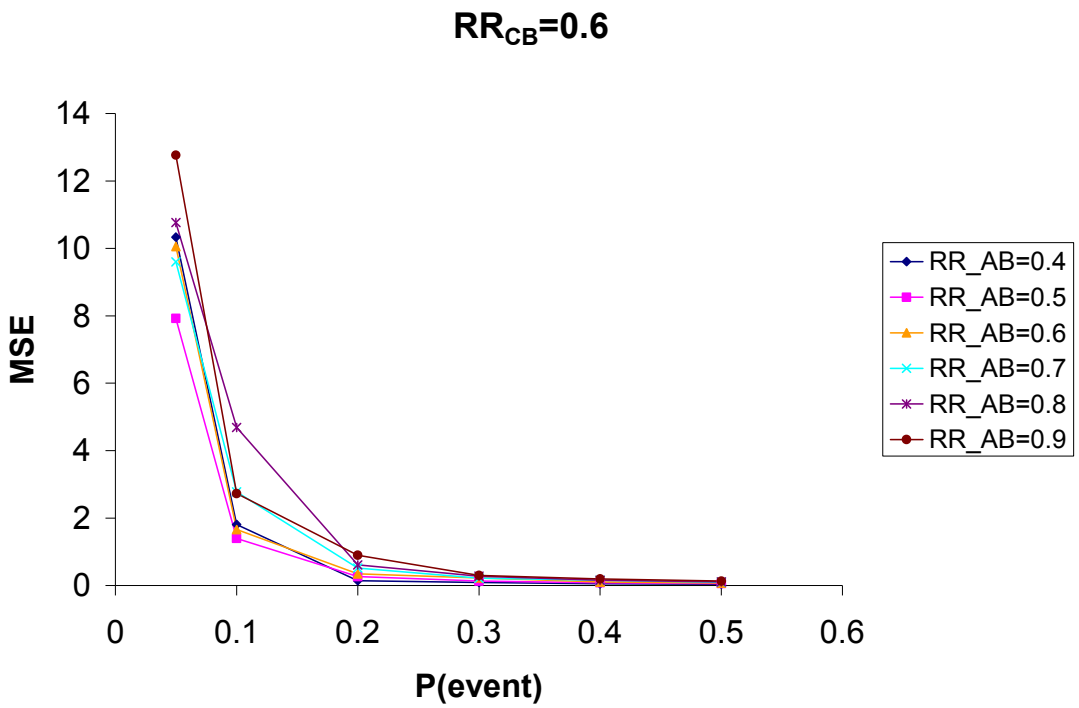
c)


RR_CB=0.5
MSE
P(event)
0
5
10
15
20
25
30
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

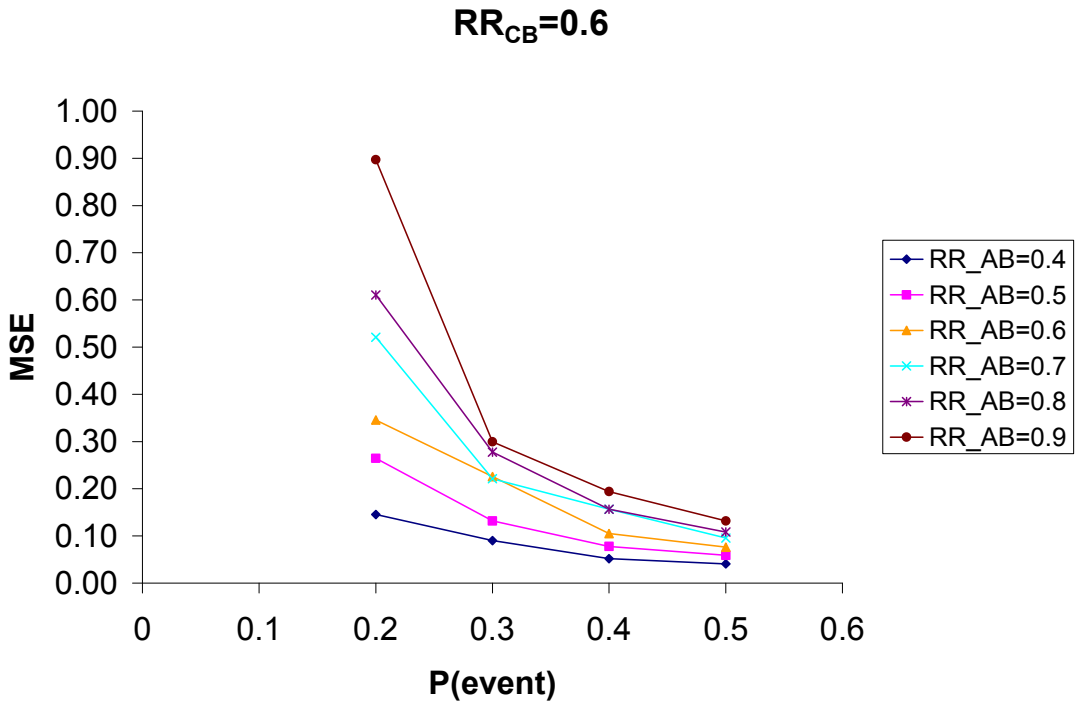
d)


RR_CB=0.5
MSE
P(event)
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

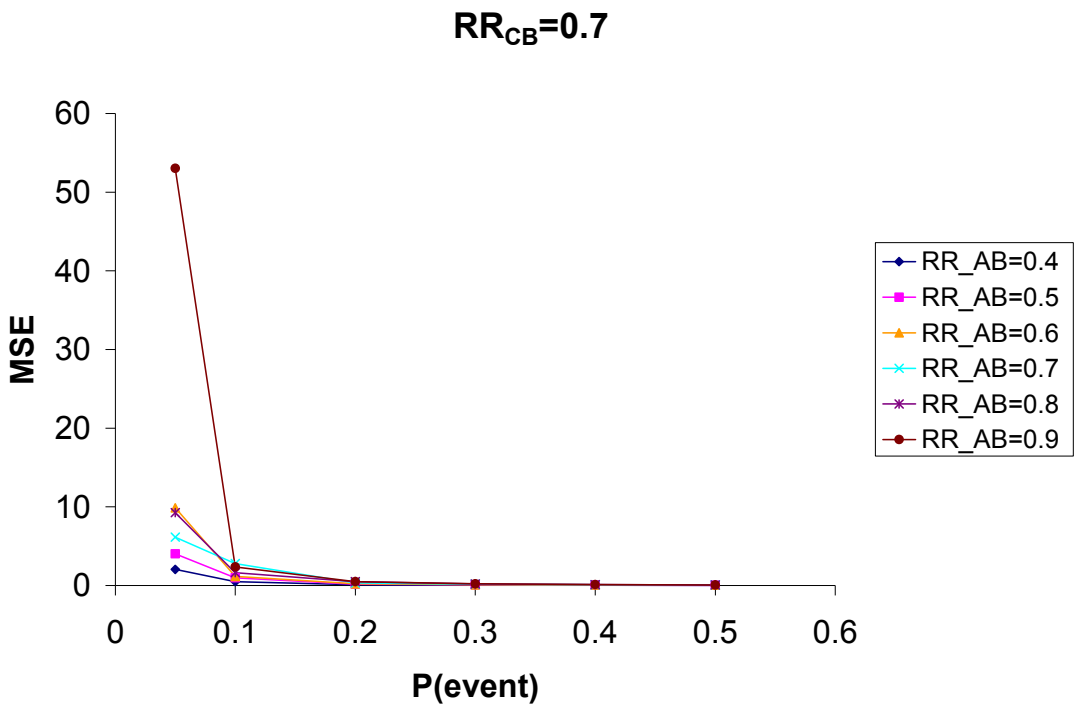
e)


RR_CB=0.6
MSE
P(event)
0
2
4
6
8
10
12
14
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

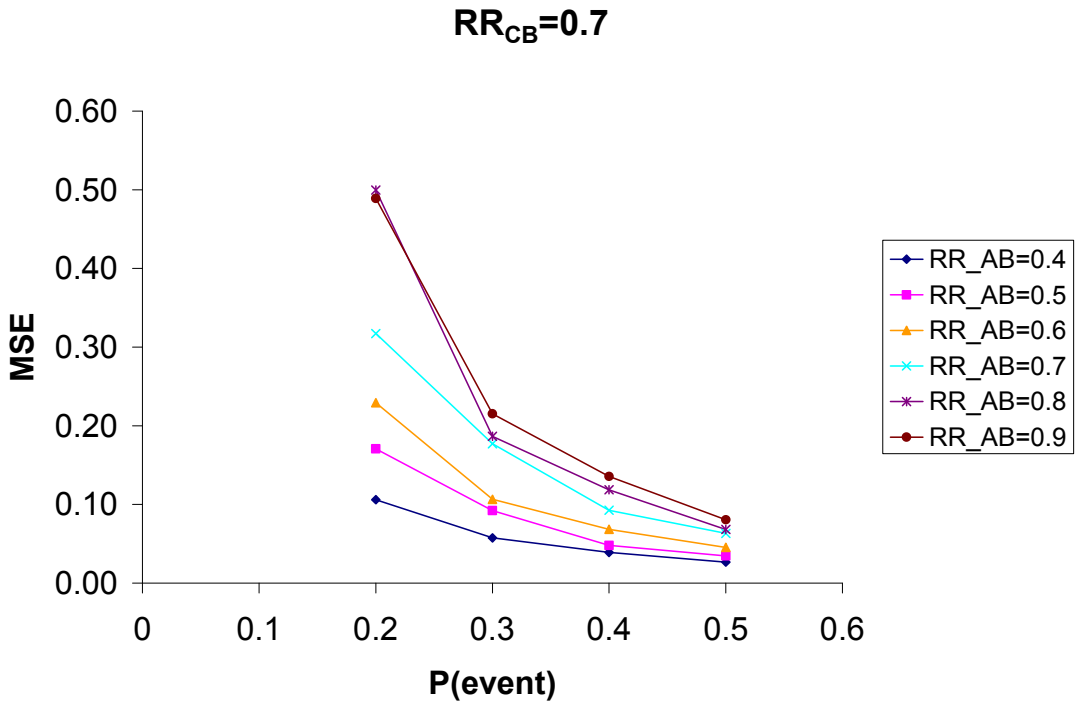
f)


RR_CB=0.6
MSE
P(event)
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

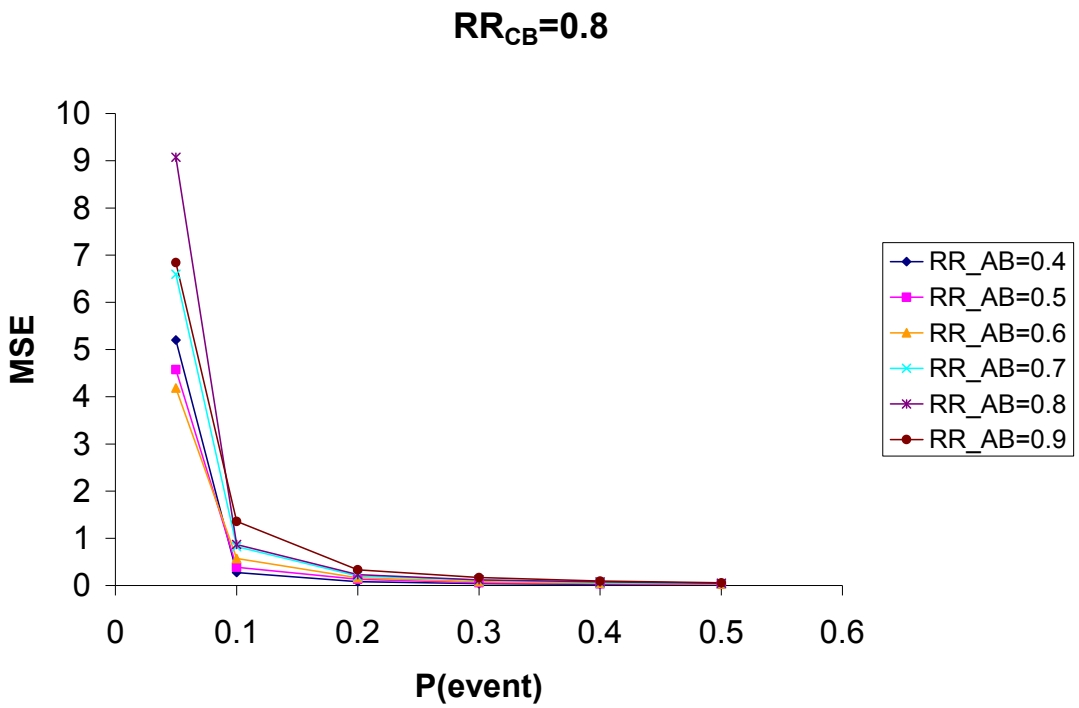
g)


RR_CB=0.7
MSE
P(event)
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

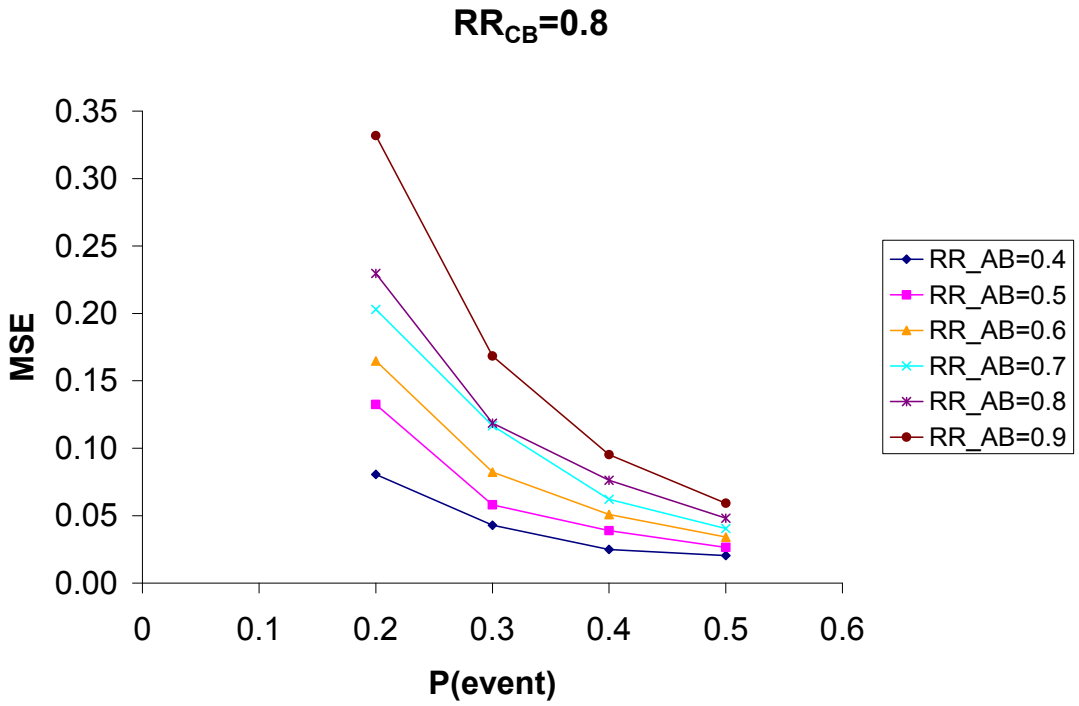
h)


RR_CB=0.7
MSE
P(event)
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

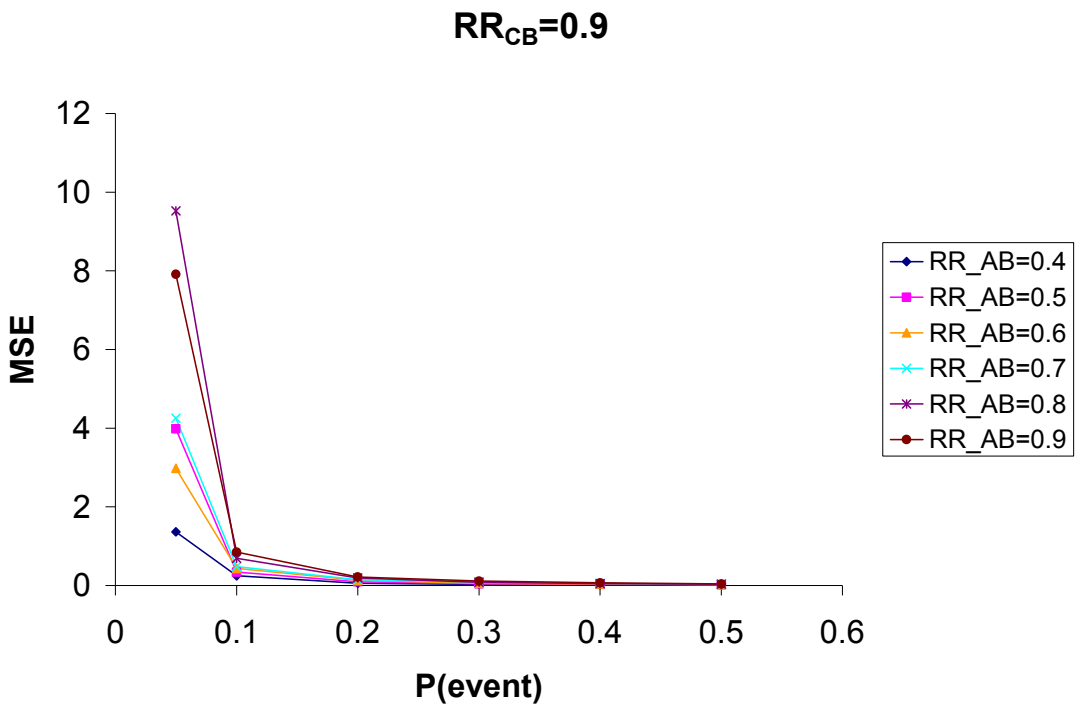
i)


RR_CB=0.8
MSE
P(event)
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

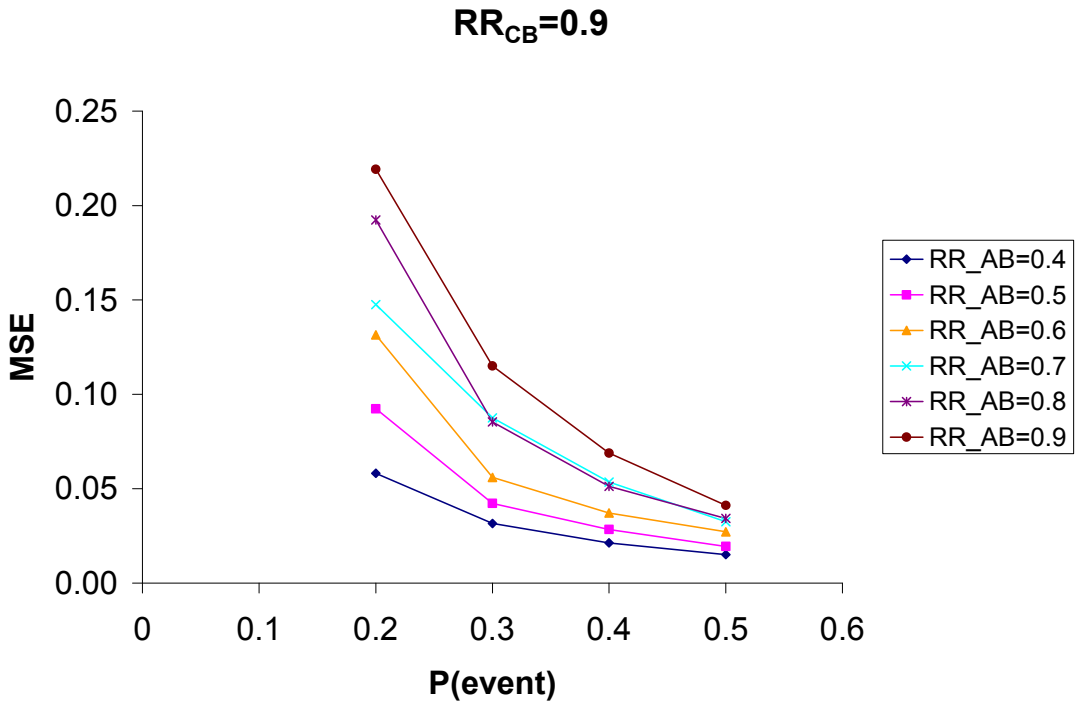
j)


RR_CB=0.8
MSE
P(event)
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

k)


RR_CB=0.9
MSE
P(event)
0
2
4
6
8
10
12
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

l)


RR_CB=0.9
MSE
P(event)
0.00
0.05
0.10
0.15
0.20
0.25
0
0.1
0.2
0.3
0.4
0.5
0.6
RR_AB=0.4
RR_AB=0.5
RR_AB=0.6
RR_AB=0.7
RR_AB=0.8
RR_AB=0.9

B.3 Simulation Results for the Odds Ratio

For each of the three populations (A, B and C) in the Monte Carlo simulation for the odds ratio, the outcome risk level for each population was selected according to a specific combination of the parameters OR_{AB}, OR_{CB}, and $P(E \mid B)$, as follows:

OR_{AB}: 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
OR_{CB}: 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
$P(E \mid B)$: 0.05, 0.1, 0.2, 0.3, 0.4, 0.5

For the various combinations of these parameters, the results of the simulation for the bias, variance and MSE for the direct and indirect estimators of the OR are provided in Table B.3.1.

As an illustration of the frequency distribution of the estimators, the frequency distributions for the parameter settings OR_{AB}=0.6, OR_{CB}=0.8 and $P(E \mid B)$=0.05,0.3,0.5 for the direct and indirect estimators are presented graphically in Figure B.3.1 on the logarithmic scale. It is apparent from these figures that both estimators have a mound shape, symmetric distribution (on the logarithmic scale). The indirect estimator has a larger variance and bias.

Table B.3.1: Bias, variance and mean square error (MSE) of direct and indirect odds ratio (OR) estimators for different settings of the indicated parameters (k=3 treatments)

Odds Ratio Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
OR_{AB}=0.9	**0.05**	0.292	0.100	0.302	4.521	0.370	4.658
OR_{CB}=1.5	**0.1**	0.098	0.051	0.101	0.309	0.108	0.321
	0.2	0.044	0.026	0.045	0.103	0.058	0.107
	0.3	0.035	0.025	0.035	0.068	0.042	0.070
	0.4	0.032	0.023	0.032	0.077	0.047	0.080
	0.5	0.029	0.025	0.030	0.067	0.034	0.068
OR_{AB}=0.9	**0.05**	0.687	0.156	0.711	2.458	0.429	2.642
OR_{CB}=1.3	**0.1**	0.161	0.086	0.168	0.394	0.152	0.417
	0.2	0.065	0.038	0.066	0.173	0.086	0.180
	0.3	0.054	0.035	0.055	0.107	0.056	0.110
	0.4	0.042	0.032	0.043	0.099	0.045	0.101
	0.5	0.037	0.027	0.037	0.096	0.066	0.101
OR_{AB}=0.9	**0.05**	1.105	0.243	1.164	2.345	0.399	2.504
OR_{CB}=1.1	**0.1**	0.251	0.111	0.263	0.652	0.214	0.698
	0.2	0.098	0.044	0.100	0.244	0.113	0.257
	0.3	0.070	0.053	0.073	0.136	0.066	0.140
	0.4	0.055	0.018	0.055	0.120	0.054	0.123
	0.5	0.053	0.026	0.053	0.128	0.056	0.131
OR_{AB}=0.8	**0.05**	0.224	0.092	0.233	1.284	0.255	1.349
OR_{CB}=1.5	**0.1**	0.076	0.049	0.078	0.199	0.116	0.213

Odds Ratio Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.2	0.038	0.028	0.039	0.083	0.059	0.086
	0.3	0.026	0.026	0.027	0.058	0.047	0.061
	0.4	0.025	0.018	0.025	0.055	0.041	0.056
	0.5	0.026	0.019	0.027	0.049	0.036	0.051
OR_{AB}=0.8	**0.05**	0.742	0.173	0.772	1.628	0.358	1.756
OR_{CB}=1.3	**0.1**	0.110	0.049	0.112	0.327	0.152	0.350
	0.2	0.053	0.045	0.055	0.127	0.074	0.133
	0.3	0.038	0.027	0.039	0.079	0.060	0.083
	0.4	0.032	0.019	0.032	0.076	0.040	0.078
	0.5	0.031	0.019	0.031	0.072	0.058	0.075
OR_{AB}=0.8	**0.05**	0.600	0.192	0.637	2.134	0.359	2.263
OR_{CB}=1.1	**0.1**	0.138	0.073	0.143	0.599	0.214	0.645
	0.2	0.088	0.055	0.091	0.177	0.093	0.186
	0.3	0.057	0.038	0.058	0.120	0.077	0.126
	0.4	0.044	0.019	0.044	0.097	0.057	0.100
	0.5	0.041	0.012	0.041	0.093	0.061	0.097
OR_{AB}=0.7	**0.05**	0.235	0.112	0.248	2.192	0.295	2.279
OR_{CB}=1.5	**0.1**	0.061	0.036	0.062	0.166	0.094	0.175
	0.2	0.030	0.017	0.030	0.074	0.062	0.078
	0.3	0.024	0.022	0.025	0.056	0.042	0.058
	0.4	0.020	0.016	0.020	0.047	0.038	0.049
	0.5	0.020	0.016	0.020	0.043	0.039	0.044
OR_{AB}=0.7	**0.05**	0.279	0.112	0.292	1.947	0.281	2.026
OR_{CB}=1.3	**0.1**	0.088	0.041	0.090	0.347	0.152	0.370
	0.2	0.043	0.029	0.044	0.114	0.086	0.122
	0.3	0.028	0.013	0.028	0.061	0.036	0.063
	0.4	0.025	0.012	0.025	0.055	0.037	0.056
	0.5	0.025	0.019	0.025	0.055	0.041	0.057
OR_{AB}=0.7	**0.05**	0.959	0.211	1.004	3.053	0.428	3.237
OR_{CB}=1.1	**0.1**	0.158	0.062	0.162	0.310	0.119	0.324
	0.2	0.062	0.038	0.063	0.150	0.078	0.156
	0.3	0.039	0.033	0.040	0.095	0.052	0.098
	0.4	0.039	0.025	0.039	0.076	0.057	0.079
	0.5	0.030	0.019	0.031	0.074	0.052	0.077
OR_{AB}=0.9	**0.05**	1.418	0.302	1.509	12.063	0.836	12.763
OR_{CB}=0.9	**0.1**	0.514	0.164	0.541	1.321	0.308	1.416
	0.2	0.159	0.050	0.161	0.351	0.116	0.364
	0.3	0.112	0.050	0.115	0.240	0.096	0.249
	0.4	0.087	0.046	0.089	0.200	0.075	0.206
	0.5	0.088	0.043	0.089	0.212	0.086	0.219
OR_{AB}=0.9	**0.05**	2.634	0.450	2.836	7.954	0.870	8.711
OR_{CB}=0.8	**0.1**	0.632	0.196	0.671	1.873	0.382	2.019

Odds Ratio Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.2	0.216	0.065	0.220	0.454	0.166	0.482
	0.3	0.135	0.050	0.137	0.331	0.113	0.344
	0.4	0.114	0.039	0.115	0.237	0.066	0.241
	0.5	0.102	0.032	0.103	0.252	0.082	0.259
OR_{AB}=0.9	**0.05**	2.763	0.432	2.950	18.370	1.202	19.816
OR_{CB}=0.7	**0.1**	1.022	0.303	1.114	2.039	0.376	2.181
	0.2	0.314	0.111	0.326	0.723	0.224	0.773
	0.3	0.193	0.065	0.197	0.460	0.119	0.474
	0.4	0.142	0.048	0.145	0.405	0.134	0.423
	0.5	0.148	0.051	0.150	0.271	0.097	0.281
OR_{AB}=0.9	**0.05**	4.862	0.654	5.291	18.896	1.394	20.838
OR_{CB}=0.6	**0.1**	1.449	0.227	1.500	3.202	0.447	3.402
	0.2	0.677	0.185	0.711	1.368	0.332	1.478
	0.3	0.257	0.058	0.260	0.615	0.087	0.623
	0.4	0.224	0.093	0.233	0.458	0.164	0.485
	0.5	0.183	0.070	0.188	0.402	0.129	0.418
OR_{AB}=0.9	**0.05**	6.691	0.817	7.359	27.829	1.632	30.493
OR_{CB}=0.5	**0.1**	2.263	0.489	2.502	8.286	0.801	8.927
	0.2	0.994	0.259	1.062	2.258	0.379	2.402
	0.3	0.444	0.142	0.464	1.045	0.254	1.110
	0.4	0.357	0.100	0.367	0.731	0.187	0.766
	0.5	0.295	0.104	0.305	0.644	0.176	0.675
OR_{AB}=0.9	**0.05**	9.812	0.921	10.659	35.745	1.784	38.929
OR_{CB}=0.4	**0.1**	8.698	0.758	9.273	16.537	0.976	17.489
	0.2	1.853	0.323	1.957	3.324	0.470	3.545
	0.3	1.016	0.246	1.076	1.550	0.290	1.634
	0.4	0.542	0.103	0.553	1.150	0.210	1.194
	0.5	0.516	0.146	0.537	1.020	0.212	1.065
OR_{AB}=0.8	**0.05**	1.871	0.339	1.986	3.014	0.483	3.247
OR_{CB}=0.9	**0.1**	0.299	0.114	0.312	0.931	0.236	0.986
	0.2	0.122	0.070	0.127	0.304	0.117	0.317
	0.3	0.086	0.042	0.088	0.214	0.088	0.222
	0.4	0.069	0.035	0.070	0.161	0.088	0.169
	0.5	0.061	0.023	0.061	0.132	0.060	0.135
OR_{AB}=0.8	**0.05**	2.222	0.384	2.370	5.351	0.691	5.828
OR_{CB}=0.8	**0.1**	0.504	0.178	0.536	0.955	0.232	1.009
	0.2	0.185	0.079	0.191	0.359	0.139	0.378
	0.3	0.112	0.047	0.115	0.250	0.123	0.265
	0.4	0.092	0.051	0.094	0.173	0.076	0.179
	0.5	0.082	0.026	0.082	0.172	0.055	0.175
OR_{AB}=0.8	**0.05**	3.283	0.519	3.552	8.331	0.880	9.105
OR_{CB}=0.7	**0.1**	1.863	0.234	1.917	1.790	0.334	1.901

