INTERNATIONAL DIFFERENCES IN THE CLINICAL EFFECTIVENESS OF MEDICAL INTERVENTIONS: A STUDY USING "PANORAMIC" META-ANALYSIS

By

LOUISE CHARLOTTE HARTLEY

A thesis submitted to

The University of Birmingham

For the degree of

DOCTOR OF PHILOSOPHY

School of Health and Population Sciences

College of Medicine and Dentistry

The University of Birmingham

(November/2011)

ABSTRACT

Due to concerns about international differences in treatment effectiveness, many countries are reluctant to extrapolate overseas clinical data to form the basis of guideline recommendations and intervention approval processes. The evidence on which these concerns are based, however, comes from a limited dataset, with few studies directly assessing international differences in treatment effectiveness. This study aims to assess differences in the results of cardiovascular trials between Europe, North America, and Asia using the panoramic meta-analysis approach.

All meta-analyses containing randomised control trials for the treatment or prevention of cardiovascular diseases were searched for in The Cochrane Library (2000 to 2008) and Medline (2005-2008). Analysis was then conducted within and over the included meta-analyses by performing pair-wise comparisons of the trial results between Europe and North America, Europe and Asia, and North America and Asia and a universal comparison of all three continents' trial results together. All analyses were conducted over fatal and non-fatal endpoints.

The findings suggested that for both endpoints, interventions performed best in Asian trials. For fatal endpoints, a high proportion of positive trial results were observed for Japan. Further investigation showed that between-continent differences in treatment effect could be explained by between-continent differences in trial quality. However, the types of intervention prone to inter-continental differences could not be identified for fatal or non-fatal endpoints.

These findings suggest that those developing guidelines and approving interventions should be cautious when extrapolating overseas data. In particular, this study highlights the importance of taking trial quality into account when extrapolating and interpreting clinical trial data from different regions.

Acknowledgements

I would like to take this opportunity to thank the many people that have contributed to the development and completion of this research. First of all, I would like to express sincere thanks to my supervisors, Professor Richard Lilford, Dr Russell Bowater, and Mr Alan Girling for their invaluable help, guidance, and support throughout the research process and towards the successful completion of this thesis. Thanks must also be extended to the Multidisciplinary Assessment of Technology Centre for Healthcare (MATCH), without whose financial award I would not have been able to begin or continue this journey.

Thanks must go to my fellow PhD students, who shared their experiences, offered moral support, and encouraged me along the way. Thank you also to my friends both inside and outside academia, who kept me grounded and reminded me of the important things that life has to offer, especially (in no particular order) Paul, David, Robert, Gregory, Ben, Matthew, Laila, Nicky, and Michelle.

Finally, and most importantly, I would like to thank my family: Craig, Jo, Andy, Fiona, Billy, Gordon, Chris, Pete, and especially Mum. Thank you for always believing in me and loving me. I dedicate this thesis to them.

Authorship in each chapter

Chapter 1: Introduction
Designed by LH, written up by LH, comments by AG

Chapter 2: Background
Designed by LH, written up by LH, comments by AG, RB

Chapter 3: The Methodology of Panoramic Meta-analysis
Designed by LH, written up by LH, comments by AG

Chapter 4: Investigation Methods: Descriptions and Explanations
Designed by LH, written up by LH, comments by AG

Chapter 5: An Inter-continental Comparison of Treatment Effectiveness between Europe and North America
Designed by LH, RB, data Collected by LH, analysed by LH, RB, written up by LH, Comments by RB, AG

Chapter 6: Inter-continental Comparisons of Treatment Effectiveness for Europe, North America and Asia
Designed by LH, RB, AG data collected by LH, analysed by LH, written up by LH, comments by AG

Chapter 7: Trial Quality Investigation
Designed by LH, RL, data collected by LH, UN, analysed by LH, written up by LH, comments by AG

Chapter 8: Discussion and Conclusion
Designed by LH, written by LH, comments by AG

Contents

LIST OF TABLES, FIGURES, AND BOXES.

List of Tables.

List of Figures

List of Boxes

LIST OF ABBREVIATIONS

Medical Abbreviations

ACE = Angiotensin Converting Enzyme

AF= Atrial Fibrillation

ACEI = Angiotensin-Converting Enzyme Inhibitors

ACS = Acute Coronary Syndrome

ARB = Angiotensin Receptor Blockers

CABG = Coronary Artery Bypass Grafting

CAD = Coronary Artery Disease

CHD = Coronary Heart Disease

CHF = Congestive Heart Failure

CRT = Cardiac Resynchronisation Therapy

CVDs = Cardiovascular Diseases

EBM = Evidence Based Medicine

G-CSF = Granulocyte Colony-Stimulating Factor

HF = Heart Failure

LMWH = Low Molecular Weight Heparin

LVSD = Left Ventricular Systolic Dysfunction

MI = Myocardial Infarction

MIDCAB = Minimally Invasive Direct Coronary Bypass Grafting

NPPV = Non-invasive Positive Pressure Ventilation

PCI = Percutaneous Coronary Intervention

PE = Pulmonary Embolism

PES = Paclitaxel Eluting Stent

RCT = Randomised Controlled Trial

SES = Sirolimus Eluting Stent

STEMI = ST-Segment Elevation Myocardial Infarction

UFH = Unfractionated Heparin

Statistical Abbreviations

RR = Risk Ratio

MD = Mean Difference

OR = Odds Ratio

SMD = Standardised Mean Difference

CHAPTER 1.
INTRODUCTION

1.1: The Research Context

Over recent years, national and international communities of health policy-makers, policy-analysts, and healthcare professionals have called for medical and healthcare interventions to be based on robust evidence about an intervention's effectiveness. With national governments and international agencies expressing increased concern over healthcare provision and costs, an evidence-based approach to healthcare planning and delivery has become increasingly important. These concerns over effectiveness and costs have led to the development of 'Evidence Based Medicine' (EBM), an approach that involves systematically examining and appraising the findings of international clinical research to help achieve optimum levels of clinical care. Indeed, EBM is now considered one of the "gold standard tools" for assessing healthcare (Belsey, 2009).

In a bid to optimise healthcare and provide efficacious interventions, EBM has been used for developing clinical guidelines on the most appropriate treatment for patients. Given that these guidelines should be built on the best available international evidence, it is reasonable to expect that clinical guidelines in different countries would be based on similar studies and, therefore, that a country's clinical guidelines would have similar recommendations for practice to those in another country. However, this is not the case. Numerous studies have shown that countries tend to base their clinical guidelines on trials conducted within their own country (Burgers et al.

2002). Even when evidence from other countries is used in guideline development, the resulting recommendations still tend to differ between countries (Eisinger et al. 1999). Many reasons may account for these differences. For instance, guideline recommendations may differ due to cultural differences in the perception of risk, different interpretations of the same evidence, and honest differences in the opinion of experts (Magill & Shlim, 2011).

The effectiveness of the intervention may also account for some of this variability. Effectiveness is the degree to which an intervention has achieved its desired effect for those to whom it is offered (Abramson & Abramson, 1999). While guidelines may differ because of 'local' issues such as socio-economic, cultural, and geographical considerations or regional/national differences in intervention costs, it is also possible that they differ because of concerns about differences in treatment effectiveness between countries. Believing that the effectiveness of interventions differs between countries, even when factors such as culture or socio-economic differences are excluded, may make those developing clinical guidelines more sceptical about recommending treatments based on non-native clinical data. Indeed, the European Medicines Agency (EMA) (2006) asserts that concerns about international differences in treatment effectiveness deter many countries from relying on overseas[1] clinical data.

Although there is some evidence to suggest that treatment effectiveness may differ between countries (Pan et al. 2005; Zhang, Freemantle, & Cheng, 2011), much of this evidence derives from multinational trials whose primary aim was not the examination of inter-country differences

[1] Data from another country or continent.

in treatment effectiveness. Therefore, the difference in results may simply be an artefact of variation in treatment management strategies or implementation of the research protocol in the different countries involved in the study.

It is possible that international differences in treatment effectiveness may only exist for particular types of intervention. This could have important implications for guideline development because if intervention type A is not prone to such differences, then data regarding that intervention's effectiveness can simply be extrapolated from one country to another. Conversely, if intervention type B is known to be prone to international differences, those approving these types of interventions will know that a duplicate trial may be needed before a recommendation can be made. Data showing effectiveness differences by type of intervention could contribute to international guidance about which interventions data could be directly extrapolated and which might need nation-specific investigation.

With the above in mind, research is required that specifically explores international differences in effectiveness which is based on a robust method. This will provide countries with a strong foundation on which to base decisions about extrapolating overseas clinical data to national guidelines and for use in intervention approval processes. This study aims to provide this information for cardiovascular diseases.

1.2: Research Aims

The focus of this study is to investigate the existence of international differences in treatment effectiveness. There are two aims:

- To assess whether international differences in the effectiveness of cardiovascular interventions exist.
- To identify whether there are certain types of intervention that are more likely to exhibit differences in effectiveness between countries.

The following research questions have been posed to help achieve these aims:

- *Does treatment effectiveness for cardiovascular conditions differ between regions?*
- *What types of cardiovascular intervention are more likely to show these differences in treatment effect?*

These research questions will be explored using panoramic meta-analysis.

1.3: Panoramic Meta-Analysis

Panoramic meta-analysis is based upon the procedures and principles of standard meta-analysis and follows a systematic process within which relevant studies are identified, evaluated, and combined to give an overall result (see Chapter 3 for methodological description). Developed by Bowater, Stirling and Lilford (2009), panoramic meta-analysis enables the testing of generic hypotheses[2] and also allows for bias and heterogeneity within large datasets to be examined. Panoramic meta-analysis is the most appropriate method to answer the research questions for the following reasons:

- it enables the exploration of international differences in treatment effectiveness by specifically scrutinising heterogeneity within datasets
- it allows for trial quality to be investigated as a possible reason for such differences
- it allows for different interventions to be explored in the same investigation

[2] 'Generic' in the sense that they can apply to all members of a family of interventions or to a group of different but related medical conditions.

- it is a more efficient approach of assembling data compared to current systematic reviewing and meta-analysis methods.

The benefits of using panoramic meta-analysis are that:

- it allows for the collection of a vast amount of information about relevant trials faster than standard meta-analysis by identifying trials from meta-analyses
- it can provide a broader overview than can standard meta-analytic techniques, by testing hypotheses (over meta-analyses) about different interventions for different but related health conditions
- the required data for relevant studies can simply be extracted from meta-analyses rather than from individual studies
- it uses the robust methods of standard meta-analysis for data analysis.

1.4: Cardiovascular Diseases

Cardiovascular diseases (CVD) (i.e. disorders affecting the heart and blood vessels) were chosen as the focus for this study because they are of worldwide interest and a variety of intervention types for treatment and/or prevention are available.

Cardiovascular diseases are both a national and international concern, as although many CVDs are preventable or can be treated, it is estimated that 17.1 million people worldwide die of a CVD annually (World Health Organisation, 2010). In many countries, it is the leading cause of death (Gaziano et al. 2006). Therefore, it is important to establish appropriate interventions for the prevention and/or treatment of CVDs that can be used on the international stage. Indeed, many countries conduct extensive research on CVD treatment and prevention (Neal, Chapman, & Patel 2002). Thus, evidence on CVD interventions is available from many countries, and it is therefore an ideal topic for this study.

This investigation is important as it has the potential to impact upon the speed at which new CVD interventions are implemented and approved by aiding decisions on the extrapolation of overseas clinical data. Furthermore, by investigating whether the effectiveness of CVD interventions differs between countries, this investigation may impact upon the types of intervention that are recommended for CVD treatment both at national and international levels.

1.5: The Structure of the Thesis.

The study is organised into eight chapters divided into three parts. **Part One** focuses on the methodology and data collection, **Part Two** on the statistical analyses and results of the investigation and **Part Three** on the overall conclusions that can be reached.

Chapter 1 provides the introduction to the study. It presents an overview of the study area and explains why the research is important. The introduction provides a preliminary insight into the research context, problems, questions and the methodological approach to be deployed in the thesis.

PART ONE: There are three chapters in Part One organised under the heading "Methodology and Data Collection" (Chapters 2 to 4).

Chapter 2 provides the background to the thesis by providing the evidence for international differences in treatment effectiveness and evaluates this evidence.

Chapter 3 provides a description of panoramic meta-analysis and details of why and how panoramic meta-analysis was used. The chapter also discusses the issues of bias and heterogeneity.

Chapter 4 describes how the study was conducted. It details identification of evidence as well as the results from the literature search. It also provides an explanation of the methods used to analyse and synthesise the data.

PART TWO: There are three chapters in Part Two that are organised under the heading "The investigation of international differences in treatment effectiveness" (Chapters 5 – 7).

Chapter 5 provides the results from the investigation into treatment effect differences between Europe and North America. It draws together the results from pair-wise comparisons as well as global estimates of continental difference over both fatal and non-fatal endpoints.

Chapter 6 provides the results from the investigations concerning Europe, North America, and Asia. It provides the results from the pair-wise comparisons conducted between Europe and Asia and between North America and Asia and provides the results of the global estimates of continental difference. It also provides a universal comparison of international differences in treatment effect that analyses all regions together and reports the investigation into the types of intervention prone to international differences in treatment effect.

Chapter 7 provides a description of the methods used to assess trial quality and reports the results of this assessment.

PART THREE: There is one chapter within Part Three organised under the heading "Evaluation and Conclusion". This is Chapter 8.

Chapter 8 integrates the results from all investigations to show whether international differences in treatment effectiveness for cardiovascular diseases exist and the extent of any differences. The chapter provides the limitations and strengths of the thesis before providing suggestions on future research.

1.6: The Importance of the Study

Through the use of panoramic meta-analysis, this thesis advances beyond the existing literature and directly investigates the existence of international differences in treatment effectiveness. Methodologically, this study applies a novel research method and has advanced beyond previous studies that have used panoramic meta-analysis in two main ways. First, it has been used to critically compare the clinical effectiveness of CVD interventions between regions. Second, panoramic meta-analysis was enhanced in this study by examining differences in trial quality. Empirically, this study investigates directly the existence of international differences in the clinical effectiveness of CVD interventions and also examines differences over a variety of intervention types. Combining these two approaches provides a novel investigation into international differences. This research could, therefore, have important implications for decisions and recommendations made by both national and international agencies about data extrapolation and guideline development. As such, this study has importance for policy makers, practitioners, and researchers and would be of interest to national and international agencies involved in producing guideline recommendations. It would also be useful for those who

extrapolate (or are sceptical about extrapolating) overseas clinical data from one country to another as well as to those conducting multinational trials by bringing to the forefront the importance of considering international differences in treatment effect when analysing clinical data.

PART ONE: METHODOLOGY AND DATA COLLECTION

CHAPTER 2.
BACKGROUND

Several studies looking at clinical guideline recommendations between countries have recently been conducted. They conclude that guideline recommendations differ between countries and that recommendations are commonly based upon evidence generated in the country producing the guideline (Burgers et al. 2002; Matthy et al. 2007). One would not expect this given the advent of EBM and technological advances allowing global accessibility to trial databases and clinical trial results. While many explanations have been proposed for these differences, it is possible that disparities in guideline recommendations result from the belief that the effectiveness of interventions differs between countries (European Medicines Agency, 2006); a possible explanation for why recommendations are usually based upon trials conducted in the country constructing the guideline. This thesis will address this issue by examining whether inter-country differences in treatment effectiveness exist.

This chapter provides an overview of the available evidence about international differences in clinical effectiveness. It is organised into four main sections. It begins by exploring EBM and its impact upon clinical guideline recommendations. Section two discusses the reasons why recommendations differ between countries. Section three explores evidence of international differences in treatment effectiveness and the factors that might have contributed to these

findings. The chapter concludes by summing up the evidence and explaining why this study is important for national and international healthcare systems and the people who benefit from them.

2.1: Evidence Based Medicine and Clinical Guidelines

Although the practice of EBM is relatively new, the concept behind it is well established. Professor Archie Cochrane's book "Effectiveness and Efficiency: Random Reflections on Health Services" (Ashcroft, 2004) is widely recognised as the catalyst for EBM. Since then, the concept and its development into a practical methodology have become accepted worldwide, with EBM now considered one of the "gold standard tools" in assessing health care (Belsey, 2009) and helping clinicians decide on the best treatment available for their patients. EBM is described by Sackett (2000) as "the integration of the best research evidence with clinical expertise and patient values" and involves systematically examining and appraising the findings of clinical research to ensure that patients receive optimum clinical care. It influences not only clinical practice but also health policy, medical education, and patient information (Straus & Jones, 2004) and is often seen as objective (Turkelson & Hughes, 2006) in that its recommendations and results are based on the highest quality research from around the world (Rosenberg & Donald, 1995).

EBM has brought about an increase in the formulation of evidence based clinical guidelines. These are systematically developed recommendations for practitioners and patients on the appropriate treatment and care for specific medical conditions (Field & Lohr, 1990). They provide a guide to best practice and can be used as a standard against which clinical practice can be assessed. Moreover, they provide a framework within which clinical decisions can be made

(Turner et al. 2008). In the construction of such guidelines, comprehensive literature searches are conducted and the quality of the evidence is assessed. Patients' preferences and values are also considered. These processes ensure that the recommendations made in such guidelines are based upon the best available evidence (Lim et al. 2008) and, by including patient views, are likely to be acceptable to people affected by the condition.

With this in mind, and with increased access to clinical trial results, it might be expected that guideline recommendations, and the evidence on which they were based, would be similar worldwide. However, this is not the case. A number of studies have found that both guideline recommendations and the studies on which they are based differ between countries (Eisinger et al. 1999; Matthy, DeMeyere, van Driel, & DeSutter, 2007). One such study, conducted by Burgers et al (2002),examined diabetes guidelines in thirteen countries to explore whether they were based on similar research. They concluded that the evidence on which the guidelines were based differed between countries because many of the countries based their guidelines on evidence that had originated in their own country. Further evidence is provided by Matthys et al (2007) who found that guideline recommendations regarding the use of rapid antigen tests and throat cultures differed between North America and Europe; the use of the rapid antigen test was recommended in most US, French, and Finnish guidelines but not in Belgian, Dutch, or Scottish guidelines. Throat cultures were advised within US, Canadian, and Finnish guidelines but not in those of Belgium, France, Scotland, England, or the Netherlands.

Why, then, do guideline recommendations and the evidence on which they are based differ between countries?

2.2: Reasons for Global Differences in Clinical Guideline Recommendations

Many explanations have been provided to explain the differences in clinical guidelines between countries, including methodological and interpretation issues and issues linked to effectiveness.

The differences in guideline recommendations found between countries may be explained by different interpretations of the evidence (Fahey & Peters 1996; Ravago, Mosniam, & Alem 2000; Vogel et al. 2000). However, others suggest that the between-country differences can be explained by the use of unsystematic guideline development methods (Matthy, DeMeyere, van Driel, & DeSutter, 2007). As there is no international standardised procedure for the construction of guidelines, the methods used to produce them are likely to differ between countries, which may result in different evidence being found and different recommendations being made. Thomson, McElroy, and Sudlow (1998) found that not all the guidelines they examined showed clear links between the quality of the supporting evidence and the recommendations. Furthermore, they found that in the majority of cases, literature reviews and the appraisal of evidence on which recommendations were based were unsystematic, as were methods for incorporating evidence and opinions. They conclude that this results in guidelines that are different not only in content, but also in impact.

Others, however, propose that between-country disparities could be due to the relative influence of professional bodies over guideline development (Burgers et al. 2002; Matthy, DeMeyere, van Driel, & DeSutter, 2007). Littlejohns et al (1999) found that the views of the Royal College of General Practitioners and the Royal College of Physicians strongly influenced recommendations

made in nine UK guidelines on the treatment of depression in primary care. After comparing national diabetes guidelines in 13 countries, Burgers et al (2002) agree. They found that all, with the exception of Scotland and England, were strongly influenced by the American Diabetes Association. As such, it is likely that the guidelines in these two countries differ from those in the other 11 countries.

Studies suggest other factors such as patient preference (van Driel et al. 2006), socio-economic issues, characteristics of healthcare systems (DeMaeseneer & Derese, 1999; Eccles 2003), and culture (Christiaens et al. 2004) may offer some explanation for international variation in guideline recommendations. Fervers et al (2006), for example, claim cultural differences between countries may create legitimate variations in clinical guidelines, even when based on the same evidence. Eisinger et al.'s (1999) study exemplified this, proposing cultural differences in the symbolic and aesthetic values placed upon breasts and the emphasis placed on patient control and autonomy as possible explanations for guideline differences between the US and France on the clinical management of women with an inherited predisposition to ovarian and breast cancer.

It may be the case that clinical guideline recommendations differ due to the belief of those developing guidelines that treatment effectiveness differs between countries. This would make countries more sceptical about recommending treatments based on overseas clinical data and using such data as the basis for their guidelines. Indeed, the European Medicines Agency (EMA) (2006) suggests that clinical effectiveness is an important factor to consider as it impacts on the extrapolation of clinical trial results between countries. According to the EMA, concerns about international differences in clinical effectiveness have limited the willingness of many countries

to rely on overseas clinical data. This has led to many countries requesting that duplicate studies be conducted before approving a new treatment, thus increasing the cost of such treatments and delaying their approval. But to what extent does evidence support the belief that differences in treatment effectiveness exist between countries?

2.3: Evidence of International Differences in Treatment Effectiveness

Vickers et al.'s (1998) research, one of the few studies directly assessing the existence of international differences in treatment effectiveness, found that studies conducted in Eastern Asia or Eastern Europe had a higher proportion of positive[3] results than those conducted in other countries, which suggests that interventions were more effective in these regions. Pan et al. (2005) found that Chinese trials on human genome epidemiology showed significantly more positive results than non-Chinese studies. A more recent study by Zhang, Freemantle, and Cheng (2011) also found that Chinese trials produced more positive results than those conducted in India or Western countries, suggesting that interventions were more effective in Chinese trials than in the other regions.

An early piece of direct evidence for differences in clinical effectiveness between countries examined isoniazid (INH) preventive therapy for Tuberculosis in seven Eastern European countries (Thompson, Snider and Farer, 1985). This trial showed that variation between countries was larger than variation within countries for all the variables being studied. For instance, one

[3] A finding that the intervention was more effective than its comparator

country had an incidence rate of drug-related adverse reactions that was almost 50% greater than the other six countries. However, they argue that these differences might result from differences in healthcare practice, behaviours, and methodological factors.

Other studies, however, have found little evidence that treatment effect differs between countries. Chang et al (2005) investigated whether inter-country differences in the following outcomes –one year all-cause mortality, thirty-day death, or post-admission myocardial infarction (MI) –could be explained by patient, hospital and country-level factors by using data from the Global Utilisation of Strategies to Open Occluded Coronary Arteries IV Acute Coronary Syndromes (GUSTO IV ACS) trial, which involved 7800 patients from 458 hospitals in 24 countries. They found that 96-99% of total variance in the results could be explained by patient-level factors rather than hospital or country-level factors. It is possible, therefore, that the inter-country differences found in previous studies are the consequence of patient-level differences rather than country-level differences.

Because studies assessing international differences in therapeutic effectiveness are scarce, evidence of these differences must be looked for in the results of multinational trials. Although the investigation of country differences in treatment effect is not the primary aim of these trials, evidence of such disparities can be found by examining the differences in outcomes between the countries studied. It should be noted, however, that while disparities in outcomes may not necessarily represent differences in treatment effect, because they may also apply to the control group, disparities in effectiveness will certainly lead to such differences in outcome. Therefore, country differences in outcome can provide *prima facie* evidence of an interaction between

country and treatment effect. Table 2.1 provides the details of the seven multinational trials that provided evidence of international differences in treatment effect.

Table 2.1: Multinational Trials That Provided Evidence of International Differences in Treatment Effectiveness.

Author	Name of Trial	Acronym	Aims of Study	Regions Involved	Findings	Explanations for Findings
Senn and Harrell (1997) O'Shea and DeMets (2001)	Beta-Blocker Heart Attack Trial	BHAT	Evaluate beta-blockers as the immediate treatment for MI patients	US and Canada	A difference in the direction and size of treatment effect between countries	Random variation (Senn and Harrell, 1997) Small sample sizes in some trial centres' (O'Shea and DeMets, 2001)
O'Shea and Califf (2001)	Flolan International Randomised Survival Trial	FIRST	Evaluate effectiveness of Epoprostenol	US, Canada and Europe	Difference in treatment effect between North America and Europe	Due to patients in Europe receiving the placebo having improved performance
O'Shea and Califf (2001)	Platelet IIb/IIIa in Unstable Angina: Receptor Suppression Using Integrilin Therapy	PURSUIT	Evaluate effectiveness of Eptifibatide	North America, Western Europe, Eastern Europe and Latin America	Treatment benefit was greater in North America. Fewer composite events in Western Europe for Eptifibatide group than in Latin America and Eastern Europe	Different baseline demographics Different adjunctive treatment strategies Different definitions of MI (Akkerhuis, Deckers, Boersma et al, 2000)
Van de Werf, Topol, Lee et al (1995)	Global Use of strategies to open Occluded Coronary Arteries	GUSTO	Evaluate Effect of 4 thrombolytic strategies on mortality rates	US, Australia, Belgium, Canada, France, Germany,	In non-US patients, the combination strategy worked as well as	Country differences in the rate of Haemorrhagic stroke (O'Shea and

				Ireland, Israel, Luxembourg , The Netherlands, New Zealand, Poland, Spain, Switzerland and UK	accelerated tPA regimen, while in US patients it had the worst outcomes. Accelerated tPA had greatest benefit in US patients	Califf, 2001)
Madan, Labinaz, Cohen, Buller et al (2004)	Enhanced Suppression of the Platelet IIb/IIIa Receptor with Integrilin Therapy	ESPRIT	The 30-day and one-year rates of death, myocardial infarction (MI) and target vessel revascularisation (TVR) were compared between Canadian and US patients	US and Canada	While the beneficial effect of Eptifibatide was observed in both countries, Canadian patients had better 30 day and 1 year clinical outcomes then North American patients after PCI	
Cohen et al (2001)	Evaluation of IIb/IIIa Platelet Inhibition for Stenting	EPISTENT	Compared clinical outcomes between Canada and US after PCI	US and Canada	Canadian patients had a lower incidence of death, MI, TVR	
Kang, Ohtsu, Van Cutsem et al (2010)	Unknown	AVAGAST	Compared overall survival between Asia, Europe, and North and South America to evaluate effect of Bevacizumab plus chemotherapy in Gastric Cancer patients	Asia, Europe, North and South America	Overall Survival was heterogeneous across regions for the intervention under investigation	Regional Differences in patient characteristics and treatment strategy (Saijo, 2010)

Although finished prematurely, one study looking at beta-blockers as an immediate treatment for those surviving MI, showed a difference in both the direction and the size of treatment effect

between the US and Canada (Senn and Harrell, 1997; O'Shea and DeMets, 2001). Senn and Harrell (1997), when analysing this data, concluded that the variation could be explained by random variation. However, O'Shea and DeMets (2001) suggest that these differences cannot be explained by differences in the quality of care, because all participating centres were chosen by virtue of their ability to recruit patients and their outstanding academic records. Despite sample sizes from some of the centres being small, which possibly explains some of these differences, their work indicates that an interaction between country-level factors and treatment existed, supporting the notion of international differences in treatment effectiveness.

Another multinational trial with similar results is reported by O'Shea and Califf (2001), who examined the results from the FIRST trial. Findings showed that patients in North America who received Epoprostenol experienced fewer major events than those receiving a placebo and showed that there was no difference in unadjusted survival probability between patients who did and did not receive the intervention. This contrasted with the findings from Europe, where those treated with the intervention showed an increased risk of death, suggesting that there was some interaction between region and treatment. However, O'Shea and Califf argue that these differences may have been the consequence of a *perceived* clinical improvement by those treated with placebo in Europe compared with their North American counterparts.

O' Shea and Califf (2001) also analysed the data from the PURSUIT multinational trial, in which 10,948 patients with acute coronary syndrome (ACS) from 28 countries in four geographic regions (North America, Western Europe, Eastern Europe and Latin America) were randomised to either Eptifibatide or placebo in addition to standard care. It was found that treatment benefit

with regard to death or MI at 30 days was greatest in North America. Furthermore, patients in Western Europe were found to have fewer composite events when treated with Eptifibatide compared to Latin America and Eastern Europe, where no apparent treatment effect was found. This implies that there was an interaction between country and effect. However, Akkerhuis et al (2000), analysing the same data, suggested that inter-country differences could be explained by different baseline demographics, adjunctive treatment strategies, and definitions of MI.

The GUSTO trial reported by Van de Werf, Topol, Lee et al (1995) provides further evidence of these differences. This trial randomised 41,012 patients in 15 countries to four thrombolytic strategies to investigate their effect on mortality rates. The four thrombolytic therapies were Streptokinase with subcutaneous Heparin, Streptokinase with intravenous Heparin, Accelerated tPA with intravenous Heparin, and a combination of Streptokinase with tPA and intravenous Heparin. In non-US patients, the combination strategy worked as well as the accelerated tPA strategy, while in US patients this combination was associated with the worst clinical outcomes. Furthermore, accelerated tPA had the greatest benefit in US patients over the other regimens, but this benefit was not observed in the other countries involved in the study. However, O'Shea and Califf (2001) suggest that these differences result from a greater occurrence of haemorrhagic strokes in patients who received the combination regimen in the US compared with the other participating countries and those treated with the other three regimens. This, they suggest, may have impacted the perceived effectiveness of this regimen, making it seem less effective in the US patients.

Data from the ESPRIT multinational trial (Madan et al. 2004) shows differences between Canadian and US patients undergoing non-urgent Percutaneous Coronary Interventions (PCI). Results showed that Canadian patients had better thirty-day and one-year clinical outcomes than their US counterparts, differences that persisted even after adjustment for known baseline differences. This suggests the existence of an interaction between country and treatment and is consistent with the findings of the EPISTENT trial, which also compared outcomes between Canada and the US after PCI (Cohen et al. 2001). Cohen et al. found that at one year, Canadian patients had a lower incidence of target vessel revascularisation, MI, and death after undergoing PCI than their US counterparts. Madan et al. (2004) suggest that consistency in the findings between these trials means that observed differences cannot be fully explained by chance, but are a consequence of differences in effectiveness between countries.

Although trials involving patients with conditions other than cardiovascular diseases also show this interaction between country and treatment,[4] these are not explored here because this thesis is exclusively concerned with cardiovascular disease.

[4]For example, the AVAGAST trial explored the effect of Bevacizumab versus placebo as an addition to chemotherapy on overall survival in 774 patients with inoperable Gastric Cancer in three regions: Asia, Europe and the Americas (Kang et al. 2010). It showed that although Bevacizumab failed to significantly improve overall survival, its effect was heterogeneous between all regions investigated. In Asia, the overall survival rate was found to be 12.1 months with Bevacizumab compared to 13.9 months with placebo. This contrasted with Europe where overall survival was 8.6 months with Bevacizumab and 11.1 months with placebo and the Americas where overall survival was 6.8 months and 11.5 months respectively. However, the regional differences found in this trial could be due to regional differences in patient characteristics and treatment strategies (Saijo, 2010).

The evidence presented above suggests the existence of international differences in treatment effectiveness, but the majority of this originates from data that were not collected for this purpose. As such, much of the evidence is *prima facie* evidence of an interaction between country and treatment or country and outcomes, with little direct evidence of international differences in treatment effectiveness. Furthermore, authors commonly suggest that these inter-country differences can be explained by differences in patient characteristics or clinical management, factors which had not been accounted for when the trial was conducted. However, it is clear from the studies presented here that some of the variation cannot be explained by these two factors alone. The next section discusses the factors that may explain international differences in treatment effectiveness found in multinational trials.

2.4: Explanations for Treatment Effectiveness Differences

There are many factors that might explain the interaction found between country and treatment. As with guideline recommendations, differences could derive from culture or ethnicity. In the case of multinational trials, they may also be a consequence of participating countries implementing the trial protocol differently (see Box 2.1).

BOX 2.1: FACTORS THAT MAY EXPLAIN INTERNATIONAL DIFFERENCES IN TREATMENT EFFECT FOUND IN MULTINATIONAL TRIALS (DISPLAYED USING AN ADJUSTED PICO[5] FRAMEWORK).

Patients/ Population	**Differences in Patient Selection Factors:** • Patient baseline demographics • Disease severity levels • Patient risk level • Compliance • Ethnicity • Healthcare System • Socio-economic status
Intervention	**Differences in Timeliness of Treatment:** • Time taken to receive treatment
Comparator centre/site	**Differences in centre or site:** • Teaching or non-teaching hospitals • Staffing • Length of Hospital Stay • Access to facilities • Thresholds for admission to centre
Outcome	**Differences in Outcome Ascertainment and Accuracy:** • Differences in how centres ascertain an endpoint • Differences in the accuracy of laboratory results used to ascertain an endpoint

[5] Participants, Intervention, Comparator and Outcomes

Adjunct Treatment	**Differences in Treatment Management:** • Adjunctive surgical procedure use • Adjunctive medication use • Route of administration for adjunctive medications • Dosage of medications
Trial Quality	**Differences in Trial Quality:** Differences in the degree to which centres follow trial procedures, i.e. randomisation procedures

2.4.1: Patient Differences

Any international differences in treatment that multinational trials have reported may be a consequence of four main patient baseline differences between the countries in the trial.

2.4.1.1: Baseline demographics

Trials reporting an interaction between treatment and country have found differences between countries in patient baseline demographics (Domanski et al. 2004; Fox et al. 2000).White (2000) suggests that while patient selection factors can significantly affect treatment outcome in multinational trials, their effect is often ignored by researchers. O'Shea and Califf (2001) point out that clinical baseline demographics also differ between countries in such trials. They found that in trials examining fibrinolytic therapy, non-US patients are more likely than US patients to have suffered an anterior MI prior to randomisation. Other factors that may create differences in patient baseline demographics are social class, income, and education (Dragano et al. 2007; Galobardes, Lynch, & Smith, 2007).

It is possible, therefore, that patient baseline differences may explain the interaction found between country and treatment in multinational trials, differences that Roberts and Torgerson (1999) term “chance bias”. Nevertheless, they note that many RCTs already protect against this type of bias by implementing appropriate randomisation techniques or adjusting statistical analysis for baseline characteristics that are known to influence the outcome of a treatment. Each of these procedures goes some way to ensuring that the outcomes of the trial can be assumed to be a consequence of treatment.

Patient baseline demographics also include a number of physiological and environmental factors that could explain the country-treatment interaction. One such factor is ethnicity, and studies looking at many different conditions have shown that care seeking behaviour, type and quality of care received, and the incidence of diseases vary between ethnic groups (Broderick et al. 2010; Jha et al. 2005: Shen et al. 2007) as do outcomes after treatment (Saha et al. 2008). The Veterans Administration Cooperative Study Group on Antihypertensive Agents (1982), for example, found that those of white European descent had a better antihypertensive response to beta-blockers than those of black African descent. It is possible, therefore, that the interaction between country and treatment found in many multinational trials is due to differences in ethnicity. Different ethnic groups receive different types and levels of treatment, react differently to treatment, and are exposed to different disease associated risk factors, all of which are known to influence treatment outcomes.

The interaction found in multinational trials between country and treatment might be explained by taking into account socio-economic status (SES). This not only affects individual patient

health and risk factors (Marmot, 2003; Marmot, Shipley, & Rose, 1984), but also influences treatment outcomes (Brechner et al. 1993). Evidence suggests that those in the lowest socio-economic groups have greater risk factors for certain diseases (Almqvist, Pershagen, & Wickman, 2005; Millar & Wigle, 1986), are less likely to receive some treatments, wait longer for treatment, and have worse outcomes than their counterparts in higher socio-economic groups (Gornick et al. 2009; Potosky et al. 1998; Kapral et al. 2002). The interaction between country and treatment found in a number of multinational trials, therefore, may be a consequence of participating countries having different SES compositions, which then affects not only how, when, and where patients receive treatment, but also their outcomes.

Another factor that should be considered is variation in healthcare systems. Healthcare may be provided in the private or public sector. For instance, in 2001, Canada spent only 57% as much per capita on healthcare expenditure as the US (Reinhardt, Hussey, & Anderson, 2004). This was because the US was able to utilise more expensive technology and had larger administrative expenses than Canada, which had imposed healthcare budgetary restraints (O'Shea et al. 2001). This led to Canada having less access to highly technological procedures such as magnetic resonance imaging. Such differences in healthcare systems may result in differing patient outcomes and should be accounted for when adjustments are made in the analysis of multinational trials.

Countries also differ on how healthcare budgets are spent. Some countries may have the resources to procure treatments, equipment, and staff that have a beneficial effect on patients in some areas of their healthcare system, while in other parts of their system, interventions and

equipment may be less advanced. The areas that each country prioritises also differ. For example, Heijink et al. (2008) found that while Australia spent a lot of money on ambulatory care, France spent the majority of its health budget on medical goods. If a country in a multinational trial spends more of their healthcare budget in the clinical area studied, this may have implications for the outcome of the study intervention and needs to be considered when comparing the results of trials across the countries involved.

The mechanisms through which healthcare systems are funded may influence treatment outcomes. Funding of healthcare systems differs internationally, and countries with differently funded systems are usually involved in multinational trials. In the UK, for example, healthcare costs are met through funds raised through general taxation; in others, healthcare systems are privately run but receive some funding from government (e.g. Canada) or are funded jointly by beneficiaries and employers, with contributory insurance rates to a person's wage (e.g. Japan). Of course, the way in which healthcare systems are funded can have a direct impact on whether new technologies can be purchased and the quality of care a patient receives. All these factors can affect treatment and services and could lead to treatment outcomes differing between countries.

2.4.1.2: Disease severity

International differences may be a consequence of one country in the trial enrolling patients with a high disease severity, while another enrols patients with a low disease severity, with patient outcomes differing according to the patients' severity levels. White (2000) contends that in many multinational trials examining patients with coronary disease, disease severity will vary between populations, and this may influence treatment outcomes. Indeed, there is evidence to suggest that

responses to treatment differ in line with this factor (Anderson et al. 2000). Lindsay, Zaman and Cowan (1995) propose that from post-infarction trials, it can be argued that the beneficial effect of ACE inhibitors is heterogeneous, as they appear to be most beneficial in patients who have clinical evidence of heart failure, that is, those who have the greatest level of disease severity.

Closely related to disease severity is the risk level of patients enrolled in multinational trials. As White (2000) points out, the risk of events among patients will differ as a consequence of differences in the degree of the disease. For example, as the level of disease severity increases, so does the likelihood of experiencing a significant event such as an MI. This may create between-regional differences in treatment effect when treatment effect is measured in absolute terms such as the risk difference. This is because absolute treatment effect measures are sensitive to baseline risk, with absolute estimates of treatment effect being higher in centres with higher base rates (Borenstein et al. 2011).

2.4.1.3: Patient Behaviour

Evidence suggests that exposure to disease related risk factors, such as smoking or diet behaviours, and beneficial behavioural factors, such as moderate amounts of exercise, differs between countries (Bergovec et al. 2008; Kromhout et al. 2002; Park D et al. 1998; Myers, 2003; Rimm et al. 1999). It is unsurprising, therefore, international differences in response to treatment have been found in multinational trials. However, many trials do not examine patients' differing exposures to associated risk factors across countries and are, in effect, comparing heterogeneous groups of patients. National behavioural characteristics therefore need to be considered before international differences in treatment outcomes and effectiveness can be explained.

2.4.1.4: Treatment Compliance

Compliance, defined by Hayes (cited in Bosworth, Oddone, & Weinberger, 2005) as the extent to which patients adhere to medical and health advice, has been found to influence the overall effectiveness of a treatment (Cramer, 2002; Psaty et al. 1990). Bleyer et al (1999) suggest that differences in compliance between countries could significantly alter the efficacy of some treatments and could, therefore, explain international differences. Indeed, when low levels of compliance are not considered in trials, it is thought that treatment effect is underestimated (Pullar, Kumar, & Feely, 1989). As such, inter-country differences in treatment effectiveness may be a consequence of differing levels of compliance between countries.

However, multinational trials may analyse data using the intention to treat (ITT) principle, which includes all randomised patients in the analysis according to the group to which they were allocated, regardless of the treatment they received, whether they completed treatment, or whether they complied with treatment (Newell, 1992: Fisher et al. 1990). In doing this, ITT analysis deliberately ignores differences in compliance, meaning that biases related to such differences cannot be introduced into the analysis. This is important, since patients who comply with treatment tend to have better outcomes regardless of group assignment (intervention or control) (Montori & Guyatt, 2001).Therefore, in trials that have used ITT analysis, inter-country differences in treatment effectiveness cannot be the consequence of differing levels of compliance between countries.

2.4.2: Differences in the Timeliness of the Investigative Intervention.

International differences in outcomes that multinational trials have reported may be a consequence of differences in the time it took to administer the investigative treatment.

2.4.2.1: Timeliness of Treatments

Time to treatment has been found to differ internationally (Fu et al. 2000; Gupta et al. 2003) and can affect event rates for a variety of conditions (Besselink et al. 2007; De Luca et al. 2004) as well as determining the benefits of certain interventions. The Fibrinolytic Therapy Trialist Collaborative Group (1994) found that the benefit of intravenous thrombolysis was influenced by the time between symptom onset and start of treatment. It may be, therefore, that inter-country differences in multinational trials can be explained as a consequence of time to treatment differences between the countries participating. However, Welsh et al. (2005) found that in the ASSENT-3 PLUS trial, event rates differed, even though time from randomisation to treatment (including any adjunctive medications) was homogenous over all participating countries.

2.4.3: Comparator Centres or Sites

Between-country differences in the types of centre or site enrolled in multinational trials may also lead to differences in outcomes.

2.4.3.1: Types of Institutions Included in Multinational Trials

As many multinational trial protocols do not stipulate the type of site that should be included in the study, different types of institutions may be recruited in each country, leading to patients receiving different approaches to treatment management. Evidence suggests that the type of

institution affects treatment management and patient outcomes (Jensen, Webster, & Witt, 2009), and differences in types of site may therefore help explain the interaction between country and treatment. For example, patients admitted to academic hospitals have better outcomes than those admitted to community hospitals (Allison et al. 2000). Chaudhry, Goal and Sawaka (2001) state that the teaching status of a hospital may affect a patient's outcome both directly and indirectly, as teaching status not only leads to better knowledge and skills, but also to improved processes of care. Other studies, however, suggest that hospital type has little impact upon patient outcome. Papanikolaou, Christidi and Ioannidis (2006), for instance, found that there was insufficient evidence that teaching hospital status improved patient outcome. This does not mean that international differences in patient outcome are in no way due to the types of sites enrolled, but rather that it may not be the most important factor, especially given that the lack of reporting on the difference in the types of institutions makes it difficult to assess its impact.

2.4.3.2: Staffing of Sites Recruited

Evidence suggests that patient outcomes are influenced by the type of staff caring for the patients; mortality rates are lower for patients treated by specialists (Jollis et al. 2009; Nash, Nash, & Fuster, 1997). Brevetti et al. (2007) found that survival rates for patients with peripheral arterial disease were higher when patients were managed by a specialist. It is argued this is because specialists are more likely to use more effective medicines and are less likely to use treatments with limited benefits (Abubakar et al. 2004; Brevetti et al. 2007; White 2000). Welsh et al. (2005) used the assessment of the safety and efficacy of a new thrombolytic (ASSENT-3 PLUS) trial to assess global variations on the impact of pre-hospital care and found that pre-hospital treatment for patients in Sweden, Canada, and the Netherlands was provided by advanced cardiac life

support trained nurses and paramedics, while in France, Spain, and Germany it was provided by physicians. In the UK and Norway, pre-hospital treatment was provided by physicians in some sites and not in others. Overall, only 63.8% of patients received pre-hospital treatment from a physician. They found that the presence of a physician was related to fewer patients experiencing recurrent ischemia and repeated MI at thirty days. This suggests that how sites or hospitals are staffed influences patient outcomes and may be an important factor in explaining the interaction between country and treatment found in multinational trials.

2.4.3.3: Length of Hospital Stay

Evidence that length of hospital stay differs between countries in multinational trials comes from Kaul et al. (2004). They examined early discharge and length of hospital stay for patients with uncomplicated acute MI in The Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries (GUSTO-I), the Global Use of Strategies to Open Occluded Coronary Arteries (GUSTO-III), and the Assessment of the Safety and Efficacy of a New Thrombolytic (ASSENT-2) trials, which took place between 1990–1998. They found that in the ASSENT-2 trial, the median length of hospital stays was lower than that from the GUSTO-I trial conducted six years earlier (See Table 2.2).

Table 2.2: Median Length of hospital stays (in days) across GUSTO –I and ASSENT -2.

Country	GUSTO-I	ASSENT-2
Germany	24	17
Spain	12	10
France	12	11
Belgium	12	11
US	8	5
Australia	8	6
New Zealand	7	5

They also found that the decrease in median length of hospital stay over all countries was significant ($p<0.0001$), with the median length of hospital stay of nine days in the GUSTO-I trial decreasing to seven days in the ASSENT–2 trial.[6]

Differences in the length of hospital stay can affect overall outcomes for patients. Kondo, Zierler and Hagino (2010) found that patients who had undergone hip fracture surgery and who were discharged within two weeks had a significantly higher risk of mortality than patients who had undergone the same type of surgery but remained in hospital for more than 40 days. The interaction between country and treatment reported by some multinational trials may therefore be a consequence of these differences in hospital stay. White (2000) suggests that an ascertainment bias may have existed in the PURSUIT trial owing to longer hospital stays in Eastern Europe compared to North America. Longer hospital stays meant patients were more likely to have an electrocardiogram and to have cardiac enzymes monitored. Furthermore, longer hospital stays made extra monitoring possible. However, it should be pointed out that while the length of hospital stay may be an important factor in explaining between-country differences in patient outcome, it may be part of an interplay of factors. Indeed, Christensen et al. (2009) found that while there was a strong correlation between the length of hospital stay and patient mortality, the

[6]Further evidence is provided by Christensen et al (2009), who, when investigating data from the FAST trial, found that the average length of hospital stay differed between the participating countries. The average length of hospital stay was 27 days in Sweden and Israel, 14 days in the Netherlands, 25 days in China, and just 16 days in Australia.

same was true for other factors such as the proportion of patients receiving magnesium and whether patients received treatment for oedema.

2.4.3.4: Access to Facilities

Sites that have better access to facilities such as catheterization laboratories are more likely to perform more procedures, have shorter waiting times for treatment, and therefore increase the benefits of treatment overall than sites with less access to these facilities (White, 2000; Yusuf et al. 1998). Every et al. (1993) found that patients admitted to hospitals with on-site catheterisation facilities were more likely to have coronary angiography than those admitted to hospitals without these facilities. This is important because coronary angiography can increase the benefits of treatments that reduce the incidence of periprocedural infarction (White, 2000). Furthermore, Fu et al. (2000), examining data from the US and Canadian sites of the GUSTO-IIb trial, found that US hospitals were more likely to have on-site facilities for revascularisation and angiography than Canadian hospitals. They argue that the availability of on-site facilities within the US resulted in shorter waiting times for these procedures. This may then affect patient outcomes.[7]

2.4.3.5: Thresholds for Admission

Different countries may have differing thresholds for admission, which the multinational study protocol may fail to control for. Rouleau et al. (1993), using data from the Survival and

---

[7]Evidence can also be found in multinational trials involving stroke patients. Gray et al. (2006) examined data from the Tinzaparin in acute ischaemic stroke (TAIST) multinational trial, which investigated the effectiveness and safety of Tinzaparin at different doses in patients with acute ischemic stroke. They found that admission to a stroke care unit differed between countries. In Norway 67.1% of patients were admitted to a stroke care unit compared with no patients in Finland and only 4.9% in Ireland.

Ventricular Enlargement (SAVE) trial, investigated differences in admission thresholds in the US and Canada for patients with chest pain and found that Canada had the higher admission threshold. In Canada, 51% of patients admitted to one of the coronary care units had clinical MIs as compared to 35% of US patients. This is important because different thresholds for admission mean that different patients with different risk factors could be admitted to a study, meaning that patient outcomes are likely to differ. By ignoring this factor, multinational trials may be unwittingly comparing heterogeneous patient groups, which may contribute to the international variation in outcomes (White, 2000).

However, it should be noted that most multinational trials will adjust any analysis by centre. This is to account for any between-centre variation caused by known and unknown factors such as inconsistent operating procedures, staffing differences, different lengths of hospital stay, and different patient baseline demographics (Chow & Liu, 2004). In doing this, an unbiased estimate of treatment effect can be calculated, which can be generalised over centres. This means that while between-country differences in centre type, staffing, and admission thresholds may impact upon patient outcomes, they are unlikely to influence estimates of treatment effect.

2.4.4: Outcomes

Other factors may play a role in producing the international differences found in multinational trials. Some of these may relate to outcome definition and measurement.

2.4.4.1: Definition of Endpoint

The criteria used to define an endpoint of MI may differ between countries and may result in differences in the number of MIs recorded. However, this factor may be less important than

others in explaining the occurrence of country-treatment interactions, since in many multinational trials, clinical event committees are used to adjudicate suspected endpoints and exclude events that do not adhere to the strict trial definitions (Mahaffey et al. 2001; White, 2000). Furthermore, over 100 countries use the classification of diseases (ICD) system to define and report clinical events. This means that definitions of clinical events, such as an MI, should be similar between countries. Despite this, the use of these committees and the ICD is not a standard requirement in multinational trials, and the means of specifying endpoints may still differ between countries. Moreover, the country-treatment interaction found may be produced by the varying accuracy of laboratory results between countries. In multinational trials, laboratory reports and results may be used to diagnose a clinical event. For instance, laboratory reports of cardiac enzymes may be used to define the occurrence of an MI. Differences between countries would be apparent if the accuracy of these results differed between countries. Indeed, if reproducibility is weak, then the specificity and sensitivity of the endpoint diagnosis may vary (White, 2000).

2.4.5: Adjunct Treatment

Treatment management varies between countries, meaning that adjunctive treatments also vary, and this may explain the apparent interaction between country and treatment in multinational trials. Many multinational trials do not take into account the treatment management differences between the participating centres and countries that can occur after randomisation (Fox et al. 2000), even though such differences may lead to differences in the use of adjunctive treatments, which may influence patient outcomes. Reed et al (2006) suggest that multinational trials may introduce a bias into their dataset by comparing patients who have been unequally treated,

especially in cases where treatment management differs between centres and where an interaction exists between the investigative treatment and a background treatment, which may give rise to the apparent differences in patient response to treatment.

2.4.5.1: Surgical Interventions

The rate of surgical procedures varies between countries included in multinational trials. For example, the PURSUIT trial showed regional variations both in the number of PCIs performed, with only 2% of Eastern European patients receiving this intervention as compared to 25% of North American patients (White, 2000), and in the rate of CABG during hospitalisation (11.3% of cases in Latin America as compared to 19.4% in North America) (Cohen et al, 2001).[8]

The association found between country and treatment in many multinational trials may be due to the different rates at which surgical procedures are used alongside the trial interventions. If these surgical procedures are beneficial, then their use will influence patients' responses. In such cases, patient outcomes may be better in some countries not because of the intervention *per se*, but because of the adjunctive use of a beneficial surgical procedure.

Indeed, Gupta et al. (2003) found that the Assessment of the Safety and Efficacy of a New Thrombolytic (ASSENT-2) trial showed that increased revascularisation in some countries was related to better outcomes. They concluded that adjunctive surgical procedures better explain inter-country variation in mortality rates and outcomes than patient baseline characteristics.

---

[8]Similar results were also reported from the ESSENCE trial (Fox et al. 2000), where participating countries had different revascularisation rates, with rates being high in Argentina but low in the UK.

However, results from the Intravenous nPA for Treatment of Infarcting Myocardium Early II (InTIME-II) trial challenges this, finding that Latin America had a higher mortality rate following Fibrinloysis than Western Europe, despite having similar revascularisation rates (Giugliano et al. 2001). This suggests that international differences in patient response exist regardless of the use of adjunctive beneficial surgical procedures.

2.4.5.2: Medicinal Treatments

The types of medication used alongside the intervention under investigation also differ between countries involved in multinational trials, possibly because the protocols of such trials do not always stipulate which adjunctive medications should be used, and may, therefore, explain between-country differences in patient outcomes.

2.4.5.2.1: Type of Medicinal Treatment

Evidence that countries participating in multinational trials use different adjunctive medications comes from a number of trials. The Factor Seven for Acute Hemorrhagic Stroke (FAST) trial (Christensen et al. 2009) showed that 95% of Chinese patients received medication for oedema compared to 21% of US patients, 9% of Canadian patients, and 0% of patients in Finland. The ESSENCE trial (Fox et al. 2000) found that, of the countries involved, UK patients received the lowest percentage of nitrates and beta-blockers and a higher percentage of calcium channel blockers compared to those in Argentina.[9]

[9] In another trial, the use of low molecular weight heparin in patients with ACS has been found to be up to three times lower in the U.S when compared to other countries (Sinnaeve& Van de Werf, 2004).

As with adjunctive surgical procedures, adjunctive medications may influence patient outcomes. However, not all studies support this explanation for international differences. For instance, Rouleau et al. (1993) found that variation in the number of recurrent MIs and deaths between Canada and the U.S was not associated with a greater use of adjunctive medicinal therapies.

2.4.5.2.2: Route of Medicinal Treatment

The administration route of adjunctive medications may also account for the interaction between country and treatment in multinational trials. White (2000) found that the route of heparin administration differed between countries in the PURSUIT trial. If some routes of administration are more effective than others, this may influence the response of patients to conventional treatment and the investigative intervention (Butkiewicz et al. 1995; Fisher, Shahshahani, & Kitabchi, 1977; Schwartzman & Morgan, 2004).

2.4.5.2.3: Dose

Dosages of adjunctive medications are rarely stipulated in multinational trial protocols and may differ across participating countries, thus contributing to international differences. In the Scandinavian Simvastatin Survival Study (4S), Faergeman et al. (1998) found that aspirin dosages given as an adjunctive medication differed across the participating countries. For example, Norwegian patients were given less aspirin than their Danish counterparts. This means that multinational trials may be comparing patients who are treated differently, which might produce differences in patient outcome, particularly if one dosage regime was more effective than another. Indeed, there is evidence to suggest that differences in the dosage of adjunctive medications affect the overall outcome of a treatment (Berger et al. 2008).

2.4.6: Trial Quality

One of the most important factors to consider when investigating international differences in treatment effectiveness is trial quality. It may be that the differences found between countries in outcomes and treatment effect are a result of the trial centres enrolled in multinational trials differing in the degree to which they follow trial procedures. For instance, if the centres in one country do not strictly follow randomisation procedures, then treatment effect estimates may be higher than a country in which the centres strictly follow procedures since treatment assignment will no longer be masked. Therefore, countries showing the most positive results may be producing lower quality trials. Indeed, studies suggest that low quality trials tend to exaggerate the effect of treatment (Egger et al. 2003; Kjaergard, Villumsen, & Gluud, 2001; Noseworthy et al. 1994; Peduzzi et al. 1993). Moher et al. (1998) examined 127 trials from 11 meta-analyses concerning circulatory and digestive diseases, mental health, pregnancy and childbirth. They found that when trial results from low quality trials were pooled, there was a statistically significant exaggeration of treatment effect by 30-50%. Furthermore, there was an association between inadequate allocation concealment and treatment benefit estimates, with trials that had inadequate allocation concealment tending to show increased treatment benefit. However, these large distorted treatment effects were not evident when trials were not double-blinded. According to Juni, Altman, and Egger (2001), this is because the impact that a lack of double blinding has on the distortion of the effect estimate will, to some extent, depend on the outcomes assessed. For instance, when investigating treatment effect with regards to overall mortality, the blinding of outcome assessment will be immaterial.

Further evidence of treatment effect estimates being affected by trial quality is provided by Schulz et al. (1995). They used a database of systematic reviews of randomised controlled trials concerned with pregnancy and childbirth to investigate whether biases were related to inadequate methodological approaches. Trial results were examined with respect to randomisation, blinding, and attrition. They found that inadequate methodological approaches exaggerated treatment effect. For randomisation, odds ratios were exaggerated by around 41% when trials inadequately concealed treatment allocation, while trials that did not report how they concealed treatment allocation exaggerated the treatment effect by around 30%. Trials that were not double-blind gave larger effect estimates, exaggerating the effect of a treatment by around 17%. However, when attrition was examined, trials that had excluded patients after randomisation did not result in exaggerated treatment effect estimates, but this lack of association may be a consequence of incomplete reporting (Schulz et al. 1995). To conclude, Schulz et al. argue that such results highlight both the importance of adequate methodological approaches and reliable and complete reporting, without which the assessment of trial quality is not feasible.

The above evidence suggests that trial quality can have a significant impact upon the estimate of treatment effect. It may be the case, therefore, that differences in trial quality between countries create the international differences found in therapeutic effect. As such, trial quality is an important factor to investigate and is therefore examined in more detail in Chapter 7.

2.5: Conclusion

This chapter has provided an overview of the evidence for the argument that international differences in treatment effectiveness exist while also highlighting possible explanations for these

differences. It has shown that although there is little direct evidence of international differences in treatment effectiveness, *prima facie* evidence is available from multinational trials that show an interaction between country and treatment. In doing this, it has shown that trial data can be utilised in investigating international differences in treatment effectiveness. It has also presented influencing factors that may still be present, even under ideal conditions, when multinational trials are conducted and that, if not accounted for in the trial protocols and analysis, have an impact upon treatment outcomes.

With this mind, this study will look at whether international differences in treatment effectiveness exist for cardiovascular diseases. To do this, panoramic meta-analysis will be used, an approach that is described in the next chapter.

CHAPTER 3.

THE METHODOLOGY OF PANORAMIC META-ANALYSIS.

The primary aim of this study was to investigate the existence of international differences in the clinical effectiveness of medical interventions. To do this, it used the novel approach of panoramic meta-analysis. This uses the same stringent procedures and techniques as standard meta-analysis when searching for and analysing data from relevant literature but is used for collating data from existing meta-analyses rather than from individual studies. It enables testing of generic hypotheses over a set of meta-analyses concerned with different types of intervention for related health conditions.

This chapter sets out the methodological position and the novel methods that will be used to answer the research aims. Panoramic meta-analysis has been selected because it is an approach that allows for international differences in treatment effect to be directly investigated and, as such, allows a more comprehensive understanding of such differences to be crafted. It also enables generic hypotheses, such as the one in this study, to be investigated. In addition, panoramic meta-analysis is more efficient than "standard" meta-analysis as relevant studies can be identified more swiftly by searching for meta-analyses rather than the studies themselves.

3.1: Panoramic Meta-analysis

Panoramic meta-analysis has its foundations in the principles and procedures of standard systematic reviewing and meta-analysis, and likewise it follows a systematic process in which relevant studies are identified, evaluated, and combined. However, it uses information from published systematic reviews containing meta-analyses to investigate the validity of "generic" hypotheses.[10] So, for instance, panoramic meta-analysis is applicable to any situation in which a hypothesis involves different types of intervention used for similar medical conditions. It can also be used when similar interventions are employed to manage or cure different, but related, medical conditions or when the same intervention is used to treat or cure the same condition in different, but related, clinical scenarios.

3.2: Why is Panoramic Meta-analysis Needed?

It is estimated that studies of healthcare interventions are growing at a rate of tens of thousands per year (Ghersi & Pang, 2009; Hawker et al. 2002). While these studies are essential for evidence based medicine, the sheer number of studies can create problems for health policy-makers, analysts, and professionals when making decisions about the most appropriate clinical treatment to use. This issue has long been recognised, with many researchers preparing reviews to collate, appraise and summarise the individual studies (Smith et al. 2011). As far back as 1753, James Lind carried out a review of the reports on the prevention and treatment of scurvy (The James Lind Library, 2007). Egger, Smith and O'Rourke (2001) state that one of the earliest noted

[10] generic in the sense that they can apply to all members of a family of interventions or to a group of medical conditions

meta-analyses was carried out in 1904 by Pearson, who integrated the results of five studies to examine the effectiveness of inoculation on typhoid. Today, there are many national and international organisations that are committed to conducting systematic reviews and meta-analyses, including the National Institute of Health and Clinical Excellence (NICE),[11] the NHS Centre for Reviews and Dissemination,[12] the Institute for Work and Health,[13] and The Cochrane Collaboration.[14]

Due to the number of organisations and researchers conducting systematic reviews and meta-analysis, the healthcare literature is now overrun with such types of research. For example, The Cochrane Library contains around 4,500 systematic reviews and around 2000 protocols for new or updated systematic reviews (The Cochrane Collaboration, 2011). The sheer quantity of literature could prove overwhelming to those involved in decision making, a problem exacerbated by there usually being more than one review on a topic, with reviews sometimes reaching different conclusions and being of varied quality (Cook et al. 1996; Katerndahl & Lawler, 1999; Smith, Devane, Begley, & Clarke, 2011). Furthermore, the steps required to produce and report systematic reviews and meta-analyses have increased in recent years, producing longer reports (Bastian, Glasziou, & Chalmers, 2010). For instance, Bastian, Glasziou and Chalmers (2010) note that early Cochrane reviews were usually only 10-20 pages long, whereas today they can be over several hundred. This makes reading and evaluating these publications time consuming for those involved in decision-making processes.

---

[11] Information on NICE can be found at: http://www.nice.org.uk/
[12] Information on the NHS Centre for Reviews and Dissemination can be found at: http://www.york.ac.uk/inst/crd/
[13] Information on the Institute for Work and Health can be found at: http://www.iwh.on.ca/
[14] Information on the Cochrane Collaboration can be found at: http://www.cochrane.org/

Consequently, leaner, more efficient, and innovative methods are required to compare and contrast the findings of systematic reviews and meta-analyses and provide an overview of the best available evidence for health policy-makers, analysts, and healthcare professionals. Panoramic meta-analysis is one method for achieving this.

However, panoramic meta-analysis is distinguished from other meta-epidemiological approaches in that unlike other approaches, the trial data from published meta-analyses can be used to investigate the validity of generic hypotheses. For example, it can be used to investigate and compare the effectiveness of different intervention types for different, but related, medical conditions within the same analysis and has been used to investigate the effectiveness of prophylactic antibiotics in surgery for a variety of medical conditions (Bowater, Stirling, and Lilford, 2009). In doing this, it can provide a broader overview of intervention effectiveness and provide a good foundation on which healthcare decisions can be made.

3.3: A History of Panoramic Meta-analysis

One of the fundamental aspects of panoramic meta-analysis is the systematic reviewing of published systematic reviews containing meta-analysis. This concept is not new and has been used in a number of past studies. For example, using this approach, Ernst (2002) investigated the effectiveness of homeopathy by conducting a critical analysis of seventeen systematic reviews involving meta-analyses of controlled trials on homeopathic treatments in human patients or volunteers from four relevant databases. Of these, 11 were independent systematic reviews, while six were related to the re-analysis of a landmark meta-analysis in this area that had suggested that homeopathy was efficacious and was more than a mere placebo. Ernst found that

the six systematic reviews that had re-analysed the landmark meta-analysis did not support its positive findings. Furthermore, when all the included systematic reviews were analysed collectively, they failed to provide any evidence that homeopathic treatments were effective.

More recently, Derry et al. (2006) conducted a systematic review of systematic reviews to establish the usefulness of systematic reviews in assessing the evidence for acupuncture, as while reviews of acupuncture supported its use, they were based on imprecise inclusion criteria which may have influenced their overall conclusions. They found that when excluding trials from the included reviews that were known to have sources of bias, none of the 35 systematic reviews provided strong support for the effectiveness of acupuncture. They concluded that by including studies that showed bias, systematic reviews of acupuncture had exaggerated the effectiveness of this intervention and that there was no robust evidence that acupuncture worked for any condition.

Tennant et al. (2007) also performed a systematic review of systematic reviews to evaluate the effectiveness of interventions promoting mental health and preventing mental illness. From ten electronic databases, they found 27 relevant systematic reviews on mental illness prevention and mental health promotion interventions for infants, children, and young people up to the age of 19. They then categorised each systematic review according to the type of intervention assessed: parenting interventions, programmes for prevention of anxiety and depression, programmes to promote self-esteem, violence and aggression prevention programmes, school based programmes, and general reviews. The evidence from each category was then critically appraised using the Critical Appraisal Skills Programme Appraisal Tool for Systematic Reviews (Critical Appraisal

Skills Programme, 2002). Where possible, effect sizes and 95% confidence intervals (CI) were calculated for each review to assess the effect of the intervention. They found that while many of the systematic reviews lacked methodological rigour, they did provide evidence of the effectiveness of a range of interventions for preventing and promoting mental health, even though the effect sizes of many of the reviews were relatively modest. In concluding, Tennant et al. argue that systematically reviewing systematic reviews is beneficial since it enables the inclusion of a range of different interventions while comparing the relative effectiveness of different approaches to mental health. However, they point out that with this approach, it is inevitable that some form of double-counted data will exist, as the trials in each systematic review may be included in more than one of the included systematic reviews.

Nonetheless, systematic reviews of systematic reviews are becoming increasingly popular. In 2009, The Cochrane Collaboration introduced a new type of review: the overview of Cochrane reviews (Becker & Oxman, 2009). These summarise the findings of multiple Cochrane reviews and address the effectiveness of two or more different interventions for the same condition. They have a similar structure to standard systematic reviews, but instead of being based on primary studies, are based on systematic reviews. Although two of these new overviews should have been published by the end of October 2010, in fact only two protocols were in print by this time.

However, panoramic meta-analysis goes beyond a simple systematic review of systematic reviews by using meta-analytic techniques to create pooled effect estimates. This is similar to meta-epidemiological studies, which investigate the relationship between the treatment effect estimates of a group of meta-analyses and a particular study characteristic. For instance, Schulz et

al. (1995) conducted a meta-epidemiological study to investigate the impact of trial quality on the effect estimates of meta-analyses from the Cochrane pregnancy and childbirth database. They classified the component trials of the included meta-analyses according to whether they had used appropriate randomisation and blinding methods and then examined whether those that had used inadequate methods had provided inflated effect estimates.

Panoramic meta-analysis also uses a similar strategy to meta-epidemiological studies when assembling study data, in that it identifies relevant studies through identifying relevant meta-analyses. However, panoramic meta-analysis differs from meta-epidemiological studies in that it facilitates the assessment of the validity of generic hypotheses that apply across different medical conditions and interventions. As such, it is able to investigate whether "country" is an important factor to consider when extrapolating data for multiple intervention types and medical conditions. It was first used by Bowater, Stirling and Lilford (2009), who wanted to use the information from meta-analyses in published systematic reviews to assess the validity of generic hypotheses about the effectiveness of antibiotic prophylaxis in preventing postoperative wound infection across various types of surgery. Bowater et al. searched Medline and the Cochrane Database of Systematic Reviews for systematic reviews that:

- were published in English between 1990 and 2006
- reported the effect of antibiotic prophylaxis in preventing wound infection in surgery that involved an incision to the skin
- included studies reported in a meta-analysis
- had studies that reported outcomes based on wound infection rates
- had studies with control groups where patients received no antibiotics

The literature search identified 21 systematic reviews containing meta-analysis that reported the combined estimates of the treatment effects of 21 different types of surgery. To analyse the data, the pooled treatment effect estimate calculated for each of the systematic reviews was placed onto a forest type plot along with estimates of confidence bands. From observing these plots, Bowater et al. were able to conclude that antibiotic prophylaxis was effective in all types of surgery and suggested that it be used routinely across all surgical types.

Hemming, Lilford and Bowater (2010) have augmented Bowater et al.'s panoramic meta-analysis by re-analysing the original data using a newly proposed meta-analytic model based upon the Bayesian approach. They proposed a hierarchical model using the random effects approach to combine pooled effect estimates that incorporates a "between disease" component of variance and a "between study" (within disease) component in its overall effect estimate (pooled over all diseases). In applying this model to the data, they found that the risk of post-surgery infection was lower in almost all surgical types, except one, in the treatment groups.

Panoramic meta-analysis, therefore, is a very useful approach in drawing together the extensive systematic review literature on a topic. To do this, a systematic approach to data collection, management, and analysis has to be followed. This is discussed in the following sections.

3.4: Methodology of Panoramic Meta-analysis

As in standard meta-analysis, the panoramic meta-analytic approach aims to achieve objectivity, precision, and generalisability (Dickersin & Berlin, 1992) by including all available, relevant evidence and following these five main stages in order:

1. Formulating the scope of the panoramic meta-analysis
2. Data collection
3. Data Evaluation
4. Analysis and interpretation of data from included meta-analyses
5. Presenting findings of the analysis.

These will now be discussed in turn.

3.4.1: Formulating the Scope of Panoramic Meta-analysis

The first stage in any panoramic meta-analysis is to consider the scope of the research. The aims and research questions should be formulated with their parameters defined. The population to be investigated and the medical condition of concern also need to be defined. As with standard meta-analysis, these are all considered prior to conducting the literature search to ensure that selection and inclusion biases are minimised. This may also guarantee greater efficiency since establishing detailed, clear objectives prior to conducting the research eliminates retrieving irrelevant literature (Torgerson, 2003).

It is also at this stage that eligibility criteria are established. These are concerned with the type of meta-analyses to be included or excluded in the panoramic meta-analysis. For instance, panoramic and standard meta-analyses may contain eligibility criteria pertaining to participants, the intervention under investigation, and its comparators (Higgins & Green, 2011). In addition, in some meta-analyses, eligibility can be restricted to certain outcomes.

3.4.2: Data Collection

In panoramic meta-analysis, the methods used to source potentially relevant literature are comparable to those used in systematic reviewing and standard meta-analysis. In short, data

collection entails conducting a systematic literature search over relevant databases using a search string of relevant terms and checking the reference lists of the publications found for further relevant studies. Authors may also be contacted directly during this phase in order to gain more information on an individual study, systematic review, or meta-analysis.

As with standard meta-analysis, the panoramic approach requires an objective, comprehensive, and reproducible search of a range of sources so that as many relevant meta-analyses as possible can be identified. Doing this minimises bias and consequently assists in producing reliable effect estimates (Higgins & Green, 2011). However, data collection methods in panoramic meta-analysis differ from standard meta-analytic practices in several ways. First, rather than having to identify many different individual studies, panoramic meta-analysis needs only to identify relevant meta-analyses. The data from the studies comprising these meta-analyses are extracted and used in the panoramic meta-analysis. This means that it is a more efficient approach and can provide evidence on a broad topic quickly. Second, when trying to identify relevant studies for a panoramic meta-analysis, searches can be limited to a few databases dedicated to systematic reviews and meta-analysis, such as the Cochrane Database of Systematic Reviews. This is because all relevant studies can be captured through identifying relevant, up-to-date meta-analyses on the issue under consideration. This is not the case for standard meta-analysis, where it is recommended that many different electronic databases are searched to ensure that all relevant studies are included so as to minimise reporting bias (Critical Reviews Advisory Group, 1996). Third, the searches used in panoramic meta-analysis can be limited to certain time periods. As the widespread use of meta-analysis is relatively new, limiting the search to meta-analyses conducted within specific time periods (for example, searching only for meta-analyses from 1990 onwards)

will still identify almost all relevant up-to-date meta-analyses (Smith, Devane, Begley, &Clarke, 2011). This cannot be said for standard meta-analysis, as narrowing the search by placing limitations on search periods could result in some relevant studies not being identified, which may then bias its findings.

3.4.3: Data Evaluation

In the data evaluation stage of panoramic meta-analysis, the potentially relevant meta-analyses that have been identified are assessed for inclusion. Processes similar to standard meta-analysis are used, but the panoramic meta-analysis approach differs slightly in that it assesses both the meta-analyses and the studies that they contain. The process is as follows:

1) Initial screening of meta-analyses 'titles and abstracts to assess relevance and duplication
2) Obtain the full reports of potentially relevant meta-analyses
3) Assess the full reports of meta-analyses for relevance
4) Screening studies in meta-analyses for inclusion.

As with standard systematic reviews and meta-analysis, data evaluation should be carried out by at least two researchers using predefined and agreed criteria in order to reduce review and selection bias (Smith, Devane, Begley, & Clarke, 2011). Disagreements about study inclusion should be resolved through discussion with a third party (Bhandari & Joensson, 2009).

In standard meta-analysis, the quality of included studies is also assessed at this stage. This is important to this approach as study quality may influence the overall effect estimate it produces. Indeed, study quality has been found to affect the size of treatment effect estimates (Khan, Daya,

& Jadad, 1996; Moher et al. 1998). This is also the case in panoramic meta-analysis, where analysis is based not on the meta-analyses included, but on the trials they contain.

3.4.4: Analysis and Interpretation of Data

In the fourth stage, data is analysed and interpreted. Panoramic meta-analysis uses the same robust techniques as standard meta-analysis to combine data while also examining bias and heterogeneity that may affect the overall treatment effect estimate. Like standard meta-analysis, panoramic meta-analysis also investigates any trial characteristics that are considered important in relation to the overall treatment effect estimate. However, panoramic meta-analysis differs from standard meta-analysis in that in order to provide a broader view, it draws data together for analysis from studies that have investigated different interventions for different but related health conditions.

Issues of heterogeneity and bias will be discussed first since these may impact upon the overall analysis and what conclusions can be drawn from a study using panoramic meta-analysis.

3.4.4.1: Heterogeneity and Bias

As with standard meta-analysis, heterogeneity and bias are concerns in panoramic meta-analysis. Heterogeneity and bias are quite distinct concepts that can exist simultaneously in any meta-analytic approaches. While heterogeneity pertains to the variability of the true underlying effects between the studies, bias relates to a systematic difference between the study results and the true results arising from how the primary studies or meta-analysis were conducted (British Medical

Journal, 2011). Although both are investigated independently, bias can create heterogeneity in any set of data, and this should be considered when conducting any type of meta-analysis.

For this research, heterogeneity is important because this study investigates country as a component of heterogeneity. It is also necessary to consider both heterogeneity and bias because they may impact upon the extent to which studies can be combined and the effect estimate that is calculated. These concepts are now considered in turn.

3.4.4.1.1: Heterogeneity

In any form of meta-analysis, estimates of treatment effect will vary between the studies that have been included (Song et al. 2001). It is important to investigate between-study differences because such variations can influence the overall results of the analysis. Study results may vary by chance due to sampling error, but when there is a variation between study results greater than that expected by chance alone, this is termed statistical heterogeneity. Thompson (1994) argues that statistical heterogeneity can be a consequence of clinical heterogeneity and/or methodological heterogeneity. Clinical heterogeneity occurs when the studies included in any meta-analysis contain clinical differences. These may include differences in patient management, different types of patient, and different treatment regimes. Methodological heterogeneity refers to differences in the methods and analyses that studies have used, such as variation in how patients have been randomised, how attrition was handled, and how issues of patient blinding were tackled. However, it should also be pointed out that heterogeneity can also be caused by unmeasured study characteristics, which should be investigated when conducting a meta-analysis (Thompson, 1994) or a panoramic meta-analysis.

To investigate whether statistical heterogeneity is present in meta-analyses, statistical tests can be conducted. The Q test, as defined by Cochran in 1954 (Huedo-Medina et al. 2006), assesses whether studies are heterogeneous. It tests if the deviation of the results of the individual studies from the overall effect estimate is beyond that expected by chance. Cochran's Q test is based upon the squared difference between each study's estimated treatment effect (θ_i) and the overall combined treatment effect (θ). This is weighted by the inverse of the estimated variance of the treatment effect in each study ($w_i^{(IV)}$). This means that large, more accurate studies are given greater weight in the Q statistic than small studies.

A high Cochran's Q value indicates large differences between studies and shows, therefore, that the effects from the included studies are heterogeneous. Consequently, the null hypothesis of homogeneity should be rejected in this case. However, to validate this and ascertain its significance, a p-value from the χ^2 distribution is also often given with the Cochran's Q statistic. This is an indication of the extent of between-study variability, with a p-value of <0.10 being commonly used as the cut-off point (The Cochrane Collaboration, 2002). The equations used to calculate Cochran's Q can be seen in Box 3.1.

BOX 3.1: THE EQUATION FOR COCHRAN'S Q.

$$Q = \sum_{i=1}^{k} w_i^{(IV)} (\theta_i - \theta)^2$$

where k is the number of studies combined in the meta-analysis and where θ is calculated using this formula:

$$\theta = \frac{\sum w_i \theta_i}{\sum w_i}$$

Cochran's Q also has a chi-squared (χ^2) distribution with $k-1$ degrees of freedom (df)

Another statistical measure of heterogeneity is the I^2 Index developed by Higgins and Thompson (2002). This gives a measure of the degree of inconsistency between the study results and describes the percentage of total variability across studies that is a consequence of true heterogeneity rather than chance. The equations used in its calculation can be seen in Box 3.2.

BOX 3.2: THE EQUATIONS USED FOR THE I^2 INDEX.

$$I^2 = 100(Q - df)/Q$$

where Q is the value of Cochran's Q and where:

$$df = k - 1$$

Negative I^2 values are recorded as equal to zero so that I^2 values can range between 0 and 100%, with zero indicating that no observed heterogeneity exists across the included studies. I^2 values of 25%, 50%, and 75% are regarded as showing low, moderate and high heterogeneity, respectively (Higgins et al. 2003).

The I^2 Index has many advantages. Unlike Cochran's Q, it is not heavily dependent upon the number of studies included in a meta-analysis (Higgins et al. 2003). It can also be accompanied by an uncertainty interval, which is usually expressed with 95% confidence intervals and can be interpreted in the same way irrespective of the type of effect measure and outcome data used. It is also easy to interpret and can be easily calculated for any meta-analysis (Higgins et al. 2003).

However, when meta-analyses contain few studies, the above statistical tests have low power to detect heterogeneity. Conversely, where many studies are included, these tests can indicate differences when heterogeneity is in fact negligible (Alexander, Scozzaro, & Borodkin, 1989; Cornwell, 1993; Cornwell & Ladd, 1993; Hardy & Thompson, 1998; Harwell, 1997). Consequently, non-significant results from such tests cannot be assumed to show homogeneity of studies, and significant results cannot be assumed to show heterogeneity between studies. Indeed, Egger and Smith (2001) suggest that low power can frequently result in failing to reject the null hypothesis about the existence of homogenous results, even if inter-study differences are present in the dataset. As a result, Thompson (2001) argues that it is reasonable to dispute the usefulness of heterogeneity testing, suggesting that such tests are irrelevant. This is because all studies

within a meta-analysis will differ, both in clinical and methodological terms. This is particularly the case in panoramic meta-analysis, where, to provide a broad overview, studies that differ in clinical and methodological terms are intentionally analysed together. For this reason, a statistical test of heterogeneity was not conducted in this study.

Thompson (2001) suggests that it may be better to examine the influences of specific differences between the trials (for example, patient management) rather than depending upon a statistical test that shows the existence of heterogeneity. He suggests that if specific differences between studies can be identified, the true dissimilarities between the studies will be detected. This is supported by Colditz, Burdick, and Mosteller (1995) and Berlin (1995), who suggest that the sources of heterogeneity in a meta-analysis should be examined so as to understand the reasons for differences. Indeed, Berlin suggests that investigating the reasons why results differ between studies could lead to new insights about the relationships between study results and study protocol. This is the foundation of this study, which aims to examine country as a component of between-study variability.

3.4.4.1.2: Techniques to Investigate Heterogeneity

For any meta-analytic approach, there are two main ways to investigate heterogeneity between studies: meta-regression and stratified analysis. Meta-regression is a statistical assessment that examines whether certain study characteristics influence the size of the treatment effect across studies (Lau, Ioannidis, & Schmid, 1997), with the size of these effects being calculated using the same type of effect measures as in standard meta-analysis. Study characteristics that could be investigated may be the dose of a drug, aspects of the intervention, and/or the length of treatment.

Stratified analysis, on the other hand, is used when a particular characteristic is to be investigated (e.g. when comparing studies with old versus young patients) (The Cochrane Collaboration, 2010). In this approach, studies are sub-grouped according to a particular characteristic. Separate meta-analyses are then conducted on each group to provide a summary of treatment effect per group, its variance, and an estimate of the heterogeneity observed in each group. The overall summaries of treatment effect can then be examined to see whether there are variations in the clinical effectiveness of the intervention across sub-groups.

In this approach, all assumptions that treatment effectiveness differs between groups of studies are based upon a recognised test of statistical significance (Deeks, Altman, & Bradburn, 2001). If there are only two groups in the analysis, then the significance of the differences between groups can be assessed by comparing the z statistic (the difference between the two groups' estimates) with critical values of the normal distribution. See Box 3.3 for the equation for calculating z.