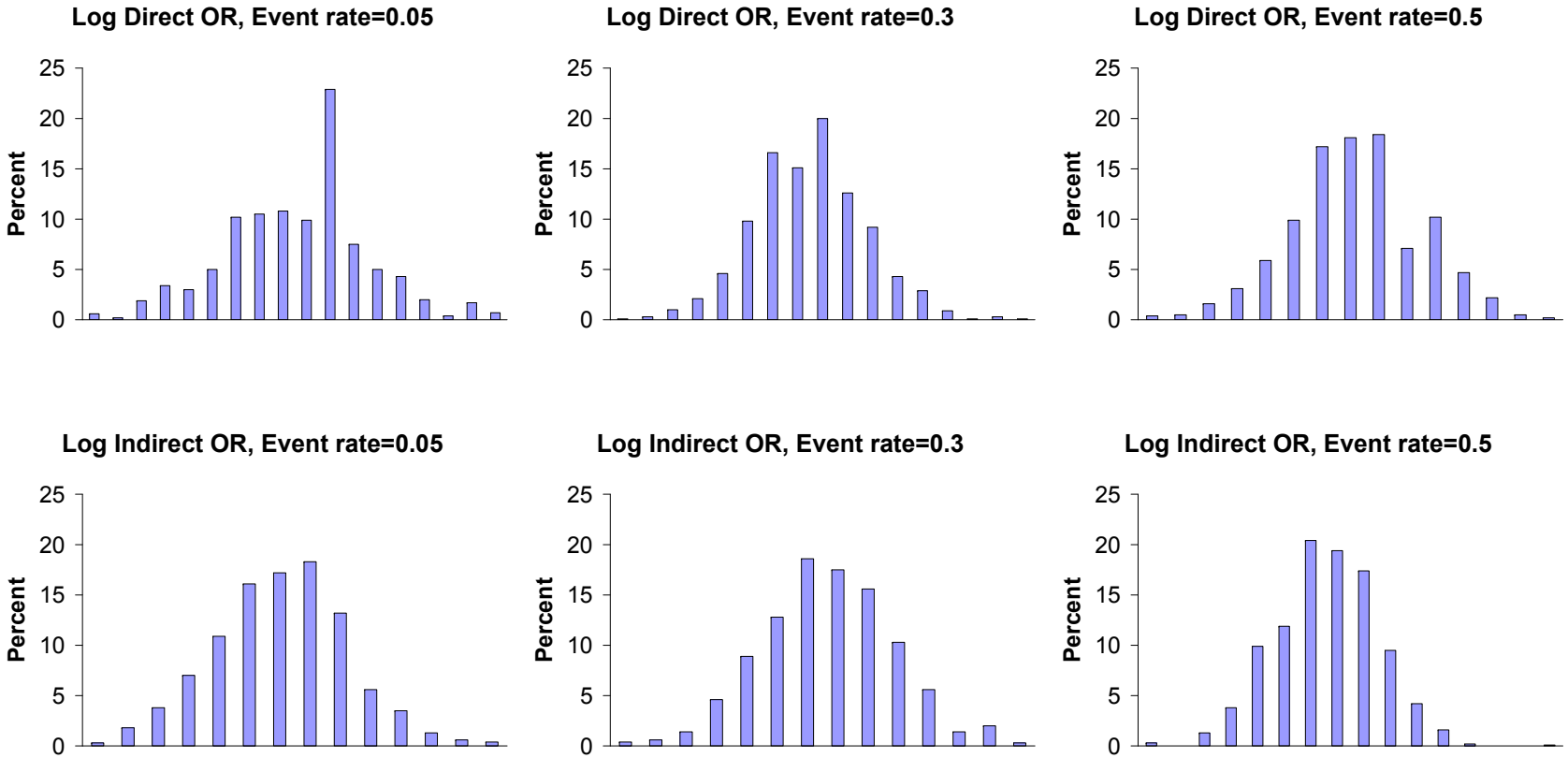
Odds Ratio Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.2	0.273	0.113	0.286	0.600	0.198	0.639
	0.3	0.177	0.079	0.183	0.390	0.133	0.408
	0.4	0.126	0.043	0.128	0.242	0.087	0.250
	0.5	0.097	0.025	0.098	0.248	0.098	0.258
OR_{AB}=0.8	**0.05**	3.879	0.596	4.235	9.488	1.012	10.513
OR_{CB}=0.6	**0.1**	1.793	0.349	1.915	3.953	0.481	4.184
	0.2	0.407	0.147	0.428	0.772	0.189	0.808
	0.3	0.232	0.074	0.237	0.452	0.125	0.468
	0.4	0.180	0.068	0.185	0.357	0.104	0.368
	0.5	0.153	0.071	0.158	0.336	0.078	0.342
OR_{AB}=0.8	**0.05**	4.360	0.644	4.774	25.852	1.626	28.496
OR_{CB}=0.5	**0.1**	3.915	0.469	4.135	5.737	0.718	6.253
	0.2	0.589	0.170	0.618	1.215	0.332	1.326
	0.3	0.356	0.114	0.369	0.696	0.191	0.733
	0.4	0.257	0.071	0.262	0.673	0.187	0.708
	0.5	0.246	0.057	0.249	0.489	0.127	0.505
OR_{AB}=0.8	**0.05**	7.834	1.017	8.868	27.979	1.679	30.799
OR_{CB}=0.4	**0.1**	7.550	0.792	8.177	12.854	1.161	14.201
	0.2	1.332	0.256	1.398	3.225	0.580	3.561
	0.3	0.660	0.137	0.678	1.236	0.210	1.280
	0.4	0.531	0.128	0.548	0.999	0.230	1.052
	0.5	0.433	0.136	0.451	0.797	0.097	0.807
OR_{AB}=0.7	**0.05**	1.146	0.254	1.211	6.419	0.766	7.006
OR_{CB}=0.9	**0.1**	0.265	0.072	0.270	1.785	0.273	1.860
	0.2	0.095	0.041	0.096	0.272	0.106	0.284
	0.3	0.078	0.042	0.080	0.162	0.070	0.167
	0.4	0.055	0.023	0.056	0.131	0.064	0.135
	0.5	0.052	0.032	0.053	0.129	0.065	0.133
OR_{AB}=0.7	**0.05**	1.250	0.316	1.350	6.023	0.714	6.533
OR_{CB}=0.8	**0.1**	0.382	0.124	0.398	1.096	0.296	1.183
	0.2	0.156	0.076	0.162	0.277	0.096	0.286
	0.3	0.086	0.038	0.088	0.231	0.102	0.242
	0.4	0.076	0.043	0.077	0.178	0.087	0.186
	0.5	0.064	0.025	0.064	0.161	0.079	0.167
OR_{AB}=0.7	**0.05**	2.140	0.421	2.318	10.296	0.841	11.004
OR_{CB}=0.7	**0.1**	0.654	0.229	0.706	1.767	0.363	1.899
	0.2	0.264	0.091	0.272	0.476	0.156	0.501
	0.3	0.132	0.052	0.135	0.307	0.130	0.324
	0.4	0.094	0.051	0.097	0.224	0.096	0.233
	0.5	0.075	0.029	0.076	0.204	0.081	0.211
OR_{AB}=0.7	**0.05**	3.180	0.510	3.440	22.883	1.185	24.288
OR_{CB}=0.6	**0.1**	2.043	0.362	2.174	6.431	0.423	6.611

Odds Ratio Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.2	0.276	0.098	0.286	0.717	0.209	0.760
	0.3	0.192	0.062	0.196	0.468	0.132	0.485
	0.4	0.135	0.040	0.136	0.321	0.117	0.335
	0.5	0.124	0.030	0.125	0.295	0.108	0.307
OR_{AB}=0.7	**0.05**	4.766	0.701	5.257	15.102	1.295	16.778
OR_{CB}=0.5	**0.1**	2.429	0.383	2.576	4.509	0.607	4.878
	0.2	0.537	0.141	0.557	0.915	0.224	0.965
	0.3	0.270	0.105	0.281	0.646	0.186	0.681
	0.4	0.237	0.076	0.243	0.432	0.121	0.447
	0.5	0.168	0.063	0.172	0.400	0.125	0.416
OR_{AB}=0.7	**0.05**	5.760	0.830	6.450	45.917	1.923	49.614
OR_{CB}=0.4	**0.1**	6.011	0.701	6.503	9.214	0.805	9.862
	0.2	1.190	0.253	1.254	1.607	0.326	1.713
	0.3	0.551	0.166	0.579	0.972	0.180	1.004
	0.4	0.431	0.113	0.444	0.632	0.148	0.654
	0.5	0.292	0.100	0.302	0.601	0.131	0.618
OR_{AB}=0.6	**0.05**	1.002	0.279	1.080	2.456	0.464	2.671
OR_{CB}=0.9	**0.1**	0.469	0.104	0.480	0.525	0.198	0.564
	0.2	0.089	0.043	0.091	0.182	0.098	0.192
	0.3	0.055	0.042	0.057	0.123	0.073	0.128
	0.4	0.042	0.027	0.043	0.092	0.052	0.095
	0.5	0.038	0.028	0.039	0.080	0.043	0.082
OR_{AB}=0.6	**0.05**	1.688	0.381	1.833	5.183	0.702	5.676
OR_{CB}=0.8	**0.1**	0.257	0.097	0.267	0.827	0.218	0.874
	0.2	0.126	0.059	0.130	0.231	0.096	0.240
	0.3	0.078	0.030	0.078	0.154	0.094	0.162
	0.4	0.055	0.024	0.056	0.135	0.068	0.139
	0.5	0.045	0.034	0.046	0.108	0.061	0.112
OR_{AB}=0.6	**0.05**	2.040	0.375	2.180	8.042	0.851	8.765
OR_{CB}=0.7	**0.1**	1.022	0.235	1.077	1.680	0.299	1.769
	0.2	0.155	0.067	0.159	0.316	0.119	0.330
	0.3	0.096	0.028	0.097	0.190	0.071	0.195
	0.4	0.074	0.037	0.075	0.135	0.061	0.138
	0.5	0.065	0.033	0.066	0.137	0.074	0.143
OR_{AB}=0.6	**0.05**	3.463	0.571	3.790	12.877	0.894	13.677
OR_{CB}=0.6	**0.1**	0.986	0.261	1.054	2.732	0.434	2.921
	0.2	0.240	0.132	0.257	0.462	0.148	0.484
	0.3	0.131	0.031	0.132	0.242	0.109	0.254
	0.4	0.115	0.053	0.118	0.206	0.071	0.211
	0.5	0.092	0.057	0.095	0.182	0.079	0.188
OR_{AB}=0.6	**0.05**	3.975	0.628	4.370	16.601	1.236	18.129
OR_{CB}=0.5	**0.1**	2.048	0.337	2.162	3.828	0.572	4.156

Odds Ratio Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.2	0.409	0.142	0.430	0.747	0.233	0.801
	0.3	0.244	0.090	0.252	0.508	0.164	0.535
	0.4	0.156	0.066	0.161	0.292	0.111	0.304
	0.5	0.150	0.067	0.155	0.290	0.093	0.299
OR_{AB}=0.6	**0.05**	5.720	0.817	6.387	24.614	1.600	27.175
OR_{CB}=0.4	**0.1**	3.976	0.536	4.264	8.805	0.902	9.619
	0.2	0.673	0.149	0.696	1.747	0.293	1.833
	0.3	0.461	0.142	0.481	0.739	0.177	0.770
	0.4	0.302	0.086	0.310	0.542	0.156	0.566
	0.5	0.231	0.062	0.235	0.474	0.130	0.491
OR_{AB}=0.5	**0.05**	0.745	0.267	0.816	3.311	0.497	3.558
OR_{CB}=0.9	**0.1**	0.190	0.074	0.195	0.686	0.200	0.727
	0.2	0.066	0.033	0.067	0.126	0.060	0.129
	0.3	0.036	0.021	0.037	0.076	0.036	0.077
	0.4	0.031	0.016	0.032	0.067	0.036	0.069
	0.5	0.027	0.017	0.027	0.062	0.050	0.065
OR_{AB}=0.5	**0.05**	1.452	0.367	1.587	14.086	0.778	14.691
OR_{CB}=0.8	**0.1**	0.255	0.126	0.271	0.699	0.220	0.748
	0.2	0.088	0.038	0.089	0.152	0.090	0.160
	0.3	0.049	0.031	0.050	0.109	0.059	0.113
	0.4	0.039	0.019	0.040	0.086	0.050	0.088
	0.5	0.036	0.023	0.036	0.074	0.059	0.077
OR_{AB}=0.5	**0.05**	1.189	0.290	1.273	3.785	0.623	4.174
OR_{CB}=0.7	**0.1**	0.318	0.127	0.334	0.950	0.276	1.026
	0.2	0.144	0.065	0.148	0.236	0.114	0.249
	0.3	0.068	0.036	0.069	0.158	0.089	0.165
	0.4	0.054	0.030	0.055	0.110	0.065	0.114
	0.5	0.049	0.032	0.050	0.103	0.059	0.106
OR_{AB}=0.5	**0.05**	1.779	0.429	1.963	5.379	0.741	5.929
OR_{CB}=0.6	**0.1**	0.600	0.159	0.625	1.804	0.297	1.892
	0.2	0.182	0.077	0.188	0.378	0.176	0.409
	0.3	0.094	0.027	0.095	0.195	0.076	0.201
	0.4	0.078	0.044	0.080	0.171	0.078	0.177
	0.5	0.063	0.020	0.064	0.133	0.070	0.138
OR_{AB}=0.5	**0.05**	2.307	0.433	2.494	11.514	0.941	12.399
OR_{CB}=0.5	**0.1**	1.894	0.301	1.985	3.678	0.576	4.009
	0.2	0.303	0.092	0.311	0.521	0.163	0.548
	0.3	0.163	0.074	0.168	0.283	0.134	0.301
	0.4	0.114	0.037	0.115	0.269	0.086	0.276
	0.5	0.098	0.051	0.100	0.188	0.079	0.195
OR_{AB}=0.5	**0.05**	4.433	0.757	5.007	25.543	1.410	27.533
OR_{CB}=0.4	**0.1**	2.816	0.461	3.029	5.591	0.746	6.148

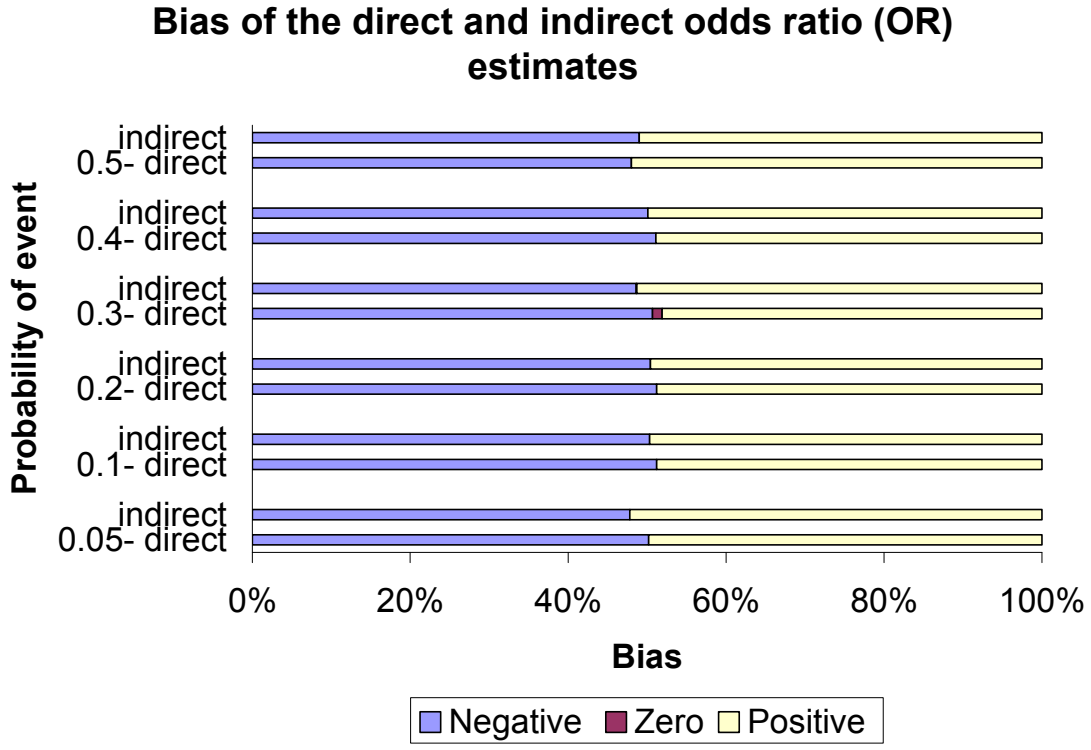
Odds Ratio Settings	Event Rate P(E|B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.2	0.596	0.164	0.623	1.140	0.283	1.220
	0.3	0.326	0.131	0.343	0.595	0.180	0.628
	0.4	0.196	0.061	0.200	0.337	0.095	0.346
	0.5	0.154	0.059	0.158	0.339	0.139	0.358
OR_{AB}=0.4	**0.05**	0.474	0.196	0.512	1.529	0.365	1.662
OR_{CB}=0.9	**0.1**	0.107	0.049	0.109	0.303	0.128	0.319
	0.2	0.042	0.023	0.043	0.093	0.049	0.096
	0.3	0.029	0.018	0.029	0.058	0.041	0.060
	0.4	0.023	0.021	0.023	0.046	0.044	0.048
	0.5	0.019	0.014	0.019	0.040	0.029	0.041
OR_{AB}=0.4	**0.05**	0.669	0.264	0.739	6.020	0.578	6.354
OR_{CB}=0.8	**0.1**	0.154	0.055	0.157	0.379	0.163	0.406
	0.2	0.057	0.051	0.060	0.122	0.071	0.127
	0.3	0.036	0.021	0.036	0.071	0.047	0.074
	0.4	0.026	0.017	0.027	0.051	0.040	0.053
	0.5	0.023	0.022	0.023	0.049	0.025	0.050
OR_{AB}=0.4	**0.05**	0.965	0.268	1.037	7.234	0.693	7.714
OR_{CB}=0.7	**0.1**	0.336	0.127	0.352	0.721	0.204	0.762
	0.2	0.088	0.057	0.091	0.156	0.094	0.164
	0.3	0.046	0.030	0.047	0.098	0.049	0.101
	0.4	0.034	0.022	0.035	0.066	0.029	0.067
	0.5	0.034	0.030	0.035	0.060	0.044	0.062
OR_{AB}=0.4	**0.05**	1.664	0.343	1.782	5.203	0.714	5.712
OR_{CB}=0.6	**0.1**	0.708	0.171	0.737	1.272	0.299	1.361
	0.2	0.132	0.075	0.137	0.227	0.087	0.235
	0.3	0.073	0.038	0.074	0.140	0.079	0.146
	0.4	0.053	0.027	0.054	0.096	0.045	0.098
	0.5	0.047	0.040	0.048	0.087	0.049	0.090
OR_{AB}=0.4	**0.05**	2.013	0.430	2.198	11.629	0.873	12.391
OR_{CB}=0.5	**0.1**	0.866	0.204	0.908	2.271	0.365	2.404
	0.2	0.196	0.084	0.203	0.342	0.132	0.360
	0.3	0.110	0.041	0.111	0.223	0.114	0.236
	0.4	0.080	0.043	0.082	0.158	0.089	0.166
	0.5	0.066	0.040	0.068	0.125	0.065	0.129
OR_{AB}=0.4	**0.05**	2.841	0.586	3.185	9.868	1.096	11.070
OR_{CB}=0.4	**0.1**	1.622	0.337	1.736	7.154	0.628	7.549
	0.2	0.370	0.118	0.384	0.569	0.178	0.600
	0.3	0.170	0.075	0.175	0.484	0.165	0.511
	0.4	0.129	0.073	0.134	0.299	0.113	0.312
	0.5	0.116	0.049	0.118	0.195	0.073	0.200

Figure B.3.1: Frequency distribution of direct and indirect odds ratio (OR) estimators for the parameter settings OR_{AB}=0.6, OR_{CB}=0.8, P(E | B)=0.05, 0.3, 0.5 (k=3 treatments)



For the parameter settings OR_{AB}=0.6, OR_{CB}=0.8, the direction of the biases over the 1000 samples are illustrated in Figure B.3.2 for the direct and indirect approaches under the different settings of the event rate P(E | B)): 0.05, 0.1, 0.2, 0.3, 0.4, 0.5.

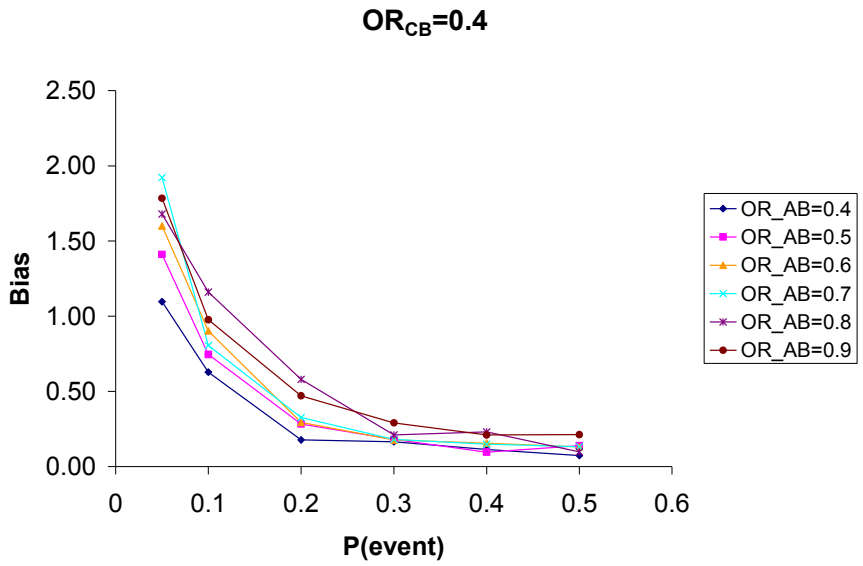
Figure B.3.2: Direction of the bias of the direct and indirect odds ratio (OR) estimates for the parameter settings RR_{AB}=0.6, RR_{CB}=0.8, P(E | B) =0.05, 0.1, 0.2, 0.3, 0.4, 0.5 (k=3 treatments)



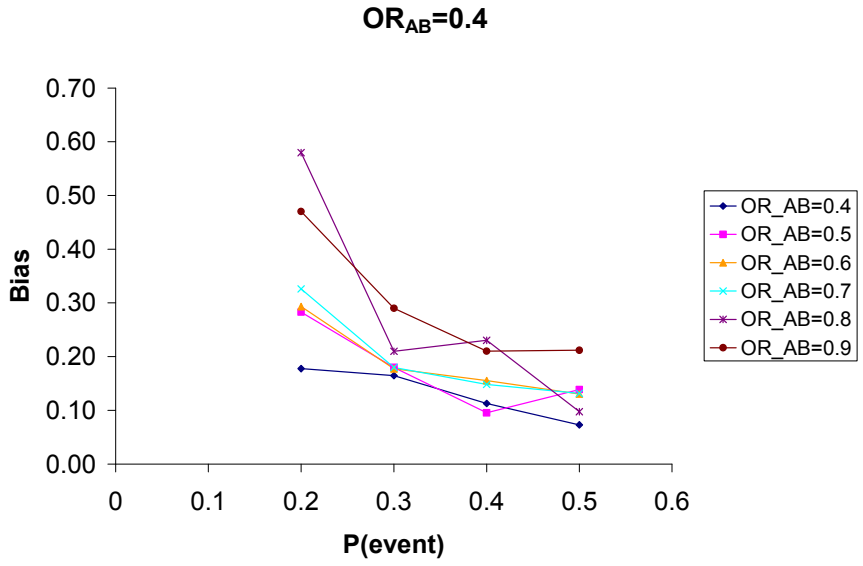
Patterns of bias and MSE for the indirect odds ratio (OR) estimates are displayed for different settings of OR_{AB}, OR_{CB}, and the event rates (Figure B.3.3 and Figure B.3.4). The patterns are shown for the event rate 0.05 to 0.5 and the results are symmetric about 0.5. As such, patterns for event rates ranging from 0.5 to 0.95 are not shown. For each of the settings, the figures are displayed for event rates that start at 0.05. The figures are also displayed for event rates that start at 0.2 in order to improve the resolution of the graphs.

Figure B.3.3: Bias for indirect odds ratio (OR) estimates for various parameter settings

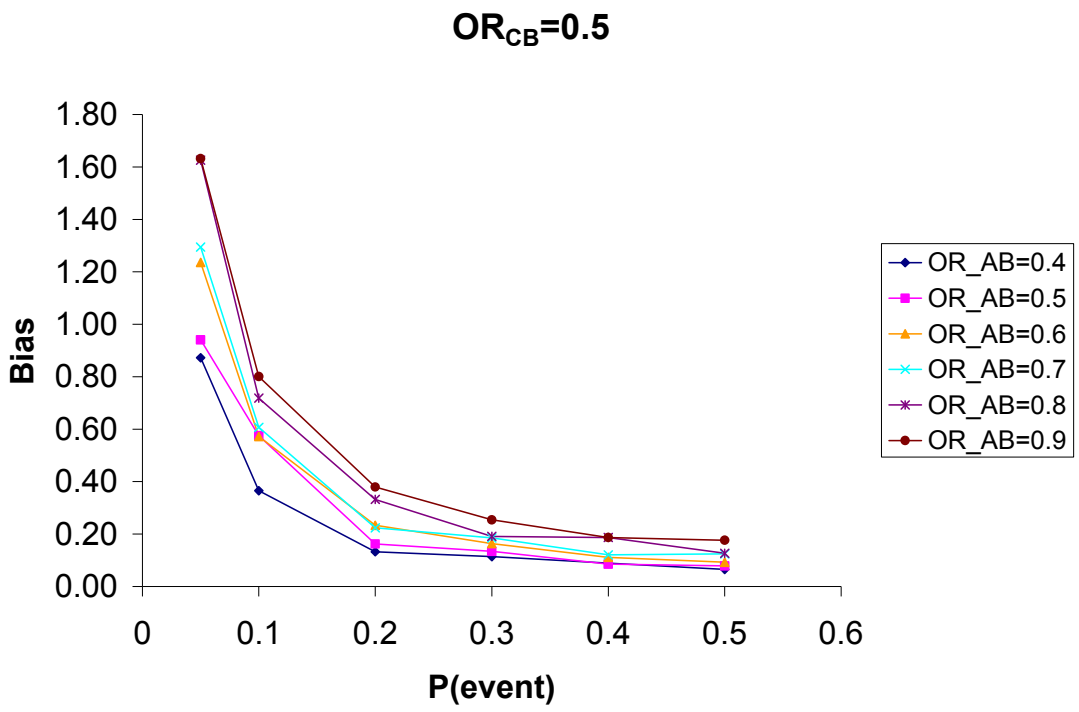
a)



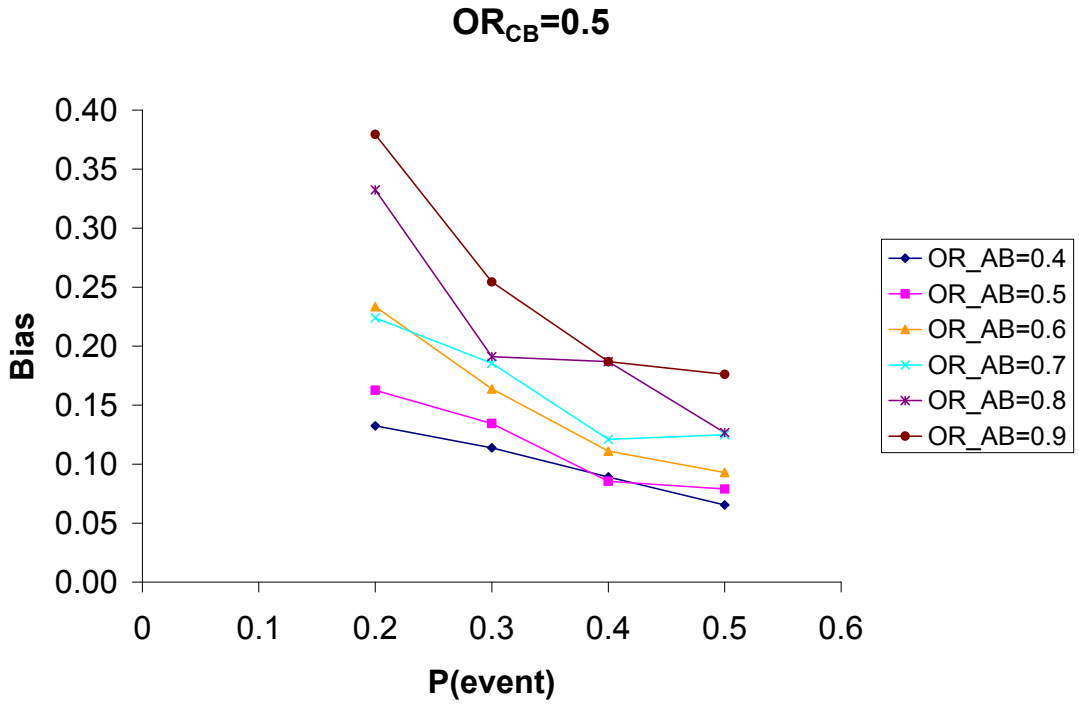
b)



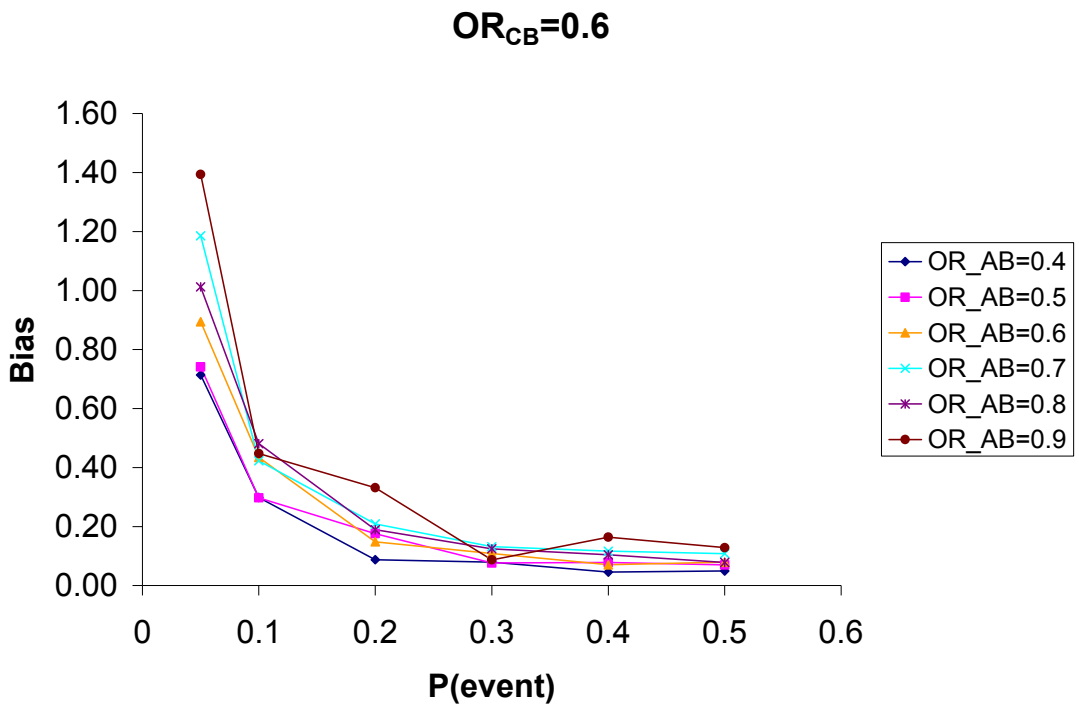
c)


OR_CB=0.5
Bias
P(event)
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

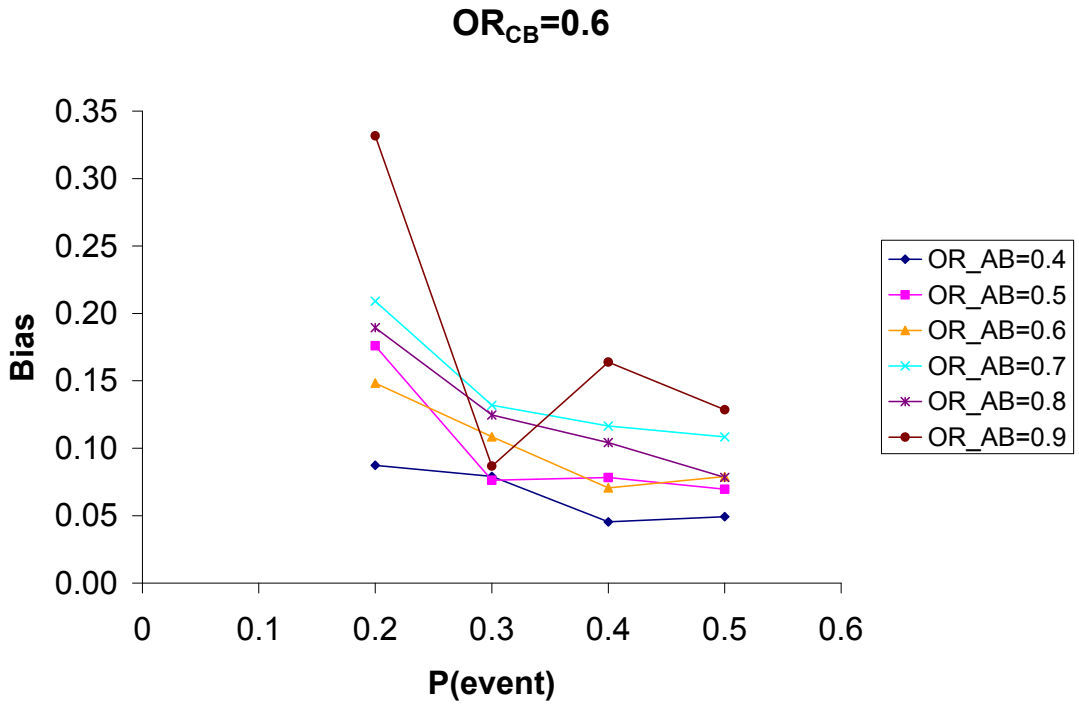
d)