BOX 3.3: Z STATISTIC EQUATION.

$$z = \frac{\theta^1 - \theta^2}{\sqrt{[SE(\theta^1)]^2 - [SE(\theta^2)]^2}}$$

Alternatively, if there are more than two sub-groups in the stratified analysis, the significance of the difference between sub-groups can be investigated by separating the overall heterogeneity into that which can be accounted for by the differences between the sub-groups and that which remains unexplained within the sub-groups. The process for this is shown in Box 3.4.

BOX 3.4: CALCULATING THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MORE THAN TWO GROUPS IN A STRATIFIED ANALYSIS.

According to Deeks, Altman and Bradburn (2001), if the heterogeneity of the overall unstratified analysis is Q_T, the heterogeneity explained by differences between sub-groups, Q_B, is given by:

$$Q_B = Q_T - \sum_k Q_k$$

where Q_k is the heterogeneity observed within each group and where k is the number of studies. After this has been calculated Q_B can be compared with the critical values of the Chi-squared distribution with $k - 1$ degrees of freedom.

Stratified analysis can also be conducted when individual patient data is obtainable from the studies included in a meta-analytic approach, that is, when event rates in each study are available for every patient who participated. When this type of data is available, characteristics such as patients' age or ethnicity can be investigated for heterogeneity. However, since individual patient data is rarely available or readily accessible when conducting a meta-analysis, stratified analysis using this type of data cannot always be performed.

There are limitations to each of these approaches. For instance, when meta-regression is based upon a small number of studies, it will have insufficient power to detect any important effects (Petrie & Sabin, 2009), meaning that it is difficult to draw robust conclusions (Thompson & Higgins, 2002). Furthermore, in meta-regression, any relationship found is observational since, by being conducted across studies, it does not have the benefits associated with randomisation. As such, it can suffer from the same disadvantages as other observational investigations, namely,

confounding[15] and bias (Thompson & Higgins, 2002). For example, if studies were investigating the effectiveness of a new treatment and some happened to enrol patients with a higher severity of a disease, it would be difficult to tell which aspect (the treatment itself or the severity of the disease) was to cause of any difference in treatment effect between these studies and the others examined. This is also the case with stratified analysis, where the investigation into differences in treatment effectiveness between sub-groups is not a randomised comparison and so is susceptible to all the issues that arise in inferring causality when observational studies are considered.

Heterogeneity may also not be fully accounted for by meta-regression, as there may be many study and patient characteristics that could explain heterogeneity that are not considered in meta-regression (Thompson & Higgins, 2002). As with meta-regression, the findings from stratified analysis should be treated with care since any differences found can be due to factors unaccounted for or due to chance (Deeks, Altman, &Bradburn, 2001). Moreover, both meta-regression and stratified analysis can only include available data. To conduct a meta-regression or a stratified analysis, information concerning the topic of interest, the estimated treatment effect, and its variance have to be extracted from each included study, but such information is not always available or readily accessible.

In addition, stratified analysis has specific limitations. Lau, Ioannidis and Schmid (1997) state that as the analysis of sub-groups is a post-hoc exercise, it has the potential to turn into what they call a "fishing expedition". They suggest that stratified analysis can sometimes be pernicious, as

[15] Characteristics are confounded if their influence on treatment effect cannot be separated.

differential treatment effects between sub-groups may have been taken from small sub-divisions of the aggregated data and so are unlikely to represent the "truth", resulting instead in a high probability of drawing false positive conclusions. Moreover, in meta-analytic approaches, estimating the effects in sub-groups is restricted by the lack of standardised reporting of data between studies. Consequently, any analysis involving sub-groups should be based upon a scientific rationale (Hahn et al. 2000) or should be considered an exploratory method useful in generating hypotheses (Lau, Ioannidis, &Schmid, 1997).

Nevertheless, for this study, stratified analysis was chosen as the main method to investigate heterogeneity because it was considered the most appropriate approach to use when making indirect comparisons between sub-groups (Higgins, Deeks, and Altman, 2011). Moreover, its utility is further enhanced because stratified analysis is a relatively simple method for investigating heterogeneity (Madhukar, 2009).

To tackle some of the issues that arise when heterogeneity is investigated, possible sources of heterogeneity should be pre-specified when conducting stratified analysis and meta-regression. This is because pre-specification increases the credibility of any statistically significant findings, as it can be seen that such findings are not data-derived (Colditz, Burdick, & Mosteller, 1995; Deeks, Altman, & Bradburn, 2001). Furthermore, Thompson and Higgins (2002) suggest that pre-specification can protect against drawing false positive conclusions from investigations into heterogeneity. On the basis of this advice, country and continent were pre-specified as a possible source of heterogeneity in this study.

3.4.4.1.3: Bias

When using any meta-analytic approach it is important to consider the impact of any types of bias upon the results. This is because bias has the potential to seriously distort any efforts made to estimate the effect under investigation (Thornton & Lee, 2000), leading to erroneous conclusions about what the body of evidence shows.

Differential dissemination of research findings is one cause of bias in meta-analytic approaches. According to Dickersin (1990), since only a small proportion of research ever reaches publication in an indexed journal (and is therefore easily identifiable by systematic reviewers), information that may be potentially important to the results of a review remains unseen by reviewers, which may consequently lead to that review being biased. Moreover, the dissemination of research findings is greatly influenced by the nature and significance of a study's result (Egger, Dickersin, & Davey Smith, 2001). For example, research showing a statistically significant result for the beneficial effect of a treatment is more likely to be published than research showing non-significant results (Egger, Dickersin, & Davey Smith, 2001). Such occurrences are referred to as reporting biases, and these have the strongest influence upon which studies are likely to be identified for both standard and panoramic meta-analyses. Types of reporting bias are discussed below.

3.4.4.1.3.1: Publication Bias

One type of reporting bias is publication bias. This is defined in the dictionary of epidemiology (Thornton & Lee, 2000) as "an editorial predilection for publishing particular findings e.g.

positive results, which leads to the failure of authors to submit negative findings for publication." (p. 207).

Publication bias has long been recognised within the sciences, with studies from many disciplines reporting that statistically significant results are more likely to be published than non-significant results (Easterbrook et al. 1991). Indeed, Sterling, Rosenbaum and Weinkam (1995) found that in psychology journals, 95.6% of the published studies reported statistically significant results. This suggests that studies presenting results showing treatments to be beneficial are more likely to be published than studies showing no treatment benefit (or that studies showing no benefit are rarely submitted for publication to journals). In such cases, a meta-analysis based on only published research could show a false benefit of a treatment or fail to spot important adverse events related to it. This is supported by Simes (1986) who investigated whether a meta-analysis containing only published trials would produce a different treatment effect estimate than a meta-analysis containing both published and unpublished trials from an international trial registry. Using trials of different cancer chemotherapies, Simes found that when meta-analysis was based solely on published trials, the survival of patients with advanced ovarian cancer was improved by using combination chemotherapy. However, when the meta-analysis included trials from the international registry, it was found that combination chemotherapy had only a slight beneficial effect that was not statistically significant. When conducting meta-analysis, therefore, it is essential that published and unpublished literature is sourced to ensure that bias is minimised.

A factor contributing to publication bias is source funding. According to Thornton and Lee (2000) the source of funding for a trial may influence whether a trial is published or not.

Evidence suggests that trials funded by voluntary organisations or a national government are more likely to be published than those funded by the pharmaceutical industry (Easterbrook et al. 1991). Easterbrook, et al. (1991) argue that this is due to scepticism surrounding data management of trials funded by pharmaceutical companies. For example, Davidson (1986) found that many pharmaceutical company trials produce results favouring the intervention over the control. Davidson conducted a review of clinical trials to examine the difference in results between trials funded by the pharmaceutical industry and those funded by other means. He found that 89% of trials funded by the pharmaceutical industry favoured the treatment under investigation compared to only 61% of trials funded by other means. In a similar study, Bourgeois, Murthy and Mandl (2010) discovered that 85.4% of trials funded by the pharmaceutical industry reported the trial drug as having greater effect than its comparator. Etter, Burri and Stapleton (2007) suggest that the greater effectiveness of interventions found in industry-sponsored trials may be due to the amount of funding that such sponsors provide. They argue that since more resources are available, this may lead to higher levels of compliance and thus greater treatment effect. However, Egger and Smith (1998) suggest that the greater levels of intervention effectiveness found in industry-sponsored trials are due to pharmaceutical companies actively discouraging the publication of trials that they have funded and that have shown their interventions to have no effect. Consequently, a meta-analysis or a panoramic meta-analysis based only upon published studies funded by the pharmaceutical industry would be misleading since it would produce an over-stated overall effect estimate that would not occur if both published and unpublished studies were included.

3.4.4.1.3.2: Time-Lag Bias

Time-lag bias is the rapid or delayed publication of study results owing to the nature and significance of the results. It has been argued that such a bias exists, whereby studies showing positive results are more likely to be published faster (Stern & Simes, 1997). For instance, Ioannidis (1998) found that trials relating to HIV infection in the US took 4.2 years after the start of patient enrolment to be published if results were statistically significant but 6.4 years if results were non-significant. According to Egger, Dickersin and Smith (2001), time-lag bias can occur in meta-analyses even in circumstances when a large number or all studies will eventually be published. This is because trials showing treatment benefit will dominate the literature until trials showing no or poor treatment benefit are finally published. This may then lead to interventions being accepted as effective, with meta-analyses overestimating the effect of treatment, even though they are not. In these instances, panoramic meta-analysis would also over-estimate the effect of interventions as it would be pooling information from biased meta-analyses.

3.4.4.1.3.3: Language Bias

Language bias describes the situation whereby statistically significant results are more likely to be published in particular languages (Egger et al. 1997).Egger et al. (1997) found that studies that have statistically significant results are more likely to be published in English than studies that do not have such results. It has also been found that statistically significant results are more likely to be published in English language based international journals and non-significant results published in local journals (Egger & Davey Smith, 1998; Gregoire, Derderian, & Le Lorier, 1995). Bias leading to an overestimation of treatment effect could thus be introduced into meta-analyses that are solely based upon studies written in English. However, some research does

suggest that language bias does not influence the overall outcome of a meta-analysis. A retrospective analysis conducted by Juni et al. (2002) found that excluding studies that were not published in English had little impact upon the overall treatment effect estimates given by meta-analyses.

3.4.4.1.3.4: Database Bias

Database bias refers to the biased indexing of published studies within databases (Felson, 1992). While databases such as Medline and EMBASE index the majority of non-English language European journals, they do not index journals that are published in less developed or developing countries (Singh & Singh, 1994). Indeed, Egger and Davey Smith (1998) point out that of the 3861 journals published in Medline, only 30 are from India, even though this country has one of the largest research outputs. Due to this, published studies in journals not indexed by a major database are likely to be hidden from researchers undertaking systematic reviews and are therefore unlikely to be used in meta-analysis. This may then have a detrimental impact upon the overall accuracy of the meta-analysis, especially if indexed studies and non-indexed studies have different results and characteristics. For example, Song (2000) suggests that if a meta-analysis is based upon a database where the indexed studies' results systematically differ to non-indexed studies, then its literature search is likely to be biased and its overall effect estimate inaccurate. However, database bias is minimised in panoramic meta-analysis because data is gathered through relevant meta-analyses. These would have utilised many different databases in their literature searches, making the sample of studies in a panoramic meta-analysis more likely to include both indexed and non-indexed studies.

3.4.4.1.3.5: Citation Bias

Citation bias occurs when the citation or non-citation of study results is dependent upon their nature and significance (Egger, Dickersin, & Davey Smith, 2001). Citation bias is introduced into a meta-analysis when database searches are enhanced by checking the reference lists of other studies and by contacting authors in the data collection stage. This is because the citing of previous work is not necessarily objective, as studies that report a positive beneficial effect of a treatment are likely to be cited more often than those reporting null or non-beneficial effects (Egger, Dickersin, & Davey Smith, 2001; Ravnskov, 1992). Indeed, Kjaergard and Gluud (2002) found that studies that had positive results were more frequently cited in other research and reviews than those with non-significant or negative results. Since locating supportive studies is more likely when using the reference lists of other studies, citation bias may negatively affect the results of any meta-analysis by biasing the results of meta-analyses in favour of treatment (Egger, Dickersin, & Davey Smith, 2001). Furthermore, citation bias can also influence the conclusions of panoramic meta-analysis. If positive studies are more likely to be cited, then they may be more likely to be located, making them more likely to be included in meta-analyses and thus more likely to be included in panoramic meta-analysis, which in using the data from the meta-analyses as the basis of its analyses will have biased findings.

3.4.4.1.3.6: Multiple Publication Bias

Multiple publication bias occurs when a single study is published more than once and is thought to lead to problems within meta-analysis in a variety of ways (Huston & Moher, 1996). For instance, research has found that studies that give a positive result are more likely to be published more than once (Easterbrook et al., 1991), meaning that such studies are more likely to be

identified and included in a meta-analysis. In addition, it is not always clear whether the multiple publications are reporting results from the same study. This may lead to duplicate data being placed into a meta-analysis or panoramic meta-analysis, which, if the data favours the intervention, may lead to an overestimation of the intervention's clinical effectiveness. This is particularly the case when multi-centre trials are considered (Egger & Davey Smith, 1998; Leizorovicz, Haugh, & Boissel, 1992) as a meta-analysis or panoramic meta-analysis may include the data from a multi-centre trial while also including a subset of the same data reported separately by one of the centres within the multi-centre trial. However, unless the author can be contacted, it is difficult to eradicate this type of bias because it is not always easy for reviewers to discover whether two papers are duplicate publications of one trial or are two distinct trials (Egger & Davey Smith, 1998); they may report the same trial but may not have an author in common (Felson, 1992; Tramer et al. 1997).

3.4.4.1.3.7: Additional Factors

Additional factors may also introduce bias into a standard or panoramic meta-analysis. One such factor may be the provision of data by investigators. When using meta-analytic approaches, additional data that is not available in the published report is sometimes required. This, however, can be difficult to obtain, and there may be many reasons why an investigator is unable to give the information needed. For instance, investigators may no longer have the information to hand or are prohibited from providing the data by funders. In this instance, studies for which additional data was not provided would be excluded from the meta-analysis or panoramic meta-analysis, which would put these meta-analyses at risk of producing a biased effect estimate.

Finally, study quality may introduce bias into meta-analytic approaches. Many studies have shown that the size of the treatment effect estimate is greatly affected by the quality of the included studies (Egger et al. 2003; Moher et al. 1998; Wood et al. 2008c) . For instance, Khan, Daya and Jadad (1996) found that high quality studies tended to show no overall treatment benefit, whereas low quality studies showed a positive treatment effect. Therefore, it is important to consider the study quality when conducting a systematic review, meta-analysis, or panoramic meta-analysis. Indeed, Juni, Altman and Egger (2001) suggest that if the studies included in a review or meta-analysis are flawed, then the conclusions that can be drawn from the review or meta-analysis will be weakened and may be considered invalid.

To counter this problem, many researchers will restrict inclusion criteria so that only randomised controlled trials are included (Berlin & Rennie, 1999). This assumes that bias does not exist within RCTs, but Moher et al.(1998) and Schulz et al.(1995) show that this is not the case. Thus, additional steps may be required to eradicate the influence of study quality on effect estimates.

As panoramic meta-analysis can base its analysis on both data from the meta-analytic level and trial data level, it may also be important to assess the quality of the meta-analyses it includes. This should ensure that any analysis is grounded in the best available evidence. However, while several researchers have proposed guidelines and instruments for assessing the quality of meta-analyses (Cook et al. 1997; Shea et al. 2007), there is as yet no consensus across disciplines about how this should be done or which instrument to use (Moher et al. 1999). Moreover, none of the instruments developed for this purpose assess the quality of the statistical methods used in the meta-analysis (Jude-Eze. 2011). Therefore, until a consensus is reached on how meta-analysis

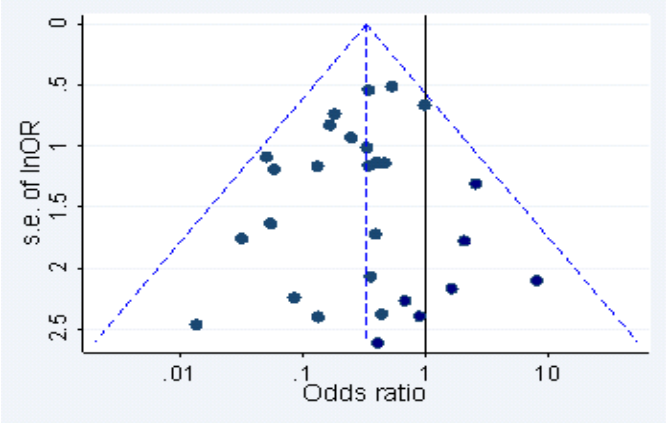
quality should be assessed and until instruments become more robust and effective, the quality of meta-analyses cannot be accurately assessed in panoramic meta-analysis.

3.4.4.1.4: Techniques to Investigate Bias

As bias may have a detrimental impact upon the estimate of treatment effect that a meta-analysis or panoramic meta-analysis produces, it is important to know how to investigate its existence.

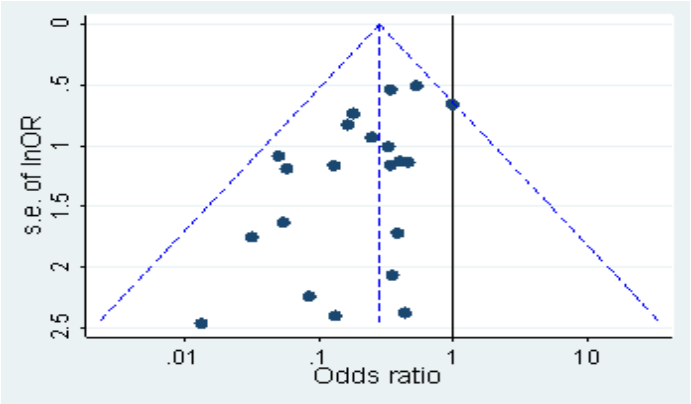
Investigating publication bias in a meta-analysis is usually conducted using funnel plots in conjunction with statistical analyses. Funnel plots are simple scatter plots that plot the treatment effect estimate of each included study (on the x axis) against a measure of precision (on the y axis) (Sterne & Harbord, 2008). In general, the precision of the estimated underlying treatment effect increases as the sample size of the included studies increases. This causes large studies to have a narrow spread and be placed at the top of the graph and small studies to scatter widely around the bottom, thus creating a funnel-type shape. If bias is not present, then the graph should look like an inverted symmetrical funnel, as in Figure 3.1. If bias is present, then the funnel plot might look like Figure 3. 2.

FIGURE 3.1: HYPOTHETICAL SYMMETRICAL FUNNEL PLOT IN THE ABSENCE OF BIAS*



*Treatment effect versus a measure of study size from Sterne, Egger & Moher (2008)

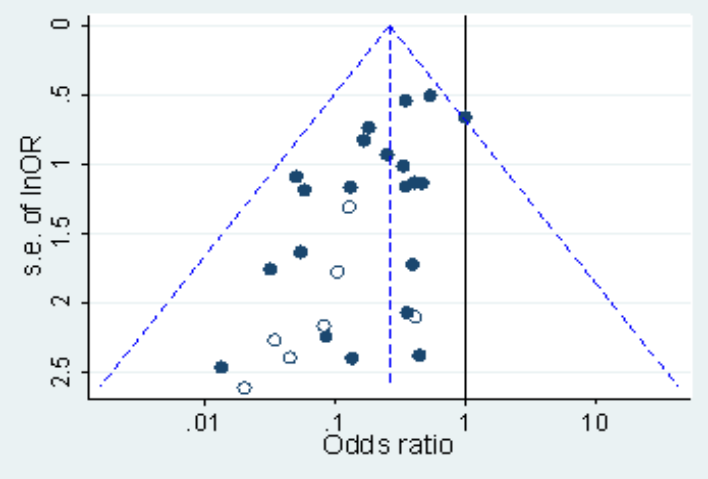
FIGURE 3.2: HYPOTHETICAL ASYMMETRICAL FUNNEL PLOT IN PRESENCE OF PUBLICATION BIAS*



*plotting the estimates of treatment effect against a measure of study size from Sterne, Egger & Moher (2008)

A funnel plot may also be asymmetrical in the presence of publication bias where smaller studies that have no significant effect remain unpublished and, therefore, are not placed on to the funnel plot. Biases due to language, citation, and multiple publication can also cause this asymmetry (Egger et al. 1997), as can chance. In addition, funnel plots can detect bias due to low methodological quality, where small trials in a meta-analysis are exaggerating the treatment effect estimate. Figure 3.3, for instance, shows a hypothetical funnel plot in which the small low quality trials (shown as open circles) are contributing to asymmetry by exaggerating the treatment effect. For both types of asymmetric funnel plot, the more prominent the asymmetry, the more significant the bias and, therefore, the higher the probability that the combined effect estimate for the meta-analysis in question will be overestimated.

FIGURE 3.3: HYPOTHETICAL ASYMMETRICAL FUNNEL PLOT IN THE PRESENCE OF BIAS DUE TO LOW QUALITY SMALL STUDIES*



* The estimates of treatment effect versus a measure of study size from Sterne, Egger & Moher (2008).

However, asymmetry in a funnel plot may not be proof of bias. The shape of a funnel plot may also be affected by the choices made in relation to its construction. Tang and Liu (2000) suggest that the shape of a funnel plot depends, to a great extent, upon what effect measure is used and how precision is defined within its construction. From their study, they found that selection bias was suggested by asymmetry in 21.7% of all the meta-analyses they examined. However, 86% of these asymmetric funnel plots became symmetrical when different definitions for precision and/or effect measure were employed, showing that the shape of a funnel plot is associated with the approach used to construct it. Furthermore, as funnel plots must be examined visually they are open to subjectivity. Therefore, funnel plots were not utilised in this study.

More formal statistical methods have been developed that examine the association between study size and its effect measure (Begg & Mazumdar, 1994; Egger et al. 1997). One such method is the Trim and Fill method proposed by Duval and Tweedie (2000). It works by omitting small studies one at a time until the funnel plot becomes symmetrical. This is called trimming. Once these studies have been removed, the trimmed plot is then used to estimate the adjusted overall effect estimate. Omitted studies are then reinstated into the plot, around the centre, along with their missing counterparts (i.e. "missing" studies). This is called filling. A new adjusted estimate of intervention effect is then calculated, along with its variance based upon this "filled" funnel plot. While this method may be useful as it provides an estimate of how many trials are missing and allows an adjusted treatment effect to be calculated, it cannot account for asymmetry other than by publication bias (Higgins and Green, 2011). Therefore, the adjusted estimate of intervention effect should be treated with caution. Furthermore, the Trim and Fill method performs poorly when considerable between-study heterogeneity is present in the dataset (Peters et al. 2007;

Terrin et al. 2003). Therefore, it would not be appropriate to use this in panoramic meta-analysis because the approach is based on studies that differ in clinical and methodological terms.

3.4.4.2: Meta-analytic Techniques

The panoramic meta-analytic approach uses similar techniques to standard meta-analysis to analyse and interpret data. Both approaches entail calculating individual treatment effect estimates for each of the included studies. The treatment effect estimate can be calculated for both binary (event-based) and continuous outcomes (i.e. weight). When data is event based, summary statistics are calculated in the form of odds ratios (OR), risk ratios (RR), or risk differences, while continuous data is usually given as the difference between the means (MD) or as a standardised mean difference (SMD). Hazard ratios are the summary statistic for survival time data. The effect estimates of each study can then be pooled to provide an overall treatment effect estimate. This can be done using a fixed or random effects approach.

3.4.4.2.1: Calculating Individual Treatment Effect Estimates

As panoramic meta-analysis involves calculating treatment effect estimates for studies from many meta-analyses that have assessed different types of outcome, all or some of these summary statistics will need to be calculated. The different summary statistics – odds ratios, risk ratios, risk difference, and mean difference – will be discussed in turn. Hazard ratios will not be discussed in this chapter because none of the meta-analyses included in this study reported outcomes in terms of Hazard ratios[16] (see Chapter 4).

[16] For further information about hazard ratios see Deeks et al. (2001).

3.4.4.2.1.1: Binary Outcomes

An outcome that takes 1 of 2 values is described as a binary outcome. For example, if the endpoint measured for a given intervention is mortality, the outcome for the patient is either death or no death and can be presented as shown in Table 3.1. In this example, if someone in the intervention group dies, they will be added to the intervention/event cell (cell a). If someone in the control group does not die, they will be added to the control/no event cell (cell d). The data from this table is then used to calculate the summary statistic for each trial.

Table 3.1. Summary information for a Binary Outcome.

	Event	**No Event**	**Group Size**
Intervention	a	b	n_1
Control	c	d	n_2
Total			N_i

The summary statistics can be expressed as either a relative or absolute measure. Odds ratios and risk ratios are relative measures, whereas the risk difference is an absolute measure. However, absolute measures of effect[17] will not be discussed in this chapter because none of the meta-analyses included in this study reported outcomes in terms of risk difference or numbers needed to treat.

[17] For further information about absolute measures of effect see Deeks et al. (2001).

The odds ratio is the ratio between the odds that an event will happen in one group and the odds of the event happening in another group. The equations used to calculate an odds ratio can be seen in Box 3.5.

BOX 3.5: THE EQUATIONS USED TO CALCULATE AN ODDS RATIO.

$$OR_i = \frac{a_i d_i}{b_i c_i}$$

with the standard error of the log odds ratio, where ln denotes the logarithms to base e, being given by:

$$SE\left[\ln\left(OR_i\right)\right] = \sqrt{\frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}}$$

Risk Ratios represent the ratio of the probability of the event occurring in the intervention group to the probability of the event occurring in the control group. These are calculated using the formula in Box 3.6.

BOX 3.6: THE FORMULA USED TO CALCULATE RISK RATIOS.

$$RR_i = \frac{a_i / n_{1i}}{c_i / n_{2i}}$$

with the standard error of the log risk ratio being provided by;

$$SE\left[\ln\left(RR_i\right)\right] = \sqrt{\frac{1}{a_i} + \frac{1}{c_i} - \frac{1}{n_{1i}} - \frac{1}{n_{2i}}}$$

When conducting a meta-analysis, the standard errors of these summary statistics are required, along with the actual summary statistics, in order to combine studies. However, problems can arise when calculating the standard error if the event rate in both the intervention and control group is zero. When this occurs, 0.5 is added to each of the cells in the table (Table 3.1) so that calculations can be made (Deeks, Altman, &Bradburn, 2001).

3.4.4.2.1.2: Continuous Outcomes

When a study has taken data from a range of values, for instance body weight, this is a continuous outcome and can be presented as shown in Table 3.2.

Table 3.2. Summary Information for Continuous Outcomes.

	Mean Response	**Standard Deviation**	**Group Size**
Intervention	m_1	SD_1	n_1
Control	m_2	SD_2	n_2
Total			N

As can be seen, continuous measures are calculated using different information than that needed for binary outcomes. In order to calculate a summary statistic from continuous outcomes, the number of participants, the mean response, and its standard deviation for both intervention and control groups are needed. Summary statistics are given as either a mean difference[18] (Box 3.7) or as a standardised mean difference (Box 3.8).

BOX 3.7: THE FORMULAS TO CALCULATE THE MEAN DIFFERENCE.

$$MD_i = m_{1i} - m_{2i}$$

with its standard error given by;

$$SE(MD_i) = \sqrt{\frac{SD_{1i}^2}{n_{1i}} + \frac{SD_{2i}^2}{n_{2i}}}$$

The mean difference measures the difference between the means of the two groups and estimates by how much, on average, the intervention has changed the outcome relative to the control. This is similar to the standardised mean difference, except that this expresses the treatment effect in terms of standard units rather than the original units of measurement. For the standardised mean difference, there are three main formulations of effect size, which differ according to the standard deviation used in their calculations (see Box 3.8).

[18] Analyses using this treatment effect measure have historically called it the weighted mean difference. This can be confusing as although a meta-analysis will calculate a weighted average of the mean difference, calculating the individual study summary statistic would not involve weighting.

BOX 3.8: CALCULATIONS FOR THE STANDARDISED MEAN DIFFERENCE.

Cohen's d:

$Cd_i = \frac{m_{1i} - m_{2i}}{S_i}$ with the standard error given by:

$$SE(Cd_i) = \sqrt{\frac{N_i}{n_{1i}n_{2i}} + \frac{d_i^2}{2(N_i - 2)}}$$

Hedges' adjusted g:

$g_i = \frac{m_{1i} - m_{2i}}{S_i}\left[1 - \frac{3}{4N_i - 9}\right]$ with its standard error provided by:

$$SE(g_i) = \sqrt{\frac{N_i}{n_{1i}n_{2i}} + \frac{g_i^2}{2(N_i - 3.94)}}$$

Unlike the Cohen's d effect estimate, this treatment effect estimate includes an adjustment in order to correct for the small sample bias.

Glass's Δ.

$\Delta_i = \frac{m_{1i} - m_{2i}}{SD_{2i}}$ with the standard error as:

$$SE(\Delta_i) = \sqrt{\frac{N_i}{n_{1i}n_{2i}} + \frac{\Delta_i^2}{2(n_{2i} - 1)}}$$

This method differs from the others as it uses the standard deviation of the control group as its scaling factor.

They also differ in whether they make corrections for small sample bias, which Deeks, Altman and Bradburn (2001)define as "the difference between the expected value of an estimate given a

small sample and the expected value if the sample is infinite" (p. 290). Because the standardised mean difference is not dependent upon the measurement scale, it can be extensively utilised as a treatment effect measure (Tian, 2007) and is particularly useful when conducting panoramic meta-analysis, where the effectiveness of different interventions needs to be compared.

3.4.4.2.2: Calculating a Pooled Effect Estimate

In meta-analysis, it may be appropriate to calculate a pooled effect estimate by combining the data from each study. This pooled effect estimate will represent the weighted mean of the included studies' treatment effects. To do this, consideration need to be given to the most appropriate summary statistic and the best approach to use to combine the studies.

3.4.4.2.2.1.: Binary Data

For binary data, studies can be combined using both relative and absolute measures of effect, namely, odds ratios, risk ratios, or risk difference. For relative measures, the individual trial's treatment effect estimates are combined on the log scale. Since none of these summary statistics is necessarily better than the others for summarising treatment effect (Higgins & Green, 2011), three criteria should be considered when deciding on which to use: consistency, mathematical properties, and ease of interpretation (Deeks, 2002).

3.4.4.2.2.1.1: Consistency

Evidence suggests that absolute measures, such as risk difference, are less consistent that relative measures (Engels et al. 2000). As such, meta-analyses should not be conducted using absolute measures unless there is strong evidence to suggest that this measure will be consistent in a

particular context (Higgins & Green, 2011). Between OR and RR, however, there is considered to be little difference in consistency (Deeks, 2002).

3.4.4.2.2.1.2: Mathematical Properties

According to Higgins and Green (2011), the most vital mathematical property for any summary statistic is the availability of a reliable variance estimate (i.e. standard error). While most summary statistics have a variance estimator and so could be used, the OR is considered to have the most mathematical properties that facilitate data combination and testing statistical significance (Deeks & Altman, 2001). This is because it is unbounded and does not, unlike the RR, rely on which of the two outcome conditions is defined as the "event".

3.4.4.2.2.1.3: Ease of Interpretation

Of all the summary statistics, the OR is regarded as the hardest to understand (Higgins & Green, 2011),whereas RR are considered to be the most intuitively comprehensible (Deeks & Altman, 2001). Absolute measures are thought to be more comprehensible to clinicians (Sinclair & Bracken, 1994) but are less generalisable than relative measures.

3.4.4.2.2.2: Continuous Data

If the outcomes of the included trials are continuous, data can be combined to provide a mean difference or a standardised mean difference. The mean difference can be used when all the studies have used the same scale for their outcome measurements. The standardised mean difference is used when all studies assess the same outcome but have measured it in different ways. In this method, each study's effect estimate has to be standardised to a uniform scale before all studies can be pooled to give an overall treatment estimate. Therefore, the size of each

study's treatment effect estimate is expressed relative to the variability observed in that study (Deeks, Altman, & Bradburn, 2001).

It should be pointed out that both the mean difference and standardised mean difference assume that the outcome measurements of each study are normally distributed. If the data is skewed, then misleading results could occur. Furthermore, standardised mean differences tend to overestimate the treatment effect when limited samples are used. However, only when small sample sizes are used is this bias substantial.

In this study, the RR was chosen as the summary statistic when event-based data was available. This was because it would be more comprehensible to most people. However, if event rates were not available or the outcome was a continuous measure, then the summary statistic from the original meta-analysis was used instead. For example, if the original meta-analysis used OR then the summary statistic used for that meta-analysis in this analysis was also an OR, and so on.

3.4.4.2.3: Approaches to Calculating Pooled Effect Estimates

In a meta-analysis, it may be considered appropriate to combine the individual study effect estimates to provide an overall treatment effect estimate (denoted by θ). If all studies collected were equally precise, then combining them would only involve computing a mean of the effect size (perhaps on a log scale). However, this is generally not the case because some studies provide more information than others. In meta-analysis, therefore, it is usual to compute a weighted mean so that those studies that provide the greatest amount of information are assigned greater weights. In these cases, each study is given a weight (w_i) related to the standard error of

the overall effect estimate ($SE(\theta)$).[19]However, how the weights are assigned to each study depends upon the way the researchers define "combined effect" (Borenstein, Hedges, & Rothstein, 2007). This is because meta-analysis has two main models for combining studies, each with different methods: the fixed effect approach and the random effects approach. The different assumptions that the models make about the nature of the individual studies lead to different definitions for the combined estimate and, therefore, different ways of assigning weights to the studies (Borenstein, Hedges, & Rothstein, 2007).

3.4.4.2.3.1: Fixed Effect Approach

The fixed effect model is based upon the mathematical assumption that there is one true effect size that all the included studies of a meta-analysis share, with no between-study heterogeneity. Any differences found in effect size between studies are considered purely due to chance. Therefore, the overall combined effect estimate produced by a fixed effect model is the estimate of this one true effect size. Weights are assigned to each study according to the amount of information that they provide: a large study would be given a large weight, and a small study would be assigned a small weight.

There are two main methods of combining studies using this model. The first is the inverse variance method, which can be utilised when the data that is needed to be combined is continuous or binary (Deeks, Altman, & Bradburn, 2001). As with all meta-analytic techniques, the individual studies' effect estimates are combined to produce a pooled effect estimate with

[19]It should also be pointed out that individual studies can also be weighted to reflect not only the quality of the trial, but also its sample size or other factors such as attrition.

weighted averages calculated for each of the individual study's treatment estimates. However, for the inverse variance, the weight assigned to each study is the reciprocal of the study's squared standard error (Deeks, Altman, & Bradburn, 2001). This means that small studies, which have large standard errors, are assigned less weight than large trials. The pooled effect estimate (θ^{IV}) within the inverse variance method can be seen in Box 3.9.

BOX 3.9: CALCULATIONS FOR THE POOLED EFFECT ESTIMATE IN THE INVERSE VARIANCE METHOD.

$$\theta^{IV} = \frac{\sum w_i \theta_i}{\sum w_i}$$

where θ_i is the individual study effect estimate and where weights (w_i^{IV}) are given by:

$$w_i^{IV} = \frac{1}{SE(\theta_i)^2}$$

The standard error of θ^{IV} ($SE(\theta^{IV})$) is given by:

$$SE(\theta^{IV}) = \frac{1}{\sqrt{\sum w_i}}$$

This method of combining individual studies has advantages in that it can be used to pool any treatment effect estimates that have standard errors available (Deeks, Altman, & Bradburn, 2001). Furthermore, by using weights that are equal to the precision of the individual studies, the variance of the pooled effect estimate is reduced (Van Den Noortgate & Onghena, 2003).

However, it should be pointed out that when event rates are low and studies are small, the standard error estimates of the individual treatment effects that this method uses to weight studies may be weak, making the associated weights unstable (Deeks, Altman, & Bradburn, 2001) and, therefore, creating misleading results.

Another fixed effect model used within meta-analysis is the Mantel-Haenszel method. This uses a different weighting scheme that is considered to be more robust when data is sparse (Deeks, Higgins, & Altman, 2008).[20]The Mantel-Haenszel method assumes that the true effect is the same across all the studies included in the meta-analysis and can only be used when data is binary. It employs the same table used to calculate the individual study's treatment effect estimates (Table 3.1) to calculate each study's weight and the overall effect estimate. Boxes 3.10, 3.11, and 3.12 show the formulas used to calculate a pooled treatment effect estimate in this approach.

[20] In situations when data is not sparse, both the inverse variance method and the Mantel-Haenszel method provide similar estimates.

BOX 3.10: FORMULAS USED BY THE MANTEL-HAENSZEL APPROACH TO CALCULATE A POOLED TREATMENT EFFECT ESTIMATE FOR ODDS RATIOS.

The pooled treatment effect estimate (θ^{MH}) is given by:

$$\theta^{MH} = \frac{\sum w_i^{MH} \theta_i}{\sum w_i^{MH}}$$

To combine all trails using odds ratios, each trial's odds ratio is given weight by using the following formula:

$$w_i^{OR} = \frac{b_i c_i}{N_i}$$

The logarithm of OR_{MH} then has the standard error given by:

$$SE[\ln(OR_{MH})] = \sqrt{\frac{1}{2}\left(\frac{E}{R^2} + \frac{F+G}{R \times S} + \frac{H}{S^2}\right)}$$ where:

$$R = \sum \frac{a_i d_i}{N_i}$$

$$S = \sum \frac{b_i c_i}{N_i}$$

$$E = \sum \frac{(a_i + d_i) a_i d_i}{N_i^2}$$

$$F = \sum \frac{(a_i + d_i) b_i c_i}{N_i^2}$$

$$G = \sum \frac{(b_i + c_i) a_i d_i}{N_i^2}$$

$$H = \sum \frac{(b_i + c_i) b_i c_i}{N_i^2}$$

BOX 3.11: FORMULAS USED BY THE MANTEL-HAENSZEL APPROACH TO CALCULATE A POOLED TREATMENT EFFECT ESTIMATE FOR RISK RATIOS.

Risk Ratios, on the other hand are weighted using the following formula:

$$w_i^{RR} = \frac{c_i n_{1i}}{N_i}$$

with the logarithm of RR_{MH} having its standard error calculated by:

$$SE\left[\ln\left(RR_{MH}\right)\right] = \sqrt{\frac{P}{R \times S}} \text{ where:}$$

$$P = \sum \frac{\left(n_{1i} n_{2i} \left(a_i + c_i\right) - a_i c_i N_i\right)}{N_i^2}$$

$$R = \sum \frac{a_i n_{2i}}{N_i}$$

$$S = \sum \frac{c_i n_{1i}}{N_i}$$

BOX 3.12: FORMULAS USED BY THE MANTEL-HAENSZEL APPROACH TO CALCULATE A POOLED TREATMENT EFFECT ESTIMATE FOR RISK DIFFERENCES.

Risk Differences are also weighted differently from Odds Ratios and Relative risks. Each trial's Risk Difference estimate is weighted by using the following calculation:

$$w_i^{RD} = \frac{n_{1i} n_{2i}}{N_i}$$

with the standard error of RD_{MH} provided by;

$SE(RD_{MH}) = \sqrt{J / K^2}$ where;

$$J = \sum \left(\frac{a_i b_i n_{2i}^3 + c_i d_i n_{1i}^3}{n_{1i} n_{2i} N_i^2} \right)$$

$$K = \sum \left(\frac{n_{1i} n_{2i}}{N_i} \right)$$

3.4.4.2.3.2: Random Effects Approach

The random effects model is based upon the assumption that the actual individual effects vary around an overall average. It assumes that the true treatment effect estimate can differ between individual studies. As with the previous model, the weight assigned to each study is the inverse of that study's variance, but unlike the previous model, it takes into account both the within- and between-study components of variance. This ensures that study weights are more balanced, prevents large studies from dominating the results, and small studies from effectively being ignored.

The method used to combine studies under the random effects approach is the DerSimonian and Laird method. To assign weights in this method, the heterogeneity statistic (Cochran's Q) needs to be broken down into the amount that can be accounted for by within-study differences and the amount due to between-study differences. This is done by firstly calculating Cochran's Q and then isolating the within-studies variance. The difference between Cochran's Q and the within-study variance will then provide the between-study variance (τ^2) (Borenstein, Hedges, & Rothstein, 2007). The formula to calculate the estimate of inter-study variance is provided in Box 3.13.