OR_CB=0.5
Bias
P(event)
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

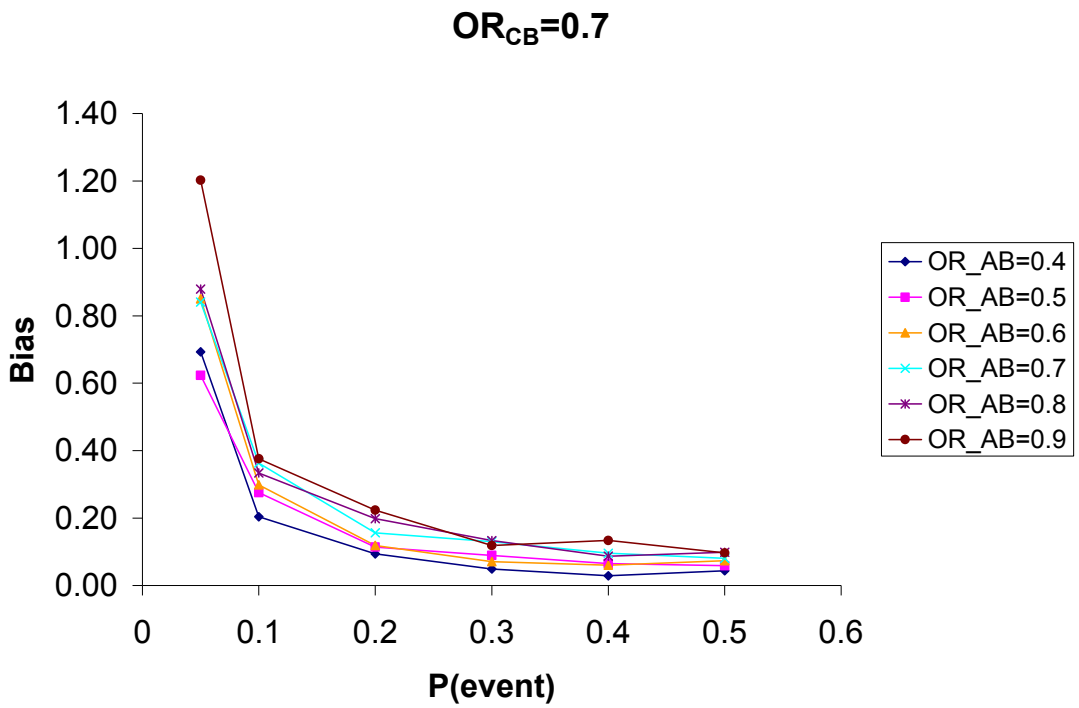
e)


OR_CB=0.6
Bias
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

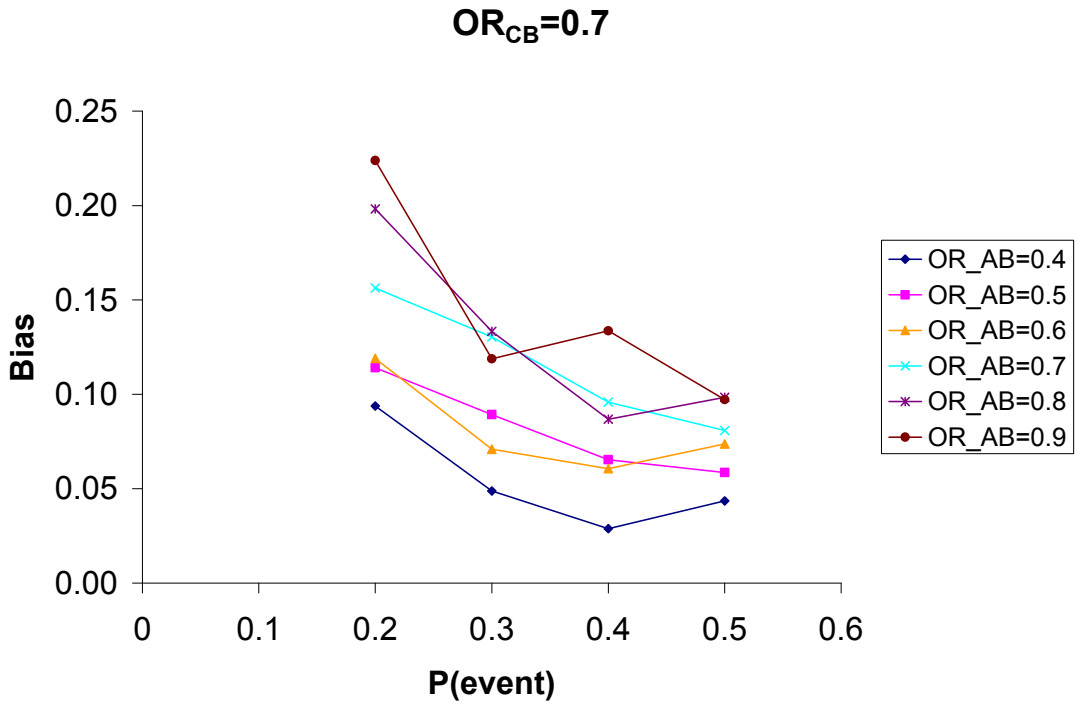
f)


OR_CB=0.6
Bias
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

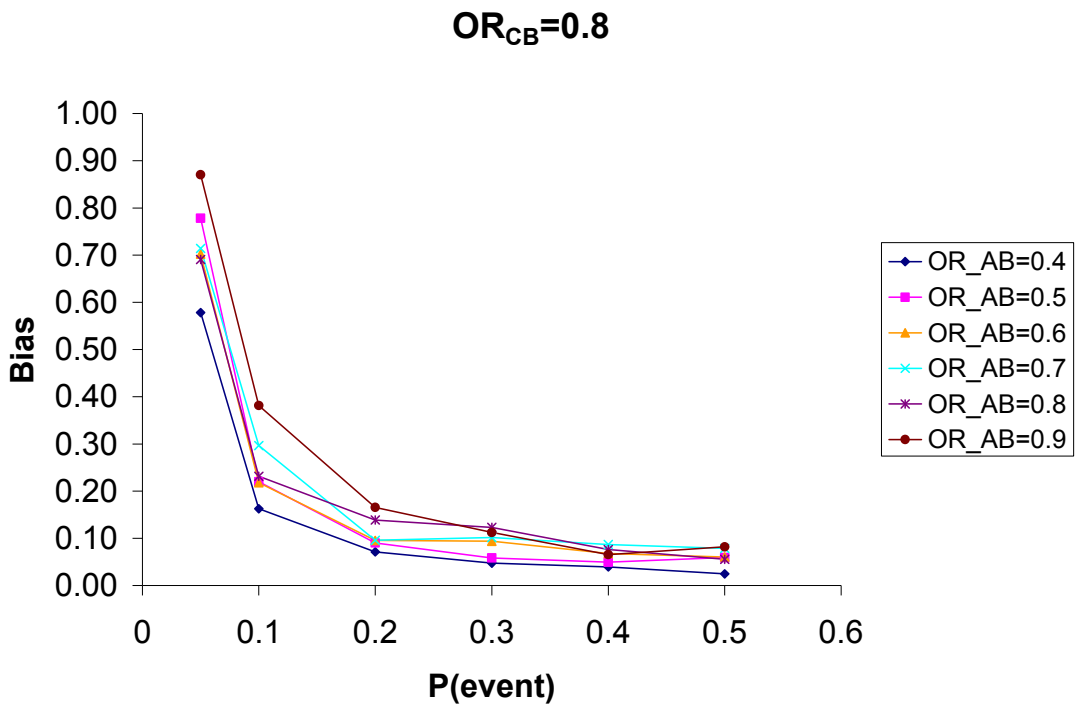
g)


OR_CB=0.7
Bias
P(event)
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

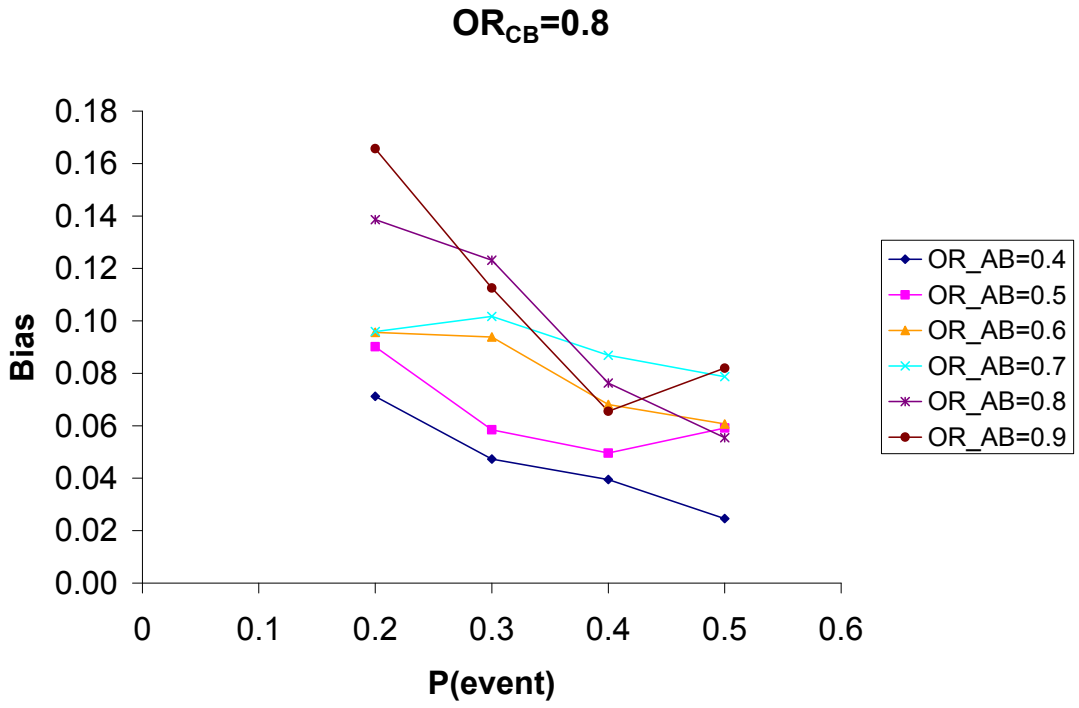
h)


OR_CB=0.7
Bias
P(event)
0.25
0.20
0.15
0.10
0.05
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

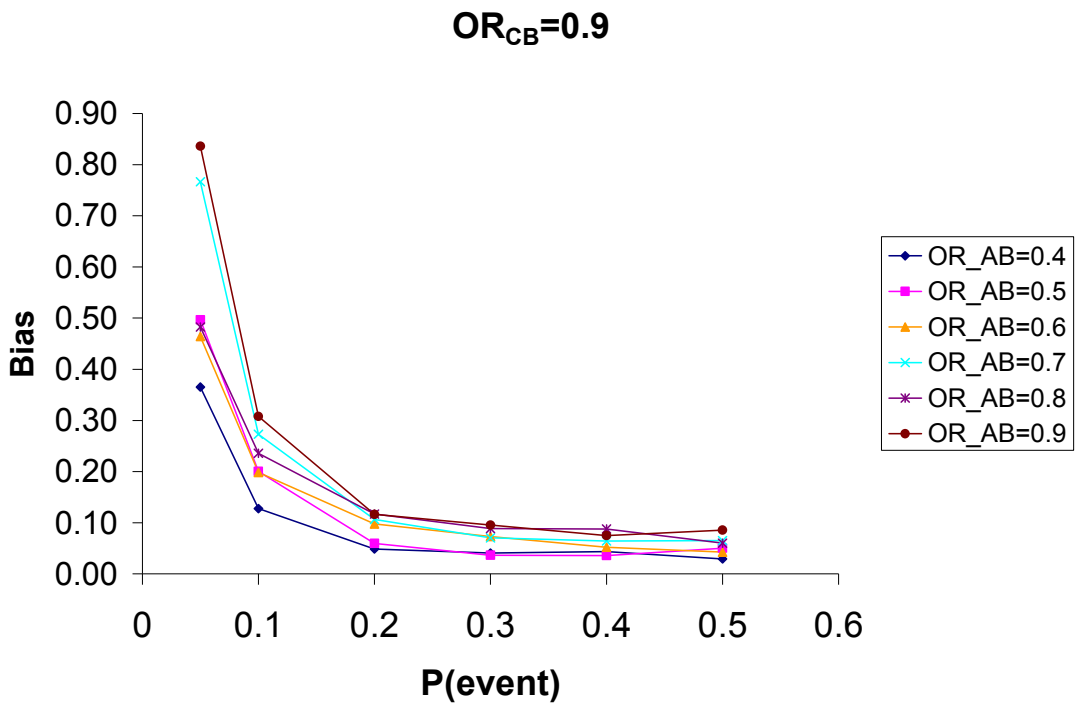
i)


OR_CB=0.8
Bias
P(event)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

j)


OR_CB=0.8
Bias
P(event)
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

k)


OR_CB=0.9
Bias
P(event)
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

l)

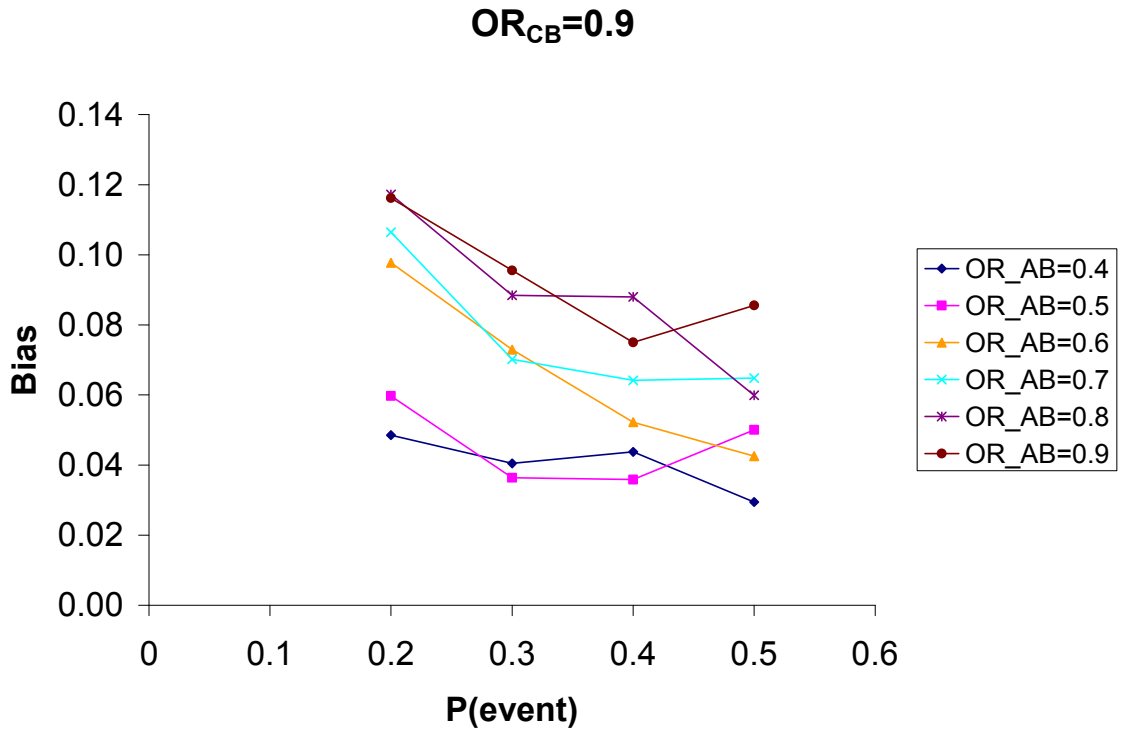
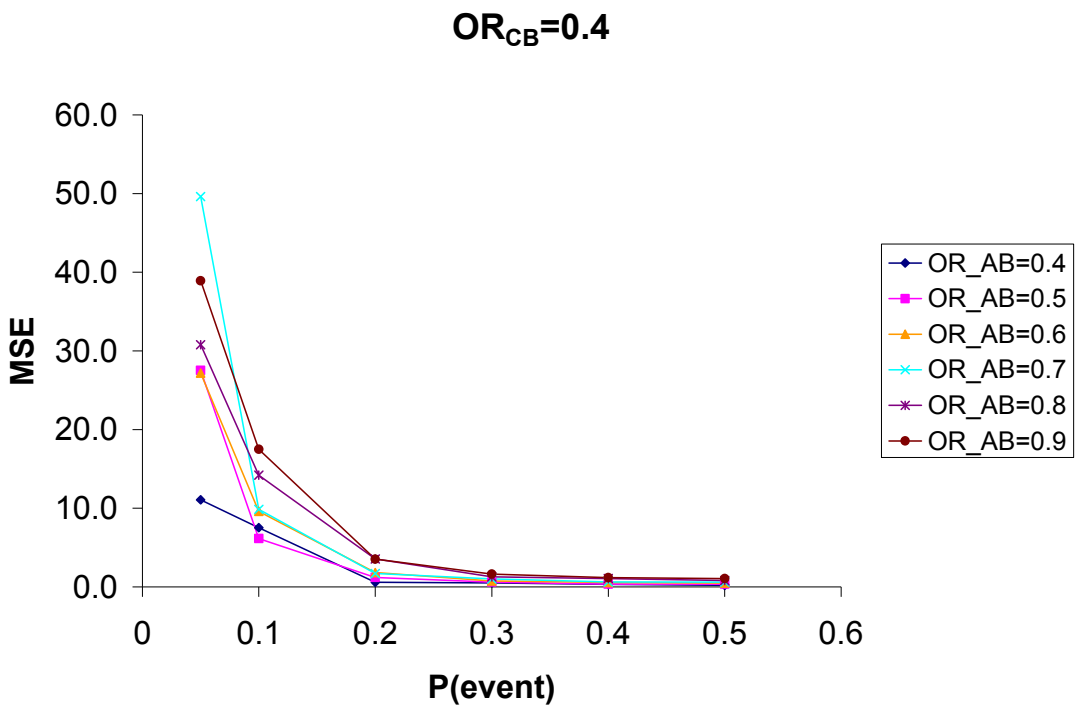
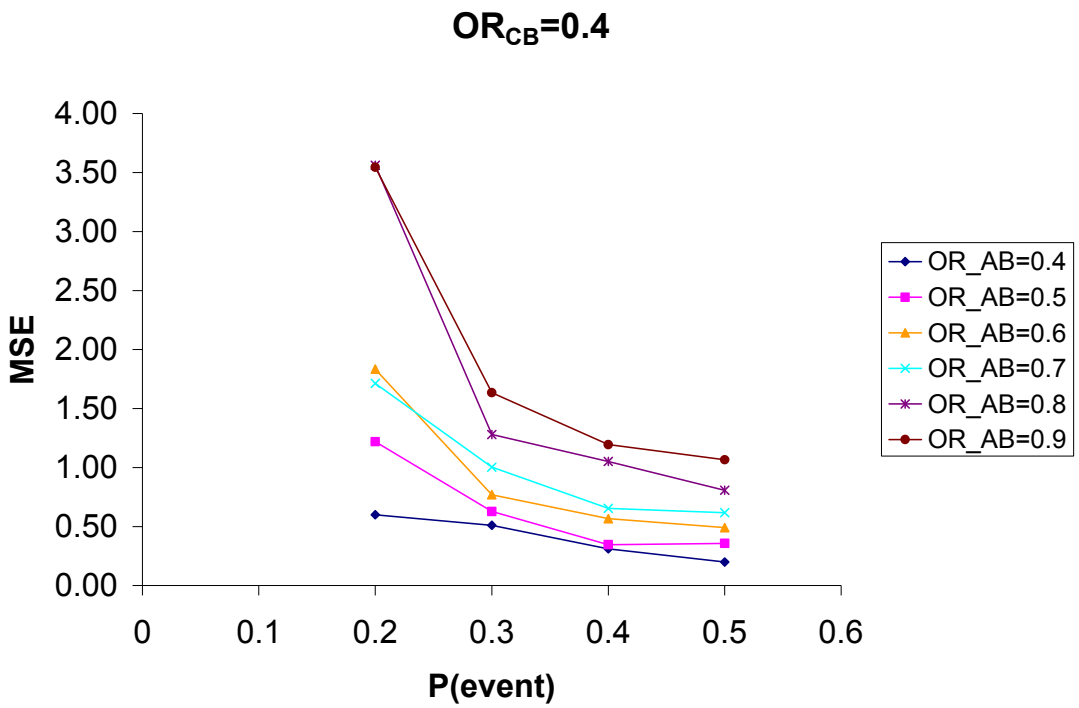

OR_CB=0.9
Bias
P(event)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

Figure B.3.4: Mean square error (MSE) for indirect odds ratio (OR) estimates for various parameter settings

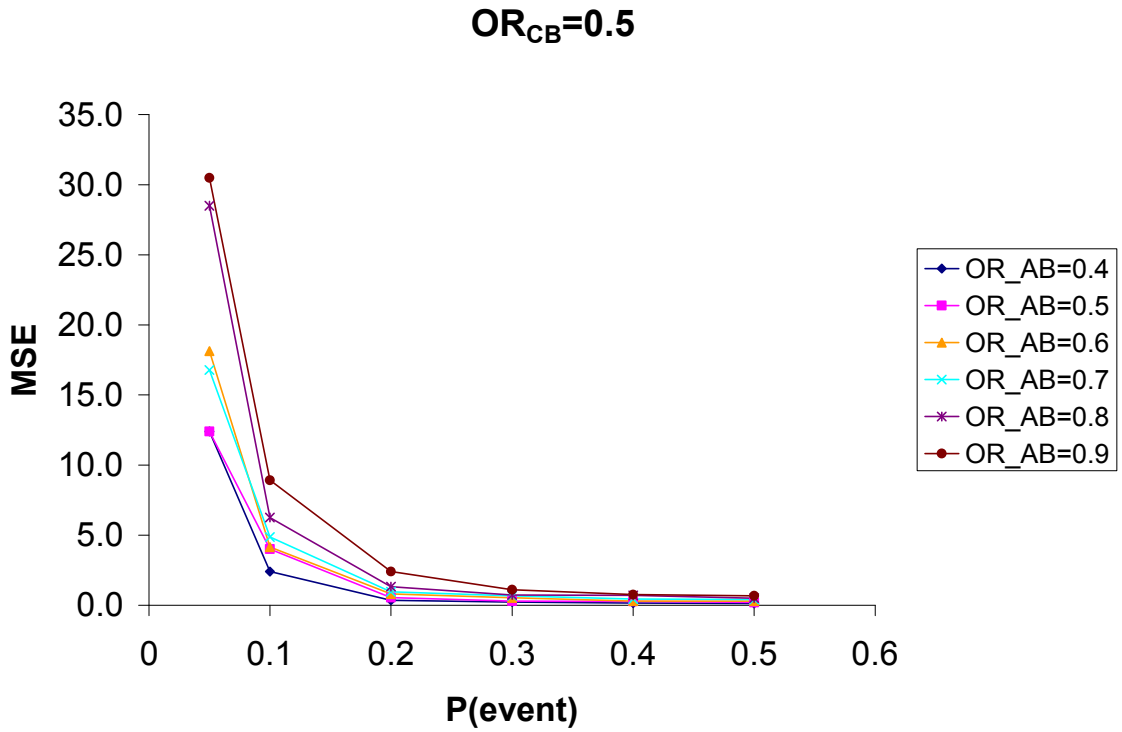
a)



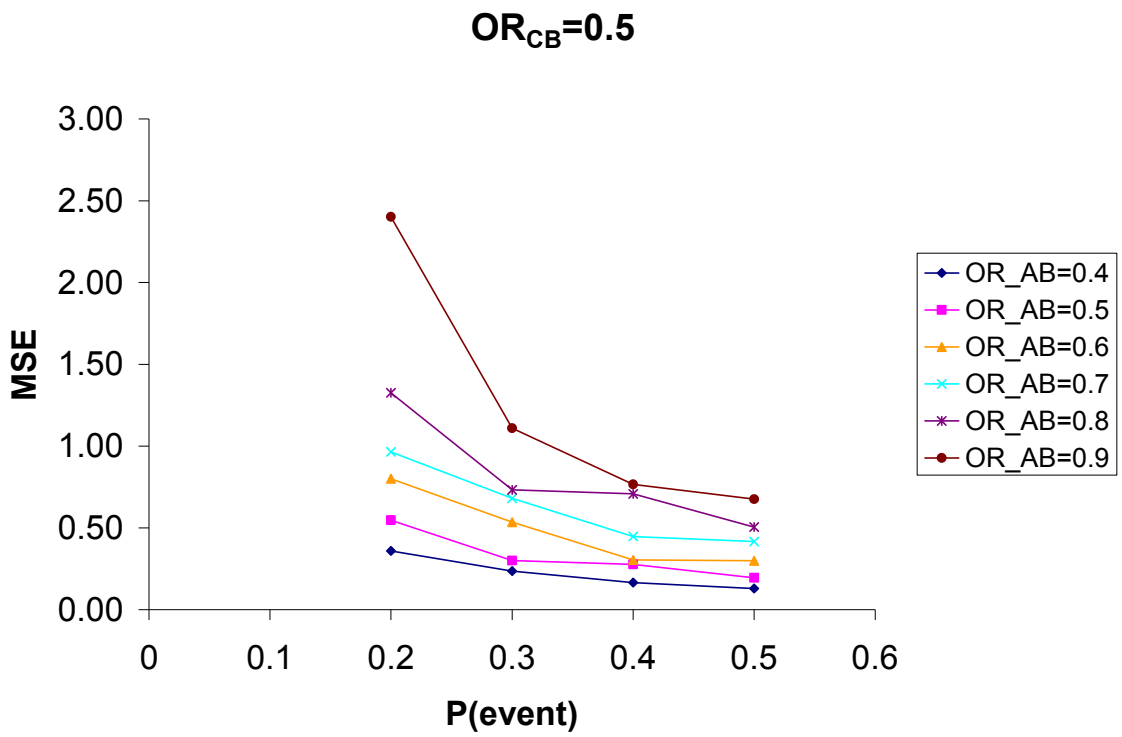
b)


OR_{CB}=0.4
MSE
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

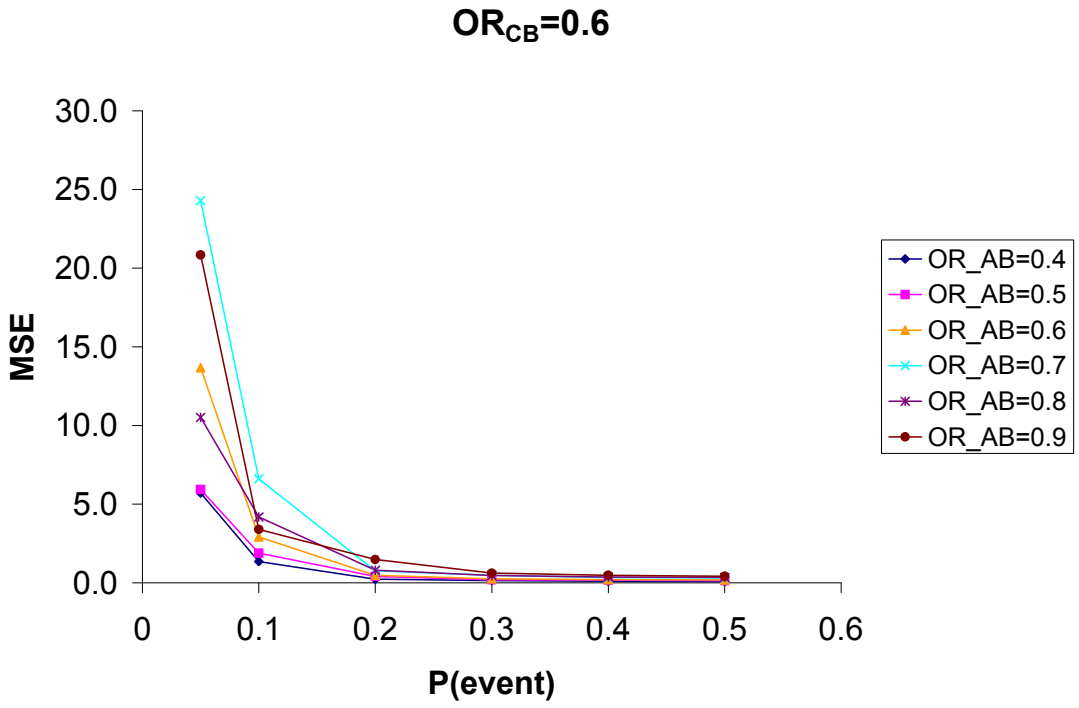
c)


OR_{CB}=0.5
MSE
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

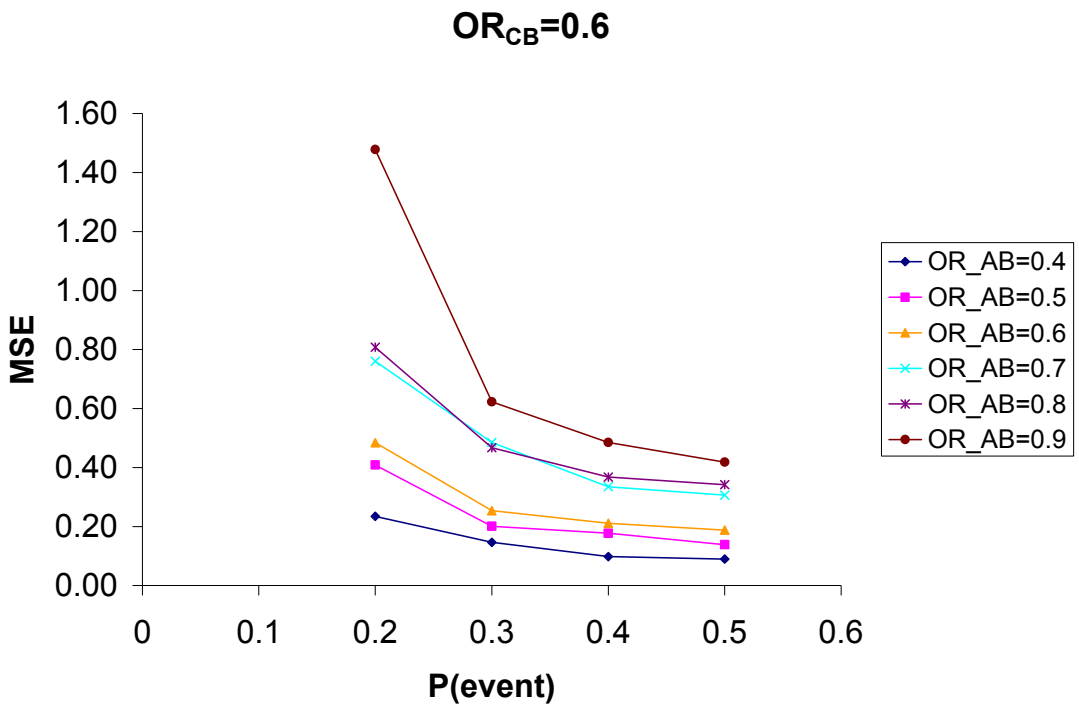
d)


OR_CB=0.5
MSE
3.00
2.50
2.00
1.50
1.00
0.50
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

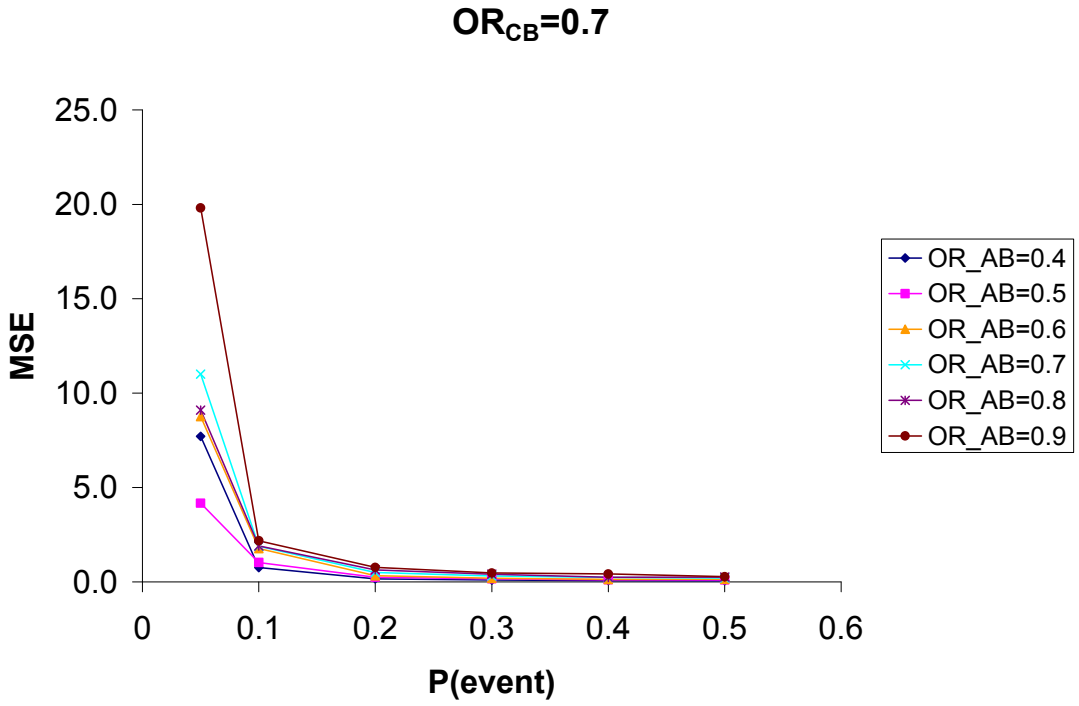
e)