BOX 3.13: THE FORMULA FOR INTER-STUDY VARIANCE.

$$\tau^2 = \frac{Q-(k-1)}{C} \text{ if Q} > k-1$$

Or

$$\tau^2 = 0 \text{ if } Q \leq \kappa - 1$$

Where κ is the number of studies and where C is given by:

$$C = \sum w_i - \frac{\sum w_i^2}{\sum w_i}$$

As with the previous method of combining studies, this method then assigns weights to each study but includes the estimate of between-study variance (see Box 3.14 for the weighting formula).

BOX 3.14: FORMULA FOR CALCULATING THE STUDY WEIGHT IN THE RANDOM EFFECTS APPROACH.

$$w_i^{DL} = \frac{1}{SE(\theta_i)^2 + \tau^2}$$

If the between-study variance (τ^2)is larger than zero, then the weights that this model assigns to each study will be smaller and more similar to each other than the weights that would be assigned

under a fixed effect model (Deeks, Altman, & Bradburn, 2001). Therefore, this method assigns more weight to smaller studies than the fixed effect approach.

A pooled treatment effect estimate can then be calculated. The formula for this can be seen in Box 3.15.

BOX 3.15: THE FORMULA TO CALCULATE THE POOLED TREATMENT EFFECT ESTIMATE IN THE RANDOM EFFECTS APPROACH.

$$\theta_{DL} = \frac{\sum w_i^{DL} \theta_i}{\sum w_i^{DL}}$$

with its standard error given by;

$$SE(\theta_{DL}) = \frac{1}{\sqrt{\sum w_i^{DL}}}$$

As with the inverse-variance method, this model is widely applicable and can be used to combine any kind of estimates as long as standard errors are available. This model is also considered to be more conservative than the fixed effect approaches by tending to have higher estimated variances and, therefore, wider confidence intervals. However, Poole and Greenland (1999) suggest that, in some situations, the opposite can be the case because the pooled estimate can lie further away from the null value or because it provides a lower null p-value. This causes the pooled effect estimate produced by a random effects model to appear more strongly supportive of an intervention than that produced by a fixed effect model. Such a situation can arise when small or

imprecise studies are combined, as the random effects model gives greater relative weight to these studies.

In this study, both random and fixed effect approaches have been used. The reasons for this are discussed in Chapter 4 (section 4.7.1.1). The fixed effect approach selected for this study was the Mantel- Haenszel approach. This was chosen because it is more robust than the inverse variance approach when data is sparse or small trials are included in a meta-analysis, which was the case for this study. The random effects model chosen was the DerSimonian and Laird approach, which was selected because it is the most commonly used random effects approach in meta-analysis. For the universal comparison, the Mantel-Haenszel approach was selected to calculate the overall treatment effect estimate for each meta-analysis. While there is no agreement on which method is better (The Cochrane Collaboration, 2002), this fixed effect approach was chosen because it is thought to be more reliable when analysis includes small or few studies (The Cochrane Collaboration, 2002), as was the case here.

3.4.5: Presenting Findings

The final stage in a meta-analysis or panoramic meta-analysis is presenting the findings. All findings should be reported in a clear and systematic way with tables and figures used to illustrate points where appropriate. A number of factors should be considered and reported:

- The limitations of the study
- The strengths of the evidence gained
- The implications of the research
- Recommendations and future research

3.5: Why is Panoramic Meta-analysis a Good Approach to Use?

To summarise, this chapter has presented the four key reasons for using panoramic meta-analysis for this study. Because panoramic meta-analysis collects and collates a vast amount of data, it enables the testing of generic hypotheses over a set of meta-analyses concerned with different types of intervention for related health conditions. It therefore provides a broad overview of intervention types and effectiveness that can be used to inform those involved in making healthcare decisions. In addition, panoramic meta-analysis employs a more efficient approach to data collection than standard meta-analytic practices because it searches for and collects data from meta-analyses in published systematic reviews. This means that a vast number of relevant studies can be identified more quickly, and individual study data can simply be extracted from the meta-analyses. By drawing together multiple meta-analyses, there will always be sufficient information from which to reach meaningful conclusions. Moreover, panoramic meta-analysis ensures that both published and grey literature will be included in any analysis since published meta-analyses should contain both published and unpublished literature. However, if this is not the case, then panoramic meta-analysis will be limited by the quality of the included meta-analyses.

Finally, as panoramic meta-analysis has its foundations in the meta-analytic approach, robust standard meta-analysis methods can be used to explore issues of heterogeneity and bias. This is essential for this study since it facilitates an investigation into "country" as a component of between-study heterogeneity.

3.6: Conclusion

This chapter has set out the methodological position and novel approach that this study uses to achieve its research aims. Panoramic meta-analysis has been explained in its historical context, and the key differences between standard meta-analysis and panoramic meta-analysis have been presented. Panoramic meta-analysis has been selected for this study because it allows treatment effectiveness differences between countries or continents to be assessed directly and can provide a comprehensive understanding of these differences. In addition, it is more efficient than standard meta-analysis, as relevant studies are identified more quickly by searching for published systematic reviews containing meta-analyses. Furthermore, in using the statistical methods associated with meta-analysis, it can explore the issues of heterogeneity and bias.

The next chapter describes the implementation of this methodological approach by explaining how data was collected and analysed for this study.

CHAPTER 4.

INVESTIGATION METHODS: DESCRIPTIONS AND EXPLANATIONS.

As described in the previous chapter, meta-analysis requires choices to be made about search strategies, data collection, data selection, and analysis. This chapter explains how data relevant for this study was identified and provides a thorough description of how data was extracted. It also presents the reasoning for the choices that were made about data synthesis and analysis and describes how these processes were conducted.

This chapter is structured around the guidelines proposed by QUORUM[21](Moher et al. 1999), which require that authors provide information about literature searches, selection of relevant data, validity assessment, data abstraction, study characteristics, quantitative data synthesis, and trial flow, and is divided into three main sections. Section One provides details of how relevant data was identified and summarises the inclusion and exclusion criteria. Section Two presents the results of the literature search, detailing the number of meta-analyses found and data extraction processes. Section Three explains the methods of data synthesis and analysis utilised in this study and discusses the decisions made regarding these processes.

---

[21] Quality of Reporting of Meta-Analysis

4.1: The Study Topic

To meet the aims of this study, namely, to explore international differences in treatment effectiveness and ascertain whether differences existed in the effect of certain types of intervention, the research focused on cardiovascular diseases (CVDs).[22] CVDs are of increasing national and international concern and are, therefore, an appropriate clinical area through which to pursue the aims of this study.

CVDs are the leading cause of mortality in developed and developing countries (Bonneux et al. 1994; Rayner et al. 2001). The World Health Organisation (2011) estimates that CVDs cause 17.1 million deaths globally each year. An estimated 198,000 of these deaths are in the UK (British Heart Foundation, 2011). It is projected that by 2020 the two leading causes of death worldwide will continue to be ischemic heart disease and cerebrovascular disease (Murray & Lopez, 1997). Being the most common of all chronic conditions, CVDs are also a leading cause of morbidity across the world (Gaziano, Opie, & Weinstein, 2006; Poole-Wilson, 2005). Neal, Chapman and Patel (2002) suggest that the number of disability-adjusted life-years[23] (DALY's) worldwide attributed to CVDs will have increased from 134 million in 1994 to around 204 million by 2020.

Some evidence does suggest, however, that mortality and morbidity associated with CVDs are decreasing, especially in developed countries such as the UK (British Heart Foundation, 2011;

---

[22] For the purpose of this research, CVD included any condition that primarily affected the heart and/or blood vessels in the Ventral body cavity.

[23] A measure of the total burden of disease where the total burden is caused by the combination of non-fatal events and premature death.

Kuulasmaa et al. 2000). This is because the major risk factors associated with CVDs have been identified, and effective control strategies are used in health promotion and clinical practice (Reddy & Yusuf, 1998). Nonetheless, uncertainty remains about whether this decrease will continue in the future; while major risk factors such as smoking are now better prevented, detected, and treated, other risk factors such as obesity are increasing and are likely to decelerate this decrease in CVD morbidity and mortality (Lyratzopoulos, 2006).

CVDs also impose a significant burden on healthcare costs (Bonneux et al. 1994). It is estimated that CVDs cost the EU economy around €192 billion per annum (British Heart Foundation, 2008). Of this total, productivity losses due to mortality and morbidity account for 21%, 22% is attributed to the cost of informal care,[24] and 55% is direct healthcare costs. In the UK, around £29.1 billion is spent on CVDs, 60% of which is spent on direct healthcare. The remaining 40% is split between productivity losses (23%) and informal care costs (17%). The costs associated with CVDs are already very high and are likely to increase in the future (Lyratzopoulos, 2006) due to an ageing population[25] combined with increases in medical spending on new (and hence more costly, even if only in the short term) and more effective procedures for dealing with CVDs (Heidenreich et al., 2011).

Given the continued international concern about the increasing prevalence of CVDs and their associated costs, it is unsurprising that there has been an abundance of research into CVD on the worldwide stage looking at the different types of intervention used to prevent and treat CVD.

---

[24] Care provided by those who do not have a formal paid job in healthcare and in social care.
[25] Elderly people are more prone to cardiovascular diseases (National Academy on an Ageing Society, 2000)

This wealth of research makes CVD a particularly appropriate topic on which to base a panoramic meta-analysis.

4.2: The Continents under Investigation

This study had originally intended to examine differences in treatment effectiveness between countries. However, several difficulties arose that prevented country based analysis. In most cases, specific country details were not available in study reports or in meta-analyses, with many only reporting continent-level information. In such cases, it would have been necessary to contact the authors of the study reports and meta-analyses directly to collect country-level data. This would have led to additional problems: it would have been time-consuming, and where authors did not respond or where they were not able to provide this level of data, it may have led to otherwise relevant studies being excluded. Furthermore, the statistical power of any analysis based on country-level information would have been low due to the small number of studies reporting country information. Due to these limitations, but also to allow a broader picture to develop, it was decided to investigate between-continent differences in treatment effectiveness instead.

The continents included were Europe, North America, and Asia. North America included the US and Canada, and European countries were those included in the United Nations (UN) definition (United Nations Statistics Division, 2008).[26] For Asia, the UN definition was also used,[27] but

[26] Belarus, Bulgaria, Czech republic, Hungary Poland, Republic of Moldova, Romania, Russia, Slovakia, Ukraine, Aland Islands, Channel Islands, Denmark, Estonia, Faeroe Islands, Finland, Iceland, Ireland, Latvia, Lithuania, Norway, Sweden, UK, Albania, Andorra, Bosnia and Herzegovina, Croatia, Greece, Gibraltar, Italy, Malta,

Middle Eastern countries were excluded because few trials are conducted in the region and little information would be have been available about intervention effectiveness there.

These continents were chosen because they allowed for the exploration of international differences in treatment effectiveness over continents that have different healthcare systems and that are culturally diverse. This investigation is essential, as a country may be less willing to extrapolate data from one that has a dissimilar healthcare system and culture due to fears that differences in these factors has an impact on treatment effectiveness.

4.3: The Selection of Systematic Reviews Containing Meta-analyses.

As with standard meta-analysis, panoramic meta-analysis is based on data collected from relevant randomised controlled trials (RCTs). However, as stated in Chapter 3 (Section 3.4.2), the processes for identifying relevant RCTs are different. In the panoramic meta-analytic approach, systematic reviews containing meta-analyses related to the topic of interest are identified and used to identity relevant and appropriate RCTs. This enables relevant RCTs to be identified more quickly. Furthermore, it enables RCT data to be collected and collated more efficiently because relevant RCT data can simply be extracted directly from the meta-analyses in the identified systematic reviews. Before searches can be conducted for relevant systematic reviews containing meta-analysis, inclusion and exclusion criteria need to be defined.

Montenegro, Portugal, Serbia, Slovenia, Spain, The former Yugoslav Republic of Macedonia, Austria, Belgium, France, Germany, Luxembourg, Monaco, Netherlands, Switzerland.

[27]China, China Hong Kong Special Administrative Region, China Macao Special Administrative Region, Democratic People's Republic of Korea, Japan, Mongolia, Republic of Korea, Brunei Darussalam, Cambodia, Indonesia, Lao People's Democratic Republic, Malaysia, Myanmar, Philippines, Singapore, Thailand, Timor-Leste, Viet Nam.

4.3.1: Inclusion Criteria for Systematic Reviews

For the purpose of this study, published systematic reviews involving meta-analysis were eligible for inclusion if they satisfied the following criteria:

- they involved randomised controlled trials
- they were published in English - *those searching and assessing reviews for inclusion were not competent in other languages and no translator was available to assist*
- they involved a comparator group - *comparators could vary according to the type of intervention that the systematic review investigated. For instance, if the review examined a new surgical intervention, the comparator could be usual care, while for a systematic review examining a drug intervention, the comparator could be a placebo*
- they involved adult patients who had a cardiovascular disease or who were in a high risk group for this condition
- they involved therapeutic or preventative interventions that were aimed at treating a problem with the heart or blood vessels within the ventral body cavity[28]
- they reported a meta-analysis conducted over at least one fatal and/or one non-fatal endpoint.

It should be noted that fatal (and non-fatal) endpoints could differ between the included meta-analyses. For example, one meta-analysis might examine mortality at 30 days, while another might examine mortality at 6 months. This was not a concern, because these differences were taken into account when conducting this study's analysis. For the pair-wise comparisons, each continent's trial results were compared only within meta-analyses. This meant that trials from the same meta-analysis were compared only over the same fatal and/or non-fatal endpoint. For the

[28] The ventral body cavity is located at the front of the human body and is divided into three main cavities. The thoracic cavity (which is located in the chest and contains the heart, lungs, and large blood vessels), the abdominal cavity (containing the intestines and stomach), and the pelvic cavity (containing the bladder and reproductive organs).

universal comparison, each continent's trial results were standardised, allowing the comparison of results from trials that had used different outcomes and that had investigated different interventions (see Section 4.7.2 for further information on this process).

4.3.2: Excluded Systematic Reviews

Systematic reviews were excluded from this study if they:

- were only narrative reviews
- were editorials
- were opinions and reports that did not contain any outcome data
- involved problems with blood vessels outside of the ventral chest cavity

4.4: The Selection of Randomised Controlled Trials in Meta-analyses.

Data from RCTs found within the included meta-analyses were used as the basis for analysis. RCTs are considered the "gold standard" for the evaluation of clinical effectiveness because the processes involved make them less prone to bias (Haghdoost et al. 2007).

4.4.1: Inclusion Criteria for RCTs

Data from RCTs contained in meta-analyses were only used when the RCT fulfilled the following eligibility criteria:

- they were published in or after 1990 - *this was because it would be difficult to obtain the full trial reports of RCTs published before this date, which may have been needed to obtain country or continent-specific information. Furthermore, restricting the inclusion of RCTs to only those published from 1990 meant that this study could concentrate on the measurement of current international differences in treatment effectiveness, rather than on past differences.*

- they were conducted in countries in Europe, North America, or Asia
- they were published in any language – *by placing no language restriction on RCTs, reassurance could be provided, to some extent, that no language bias was introduced into the findings of this study.*

Multi-centre trials were included in this study only where all the centres involved were based in one continent. For example, a multi-centre trial with all participating centres based in Europe would be included, whereas a multi-centre trial with participating centres in Europe and Asia would not. Since data from multi-centre trials are not usually provided on a centre-by-centre basis, it would be difficult, therefore, to distinguish which data related to which continent's centres. Given the aims of this research, it was essential that data could be disaggregated at the continent level.

4.4.2: Excluded Randomised Controlled Trials

RCTs were excluded if:

- they were multinational trials
- country or continent information could not be obtained
- if they had no outcome data for both the intervention and control group.

4.5: Literature Search

A literature search was developed and conducted. This section describes the search strategy implemented in this study, the screening strategies used, and the results of the literature search.

4.5.1: Search strategy

A comprehensive search strategy was developed to ensure that all relevant systematic reviews containing meta-analyses were located. This involved systematically searching two electronic databases and limiting searches by language, dates of publication, study design, and human-only studies so that the resulting literature complied with the eligibility criteria described above. Since data from existing meta-analyses is collated for panoramic meta-analysis, searches need only be conducted on those databases that are most likely to hold systematic reviews containing meta-analysis. Panoramic meta-analysis is not the only approach in which this limited approach to literature searching has been recommended. For example, Smith et al. (2011) contend that this is also appropriate when conducting an overview of systematic reviews, and this approach has been confirmed by several studies in which literature searching has been limited to a few databases (Bowater, Stirling, & Lilford, 2009; Derry et al. 2006; van der Wouden, Bueving, & Poole, 2005; Zed, Loewen, & Robinson, 1999).

For this study, Medline and The Cochrane Library were searched for relevant published literature. The Cochrane Library was chosen because it reportedly contains the best systematic reviews produced by the world's leading systematic reviewing organisation, The Cochrane Collaboration (Fedorowicz & Al-Muharraqi, 2009). The reviews contained within The Cochrane Library are regarded as the "gold standard" (Fedorowicz & Al-Muharraqi, 2009) and have been found to have greater methodological rigour and to be updated more often than reviews found in paper-based journals (Jadad et al. 1998). Medline was chosen because it allowed identification of systematic reviews not written by The Cochrane Collaboration and because it is seen as the single most important database for those in medical and healthcare academia and practice (Pena, 1994).

It is designed to have worldwide coverage of journals and articles and is considered by some to be "the world's most comprehensive source of life sciences and biomedical bibliographic information" (Rai, 2006 p. 2).

The original searches of the databases were undertaken between September 2008 and February 2009. The Medline search covered the period January 2005 – December 2008, and The Cochrane Library search covered the period January 2000 –December 2008. The search periods for each database differed because the initial Medline search, which had used the same time period as the Cochrane search, produced over 4000 systematic reviews. Due to the extensive work involved in extracting data when using panoramic meta-analysis, and the high quality standards for data management and analysis associated with this method, there would not have been time to review all these systematic reviews and extract the required data from each. Nevertheless, given that the Medline search revealed that some systematic reviews were updates of reviews published prior to 2005 and that some overlaps existed between systematic reviews identified through Medline and The Cochrane Library, (see figure 4.1), it is reasonable to assume that few systematic reviews were overlooked as a result of using different search periods between the two databases. Furthermore, as this study used trial data from the meta-analyses of these systematic reviews, which could be expected to have included all relevant up-to-date trials, the different search date used for each database was less of a concern than would be the case for standard meta-analysis. Inclusion of older systematic reviews from Medline (i.e. published before 2005) may not necessarily have enhanced the data for this research since the older trials included in newer reviews may well have been included in earlier systematic reviews.

A different search strategy was used for each database. The Cochrane Library gave direct access to all systematic reviews indexed under the Medical Subject Heading (MeSH) of *Cardiovascular Diseases* simply by conducting a MeSH search. Thus using "Cardiovascular Diseases" in the MeSH search box produced not only the MeSH tree for that term, but also an option for viewing the results of all systematic reviews that were indexed under this MeSH descriptor. The required limitations could then be placed on these viewed results to find all relevant systematic reviews involving meta-analysis.

To identify relevant systematic reviews from Medline, search strings were created. Search strings were developed by consulting the MeSH tree for the MeSH term *Cardiovascular Diseases* within Medline. To do this, the term "*Cardiovascular Diseases*" was placed into the MeSH Database search box and the results scanned to find the terms that would provide systematic reviews on all conditions indexed under *Cardiovascular Diseases*. This process indicated that four sub-headings of the *Cardiovascular Diseases* MeSH tree would be sufficient to provide all relevant reviews pertaining to cardiovascular conditions (see Table 4.1).

As some cardiovascular diseases may have been indexed under the MeSH term *Vascular Diseases,* another search string was created to ensure that no systematic reviews concerned with cardiovascular illnesses had been missed (see Table 4.1). A third search string was created to ensure that no systematic reviews containing meta-analysis had been missed in the Medline searches. This time, filters such as [ti], for title, and [pt], for publication type, were used to search for publications that had words relating to meta-analysis in the title but were not indexed by the publication type of meta-analysis (see Table 4.1). This search string was "Cardiovascular

diseases" AND ("Meta – analysis" OR "Metaanalysis" OR "Metanalysis" OR "Systematic review" OR "Pooled analysis") [ti] NOT meta – analysis [pt].

Table 4.1: Individual Search Strings used in Medline.

Search Strings used in Medline
"Cardiovascular abnormalities" OR "Cardiovascular infections" OR "Heart diseases" OR "Pregnancy complications, Cardiovascular"
"Vascular diseases" NOT ("Cardiovascular abnormalities" OR "Cardiovascular infections" OR "Heart diseases" OR "Pregnancy complications, cardiovascular").
"Cardiovascular diseases" AND ("Meta – analysis" OR "Metaanalysis" OR "Metanalysis" OR "Systematic review" OR "Pooled analysis") [ti] NOT meta – analysis [pt].

4.5.2: Screening Strategies

A three-stage screening process was adopted to determine which systematic reviews were included in this study. This involved two reviewers, LH and RB. LH reviewed the titles and abstracts of each systematic review identified in the literature search, while both LH and RB independently reviewed the full text of any systematic review thought relevant by LH. LH then screened the trials in each meta-analysis. Any systematic reviews that appeared to be duplicates were independently screened by both LH and RB, and the most recent publication of the review was retained. Any disagreements between the two reviewers were resolved by discussion. This was the case in two instances where it was unclear whether the systematic reviews were duplicates. After further investigation, it was determined that both these reviews were duplicates and were therefore removed from the dataset.

Stage one of the screening process involved initial screening of the titles and abstracts of the systematic reviews identified via the database searches in order to exclude publications that were not pertinent to this study. At this stage, publications were excluded if they were not systematic

reviews, did not include the previously specified types of patients, or did not investigate an intervention aimed at treating or preventing CVDs. In stage two, a more detailed examination of the potentially relevant publications found in stage one was conducted. This involved examining the full text of the systematic reviews and excluding any that did not involve meta-analysis or RCTs. Stage three involved the screening of all the trials included in the systematic reviews to ensure that all trials in each review satisfied the inclusion criteria for RCTs outlined above. At this third stage, systematic reviews were excluded if:

- all the trials they contained were published before 1990
- all the trials they contained were conducted in one continent[29]
- all trials they contained were not from the continents of interest
- all trials were multinational[30]
- the continents for all trials were unknown
- they did not provide outcome data for both intervention and control groups for all trials contained in the meta-analysis.

4.5.3: Literature Search Results

A total of 2269 relevant systematic reviews involving meta-analysis were identified in the literature search: 2156 via Medline and 113 through The Cochrane Library. Thirty-seven systematic reviews appeared in both Medline and The Cochrane Library (see Table 4.2)

.

[29]To conduct between-continental comparisons within meta-analyses, as well as over meta-analyses, the trials in each meta-analysis could not all be from one continent. Furthermore, conducting analysis within meta-analyses meant that trials that had investigated the same intervention but were from different continents could be compared.
[30] Multinational trials were not included in this study as they did not provide country by country or continent by continent breakdowns of results which were essential for this analysis.

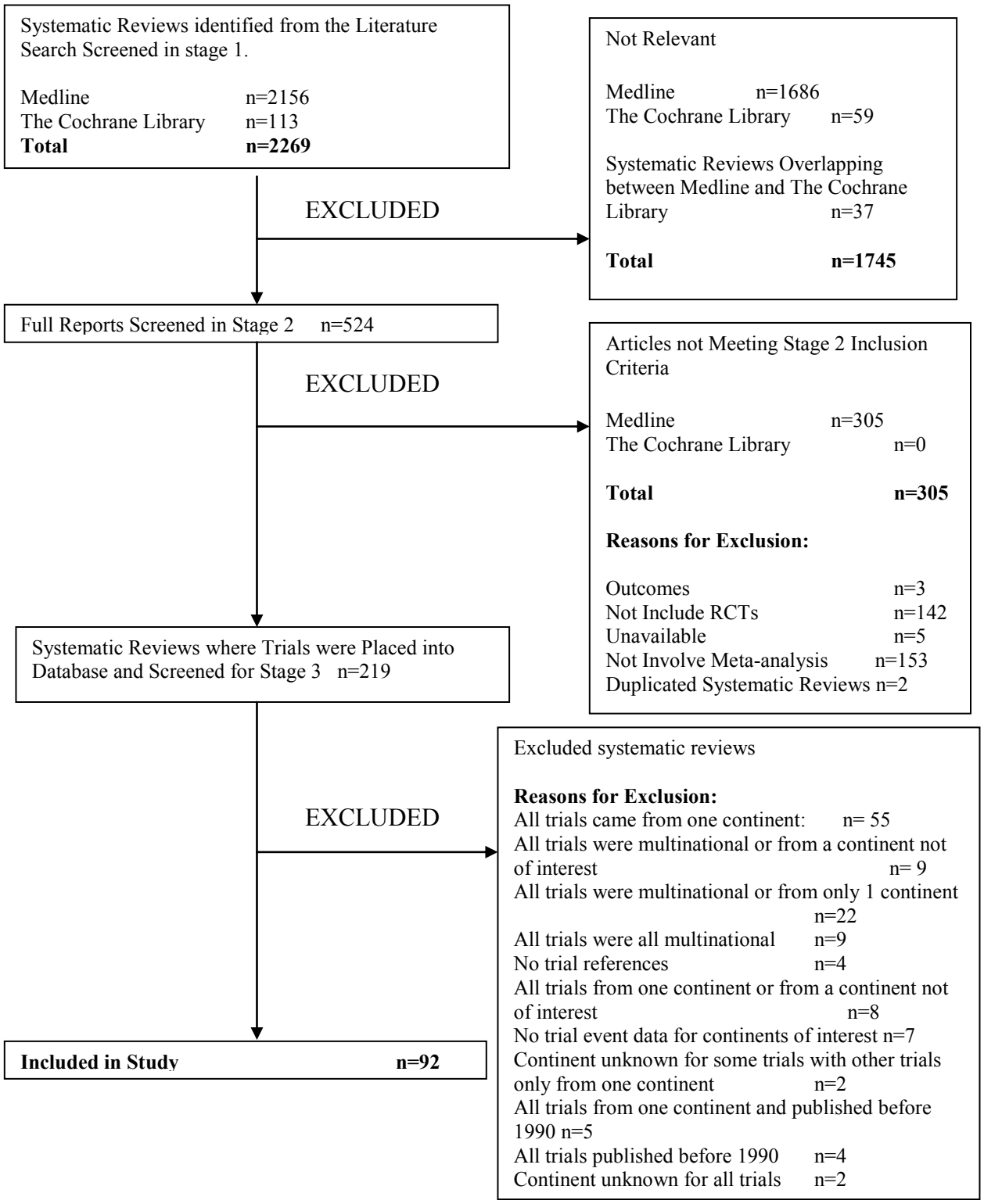
Table 4.2: The Individual Yields for Each Search String

Database	Search String	Yield
Medline	"Cardiovascular abnormalities" OR "Cardiovascular infections" OR "Heart diseases" OR "Pregnancy complications, Cardiovascular"	1360
Medline	"Vascular diseases" NOT ("Cardiovascular abnormalities" OR "Cardiovascular infections" OR "Heart diseases" OR "Pregnancy complications, cardiovascular").	754
Medline	"Cardiovascular diseases" AND ("Meta – analysis" OR "Metaanalysis" OR "Metanalysis" OR "Systematic review" OR "Pooled analysis") [ti] NOT meta – analysis [pt].	42
The Cochrane Library	*Cardiovascular Diseases*	113
	Overlap between Medline and The Cochrane Library	37

Of these 2269 publications, 1745 were excluded in the first round through the first-stage screening process outlined above (see Figure 4.1). These reviews did not include relevant patient populations, relevant conditions, or were not systematic reviews. The full texts of the remaining 524 publications were then screened by both reviewers.

As shown in Figure 4.1, the second stage of screening resulted in the exclusion of 303 systematic reviews. The remaining 221 systematic reviews were then entered into the database for this study and screened further for inclusion (see Section 4.3). This confirmed whether all the RCTs included in the meta-analyses of the included systematic reviews were relevant and satisfied the inclusion criteria described above.

FIGURE 4.1: FLOW CHART OF STUDY SELECTION.



A further 127 systematic reviews were excluded at this point because the trials included in the meta-analysis did not fulfil the eligibility criteria listed earlier. Systematic reviews were excluded if trials in their meta-analysis met one of the following criteria:

- they contained trials from only one continent
- they contained trials from only one continent of interest, with other trials from continents not of interest
- some trials in the meta-analysis were multinational, with all other trials from continents not of interest
- they contained multinational trials, with other trials from only one continent of interest
- all trials were multinational
- included trials were only from one continent of interest, with other trials published before 1990
- all trials were published before 1990
- no references were given for included trials
- no outcome data was provided for relevant trials
- the continents of the included trials were unknown
- the continents of the included trials were unknown and other trials were only from one continent.

Additional details are provided in Figure 4.1. Following the third round of exclusions, 92 systematic reviews containing meta-analysis were eligible for this study.

4.6: Data extraction and Management

Data extraction was carried out by LH and was independently checked by RB. Data from the relevant systematic reviews was extracted into a database developed by LH. Data relating to the systematic reviews, their meta-analyses, and the trials contained within them were extracted.

For each systematic review, reference data was extracted (i.e. author, title, publication year, and journal). Information was also recorded about the intervention investigated and was categorised into one of the five following intervention types: drug, device, surgery, lifestyle, and management. The comparator treatment and the cardiovascular disease explored were recorded, as was the fatal and/or non-fatal outcome assessed in each systematic review. Recording fatal and non-fatal details was relatively easy in some cases as the systematic review contained only one fatal and/or non-fatal meta-analysis. In other systematic reviews, however, more than one fatal and/or non-fatal meta-analysis was reported. In these cases, the most commonly reported fatal and/or non-fatal meta-analysis in each systematic review was chosen, that is, the fatal (or non-fatal) endpoint whose meta-analysis had combined the highest number of trials. For instance, in a systematic review where there were two meta-analyses for fatal endpoints and one combined seven trials and the other five, then the former was identified as the most commonly reported fatal endpoint. Where the same number of trials were combined in fatal and/or non-fatal meta-analyses, the endpoint chosen was the one whose meta-analysis combined the greatest number of patients.

Trial information for each trial contained in included systematic reviews' meta-analyses was recorded in the database. For each trial, the following data was recorded:

- author
- publication year
- event rate for fatal and/or non-fatal endpoints
- trial effect estimate and its variance estimate
- the type of effect measure used in each fatal and/or non-fatal meta-analysis
- the country or continent in which each trial was conducted.

Information about the country or continent in which each trial was conducted was initially sought from the systematic reviews. Where this was not specified, the trial's abstract was examined, and if still not found, the full text of the original trial report was examined. However, the countries or continents in which the trials were conducted were not always reported. In such cases, trial authors were contacted directly, but the majority of authors did not respond or did not provide the country or continent information required. To overcome this problem, when all authors were based in the same country or continent, trial country or continent was determined by the addresses of the authors themselves. If no country or continent information could be found for a trial, then that trial was excluded from the database and, therefore, from analysis.

4.7: Data Synthesis and Analysis

In panoramic meta-analysis, as in standard meta-analysis, data needs to be synthesised and analysed. In the panoramic meta-analytic approach, this can be done using pair-wise comparisons and a universal comparison. This section first explains the decisions that had to be made prior to data synthesis and analysis. Section 4.7.2 then describes the methods used for the comparisons.

4.7.1: Decisions Surrounding Data Synthesis and Analysis

In this study, three main decisions had to be made:

1. Which meta-analyses should be included in each pair-wise comparison?
2. Which type of effect measure should be used?
3. What should be done with overlapping meta-analyses?

4.7.1.1: Meta-analyses Included in Each Pair-wise Comparison

The continents studied in this research (Europe, North America, and Asia) were compared in pairs – Europe and North America, Europe and Asia, and North America and Asia – before a universal comparison that considered all three continents together was conducted. For these comparisons, it was important to decide which meta-analyses from the relevant systematic reviews would be included.

For the comparison of treatment effectiveness between Europe and North America, meta-analyses had to contain results from at least two recent (since 1990) trials conducted in Europe and at least two recent trials conducted in North America. Trials had to be recent so that this study could investigate differences that exist currently. Including meta-analyses that had two trials from each continent enabled heterogeneity within continents to be taken into account by using a random effects model. The analysis was also carried out using a fixed effects approach because this is considered to have increased statistical power and yields a more precise estimate of the pooled treatment effect than a random effects model (Cohn & Becker, 2003). In doing this, however, the overall treatment effect estimate calculated for each continent was found to differ between the random and fixed effect models, and so it was decided to report the results of the comparisons from the random effects model.

For the pair-wise comparisons between Europe and Asia and between North America and Asia, meta-analyses only needed to contain at least one recent trial from each continent. This was due to the lack of Asian trials in some meta-analyses, with most of the included systematic reviews containing only one trial from Asia. If the selection criteria for meta-analyses had been the same

as in the previous comparison, (i.e. two trials from each continent), there would have been too few meta-analyses containing Asian trial data to make meaningful comparisons.

For these pair-wise comparisons, trials from each continent were combined using a fixed effect model: the Mantel-Haenszel method. In doing this, however, the between-study heterogeneity for meta-analyses that included continents with more than two trials could not be included in the analysis and the impact of its inclusion studied. Therefore, all results for pair-wise comparisons obtained using the fixed effect model were checked and verified by re-running the analysis using a random effects model. As expected, the standard error and confidence intervals calculated using this model were larger than those using the fixed effect model. However, the overall treatment effect estimates for each continent calculated by both analytic models were similar, meaning that, for these comparisons, the fixed effect model was appropriate to report.

4.7.1.2: Type of Effect Measure

Decisions had to be made about the type of effect measure to use. To allow for a "cleaner"[31] comparison of treatment effectiveness between continents, trials were combined over RR, where possible, for the pair-wise comparisons and for calculating the best treatment effect estimates for each meta-analysis in the universal comparison. RR were used for each meta-analysis even if the original meta-analysis had combined trials using OR or MD. However, this was only possible where the endpoint reported in the meta-analysis was event-based and event rates were available.

[31] Fewer meta-analyses measuring treatment effect in different ways would make the comparisons more comprehensible.

If no event rates were reported and the results were given as OR, then trial results were combined using OR. When neither RR nor OR were available, MD were used.

Before being used in analysis, ratio summary statistics, that is, RR and OR, underwent natural log transformations. The common feature of ratio summary statistics is that they take on values between zero and infinity, with a value of one corresponding to a neutral value.[32] This means that they are easy to interpret. For example, a RR of 0.5 shows the intervention to be effective, whereas a RR of 1.5 shows the control group to be better. However, the number scale for these statistics is not symmetrical. For instance, a RR of 0.5 is opposite to a risk ratio of 2, such that when averaged they should produce 1. However, this is not the case, as when averaged they produce a RR of 1.25. Therefore, to make the scale symmetrical, natural log transformations are performed on ratio summary statistics. This enables ratio summary statistics of the same size but in opposite directions to be equidistant from one while also making the confidence intervals of the combined treatment effect estimates symmetrical around the point estimate (Egger & Davey Smith, 2001). In doing so, they ensure the results of this study are more easily comprehended.

4.7.1.3: Overlapping Meta-analysis

A decision had to be made about how to deal with overlapping meta-analyses. It was decided that overlapping meta-analyses were those that shared at least two trials with at least one other meta-analysis in the database. This occurred where different systematic reviews had addressed similar clinical questions and was a concern for this research because it meant that some trial results

[32] A neutral value means no difference between groups. For a meta-analysis examining an intervention, 1 corresponds to no intervention effect.

would be duplicated in the panoramic meta-analysis. This may be problematic if duplicated trial results favour the intervention in one continent over another. When meta-analyses are pooled, therefore, evidence for international differences in treatment effectiveness may appear stronger than the evidence really suggests, with interventions appearing to perform better in the continent with the most duplicated trials.

Despite these concerns, overlapping meta-analyses were included in the dataset because it might otherwise have been too small to detect international differences in treatment effectiveness. Nevertheless, the impact of including overlapping meta-analyses upon the results of this study was considered. For each type of endpoint in each pair-wise comparison, the number of overlapping and non-overlapping meta-analyses was noted. The percentage of overlapping and non-overlapping meta-analyses that favoured the stronger[33] continent was then calculated. The percentages were then compared to see if the inclusion of overlapping meta-analyses had exaggerated the percentage of meta-analyses favouring the stronger continent. For the universal comparison, the impact of including overlapping meta-analyses was taken into account by running the analysis both with and without the duplicated trial results from overlapping meta-analyses.

4.7.2: Methods of Synthesis and Analysis

All comparisons were conducted on data from the meta-analyses in the 92 included systematic reviews using the Comprehensive Meta-analysis Version 2 Package, SPSS, STATA, and Excel.

---

[33]The continent that was favoured by most of the included meta-analyses.

4.7.2.1: Conducting Pair-wise Comparisons

The pair-wise comparisons conducted between Europe and North America, Europe and Asia, and North America and Asia were carried out for both fatal and non-fatal endpoints, with analysis conducted both within and over meta-analyses.

The first task in conducting these pair-wise comparisons was to calculate individual estimates of treatment effectiveness for each continent in the comparison in each of the identified meta-analyses. Random and fixed effect approaches were used for combining the trial results for each continent in each meta-analysis. For both approaches and in each meta-analysis, each continent's treatment effect estimate was calculated using the event rates of each continent's trials to produce an overall RR measure of treatment effect for that continent. This was done using the log scale, as described in the previous section. When data was binary and no event data was available for either the intervention or control group in a trial, 0.5 was added to all cells so that the treatment effect estimate could be calculated in line with the technique recommended by Deeks et al. (2001). This is done because computational problems in calculating the treatment effect estimate can occur when one or both groups contain zero events. Where event rates were unavailable, trials were combined using their original effect measure.[34]

Next, for each meta-analysis, the difference between the two continent's effect estimates was calculated. This was calculated using the formula in Box 4.1.

---

[34] As stated in Section 4.4.1.

BOX 4.1: THE FORMULA TO CALCULATE THE BETWEEN-CONTINENT DIFFERENCE IN TREATMENT EFFECT.

$$DE = \theta^{c1} - \theta^{c2}$$

where θ^{c1} is the treatment effect estimate of continent one and θ^{c2} is the treatment effect estimate of continent two.

A positive (+) or negative (-) symbol was used to indicate the continent in which the intervention performed best because if continent two had a larger treatment effect estimate than continent one (when continent two's estimate was subtracted from continent one's estimate), the sign of the calculated difference would be negative.

The standard error for the between-continent difference in effect estimates was also calculated using the formula in Box 4.2.

BOX 4.2: THE FORMULA TO CALCULATE THE STANDARD ERROR FOR THE BETWEEN-CONTINENT DIFFERENCE IN TREATMENT EFFECT.

$$SE(DE) = \sqrt{[SE(\theta^{c1})^2 + SE(\theta^{c2})^2]}$$

where the $SE(\theta^{c1})$ is the standard error for the effect estimate of continent one and $SE(\theta^{c2})$ is the standard error for the effect estimate of continent two.

The z-score, confidence intervals, and 2-sided p-values were also calculated for the between-continent difference in treatment effect for each of the relevant meta-analyses.[35] The number of meta-analyses favouring[36] each continent was then summed. Binomial sign tests were then conducted to discover whether the difference in the number of meta-analyses favouring one continent over the other was statistically significant at the 5% level.

For each pair-wise comparison, the types of intervention prone to inter-continental differences in treatment effect were also investigated. For each pair-wise comparison, the included meta-analyses were grouped according to the type of intervention investigated (Drug, Device, Surgery, Lifestyle or Management). Then, within each intervention group, the Z-statistic and 2-sided p-value calculated for the between-continent difference in treatment effect for each meta-analysis

[35] These were calculated in Excel using the formula for a z-test
[36] Where the intervention performed better in one continent over the other – significance not required

were observed to find any differences that were statistically significant at the 5% level. This was done in order to see if there was any clear pattern of statistically significant differences for any intervention type. This would show the types of intervention that were more or less prone to international differences in treatment effect over all of the included meta-analyses.

Global estimates of continental difference were also produced for each pair-wise comparison. These showed the average between-continent difference in treatment effect over all of the included meta-analyses. They were calculated by combining the estimates of between-continent difference of all the included meta-analyses, on the log scale, using the fixed effects model. This was done for each type of effect measure, for example, a global estimate was calculated for all meta-analyses where risk ratios had been used, while a separate global estimate was calculated for those meta-analyses that had provided odds ratios.