OR_CB=0.6
MSE
30.0
25.0
20.0
15.0
10.0
5.0
0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

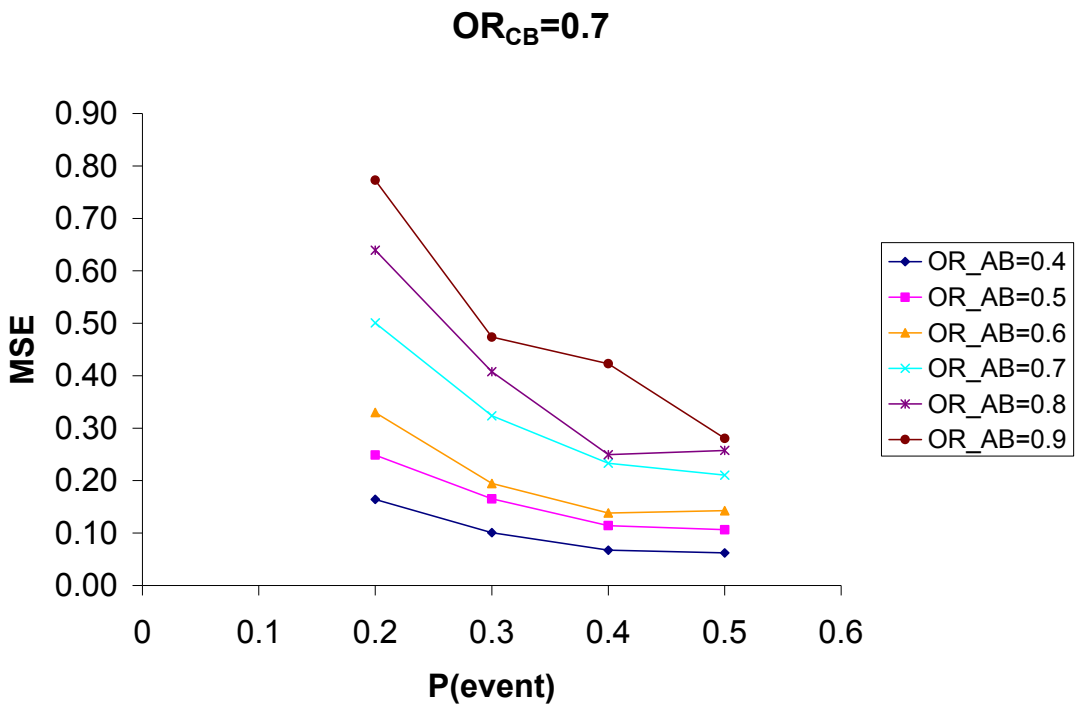
f)


OR_CB=0.6
MSE
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

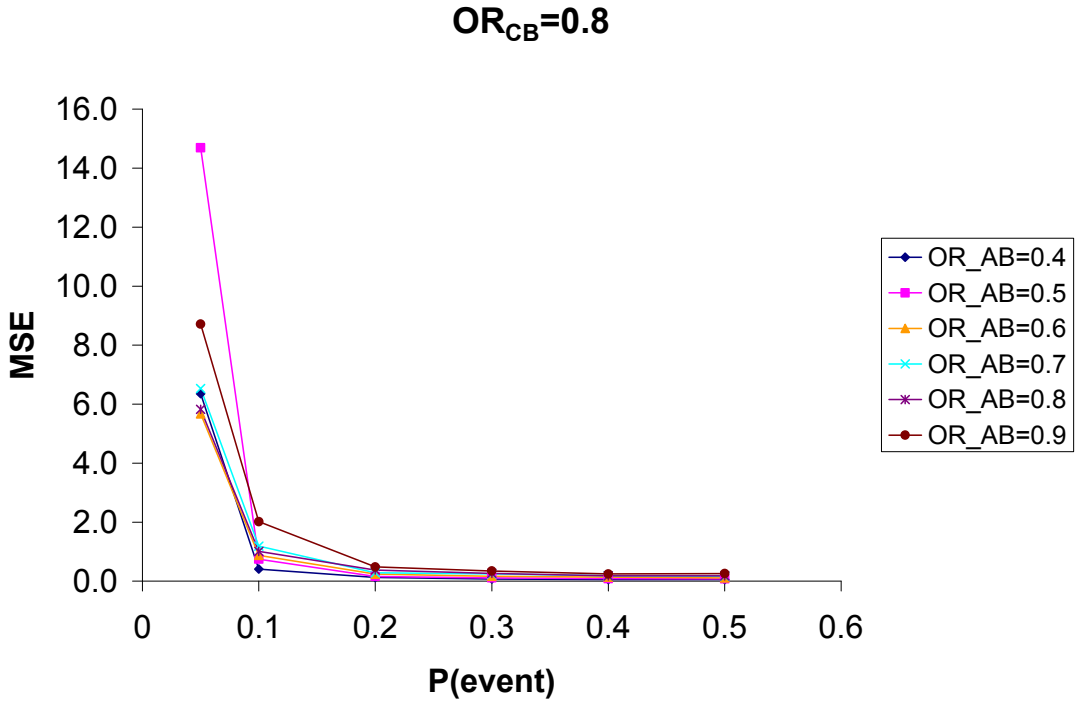
g)


OR_CB=0.7
MSE
25.0
20.0
15.0
10.0
5.0
0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

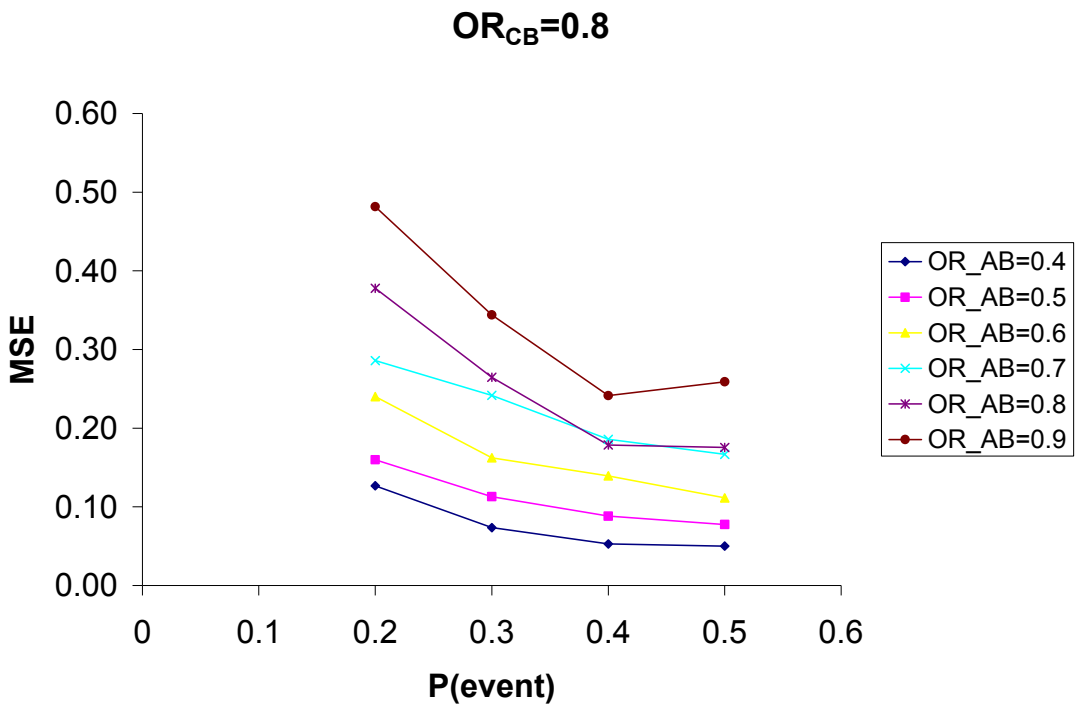
h)


OR_CB=0.7
MSE
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

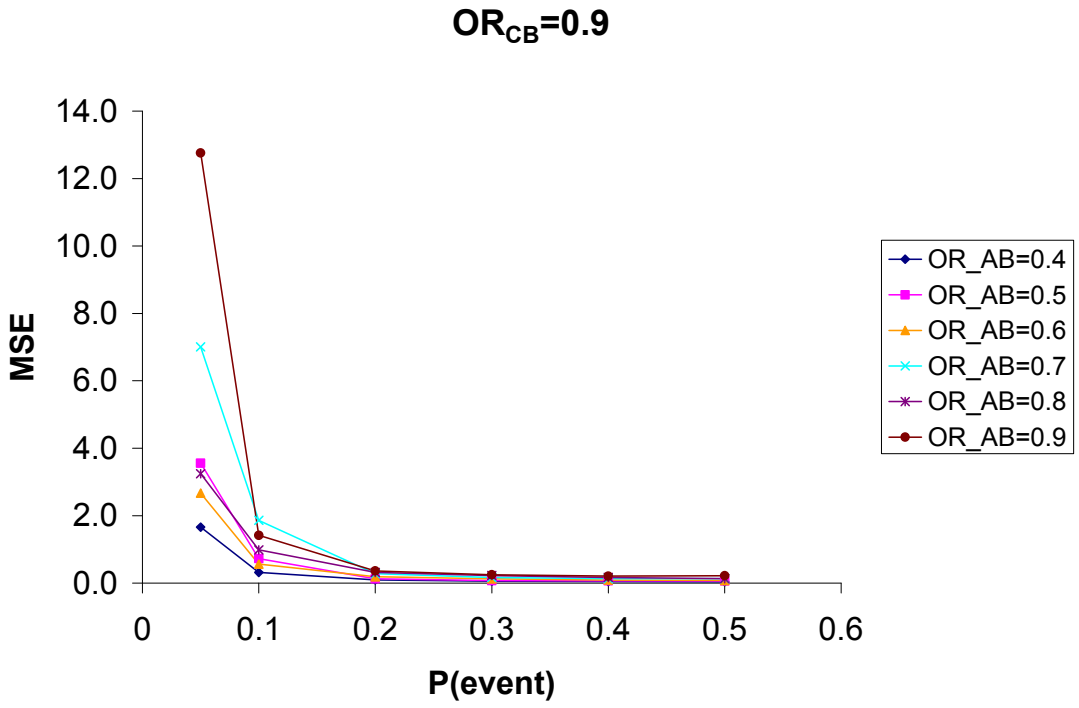
i)


OR_CB=0.8
MSE
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

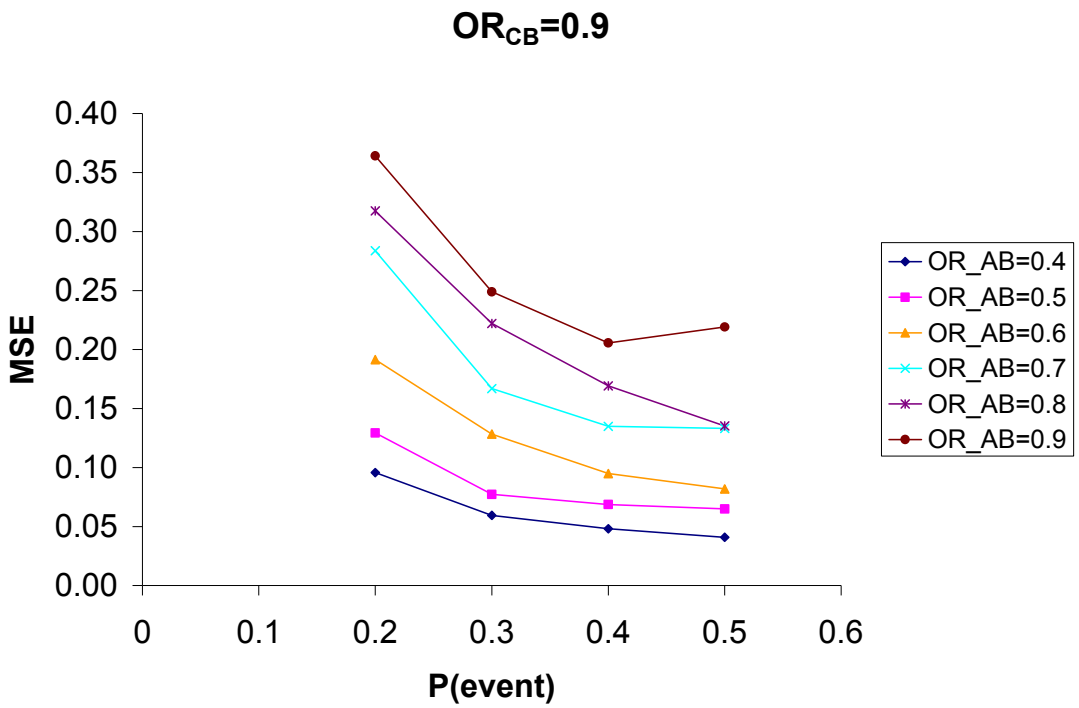
j)


OR_CB=0.8
MSE
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

k)


OR_CB=0.9
MSE
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
P(event)
OR_AB=0.4
OR_AB=0.5
OR_AB=0.6
OR_AB=0.7
OR_AB=0.8
OR_AB=0.9

l)



B.4 Simulation Results for the Risk Difference

For each of the three populations (A, B and C) in the Monte Carlo simulation for the risk difference, the outcome risk level for each population was selected according to a specific combination of the parameters RD_{AB}, RD_{CB}, and $P(E \mid B)$), as follows:

OR_{AB}: -0.1, -0.2, -0.3, -0.4, -0.5, -0.6
OR_{CB}: -0.1, -0.2, -0.3, -0.4, -0.5, -0.6
$P(E \mid B)$): 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9

For the various combinations of these parameters, the results of the simulation for the bias, variance and MSE for the direct and indirect estimators of the OR are provided in Table B.4.1.

As an illustration of the frequency distribution of the estimators, the frequency distributions for the parameter settings OR_{AB}=0.6, OR_{CB}=0.8 and $P(E \mid B)$=0.05,0.3,0.5 for the direct and indirect estimators are presented graphically in Figure B.3.1 on the logarithmic scale. It is apparent from these figures that both estimators have a mound shape, symmetric distribution (on the logarithmic scale). The indirect estimator has a larger variance and bias.

Table B.4.1: Bias, variance and mean square error (MSE) of direct and indirect risk difference (RD) estimators for different settings of the indicated parameters (k=3 treatments)

Risk difference Settings	Event Rate P(E | B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
RD_{AB}=-0.6	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.6	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0015	0.0019	0.0015	0.0055	-0.0035	0.0055
	0.8	0.0024	0.0001	0.0024	0.0052	-0.0019	0.0052
	0.9	0.0032	-0.0013	0.0032	0.0051	0.0023	0.0051
RD_{AB}=-0.6	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.5	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0019	-0.0024	0.0019	0.0055	-0.0007	0.0055
	0.8	0.0033	-0.0003	0.0033	0.0056	-0.0024	0.0056
	0.9	0.0036	0.0012	0.0036	0.0050	0.0016	0.0050
RD_{AB}=-0.6	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.4	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0023	-0.0008	0.0023	0.0062	-0.0005	0.0062
	0.8	0.0033	0.0013	0.0033	0.0063	-0.0032	0.0064
	0.9	0.0040	0.0018	0.0040	0.0052	0.0014	0.0052
RD_{AB}=-0.6	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.3	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0028	0.0022	0.0028	0.0059	-0.0028	0.0059
	0.8	0.0034	0.0001	0.0034	0.0063	-0.0002	0.0063
	0.9	0.0039	-0.0011	0.0039	0.0054	-0.0046	0.0054
RD_{AB}=-0.6	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.2	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0030	-0.0008	0.0030	0.0058	0.0031	0.0058
	0.8	0.0032	-0.0030	0.0032	0.0056	0.0025	0.0056
	0.9	0.0038	-0.0019	0.0038	0.0044	0.0006	0.0044
RD_{AB}=-0.6	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.1	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0028	0.0015	0.0028	0.0057	0.0010	0.0057
	0.8	0.0031	-0.0007	0.0031	0.0060	0.0022	0.0060
	0.9	0.0030	-0.0022	0.0030	0.0043	-0.0013	0.0043
RD_{AB}=-0.5	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.6	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-

Risk difference Settings	Event Rate P(E | B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0020	-0.0003	0.0020	0.0056	-0.0043	0.0056
	0.8	0.0025	0.0016	0.0025	0.0058	0.0004	0.0058
	0.9	0.0034	0.0011	0.0034	0.0054	-0.0004	0.0054
RD_{AB}=-0.5	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.5	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	0.0011	-0.0003	0.0011	0.0056	-0.0010	0.0056
	0.7	0.0020	0.0019	0.0020	0.0061	0.0029	0.0061
	0.8	0.0028	0.0021	0.0028	0.0063	-0.0024	0.0063
	0.9	0.0031	-0.0010	0.0031	0.0057	-0.0022	0.0057
RD_{AB}=-0.5	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.4	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	0.0018	0.0037	0.0018	0.0058	-0.0051	0.0059
	0.7	0.0025	0.0003	0.0025	0.0057	0.0013	0.0057
	0.8	0.0033	-0.0004	0.0033	0.0064	-0.0027	0.0064
	0.9	0.0034	0.0023	0.0034	0.0055	-0.0029	0.0056
RD_{AB}=-0.5	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.3	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	0.0025	-0.0024	0.0025	0.0057	0.0006	0.0057
	0.7	0.0029	-0.0002	0.0029	0.0060	-0.0015	0.0060
	0.8	0.0034	-0.0020	0.0034	0.0062	-0.0034	0.0062
	0.9	0.0034	0.0004	0.0034	0.0054	0.0006	0.0054
RD_{AB}=-0.5	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.2	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	0.0025	0.0013	0.0025	0.0067	-0.0017	0.0067
	0.7	0.0032	-0.0002	0.0032	0.0063	-0.0053	0.0063
	0.8	0.0034	-0.0010	0.0034	0.0056	0.0051	0.0056
	0.9	0.0032	-0.0006	0.0032	0.0049	-0.0007	0.0049
RD_{AB}=-0.5	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.1	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	0.0025	0.0021	0.0025	0.0057	0.0042	0.0057
	0.7	0.0031	-0.0001	0.0031	0.0058	0.0022	0.0058
	0.8	0.0031	0.0007	0.0031	0.0048	-0.0030	0.0048
	0.9	0.0031	-0.0005	0.0031	0.0044	0.0010	0.0044
RD_{AB}=-0.4	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.6	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0022	0.0022	0.0022	0.0055	-0.0010	0.0055

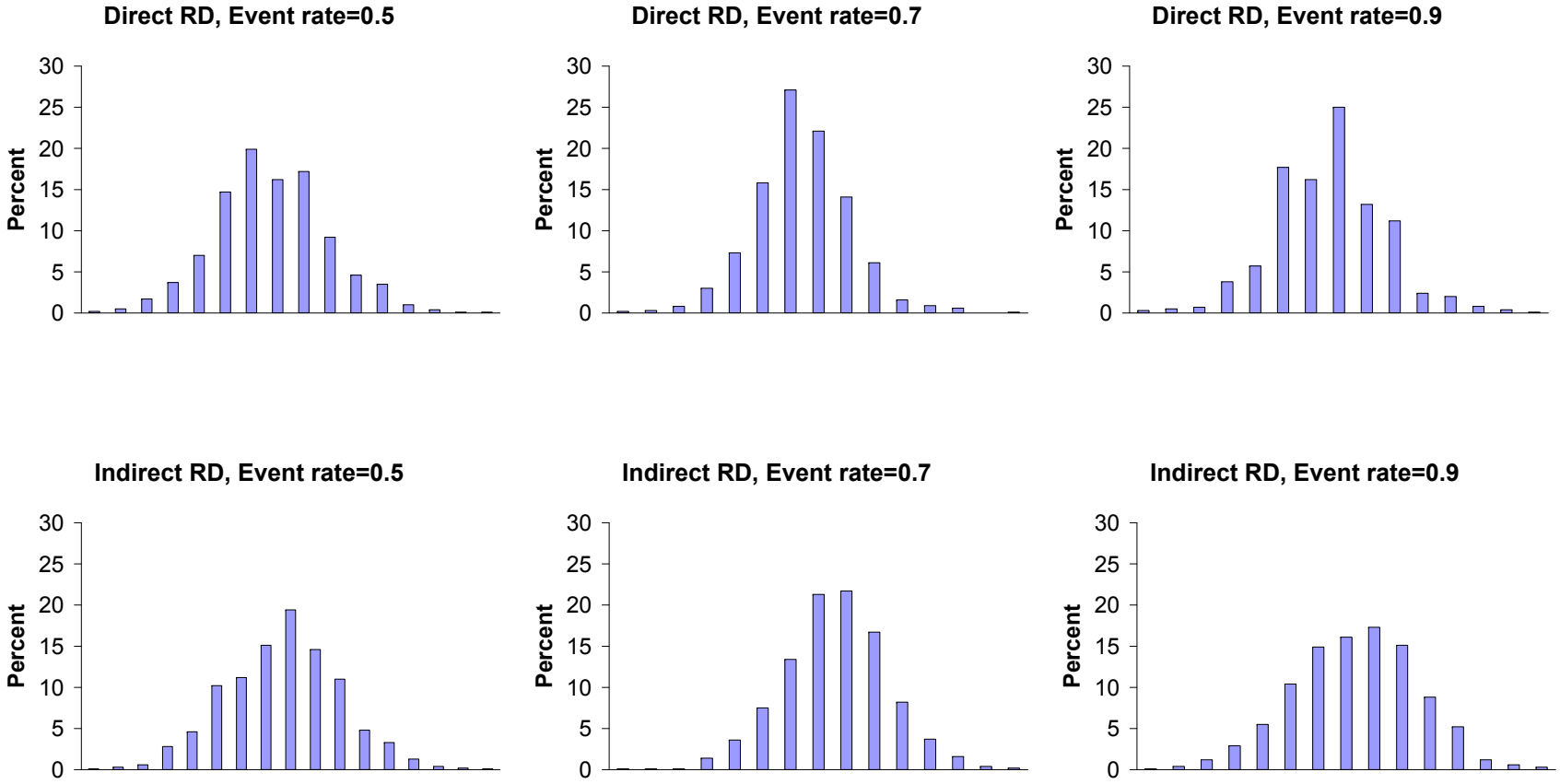
Risk difference Settings	Event Rate P(E | B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.8	0.0029	-0.0004	0.0029	0.0055	-0.0014	0.0055
	0.9	0.0031	-0.0005	0.0031	0.0056	0.0003	0.0056
RD_{AB}=-0.4	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.5	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	0.0016	-0.0008	0.0016	0.0058	0.0000	0.0058
	0.7	0.0022	-0.0013	0.0022	0.0062	0.0024	0.0062
	0.8	0.0029	-0.0009	0.0029	0.0063	-0.0048	0.0063
	0.9	0.0030	-0.0027	0.0030	0.0056	-0.0009	0.0056
RD_{AB}=-0.4	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.4	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	0.0009	0.0012	0.0009	0.0052	-0.0020	0.0052
	0.6	0.0015	-0.0006	0.0015	0.0058	0.0006	0.0058
	0.7	0.0021	-0.0014	0.0021	0.0060	-0.0011	0.0060
	0.8	0.0027	-0.0027	0.0027	0.0062	-0.0010	0.0062
	0.9	0.0027	-0.0006	0.0027	0.0048	-0.0031	0.0049
RD_{AB}=-0.4	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.3	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	0.0016	-0.0021	0.0016	0.0056	-0.0004	0.0056
	0.6	0.0023	0.0029	0.0024	0.0056	0.0024	0.0056
	0.7	0.0026	-0.0025	0.0026	0.0058	0.0032	0.0058
	0.8	0.0027	-0.0007	0.0027	0.0054	0.0006	0.0054
	0.9	0.0027	-0.0026	0.0027	0.0053	0.0031	0.0053
RD_{AB}=-0.4	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.2	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	0.0022	-0.0014	0.0022	0.0053	-0.0026	0.0053
	0.6	0.0024	0.0013	0.0024	0.0059	-0.0012	0.0059
	0.7	0.0030	-0.0021	0.0030	0.0055	-0.0014	0.0055
	0.8	0.0032	0.0011	0.0032	0.0056	0.0008	0.0056
	0.9	0.0027	0.0010	0.0027	0.0046	0.0000	0.0046
RD_{AB}=-0.4	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.1	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	0.0025	0.0011	0.0025	0.0053	-0.0013	0.0053
	0.6	0.0029	-0.0027	0.0030	0.0054	0.0021	0.0054
	0.7	0.0029	-0.0020	0.0030	0.0053	0.0009	0.0053
	0.8	0.0031	-0.0006	0.0031	0.0048	-0.0028	0.0048
	0.9	0.0030	-0.0043	0.0030	0.0041	-0.0012	0.0041
RD_{AB}=-0.3	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.6	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0024	0.0009	0.0024	0.0055	0.0038	0.0055
	0.8	0.0028	-0.0022	0.0028	0.0050	-0.0017	0.0050
	0.9	0.0028	0.0011	0.0028	0.0047	-0.0009	0.0047
RD_{AB}=-0.3	**0.2**	-	-	-	-	-	-

Risk difference Settings	Event Rate P(E | B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
RD_{CB}=-0.5	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	0.0021	0.0004	0.0021	0.0053	0.0027	0.0053
	0.7	0.0027	-0.0019	0.0027	0.0054	0.0049	0.0054
	0.8	0.0028	-0.0015	0.0028	0.0056	0.0015	0.0056
	0.9	0.0026	-0.0020	0.0026	0.0050	-0.0003	0.0050
RD_{AB}=-0.3	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.4	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	0.0016	0.0005	0.0016	0.0053	-0.0024	0.0053
	0.6	0.0021	-0.0022	0.0021	0.0057	-0.0017	0.0057
	0.7	0.0023	-0.0014	0.0023	0.0058	0.0046	0.0059
	0.8	0.0025	0.0034	0.0025	0.0055	0.0009	0.0055
	0.9	0.0024	-0.0007	0.0024	0.0053	-0.0003	0.0053
RD_{AB}=-0.3	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.3	**0.3**	-	-	-	-	-	-
	0.4	0.0007	0.0002	0.0007	0.0049	-0.0020	0.0049
	0.5	0.0013	-0.0006	0.0013	0.0051	0.0009	0.0051
	0.6	0.0019	-0.0018	0.0019	0.0057	0.0054	0.0058
	0.7	0.0022	0.0006	0.0022	0.0057	-0.0003	0.0057
	0.8	0.0019	-0.0012	0.0019	0.0053	0.0033	0.0054
	0.9	0.0018	0.0010	0.0018	0.0047	0.0051	0.0047
RD_{AB}=-0.3	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.2	**0.3**	-	-	-	-	-	-
	0.4	0.0016	0.0003	0.0016	0.0045	0.0008	0.0045
	0.5	0.0020	0.0006	0.0020	0.0054	-0.0016	0.0054
	0.6	0.0026	0.0004	0.0026	0.0056	0.0027	0.0056
	0.7	0.0023	0.0027	0.0023	0.0052	0.0012	0.0052
	0.8	0.0026	0.0017	0.0026	0.0049	0.0002	0.0049
	0.9	0.0027	-0.0022	0.0027	0.0046	-0.0014	0.0046
RD_{AB}=-0.3	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.1	**0.3**	-	-	-	-	-	-
	0.4	0.0020	0.0014	0.0020	0.0043	-0.0002	0.0043
	0.5	0.0025	0.0015	0.0025	0.0050	-0.0001	0.0050
	0.6	0.0028	0.0009	0.0028	0.0052	-0.0036	0.0052
	0.7	0.0027	-0.0008	0.0027	0.0047	-0.0015	0.0047
	0.8	0.0026	0.0044	0.0026	0.0043	-0.0025	0.0043
	0.9	0.0025	-0.0011	0.0025	0.0039	-0.0004	0.0039
RD_{AB}=-0.2	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.6	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0026	-0.0003	0.0026	0.0048	0.0015	0.0048
	0.8	0.0027	-0.0003	0.0027	0.0045	0.0007	0.0045
	0.9	0.0027	0.0017	0.0027	0.0042	0.0009	0.0042
RD_{AB}=-0.2	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.5	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-

Risk difference Settings	Event Rate P(E | B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.6	0.0020	-0.0002	0.0020	0.0053	0.0019	0.0053
	0.7	0.0027	0.0038	0.0027	0.0051	0.0009	0.0051
	0.8	0.0027	0.0033	0.0027	0.0051	0.0006	0.0051
	0.9	0.0027	0.0022	0.0027	0.0044	-0.0018	0.0044
RD_{AB}=-0.2	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.4	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	0.0019	0.0001	0.0019	0.0049	0.0005	0.0049
	0.6	0.0025	0.0010	0.0025	0.0052	-0.0035	0.0052
	0.7	0.0026	0.0019	0.0026	0.0056	-0.0009	0.0056
	0.8	0.0024	-0.0013	0.0024	0.0050	-0.0001	0.0050
	0.9	0.0024	0.0013	0.0024	0.0044	0.0005	0.0044
RD_{AB}=-0.2	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.3	**0.3**	-	-	-	-	-	-
	0.4	0.0013	0.0000	0.0013	0.0039	-0.0013	0.0039
	0.5	0.0019	0.0009	0.0019	0.0052	0.0021	0.0052
	0.6	0.0019	-0.0017	0.0019	0.0048	-0.0020	0.0048
	0.7	0.0025	0.0015	0.0025	0.0054	-0.0006	0.0054
	0.8	0.0025	-0.0008	0.0025	0.0050	0.0028	0.0050
	0.9	0.0019	-0.0022	0.0019	0.0040	-0.0068	0.0040
RD_{AB}=-0.2	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.2	**0.3**	0.0006	-0.0001	0.0006	0.0038	-0.0010	0.0038
	0.4	0.0012	0.0007	0.0012	0.0044	-0.0021	0.0044
	0.5	0.0014	0.0010	0.0014	0.0050	-0.0015	0.0050
	0.6	0.0016	0.0013	0.0016	0.0052	-0.0031	0.0052
	0.7	0.0016	-0.0009	0.0016	0.0047	-0.0012	0.0047
	0.8	0.0018	-0.0030	0.0018	0.0043	-0.0034	0.0043
	0.9	0.0015	-0.0027	0.0015	0.0040	-0.0006	0.0040
RD_{AB}=-0.2	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.1	**0.3**	0.0013	-0.0010	0.0013	0.0036	0.0019	0.0036
	0.4	0.0016	0.0006	0.0016	0.0040	0.0020	0.0040
	0.5	0.0020	-0.0025	0.0020	0.0044	-0.0002	0.0044
	0.6	0.0021	0.0017	0.0021	0.0043	0.0029	0.0044
	0.7	0.0020	-0.0014	0.0020	0.0045	0.0010	0.0045
	0.8	0.0021	0.0003	0.0021	0.0040	0.0007	0.0040
	0.9	0.0016	0.0005	0.0016	0.0031	0.0029	0.0031
RD_{AB}=-0.1	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.6	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	-	-	-	-	-	-
	0.7	0.0026	-0.0024	0.0026	0.0040	0.0003	0.0040
	0.8	0.0025	-0.0010	0.0025	0.0041	0.0013	0.0041
	0.9	0.0025	0.0002	0.0025	0.0038	-0.0003	0.0038
RD_{AB}=-0.1	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.5	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	-	-	-	-	-	-
	0.6	0.0025	0.0001	0.0025	0.0044	0.0003	0.0044
	0.7	0.0028	-0.0009	0.0028	0.0047	0.0023	0.0047
	0.8	0.0028	0.0002	0.0028	0.0047	-0.0006	0.0047

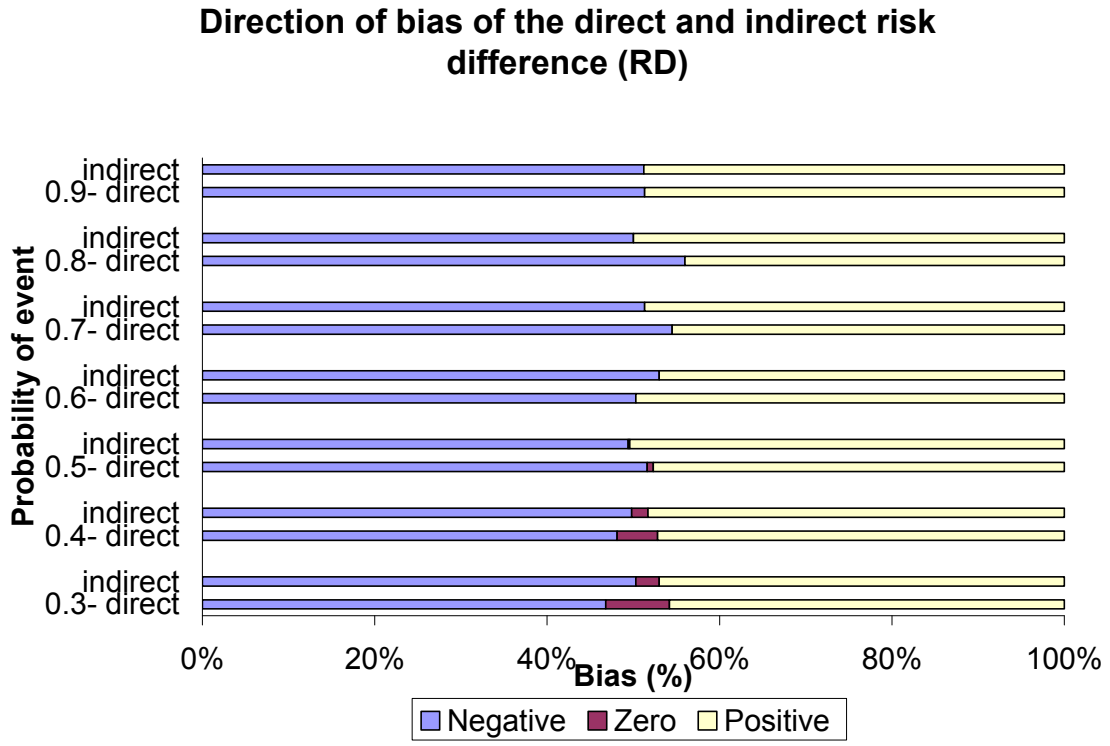
Risk difference Settings	Event Rate P(E | B)	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
	0.9	0.0027	0.0031	0.0027	0.0038	-0.0032	0.0038
RD_{AB}=-0.1	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.4	**0.3**	-	-	-	-	-	-
	0.4	-	-	-	-	-	-
	0.5	0.0022	0.0028	0.0023	0.0048	0.0002	0.0048
	0.6	0.0024	-0.0005	0.0024	0.0048	-0.0007	0.0048
	0.7	0.0025	0.0004	0.0025	0.0045	0.0005	0.0045
	0.8	0.0027	0.0004	0.0027	0.0043	-0.0045	0.0043
	0.9	0.0026	-0.0020	0.0026	0.0037	-0.0017	0.0037
RD_{AB}=-0.1	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.3	**0.3**	-	-	-	-	-	-
	0.4	0.0018	0.0007	0.0018	0.0043	0.0011	0.0043
	0.5	0.0021	-0.0033	0.0021	0.0048	-0.0024	0.0048
	0.6	0.0025	-0.0009	0.0025	0.0043	-0.0027	0.0043
	0.7	0.0025	-0.0006	0.0025	0.0048	-0.0018	0.0048
	0.8	0.0024	0.0000	0.0024	0.0042	0.0012	0.0042
	0.9	0.0021	-0.0020	0.0022	0.0041	-0.0014	0.0041
RD_{AB}=-0.1	**0.2**	-	-	-	-	-	-
RD_{CB}=-0.2	**0.3**	0.0013	0.0021	0.0013	0.0035	-0.0010	0.0035
	0.4	0.0017	0.0010	0.0017	0.0042	-0.0004	0.0042
	0.5	0.0019	0.0003	0.0019	0.0049	0.0031	0.0050
	0.6	0.0020	-0.0006	0.0020	0.0045	-0.0026	0.0045
	0.7	0.0020	-0.0005	0.0020	0.0038	0.0011	0.0038
	0.8	0.0020	-0.0013	0.0020	0.0035	0.0033	0.0035
	0.9	0.0016	0.0023	0.0016	0.0028	-0.0001	0.0028
RD_{AB}=-0.1	**0.2**	0.0005	0.0007	0.0005	0.0028	0.0007	0.0028
RD_{CB}=-0.1	**0.3**	0.0009	0.0001	0.0009	0.0031	0.0009	0.0031
	0.4	0.0013	0.0009	0.0013	0.0038	0.0016	0.0038
	0.5	0.0014	0.0007	0.0014	0.0040	0.0026	0.0040
	0.6	0.0016	-0.0004	0.0016	0.0039	-0.0025	0.0039
	0.7	0.0015	-0.0005	0.0015	0.0035	0.0018	0.0035
	0.8	0.0012	-0.0021	0.0012	0.0033	-0.0013	0.0033
	0.9	0.0010	-0.0012	0.0010	0.0023	-0.0001	0.0023

Figure B.4.1: Frequency distribution of direct and indirect risk difference (RD) estimators for the parameter settings RD_{AB}=-0.2, RD_{CB}=-0.4, P(E | B)=0.5, 0.7, 0.9 (k=3 treatments)



For the parameter settings RD_{AB}=-0.1, RD_{CB}=-0.2, the direction of the biases over the 1000 samples are illustrated in Figure B.4.2 for the direct and indirect approaches under the different settings of the event rate P(E | B)): 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9. As noted, in most instances the direct and indirect estimators equally overestimate and underestimate the bias.

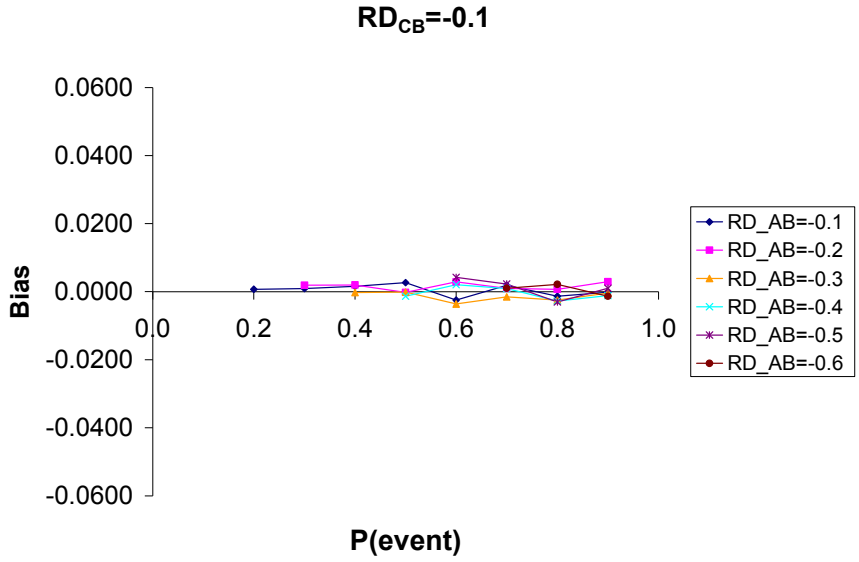
Figure B.4.2: Direction of the bias of the direct and indirect risk difference (RD) estimates for the parameter settings RD_{AB}=-0.1, RD_{CB}=-0.2, P(E | B) =0.3, 0.4, 0.5, 0.6, 0.7, 0.8 (k=3 treatments)



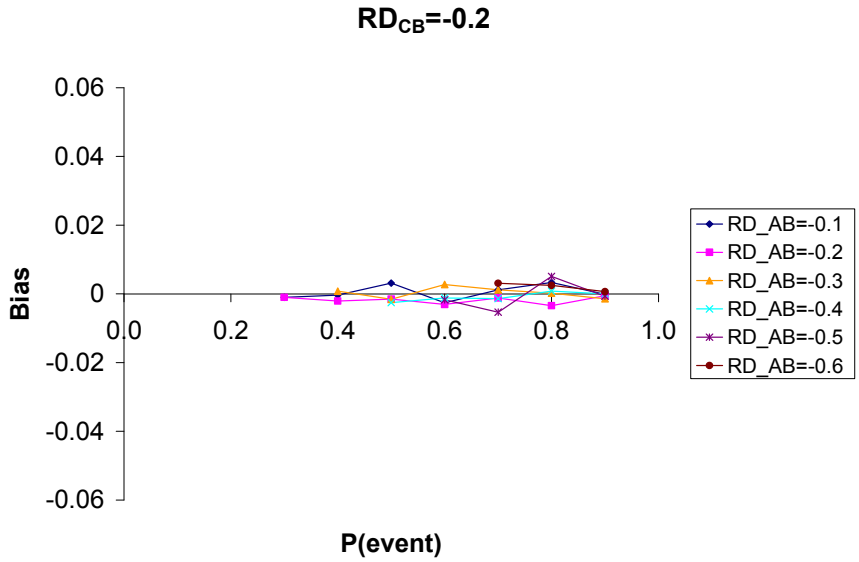
Patterns of bias and MSE for the indirect risk difference (RD) estimates are displayed for different settings of RD_{AB}, RD_{CB}, and the event rates (Figure B.4.3 and Figure B.4.4). The patterns are shown for the event rate 0.2 to 0.9. For the graphs displaying the patterns of bias for the indirect RD estimates, the upper and lower bound for the y-axis scale was selected by determining the value that corresponds to $1/10^{th}$ of the largest RD_{AB}.

Figure B.4.3: Bias for indirect risk difference (RD) estimates for various parameter settings.

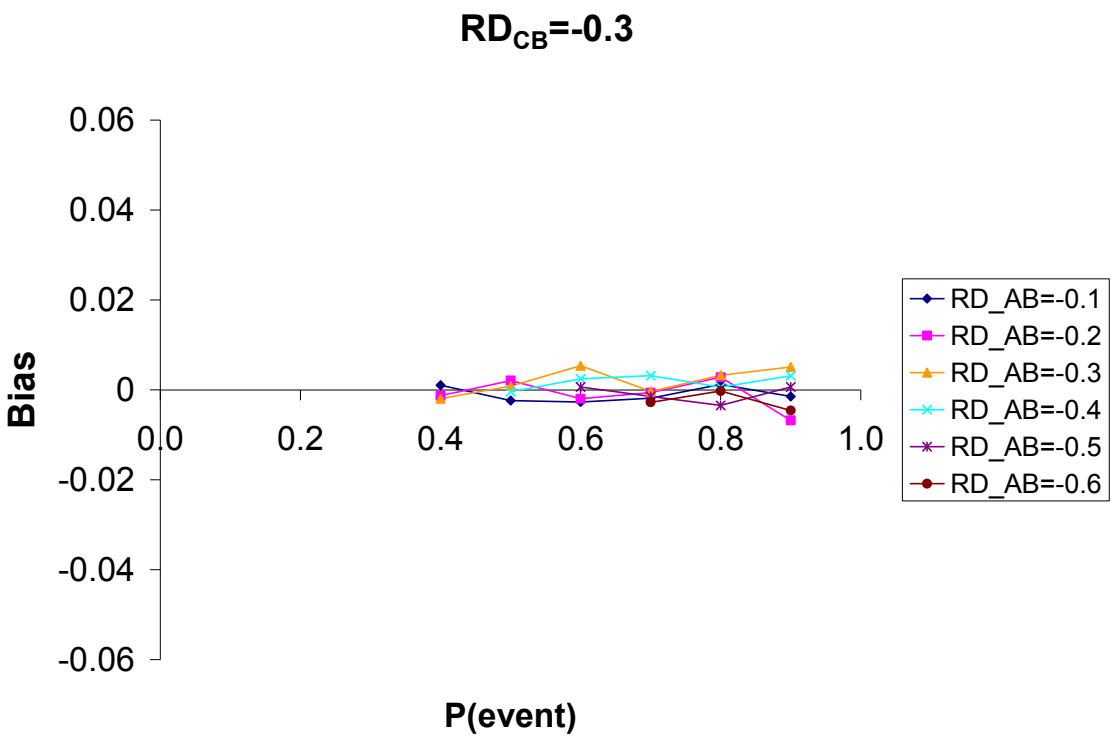
a)



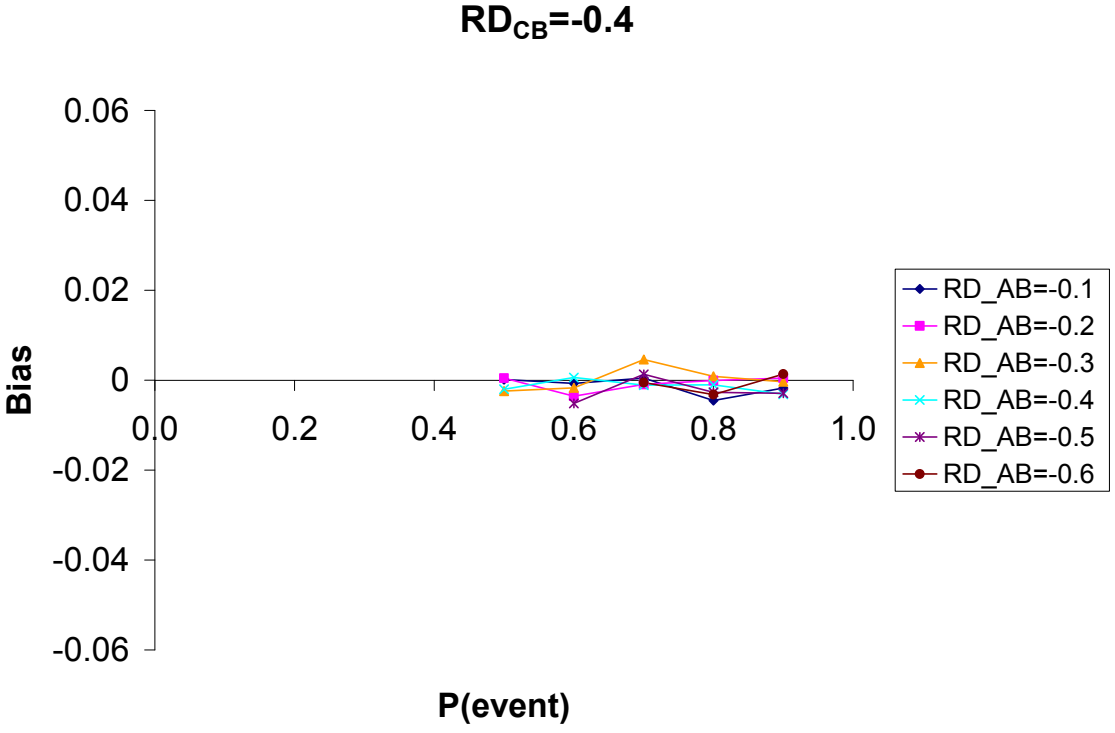
b)



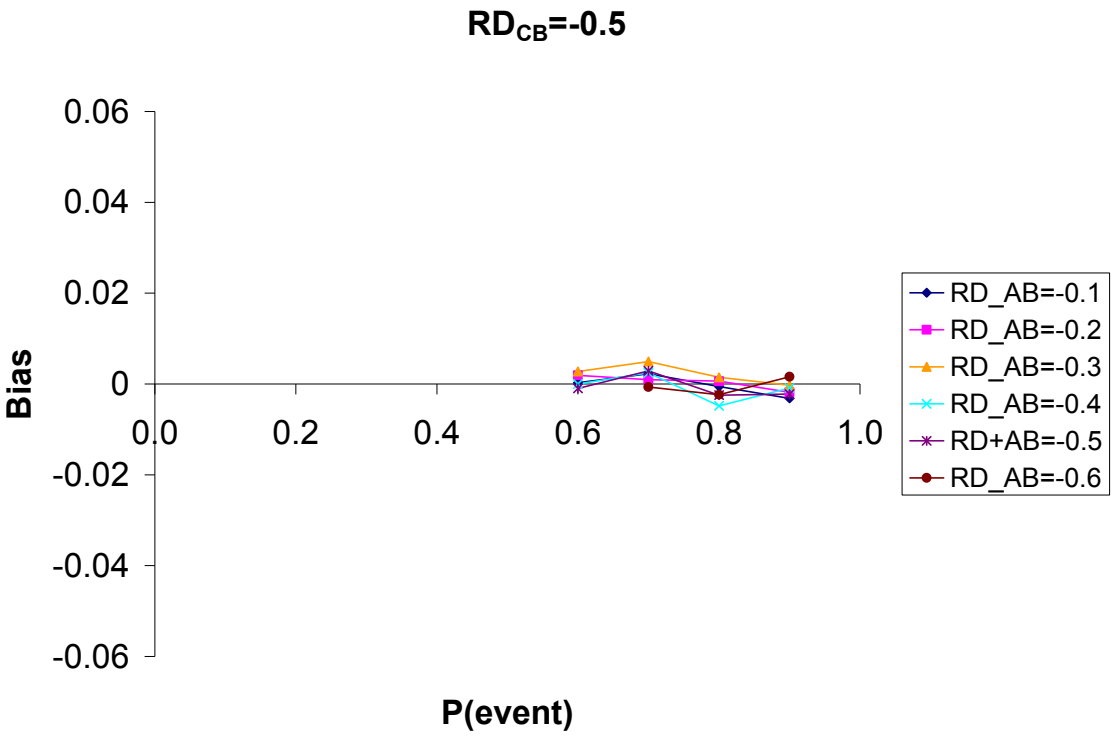
c)


RD_CB=-0.3
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
Bias
0.0
0.2
0.4
0.6
0.8
1.0
P(event)
RD_AB=-0.1
RD_AB=-0.2
RD_AB=-0.3
RD_AB=-0.4
RD_AB=-0.5
RD_AB=-0.6

d)


RD_CB=-0.4
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
Bias
0.0
0.2
0.4
0.6
0.8
1.0
P(event)
RD_AB=-0.1
RD_AB=-0.2
RD_AB=-0.3
RD_AB=-0.4
RD_AB=-0.5
RD_AB=-0.6

e)


RD_CB=-0.5
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
Bias
0.0
0.2
0.4
0.6
0.8
1.0
P(event)
RD_AB=-0.1
RD_AB=-0.2
RD_AB=-0.3
RD_AB=-0.4
RD+AB=-0.5
RD_AB=-0.6

f)

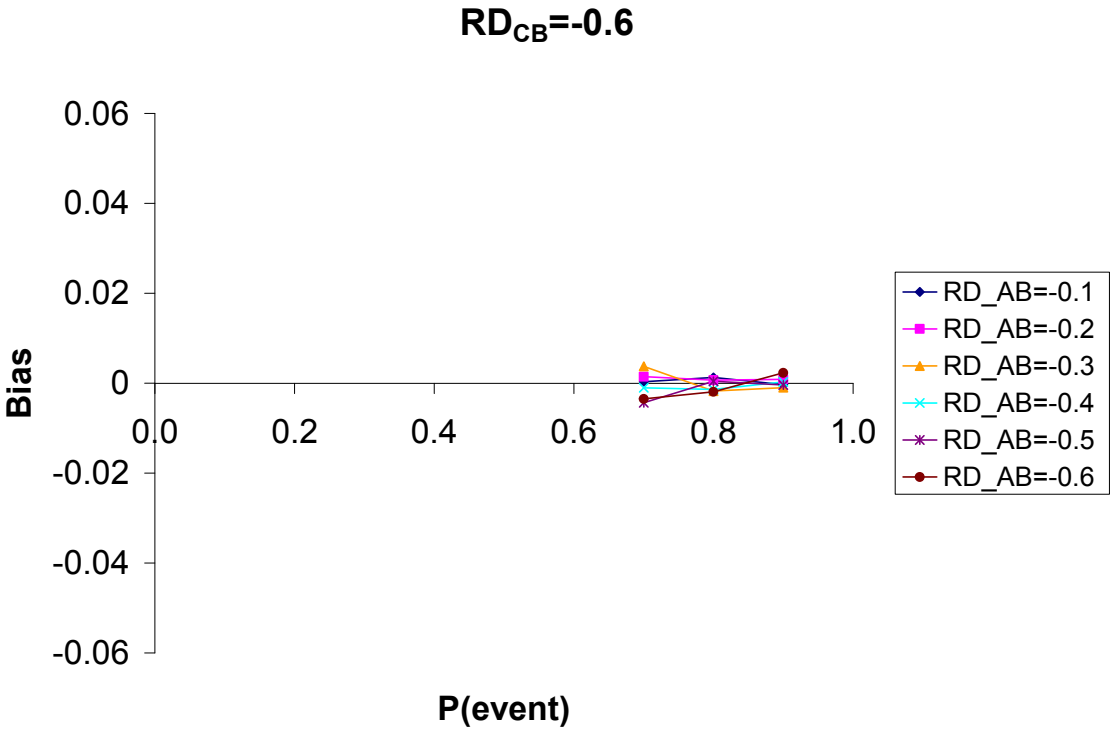
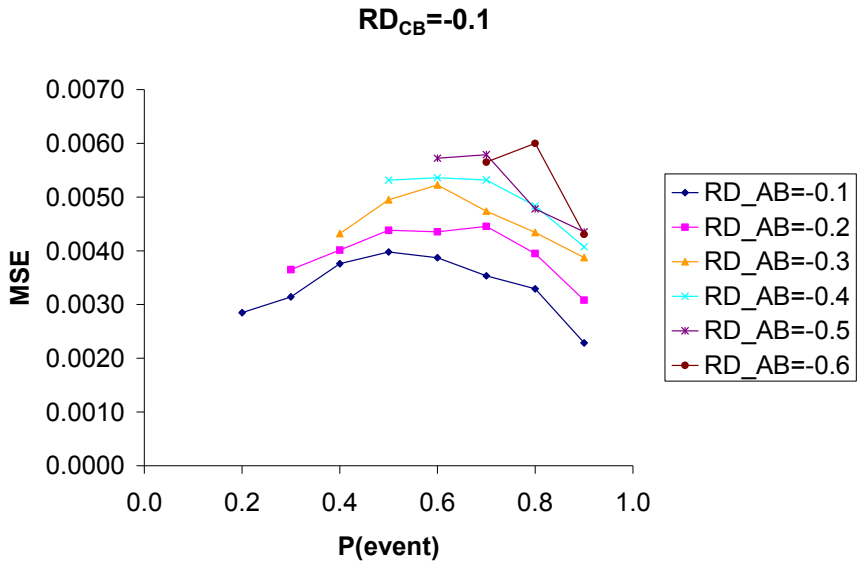
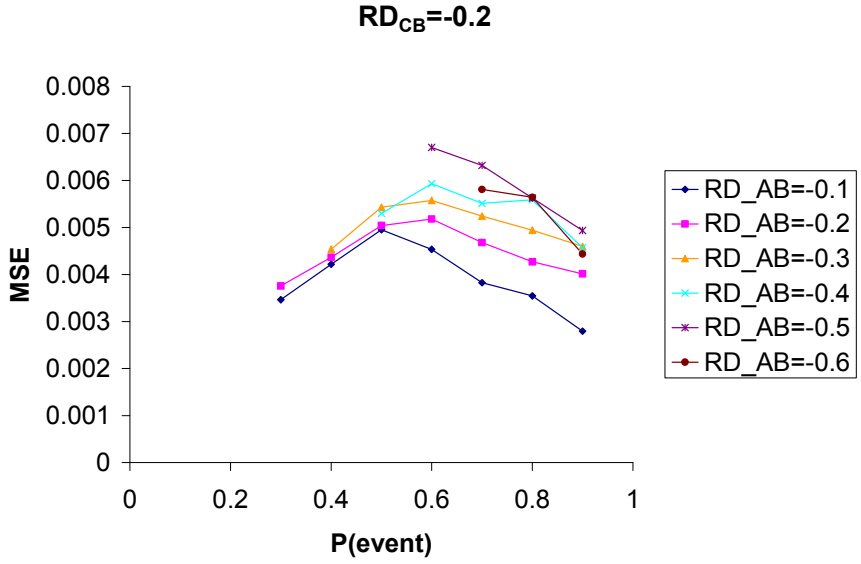

RD_CB=-0.6
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
Bias
0.0
0.2
0.4
0.6
0.8
1.0
P(event)
RD_AB=-0.1
RD_AB=-0.2
RD_AB=-0.3
RD_AB=-0.4
RD_AB=-0.5
RD_AB=-0.6

Figure B.4.4: Mean square error (MSE) for indirect risk difference (RD) estimates for various parameter settings.

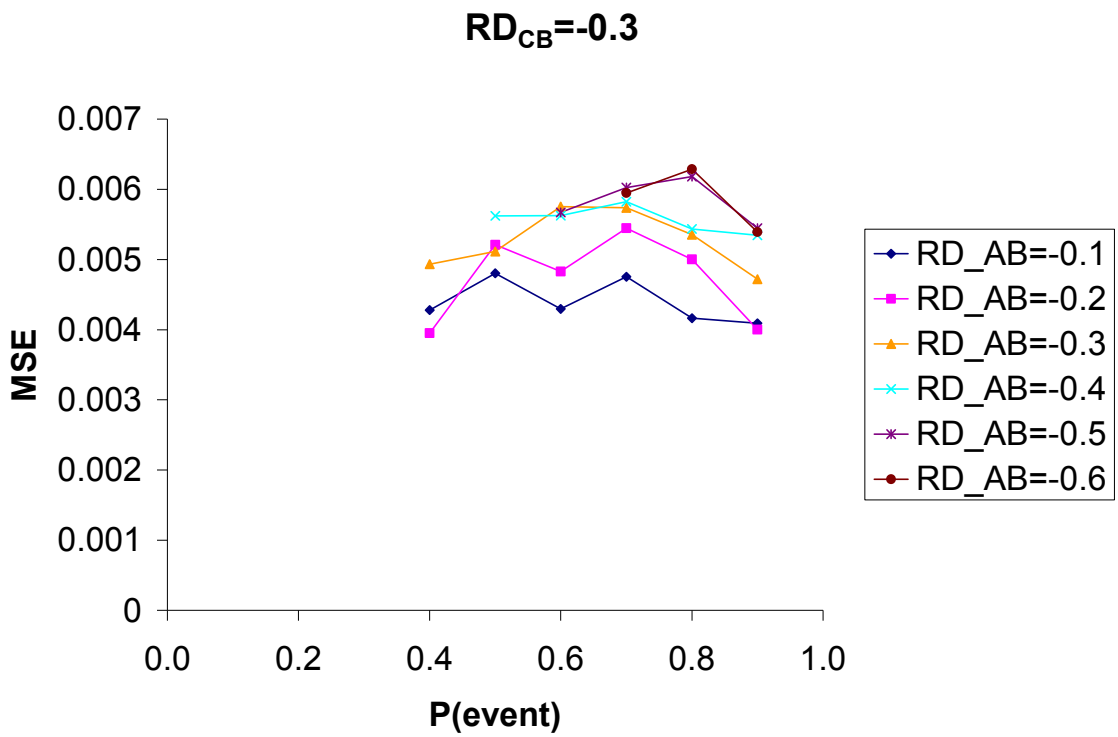
a)



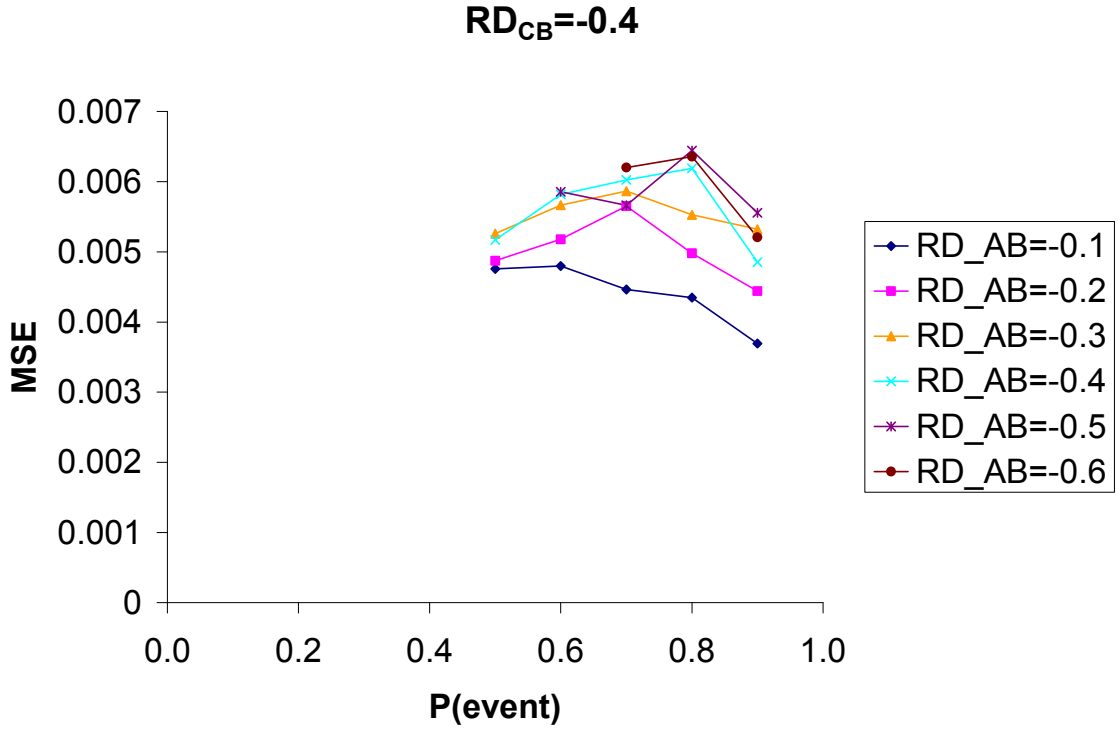
b)



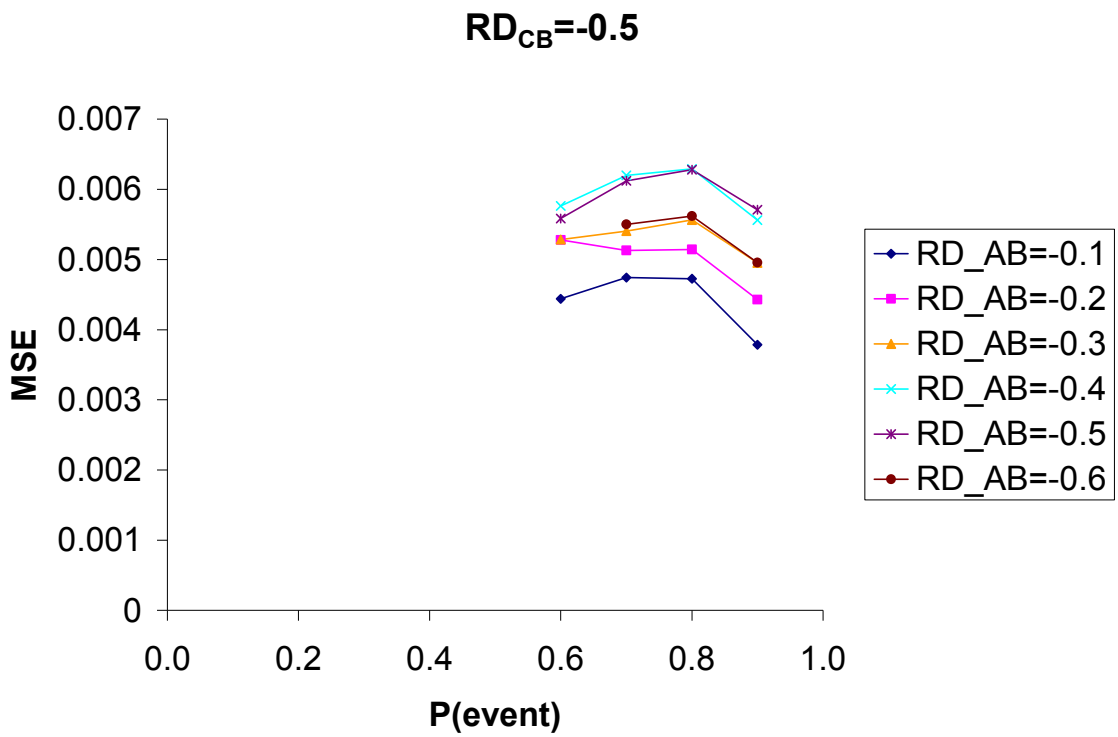
c)


RD_CB=-0.3
MSE
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
0.0
0.2
0.4
0.6
0.8
1.0
P(event)
RD_AB=-0.1
RD_AB=-0.2
RD_AB=-0.3
RD_AB=-0.4
RD_AB=-0.5
RD_AB=-0.6

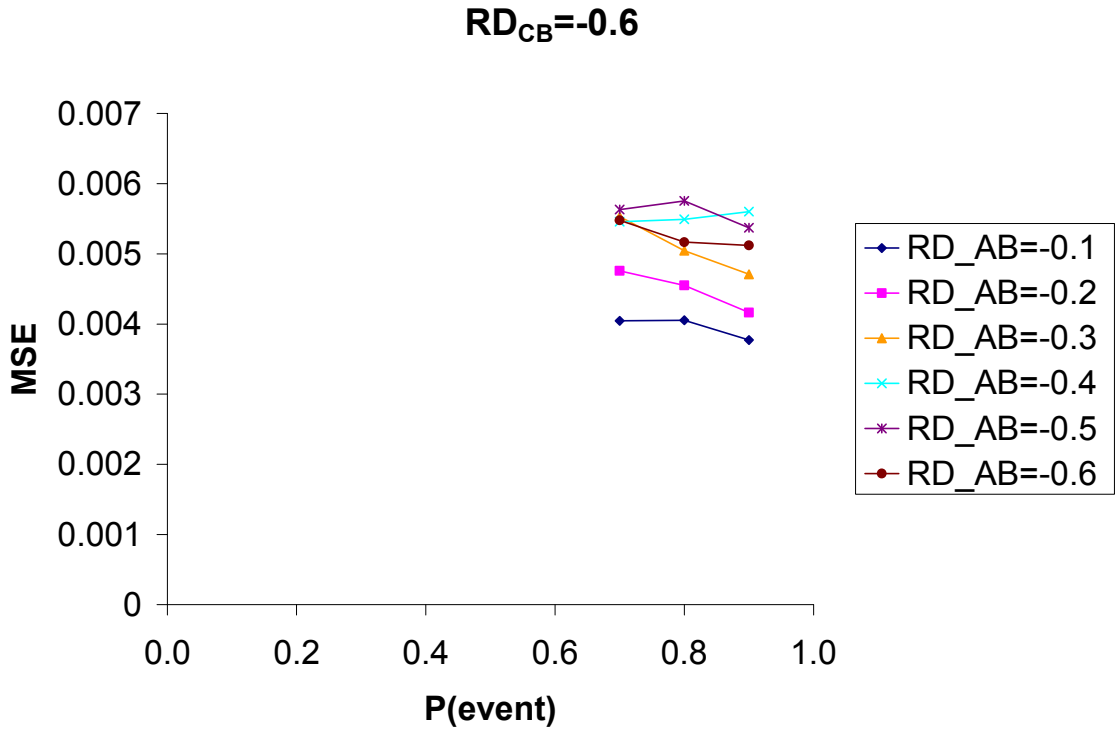
d)


RD_CB=-0.4
MSE
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
0.0
0.2
0.4
0.6
0.8
1.0
P(event)
RD_AB=-0.1
RD_AB=-0.2
RD_AB=-0.3
RD_AB=-0.4
RD_AB=-0.5
RD_AB=-0.6

e)



f)



B.5 Mean Difference

For each of the three populations (A, B and C) in the Monte Carlo simulation for the mean difference, the level of the outcome of interest for each population was selected according to a specific combination of the parameters ES_{AB}, ES_{CB}, M_B (where M_B is the mean of the outcome of interest in population B) and CV_B, as follows:

ES_{AB}: .2, .5, .8
ES_{CB}: .2, .5, .8
M_B: 10, 20, 30
CV_B=0.1, 0.3, 0.5

For the various combinations of these parameters, the results of the simulation for the bias, variance and MSE for the direct and indirect estimators of the ES are provided in Table B.5.1.