4.7.2.2: Conducting the Universal Comparison

In addition, a universal comparison was conducted to examine the spread of each continent's individual trial results. To do this, the result of each European, North American and Asian trial in each included meta-analysis was standardised.[37] The standardised trial results were then extracted from their meta-analysis and pooled together. They were then plotted onto graphs according to the continent in which they were conducted. This was done for fatal and non-fatal endpoints.

---

[37]This was done within meta-analyses and meant that each trial's intervention effect estimate was only being compared to trials that had investigated the same intervention using the best available effect estimate for the intervention investigated.

Standardising the effect measure allowed for different interventions that had measured treatment effect in different ways to be compared. Furthermore, standardised measures always follow a standard normal distribution, which has a mean of zero and a standard deviation equal to one. Consequently, the two standard deviations rule (Anderson & Finn, 1996) was used to look for outliers, that is, to identify trials that fell over two standard deviations either side of the mean. The standardisation formula used in this study is shown in Box 4.3.

BOX 4.3: STANDARDISATION FORMULA.

$$Y_i = \frac{\theta_i - \mu}{SE_i}$$

where θ_i is the ith trial's effect estimate, μ is the combined estimate of all of the trials from all three continents in each meta-analysis (combined using the Mantel Haenszel approach) and the SE_i is the standard error of the deviation between the two estimates. The standard error is given by:

$$SE_i = \sqrt{s_i^2 - s^2}$$

where s_i is the standard error of the effect estimate of the ith trial and s is the standard error of μ.

NOTE: *this formula is only exact when the Inverse Variance method is used to estimate* μ. *Therefore, it is only approximately correct for use with the Mantel Haenszel method that was utilised for this part of the investigation.*

The standardisation of each trial's treatment effect estimate and plotting these onto graphs according to the continent in which they originated facilitated an investigation into whether trials

from some continents were more likely to provide positive results favouring the intervention over the control. Binomial sign tests were then conducted for each continent to see whether the difference in the number of trials favouring the intervention relative to the control was statistically significant.

For each continent, the mean effect size was also calculated along with a z-score and a p-value. From these, the difference in effect size between each continent could be observed. While binomial sign tests make few assumptions about the nature of the distributions being tested, meaning they have general applicability, they throw away useful data and as such, may lack statistical power. Therefore, t-tests[38] were conducted between Europe and Asia and between North America and Asia to observe the direction and significance of results, as they are considered to have more statistical power and, therefore, to be more robust (Searle, 1999).

The standardised effect measures were used to further investigate and identify the types of intervention prone to inter-continental differences. This was done by grouping each trial's standardised effect estimates by continent and then sub-grouping these into intervention type. T-tests were then performed between continents for each type of intervention. This meant that comparisons of the effect estimate of each intervention type could be made between Europe and Asia and between North America and Asia.

[38] T-tests could not be conducted in the pair-wise comparison (that had also used the binomial sign test) since this investigation was conducted at the meta-analysis level and not the trial data level.

4.8: Conclusion

This chapter has established the importance of this study and explained and justified the approaches taken.

This investigation is important as it may have international implications since it will provide evidence about when it is appropriate to directly extrapolate overseas clinical data. As such, it has the potential to impact upon the speed at which new cardiovascular interventions are approved across countries and save money when using such information as the basis for decisions on the implementation of a new intervention. In addition, this may influence the types of CVD intervention that are recommended at both national and international levels. For instance, if it is known that a certain type of intervention is less effective in one country than another, then the country in which it is less effective may be less likely to recommend this type of intervention in its guidelines and treatment management strategies. If, on the other hand, international differences in effectiveness are known not to exist in a certain type of intervention then an international guideline can be developed and used.

This chapter has provided a thorough description of the methods used for identifying relevant systematic reviews containing meta-analysis and the processes and tests used to conduct panoramic meta-analysis. In addition, the reasons that particular methodological decisions were made have been explained, and the processes used to counter any limitations of these decisions and methods have been described and justified.

Part one of this thesis, then, has identified and explained the importance of this research and the novel methodological approach that will be used. It has also described the processes around data collection for this study and explained how data was analysed.

Part Two turns to the actual results of the study. The next chapter presents the results of the first pair-wise comparison conducted between trials from Europe and North America.

PART TWO: THE INVESTIGATION OF INTERNATIONAL DIFFERENCES IN TREATMENT EFFECTIVENESS

CHAPTER 5.

AN INTER–CONTINENTAL COMPARISON OF TREATMENT EFFECTIVENESS BETWEEN EUROPE AND NORTH AMERICA.

To investigate the existence of inter-continental differences in treatment effectiveness three pair-wise comparisons were conducted using secondary data from meta-analyses of cardiovascular interventions as described in the previous chapter. This chapter presents the findings from the first of these pair-wise comparisons: treatment effectiveness differences between Europe and North America for both fatal and non-fatal endpoints. The results are separated into two parts. In the first part, findings are presented for fatal endpoints, and in the second part, the findings for non-fatal endpoints are presented. The chapter then concludes by summarising the overall findings for this pair-wise comparison.

5.1: The Pair-wise Comparison.

The pair-wise comparisons allowed for between-continent differences in treatment effectiveness to be investigated at the meta-analysis data level. That is, data from each of the continents was compared both within and over meta-analyses.

As stated in Chapter 4 (Section 4.4.1), each meta-analysis, for the pair-wise comparison between Europe and North America, had to contain at least two trials in each continent. Of the 92 systematic reviews that contained meta-analyses relevant for this study, only 59 had a meta-analysis that satisfied this criterion. The findings in this chapter are based on these 59 meta-analyses.

5.1.1: Fatal Endpoints.

Only 47 of these 59 meta-analyses contained two trials from each continent with fatal endpoint data that allow for separate estimates of treatment effect to be calculated for each continent. Details of these meta-analyses, including details of the intervention and control groups per meta-analysis, endpoint reported and number of trials each meta-analysis contained from Europe and from North America are shown in Table 5.1 to 5.4. The effect estimate for each continent in each meta-analysis is also provided, as is the identity of the continent in which the intervention performed best.

5.1.1.1: Meta-analyses Grouped by Treatment Effect Measure for Fatal Endpoints

The 47 meta-analyses that contained fatal endpoint data have been grouped on how treatment effect was measured. For example, Table 5.1 details the 37 meta-analyses that measured treatment effect using RR and Table 5.3 details the six meta-analyses that used the OR.

5.1.1.1.1: Fatal Endpoints Measured as Risk Ratios

37 of the 47 meta-analyses that reported a fatal endpoint did so using a RR. The details of these studies are summarised in Table 5.1.[39]

Table 5.1: Meta-analyses with Fatal Endpoint Data. Treatment Effect Measured by the RR.

Article Ref. no	Intervention	Control	End point	No. of trials (Europe/ N. Amer.)	RR estimate for Europe *(95% CI)*	RR estimate for N. Amer.*(95% CI)*	Best Region (p-value of between-continent difference)
178 [C22]	Digitalis for CHF	Placebo	All-Cause Mortality	2/3	0.55 *(0.12, 2.53)*	0.99 *(0.93,1.06)*	Europe (0.447)
179 [C23]	Omega 3 fatty acids for CVD	Placebo, no supplement or usual diet	Cardiovascular mortality	8/2	1.00 *(0.73,1.38)*	0.25 *(0.03, 2.24)*	N. Amer. (0.219)
304 [C25]	Psychological interventions for CHD	Usual care or no intervention	Cardiovascular mortality	2/6	0.70 *(0.32, 1.52)*	0.93 *(0.67, 1.28)*	Europe (0.505)
18 [C12]	Phosphodiesterase III inhibitors for HF	Placebo	Cardiovascular mortality	2/5	1.55 *(1.01, 2.37)*	1.03 *(0.72, 1.49)*	N. Amer. (0.160)
223 [P96]	Beta-blockers without ACE inhibitors	Placebo	All-cause mortality	2/3	0.93 *(0.35, 2.47)*	0.72 *(0.45,1.17)*	N. Amer. (0.641)
313[P75]	Warfarin plus aspirin after MI or ACS	Aspirin only	All-cause mortality	3/2	0.93 *(0.71, 1.22)*	1.63 *(0.33,8.02)*	Europe (0.498)
216 [P62]	Facilitated PCI for ST-elevation MI	Primary PCI	Short-term death (all cause)	8/2	1.07 *(0.57, 1.99)*	0.64 *(0.08, 5.10)*	N. Amer. (0.644)
103 [P133]	Combined aspirin-oral anticoagulation therapy	Oral anticoagulation therapy alone	All-cause mortality	4/2	1.13 *(0.88, 1.45)*	0.61 *(0.15, 2.41)*	N. Amer. (0.382)
131 [P114]	Implantable cardioverter defibrillators for LVSD	Usual care	All-cause mortality	2/3	0.82 *(0.62, 1.08)*	0.75 *(0.63, 0.89)*	N. Amer. (0.584)
227	Anti-arrhythmics	Placebo, drugs	All-cause	6/4	1.00	1.43	Europe

[39]Best Region is defined as the continent in which the intervention performed best.

[C4]	after cardioversion of AF	for rate control or no treatment	mortality		*(0.78, 1.29)*	*(0.76, 2.70)*	(0.301)
389 [P41]	Rescue PCI after failed fibrinolytic therapy for MI	Conservative therapy	All-cause mortality	2/2	0.68 *(0.38, 1.22)*	0.80 *(0.04, 14.5)*	Europe (0.918)
93 [P105]	Early IIb/IIIa inhibitors for primary PCI	Late IIb/IIIa inhibitors	All-cause mortality	6/2	0.64 *(0.33, 1.23)*	1.19 *(0.38, 3.67)*	Europe (0.357)
320 [P99]	PCI-based invasive strategy for stable CAD	Usual medical treatment	Cardiovascular mortality	5/2	0.51 *(0.26, 1.00)*	0.92 *(0.53, 1.59)*	Europe (0.187)
388 [P101]	Prophylactic steroids for cardiopulmonary bypass	Placebo or standard care	In-hospital mortality (all cause)	4/6	0.67 *(0.22, 2.09)*	0.65 *(0.23, 1.86)*	N. Amer. (0.967)
390 [P1]	Early invasive strategy after fibrinolytic therapy for MI	Ischemia-guided management	All-cause mortality	3/2	0.58 *(0.37, 0.94)*	0.60 *(0.16, 2.22)*	Europe (0.976)
51 [C11]	Glycoprotein IIb/IIa inhibitors for PCI	Placebo	Death within 30 days (all cause)	7/8	0.60 *(0.32, 1.11)*	0.72 *(0.50, 1.06)*	Europe (0.598)
21 [P60]	Antichlamydial antibiotic therapy for CAD	Placebo	All-cause mortality	4/2	1.11 *(0.70, 1.76)*	1.08 *(0.86, 1.35)*	N. Amer. (0.898)
71 [P64]	Secondary prevention programmes for CAD	Usual care	All-cause mortality	17/10	0.81 *(0.70, 0.94)*	1.08 *(0.83, 1.40)*	Europe (0.066)
148 [C48]	Hormone replacement therapy for preventing CVD	Placebo	Cardiovascular mortality	2/4	0.70 *(0.41, 1.20)*	1.11 *(0.88, 1.39)*	Europe (0.130)
177 [P78]	Multidisciplinary interventions for HF	Usual care	All-cause mortality	13/10	0.80 *(0.61, 1.06)*	0.82 *(0.66, 1.01)*	Europe (0.922)
309 [P138]	Disease management programmes for HF	Usual care	All-cause mortality	12/13	0.85 *(0.65, 1.11)*	0.84 *(0.67, 1.05)*	N. Amer. (0.967)
354 [C13]	Case management intervention for HF	Usual care	All-cause mortality	4/5	1.12 *(0.85, 1.47)*	0.75 *(0.53, 1.05)*	N. Amer. (0.074)
5 [P148]	CRT for LSVD	No CRT	All-cause mortality	5/3	0.69 *(0.54, 0.87)*	0.82 *(0.66, 1.03)*	Europe (0.287)
82 [P162]	Systematic early PCI after fibrinolytic therapy	Ischemia-guided PCI	All-cause mortality	2/2	0.52 *(0.27, 1.00)*	0.92 *(0.24, 3.61)*	Europe (0.460)

109 [C43]	Thrombolytic therapy for PE	Heparin or placebo plus heparin	All-cause mortality	2/3	1.57 *(0.45, 5.45)*	0.91 *(0.16, 5.32)*	N. Amer. (0.620)
155 [P55]	Disease management programmes for CHF	Standard care	All-cause mortality	12/15	0.79 *(0.59, 1.06)*	0.86 *(0.71, 1.04)*	Europe (0.653)
310 [P67]	Drug eluting stents	Bare metal stents	All-cause mortality	3/5	0.47 *(0.06, 3.64)*	0.98 *(0.52, 1.84)*	Europe (0.506)
328 [P108]	Off pump CABG surgery	On pump CABG surgery	All-cause mortality	9/3	0.92 *(0.44, 1.93)*	1.21 *(0.41, 3.58)*	Europe (0.677)
75 [P139]	Secondary prevention programmes for CHD	Usual care	All-cause mortality	9/7	0.81 *(0.70, 0.94)*	0.99 *(0.71, 1.38)*	Europe (0.278)
76 [P34]	Remote monitoring programmes for chronic HF	Usual care	All-cause mortality	3/9	0.63 *(0.43, 0.92)*	0.78 *(0.60, 1.00)*	Europe (0.363)
211 [P31]	Sirolimus-eluting stents	Bare metal stents	All-Cause mortality	5/2	1.00 *(0.69, 1.44)*	0.95 *(0.65, 1.39)*	N. Amer. (0.837)
226 [P125]	Glycoprotein IIb/IIIa inhibitors for PCI	Placebo or standard care	Death within 30 days (all cause)	4/7	0.44 *(0.21, 0.89)*	0.72 *(0.50, 1.04)*	Europe (0.221)
250 [P115]	CRT for LVSD	No CRT	All-cause mortality	5/4	0.72 *(0.57, 0.90)*	0.83 *(0.67, 1.03)*	Europe (0.349)
298 [P168]	Thrombolytic therapy for submassive PE	Placebo	Short-term death (all cause)	2/2	1.57 *(0.45, 5.45)*	0.70 *(0.08, 6.44)*	N. Amer. (0.534)
308 [P126]	Statin monotherapy for CVD in older adults	Placebo	All-cause mortality	5/6	0.89 *(0.84, 0.94)*	0.97 *(0.43, 2.18)*	Europe (0.828)
385 [P160]	Statins for CHD	Placebo	CHD Mortality	12/5	0.81 *(0.74, 0.89)*	0.82 *(0.64, 1.05)*	Europe (0.900)
259 [P102]	Off pump CABG surgery	On pump CABG surgery	All-cause mortality	14/4	0.96 *(0.60, 1.53)*	1.55 *(0.72, 3.36)*	Europe (0.290)

Table 5.1 shows, in addition to study details and the favoured continent, the significance of any difference in how the intervention performed between the two continents. For example, in P160 (385) the intervention group received statins for Chronic heart disease (CHD), while the control

group received a placebo. The endpoint reported was CHD mortality and 12 European and 5 North America trials contributed to this endpoint. The estimate for treatment effect for Europe was a RR of 0.81 while for North America the RR was 0.82. The intervention did perform better in Europe but a 2-sided p-value of 0.900[40] showed that the difference between continents in effect was not significant at the 5% level.

Some meta-analyses that reported fatal endpoints as RR did not report which group was the intervention under investigation and which was the control group. Therefore, it was not possible in these cases to calculate which region the intervention favoured, as it was not clear what the intervention was. This was only the case for three meta-analyses as shown in Table 5.2.

[40] The p-values reported are those given by the Z-tests calculated for each meta-analysis for the between-continent difference.

Table 5.2: Meta-analyses with Fatal Endpoints for Europe vs. North America that do not Identify which Group is the Intervention under Investigation. Effectiveness measured by RR.

Article Ref. no	**Intervention 1**	**Intervention 2**	**End point**	**No. of trials (Europe/ N. Amer.)**	**RR estimate for Europe *(95% CI)***	**RR estimate for N. Amer. *(95% CI)***	**Best Region (p-value of between-continent difference)**
56 [P24]	PCI	CABG surgery	All-cause mortality	7/3	0.61 *(0.33,1.13)*	0.74 *(0.41, 1.34)*	Not Applic. (0.652)
252 [P54]	Routine invasive strategies for ACS	Selective invasive strategies	All-cause mortality	2/3	0.55 *(0.17, 1.83)*	1.18 *(0.93, 1.51)*	Not Applic. (0.221)
296 [P16]	Routine invasive strategy for ACS	Selective invasive strategy	All-cause mortality	4/3	0.74 *(0.47, 1.16)*	1.18 *(0.93, 1.51)*	Not Applic. (0.0728)

5.1.1.1.2: Fatal Endpoints Measured as Odds Ratios

Six of the 47 meta-analyses that reported fatal endpoints did so using OR. The details of these meta-analyses are summarised in Table 5.3.

Table 5.3: Meta-analyses with Fatal Endpoint Data. Treatment Effect measured by OR.

Article Ref. no	**Intervention**	**Control**	**End point**	**No. of trials (Europe/ N. Amer.)**	**OR estimate for Europe *(95% CI)***	**OR estimate for N. Amer. *(95% CI)***	**Best Region (p-value of between-continent difference)**
7 [P152]	ACE inhibitors for LSVD	Placebo	All-cause mortality	3/4	0.50 *(0.19, 1.33)*	0.87 *(0.73, 1.02)*	Europe (0.280)
286 [P153]	Amiodarone for Preventing Post-operative AF	Placebo or routine treatment	Death within 30 days (all cause)	3/6	1.79 *(0.34, 9.43)*	0.82 *(0.43, 1.59)*	N. Amer. (0.397)
65 [P93]	Immediate PCI after thrombolytic therapy	Delayed PCI	Death within 12 months (all cause)	3/2	0.55 *(0.32, 0.92)*	1.42 *(0.73, 2.76)*	Europe (0.027)
68 [P159]	Hormone replacement therapy for preventing CVD	Placebo	Cardiovascular mortality	2/3	1.41 *(0.26, 7.62)*	1.10 *(0.88, 1.38)*	N. Amer. (0.776)
274 [P49]	Drug-eluting stents for CAD	Bare metal stents	Cardiovascular death within 1 year	4/4	0.85 *(0.37, 1.97)*	0.94 *(0.50, 1.77)*	Europe (0.849)
28 [P154]	Azithromycin for secondary prevention of CAD	Placebo	All-cause mortality	2/2	0.90 *(0.22, 3.71)*	0.94 *(0.62, 1.42)*	Europe (0.952)

One meta-analysis using OR did not report which was the intervention group and which was the control group. This is shown in Table 5.4. As with fatal endpoints reported as RR, no calculation was able to be made about which region the intervention favoured for this meta-analysis.

Table 5.4: The Meta-analysis with Fatal Endpoints for Europe vs. North America that does not Identify which group is the Intervention and which the Control. Effectiveness measured by OR.

Article Ref. no	Intervention 1	Intervention 2	End point	No. of trials (Europe/ N. Amer.)	OR estimate for Europe *(95% CI)*	OR estimate for N. Amer. *(95% CI)*	Best Region (p-value of between-continent difference)
69 [P81]	Invasive management strategy for angina and MI	Non-invasive management strategy	All-cause mortality	3/3	0.74 *(0.40, 1.36)*	1.22 *(0.92, 1.61)*	Not Applic. (0.147)

5.1.1.2: Examination over all Meta-analyses for Fatal Endpoints

Table 5.5 shows the number of meta-analyses identifying their intervention and control groups that favoured each continent.

Table 5.5: The Number of Meta-analyses Favouring Each Continent for Fatal Data

Total Number of Meta-analyses (%)	Number of Meta-analyses Favouring Europe (%)	Number of Meta-analyses Favouring North America (%)
43 (100%)	28 (65%)	15 (35%)

As can be seen, 28 of the 43 meta-analyses showed the intervention to perform best in Europe and in 15 the intervention to perform best in North America. To find the significance of this, a binomial sign test was conducted under the null hypothesis that the effect of interventions, relative to controls, would be the same in both continents. This showed that the effect of interventions, relative to controls, was not significantly different between Europe and North America, (2-sided p-value = 0.066).

5.1.1.3: Global Estimates of Continental Differences for Fatal Endpoints

Global estimates of continental difference were also produced. These were the calculated weighted average difference between the treatment effect estimates of Europe and North America over all of the meta-analyses. These estimates were produced for RR and OR. From these global estimates, it was found that for RR (37 meta-analyses), the mean log difference over all of the meta-analyses was -0.0839 with 95% confidence intervals of -0.172 and 0.0046 (2-sided p-value = 0.06), while for OR (6 meta-analyses), the mean log difference over all of the meta-analyses was -0.378 with 95% confidence intervals of -0.852 and 0.0966 (2-sided p-value = 0.119).These results showed that there was insufficient evidence for a difference in overall treatment effect between Europe and North America over all of the meta-analyses that had identified their intervention and control groups, as neither of these global estimates was statistically significant at the 5% level.

5.1.1.4: Inter-Continental Differences by Intervention Type – Fatal Endpoints

Even though statistically significant differences were not found in treatment effect between Europe and North America, it was considered important to establish whether specific intervention types were prone to inter-continental differences in treatment effect. This was investigated by grouping the meta-analyses according to intervention type and then, within each group, observing the Z-statistic and 2-sided p-value of each meta-analysis. This was to see whether, over all meta-analyses in that intervention type, there was any clear pattern of statistically significant differences in treatment effect. This was done by using Tables 5.6 to 5.10.

Table 5.6 provides details of the meta-analyses that examined drug therapies. For each meta-analysis, information is provided on the intervention examined, the number of trials included from Europe and North America, the pooled treatment effect estimate for each continent, the between-continental difference, and the result of the Z-test.

Table 5.6: Meta-analyses Investigating Drugs with the Z-statistic for the Difference in Treatment Effect between Europe and North America for Fatal Endpoints.

Meta-analysis Ref. No.	**Intervention**	**Treatment Effect Measure**	**No. of Trials (Europe / N. Amer.)**	**Pooled Estimate Europe *(95% CI)***	**Pooled Estimate North America *(95% CI)***	**Difference in Treatment Effect *(95% CIs)***	**Z-Statistic for difference**	**2-sided p-value of difference**
298 [P168]	Thrombolytic Therapy	RR	2/2	0.89 *(0.84,0.94)*	0.97 *(0.43, 2.18)*	0.8062 *(-1.74, 3.35)*	0.622	0.534
103 [P133]	Aspirin plus Oral Anti-coagulation Therapy	RR	4/2	1.13 *(0.88, 1.45)*	0.61 *(0.15, 2.41)*	0.6236 *(-0.77, 2.02)*	0.874	0.382
93 [P105]	Glycoprotein IIb/IIIa	RR	6/2	0.64 *(0.33, 1.23)*	1.19 *(0.38, 3.67)*	-0.6128 *(-1.91, 0.69)*	-0.921	0.357
178 [C22]	Digitalis	RR	2/3	0.55 *(0.12/2.53)*	0.99 *(0.93/1.06)*	-0.5950 *(-2.13, 0.94)*	-0.761	0.447
313 [P75]	Warfarin plus Aspirin	RR	3/2	0.93 *(0.71, 1.22)*	1.63 *(0.33, 8.02)*	-0.5587 *(-2.17, 1.06)*	-0.678	0.498
109 [C43]	Thrombolytic Therapy	RR	2/3	1.57 *(0.45, 5.45)*	0.91 *(0.16, 5.32)*	0.5465 *(-1.61, 2.71)*	0.496	0.620
226 [P125]	Glycoprotein IIb/IIIa	RR	4/7	0.44 *(0.21, 0.89)*	0.72 *(0.50, 1.04)*	-0.5017 *(-1.30, 0.30)*	-1.225	0.221
18	Phosphodiesterase III Inhibitors	RR	2/5	1.55	1.03	0.4014	1.407	0.160

[C12]				*(1.01,2.37)*	*(0.72,1.49)*	*(-0.16, 0.96)*		
227 [C4]	Anti-arrhythmics	RR	6/4	1.00 *(0.78,1.29)*	1.43 *(0.76, 2.70)*	0.3591 *(-1.04, 0.32)*	-1.035	0.301
223 [P96]	Beta-blockers	RR	2/3	0.93 *(0.35,2.47)*	0.72 *(0.45,1.17)*	0.2578 *(-0.83. 1.34)*	0.466	0.641
51 [C11]	Glycoprotein IIb/IIIa	RR	7/8	0.60 *(0.32,1.11)*	0.72 *(0.50,1.06)*	-0.1969 *(-0.93, 0.53)*	-0.528	0.598
308 [P126]	Statins	RR	5/6	0.89 *(0.84,0.94)*	0.97 *(0.43,2.18)*	-0.0899 *(-0.90, 0.72)*	-0.218	0.828
21 [P60]	Antibiotic Therapy	RR	4/6	1.11 *(0.70, 1.76)*	1.08 *(0.86,1.35)*	0.0333 *(-0.48, 0.54)*	0.128	0.898
388 [P101]	Prophylactic Steroids	RR	4/6	0.67 *(0.22, 2.09)*	0.65 *(0.23,1.86)*	0.0324 *(-1.51, 1.57)*	0.041	0.967
385 [P160]	Statins	RR	12/5	0.81 *(0.74, 0.89)*	0.82 *(0.64, 1.05)*	-0.0169 *(-0.28, 0.25)*	-0.126	0.900
286 [P153]	Amiodarone	OR	3/6	1.79*(0.34, 9.43)*	0.82*(0.43, 1.59)*	0.7727*(-1.02, 2.56)*	0.847	0.397
7[P152]	ACE Inhibitors	OR	¾	0.50*(0.19, 1.33)*	0.87*(0.73, 1.02)*	-0.5424*(-1.53, 0.44)*	-1.079	0.280
28[P154]	Azithromycin	OR	2/2	0.90*(0.22, 3.71)*	0.94*(0.62, 1.42)*	-0.0451*(-1.53, 1.44)*	-0.060	0.952

As can be seen, none of the meta-analyses investigating drugs showed a statistically significant difference in treatment effect between Europe and North America. Therefore, no clear pattern of statistically significant differences could be identified to suggest that this type of intervention is prone to inter-continental differences in treatment effect.

Table 5.7 provides the above details for the meta-analyses that investigated medical devices.

Table 5.7: Meta-analyses Investigating Medical Devices with the Z-statistic for the Difference in Treatment Effect between Europe and North America for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ N. Amer.)	Pooled Estimate Europe *(95% CIs)*	Pooled Estimate North America *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic for difference	2-sided p-value of difference
310 [P67]	Stent	RR	3/5	0.47 *(0.06, 3.64)*	0.98 *(0.52, 1.84)*	-0.7255 *(-2.87, 1.41)*	-0.665	0.506
5 [P148]	Cardiac Resynchronisation Therapy	RR	5/3	0.69 *(0.54,0.87)*	0.82 *(0.66, 1.03)*	-0.1754 *(-0.50, 0.15)*	-1.065	0.287
250 [P115]	Cardiac Resynchronisation Therapy	RR	5/4	0.72 *(0.57,0.90)*	0.83 *(0.67,1.03)*	-0.1477 *(-0.48, 0.17)*	-0.936	0.349
131 [P114]	Implantable Cardioverter Defibrillators	RR	2/3	0.82 *(0.62, 1.08)*	0.75 *(0.63,0.89)*	0.0896 *(-0.23, 0.41)*	0.548	0.584
211 [P31]	Stents	RR	5/2	1.00 *(0.69,1.44)*	0.95 *(0.65,1.39)*	0.0557 *(-0.48, 0.59)*	0.206	0.837
274 [P49]	Stents	OR	4/4	0.85*(0.37, 1.97)*	0.94*(0.50, 1.77)*	-0.1023*(-1.15, 0.95)*	-0.191	0.849

As Table 5.7 shows, none of the medical devices investigated showed a statistically significant difference in treatment effect between Europe and North America. Therefore, for medical devices, no clear pattern could be identified to show that this intervention type was prone to inter-continental differences.

Table 5.8 shows the details of the meta-analyses that investigated surgical interventions. As before, it states the intervention investigated, the number of European and North American trials

included, the treatment effect estimate for each continent, and the Z- statistic for the between-continental difference with its 2-sided p-value.

Table 5.8: Meta-analyses Investigating Surgery with the Z-statistic for the Difference in Treatment Effect between Europe and North America for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ N. Amer.)	Pooled Estimate Europe *(95% CI)*	Pooled Estimate North America *(95% CI)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic for difference	2-sided p-value of difference
216 [P62]	Primary PCI	RR	8/2	1.07 *(0.57, 1.99)*	0.64 *(0.08, 5.10)*	0.5102 *(-1.65, 2.67)*	0.462	0.644
259 [P102]	Off-pump CABG	RR	14/4	0.96 *(0.60, 1.53)*	1.55 *(0.72, 3.36)*	-0.4869 *(-1.39, 0.42)*	-1.058	0.290
328 [P108]	Off-pump CABG	RR	9/3	0.92 *(0.44, 1.93)*	1.21 *(0.41, 3.58)*	0.2789 *(-1.59, 1.03)*	-0.416	0.677
389 [P41]	Rescue PCI	RR	2/2	0.68 *(0.38,1.22)*	0.80 *(0.04,14.5)*	-0.1555 *(-3.11, 2.80)*	-0.103	0.918

As can be seen, the difference in treatment effect between Europe and North America was statistically significant in none of the meta-analyses at the 5% level. Consequently, no clear pattern was found to suggest that surgical interventions were prone to between-continental differences in treatment effect.

Table 5.9 shows the details of the meta-analyses that investigated lifestyle interventions.

Table 5.9: Meta-analyses Investigating Lifestyle Interventions with the Z-statistic for the Difference in Treatment Effect between Europe and North America for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ N. Amer.)	Pooled Estimate Europe *(95% CI)*	Pooled Estimate North America *(95% CI)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic for difference	2-sided p-value of difference
179 [C23]	Omega 3 Fatty Acids	RR	8/2	1.00 *(0.73,1.38)*	0.25 *(0.03, 2.24)*	1.3885 *(-0.83, 3.60)*	1.229	0.219
148 [C48]	Hormone Replacement Therapy	RR	2/4	0.70 *(0.41, 1.20)*	1.11 *(0.88, 1.39)*	-0.4514 *(-1.04, 0.13)*	-1.516	0.130
304 [C25]	Psychological Interventions	RR	2/6	0.70 (0.32,1.52)	0.93 (0.67,1.28)	0.2854 *(-1.12, 0.55)*	-0.666	0.505
75 [P64]	Secondary Prevention Programmes	RR	17/10	0.81 *(0.70,0.94)*	1.08 *(0.83, 1.40)*	0.2829 *(-0.58, 0.02)*	-1.839	0.066
74 [P139]	Secondary Prevention Programmes	RR	9/7	0.81 *(0.70,0.94)*	0.99 *(0.71, 1.38)*	-0.2007 *(-0.56, 0.16)*	-1.086	0.278
68 [P159]	Hormone Replacement Therapy	OR	2/3	1.41*(0.26, 7.62)*	1.10*(0.88, 1.38)*	0.2468*(-1.46, 1.95)*	0.284	0.776

Again, all meta-analyses found the difference in treatment effect between Europe and North America to be statistically non-significant. As such, no clear pattern of statistically significant differences was found to suggest that lifestyle interventions were prone to between-continental differences.

The next table provides the details of the meta-analyses that investigated management of cardiovascular disease.

Table 5.10: Meta-analyses Investigating Management with the Z-statistic for the Difference in Treatment Effect between Europe and North America for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ N. Amer.)	Pooled Estimate Europe *(95% CI)*	Pooled Estimate North America *(95% CI)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic for difference	2-sided p-value of difference
354 [C13]	Case Management	RR	4/5	1.12 *(0.85,1.47)*	0.75 *(0.53,1.05)*	0.4012 *(0.04, 0.84)*	1.789	0.074
76 [P34]	Telemonitoring	RR	3/9	0.63 *(0.43,0.92)*	0.78 *(0.60,1.00)*	-0.2139 *(-0.67, 0.25)*	-0.910	0.363
155 [P55]	Disease Management Programmes	RR	12/15	0.79 *(0.59, 1.06)*	0.86 *(0.71,1.04)*	-0.0804 *(-0.43, 0.27)*	-0.450	0.653
390 [P1]	Early Invasive Strategy	RR	3/2	0.58 *(0.37, 0.94)*	0.60 *(0.16, 2.22)*	-0.0211 *(-1.41, 1.37)*	-0.030	0.976
177 [P78]	Multi-disciplinary Interventions	RR	13/10	0.80 *(0.61, 1.06)*	0.82 *(0.66, 1.01)*	-0.0175 *(-0.37, 0.33)*	-0.098	0.922
309 [P138]	Disease Management Programmes	RR	12/13	0.85 *(0.65, 1.11)*	0.84 *(0.67, 1.05)*	0.0073 *(-0.34, 0.36)*	0.041	0.967
82 [P162]	Early Invasive PCI	RR	2/2	0.52 *(0.27, 1.00)*	0.92 *(0.24, 3.61)*	-0.5698 *(-2.08, 0.94)*	-0.739	0.460
320 [P99]	PCI-based Invasive Strategy	RR	5/2	0.51 *(0.26, 1.00)*	0.92 *(0.53, 1.59)*	-0.5830 *(-1.45, 0.28)*	-1.318	0.187
65[P93]	Immediate PCI	OR	3/2	0.55*(0.32, 0.92)*	1.42*(0.73, 2.76)*	-0.9561*(-1.80,-0.11)*	-2.213	0.027

As can be seen, one meta-analysis investigating management found the difference in treatment effect between Europe and North America to be statistically significant. This was P93 (65), a

meta-analysis comparing immediate PCI after thrombolytic therapy (Intervention) with delayed PCI (Control). The immediate PCI performed better, with respect to all-cause death within 12 months, in Europe compared with North America. The 2-sided p-value was 0.027. However, as no other meta-analyses in this group found a statistically significant difference between Europe and North America, no clear pattern could be identified to suggest that management interventions were prone to inter-continental differences in treatment effect.

Since no group of interventions showed a clear pattern of statistically significant differences in treatment effect between Europe and North America, the types of intervention prone to such differences for fatal endpoints could not be identified. Furthermore, given that there were 47 meta-analyses overall that provided fatal endpoint data, finding a between-continent difference statistically significant at the 5% level in only one of these is not more than would be expected by random chance. As such, there was insufficient evidence to conclude which types of intervention were prone to differences in treatment effect between Europe and North America for fatal endpoints.

5.1.1.5: The Impact of Including Overlapping Meta-analyses on Fatal Endpoint Results

Differences in treatment effect may have been diminished or enhanced by the inclusion of duplicated trial results from overlapping meta-analyses. As mentioned in Chapter 4 (Section 4.4.1) some of the systematic reviews on which this comparison was based were investigating the same clinical question and, as such, their meta-analyses may have contained some of the same trials. This may create a problem if an excessively large number of such trials favoured Europe

over North America, as when meta-analyses comprising these trials were pooled there would be unreasonably strong support for interventions favouring Europe compared to North America.

Of the 43 included meta-analyses, 26 overlapping meta-analyses were identified (C11,C13,C43, C48,P125, P34, P55, P78, P138, P168, P159, P31, P49, P67, P60, P154, P64, P139, P93, P162, P102, P108, P115, P148, P126, P160) and 17 non-overlapping meta-analyses (C4, C12, C22, C23, C25, P1, P41, P62, P75, P96, P99, P101, P105, P114, P133, P152, P153). Of the 26 overlapping meta-analyses, there were nine pairs of meta-analyses that had one or more trials in common with each other (C11 and P125, C43 and P168, C48 and P159, P60 and P154, P64 and P139, P93 and P162, P102 and P108, P115 and P148 and P126 and P160). Furthermore, there was one group of five meta-analyses which had one or more trials in common (C13, P34, P55, P78, P138) and one group of three meta-analyses where this was the case (P31, P49 and P67).

More importantly, it was found that of the meta-analyses considered to be overlapping, 73% (19 out of 26 meta-analyses) favoured Europe over North America, while for non-overlapping meta-analyses only 53% (9 out of 17 meta-analyses) favoured Europe. This suggests that, with regard to fatal endpoints, the large number of meta-analyses that had favoured Europe over North America (65% or 28 out of 43 meta-analyses) may have been inflated as a consequence of including overlapping meta-analyses.

5.1.2: Non-Fatal Endpoints.

Of the 59 meta-analyses relevant for the Europe versus North America comparison, only 44 had at least two trials from each continent with non-fatal data that would allow for separate estimates

of treatment effect to be calculated for each continent. Tables 5.11 to 5.15 provide the details of these meta-analyses.

5.1.2.1: Meta-analyses Grouped by Treatment Effect Measure for Non-Fatal Endpoints

Table 5.8 shows the details for the included meta-analyses that measured treatment effect by the RR for non-fatal endpoints. It includes the details about the intervention group, control group, and endpoint reported for each meta-analysis as well as the number of trials based in each continent. The treatment effect estimates are also given for each continent along with the details of the continent where the intervention performed best (Best region and 2-sided p-value).

5.1.2.1.1: Non-Fatal Endpoints Measured as Risk Ratios.

Thirty-five meta-analyses reported treatment effect in terms of RR. The details of these are summarised in Table 5.11.

Table 5.11: Meta-analyses with Non-Fatal Endpoint Data. Treatment Effect Measured by RR.