As an illustration of the frequency distribution of the estimators, the frequency distributions for the parameter settings ES_{AB}=0.2, ES_{CB}=0.5, CV_B=0.3 and $Mean_B$ =10, 20, 30 for the direct and indirect estimators are presented graphically in Figure B.5.1. It is apparent from these figures that both estimators have a mound shape and symmetric distribution. The indirect estimator has a larger variance and bias.

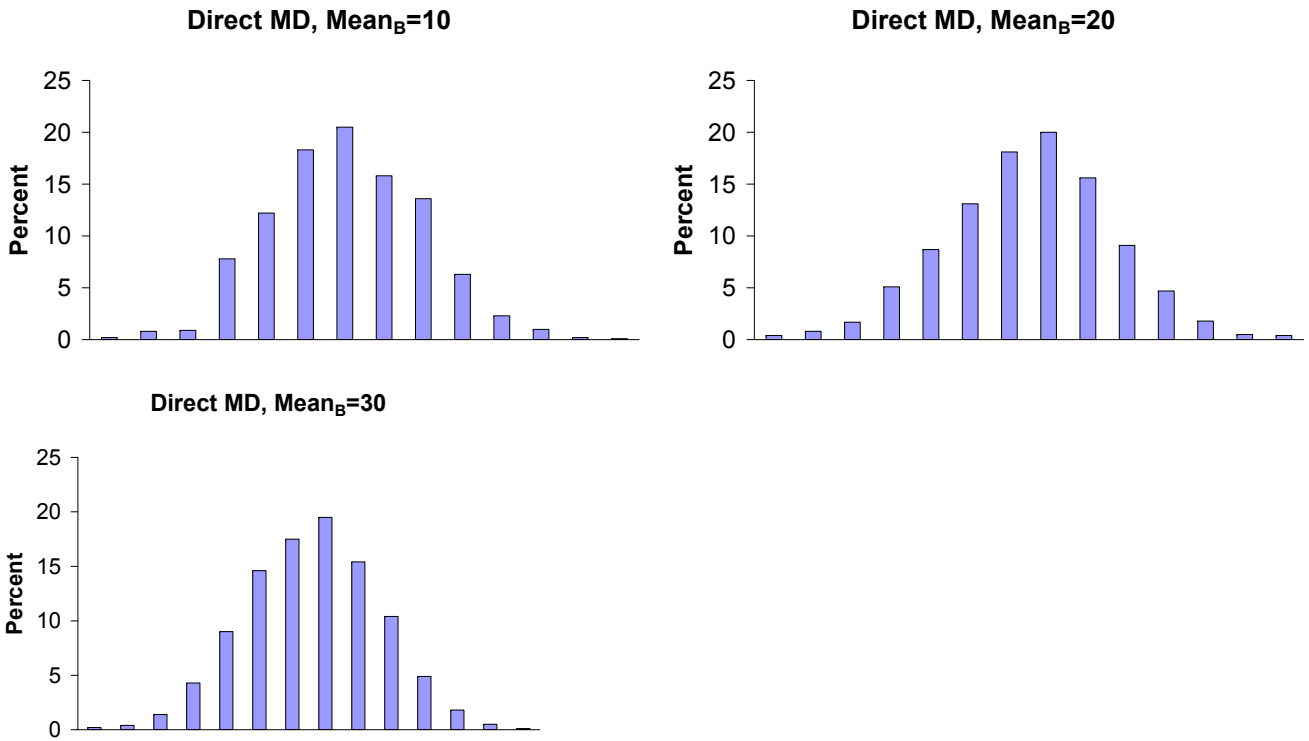
Table B.5.1: Bias, variance and mean square error (MSE) of direct and indirect mean difference (MD) estimators for different settings of the indicated parameters (k=3 treatments)

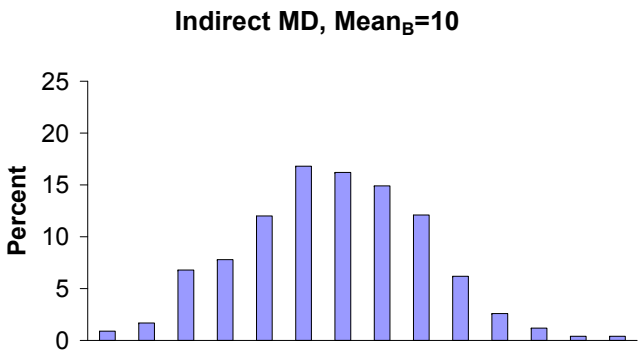
Mean difference Settings	CV_B	Mean_B	Direct Estimator			Indirect Estimator		
			Variance	Bias	MSE	Variance	Bias	MSE
ES_{CB}=0.8	**0.5**	**10**	0.467	-0.304	0.559	0.990	-0.348	1.111
HR_{AB}=0.8		**20**	1.606	-0.634	2.008	3.874	-0.659	4.309
		30	4.064	-0.930	4.929	8.427	-0.922	9.276
ES_{CB}=0.8		**10**	0.441	-0.319	0.542	0.992	-0.320	1.095
ES_{AB}=0.5		**20**	1.857	-0.611	2.231	3.733	-0.678	4.193
		30	3.786	-1.018	4.822	8.281	-1.017	9.314
ES_{CB}=0.8		**10**	0.455	-0.300	0.545	0.916	-0.304	1.008
ES_{AB}=0.2		**20**	1.737	-0.603	2.100	3.884	-0.596	4.240
		30	4.055	-0.898	4.861	8.178	-0.846	8.894
ES_{CB}=0.8	**0.3**	**10**	0.150	-0.177	0.182	0.320	-0.165	0.348
HR_{AB}=0.8		**20**	0.606	-0.369	0.742	1.357	-0.343	1.475
		30	1.527	-0.611	1.899	3.067	-0.513	3.330
ES_{CB}=0.8		**10**	0.157	-0.176	0.188	0.325	-0.182	0.358
ES_{AB}=0.5		**20**	0.671	-0.443	0.868	1.227	-0.399	1.386
		30	1.422	-0.605	1.788	3.058	-0.599	3.417
ES_{CB}=0.8		**10**	0.159	-0.189	0.195	0.325	-0.175	0.355
ES_{AB}=0.2		**20**	0.630	-0.414	0.802	1.349	-0.370	1.486
		30	1.416	-0.558	1.727	2.805	-0.559	3.117
ES_{CB}=0.8	**0.1**	**10**	0.017	-0.062	0.021	0.037	-0.052	0.039
HR_{AB}=0.8		**20**	0.069	-0.118	0.083	0.143	-0.135	0.161
		30	0.154	-0.194	0.191	0.334	-0.168	0.362
ES_{CB}=0.8		**10**	0.016	-0.060	0.020	0.039	-0.076	0.045
ES_{AB}=0.5		**20**	0.070	-0.132	0.088	0.146	-0.112	0.158

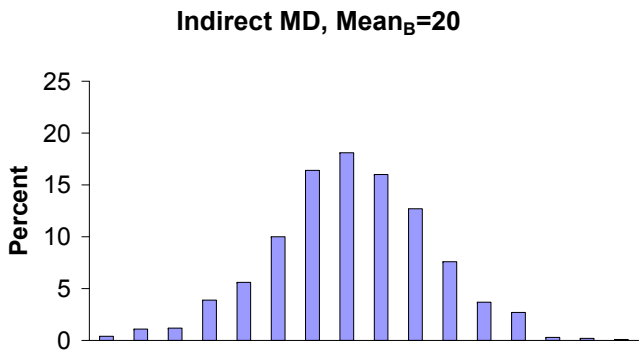
Mean difference Settings	CV_B	Mean_B	Direct Estimator			Indirect Estimator		
			Variance	Bias	MSE	Variance	Bias	MSE
		30	0.153	-0.183	0.187	0.353	-0.216	0.399
ES_{CB}=0.8		**10**	0.019	-0.063	0.023	0.033	-0.071	0.038
ES_{AB}=0.2		**20**	0.073	-0.120	0.087	0.142	-0.112	0.155
		30	0.157	-0.209	0.201	0.333	-0.208	0.376
ES_{CB}=0.5	**0.5**	**10**	0.437	-0.317	0.537	0.946	-0.349	1.068
HR_{AB}=0.8		**20**	1.686	-0.534	1.971	3.506	-0.658	3.939
		30	4.115	-1.066	5.252	8.261	-0.959	9.182
ES_{CB}=0.5		**10**	0.433	-0.322	0.537	0.866	-0.282	0.946
ES_{AB}=0.5		**20**	1.704	-0.611	2.077	3.638	-0.657	4.070
		30	3.937	-1.066	5.073	8.155	-0.933	9.025
ES_{CB}=0.5		**10**	0.421	-0.291	0.505	0.862	-0.263	0.932
ES_{AB}=0.2		**20**	1.859	-0.677	2.318	3.431	-0.577	3.764
		30	4.097	-0.892	4.892	8.077	-0.944	8.969
ES_{CB}=0.5	**0.3**	**10**	0.162	-0.212	0.208	0.335	-0.187	0.369
HR_{AB}=0.8		**20**	0.671	-0.381	0.817	1.375	-0.399	1.533
		30	1.504	-0.489	1.742	3.079	-0.520	3.349
ES_{CB}=0.5		**10**	0.150	-0.202	0.191	0.308	-0.197	0.347
ES_{AB}=0.5		**20**	0.653	-0.360	0.782	1.249	-0.331	1.358
		30	1.413	-0.515	1.678	2.986	-0.663	3.426
ES_{CB}=0.5		**10**	0.150	-0.194	0.188	0.333	-0.165	0.360
ES_{AB}=0.2		**20**	0.701	-0.390	0.853	1.346	-0.389	1.498
		30	1.461	-0.536	1.748	3.077	-0.574	3.406
ES_{CB}=0.5	**0.1**	**10**	0.017	-0.063	0.021	0.038	-0.060	0.041
HR_{AB}=0.8		**20**	0.073	-0.122	0.088	0.151	-0.138	0.170
		30	0.150	-0.205	0.192	0.345	-0.206	0.388
ES_{CB}=0.5		**10**	0.017	-0.062	0.021	0.037	-0.064	0.041
ES_{AB}=0.5		**20**	0.072	-0.129	0.089	0.144	-0.121	0.159
		30	0.168	-0.203	0.209	0.333	-0.179	0.365
ES_{CB}=0.5		**10**	0.018	-0.059	0.022	0.037	-0.062	0.040
ES_{AB}=0.2		**20**	0.071	-0.132	0.088	0.146	-0.136	0.164
		30	0.160	-0.188	0.196	0.312	-0.187	0.347
ES_{CB}=0.2	**0.5**	**10**	0.472	-0.323	0.576	0.903	-0.282	0.983
HR_{AB}=0.8		**20**	1.682	-0.673	2.134	3.564	-0.590	3.913
		30	3.768	-1.021	4.811	8.338	-0.881	9.114
ES_{CB}=0.2		**10**	0.413	-0.296	0.501	0.879	-0.317	0.980
ES_{AB}=0.5		**20**	1.814	-0.614	2.191	3.501	-0.674	3.956
		30	3.951	-0.879	4.723	8.262	-0.933	9.133
ES_{CB}=0.2		**10**	0.457	-0.311	0.554	0.967	-0.366	1.101
ES_{AB}=0.2		**20**	1.833	-0.610	2.206	3.700	-0.632	4.100
		30	4.022	-0.969	4.961	8.593	-1.068	9.734
ES_{CB}=0.2	**0.3**	**10**	0.171	-0.191	0.208	0.303	-0.166	0.330
HR_{AB}=0.8		**20**	0.657	-0.312	0.754	1.327	-0.394	1.482
		30	1.493	-0.540	1.785	2.909	-0.602	3.272
ES_{CB}=0.2		**10**	0.165	-0.193	0.202	0.362	-0.173	0.392
ES_{AB}=0.5		**20**	0.611	-0.368	0.747	1.406	-0.397	1.564
		30	1.462	-0.547	1.760	2.924	-0.655	3.353
ES_{CB}=0.2		**10**	0.167	-0.182	0.200	0.317	-0.159	0.342
ES_{AB}=0.2		**20**	0.598	-0.369	0.734	1.304	-0.387	1.454
		30	1.453	-0.633	1.853	2.914	-0.593	3.265
ES_{CB}=0.2	**0.1**	**10**	0.017	-0.068	0.022	0.036	-0.070	0.041
HR_{AB}=0.8		**20**	0.066	-0.136	0.085	0.145	-0.122	0.160

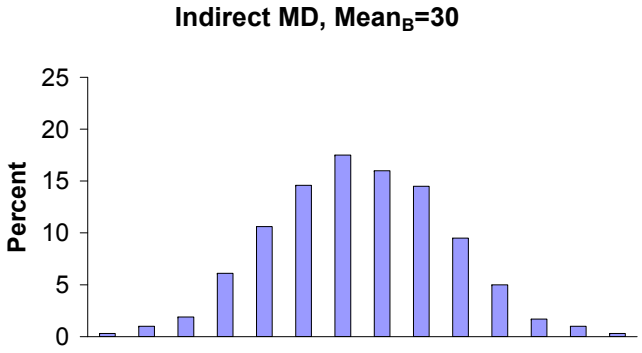
Mean difference Settings	CV_B	Mean_B	Direct Estimator			Indirect Estimator		
			Variance	Bias	MSE	Variance	Bias	MSE
		30	0.153	-0.201	0.193	0.334	-0.204	0.375
ES_{CB}=0.2		**10**	0.018	-0.063	0.022	0.037	-0.067	0.042
ES_{AB}=0.5		**20**	0.071	-0.125	0.087	0.139	-0.141	0.159
		30	0.164	-0.185	0.198	0.341	-0.191	0.378
ES_{CB}=0.2		**10**	0.018	-0.069	0.023	0.035	-0.066	0.039
ES_{AB}=0.2		**20**	0.067	-0.129	0.083	0.142	-0.127	0.158
		30	0.143	-0.204	0.185	0.319	-0.196	0.358

Figure B.5.1: Frequency distribution of direct and indirect mean difference (MD) estimators for the parameter settings ES_{AB}=0.2, ES_{CB}=0.5, CV_B=0.3 and $Mean_B$ =10, 20, 30 (k=3 treatments)



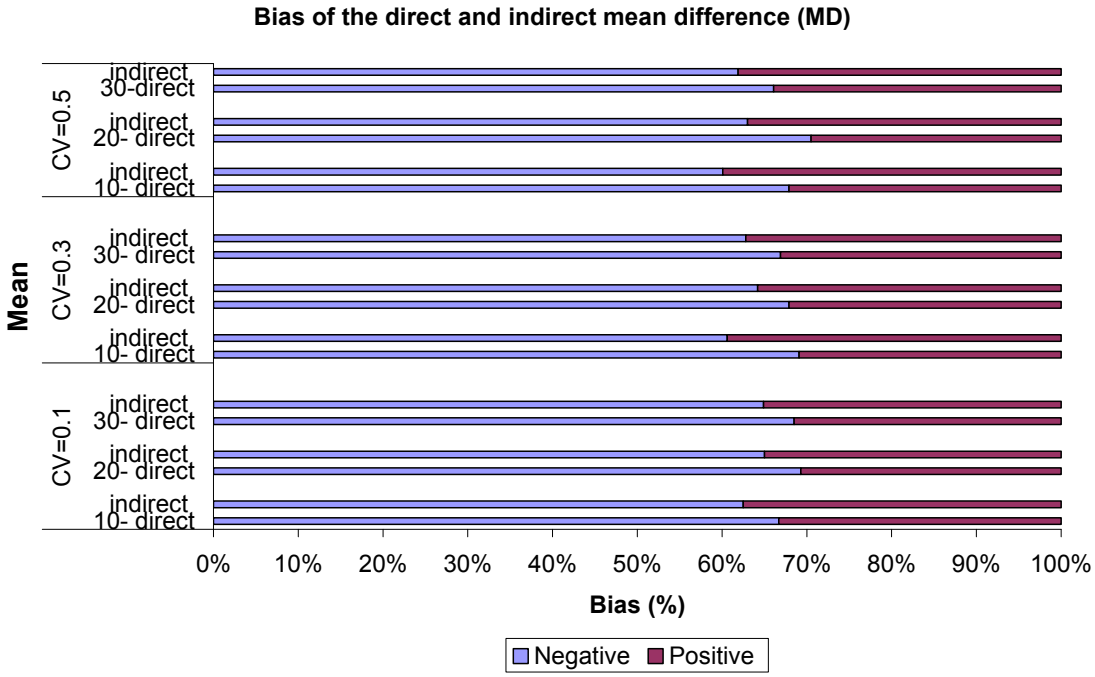

Indirect MD, Mean$_B$=10
Percent
25
20
15
10
5
0


Indirect MD, Mean$_B$=20
Percent
25
20
15
10
5
0


Indirect MD, Mean$_B$=30
Percent
25
20
15
10
5
0

For the parameter settings ES_{AB}=0.2, HR_{CB}=0.5 and CV_B=0.1, 0.3, 0.5, the direction of the biases over the 1000 samples are illustrated in Figure B.5.2 for the direct and indirect approaches under the different settings of $Mean_B$: 10, 20, 30. As noted, although the bias is small, in most instances the direct and indirect estimates underestimate the parameter.

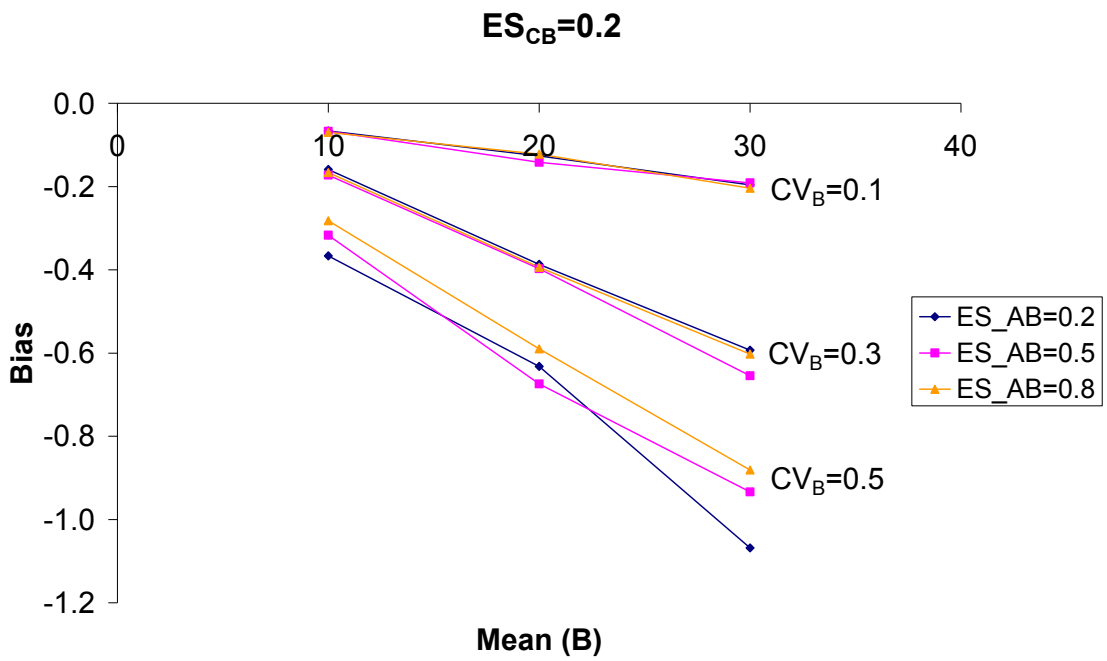
Figure B.5.2: Direction of the bias of the direct and indirect mean difference (MD) estimates for the parameter settings ES_{AB}=0.2, ES_{CB}=0.5, CVB=0.1, 0.3, 0.5 (k=3 treatments)



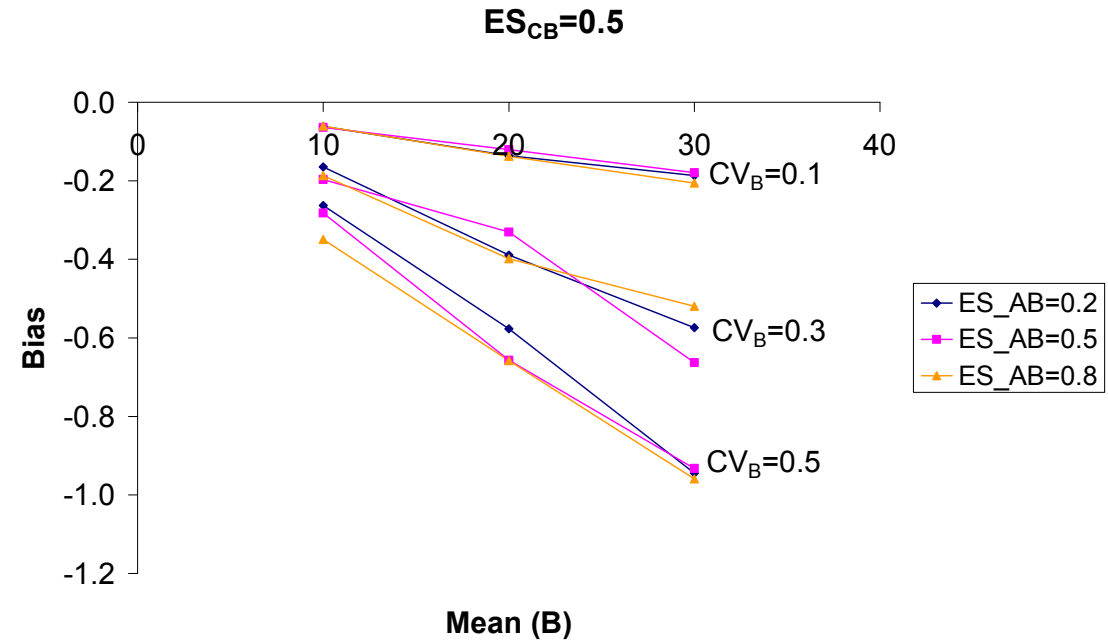
Patterns of bias and MSE for the indirect effect size (ES) estimates are displayed for different settings of ES_{AB}, ES_{CB}, and the mean in population B (Figure B.5.3 and Figure B.5.4).

Figure B.5.3: Bias for indirect mean difference (MD) estimates for various parameter settings.

a)



b)



c)

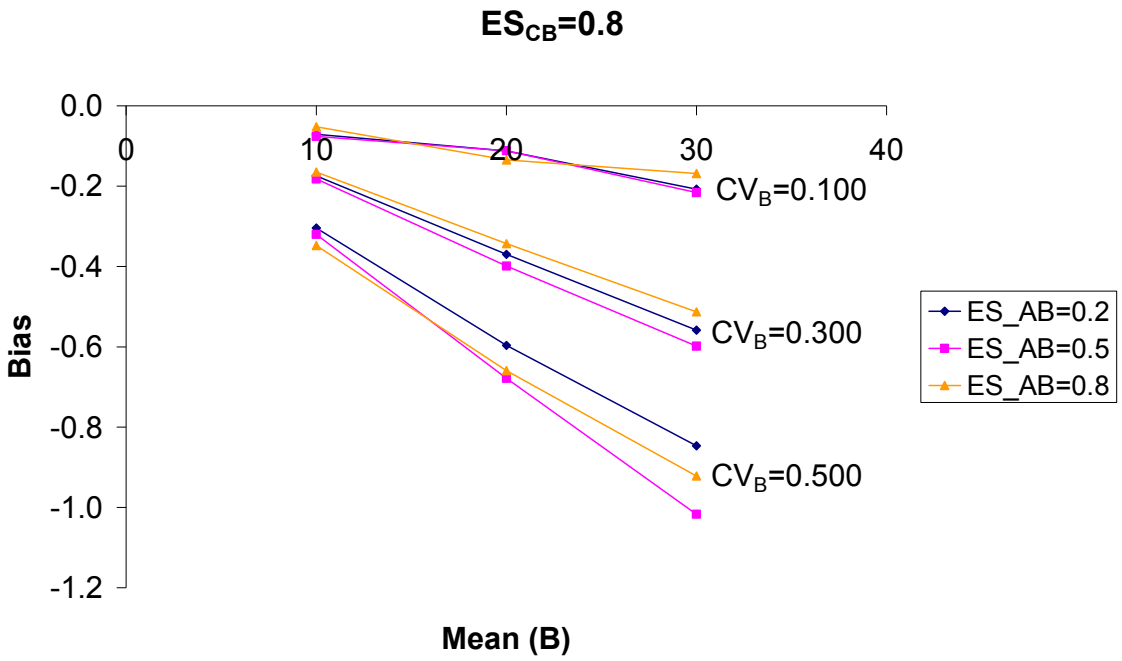
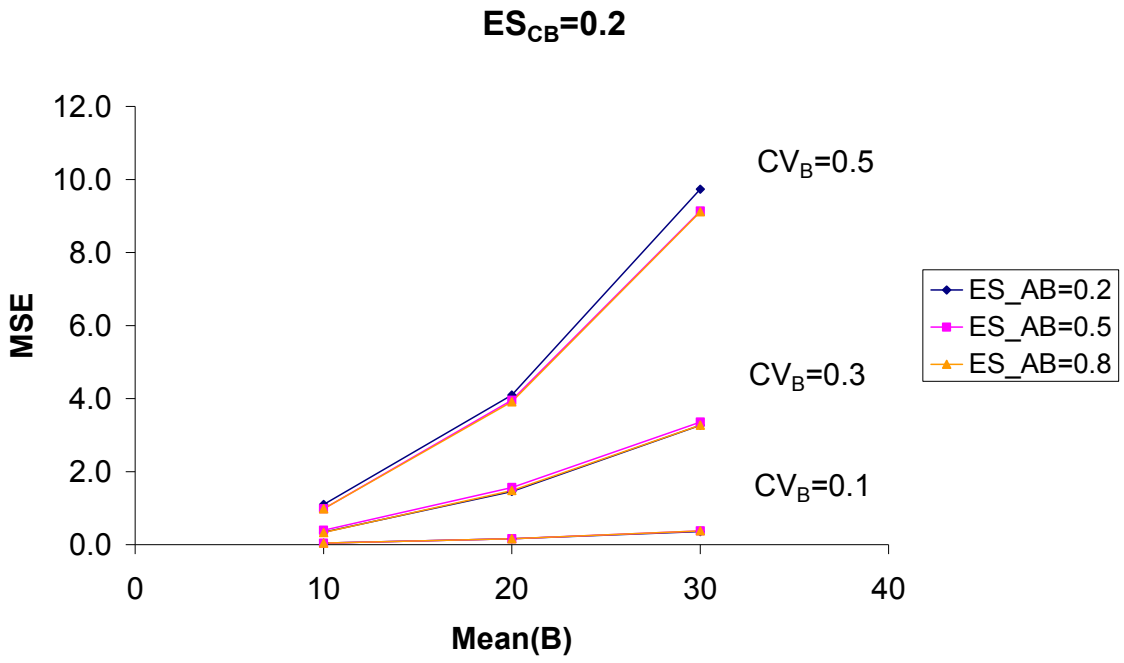
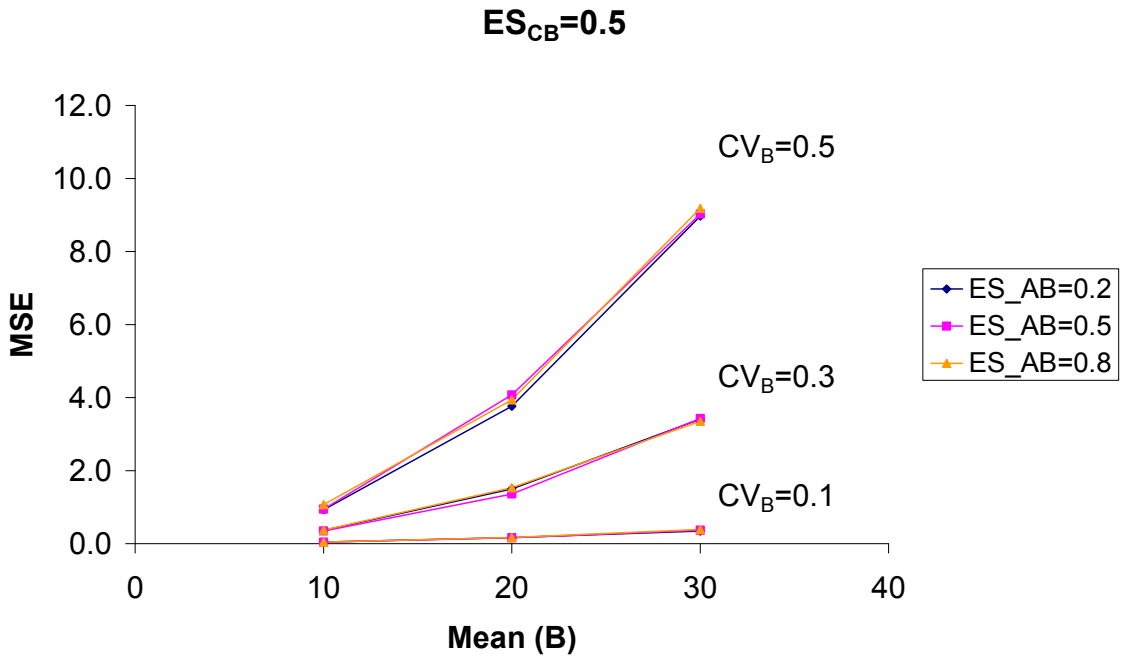

ES_CB=0.8
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
0
10
20
30
40
Bias
Mean (B)
CV_B=0.100
CV_B=0.300
CV_B=0.500
ES_AB=0.2
ES_AB=0.5
ES_AB=0.8

Figure B.5.4: MSE for indirect indirect mean difference (MD) estimates for various parameter settings.