Article Ref. no	Intervention	Control	End point	No. of trials (Europe / N. Amer.)	RR estimate for Europe *(95% CI)*	RR estimate for N. Amer. *(95% CI)*	Best Region (p-value of between-continental difference)
178 [C22]	Digitalis for CHF	Placebo	Clinical deterioration	2/2	0.19 *(0.02, 1.62)*	0.34 *(0.11, 1.07)*	Europe (0.644)
304 [C25]	Psychological interventions for CHD	Usual care or no interventio n	MI	4/7	0.74 *(0.54, 1.02)*	0.76 *(0.57, 1.02)*	Europe (0.872)
18 [C12]	Phosphodiesterase III inhibitors for HF	Placebo	Worsening heart failure	3/6	1.52 *(0.77, 3.01)*	0.85 *(0.63, 1.15)*	N. Amer. (0.130)
21 [P60]	Antichlamydial antibiotic therapy for CAD	Placebo	MI	2/2	0.61 *(0.26, 1.47)*	1.03 *(0.82, 1.29)*	Europe (0.267)
75 [P64]	Secondary prevention programmes for CAD	Usual care	Recurrent MI	13/5	0.727 *(0.60, 0.89)*	0.733 *(0.49, 1.10)*	Europe (0.971)
48 [P166]	Vitamin-mineral supplements for CAD	Placebo	Restenosis	2/2	0.84 *(0.34, 2.07)*	0.97 *(0.65, 1.44)*	Europe (0.785)
109 [C43]	Thrombolytic therapy for PE	Heparin or placebo plus heparin	Major Haemorrhagic events	2/2	0.61 *(0.13, 2.96)*	2.64 *(0.44, 15.8)*	Europe (0.230)
203 [P120]	Self-management strategy for HF	Usual manageme nt strategy	Hospital re-admission within 1 year	2/3	0.73 *(0.56, 0.95)*	0.74 *(0.58, 0.93)*	Europe (0.954)
216 [P62]	Facilitated PCI for ST-elevation MI	Primary PCI	Short-term major bleeding	8/3	1.30 *(0.83, 2.05)*	2.60 *(0.58, 11.7)*	Europe (0.388)
103 [P133]	Combined aspirin –oral anticoagulation therapy	Oral anticoagul ation therapy alone	Major bleeding	4/2	1.69 *(0.94, 3.04)*	1.28 *(0.74, 2.23)*	N. Amer. (0.502)
175 [P29]	Intravenous magnesium for acute onset AF	Placebo or other anti-arrhythmic agent	Conversion of AF to sinus rhythm	2/3	3.55 *(1.21, 10.4)*	1.07 *(0.59, 1.95)*	Europe (0.056)
227 [C4]	Anti-arrhythmics after cardioversion of AF	Placebo, drugs for rate control or no	AF recurrence	11/6	0.70 *(0.61, 0.81)*	0.80 *(0.67, 0.97)*	Europe (0.259)

		treatment					
385 [P160]	Statins for CHD	Placebo	MI	13/4	0.67 *(0.58, 0.77)*	0.78 *(0.66, 0.93)*	Europe (0.176)
93 [P105]	Early IIb/IIIa inhibitors for primary PCI	Late IIb/IIIa inhibitors	Post-procedural TIMI 3 flow	6/2	1.05 *(0.99, 1.11)*	1.01 *(0.91, 1.13)*	Europe (0.594)
30 [C54]	Psychosocial smoking cessation help for CHD	Usual care	Abstinence from smoking	6/7	1.22 *(0.95, 1.57)*	1.32 *(1.13, 1.53)*	N. Amer. (0.604)
320 [P99]	PCI based invasive strategy for stable CAD	Usual medical treatment	MI	7/5	0.78 *(0.46, 1.31)*	1.11 *(0.91, 1.37)*	Europe (0.209)
388 [P101]	Prophylactic steroids for cardiopulmonary bypass	Placebo or standard care	New onset AF	3/7	0.84 *(0.57, 1.25)*	0.60 *(0.46, 0.80)*	N. Amer. (0.172)
390 [P1]	Early invasive strategy after fibrinolytic therapy for MI	Ischemia-guided management	Major bleeding	3/2	1.56 *(0.57, 4.26)*	1.24 *(0.48, 3.24)*	N. Amer. (0.744)
51 [C11]	Glycoprotein IIb/IIIa inhibitors for PCI	Placebo	30 day urgent revascularisation	7/7	0.60 *(0.38, 0.96)*	0.48 *(0.32, 0.71)*	N. Amer. (0.453)
1 [P123]	Amiodarone prophylaxis for cardiac surgery	Placebo	Atrial fibrillation or flutter	2/6	0.74 *(0.39, 1.41)*	0.66 *(0.56, 0.78)*	N. Amer. (0.721)
34 [P65]	Sirolimus-eluting stents for PCI	Bare metal stents	Stent-associated thrombosis	2/2	0.25 *(0.07, 0.90)*	0.60 *(0.14, 2.53)*	Europe (0.376)
148 [C48]	Hormone replacement therapy for preventing CVD	Placebo	Stroke	4/2	1.71 *(0.66, 4.44)*	1.09 *(0.88, 1.34)*	N. Amer. (0.363)
153 [P149]	Amiodarone prophylaxis for cardiothoracic surgery	Placebo	Post-operative AF	6/5	0.72 *(0.56, 0.93)*	0.65 *(0.54, 0.78)*	N. Amer. (0.506)
177 [P78]	Multidisciplinary interventions for HF	Usual care	Hospital admission (all cause)	7/10	0.81 *(0.63, 1.04)*	0.88 *(0.77, 1.00)*	Europe (0.579)
309 [P138]	Disease management programmes for HF	Usual care	All cause rehospitalisation	10/12	0.79 *(0.56, 1.11)*	0.91 *(0.72, 1.16)*	Europe (0.487)
313 [P75]	Warfarin plus aspirin after MI or ACS	Aspirin only	MI	3/3	0.55 *(0.42, 0.72)*	0.65 *(0.35, 1.19)*	Europe (0.646)

354 [C13]	Case management intervention for HF	Usual care	Hospital readmission due to HF	3/4	0.50 *(0.26, 0.97)*	0.68 *(0.50, 0.91)*	Europe (0.415)
22 [P73]	Aspirin plus Warfarin after ACS	Aspirin alone	Major bleeds	4/6	2.36 *(1.53, 3.65)*	1.53 *(1.21, 1.92)*	N. Amer. (0.081)
145 [P135]	CRT for LSVD	No CRT	Hospital admission due to HF	3/2	0.52 *(0.41, 0.67)*	0.67 *(0.44, 1.05)*	Europe (0.316)
155 [P55]	Disease management programmes for CHF	Standard care	Hospital readmission (all cause)	12/16	0.75 *(0.64, 0.90)*	0.86 *(0.77, 0.97)*	Europe (0.196)
310 [P67]	Drug eluting stents	Bare metal stents	Angiographic binary restenosis	3/5	0.12 *(0.05, 0.28)*	0.37 *(0.23, 0.61)*	Europe (0.026)
328 [P108]	Off pump CABG surgery	On pump CABG surgery	Stroke	9/3	0.43 *(0.18, 1.03)*	1.22 *(0.29, 5.22)*	Europe (0.229)
226 [P125]	Glycoprotein IIb/IIIa inhibitors for PCI	Placebo or standard care	MI within 30 days	4/7	0.87 *(0.59, 1.30)*	0.62 *(0.50, 0.77)*	N. Amer. (0.132)
250 [P115]	CRT for LSVD	No CRT	Hospital admission due to HF	2/3	0.53 *(0.41, 0.68)*	0.82 *(0.55, 1.20)*	Europe (0.069)
259 [P102]	Off pump CABG surgery	On pump CABG surgery	MI	14/4	0.75 *(0.45, 1.27)*	1.23 *(0.40, 3.77)*	Europe (0.437)

As for fatal endpoints, the best region could not be identified for some of the meta-analyses using RR as these meta-analyses had not specifically identified which was the intervention group and which was the control group. As such, it was not possible to calculate the continent in which the intervention performed best. This was only the case in three meta-analyses and details of these can be found in Table 5.12.

Table 5.12: Meta-analyses with Non-Fatal Endpoints for Europe vs. North America that do not Identify which Group Was the Intervention and Which the Control Effectiveness measured by RR.

Article Ref. no	**Intervention 1**	**Intervention 2**	**End point**	**No. of trials (Europe/ N. Amer.)**	**RR estimate for Europe *(95% CI)***	**RR estimate for N. Amer. *(95% CI)***	**Best Region (p-value of between-continent difference)**
56 [P24]	PCI	CABG surgery	Stroke	4/3	0.36 *(0.09, 1.44)*	0.40 *(0.14, 1.13)*	Not Applic. (0.913)
253 [P54]	Routine invasive strategies for ACS	Selective invasive strategies	Severe angina	2/3	0.77 *(0.66, 0.89)*	0.93 *(0.69, 1.24)*	Not Applic. (0.254)
296 [P16]	Routine invasive strategy for ACS	Selective invasive strategy	MI	4/3	0.90 *(0.49, 1.66)*	0.87 *(0.70, 1.08)*	Not Applic. (0.918)

5.1.2.1.2: Non-Fatal Endpoints Measured as Odds Ratios

Only three meta-analyses reported treatment effect in terms of OR. Table 5.13 provides the details (the same as the above Table) for the included meta-analyses that reported treatment effect in terms of OR.

Table 5.13: Meta-analyses with Non-Fatal End-Point Data. Treatment Effect Measured by OR.

Article Ref. no	Intervention	Control	End point	No. of trials (Europe/ N. Amer.)	OR estimate for Europe *(95% CI)*	OR estimate for N. Amer. *(95% CI)*	Best Region (p-value of between-continent difference)
68 [P159]	Hormone replacement therapy for preventing CVD	Placebo	MI	2/3	0.70 *(0.22, 2.19)*	1.07 *(0.84, 1.36)*	Europe (0.472)
27 [P156]	Amiodarone prophylaxis for cardiac surgery	Placebo or routine treatment	Post-operative AF	5/7	0.47 *(0.32, 0.69)*	0.51 *(0.41, 0.64)*	Europe (0.697)
286 [P153]	Amiodarone for preventing post-operative AF	Placebo	Bradycardia	4/2	1.24 *(0.26, 5.85)*	1.91 *(1.05, 3.48)*	Europe (0.611)

One meta-analysis reporting treatment effect in terms of OR did not report which was the intervention group and which was the control group. This is shown in Table 5.14. As with non-fatal endpoints reported using RR, no calculation was able to be made about which continent the intervention favoured for this meta-analysis.

Table 5.14: The Meta-analysis with Non-Fatal End-Points for Europe vs. North America That Does Not Identify the Intervention under Investigation. Treatment Effect Measured by the OR.

Article Ref. no	Intervention 1	Intervention 2	End point	No. of trials (Europe/ N. Amer.)	OR estimate for Europe *(95% CI)*	OR estimate for N. Amer. *(95% CI)*	Best Region (p-value of between-continent difference)
69 [P81]	Invasive management strategy for angina and MI	Non-invasive management strategy	Hospital readmission	2/2	0.45 *(0.39, 0.53)*	0.75 *(0.60, 0.93)*	Not Applic. (0.0003)

5.1.2.1.3: Non-Fatal Endpoints Measured as Mean Difference

The next table provides the details of the meta-analyses that contained non-fatal data but reported this using MD. Two of the meta-analyses reported non-fatal endpoints as MD and Table 5.15 provides the summary information for these.

Table 5.15: Meta-analyses with Non-Fatal End-Points Data. Treatment Effect Measured by MD.

Article Ref. no	Intervention	Control	End point	No. of trials (Europe/ N. Amer.)	MD estimate for Europe *(95% CI)*	MD estimate for N. Amer. *(95% CI)*	Best Region (p-value of between-continent difference)
217 [C35]	Low glycaemic index diets for CHD	Other diets	Total cholesterol	7/3	-0.11 *(-0.34, 0.12)*	-0.06 *(-0.35, 0.23)*	Europe (0.790)
272 [P165]	Drug, biobehavioural and exercise therapies for CAD	Placebo or no intervention	Heart rate variability	13/5	0.43 *(0.25, 0.62)*	0.46 *(0.13, 0.79)*	N. Amer. (0.893)

5.1.2.2: Examination over All Meta-analyses for Non-Fatal Endpoints

Table 5.16 shows the number of meta-analyses specifically identifying their intervention and control groups that favoured each continent.

Table 5.16: The Number of Meta-analyses Favouring each Continent for Non-Fatal Data.

Total Number of Meta-analyses (%)	Number of Meta-analyses Favouring Europe (%)	Number of Meta-analyses Favouring North America (%)
40 (100%)	28 (70%)	12 (30%)

As can be seen, of the 40 meta-analyses, 28 showed the intervention performed better in Europe compared to North America, while in 12 the intervention performed better in North America. As for fatal endpoints, a binomial sign test was conducted under the null hypothesis that the effect of interventions, compared to controls, would be the same in both continents. This indicated that there was a significant difference in the effect of interventions, relative to controls, between Europe and North America (2-sided p-value = 0.017).

5.1.2.3: Global Estimates of Inter-Continental Difference for Non-Fatal Endpoints

As for the fatal-endpoint comparison, global estimates of continental difference were produced that showed the average weighted difference in treatment effect between Europe and North America. These estimates showed that the mean log difference for RR (35 meta-analyses) was -0.036 with 95% confidence intervals of -0.101 to 0.029 (2-sided p-value = 0.281) and for OR (3 meta-analyses) the mean log difference over all of the meta-analyses was -0.152 with 95% confidence intervals of -0.562 and 0.257 (2-sided p-value = 0.466).These results showed that there was insufficient evidence for any difference in the effect of cardiovascular interventions

between Europe and North America over all of the meta-analyses that had identified their intervention and control groups, as neither of the global estimates of continental difference were statistically significant at the 5% level.

5.1.2.4: Inter-Continental Differences by Intervention Type – Non-Fatal Endpoints

As there was some evidence for inter-continental differences in treatment effect for non-fatal endpoints, it was important to know if there were particular types of intervention that were likely to exhibit such differences. The same process was repeated for these non-fatal endpoints as described in the previous section on fatal endpoints.

Table 5.17 shows the meta-analysis that investigated drug therapies and included trials that had non-fatal endpoint data for Europe and North America.

Table 5.17: Meta-analyses Investigating Drugs with their Z-statistic for the Difference in Treatment Effect between Europe and North America for Non-Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe / N. Amer.)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for North America *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-statistic	2-sided p-value of difference
109 [C43]	Thrombolytic therapy	RR	2/2	0.61 *(0.13, 2.96)*	2.64 *(0.44, 15.8)*	-1.4628 *(-3.85,0.93)*	-1.201	0.230
175 [P29]	Magnesium	RR	2/3	3.55 *(1.21, 10.4)*	1.07 *(0.59, 1.95)*	1.2024 *(-0.03,2.43)*	1.912	0.056
18 [C12]	Phosphodiesterase III Inhibitors	RR	3/6	1.52 *(0.77, 3.01)*	0.85 *(0.63, 1.15)*	0.5761 *(-0.17, 1.32)*	1.514	0.130
178 [C22]	Digitalis	RR	2/2	0.19 *(0.02, 1.62)*	0.34 *(0.11, 1.07)*	-0.5678 *(-2.98, 1.84)*	-0.462	0.644
21 [P60]	Antibiotic Therapy	RR	2/2	0.61 *(0.26, 1.47)*	1.03 *(0.82, 1.29)*	-0.5119 *(-1.42, 0.39)*	-1.110	0.267
22 [P73]	Aspirin plus Warfarin	RR	4/6	2.36 *(1.53, 3.65)*	1.53 *(1.21, 1.92)*	0.4371 *(-0.05, 0.93)*	1.745	0.081
226 [P125]	Glycoprotein IIb/IIIa	RR	4/7	0.87 *(0.59, 1.30)*	0.62 *(0.50, 0.77)*	0.3476 *(-0.11, 0.80)*	1.505	0.132
388 [P101]	Prophylactic Steroids	RR	3/7	0.84 *(0.57, 1.25)*	0.60 *(0.46, 0.80)*	0.3363 *(-0.15, 0.82)*	1.364	0.172
103 [P133]	Aspirin plus Oral Anticoagulation Therapy	RR	4/2	1.69 *(0.94, 3.04)*	1.28 *(0.74, 2.23)*	0.2755 *(-0.53, 1.08)*	0.671	0.502
51 [C11]	Glycoprotein IIb/IIIa	RR	7/7	0.60 *(0.38, 0.96)*	0.48 *(0.32, 0.71)*	0.2348 *(-0.38, ,0.85)*	0.750	0.453
313 [P75]	Warfarin plus Aspirin	RR	3/3	0.55 *(0.42, 0.72)*	0.65 *(0.35, 1.19)*	-0.1560 *(-0.82,0.51)*	-0.459	0.646
385 [P160]	Statin	RR	13/4	0.67 *(0.58, 0.77)*	0.78 *(0.66, 0.93)*	-0.1555 *(-0.38, 0.07)*	-1.355	0.176
227 [C4]	Anti-arrhythmics	RR	11/6	0.70 *(0.61, 0.81)*	0.80 *(0.67, 0.97)*	-0.1321 *(-0.36, 0.10)*	-1.129	0.259
1 [P123]	Amiodarone	RR	2/6	0.74 *(0.39,1.41)*	0.66 *(0.56,0.78)*	0.1198 *(-0.54, 0.78)*	-0.778	0.721
153 [P149]	Amiodarone	RR	6/5	0.72 *(0.56,0.93)*	0.65 *(0.54,0.78)*	0.1052 *(-0.20, 0.42)*	0.665	0.506
93 [P105]	Glycoprotein IIb/IIIa	RR	6/2	1.05 *(0.99,1.11)*	1.01 *(0.91, 1.13)*	0.0329 *(-0.09, 0.15)*	0.533	0.594
286 [P153]	Amiodarone	OR	4/2	1.24 *(0.26, 5.85)*	1.91 *(1.05, 3.48)*	-0.4312 *(-2.09, 1.23)*	-0.508	0.697
27 [P156]	Amiodarone	OR	5/7	0.47 *(0.32,0.69)*	0.51 *(0.41, 0.64)*	-0.0900 *(-0.54, 0.36)*	-0.389	0.611

As can be seen, all meta-analyses found the difference in treatment effect between Europe and North America to be statistically non-significant. As such, no clear pattern of statistically significant differences was identified to suggest that drug therapies were prone to inter-continental differences in treatment effect.

Table 5.18 shows the meta-analyses that investigated medical devices.

Table 5.18: Meta-analyses Investigating Medical Devices with their Z-statistic for the Difference in Treatment Effect between Europe and North America for Non-Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ N. Amer.)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for North America *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-statistic	2-sided p-value of difference
310 [P67]	Stents	RR	3/5	0.12 *(0.05, 0.28)*	0.37 *(0.23, 0.61)*	-1.1017 *(-2.07, -0.13)*	-2.228	0.026
34 [P65]	Stents	RR	2/2	0.25 *(0.07, 0.90)*	0.60 *(0.14, 2.53)*	-0.8678 *(-2.79, 1.05)*	-0.885	0.376
250 [P115]	Cardiac Resynchronisation Therapy	RR	2/3	0.53 *(0.41, 0.68)*	0.82 *(0.55, 1.20)*	-0.4280 *(-0.89, 0.03)*	-1.815	0.069
145 [P135]	Cardiac Resynchronisation Therapy	RR	3/2	0.52 *(0.41, 0.67)*	0.67 *(0.44, 1.05)*	-0.2568 *(-0.76, 0.25)*	-1.002	0.316

Only one meta-analysis investigating medical devices found the difference in treatment effect between Europe and North America to be statistically significant at the 5% level. This was P67 (310), a meta-analysis comparing drug eluting stents (the intervention) to bare metal stents (the control). In this meta-analysis, the intervention performed better (with respect to binary angiographic restenosis) in Europe than in North America, and the 2-sided p-value for the

difference in treatment effect was 0.026. However, as no other meta-analyses found statistically significant differences in treatment effect, no clear pattern could be identified to suggest that medical devices were prone to treatment effect differences between Europe and North America.

Table 5.19 shows the meta-analyses that investigated surgical interventions and included trials that had non-fatal endpoint data for Europe and North America.

Table 5.19: Meta-analyses Investigating Surgical Interventions with their Z-statistic for the Difference in Treatment Effect between Europe and North America for Non-Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ N. Amer.)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for North America *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-statistic	2-sided p-value of difference
328 [P108]	Off-Pump CABG	RR	9/3	0.43 *(0.18, 1.03)*	1.22 *(0.29, 5.22)*	-1.0366 *(-2.72, 0.65)*	-1.204	0.229
216 [P62]	Primary PCI	RR	8/3	1.30 *(0.83, 2.05)*	2.60 *(0.58, 11.7)*	-0.6926 *(-2.26, 0.88)*	-0.864	0.388
259 [P102]	Off-Pump CABG	RR	14/4	0.75 *(0.45, 1.27)*	1.23 *(0.40, 3.77)*	-0.4898 *(-1.72, 0.74)*	-0.778	0.437

As can be seen, none of the meta-analyses investigating surgical interventions found the difference in treatment effect between Europe and North America to be statistically significant at the 5% level. Therefore, no clear pattern of statistically significant differences was identified to suggest that surgical interventions were prone to international differences.

Table 5.20 shows the details of the meta-analyses that investigated lifestyle interventions. As before, it shows the intervention investigated, the number of European and North American trials

included, the treatment effect estimate for each continent, and the Z- statistic for the between-continental difference with its 2-sided p-value.

Table 5.20: Meta-analyses Investigating Lifestyle Interventions with their Z-statistic for the Difference in Treatment Effect between Europe and North America for Non-Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ N. Amer.)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for North America *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-statistic	2-sided p-value of difference
148 [C48]	Hormone Replacement Therapy	RR	4/2	1.71 *(0.66, 4.44)*	1.09 *(0.88, 1.34)*	0.4530 *(-0.52, 1.43)*	0.910	0.363
48 [P166]	Vitamin-Mineral Supplementation	RR	2/2	0.84 *(0.34, 2.07)*	0.97 *(0.65, 1.44)*	-0.1367 *(-1.12, 0.85)*	-0.273	0.785
30 [C54]	Psychological Interventions	RR	6/7	1.22 *(0.95,1.57)*	1.32 *(1.13,1.53)*	-0.0777 *(-0.37, 0.22)*	-0.519	0.604
304 [C25]	Psychological Interventions	RR	4/7	0.74 *(0.54,1.02)*	0.76 *(0.57, 1.02)*	-0.0354 *(-0.47, 0.40)*	-0.161	0.872
75 [P64]	Secondary Prevention Programmes	RR	13/5	0.73 *(0.60,0.89)*	0.73 *(0.49,1.10)*	-0.0085 *(-0.46, 0.44)*	-0.037	0.971
68 [P159]	Hormone Replacement Therapy	OR	2/3	0.70 *(0.22, 2.19)*	1.07 *(0.84, 1.36)*	-0.4313 *(-1.61, 0.74)*	-0.720	0.472
217 [C35]	Low Glycaemic Index Diets	MD	7/3	-0.11 *(-0.34,0.12)*	-0.06 *(-0.35,0.23)*	-0.0500 *(-0.42, 0.32)*	-0.266	0.790
272 [P165]	Drug, Behavioural and Exercise Therapies	MD	13/5	0.43 *(0.25,0.62)*	0.46 *(0.13,0.79)*	-0.0255 *(-0.40, 0.35)*	-0.314	0.893

As can be seen from Table 5.20, none of the meta-analyses showed a statistically significant difference in treatment effect between Europe and North America. As such, no clear pattern of

statistically significant differences could be identified for lifestyle interventions to suggest that these types of interventions were prone to inter-continental differences in treatment effect.

Table 5.21 shows the details of the meta-analyses that investigated management strategies.

Table 5.21: Meta-analyses Investigating Management Strategies with their Z-statistic for the Difference in Treatment Effect between Europe and North America for Non-Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ N. Amer.)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for North America *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-statistic	2-sided p-value of difference
320 [P99]	PCI-based Invasive Strategy	RR	7/5	0.78 *(0.46, 1.31)*	1.11 *(0.91, 1.37)*	-0.3578 *(-0.92, 0.20)*	-1.257	0.209
354 [C13]	Case Management – Clinical Service Organisation	RR	3/4	0.50 *(0.26, 0.97)*	0.68 *(0.50, 0.91)*	-0.2985 *(-1.02, 0.42)*	-0.815	0.415
390 [P1]	Early Invasive Strategy	RR	3/2	1.56 *(0.57, 4.26)*	1.24 *(0.48, 3.24)*	0.2314 *(-1.16,1.62)*	0.327	0.744
309 [P138]	Disease Management Programmes	RR	10/12	0.79 *(0.56, 1.11)*	0.91 *(0.72,1.16)*	-0.1483 *(-0.57, 0.27)*	-0.695	0.487
155 [P55]	Disease Management Programmes	RR	12/16	0.75 *(0.64, 0.90)*	0.86 *(0.77, 0.97)*	-0.1357 *(-0.35, 0.07)*	-1.292	0.196
177 [P78]	Multi-disciplinary Interventions	RR	7/10	0.81 *(0.63,1.04)*	0.88 *(0.77,1.00)*	-0.0801 *(-0.36, 0.20)*	-0.555	0.579
203 [P120]	Self-Management Interventions	RR	2/3	0.73 *(0.56, 0.95)*	0.74 *(0.58,0.93)*	-0.0104 *(-0.36, 0.34)*	-0.058	0.954

As can be seen, it was again found that the differences in treatment effect between Europe and North America were statistically non-significant in all of the included meta-analyses. As a result,

no clear pattern of statistically significant differences was identified to suggest that management interventions were prone to inter-continental differences in treatment effect.

However, when further observing the 2-sided p-values of the between-continent difference in each meta-analysis for management interventions, it was discovered that one meta-analysis that did not identify which group was the intervention and which the control showed a significant difference between Europe and North America (2-sided p-value = 0.0003) (see Table 5.11). This was P81 (69), which simply compared an invasive management strategy for unstable angina and MI to a non-invasive management strategy.

As no group of interventions showing a clear pattern of statistically significant findings could be identified, the types of intervention prone to differences between Europe and North America could not be identified for non-fatal endpoints. Furthermore, as there were 44 meta-analyses overall that provided non-fatal data, finding a between-continent difference that was statistically significant at the 5% level in only two of the meta-analyses is not more than would be expected by random chance. Therefore, there was insufficient evidence to conclude which types of intervention were prone to differences in treatment effect between Europe and North America for non-fatal endpoints.

5.1.2.5: The Impact of Including Overlapping Meta-analyses on Non-Fatal Endpoint Results

Any differences in treatment effect found may have been the consequence of including duplicated trial results from overlapping meta-analyses. Therefore, as with the comparison using fatal endpoints, these results were investigated to see if they had been influenced by the inclusion of

overlapping meta-analyses. Of the 40 included meta-analyses, 20 overlapping meta-analyses were identified. These were C11 and P125, C48 and P159, P65 and P67, P73 and P75, P102 and P108, P115 and P135, C13, P55, P78, P138 and P123, P149, P153, P156. There were also 20 non-overlapping meta-analyses (C4, C12, C22, C25, C35, C43, C54, P1, P29, P60, P62, P64, P99, P101, P105, P120, P133, P160, P165, and P166).

Furthermore, it was found that the percentage of meta-analyses where the intervention was favoured more in Europe than North America was 70% for both overlapping and non-overlapping meta-analyses. This was because both had 14 out of 20 meta-analyses where the intervention favoured Europe over North America. This suggested that the overall percentage of meta-analyses in which the intervention favoured Europe over North America (70% or 28 out of 40 meta-analyses) had not been inappropriately exaggerated by the duplication of trial results from overlapping meta-analyses.

5.2: Conclusion

With regards to fatal endpoints, this chapter has shown that there is no evidence that inter-continental differences between Europe and North America for cardiovascular interventions exist. This is because no analysis was statistically significant at the 5 % level. This was not the case for non-fatal endpoints, where a statistically significant difference in the effect of interventions was found between Europe and North America. In particular, it was found that interventions perform better, relative to controls, in Europe than in North America. However, this was not supported by the global estimates of continental difference calculated for non-fatal endpoints, since no estimate was found to be statistically significant. It should also be noted that the types of intervention

prone to inter-continental differences in treatment effect could not be identified for either type of endpoint.

Whilst there is little evidence for treatment effectiveness differing between Europe and North America it is vital to investigate further and examine whether such differences exist between other continents. An investigation, such as this, would provide us with further knowledge of inter-continental differences in treatment effectiveness as well as helping with the extrapolation of clinical trial results between continents. This is the basis for the following chapter, which investigates inter-continental differences in treatment effectiveness by comparing Europe, North America, and Asia.

CHAPTER 6.

AN INTER-CONTINENTAL COMPARISON OF EUROPE, NORTH AMERICA AND ASIA.

To investigate the existence of international differences in treatment effectiveness further, two more pair-wise comparisons were conducted for Europe and Asia and North America and Asia, respectively, with regards to both fatal and non-fatal endpoints. A universal comparison of continents was also conducted by comparing each continent's individual trial results from the included meta-analyses.

This chapter presents the findings of these comparisons and is organised into two main sections The first section concerns the investigation of fatal endpoints and is split into four parts: the pair-wise comparison between Europe and Asia, the pair-wise comparison between North America and Asia, the universal comparison for this endpoint, and the investigation into the types of intervention prone to international differences in their effectiveness. The second section discusses the same comparisons as they pertain to non-fatal endpoints. The chapter concludes by summarising the evidence and briefly discussing its implications.

6.1: Fatal Endpoints

This section presents the comparative findings for fatal endpoints between Europe and Asia and North America and Asia and the findings of the universal comparison of all three continents.

6.1.1: The Pair-wise Comparisons

In this investigation, pair-wise comparisons of Europe and Asia and North America and Asia, respectively, were conducted for fatal endpoints. Chapter 4 (Section 4.4) describes the detailed methods used to conduct these comparisons. In brief, for each meta-analysis, each continent's trials were combined to provide a treatment effect estimate for each continent and then the difference between each continent's effect estimate was calculated to ascertain in which continent the intervention performed best. The number of meta-analyses favouring each continent was then summed to investigate whether there was a difference, over all meta-analyses, in the effect of intervention (relative to control) between the continents in the pair-wise comparison. The results of each pair-wise comparison are now presented.

6.1.1.1: Europe vs. Asia

For the pair-wise comparison between Europe and Asia, the meta-analyses on which it was based had to comprise at least one trial from Europe and one from Asia (see Chapter 4). Of the 95 systematic reviews containing meta-analyses, only 59 satisfied this criterion and could therefore be included in this comparison.

Only 20 of the 59 meta-analyses contained at least one European and one Asian trial with fatal endpoint data that allowed separate estimates of treatment effect to be calculated for each

continent. Details of each of these meta-analyses can be found in Tables 6.1, 6.2, and 6.3, which show the intervention and control group, the endpoint reported, the number of European and Asian trials in each meta-analysis, the effect estimate for both continents, and the continent in which the intervention, in each meta-analysis, performed best.

6.1.1.1.1.: Meta-analysis Grouped according to Treatment Effect Measure for Fatal Endpoints – Europe versus Asia.

The 20 meta-analyses that contained fatal endpoint data were grouped on how treatment effect was measured. For example, Table 6.1 details the meta-analyses that measured treatment effect using RR, and Table 6.3 details those using OR.

6.1.1.1.1.1: Fatal Endpoints for Europe versus Asia Measured as Risk Ratios

Thirteen of the 20 meta-analyses that reported fatal endpoints did so using RR. The details of these meta-analyses are summarised in Table 6.1.

Table 6.1 shows the meta-analysis details, the favoured continent, and the significance of any difference in how well the intervention performed between the two continents. For example, meta-analysis C12 (18) investigated two European and one Asian trial comparing Phosphodiesterase III inhibitors against a placebo for cardiovascular mortality. The treatment effect estimate was a RR of 1.562 for Europe and 0.51 for Asia. Intervention performed best in Asia, but the difference in treatment effect was not significant at the 5% level (2-sided p-value = 0.369).

Table 6.1: Meta-analyses with Fatal Endpoints for Europe versus Asia. Treatment Effect Measured by RR

Article ref. no.	Intervention	Control	Endpoint	No. of trials (Europe/ Asia)	RR estimate Europe *(95% CI)*	RR estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
18 [C12]	Phosphodiesterase III inhibitors for HF	Placebo	Cardiovascular mortality	2/1	1.562 *(1.02, 2.40)*	0.51 *(0.05, 5.60)*	Asia (0.369)
25 [P94]	Vasopressin for Cardiac arrest	Epinephrine	Death within 24hours	2/1	0.991 *(0.95, 1.03)*	0.83 *(0.68, 1.01)*	Asia (0.086)
227 [C4]	Anti-arrhythmics after cardioversion of AF	Placebo, drugs for rate control or no treatment	All-cause mortality	5/1	1.019 *(0.79, 1.31)*	0.62 *(0.03, 14.3)*	Asia (0.759)
232 [C10]	Intravenous magnesium for Acute MI	Placebo	Mortality by time of admission	5/1	0.776 *(0.61, 0.99)*	0.34 *(0.04, 3.21)*	Asia (0.475)
4 [P2]	G-CSF therapy for Acute MI	Placebo	Death	3/1	1.377 *(0.28, 6.89)*	3.63 *(0.16, 84.11)*	Europe (0.590)
383 [C41]	NPPV for Cardiogeneic pulmonary edema	Standard medical care	Hospital Mortality	6/4	0.587 *(0.37,0.92)*	0.40 *(0.18, 0.87)*	Asia (0.402)
96 [P90]	Adjunctive mechanical devices for Acute MI	Standard care or stenting	30 day mortality	5/2	1.00 *(0.47, 2.15)*	0.73 *(0.26, 2.07)*	Asia (0.628)
191 [P15]	PCI for late reperfusion after MI	Standard therapy	Death	2/1	0.96 *(0.41, 2.23)*	0.22 *(0.03, 1.82)*	Asia (0.204)
197 [P33]	MIDCAB for proximal stenosis of the left anterior descending artery	PCI	Mortality	3/1	0.93 *(0.45, 1.91)*	0.12 *(0.01, 2.43)*	Asia (0.194)
237 [P19]	Intracoronary cell therapy for Acute MI	Standard care	Death	3/1	0.51 *(0.15, 1.66)*	1.07 *(0.07, 16.33)*	Europe (0.619)
2 [P7]	Late PCI for Acute MI	Usual management	Death	5/1	0.41 *(0.26, 0.66)*	0.18 *(0.02, 1.45)*	Asia (0.446)
33 [P97]	Thrombectomy or Embolic protection devices with PCI for Acute MI	PCI alone	Mortality	9/2	0.70 *(0.16, 2.98)*	0.26 *(0.00, 281.07)*	Asia (0.786)

247 [C40]	Stem cell treatment for Acute MI	No treatment or Placebo	Mortality	4/1	0.59 *(0.21, 1.68)*	0.33 *(0.01, 7.81)*	Asia (0.739)

Six meta-analyses that used RR did not report which was the intervention group and which was the control group. In these cases, it was not possible to calculate which continent favoured intervention, as shown in Table 6.2.

Table 6.2: Meta-analyses with Fatal Endpoints for Europe vs. Asia that Do Not Identify Which is the Intervention Group and Which is the Control. Effectiveness measured by RR.

Article ref. no	**Intervention 1**	**Intervention 2**	**Endpoint**	**No. of trials (Europe/Asia)**	**RR estimate Europe *(95% CI)***	**RR estimate Asia *(95% CI)***	**Best Region (p-value of between-continent difference)**
11 [P87]	Stenting for Small Vessel CAD	PTCA	Death	6/1	0.859 *(0.46, 1.61)*	0.12 *(0.01, 2.13)*	Not applic. (0.187)
273 [C17]	Stenting for Acute MI	Balloon Angioplasty	30 day mortality	3/2	1.259 *(0.37, 4.31)*	0.40 *(0.10, 1.68)*	Not Applic. (0.234)
56 [P24]	PCI*	CABG surgery*	All-cause mortality	7/1	0.574 *(0.32, 1.03)*	1.18 *(0.11, 12.74)*	Not Applic. (0.567)
94 [P18]	LMWH for STEMI	UFH	30 day mortality	1/1	0.570 *(0.26, 1.25)*	0.80 *(0.39, 1.64)*	Not Applic. (0.526)
319 [P21]	SES for Coronary artery disease	PES	Patients experiencing death	7/3	1.010 (0.75, 1.36)	1.04 *(0.48, 2.26)*	Not Applic. (0.944)
26 [P38]	Minimally invasive left internal thoracic artery bypass for isolated lesions of left anterior descending artery	Percutaneous coronary artery stenting	Mortality at maximum follow up	3/1	2.32 *(0.51, 10.48)*	1 *(0.15, 6.82)*	Not Applic. (0.500)

6.1.1.1.1.2: Fatal Endpoints for Europe versus Asia Measured as Odds Ratios

Only one of the 20 meta-analyses that reported fatal endpoints did so using OR. The details for this meta-analysis are presented in Table 6.3.

Table 6.3: The Meta-analysis with Fatal Endpoints for Europe vs. Asia. Effectiveness measured by the OR.

Article ref. no.	Intervention	Control	Endpoint	No. of trials (Europe/ Asia)	OR estimate Europe *(95% CI)*	OR estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
286 [P153]	Amiodarone for prevention of AF	Placebo or routine treatment	Death	3/1	1.79 *(0.00, 1169.4)*	0.81 *(0.00, 153.8)*	Asia (0.853)

6.1.1.1.2: Examination over all Meta-analyses for Fatal Endpoints – Europe versus Asia.

Table 6.4 shows the number of meta-analyses on which Europe-Asia comparison calculations could be conducted.

Table 6.4: The Number of Meta-analyses favouring Each Continent – Europe versus Asia Fatal Data

Total Number of Meta-analyses (%)	Number of Meta-analyses favouring Europe (%)	Number of Meta-analyses favouring Asia (%)
14 (100%)	2 (14%)	12 (86%)

Of the 14 meta-analyses, 12 showed interventions to perform best in Asia. To find the significance of this, a binomial sign test was conducted under the null hypothesis that the effect

of interventions, relative to controls, was the same in both Europe and Asia. This showed that the effect of interventions, compared to controls, was statistically significantly different between Europe and Asia (2-sided p-value = 0.013).

6.1.1.1.3: Global Estimates of Continental Difference between Europe and Asia.

Global estimates of continental difference were calculated. These reflected the average weighted difference in effect between Europe and Asia over all of the meta-analyses that had specified their intervention and control group. From these estimates, it was found that for RR (13 meta-analyses) the mean log difference over all of the meta-analyses was 0.222 with 95% CIs of 0.0279 and 0.416 (2-sided p-value = 0.025). For OR (1 meta-analysis) the mean log difference was 0.791 with 95% CIs of -7.55 and 9.13 (2 sided p-value = 0.853). As can be seen, the global estimate of continental difference for RR was found to be statistically significant at the 5% level (p-value = 0.025). This showed that when treatment effect was measured using this summary statistic, there was a statistically significant difference in the effect of interventions between Europe and Asia, with, on average, intervention performing better, relative to control, in Asia than in Europe. The result for OR, on the other hand, was inconclusive since this did not reach the 5% level of statistical significance.

6.1.1.1.4: The Impact of Overlapping Meta-analyses

Given that overlapping meta-analyses were included (see Chapter 4), it is possible that the results summarised above are an artefact of their inclusion. The impact of overlapping meta-analyses on the pair-wise comparison between Europe and Asia was investigated, and six meta-analyses were identified as overlapping: P7 and P15, P90 and P97, and P19 and C40. Eight meta-analyses were

non-overlapping (P33, C12, P94, C4, C10, P2, C41, P153). The percentage of meta-analyses where intervention was favoured in Asia rather than Europe was 83% for overlapping meta-analyses but 88% for non-overlapping meta-analyses. This suggests that the proportion of meta-analyses favouring Asia over Europe was not exaggerated as a consequence of including overlapping meta-analyses.

6.1.1.1.5: Summary of the Europe versus Asia Pair-wise Comparison

In summary, when investigating within and over meta-analyses, there was some evidence to suggest that differences existed between Europe and Asia in treatment effect. For instance, there was a statistically significant difference in the number of meta-analyses whose interventions, relative to controls, performed better in Asia than in Europe. This was confirmed by the global estimate of continental difference for RR, which showed a statistically significant difference in treatment effect between Europe and Asia, with interventions performing better in Asia. Moreover, by investigating the impact of overlapping meta-analyses, such differences in treatment effect between Europe and Asia do not appear to have been created or exaggerated by the inclusion of overlapping meta-analyses. Therefore, there is some evidence that inter-continental differences exist between Europe and Asia when fatal endpoints are considered.

6.1.1.2: North America vs. Asia

The meta-analyses for this pair-wise comparison had to contain at least one North American trial and at least one Asian trial. Of the 95 relevant meta-analyses in the database, only 59 fulfilled this criterion and so could be used.

Of the 59 meta-analyses, only 12 contained trials with fatal endpoint data that would allow for separate estimates of treatment effect to be calculated for both continents. Details of these meta-analyses can be found in Tables 6.5 to 6.7.

6.1.1.2.1: Meta-analyses Grouped according to Treatment Effect Measure for Fatal Endpoints – North America versus Asia.

The 12 meta-analyses that reported fatal endpoints were grouped by the treatment effect measure used. Nine meta-analyses reported treatment effect using RR, one used OR, and two did not identify the intervention and control groups.

6.1.1.2.1.1: Fatal Endpoints for North America versus Asia as Risk Ratios.

Table 6.5 shows the details for the nine meta-analyses that used RR for fatal endpoints: the intervention and control group for each meta-analysis, the endpoint reported, the number of trials for each continent, the treatment effect estimates for both continents, and the details of where the intervention performed best.

Table 6.5: Meta-analyses with Fatal Endpoints for North America vs. Asia. Treatment Effect Measured by RR.

Article ref. no.	Intervention	Control	Endpoint	No. of trials (North America/ Asia)	RR estimate North America *(95% CI)*	RR estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
18 C12	Phosphodiesterase III inhibitors for HF	Placebo	Cardiovascular mortality	5/1	1.13 *(1.02, 1.26)*	0.51 *(0.05, 5.60)*	Asia (0.517)
25 P94	Vasopressin for Cardiac arrest	Epinephrine	Death within 24hours	1/1	1.02 *(0.92,1.14)*	0.83 *(0.68, 1.01)*	Asia (0.068)
96 P90	Adjunctive mechanical devices for Acute MI	Standard care or stenting	30 day mortality	1 /2	5.50 *(1.23, 24.55)*	0.73 *(0.26, 2.07)*	Asia (0.030)
227 C4	Anti-arrhythmics after cardioversion of AF	Placebo, drugs for rate control or no treatment	All-cause mortality	4/1	1.49 *(0.80, 2.77)*	0.62 *(0.03, 14.30)*	Asia (0.593)
2 P7	Late PCI for Acute MI	Usual management	Death	1/1	0.76 *(0.05, 11.39)*	0.18 *(0.02, 1.45)*	Asia (0.405)
4 P2	G-CSF therapy for Acute MI	Placebo	Death	1/1	0.18 *(0.01, 3.85)*	3.63 *(0.16, 84.19)*	North America (0.179)
33 P97	Thrombectomy or Embolic protection devices with PCI for Acute MI	PCI alone	Mortality	1 /2	2.80 *(0.05, 151.13)*	0.26 *(0.00, 281.07)*	Asia (0.562)
383 C41	NPPV	Standard medical Care	Hospital Mortality	1 /4	1 *(0.23, 4.40)*	0.40 *(0.18, 0.87)*	Asia (0.283)
55 P131	Clopidogrel plus Aspirin for the Prevention of Vascular Events	Antiplatelet Monotherapy	All-Cause Mortality	1/1	0.76 *(0.41, 1.39)*	0.93 *(0.88, 1.00)*	North America (0.500)

The best region was not calculated for two of the meta-analyses using RR, as these meta-analyses did not specifically identify which group was the intervention and which the control. It was not possible, therefore, to calculate the continent favouring intervention. Details are provided in Table 6.6.