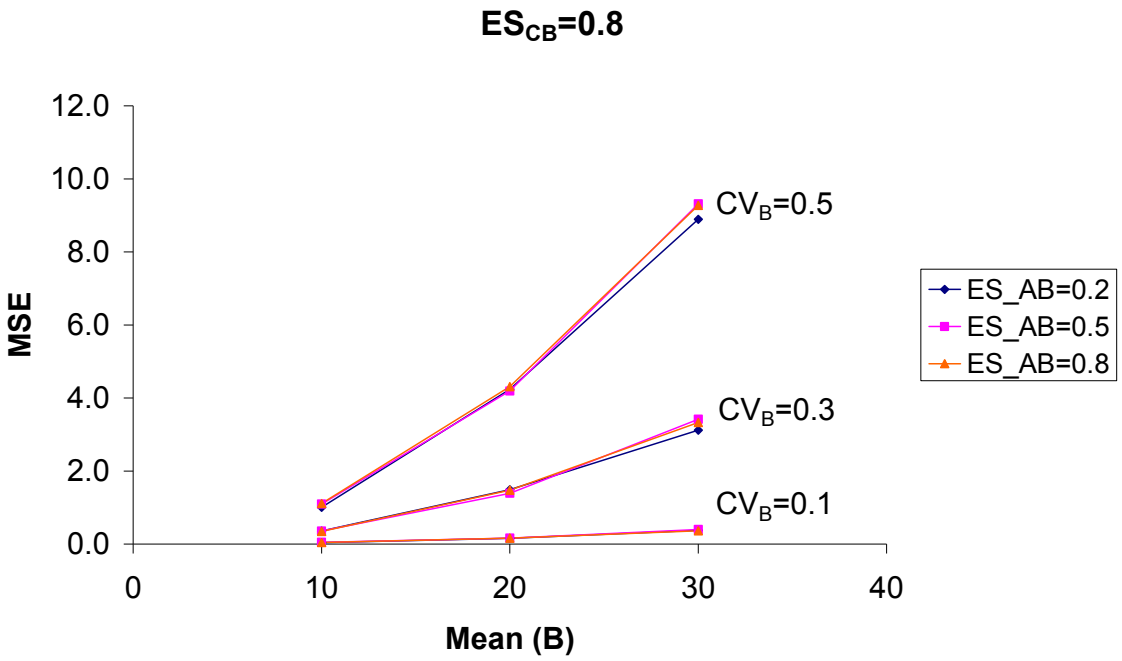
a)



b)



c)


ES_CB=0.8
12.0
10.0
8.0
6.0
4.0
2.0
0.0
MSE
0
10
20
30
40
Mean (B)
CV_B=0.5
CV_B=0.3
CV_B=0.1
ES_AB=0.2
ES_AB=0.5
ES_AB=0.8

B.6 Simulation Results for the Hazard Ratio

For each of the three populations (A, B and C) in the Monte Carlo simulation for the hazard ratio, the outcome risk level for each population was selected according to a specific combination of the parameters HR_{AB}, HR_{CB}, and the hazard rate in population B, as follows:

HR_{AB}: 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
HR_{CB}: 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
Hazard rate in population B: 0.05, 0.1, 0.2, 0.3, 0.4, 0.5

For the various combinations of these parameters, the results of the simulation for the bias, variance and MSE for the direct and indirect estimators of the OR are provided in Table B.6.1.

As an illustration of the frequency distribution of the estimators, the frequency distributions for the parameter settings HR_{AB}=0.6, HR_{CB}=0.8 and hazard rate in population B=0.05, 0.3, 0.5, for the direct and indirect estimators are presented graphically in Figure B.6.1 on the logarithmic scale. It is apparent from these figures that both estimators have a mound shape, symmetric distribution (on the logarithmic scale). The indirect estimator has a larger variance and bias.

Table B.6.1: Bias, variance and mean square error (MSE) of direct and indirect hazard ratio (HR) estimators for different settings of the indicated parameters (k=3 treatments)

Hazard Ratio Settings	Hazard rate	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
HR_{AB}=1.5	**0.05**	0.004	-0.001	0.004	0.007	0.015	0.007
HR_{CB}=1.5	**0.1**	0.003	0.003	0.003	0.008	0.003	0.008
	0.2	0.004	0.002	0.004	0.008	0.007	0.008
	0.3	0.004	0.000	0.004	0.008	0.024	0.008
	0.4	0.003	0.003	0.003	0.008	0.006	0.008
	0.5	0.003	0.000	0.003	0.008	0.005	0.008
HR_{AB}=1.5	**0.05**	0.005	0.002	0.005	0.009	-0.037	0.010
HR_{CB}=1.3	**0.1**	0.005	0.004	0.005	0.010	-0.007	0.010
	0.2	0.004	0.005	0.004	0.010	-0.001	0.010
	0.3	0.005	0.001	0.005	0.009	-0.045	0.011
	0.4	0.005	0.001	0.005	0.010	-0.016	0.010
	0.5	0.005	0.001	0.005	0.011	0.011	0.011
HR_{AB}=1.5	**0.05**	0.006	0.005	0.006	0.013	0.028	0.014
HR_{CB}=1.1	**0.1**	0.007	-0.006	0.007	0.014	-0.024	0.015
	0.2	0.006	-0.002	0.006	0.011	-0.025	0.012
	0.3	0.007	0.003	0.007	0.014	-0.065	0.018
	0.4	0.007	0.008	0.008	0.016	0.042	0.018
	0.5	0.007	0.005	0.007	0.014	-0.017	0.014
HR_{AB}=1.3	**0.05**	0.002	-0.001	0.002	0.006	0.006	0.006

Hazard Ratio Settings	Hazard rate	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
HR_{CB}=1.5	**0.1**	0.003	0.001	0.003	0.005	0.007	0.005
	0.2	0.003	0.003	0.003	0.006	0.005	0.006
	0.3	0.003	0.001	0.003	0.006	0.007	0.006
	0.4	0.003	0.002	0.003	0.005	-0.001	0.005
	0.5	0.003	0.000	0.003	0.006	-0.004	0.006
HR_{AB}=1.3	**0.05**	0.003	0.000	0.003	0.008	0.002	0.008
HR_{CB}=1.3	**0.1**	0.004	0.002	0.004	0.008	-0.003	0.008
	0.2	0.003	-0.002	0.003	0.007	0.003	0.007
	0.3	0.003	0.004	0.003	0.008	0.014	0.008
	0.4	0.004	0.001	0.004	0.007	0.011	0.007
	0.5	0.004	0.003	0.004	0.007	-0.007	0.007
HR_{AB}=1.3	**0.05**	0.005	0.003	0.005	0.010	0.030	0.011
HR_{CB}=1.1	**0.1**	0.005	0.000	0.005	0.009	-0.014	0.009
	0.2	0.005	0.000	0.005	0.012	0.047	0.014
	0.3	0.005	0.007	0.005	0.010	-0.049	0.012
	0.4	0.005	-0.001	0.005	0.009	0.012	0.010
	0.5	0.005	0.003	0.005	0.011	0.017	0.012
HR_{AB}=1.1	**0.05**	0.002	0.002	0.002	0.004	-0.005	0.004
HR_{CB}=1.5	**0.1**	0.002	0.000	0.002	0.004	0.013	0.004
	0.2	0.002	0.003	0.002	0.004	-0.019	0.004
	0.3	0.002	0.002	0.002	0.004	0.011	0.004
	0.4	0.002	0.001	0.002	0.004	-0.015	0.004
	0.5	0.002	0.000	0.002	0.004	-0.025	0.004
HR_{AB}=1.1	**0.05**	0.003	0.003	0.003	0.006	-0.025	0.006
HR_{CB}=1.3	**0.1**	0.003	0.001	0.003	0.005	0.002	0.005
	0.2	0.002	0.003	0.002	0.006	0.008	0.006
	0.3	0.003	0.001	0.003	0.006	0.040	0.007
	0.4	0.003	0.001	0.003	0.006	0.012	0.006
	0.5	0.003	0.002	0.003	0.004	-0.035	0.006
HR_{AB}=1.1	**0.05**	0.004	0.000	0.004	0.007	-0.010	0.007
HR_{CB}=1.1	**0.1**	0.003	-0.001	0.003	0.007	-0.015	0.007
	0.2	0.004	0.004	0.004	0.007	-0.024	0.007
	0.3	0.004	-0.002	0.004	0.007	-0.027	0.008
	0.4	0.004	0.003	0.004	0.008	0.010	0.008
	0.5	0.004	-0.002	0.004	0.008	0.017	0.008
HR_{AB}=0.9	**0.05**	0.004	0.003	0.004	0.008	0.003	0.008
HR_{CB}=0.9	**0.1**	0.004	0.003	0.004	0.007	0.005	0.007
	0.2	0.004	0.001	0.004	0.008	0.015	0.008
	0.3	0.004	0.005	0.004	0.007	0.015	0.007
	0.4	0.004	0.000	0.004	0.007	-0.006	0.007
	0.5	0.004	-0.002	0.004	0.008	0.007	0.008
HR_{AB}=0.9	**0.05**	0.005	0.000	0.005	0.010	-0.016	0.011

Hazard Ratio Settings	Hazard rate	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
HR_{CB}=0.8	**0.1**	0.004	0.001	0.004	0.009	0.048	0.011
	0.2	0.005	0.005	0.005	0.009	-0.001	0.009
	0.3	0.005	0.002	0.005	0.010	0.001	0.010
	0.4	0.005	-0.001	0.005	0.010	0.011	0.011
	0.5	0.005	0.004	0.005	0.010	-0.018	0.011
HR_{AB}=0.9	**0.05**	0.006	0.005	0.006	0.012	-0.001	0.012
HR_{CB}=0.7	**0.1**	0.006	0.005	0.006	0.015	0.038	0.016
	0.2	0.006	0.007	0.006	0.011	0.008	0.011
	0.3	0.006	0.007	0.006	0.012	-0.011	0.012
	0.4	0.006	-0.001	0.006	0.012	-0.013	0.012
	0.5	0.006	0.000	0.006	0.011	-0.001	0.011
HR_{AB}=0.9	**0.05**	0.009	0.004	0.009	0.016	0.027	0.017
HR_{CB}=0.6	**0.1**	0.010	0.009	0.010	0.019	0.005	0.019
	0.2	0.008	0.001	0.008	0.015	-0.006	0.015
	0.3	0.009	0.010	0.009	0.020	0.123	0.036
	0.4	0.008	0.004	0.008	0.016	-0.034	0.017
	0.5	0.009	0.006	0.009	0.019	0.009	0.019
HR_{AB}=0.9	**0.05**	0.013	0.001	0.013	0.024	-0.042	0.026
HR_{CB}=0.5	**0.1**	0.013	0.007	0.013	0.025	-0.053	0.027
	0.2	0.013	-0.004	0.013	0.026	0.041	0.028
	0.3	0.014	0.008	0.014	0.023	-0.042	0.025
	0.4	0.011	0.001	0.011	0.023	0.019	0.023
	0.5	0.013	0.013	0.014	0.029	0.011	0.029
HR_{AB}=0.9	**0.05**	0.022	0.001	0.022	0.038	-0.013	0.039
HR_{CB}=0.4	**0.1**	0.022	0.000	0.022	0.042	-0.059	0.046
	0.2	0.021	-0.001	0.021	0.044	0.061	0.047
	0.3	0.023	0.001	0.023	0.041	0.009	0.041
	0.4	0.022	0.004	0.022	0.046	-0.007	0.046
	0.5	0.019	0.006	0.019	0.038	-0.015	0.038
HR_{AB}=0.8	**0.05**	0.003	0.002	0.003	0.006	0.015	0.006
HR_{CB}=0.9	**0.1**	0.003	0.003	0.003	0.006	0.002	0.006
	0.2	0.003	0.002	0.003	0.006	0.016	0.006
	0.3	0.003	0.003	0.003	0.006	-0.027	0.006
	0.4	0.003	-0.002	0.003	0.006	0.022	0.006
	0.5	0.003	0.002	0.003	0.006	0.022	0.006
HR_{AB}=0.8	**0.05**	0.004	0.003	0.004	0.007	0.020	0.008
HR_{CB}=0.8	**0.1**	0.004	0.001	0.004	0.009	0.023	0.009
	0.2	0.004	0.003	0.004	0.008	0.031	0.009
	0.3	0.004	0.001	0.004	0.008	0.029	0.009
	0.4	0.004	0.003	0.004	0.008	0.004	0.008
	0.5	0.003	-0.001	0.003	0.008	0.000	0.008
HR_{AB}=0.8	**0.05**	0.005	-0.003	0.005	0.010	-0.010	0.010

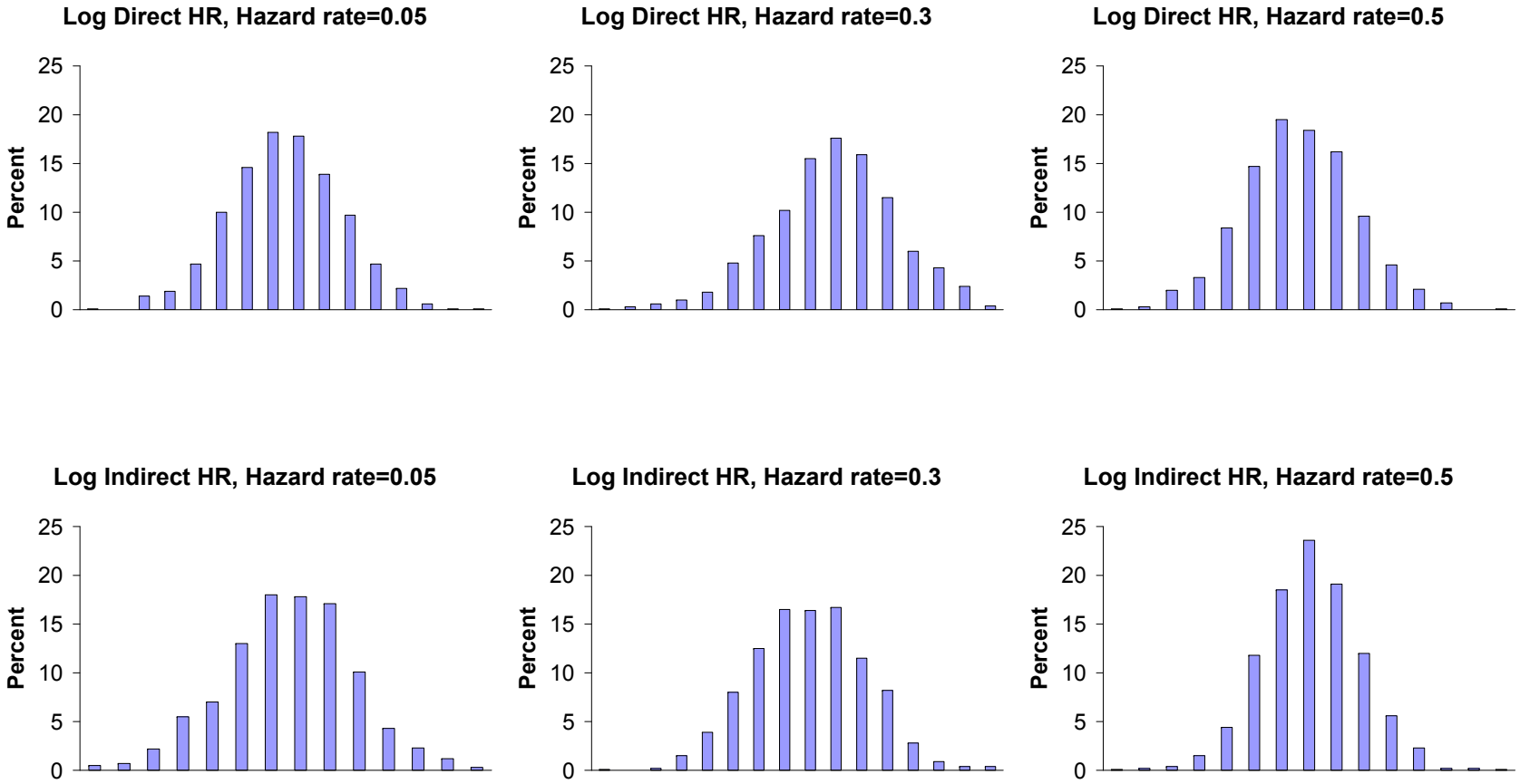
Hazard Ratio Settings	Hazard rate	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
HR_{CB}=0.7	**0.1**	0.005	0.003	0.005	0.010	-0.010	0.010
	0.2	0.005	0.005	0.005	0.010	-0.006	0.010
	0.3	0.005	0.004	0.005	0.010	0.050	0.013
	0.4	0.005	-0.001	0.005	0.009	0.020	0.010
	0.5	0.005	0.004	0.005	0.011	0.011	0.011
HR_{AB}=0.8	**0.05**	0.007	0.004	0.007	0.013	0.016	0.014
HR_{CB}=0.6	**0.1**	0.007	0.002	0.007	0.015	0.000	0.015
	0.2	0.007	0.000	0.007	0.014	0.031	0.015
	0.3	0.007	0.004	0.007	0.015	0.030	0.015
	0.4	0.006	0.006	0.006	0.012	0.011	0.012
	0.5	0.007	0.004	0.007	0.012	-0.004	0.012
HR_{AB}=0.8	**0.05**	0.010	0.009	0.010	0.021	-0.012	0.021
HR_{CB}=0.5	**0.1**	0.010	0.004	0.010	0.020	0.050	0.022
	0.2	0.010	-0.006	0.010	0.024	0.053	0.027
	0.3	0.010	0.007	0.010	0.019	-0.024	0.019
	0.4	0.009	0.007	0.009	0.019	0.042	0.021
	0.5	0.011	0.003	0.011	0.019	-0.039	0.021
HR_{AB}=0.8	**0.05**	0.018	0.014	0.018	0.029	-0.021	0.030
HR_{CB}=0.4	**0.1**	0.016	-0.001	0.016	0.034	0.050	0.036
	0.2	0.017	0.008	0.017	0.029	-0.084	0.036
	0.3	0.017	0.003	0.017	0.035	0.107	0.047
	0.4	0.015	0.007	0.015	0.027	-0.077	0.033
	0.5	0.015	0.010	0.015	0.030	-0.024	0.030
HR_{AB}=0.7	**0.05**	0.002	0.002	0.002	0.006	0.026	0.006
HR_{CB}=0.9	**0.1**	0.002	0.002	0.002	0.004	-0.014	0.005
	0.2	0.002	0.001	0.002	0.004	-0.010	0.004
	0.3	0.002	0.002	0.002	0.005	0.001	0.005
	0.4	0.002	0.001	0.002	0.004	-0.014	0.005
	0.5	0.002	0.001	0.002	0.004	-0.003	0.004
HR_{AB}=0.7	**0.05**	0.003	0.001	0.003	0.006	0.011	0.006
HR_{CB}=0.8	**0.1**	0.003	-0.001	0.003	0.006	-0.010	0.006
	0.2	0.003	0.003	0.003	0.006	-0.006	0.006
	0.3	0.003	0.003	0.003	0.006	0.003	0.006
	0.4	0.003	0.003	0.003	0.005	-0.025	0.006
	0.5	0.003	0.001	0.003	0.006	0.036	0.007
HR_{AB}=0.7	**0.05**	0.004	-0.001	0.004	0.007	-0.017	0.008
HR_{CB}=0.7	**0.1**	0.004	0.002	0.004	0.007	0.009	0.007
	0.2	0.003	0.003	0.003	0.006	-0.002	0.006
	0.3	0.004	0.004	0.004	0.007	0.005	0.007
	0.4	0.004	-0.001	0.004	0.008	0.004	0.008
	0.5	0.004	0.004	0.004	0.007	-0.014	0.007
HR_{AB}=0.7	**0.05**	0.005	0.002	0.005	0.010	0.021	0.011

Hazard Ratio Settings	Hazard rate	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
HR_{CB}=0.6	**0.1**	0.005	0.004	0.005	0.010	-0.014	0.011
	0.2	0.005	0.003	0.005	0.010	0.025	0.010
	0.3	0.005	0.003	0.005	0.010	0.018	0.010
	0.4	0.005	0.001	0.005	0.011	-0.001	0.011
	0.5	0.005	0.004	0.005	0.010	0.011	0.010
HR_{AB}=0.7	**0.05**	0.008	0.004	0.008	0.014	-0.081	0.020
HR_{CB}=0.5	**0.1**	0.007	0.003	0.007	0.017	0.050	0.019
	0.2	0.007	0.001	0.007	0.016	0.007	0.016
	0.3	0.007	0.002	0.007	0.017	0.032	0.018
	0.4	0.008	0.008	0.008	0.016	-0.012	0.016
	0.5	0.007	-0.004	0.007	0.014	-0.035	0.015
HR_{AB}=0.7	**0.05**	0.012	0.003	0.012	0.024	0.018	0.025
HR_{CB}=0.4	**0.1**	0.014	0.012	0.014	0.026	-0.012	0.026
	0.2	0.011	0.004	0.011	0.026	0.064	0.030
	0.3	0.013	0.004	0.013	0.027	0.036	0.028
	0.4	0.013	0.009	0.014	0.025	0.025	0.026
	0.5	0.010	0.004	0.010	0.021	-0.029	0.022
HR_{AB}=0.6	**0.05**	0.002	0.001	0.002	0.004	0.004	0.004
HR_{CB}=0.9	**0.1**	0.002	0.001	0.002	0.003	0.007	0.003
	0.2	0.002	-0.002	0.002	0.004	0.003	0.004
	0.3	0.002	0.002	0.002	0.004	0.018	0.004
	0.4	0.002	0.003	0.002	0.004	0.010	0.004
	0.5	0.002	0.001	0.002	0.003	-0.012	0.004
HR_{AB}=0.6	**0.05**	0.002	0.001	0.002	0.004	-0.022	0.004
HR_{CB}=0.8	**0.1**	0.002	0.000	0.002	0.004	0.015	0.004
	0.2	0.002	-0.002	0.002	0.004	-0.028	0.005
	0.3	0.002	0.000	0.002	0.005	0.048	0.007
	0.4	0.002	0.001	0.002	0.004	-0.016	0.005
	0.5	0.002	-0.001	0.002	0.004	0.012	0.005
HR_{AB}=0.6	**0.05**	0.003	0.001	0.003	0.005	0.023	0.006
HR_{CB}=0.7	**0.1**	0.003	0.005	0.003	0.006	0.000	0.006
	0.2	0.003	0.003	0.003	0.005	0.006	0.005
	0.3	0.003	0.001	0.003	0.006	-0.009	0.006
	0.4	0.003	0.000	0.003	0.005	-0.003	0.005
	0.5	0.003	0.003	0.003	0.006	0.000	0.006
HR_{AB}=0.6	**0.05**	0.004	0.000	0.004	0.008	0.014	0.008
HR_{CB}=0.6	**0.1**	0.003	0.002	0.003	0.007	-0.006	0.007
	0.2	0.004	0.002	0.004	0.008	0.000	0.008
	0.3	0.004	0.002	0.004	0.007	-0.006	0.007
	0.4	0.004	0.003	0.004	0.008	0.007	0.008
	0.5	0.004	0.004	0.004	0.007	-0.028	0.008
HR_{AB}=0.6	**0.05**	0.006	0.003	0.006	0.011	-0.008	0.011

Hazard Ratio Settings	Hazard rate	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
HR_{CB}=0.5	**0.1**	0.005	0.006	0.005	0.009	-0.042	0.011
	0.2	0.005	0.002	0.005	0.012	0.031	0.013
	0.3	0.006	0.002	0.006	0.012	0.035	0.013
	0.4	0.005	0.004	0.005	0.010	-0.042	0.012
	0.5	0.005	0.003	0.005	0.011	0.023	0.011
HR_{AB}=0.6	**0.05**	0.009	-0.001	0.009	0.020	0.066	0.024
HR_{CB}=0.4	**0.1**	0.010	0.008	0.010	0.020	-0.020	0.021
	0.2	0.008	0.004	0.008	0.016	-0.013	0.017
	0.3	0.009	0.006	0.010	0.021	0.062	0.025
	0.4	0.008	0.004	0.008	0.020	0.026	0.021
	0.5	0.008	0.007	0.008	0.018	-0.035	0.019
HR_{AB}=0.5	**0.05**	0.001	0.001	0.001	0.002	-0.009	0.002
HR_{CB}=0.9	**0.1**	0.001	0.001	0.001	0.002	0.002	0.002
	0.2	0.001	0.003	0.001	0.002	0.007	0.002
	0.3	0.001	0.001	0.001	0.003	-0.001	0.003
	0.4	0.001	0.000	0.001	0.002	0.000	0.002
	0.5	0.001	0.001	0.001	0.002	-0.007	0.002
HR_{AB}=0.5	**0.05**	0.002	-0.001	0.002	0.004	0.030	0.005
HR_{CB}=0.8	**0.1**	0.002	0.000	0.002	0.003	0.006	0.003
	0.2	0.002	0.002	0.002	0.003	0.003	0.003
	0.3	0.002	-0.001	0.002	0.003	-0.002	0.003
	0.4	0.002	0.000	0.002	0.003	0.022	0.004
	0.5	0.002	0.002	0.002	0.003	-0.001	0.003
HR_{AB}=0.5	**0.05**	0.002	0.000	0.002	0.005	0.007	0.005
HR_{CB}=0.7	**0.1**	0.002	0.002	0.002	0.004	0.009	0.004
	0.2	0.002	0.004	0.002	0.004	0.006	0.004
	0.3	0.002	0.003	0.002	0.004	-0.005	0.004
	0.4	0.002	0.000	0.002	0.004	0.025	0.005
	0.5	0.002	0.000	0.002	0.004	0.005	0.004
HR_{AB}=0.5	**0.05**	0.003	0.000	0.003	0.006	0.009	0.006
HR_{CB}=0.6	**0.1**	0.002	-0.001	0.002	0.005	0.004	0.005
	0.2	0.002	0.001	0.002	0.006	-0.012	0.006
	0.3	0.002	0.001	0.002	0.005	-0.015	0.005
	0.4	0.002	0.002	0.002	0.005	-0.038	0.006
	0.5	0.003	0.002	0.003	0.006	0.018	0.006
HR_{AB}=0.5	**0.05**	0.004	0.000	0.004	0.009	-0.007	0.009
HR_{CB}=0.5	**0.1**	0.004	0.002	0.004	0.008	0.003	0.008
	0.2	0.004	0.005	0.004	0.007	0.003	0.007
	0.3	0.004	0.008	0.004	0.009	0.002	0.009
	0.4	0.004	0.005	0.004	0.008	0.026	0.009
	0.5	0.004	-0.002	0.004	0.008	0.005	0.008
HR_{AB}=0.5	**0.05**	0.005	0.001	0.005	0.013	0.012	0.013

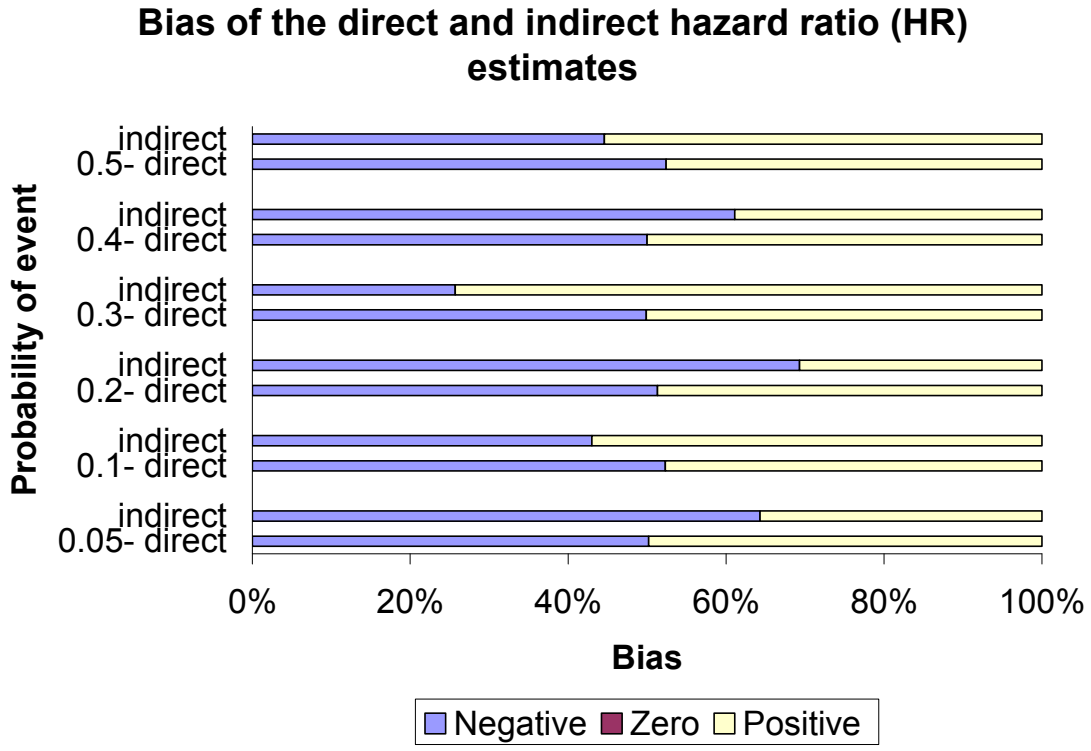
Hazard Ratio Settings	Hazard rate	Direct Estimator			Indirect Estimator		
		Variance	Bias	MSE	Variance	Bias	MSE
HR_{CB}=0.4	**0.1**	0.006	0.007	0.006	0.013	0.012	0.013
	0.2	0.006	0.001	0.006	0.013	0.004	0.013
	0.3	0.006	0.001	0.006	0.013	0.032	0.014
	0.4	0.006	0.002	0.006	0.014	-0.002	0.014
	0.5	0.006	0.004	0.006	0.013	0.010	0.013
HR_{AB}=0.4	**0.05**	0.001	0.000	0.001	0.001	-0.012	0.002
HR_{CB}=0.9	**0.1**	0.001	0.000	0.001	0.002	0.005	0.002
	0.2	0.001	0.000	0.001	0.001	0.002	0.001
	0.3	0.001	0.000	0.001	0.002	0.001	0.002
	0.4	0.001	0.000	0.001	0.002	0.006	0.002
	0.5	0.001	-0.001	0.001	0.001	-0.001	0.001
HR_{AB}=0.4	**0.05**	0.001	0.000	0.001	0.002	0.017	0.003
HR_{CB}=0.8	**0.1**	0.001	0.002	0.001	0.002	0.005	0.002
	0.2	0.001	0.000	0.001	0.002	-0.008	0.002
	0.3	0.001	0.001	0.001	0.002	-0.028	0.003
	0.4	0.001	0.000	0.001	0.002	-0.003	0.002
	0.5	0.001	-0.001	0.001	0.002	0.004	0.002
HR_{AB}=0.4	**0.05**	0.001	0.001	0.001	0.003	0.002	0.003
HR_{CB}=0.7	**0.1**	0.001	0.003	0.001	0.003	-0.008	0.003
	0.2	0.001	0.003	0.001	0.003	-0.005	0.003
	0.3	0.001	0.001	0.001	0.003	-0.027	0.003
	0.4	0.001	0.000	0.001	0.003	0.029	0.004
	0.5	0.001	0.000	0.001	0.003	0.010	0.003
HR_{AB}=0.4	**0.05**	0.002	0.001	0.002	0.004	0.014	0.004
HR_{CB}=0.6	**0.1**	0.002	0.001	0.002	0.004	0.013	0.004
	0.2	0.002	-0.001	0.002	0.004	0.012	0.004
	0.3	0.002	0.002	0.002	0.004	-0.007	0.004
	0.4	0.002	0.002	0.002	0.004	-0.002	0.004
	0.5	0.002	0.001	0.002	0.004	-0.013	0.004
HR_{AB}=0.4	**0.05**	0.003	0.001	0.003	0.006	0.023	0.007
HR_{CB}=0.5	**0.1**	0.002	0.004	0.002	0.005	-0.023	0.005
	0.2	0.003	0.001	0.003	0.005	-0.035	0.006
	0.3	0.002	0.006	0.002	0.005	-0.025	0.005
	0.4	0.002	-0.001	0.002	0.005	0.000	0.005
	0.5	0.002	0.000	0.002	0.006	0.022	0.006
HR_{AB}=0.4	**0.05**	0.004	0.003	0.004	0.008	-0.011	0.008
HR_{CB}=0.4	**0.1**	0.003	0.002	0.003	0.009	0.011	0.009
	0.2	0.004	0.001	0.004	0.010	0.009	0.010
	0.3	0.004	0.000	0.004	0.008	-0.028	0.009
	0.4	0.004	0.002	0.004	0.009	-0.006	0.009
	0.5	0.004	0.005	0.004	0.008	0.004	0.008

Figure B.6.1: Frequency distribution of direct and indirect hazard ratio (HR) estimators for the parameter settings HR_{AB}=0.6, HR_{CB}=0.8, hazard rate in population B=0.05, 0.3, 0.5 (k=3 treatments)



For the parameter settings HR_{AB}=0.6, HR_{CB}=0.8, the direction of the biases over the 1000 samples are illustrated in Figure B.6.2 for the direct and indirect approaches under the different settings of the hazard rate in population B: 0.05, 0.1, 0.2, 0.3, 0.4, 0.5. As noted, although the bias is small, for the direct estimates, the parameter is overestimated and underestimated approximately the same number of times. For the indirect estimates, the parameter is overestimated when the hazard rate in population B is 0.0.5, 0.2, and 0.4, but is underestimated when the hazard rate in population B is 0.1, 0.3 and 0.5.