Table 6.6: Meta-analyses with Fatal Endpoints for North America vs. Asia that Did Not Identify the Intervention or Control Groups. Effectiveness Measured by RR.

Article ref. no.	**Intervention 1**	**Intervention 2**	**Endpoint**	**No. of trials (North America/ Asia)**	**RR estimate North America *(95% CI)***	**RR estimate Asia *(95% CI)***	**Best Region (p-value of between-continent difference)**
56 P24	PCI	CABG surgery	All-cause mortality	3/1	0.74 *(0.41, 1.33)*	1.18 *(0.11, 12.74)*	Not Applic (0.713)
168 P140	Clopidogrel plus Aspirin for Cardiovascular Disease	Placebo plus Aspirin	Cardiovascular Mortality	1/1	0.63 *(0.29, 1.38)*	0.94 *(0.88, 1.00)*	Not Applic. (0.328)

6.1.1.2.1.2: Fatal Endpoints for North America versus Asia as Odds Ratios.

Only one of the 12 meta-analyses reported fatal endpoints in terms of OR, the details of which are provided in Table 6.7.

Table 6.7: The Meta-analysis with Fatal Endpoints for North America vs. Asia. Effectiveness measured by Odds Ratios (OR).

Article ref. no.	Intervention	Control	Endpoint	No. of trials (North America/ Asia)	OR estimate North America *(95% CI)*	OR estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
286 [P153]	Amiodarone for prevention of AF	Placebo or routine treatment	Death	6/1	0.82 *(0.06, 10.48)*	0.81 *(0.00, 153.8)*	Asia (0.997)

6.1.1.2.2: Examination over All Meta-analyses for Fatal Endpoints – North America versus Asia.

Table 6.8 shows the number of meta-analyses on which comparison calculations could be conducted favouring each continent.

Table 6.8: The Number of Meta-analyses Favouring Each Continent –North American versus Asian Fatal Data.

Total Number of Meta-analyses	Number of Meta-analyses Favouring North America	Number of Meta-analyses Favouring Asia
10	2	8

Two of the meta-analyses showed interventions performed better in North America, while eight showed interventions performed better in Asia. To assess the significance of this difference, a binomial sign test was conducted under the null hypothesis that the effectiveness of interventions, relative to controls, was the same in North America and Asia. This revealed that the effect of

interventions (relative to controls) was not significantly different between North America and Asia (2-sided p- value = 0.1094).

6.1.1.2.3: Global Estimates of the Difference between North America and Asia

Global estimates of continental difference were produced for each type of effect measure. These reflected the weighted average difference between the treatment effect estimates of North America and Asia over all of the meta-analyses that had specified their intervention and control groups. For RR (9 meta-analyses), the mean log difference was 0.241 with 95% CIs of -0.161 and 0.643 (2-sided p-value = 0.240). For OR (1 meta-analysis), the mean log difference was 0.0119, with 95% CIs of -5.82 and 5.84 (2- sided p-value = 0.9968). These findings suggested that over all of the meta-analyses and for both types of effect measure there was insufficient evidence to suggest a difference in overall treatment effect between North America and Asia, as neither estimate was statistically significant at the 5% level

6.1.1.2.4: The Impact of Overlapping Meta-analyses

Since there were no overlapping meta-analyses, they could have not have influenced the proportion of meta-analyses that favoured Asia over North America.

6.1.1.2.5: Summary of the North America versus Asia Pair-wise Comparison

In summary, when investigating both within and over meta-analyses, insufficient evidence was found to suggest that treatment effect differed between North America and Asia for fatal endpoints. This was because both the examination over all meta-analyses and the global estimates of continental difference were not statistically significant at the 5% level.

6.1.2: A Universal Comparison of Treatment Effectiveness Differences for Fatal Endpoints.

To explore the existence of inter-continental differences in treatment effectiveness further, the individual trials that made up the meta-analyses used in the previous section were analysed. For this, the effect estimate of each trial from Europe, North America, and Asia in each relevant meta-analysis was standardised using the formula described in Chapter 4 (Section 4.7.2.2). This was done within meta-analyses so that the trial effect estimate was compared only to the overall effect estimate of the meta-analysis in which it was situated. For each meta-analysis, the overall effect estimate was calculated by extracting all trials from Europe, North America, and Asia and combining them using the fixed effect approach. This provided the best available estimate of effect for the intervention in the meta-analysis since it was based on the trial data from all the continents, weighted according to the amount of information provided.

The standardised trial results from all meta-analyses were grouped by continent. Mean treatment effect estimates and standard deviations were calculated for each continent and then compared. Each continent's trial results were plotted onto graphs to facilitate the detection of outliers and to observe the spread of each continent's trial results.

Separate graphs were produced for Europe, North America, and Asia. The standardised treatment effect estimate of each trial was plotted on the horizontal axis and sample size[41] on the vertical

[41] Using the square root of the sample size rather than sample size enabled the vertical axis to be more condensed. Some trials had much larger sample sizes than the majority of trials. This meant that plotting data onto a graph using sample size would result in a large gap on the vertical axis between trials with the large samples and other trials. Using the square root of sample size reduced this gap.

axis.[42]The graphs were designed principally as visual aids for identifying outliers and the spread of trial results along the treatment effect axis. The zero on the x-axis of each graph represented a universal treatment effect estimate reflecting the best available effect estimate for the intervention in each trial.

However, as mentioned in Chapter 4, some of the systematic reviews addressed the same or similar clinical question and, as such, their meta-analyses may have contained a proportion of the same trials. Consequently, when meta-analyses were pooled, some trial results may have been duplicated and thus included in the analysis more than once. This is an important consideration because the continent reporting most support for intervention may result from that continent having the most duplicate trials with positive results for intervention.

In this study, duplications were classed as trials with the same authors, publication year, and sample size as another trial already in the meta-analysis dataset. Such evident duplications were removed from the dataset so that each trial only appeared once.[43] However, it is acknowledged that this method may not have removed all possible trial duplications. Some meta-analyses may have reported a sub-group analysis of a trial, the sample size and authors of which may have been different from the trial from which the data was taken, making it difficult to identify as trial duplication. Consequently, some trials in this analysis may have been correlated because data may have been taken from the same set of participants. Table 6.9 shows the number of trials in

---

[42] Using a measure of sample size on the vertical axis (Y axis) made the graphs easier to interpret. Furthermore, using sample size allowed assessment of whether the smaller trials in each continent were overestimating or underestimating treatment effectiveness (small-study effect).

[43] The universal comparison was also conducted with trial duplications included. However, its results became redundant once analysis had been conducted with duplications removed since the findings were qualitatively similar.

the original dataset,[44] the number of trial duplications per continent, the number of trials duplicated once and more than once, and the total number of trials per continent after duplications were removed.

Table 6.9: Duplicate and Non-duplicated Trials in each Continent that provided Fatal Data.

Continent	Number of Trials in Original Dataset	Number of Duplicates	Number of Trials Duplicated Once	Number of Trials Duplicated More than Once	Remaining Trials When Duplications Removed
Europe	311	71	53	9	240
North America	216	61	3	10	155
Asia	30	3	3	0	27
Total	557	135	89	19	422

As can be seen, there were originally 557 trials: 311 from Europe, 216 from North America, and 30 from Asia. Of these, 135 were duplicates and had to be removed from the dataset. This left 240 European trials, 155 North American trials, and 27 Asian trials to analyse.

6.1.2.1: Findings for the Universal Comparison – Fatal Endpoints

Figures 6.1-6.3 show the spread of each continent's trial results as regards fatal endpoints.

[44] Before duplications were removed.

FIGURE 6.1: THE SPREAD OF EUROPEAN TRIAL TREATMENT EFFECT ESTIMATES.

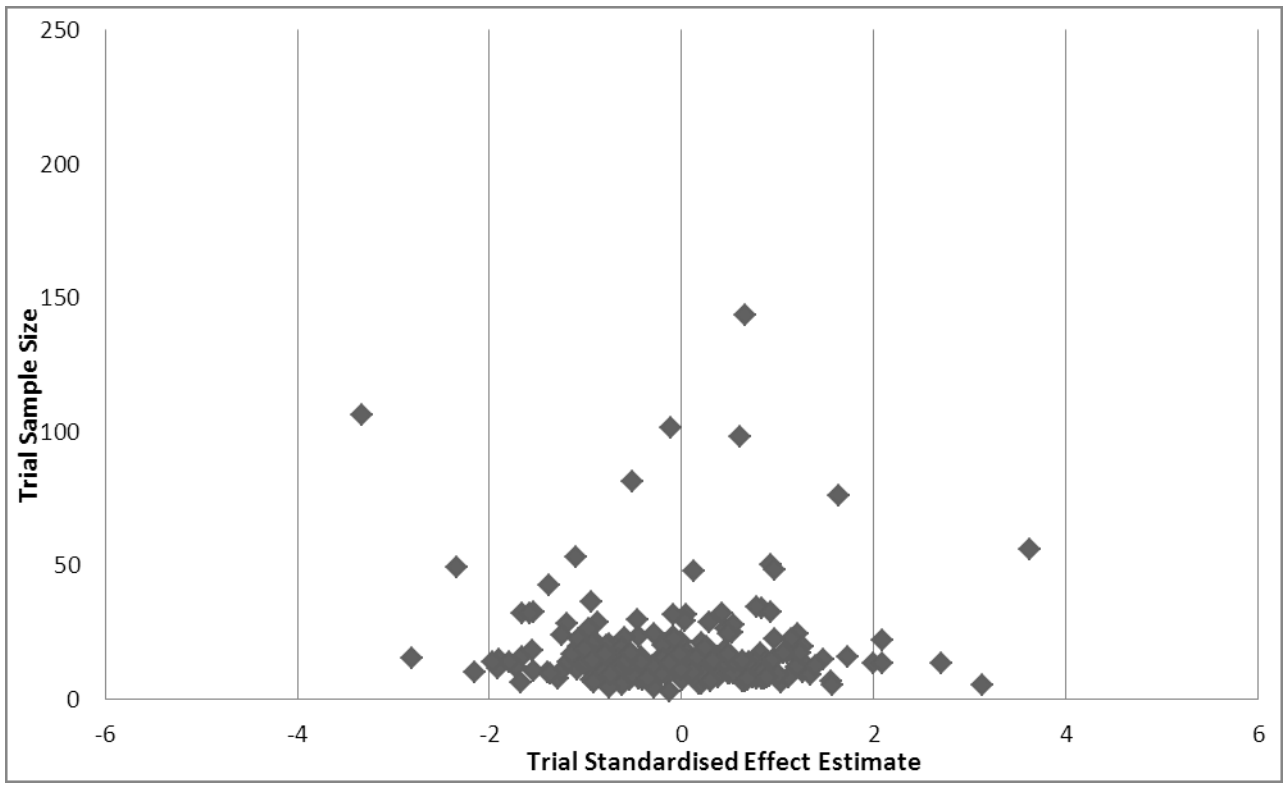


FIGURE 6.2: THE SPREAD OF NORTH AMERICAN TRIAL TREATMENT EFFECT ESTIMATES.

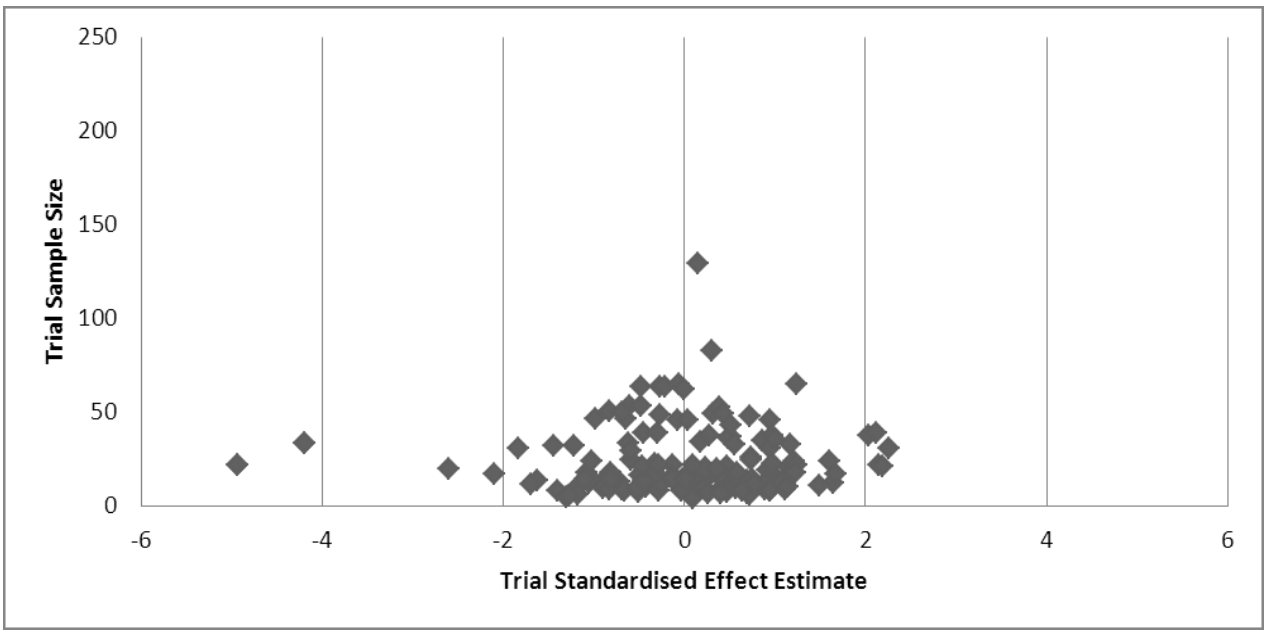
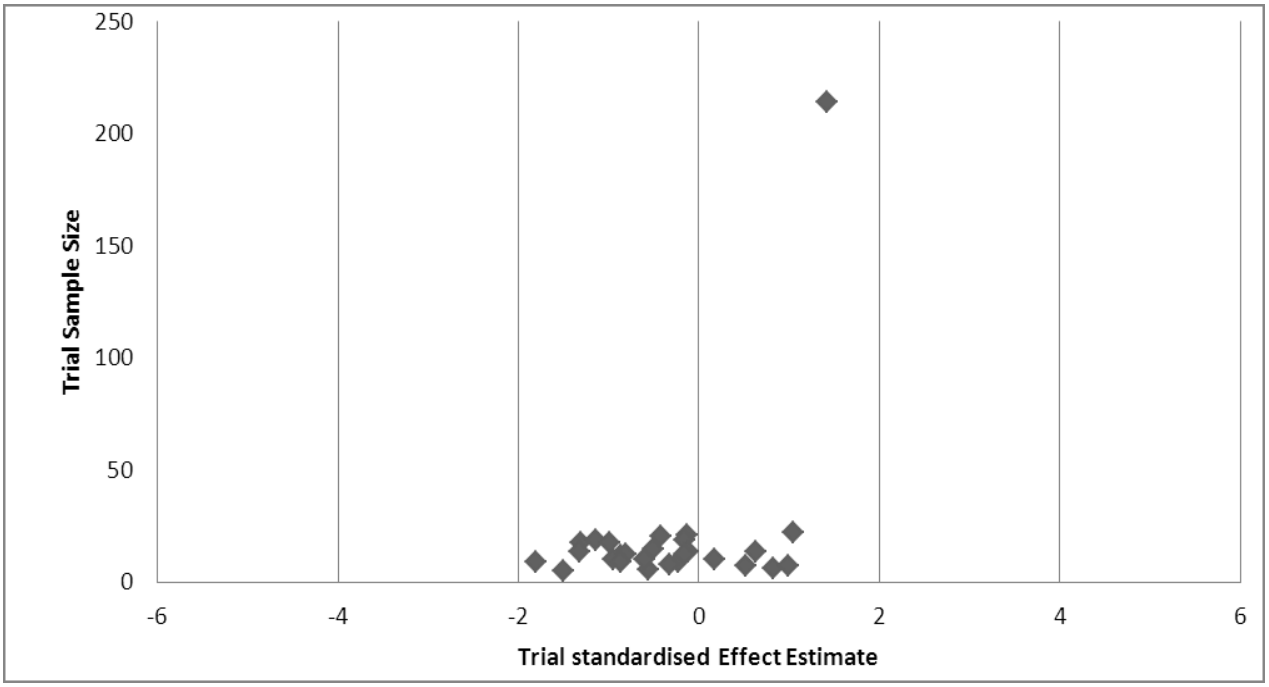


FIGURE 6.3: THE SPREAD OF ASIAN TRIAL TREATMENT EFFECT ESTIMATES.



As described earlier, Figures 6.1-6.3 were used to detect outliers for each continent. Outliers were defined as trial results deviating markedly from the others and above 2 standard errors or below -2 standard errors from the mean. However, few outliers were present in the datasets. For Europe and Asia, no outliers were detected for fatal endpoints. However, two outliers were found for North America. These were two extreme points below -4 standard errors (see Figure 6.2). Upon investigation, it was found that sponsors had discontinued both of these trials prematurely because of high mortality rates in the intervention group. As the early termination of trials has been found to distort treatment effect (Bassler et al. 2010; Schwartz, 1995; Souhami, Spiro, & Cullen, 1991),which could influence data analysis, it was decided to remove them from the dataset.

Once outliers had been detected and removed, Figures 6.1-6.3 were used to observe the spread of the individual trial results from each continent. These observations produced an interesting finding. When all three Figures were compared, it appeared that Asia produced more positive trial results, that is, more trials in Asia found interventions to be more effective (relative to controls) than trials from Europe and North America.

Table 6.10 shows the statistics calculated for each continent. It shows the mean effect estimate and its standard error, each continent's z-score,[45] its 2-sided p-value,[46] the number of positive

[45] The z-score provides an estimate of how effective the interventions, relative to controls, were in each continent. Indeed, the z-score of a standard normal distribution is seen as equivalent to an effect estimate (Coe, 2002).
[46]This, and the calculated z-score, help to determine how probable it is for the sample proportions to have happened if the null hypothesis was true.

trials found in each continent, and the 2-sided p-value for the binomial sign tests conducted by continent.

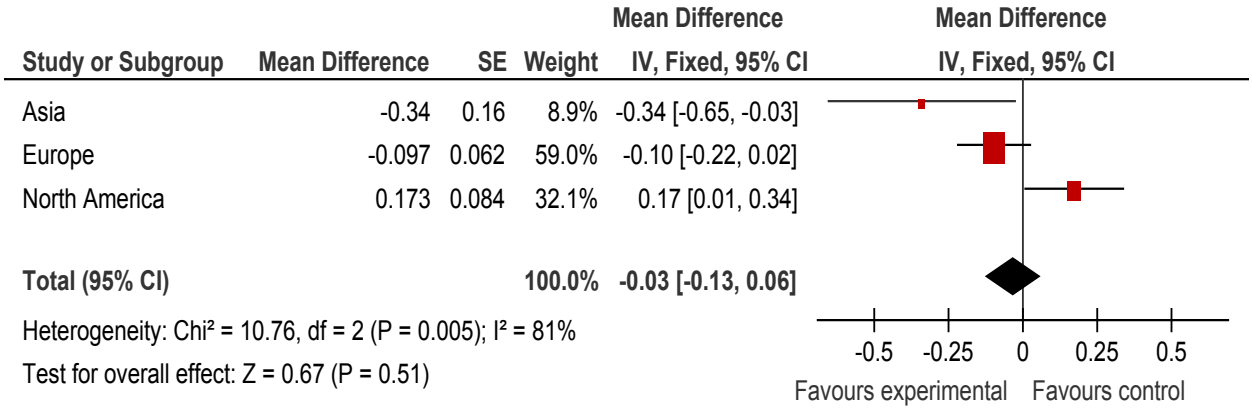
Table 6.10: Summary Statistics for Each Continent – Fatal Data

	Europe	**North America**	**Asia**
Mean Effect Estimate	-0.097	0.173	-0.340
Standard Error	0.062	0.084	0.160
Z-Score	-1.565	2.077	-2.121
p-value (2-sided)	0.118	0.038	0.034
Number of Trials (% of positive trials)	240 (53%)	153 (47%)	27 (74%)
Sign Test p-value (2-sided)	0.401	0.518	0.019

The mean effect sizes differed greatly between continents. In North America, the treatment effect estimate (Mean = 0.173, SE = 0.084) seemed to favour controls, whereas in Europe (Mean= -0.097, SE = 0.062) and Asia (Mean = -0.340, SE = 0.160) interventions were favoured. Indeed, the 2-sided p-value for the z-score of North America was statistically significant at the 5% level (2-sided p-value = 0.038), suggesting that in this instance, controls performed better relative to interventions. Asia appeared to provide the largest treatment effect estimate, suggesting that interventions performed better, relative to controls, in Asian trials than in their European and North American counterparts. The binomial sign test conducted for Asia confirmed that Asia did provide significantly more positive trial results, with a 20:27 ratio of trials showing interventions performing better relative to controls (2-sided p-value = 0.0192), a finding corroborated by the statistically significant at the 5% level 2-sided p-value calculated for the z-score in Asia (p-value = 0.034).

Figure 6.4 also shows that the effect of interventions differed greatly between continents.[47]It can be seen, for instance, that while intervention was more effective than control in Asia and Europe, control was more effective in North America. Moreover, the confidence intervals for each continent suggest the presence of a high level of heterogeneity, since there is no overlap between the confidence intervals of North America and Asia. This is supported by the I^2 statistic, which shows that the percentage of total variation between continents due to heterogeneity, rather than chance, is 81%.

FIGURE 6.4: A FOREST PLOT OF THE EFFECTIVENESS OF INTERVENTIONS IN EUROPE, NORTH AMERICA, AND ASIA FOR FATAL ENDPOINTS.



To investigate further, two independent sample t-tests were performed to compare the effectiveness of interventions between Europe and Asia and North America and Asia. The t-test for Europe versus Asia showed that the effectiveness of interventions, relative to controls, was

[47] It should be acknowledged that analysis was conducted on data within meta-analyses so that between-continent differences could not be created by the inclusion of different intervention types.

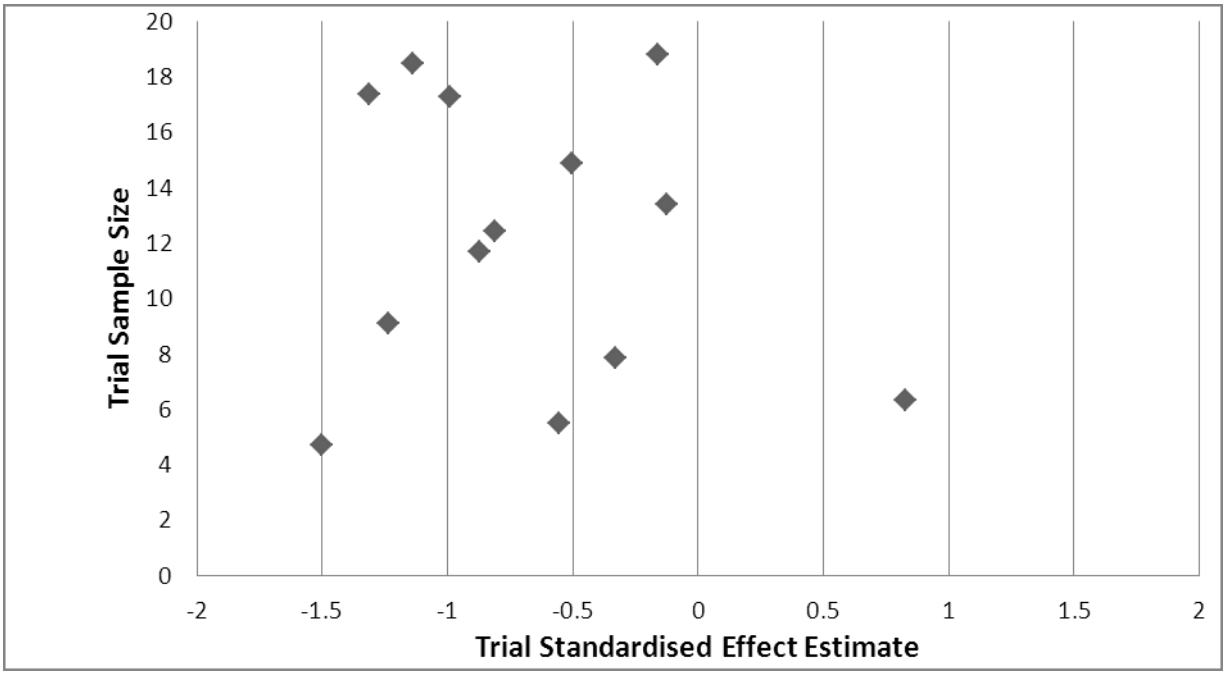
not significantly different between Europe (Mean = -0.097, SE = 0.062) and Asia (Mean = -0.340, SE = 0.160), (t (34) = -1.4161, 2-sided p-value = 0.1658). The results for North America versus Asia, however, indicated that interventions, relative to controls, were more effective in Asia (Mean = -0.340, SE=0.160) than in North America (Mean = 0.017, SE = 0.084), with this difference being statistically significant (*t* (41) = -2.839, 2-sided p-value = 0.0070).

These results suggested that there was some evidence for interventions, relative to controls, performing better in Asian trials than in their European and North American counterparts. As this finding was unexpected, it was decided to investigate further by conducting a post-hoc analysis on the Asian sample of trials. This was done by breaking down the Asian trials by country to establish if there was a particular country with a high proportion of positive trial results. The Asian trials were from China, Japan, Taiwan, and South Korea. As with the previous graphs, the spread of trial results from each country was examined by standardising each trial's effect estimate, sub-grouping trial results according to country, and plotting these onto graphs.[48]

In observing the spread of each Asian country's trial results, it was found that Japanese trials produced more positive results, with the majority favouring intervention over control. Figure 6.5 shows that the majority of the standardised effect estimates of Japanese trials were below zero.

[48]Since this was a post-hoc analysis and the sample sizes for each Asian country were small, statistical tests were considered inappropriate and so were not used.

FIGURE 6.5: THE TREATMENT EFFECT ESTIMATES OF JAPANESE TRIALS



This can be seen by examining Table 6.11, which shows the total number of trials in each country and the number of trials favouring intervention or control.

Table 6.11: The Number of Asian Trials favouring the Intervention or Control in Japan, China, Taiwan, and South Korea

Country of Trials	Number of trials favouring the intervention (%)	Number of Trials favouring the Control (%)	Total Number of Trials (%)
Japan	12 (92%)	1 (8%)	13 (100%)
China	3 (60%)	2 (40%)	5 (100%)
South Korea	2 (33%)	4 (67%)	6 (100%)
Taiwan	2 (67%)	1 (33%)	3 (100%)

As can be seen, Japan provided a high proportion of positive trial results, with 12 of 13 trials reporting that intervention was more effective than control. A high proportion of positive trial results were not observed for any other Asian country included in this sample. This suggested that interventions, relative to controls, may be more effective in Japanese trials compared with the other Asian countries included here.

6.1.3: Types of Intervention Prone to Inter-Continental Differences in Effectiveness - Fatal Endpoints

Since there was some evidence to suggest that differences in treatment effectiveness existed between Europe and Asia and North America and Asia, it was important to investigate which types of intervention were prone to these differences. First, this was investigated over and within meta-analyses. For each pair-wise comparison, the meta-analyses were grouped according to intervention type. Within these groups, the Z-statistic and 2-sided p-value of each meta-analysis were observed to see if, over all meta-analyses in that intervention type, there was a clear pattern of statistically significant differences in treatment effect. This was done using Tables 6.12 to 6.14 and Tables 6.16 to 6.18.

Second, the individual trials from the continents of interest contained in the included meta-analyses were examined. The standardised results of each trial in each meta-analysis were pooled and grouped according to continent and intervention type. T-tests were then conducted between Europe and Asia and between North America and Asia for each type of intervention: Drug, Device, and Surgery. Lifestyle and Management interventions were not included because no Asian trials involved these types of intervention for fatal endpoints. Conducting t-tests meant that

the 2-sided p-value for the between-continent difference in treatment effect could be calculated for each intervention type. This enabled the types of intervention prone to inter-continental differences to be identified.

6.1.3.1: Europe vs. Asia

First, the investigation into the types of intervention prone to inter-continental differences for fatal endpoints was conducted over and within meta-analyses. To do this, the meta-analyses were grouped according to intervention type. The Z-statistic and 2-sided p-value for each meta-analysis (now grouped by intervention type) were then "eyeballed" to see whether any difference was statistically significant at the 5% level. The number of statistically significant differences in each intervention group was then observed to see whether, over all of the meta-analyses in that intervention group, there was a clear pattern of statistically significant differences that would suggest that that intervention type was likely to show differences in treatment effect between Europe and Asia.

Table 6.12 shows the meta-analyses investigating drug therapies for the Europe versus Asia pair-wise comparison.

Table 6.12: Meta-analyses Investigating Drugs with their Z-statistic for the Difference in Treatment Effect between Europe and Asia for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe /Asia)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
18 [C12]	Phosphodiesterase III inhibitors	RR	2/1	1.56 *(1.02,2.40)*	0.51 *(0.05, 5.60)*	1.112 *(-1.32, 3.54)*	0.898	0.369
4 [P2]	GCSF Therapy	RR	3/1	1.38 *(0.28,6.89)*	3.63 *(0.16, 84.11)*	-0.970 *(-4.50, 2.56)*	-0.538	0.590
232 [C10]	Magnesium	RR	5/1	0.78 *(0.61, 0.99)*	0.34 *(0.04, 3.21)*	0.823 *(-1.43, 3.08)*	0.714	0.475
237 [P19]	Intracoronary Cell Therapy	RR	3/1	0.51 *(0.15, 1.66)*	1.07 *(0.07, 16.33)*	-0.752 *(-3.72, 2.22)*	-0.497	0.619
247 [C40]	Bone Marrow Cell Therapy	RR	4/1	0.59 *(0.21, 1.68)*	0.33 *(0.01, 7.81)*	0.565 *(-2.76, 3.89)*	0.333	0.739
227 [C4]	Anti-arrhythmics	RR	5/1	1.02 *(0.79, 1.31)*	0.62 *(0.03, 14.3)*	0.492 *(-2.65, 3.64)*	0.307	0.759
25 [P94]	Vasopressin	RR	2/1	0.99 *(0.95, 1.03)*	0.83 *(0.68, 1.01)*	0.181 *(-0.03, 0.39)*	1.716	0.086
286 [P153]	Amiodarone	OR	3/1	1.79 *(0.00, 1169.4)*	0.81 *(0.00, 153.8)*	0.791 *(-7.55, 9.13)*	0.186	0.853

As can be seen, it was found that the difference in treatment effect between Europe and Asia was statistically significant in none of the included meta-analyses that had investigated drugs. As such, no clear pattern was identified to show that this type of intervention was prone to inter-continental differences in treatment effect.

Table 6.13 shows the details of the meta-analyses that investigated medical devices.

Table 6.13: Meta-analyses Investigating Medical Devices with their Z-statistic for the Difference in Treatment Effect between Europe and Asia for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe /Asia)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
33 [P97]	Thrombectomy or Embolic Protection Devices	RR	9/2	0.70 *(0.16, 2.98)*	0.26 *(0.00, 281.07)*	0.991 *(-6.15, 8.13)*	0.272	0.786
96 [P90]	Adjunctive Mechanical Devices	RR	5/2	1.00 *(0.47, 2.15)*	0.73 *(0.26, 2.07)*	0.320 *(-0.97, 1.61)*	0.485	0.628
383 [C41]	Non-Invasive Positive Pressure Ventilation	RR	6/4	0.59 *(0.37, 0.92)*	0.40 *(0.18, 0.87)*	0.383 *(-0.51, 1.28)*	0.838	0.402

Again, no meta-analysis found the difference in treatment effect between Europe and Asia to be statistically significant at the 5% level. As a result, no clear pattern of statistically significant differences was found to suggest that medical devices were prone to between-continental differences in treatment effect.

Table 6.14 provides the details for the meta-analyses that investigated surgical interventions.

Table 6.14: Meta-analyses Investigating Surgical Interventions with their Z-statistic for the Difference in Treatment Effect between Europe and Asia for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ Asia)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
197 [P33]	Direct Coronary Bypass Grafting	RR	3/1	0.93 *(0.45,1.91)*	0.12 *(0.01,2.43)*	2.060 *(-1.05, 5.17)*	1.300	0.194
191 [P15]	PCI	RR	2/1	0.96 *(0.41, 2.23)*	0.22 *(0.03, 1.82)*	1.468 *(-0.80, 3.73)*	1.270	0.204
2 [P7]	Late PCI	RR	5/1	0.41 *(0.26, 0.66)*	0.18 *(0.02, 1.45)*	0.839 *(-1.32, 3.00)*	0.763	0.446

As with medical devices, no meta-analysis investigating surgical interventions found the difference in treatment effect between Europe and Asia to be statistically significant at the 5% level. Therefore, no clear pattern of statistically significant differences was found for surgical interventions to suggest that such interventions were prone inter-continental differences.

Since no group of interventions was found to show a clear pattern of statistically significant differences in treatment effect, the types of intervention prone to differences between Europe and Asia could not be identified for fatal endpoints.

To investigate further, analysis was conducted across the individual trials contained in the included meta-analyses. Table 6.15 shows the number of trials from each continent that investigated each intervention type, the mean effect estimate and its standard error per continent for each intervention type, and the 2-sided p-value for the t-tests that were conducted.

Table 6.15: Summary Statistics for Intervention Types – Europe versus Asia for Fatal Data.

Intervention Type	Number of European Trials	European Mean Effect Estimate	Standard Error	Number of Asian Trials	Asia Mean Effect Estimate	Standard Error	T-Test p-value
Drug	91	-0.0779	0.103	8	-0.066	0.365	0.976
Device	49	0.0054	0.106	14	-0.526	0.176	0.017
Surgery	41	-0.0110	0.154	5	-0.257	0.435	0.617

As can be seen, the number of trials investigating each intervention type differed between continents, as did each continent's mean effect estimate. The 2-sided p-values for the t-tests show that the difference in treatment effect between Europe and Asia was only statistically significant at the 5% level for device interventions. The results of the t-tests, for each intervention type, are discussed further below.

6.1.3.1.1: Drug Interventions

The mean difference between continents for drug interventions was found to be -0.012, which is a small effect size ($d = 0.022$),[49] with 95% confidence intervals of -0.887 to 0.863. An independent samples t-test revealed that the effectiveness of drug therapies did not significantly differ between Europe (Mean=-0.078, SE = 0.103) and Asia (Mean = -0.066, SE = 0.365): (t (8) = 0.0314, 2-sided p-value = 0.976). Therefore, it was concluded that there was insufficient evidence for a difference in the effect of drug therapies, relative to controls, between Europe and Asia.

[49] Cohen's *d* was used as the measure of effect size. When using this measure, Cohen (1988) considers that a *d* of 0.2 represents a small effect size, a *d* of 0.5 a medium effect size, and a *d* of 0.8 a large effect size (Zhang, 2008).

6.1.3.1.2: Device Interventions

The mean difference between continents for medical devices was 0.53, representing a large effect size (d = 1.08), with 95% confidence intervals of 0.106 to 0.956. An independent samples t-test showed that there was a statistically significant difference in the effect of device interventions, relative to controls, between Europe (Mean=0.005, SE=0.106), and Asia (Mean= -0.526, SE= 0.176), (t (23) = 2.586, 2-sided p-value =0.017). Therefore, these results indicated that a difference in the effect of medical devices existed between Europe and Asia, with such devices performing better in Asia than in Europe.

6.1.3.1.3: Surgical Interventions

The mean difference between continents for surgical interventions was 0.246, reflecting a small-medium effect size (d = 0.48), with confidence intervals of -0.940 to 1.42. An independent samples t-test indicated that there was no significant difference in the effect of surgical interventions, relative to controls, between Europe (Mean = -0.011, SE = 0.154) and Asia (Mean = -0.257, SE = 0.435): (t (5) = 0.5331, 2-sided p-value = 0.617). As a result, it was concluded that there was currently insufficient evidence that the effect of surgical interventions differed between Europe and Asia.

6.1.3.2: North America vs. Asia

As with the previous comparison, the types of intervention prone to differences in treatment effect was first examined over and within meta-analyses. As before, the meta-analyses were grouped according to the type of intervention investigated. Table 6.16 shows the meta-analyses investigating drug interventions for the pair-wise comparison between North America and Asia.

Table 6.16: Meta-analyses Investigating Drugs with their Z-statistic for the Difference in Treatment Effect between North America and Asia for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (N. Amer./ Asia)	Pooled Estimate for North America *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
4 [P2]	GCSF Therapy	RR	1/1	0.18 *(0.01, 3.85)*	3.63 *(0.16, 84.19)*	-3.007 *(-7.40, 1.38)*	-1.343	0.179
227 [C4]	Anti-arrhythmics	RR	4/1	1.49 *(0.80, 2.77)*	0.62 *(0.03, 14.30)*	0.870 *(-2.32, 4.07)*	0.534	0.593
18 [C12]	Phosphodiesterase III Inhibitors	RR	5/1	1.13 *(1.02, 1.26)*	0.51 *(0.05, 5.60)*	0.792 *(-1.60, 3.18)*	0.649	0.517
25 [P94]	Vasopressin	RR	1/1	1.02 *(0.92, 1.14)*	0.83 *(0.68, 1.01)*	0.213 *(-0.016, 0.44)*	1.826	0.068
55 [P131]	Clopidogrel plus Aspirin	RR	1/1	0.76 *(0.41, 1.39)*	0.93 *(0.88, 1.00)*	-0.210 *(-0.82, 0.40)*	-0.675	0.500
286 [P153]	Amiodarone	OR	6/1	0.82 *(0.06, 10.48)*	0.81 *(0.00, 153.8)*	0.012 *(-5.82,5.84)*	0.004	0.997

As can be seen, none of the meta-analyses investigating drug interventions found the difference in treatment effect to be statistically significant at the 5% level. As a result, no clear pattern of statistically significant differences was found to show drug interventions as being prone to inter-continental differences in treatment effect.

Table 6.17 presents the details of the meta-analyses that investigated medical devices.

Table 6.17: Meta-analyses Investigating Medical Devices with their Z-statistic for the Difference in Treatment Effect between North America and Asia for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (N. Amer./ Asia)	Pooled Estimate for North America *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
33 [P97]	Thrombectomy or Embolic Protection Devices	RR	1/2	2.80 *(0.05, 11.39)*	0.26 *(0.00, 281.07)*	2.382 *(-5.67, 10.43)*	0.580	0.562
96 [P90]	Adjunctive Mechanical Devices	RR	1/2	5.50 *(1.23, 24.55)*	0.73 *(0.26, 2.07)*	2.022 *(0.20, 3.85)*	2.173	0.030
383 [C41]	Non-Invasive Positive Pressure Ventilation	RR	1/4	1 *(0.23, 4.40)*	0.40 *(0.18, 0.87)*	0.916 *(-0.76, 2.59)*	1.074	0.283

One meta-analysis investigating medical devices found a statistically significant difference in treatment effect between North America and Asia (P90 (96)). This meta-analysis compared adjunctive mechanical devices (intervention) to usual care (control) with respect to thirty-day mortality and found that adjunctive mechanical devices performed better in Asia than in North America. The 2-sided p-value for this Z-statistic was 0.030. However, as no other meta-analyses investigating medical devices found a statistically significant difference in treatment effect, no clear pattern of statistically significant differences could be identified to indicate that medical devices were prone to differences in treatment effect between North America and Asia.

Table 6.18 shows the meta-analysis containing North American and Asian trials that investigated a surgical intervention for fatal endpoints.