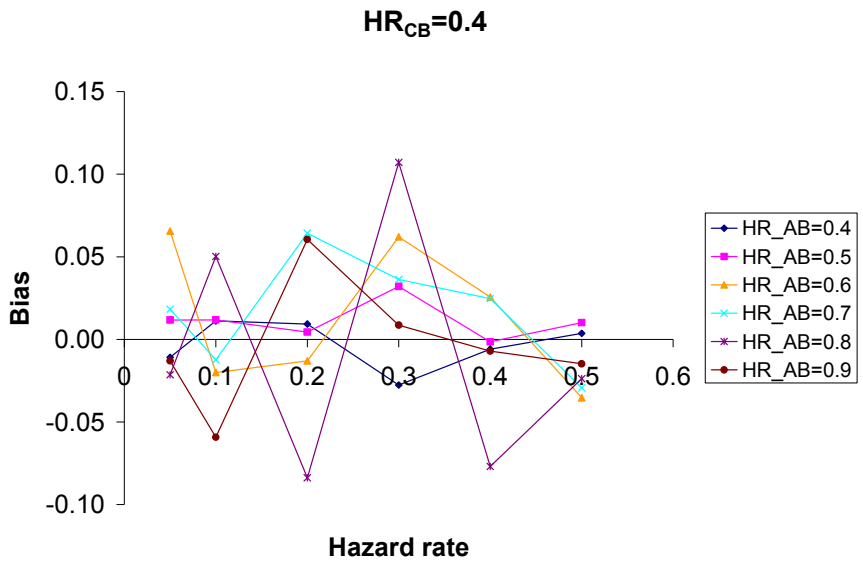
Figure B.6.2: Direction of the bias of the direct and indirect hazard ratio (HR) estimates for the parameter settings HR_{AB}=0.6, HR_{CB}=0.8, hazard rate in population B =0.05, 0.1, 0.2, 0.3, 0.4, 0.5 (k=3 treatments)



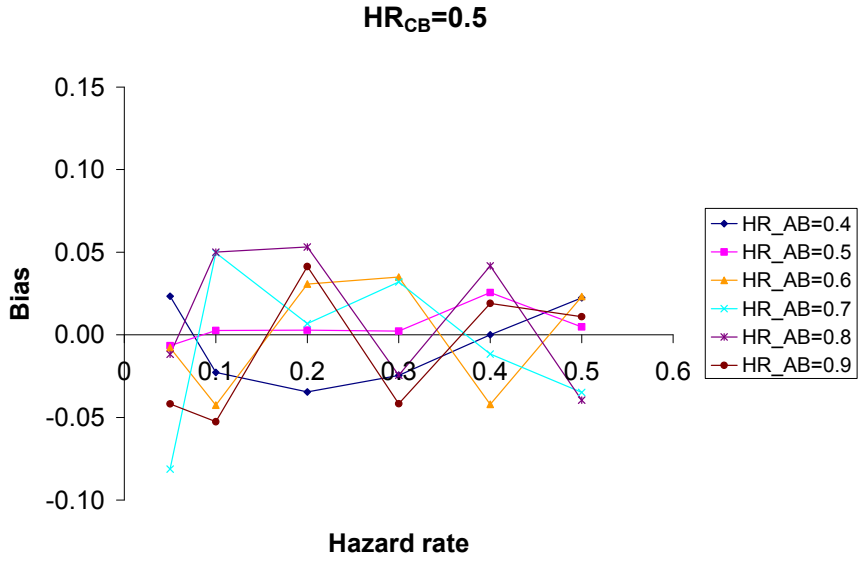
Patterns of bias and MSE for the indirect hazard ratio (HR) estimates are displayed for different settings of HR_{AB}, HR_{CB}, and the hazard rates (Figure B.6.3 and Figure B.6.4). The patterns are shown for the event rate 0.05 to 0.5 and the results are symmetric about 0.5. As such, patterns for event rates ranging from 0.5 to 0.95 are not shown. For each of the settings, the figures are displayed for event rates that start at 0.05. The figures are also displayed for event rates that start at 0.2 in order to improve the resolution of the graphs.

Figure B.6.3: Bias for indirect hazard ratio (HR) estimates for various parameter settings.

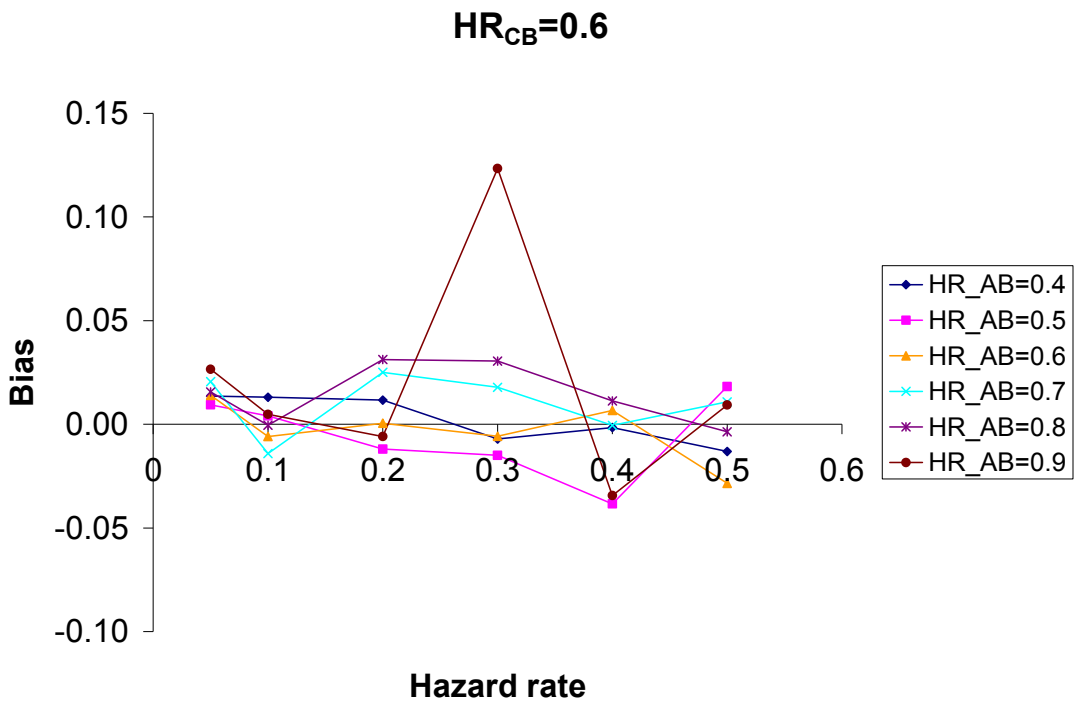
a)



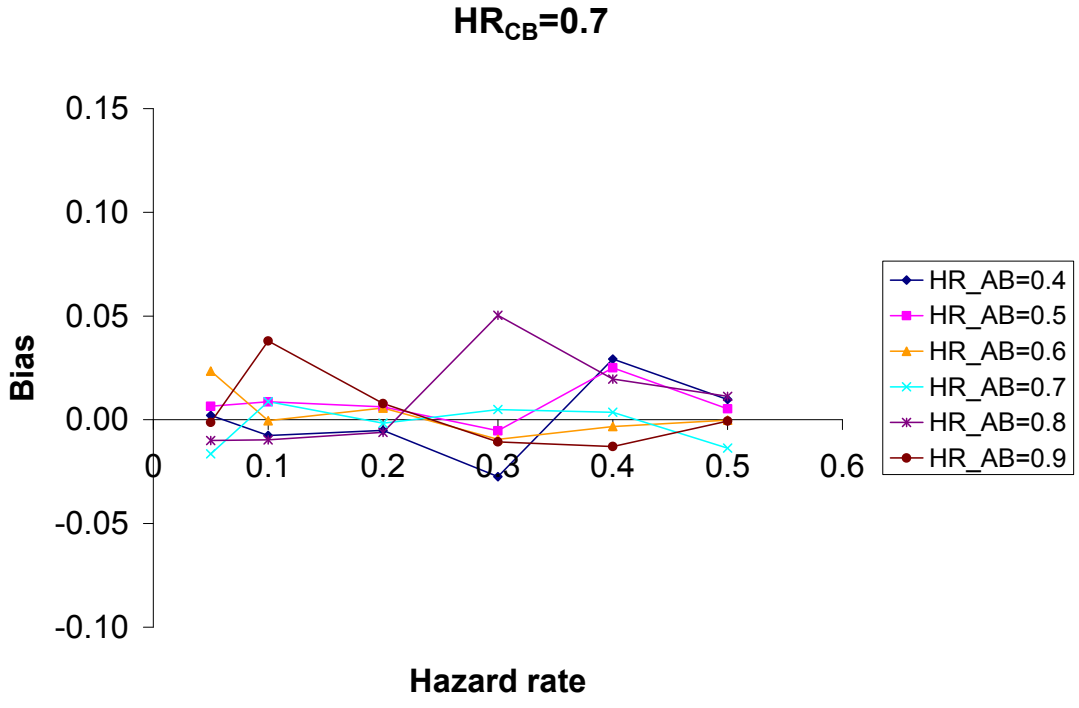
b)



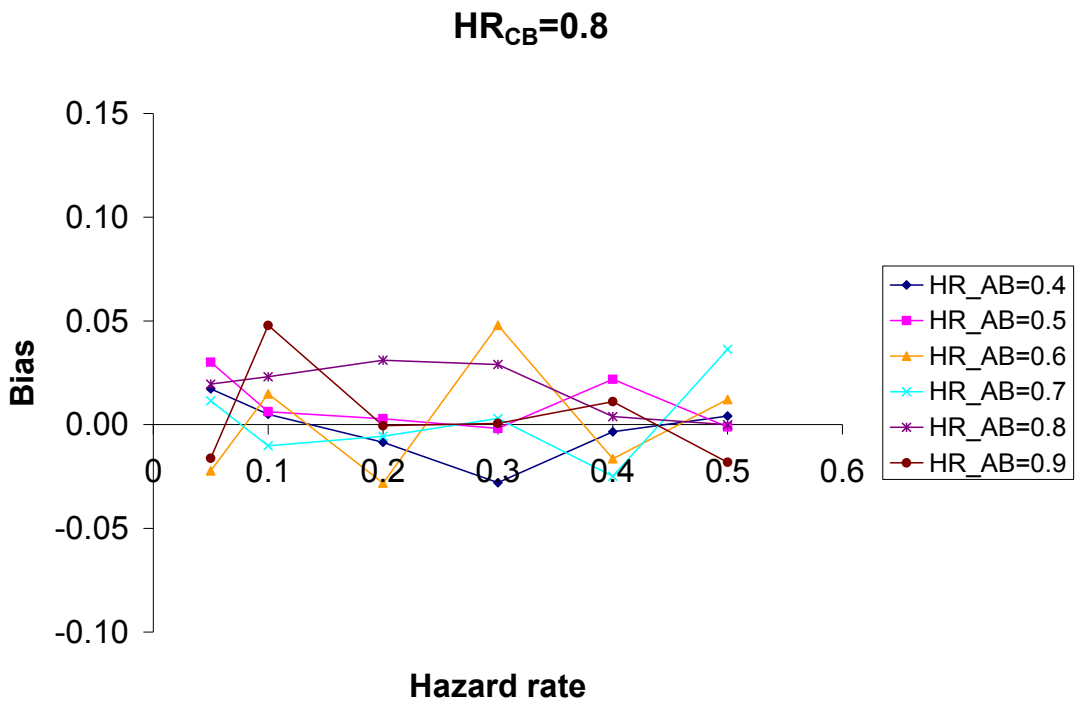
c)


HR_CB=0.6
0.15
0.10
0.05
0.00
-0.05
-0.10
Bias
0
0.1
0.2
0.3
0.4
0.5
0.6
Hazard rate
HR_AB=0.4
HR_AB=0.5
HR_AB=0.6
HR_AB=0.7
HR_AB=0.8
HR_AB=0.9

d)


HR_CB=0.7
0.15
0.10
0.05
0.00
-0.05
-0.10
Bias
0
0.1
0.2
0.3
0.4
0.5
0.6
Hazard rate
HR_AB=0.4
HR_AB=0.5
HR_AB=0.6
HR_AB=0.7
HR_AB=0.8
HR_AB=0.9

e)


HR_CB=0.8
0.15
0.10
0.05
0.00
-0.05
-0.10
Bias
0
0.1
0.2
0.3
0.4
0.5
0.6
Hazard rate
HR_AB=0.4
HR_AB=0.5
HR_AB=0.6
HR_AB=0.7
HR_AB=0.8
HR_AB=0.9

f)

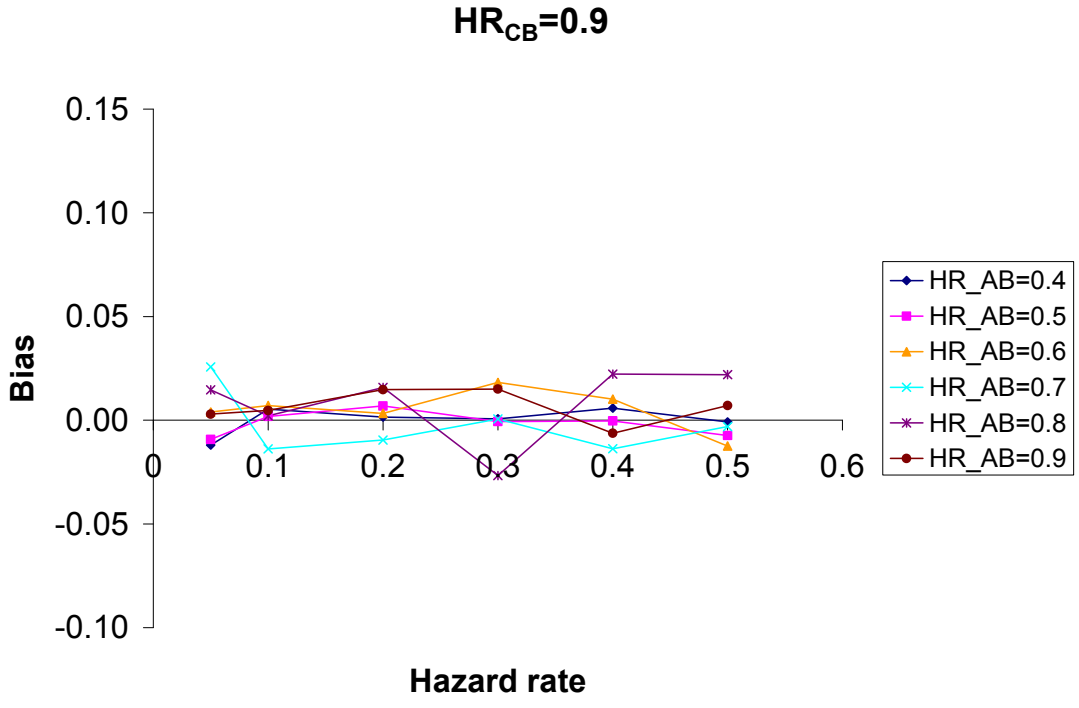
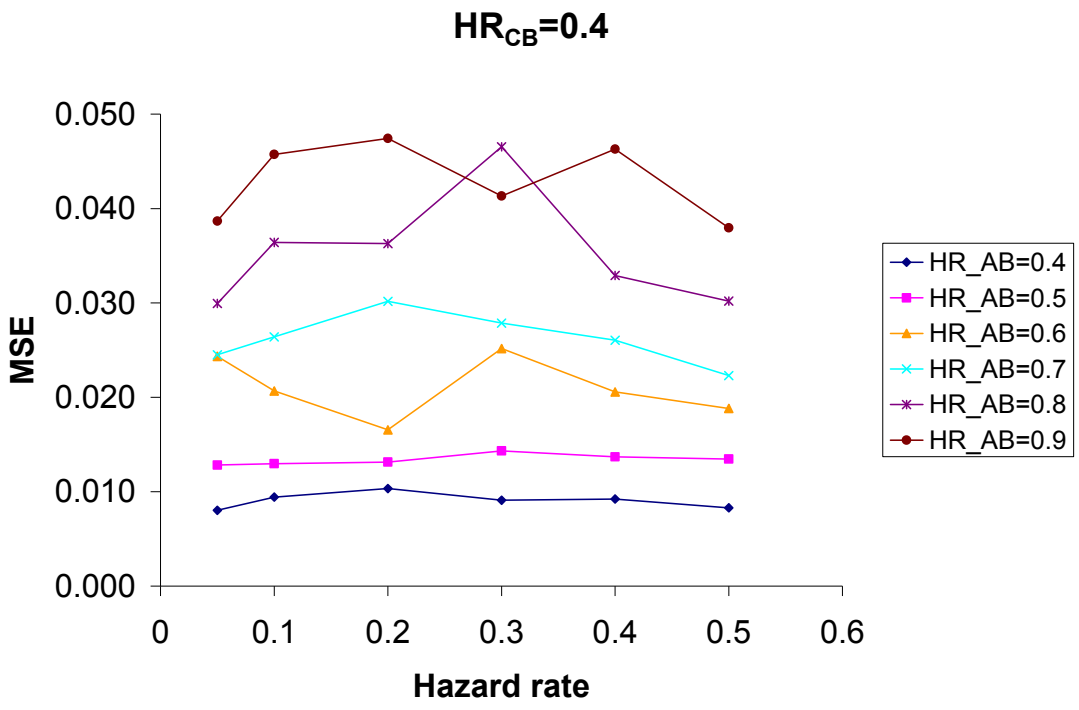
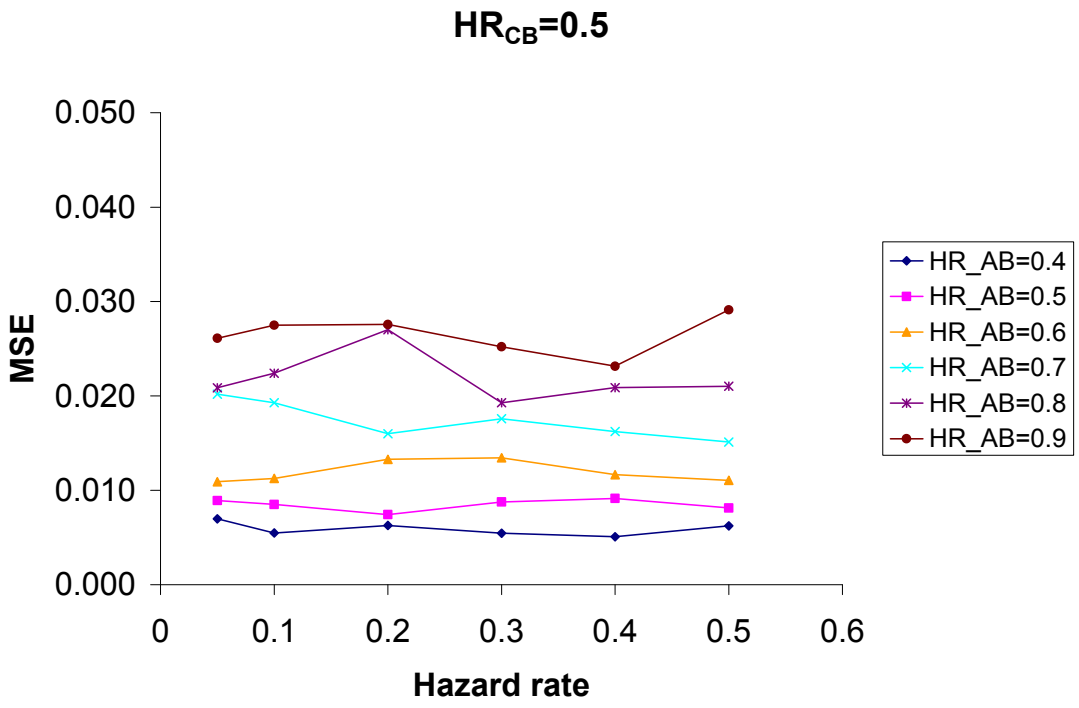

HR_CB=0.9
0.15
0.10
0.05
0.00
-0.05
-0.10
Bias
0
0.1
0.2
0.3
0.4
0.5
0.6
Hazard rate
HR_AB=0.4
HR_AB=0.5
HR_AB=0.6
HR_AB=0.7
HR_AB=0.8
HR_AB=0.9

Figure B.6.4: Mean square error (MSE) for indirect hazard ratio (HR) estimates for various parameter settings

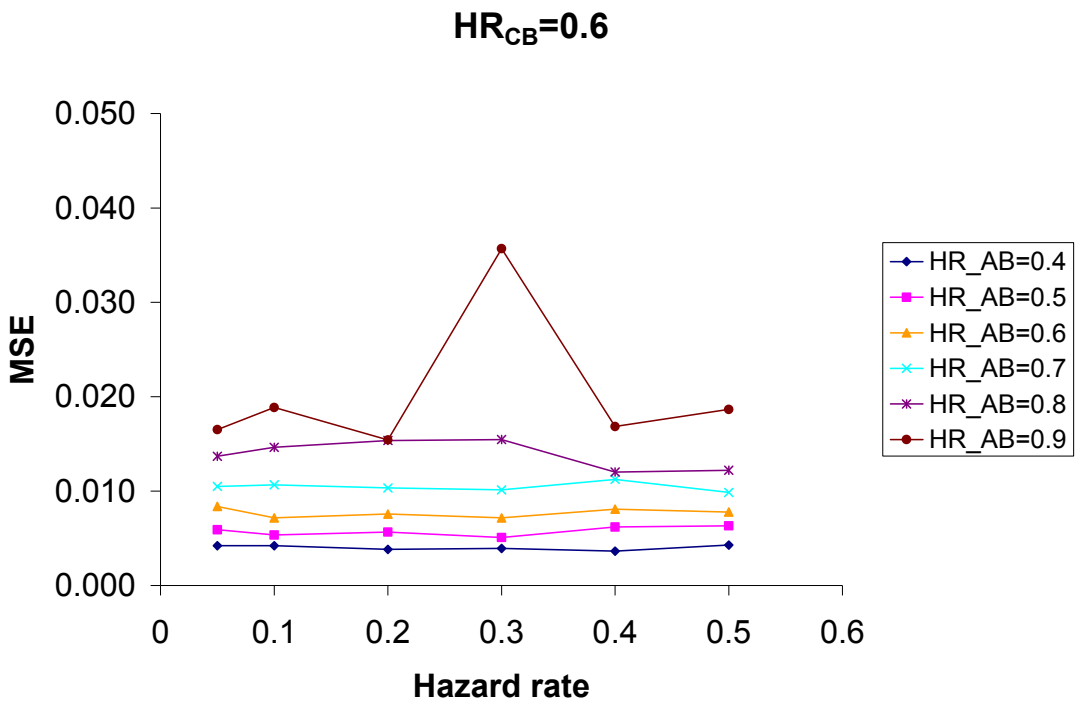
a)



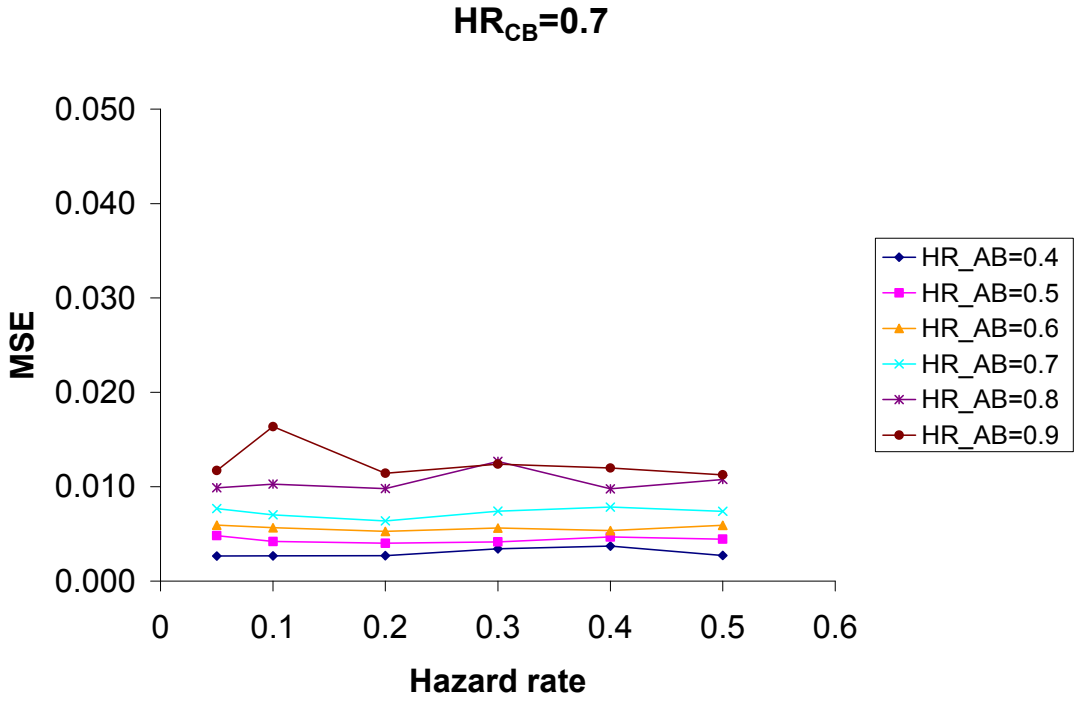
b)



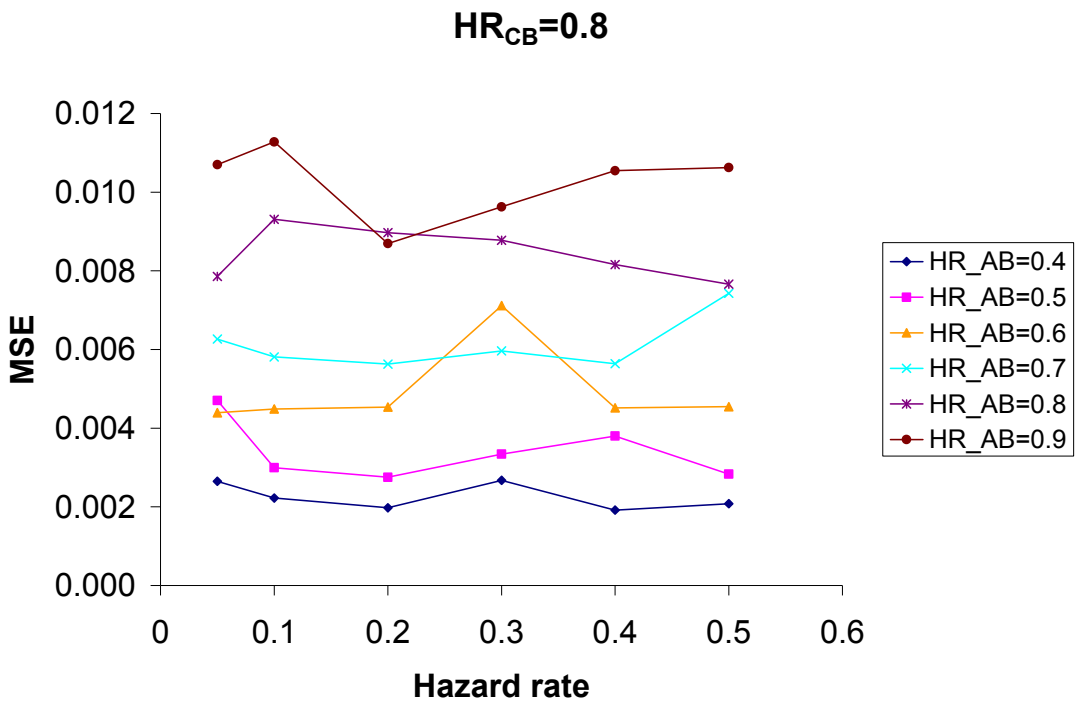
c)


HR_CB=0.6
0.050
0.040
0.030
0.020
0.010
0.000
MSE
0
0.1
0.2
0.3
0.4
0.5
0.6
Hazard rate
HR_AB=0.4
HR_AB=0.5
HR_AB=0.6
HR_AB=0.7
HR_AB=0.8
HR_AB=0.9

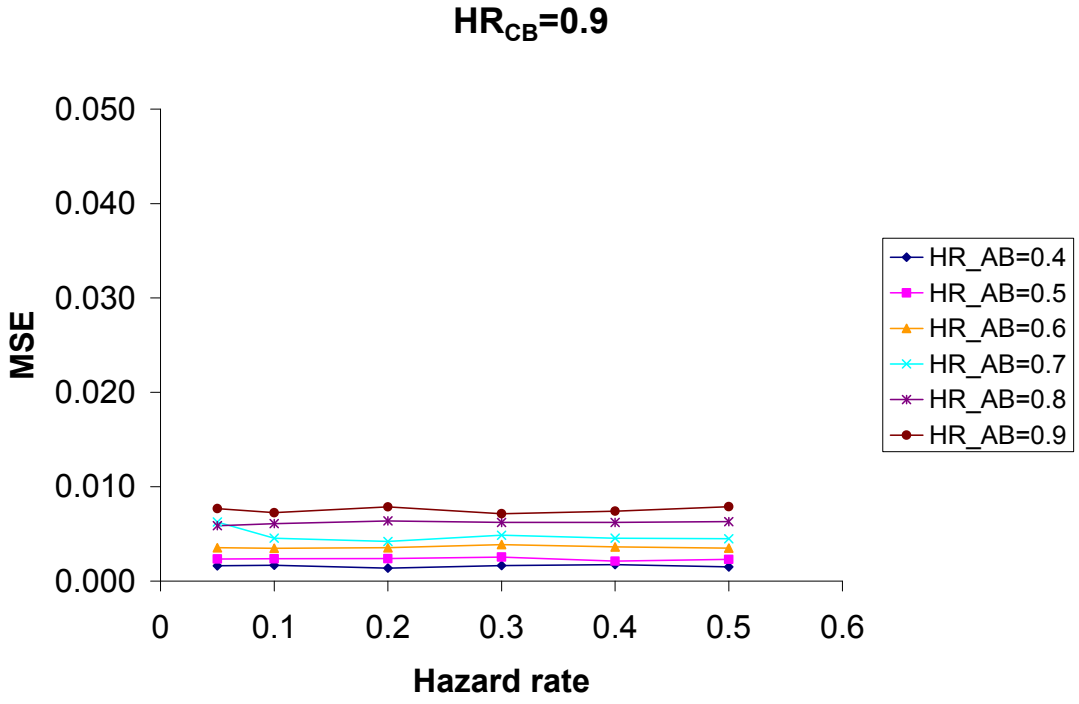
d)


HR_CB=0.7
0.050
0.040
0.030
0.020
0.010
0.000
MSE
0
0.1
0.2
0.3
0.4
0.5
0.6
Hazard rate
HR_AB=0.4
HR_AB=0.5
HR_AB=0.6
HR_AB=0.7
HR_AB=0.8
HR_AB=0.9

e)


HR_CB=0.8
0.012
0.010
0.008
0.006
0.004
0.002
0.000
MSE
0
0.1
0.2
0.3
0.4
0.5
0.6
Hazard rate
HR_AB=0.4
HR_AB=0.5
HR_AB=0.6
HR_AB=0.7
HR_AB=0.8
HR_AB=0.9

f)


HR_CB=0.9
0.050
0.040
0.030
0.020
0.010
0.000
MSE
0
0.1
0.2
0.3
0.4
0.5
0.6
Hazard rate
HR_AB=0.4
HR_AB=0.5
HR_AB=0.6
HR_AB=0.7
HR_AB=0.8
HR_AB=0.9