Table 6.18: Meta-analyses Investigating Surgical Interventions with their Z-statistic for the Difference in Treatment Effect between North America and Asia for Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (N. Amer. /Asia)	Pooled Estimate for North America *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
2 [P7]	Late PCI	RR	1/1	0.76 *(0.05, 11.39)*	0.18 *(0.02, 1.45)*	1.456 *(-1.97, 4.88)*	0.832	0.405

As Table 6.18 shows, the difference in treatment effect did not reach statistical significance in this meta-analysis. Therefore, there was no evidence that this surgical intervention was prone to inter-continental differences in treatment effect.

As no group of interventions showing a clear pattern of statistically significant differences could be identified, the types of intervention prone to treatment effect differences between North America and Asia could not be identified. Furthermore, finding only one meta-analysis that was statistically significant at the 5% level from the 10 available was no more than would be expected by random chance. Therefore, there was insufficient evidence to conclude which types of intervention were prone to differences in treatment effect between North America and Asia for fatal endpoints.

To investigate further, the North American and Asian trials from each included meta-analysis were analysed, with independent samples t-tests conducted for each type of intervention: Drug, Device, and Surgery. Table 6.19 shows the summary statistics for North America and Asia.

Table 6.19: Summary Statistics for Intervention Types– North America versus Asia for Fatal Data.

Intervention Type	Number of North American Trials	North American Mean Effect Estimate	Standard Error	Number of Asian Trials	Asia Mean Effect Estimate	Standard Error	T-Test p-value
Drug	75	-0.186	0.128	8	-0.066	0.365	0.764
Device	18	0.199	0.204	14	-0.526	0.176	0.012
Surgery	12	0.291	0.209	5	-0.257	0.435	0.308

As can be seen, for each intervention type the number of trials each continent provided differed, as did each continent's mean effect estimate. The results of the t-tests indicated that the difference in treatment effect between North America and Asia was statistically significant at the 5% level for medical devices. The results of each t-test are now discussed further.

6.1.3.2.1: Drug Interventions

The mean difference between continents for drug therapies was -0.120, representing a small effect size (d=-0.22), with 95% confidence intervals of -1.012 to 0.772. An independent sample t-test revealed that there was not a statistically significant difference in the effect of drug interventions, relative to controls, between North America (Mean= -0.186, SE = 0.128) and Asia (Mean=-0.066, SE = 0.365) ($t(8) = 0.3102$, p=0.764). As such, there was insufficient evidence for the existence of a difference in treatment effect for drug interventions between North America and Asia.

6.1.3.2.2: Device Interventions

The mean difference between continents for devices was found to be 0.7250, indicating a large effect size (d = 1.00), with 95% CIs of 0.174 to 1.276. An independent samples t-test showed that there was a statistically significant difference in the effect of medical devices, relative to controls, between North America (Mean = 0.199, SE = 0.204) and Asia (Mean = -0.526, SE = 0.176): t (29) = 2.6909, 2-sided p-value= 0.012. Therefore, these results demonstrated that there was a difference in the effect of medical devices between North America and Asia, with such devices performing better in Asia than in North America.

6.1.3.2.3: Surgical Interventions

The mean difference between continents for surgical interventions was 0.548, reflecting a large effect size (d = 0.98), with 95% CIs of -0.693 to 1.789. An independent samples t-test revealed that there was no statistically significant difference in the effect of surgical interventions, relative to controls, between North America (Mean = 0.281, SE = 0.209) and Asia (Mean= -0.257, SE= 0.435) (t(5) = 1.1355, p = 0.308). As such, there was a lack of evidence that the effect of surgical interventions differed between North America and Asia.

6.1.3.3: Summary of the Types of Intervention Prone to Inter-continental Differences.

The types of intervention prone to inter-continental differences in treatment effect could not be identified from the investigation over and within meta-analyses for either pair-wise comparison. This was because no clear pattern of statistically significant differences was found for any intervention type. However, the investigation conducted across trials indicated that medical devices may be prone to showing inter-continental differences, since in both the Europe versus

Asia and the North America versus Asia comparisons the t-tests showed statistically significant differences between continents in treatment effect.

6.2: Non-Fatal Endpoints

This section presents the findings of the pair-wise comparison between Europe and Asia and North America and Asia for non-fatal endpoints and the findings from the universal comparison for this endpoint.

6.2.1: Pair-Wise Comparison over Meta-analyses.

As with the pair-wise comparison for fatal endpoints, two comparisons were conducted for non-fatal endpoints between Europe and Asia and between North America and Asia. Each will be discussed in turn.

6.2.1.1: Europe vs. Asia

As explained previously (see Section 6.1.1.1.1), the meta-analyses on which this pair-wise comparison were based had to consist of at least one European and one Asian trial and only 59 of the 95 meta-analyses in the database fulfilled this criterion.

Of these 59 meta-analyses, 39 contained at least one European and one Asian trial with data for a non-fatal endpoint that would allow for separate estimates of treatment effect to be calculated. The details of these meta-analyses can be found in Tables 6.20–6.24.

6.2.1.1.1: Meta-analyses Grouped according to Treatment Effect Measure for Non-Fatal Endpoints – Europe versus Asia.

The 39 meta-analyses that contained non-fatal data were grouped according to the measure of treatment effect used. The following tables summarise the details of these meta-analyses, showing the intervention group, the control group, the endpoint reported in each meta-analysis, the number of trials provided by each continent, and each continent's treatment effect estimate.

6.2.1.1.1.1: Non-Fatal Endpoints for Europe versus Asia Measured as Risk Ratios.

24 of the 39 meta-analyses that reported non-fatal endpoints did so using RR. The details of these are given in Table 6.20.

Table 6.20: Meta-analyses with Non- Fatal Endpoints for Europe vs. Asia. Effectiveness Measured by RR.

Article ref. no.	Intervention	Control	Endpoint	No. of trials (Europe /Asia)	RR estimate Europe *(95% CI)*	RR estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
18 [C12]	Phosphodiesterase III inhibitors for HF	Placebo	Worsening heart failure	3/2	1.44 *(1.00, 2.06)*	0.52 *(0.27, 0.98)*	Asia (0.006)
25 [P94]	Vasopressin for Cardiac arrest	Epinephrine	Failure of return of spontaneous circulation	2/1	1.03 *(0.97, 1.11)*	0.59 *(0.38, 0.91)*	Asia (0.013)
48 [P166]	Vitamin-mineral supplements for CAD	Placebo	Restenosis	2/1	1.05 *(0.43, 2.61)*	0.56 *(0.06, 5.31)*	Asia (0.609)
176 [C39]	Calcium channel antagonists	Adenosine	Reversion Rate	2/1	1.54 *(0.58, 4.06)*	0.97 *(0.39, 2.41)*	Asia (0.497)
96 [P90]	Adjunctive mechanical devices for Acute MI	Standard care or stenting	Distal embolisation	4/2	0.65 *(0.37, 1.13)*	0.40 *(0.21, 0.75)*	Asia (0.255)
226 [P125]	Glycoprotein IIb/IIIa inhibitors for PCI	Placebo or standard care	MI within 30 days	4/1	0.86 *(0.58, 1.28)*	0.13 *(0.01, 2.38)*	Asia (0.206)
227 [C4]	Anti-arrhythmics after cardioversion of AF	Placebo, drugs for rate control or no treatment	AF Recurrence	11/2	0.75 *(0.71, 0.79)*	0.85 *(0.70, 1.03)*	Europe (0.220)
232 [C10]	Intravenous magnesium for Acute MI	Placebo	Heart Failure	4/1	0.79 *(0.65, 0.97)*	0.66 *(0.39, 1.10)*	Asia (0.512)
33 [P97]	Thrombectomy or Embolic protection devices with PCI for Acute MI	PCI alone	Target vessel revascularisation	6/2	0.94 *(0.30, 2.95)*	0.66 *(0.02, 21.39)*	Asia (0.853)
383 [C41]	NNPV for Cardiogenic Pulmonary Edema	Standard Medical Care	Endotracheal Intubation Rate	4/1	0.86 *(0.58, 1.28)*	0.13 *(0.01, 2.38)*	Asia (0.206)
388 [P101]	Prophylactic steroids for cardiopulmonary bypass	Placebo or standard care	New onset AF	3/1	0.78 *(0.62, 0.98)*	1.25 *(0.49, 3.16)*	Europe (0.338)
51 [C11]	Glycoprotein IIb/IIIa inhibitors for PCI	Placebo	30 day urgent revascularisation	5/1	0.56 *(0.35, 0.88)*	2.74 *(0.12,63. 63)*	Europe (0.327)

165 [P130]	ACEI or ARB	Placebo or alternative therapy	AF	6/1	0.93 *(0.86, 1.00)*	0.60 *(0.37,0.97)*	Asia (0.080)
328 [P108]	Off pump CABG surgery	On pump CABG surgery	Stroke	9/1	0.41 *(0.18, 0.95)*	0.35 *(0.01, 8.56)*	Asia (0.928)
2 [P7]	Late PCI for Acute MI	Usual management	Non-Fatal MI	5/1	0.49 *(0.31, 0.77)*	0.38 *(0.11, 1.37)*	Asia (0.715)
4 [P2]	G-CSF Therapy	Placebo	Instent Restenosis	4/1	0.91 *(0.58, 1.41)*	1.22 *(0.08, 18.20)*	Europe (0.830)
188 [P25]	G-CSF Administration	Placebo or No Treatment	Binary Restenosis rate	4/1	0.97 *(0.61, 1.53)*	2.75 *(0.24, 31.72)*	Europe (0.410)
200 [P155]	ACEI or ARB for prevention of AF	Placebo, other antihypertensive therapy or treatment drug plus Irbesartan	Development of AF	6/1	0.93 *(0.68, 1.27)*	0.60 *(0.10, 3.76)*	Asia (0.647)
247 [C40]	Stem cell treatment for Acute MI	No treatment or Placebo	Incidence of Restenosis	5/1	1.15 *(0.71, 1.87)*	0.20 *(0.01, 3.97)*	Asia (0.256)
259 [P102]	Off pump CABG surgery	On pump CABG surgery	MI	15/1	0.74 *(0.46, 1.19)*	1.06 *(0.15, 7.36)*	Europe (0.721)
400 [P5]	G-CSF for Acute MI	Placebo	Angiographic Restenosis	4/2	1.00 *(0.63, 1.59)*	0.83 *(0.40, 1.74)*	Europe (0.672)
153 [P149]	Amiodarone prophylaxis for cardiothoracic surgery	Placebo	Post-operative AF	6/1	0.70 *(0.55, 0.90)*	0.36 *(0.18,0.71)*	Asia (0.066)
191 [P15]	PCI for late reperfusion after MI	Standard therapy	MI	2/1	1.72 *(0.52, 5.71)*	0.38 *(0.11, 1.37)*	Asia (0.092)
237 [P19]	Intracoronary Cell Therapy	Standard medical therapy	Target vessel revascularisation	6/1	0.997 *(0.70, 1.43)*	0.36 *(0.02, 8.41)*	Asia (0.527)

As for fatal endpoints, the best region was not calculated for six of the meta-analyses using RR, as they had not specifically stated which group was the intervention and which group the control. Details are provided in Table 6.21.

Table 6.21: Meta-analyses with Non-Fatal Endpoints for Europe vs. Asia that Do Not Identify Which Group is the Intervention and Which the Control. Effectiveness Measured by RR.

Article ref. no	**Intervention 1**	**Intervention 2**	**Endpoint**	**No. of trials (Europe/ Asia)**	**RR estimate Europe *(95% CI)***	**RR estimate Asia *(95% CI)***	**Best Region (p-value of between-continent difference)**
11 [P87]	Stenting for Small Vessel CAD	PTCA	Repeat revascularisation	6/2	0.77 *(0.64, 0.92)*	0.72 *(0.41, 1.26)*	Not applic. (0.821)
273 [C17]	Stenting for Acute MI	Balloon Angioplasty	Revascularisation at 30 days	3/2	0.33 *(0.14,0.82)*	0.63 *(0.42, 0.94)*	Not Applic. (0.204)
26 [P38]	Minimally invasive left internal thoracic artery bypass	Percutaneous coronary artery stenting	Repeat Revascularisation	4/2	0.26 *(0.15, 0.47)*	0.25 *(0.06, 1.04)*	Not Applic. (0.930)
56 [P24]	PCI*	CABG surgery*	Stroke	6/1	0.36 *(0.11, 1.23)*	0.20 *(0.01, 4.78)*	Not Applic. (0.725)
94 [P18]	LMWH for STEMI	UFH	30 day Reinfarction	1/1	0.74 *(0.45, 1.23)*	1.29 *(2.45, 0.68)*	Not Applic. (0.183)
319 [P21]	SES	PES	myocardial infarction	7/3	0.87 *(0.61, 1.22)*	0.73 *(0.47, 1.16)*	Not Applic. (0.572)

6.2.1.1.1.2: Non-Fatal Endpoints for Europe versus Asia Measured as Odds Ratios.

Only one meta-analysis reported non-fatal endpoints using OR. The details for this meta-analysis are presented in Table 6.22.

Table 6.22: The Meta-analysis with Non-Fatal Endpoints for Europe vs. Asia. Effectiveness Measured by OR.

Article ref. no.	**Intervention**	**Control**	**Endpoint**	**No. of trials (Europe/ Asia)**	**OR estimate Europe *(95% CI)***	**OR estimate Asia *(95% CI)***	**Best Region (p-value of between-continent difference)**
27 [P156]	Amiodarone for prevention of AF	Placebo, Metoprolol or routine treatment	Prevention of Post cardiac surgery AF	5/1	0.47 *(0.10, 2.12)*	0.27 *(0.01, 7.34)*	Asia (0.767)

6.2.1.1.1.3: Non-Fatal Endpoints for Europe versus Asia Measured as Mean Difference.

Table 6.23 provides the details of the seven meta-analyses that contained non-fatal data and reported treatment effect using MD.

Table 6.23: Meta-analyses with Non-Fatal Endpoints for Europe vs. Asia. Effectiveness Measured by MD.

Article ref. no.	Intervention	Control	Endpoint	No. of trials (Europe /Asia)	Mean difference estimate Europe *(95% CI)*	Mean Difference estimate Asia *(95% CI)*	Best Region (p-value of between-continental difference)
107 [P95]	Beta-Blocker Therapy for Heart Failure	Placebo	Mean Change in QoL score	1/1	-0.28 *(-1.69, 1.13)*	-0.36 *(-4.19, 3.47)*	Asia (0.853)
6 [P141]	ACE inhibition	Placebo, conventional treatment or no treatment	Diastolic volume	6/1	-2.37 *(-15.95, 11.21)*	-18 *(-88.43, 52.53)*	Asia (0.669)
184 [P51]	Autologous bone marrow cells	Standard medical therapy	Change in EF	2/1	3.41 *(-11.33,18.17)*	13 *(4.96, 21.04)*	Europe (0.264)
4 [P116]	Bone Marrow Derived Cell Transplantation	Placebo, Optimal therapy or no treatment	Difference in mean improvement in left ventricular ejection fraction	7/5	2.10 *(-2.26, 6.45)*	8.15 *(-0.19, 16.48)*	Asia (0.2070
208 [P17]	G-CSF therapy	Placebo or Blank control	Change in LVEF	4/3	0.47 *(-7.24, 8.18)*	3.33 *(-9.17, 15.83)*	Asia (0.703)
23 [P14]	PCI	Optimal treatment	Change in EF from baseline to follow up	1/1	4 *(-1.55, 9.55)*	4 *(-7.41, 15.41)*	None (1)
207 [P146]	Bone marrow cells	Placebo or no treatment	LVEF change at follow up	5/1	2.60 *(-2.25, 7.45)*	6.7 *(-11.61, 25.01)*	Asia (0.671)

As with meta-analyses that had used RR, one of the meta-analyses using MD did not specifically identify which group was the intervention under investigation and which the control group. Therefore, the best region was not calculated for this meta-analysis. The details of this meta-analysis can be found in Table 6.24.

Table 6.24: The Meta-analysis with Non - Fatal Endpoints for Europe vs. Asia that does not identify which Group is the Intervention and which the Control. Effectiveness Measured by MD.

Article ref. no.	Intervention 1	Intervention 2	Endpoint	No. of trials (Europe/Asia)	Mean difference estimate Europe *(95% CI)*	Mean Difference estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
272 [P165]	Drugs, Behavioural therapy or Exercise for secondary prevention of Coronary artery disease	Detraining, low intensity exercise, rest or no treatment	Treatment outcome for time domain heart rate variability indices	13/3	0.43 *(-0.16, 1.01)*	0.33 *(-0.80, 1.46)*	Not Applic. (0.878)

6.2.1.1.2: Examination over all Meta-analyses for Non-Fatal Endpoints – Europe versus Asia

Table 6.25 shows the number of meta-analyses that identified their intervention and control groups favouring each continent.

Table 6.25: A Summary of the number of Meta-analyses Favouring either Europe or Asia for Non-Fatal Endpoints.

Total Number of Meta-analyses	Number of Meta-analyses with Intervention Performing Better in Europe (%)	Number of Meta-analyses with Intervention Performing Better in Asia (%)	Number of Meta-analyses with Intervention not Favouring either Continent (%)
32	8 (25%)	23 (72%)	1 (3%)

There were 23 meta-analyses where the intervention performed better in Asia than in Europe and eight where the opposite was the case. As with fatal endpoints, a binomial sign test was conducted under the null hypothesis that the effect of interventions, relative to controls, was the

same in both Europe and Asia. This indicated that the difference in the effect of interventions between Europe and Asia was statistically significant (2-sided p-value = 0.0201).

6.2.1.1.3: Global Estimates of Difference between Europe and Asia

To investigate the existence of inter-continental differences in treatment effectiveness for non-fatal endpoints further, global estimates of continental difference were produced. These reflected the weighted average difference between the treatment effect estimates of Europe and Asia over all of the meta-analyses that had identified their intervention and control groups. From these, it was found that for RR (24 meta-analyses) the mean log difference was 0.161 with 95% CIs of 0.021 and 0.301 (2-sided p-value = 0.025), while for OR (1 meta-analysis) the mean log difference was 0.548 with 95% CIs of -3.08 and 4.18 (2-sided p-value = 0.767). For endpoints calculated in terms of MD (7 meta-analyses), the mean difference over all of the data was found to be -1.34, with 95% CIs of -4.7 and 2.01 (2-sided p-value = 0.433). As can be seen, the global estimate of continental difference for RR was statistically significant at the 5% level (2-sided p-value = 0.025). This suggested that when treatment effect was measured using RR, there was a significant difference in the effect of interventions between Europe and Asia, with interventions on average performing better (relative to controls) in Asia than in Europe. The finding for both OR and MD, however, did not reach statistical significance, meaning that there was insufficient evidence to conclude that there was a difference in overall treatment effect between Europe and Asia over all of the included meta-analyses that used these treatment effect measures.

6.2.1.1.4: Impact of Overlapping Meta-analyses

The results found above may have been the result of the inclusion of trial duplications in overlapping meta-analyses (see Chapter 4). When the impact of overlapping meta-analyses on the pair-wise comparison between Europe and Asia was investigated, 20 meta-analyses were identified as overlapping (P125, C11, P130, P108, P7, P2, P25, P155, C40, P102, P5, P51, P116, P17, P14, P146, P149, P156,P15,C40) and 12 meta-analyses considered as non-overlapping (P97, C12, P94, P166, C39, P90, C4, C10, C41, P101, P95, P141). Of the 20 overlapping meta-analyses, there were four pairs of meta-analyses that shared a number of the same trials (P155 and P130, P108 and P102, C11 and P125, P149 and P156). In addition, there was one group of three meta-analyses which shared a proportion of the same trials (P15, P7, and P14), one group of four meta-analyses where this was the case (P25, P2, P5, P17) and one group of five (P146, C40, P51, P116, P19). More importantly, the percentage of meta-analyses where the intervention performed better in Asia than Europe was 65% for overlapping meta-analyses but 83% for non-overlapping meta-analyses. This suggests that the proportion of meta-analyses favouring Asia over Europe had not been exaggerated as a consequence of including overlapping meta-analyses.

6.2.1.1.5: Summary of Pair-wise Comparison between Europe and Asia

In summary, when investigating within and over meta-analyses, differences appear to exist between Europe and Asia in treatment effectiveness. Indeed, there was a statistically significant difference in the number of meta-analyses whose interventions, relative to controls, performed better in Asia than in Europe. This was confirmed by the global estimate of continental difference calculated for meta-analyses that reported treatment effect using RR. This also showed a statistically significant difference in treatment effect between Europe and Asia, with interventions

appearing to perform better in Asia. Moreover, by investigating the impact of overlapping meta-analyses, we can be somewhat assured that such differences in treatment effect between Europe and Asia had not been created or exaggerated by the inclusion of overlapping meta-analyses.

6.2.1.2: North America vs. Asia

The meta-analyses on which this comparison was based were required to contain at least one North America trial and at least one Asia trial. Of the 95 relevant meta-analyses in the database, only 59 fulfilled this criterion and so could be used for this comparison.[50]

Of the 59 meta-analyses, 25 had at least one trial in each continent with non-fatal data that would allow for separate estimates of treatment effect to be made for both North America and Asia. Tables 6.26–6.30 show the details of these meta-analyses.

6.2.1.2.1: Meta-analyses Grouped according to Treatment Effect Measure for Non-Fatal Endpoints – North America versus Asia.

As with the Europe versus Asia non-fatal comparison, the meta-analyses that contained non-fatal data have been grouped according to the treatment effect measure they used.

6.2.1.2.1.1: Non-Fatal Endpoints for North America versus Asia Measured as Risk Ratios

Eighteen of the 25 meta-analyses that reported non-fatal endpoints did so using RR. The details of these studies are summarised in Table 6.26.

[50] Some of these meta-analyses were used in the comparison between Europe and Asia while others were different meta-analyses.

Table 6.26: Meta-analyses with Non- Fatal Endpoints for North America vs. Asia. Effectiveness Measured by RR.

Article ref. no.	Intervention	Control	Endpoint	No. of trials (North America/Asia)	RR estimate North America *(95% CI)*	RR estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
18 [C12]	Phosphodiesterase III inhibitors for HF	Placebo	Worsening heart failure	6/2	0.92 *(0.83,1.02)*	0.52 *(0.27, 0.98)*	Asia (0.080)
25 [P94]	Vasopressin for Cardiac arrest	Epinephrine	Failure of return of spontaneous circulation	1/1	0.99 *(0.71, 1.39)*	0.59 *(0.38, 0.91)*	Asia (0.065)
251 [P167]	Angiotensin Receptor Blockers for MI	Placebo	Risk of MI	1/1	1.49 *(0.06, 35.96)*	1.03 *(0.04, 24.70)*	Asia (0.774)
48 [P166]	Vitamin-mineral supplements for CAD	Placebo	Restenosis	2/1	1.00 *(0.33,3.01)*	0.56 *(0.06, 5.31)*	Asia (0.424)
96 [P90]	Adjunctive mechanical devices for Acute MI	Standard care or stenting	Distal embolisation	1 /2	0.84 *(0.37, 1.90)*	0.40 *(0.21, 0.75)*	Asia (0.160)
227 [C4]	Anti-arrhythmics after cardioversion of AF	Placebo, drugs for rate control or no treatment	AF Recurrence	6/2	0.86 *(0.82, 0.89)*	0.85 *(0.70, 1.03)*	Asia (0.930)
312 [P124]	Thiazolidinedione Therapy for Restenosis after Coronary Stent implantation	Placebo or Standard Care	Target Lesion Revascularisation	1/3	0.67 *(0.15, 2,98)*	0.28 *(0.14, 0.55)*	Asia (0.302)
33 [P97]	Thrombectomy or Embolic protection devices with PCI for Acute MI	PCI alone	Target vessel revascularisation	1 /2	5.00 *(0.00 , 21379.8)*	0.66 *(0.02, 21.39)*	Asia (0.662)
200 [P155]	ACEI or ARB for prevention of AF	Placebo, other antihypertensive therapy or treatment drug plus Irbesartan	Development of AF	1/1	0.22 *(0.02, 2.68)*	0.60 *(0.10, 3.76)*	North America (0.530)
228 [P6]	Combination of ACEI plus Angiotensin Receptor Blocker	ACEI alone	Hypotension	1/1	3.05 *(0.39, 24.15)*	8.14 *(1.04, 63.9)*	North America (0.510)

383 [C41]	NNPV for Cardiogenic Pulmonary Edema	Standard Medical Care	Endotracheal Intubation Rate	1 /4	0.71 *(0.27, 1.89)*	0.37 *(0.23, 0.60)*	Asia (0.229)
388 [P101]	Prophylactic steroids for cardiopulmonary bypass	Placebo or standard care	New onset AF	7/1	0.58 *(0.45, 0.75)*	1.25 *(0.49, 3.16)*	North America (0.121)
51 [C11]	Glycoprotein IIb/IIIa inhibitors for PCI	Placebo	30 day urgent revascularisation	9/1	0.54 *(0.44, 0.67)*	2.74 *(0.12, 63.63)*	North America (0.313)
328 [P108]	Off pump CABG surgery	On pump CABG surgery	Stroke	3/1	1.34 *(0.36, 5.05)*	0.35 *(0.01, 8.56)*	Asia (0.449)
226 [P125]	Glycoprotein IIb/IIIa inhibitors for PCI	Placebo or standard care	MI within 30 days	7/1	0.63 *(0.56, 0.72)*	0.13 *(0.01, 2.38)*	Asia (0.287)
55 [P131]	Clopidogrel Plus Aspirin for the Prevention of Vascular Events	Antiplatelet Monotherapy	Major Coronary Events	1/1	0.74 *(0.57, 0.95)*	0.92 *(0.87,* 0.97)	North America (0.108)
259 [P102]	Off pump CABG surgery	On pump CABG surgery	MI	4/1	1.34 *(0.47, 3.82)*	1.06 *(0.15, 7.36)*	Asia (0.836)
153 [P149]	Amiodarone prophylaxis for cardiothoracic surgery	Placebo	Post-operative AF	5/1	0.64 *(0.54, 0.77)*	0.36 *(0.18, 0.71)*	Asia (0.101)

As for fatal endpoints, the best region was not calculated for some of the meta-analyses using RR as these meta-analyses had not specifically identified the intervention or control groups. As such, it was not possible to calculate the continent in which the intervention performed best since it was not clear what constituted the intervention. This was the case in two meta-analyses and details of these can be found in Table 6.27

Table 6.27: Meta-analyses with Non - Fatal Endpoints for North America vs. Asia that do not identify which Group is the Intervention and which the Control. Effectiveness Measured by RR.

Article ref. no.	Intervention 1	Intervention 2	Endpoint	No. of trials (North America/Asia)	RR estimate North America *(95% CI)*	RR estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
56 [P24]	PCI*	CABG surgery*	Stroke	3/1	0.39 *(0.14, 1.08)*	0.20 *(0.01, 4.78)*	Not Applic. (0.692)
168 [P140]	Clopidogrel plus Aspirin for Cardiovascular Disease	Placebo plus Aspirin	MI	1/1	-0.26 *(-1.51, 0.99)*	-0.15 *(-0.62, 0.32)*	Not Applic. (0.871)

6.2.1.2.1.2: Non-Fatal Endpoints for North America versus Asia Measured as Odds Ratios

Only one meta-analysis reporting a non-fatal endpoint used OR as its treatment effect measure. The details of this meta-analysis are provided in Table 6.28.

Table 6.28: The Meta-analysis with Non-Fatal Endpoints for North America vs. Asia. Effectiveness measured by the OR.

Article ref. no.	Intervention	Control	Endpoint	No. of trials (North America/ Asia)	OR estimate North America *(95% CI)*	OR estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
27 [P156]	Amiodarone for prevention of AF	Metoprolol, Placebo or Routine Treatment	Prevention of Post Cardiac Surgery AF	7/1	-0.67 *(-1.54, 0.19)*	-1.31 *(-4.61, 1.99)*	Asia (0.715)

6.2.1.2.1.3: Non-Fatal Endpoints for North America versus Asia Measured as Mean Differences.

Only three meta-analyses used the MD as their treatment effect measure. Table 6.29 shows the details of these meta-analyses.

Table 6.29: Meta-analyses with Non-Fatal Endpoints for North America vs. Asia. Effectiveness measured by the Mean Difference.

Article ref. no.	Intervention	Control	Endpoint	No. of trials (North America/ Asia)	Mean difference estimate North America *(95% CI)*	Mean Difference estimate Asia *(95% CI)*	Best Region (p-value of between-continent difference)
6 [P141]	ACE inhibition	Placebo, conventional treatment or no treatment	Diastolic volume	1/1	-13 *(-52.93, 26.93)*	-1.8 *(-88.43, 52.43)*	Asia (0.904)
107 [P95]	Beta-Blocker Therapy for Heart Failure	Placebo	Mean Change in QoL score	5/1	-0.05 *(-0.46, 0.36)*	-0.36 *(-4.19, 3.47)*	Asia (0.875)
23 [P14]	PCI	Optimal treatment	Change in EF from baseline to follow up	1/1	1.2 *(-11.85, 14.25)*	4 *(-7.41, 15.41)*	Asia (0.752)

One meta-analysis that reported non-fatal endpoints in terms of MD did not specifically identify its intervention group and control group. The details of this meta-analysis are provided in Table 6.30.

Table 6.30: The Meta-analysis with Non - Fatal Endpoints for North America vs. Asia that does not identify which Group is the Intervention and which the Control. Effectiveness Measured by MD.

Article ref. no.	Intervention 1	Intervention 2	Endpoint	No. of trials (North America/Asia)	MD estimate North America *(95% CI)*	MD estimate Asia *(95% CI)*	Best Region (p-value of difference)
272 [P165]	Drugs, Behavioural therapy or Exercise for secondary prevention of Coronary artery disease	Detraining, low intensity exercise, rest or no treatment	Treatment outcome for time domain heart rate variability indices	5/3	0.56 *(-0.31, 1.42)*	0.33 *(-0.80, 1.46)*	Not Applic. (0.750)

6.2.1.2.2: Examination over all Meta-analyses for Non-Fatal Endpoints – North America versus Asia.

Table 6.31 shows the number of meta-analyses, which had identified their intervention and control groups, which favoured each continent.

Table 6.31: A Summary of the Number of Meta-analyses favouring either North America or Asia for Non-Fatal Endpoints.

Total Number of Meta-analyses	Number of Meta-analyses with Interventions favouring North America (%)	Number of Meta-analyses with Interventions favouring Asia (%)
22	5 (23%)	17 (77%)

This table shows that there were 17 meta-analyses where the intervention performed better in Asia than in North America and 5 where the intervention performed better in North America than Asia. As before, to examine the significance of this, a binomial sign test was conducted under the

null hypothesis that the effectiveness of interventions, relative to the control, was the same in both North America and Asia. This indicated that there was a statistically significant difference in the effectiveness of interventions between North America and Asia (2-sided p value = 0.017).

6.2.1.2.3: Global Estimates of Difference between North America and Asia.

Global estimates of continental difference were calculated to see the average weighted difference in the effect of interventions between North America and Asia. For RR (18 meta-analyses), the mean log difference over all of the meta-analyses was 0.094 with 95% CIs of -0.094 and 0.182 (2-sided p-value = 0.534) while for OR the mean difference was 0.231 with 95% CIS of -1.193 and 1.655 (2-sided p-value = 0.715). For the meta-analyses using MD (3 meta-analyses) the mean difference overall was 0.174 with 95% CIs of -3.580 and 3.928 (2-sided p-value = 0.928). These results showed that there was insufficient evidence for a difference in overall treatment effect between North America and Asia over all of the included meta-analyses (which had identified their intervention and control groups), as none of these global estimates were statistically significant at the 5% level.

6.2.1.2.4: Impact of Overlapping Meta-analyses

For this comparison, it was found that there were six overlapping meta-analyses. These were C11 and P125, P102 and P108, P149, and P156. There were also 16 non-overlapping meta-analyses (C12, P94, P167, P131 P166, P90, C4, P124, P97, P155, P6, C41, P101, P141, P95, and P14). The percentage of meta-analyses identified as overlapping, where the intervention performed better in Asia than North America, was 83% but was 75% for meta-analyses that were identified as non-overlapping. It was concluded, therefore, that the proportion of meta-analyses where the

interventions performed better in Asia over North America may have been inflated as a consequence of including duplicate trial results from overlapping meta-analyses.

6.2.1.2.5: Summary of Pair-wise Comparison between North America and Asia

In summary, this investigation within and over meta-analyses found some evidence to suggest that differences between North America and Asia in treatment effectiveness existed. This is because there was a statistically significant difference in the number of meta-analyses whose intervention, relative to controls, performed better in Asia than in North America. However, this result was not confirmed by the global estimates of continental difference, as none of these were statistically significant at the 5% level. Moreover, from the investigation concerning overlapping meta-analyses it was found that of the meta-analyses considered to be overlapping, 83% (5 out of 6 meta-analyses) favoured Asia over North America, while for non-overlapping meta-analyses only 75% (12 out of 16 meta-analyses) favoured Asia. This suggests that, with regard to non-fatal endpoints, the large number of meta-analyses that had favoured Asia over North America (77% or 17 out of 22 meta-analyses) may have been inflated as a consequence of including overlapping meta-analyses.

6.2.2: A Universal Comparison of Treatment Effectiveness Differences for Non-Fatal Endpoints.

By examining the individual trials per continent that were included in the meta-analyses, the existence of inter-continental differences in treatment effectiveness for non-fatal endpoints was explored further. The effect estimate from each relevant trial was standardised using the process reported in Chapter 4 and grouped according to continent. Each continent's trial results were

plotted onto graphs to detect outliers and examine the spread of each continent's trial results. The mean treatment effect estimate and its standard error were calculated for each continent and the results compared.

As with the analysis with fatal endpoints, trial duplications were examined and removed.[51] As can be seen from Table 6.32 there were originally 692 trials: 343 from Europe, 219 from North America and 67 from Asia. Of these, 118 were duplicated trials, which were removed from the analysis leaving 511 trials to be analysed.

Table 6.32: Duplicate and Non-Duplicated Trials in each continent that had provided Non-Fatal Data.

Continent	Number of Trials in Original Analysis	Number of Duplicated Trials (%)	Number of Trials Duplicated once	Number of Trials Duplicated more than Once	Remaining Trials when Duplications Removed
Europe	343	62	30	13	281
North America	219	45	18	12	174
Asia	67	11	10	1	56

6.2.2.1: Findings for the Universal Comparison – Non-Fatal Endpoints

Figures 6.7-6.9 show the spread of each continent's individual trial results for non-fatal endpoints. Initially, the graphs were used to detect outliers in each continent but for non-fatal endpoints, no outliers were found (see Figures 6.6-6.8).

[51] As with fatal endpoints, the analysis was also conducted with trial duplications included but the results became redundant when the analysis was conducted with duplications removed. Therefore, it is not reported.

FIGURE 6.6: THE SPREAD OF EUROPEAN TRIAL TREATMENT EFFECT ESTIMATES FOR NON-FATAL ENDPOINTS

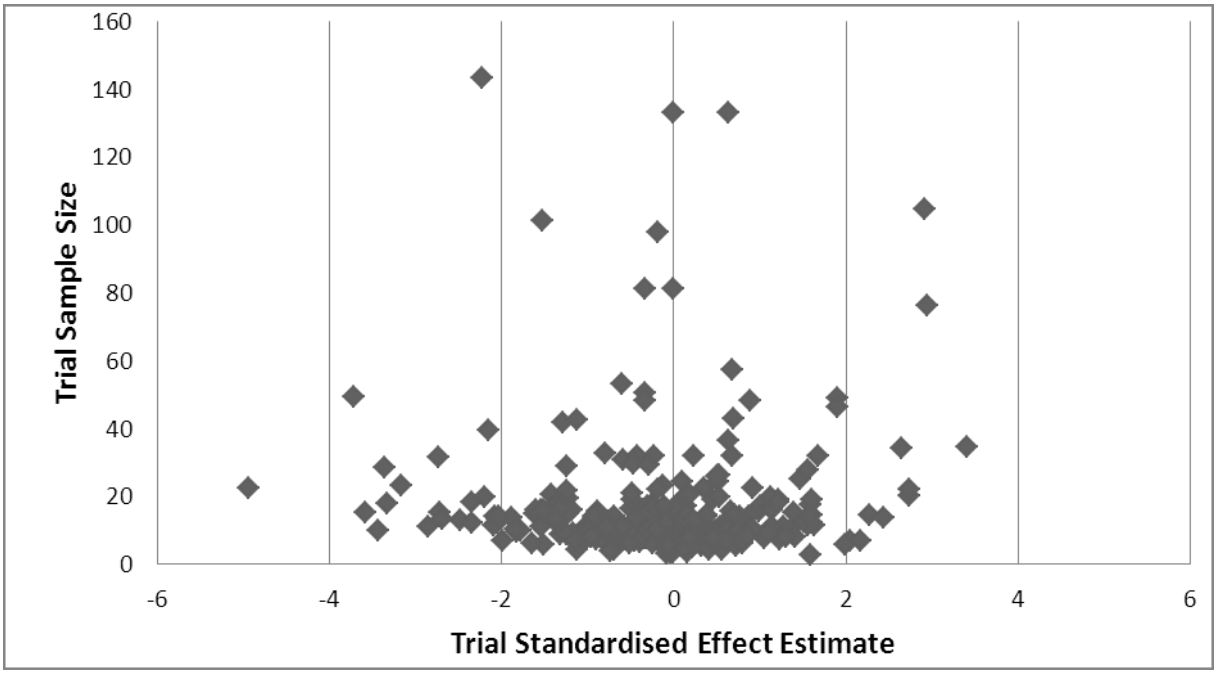


FIGURE 6.7: THE SPREAD OF NORTH AMERICAN TRIAL TREATMENT EFFECT ESTIMATES FOR NON-FATAL ENDPOINTS

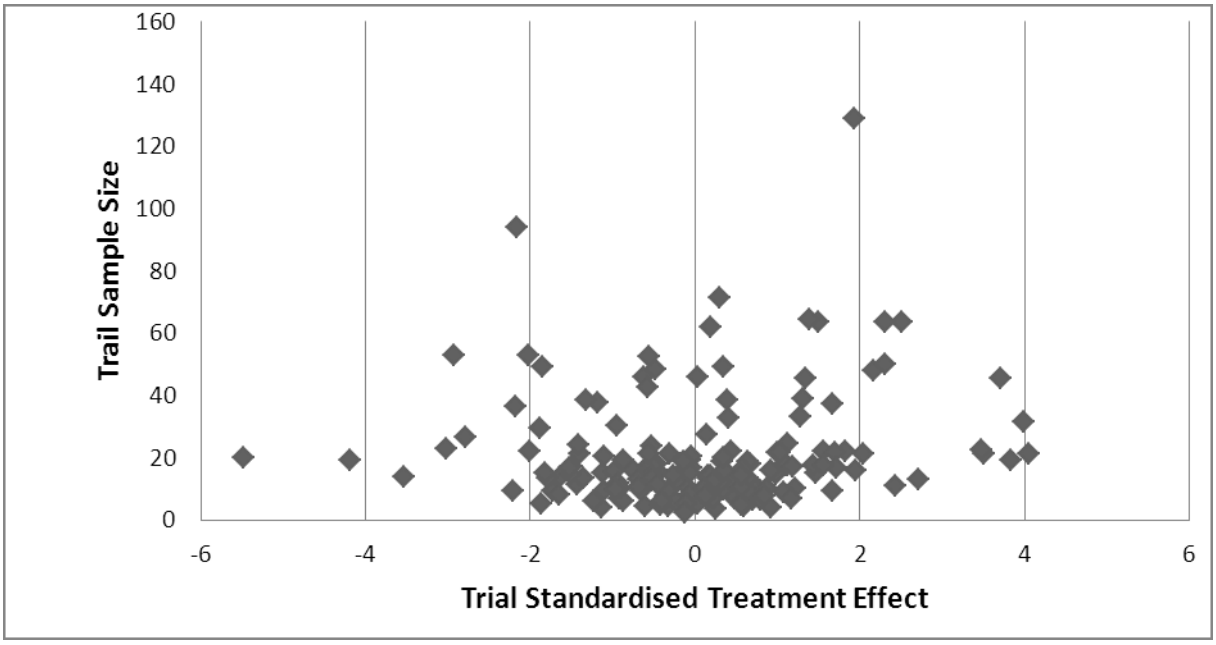
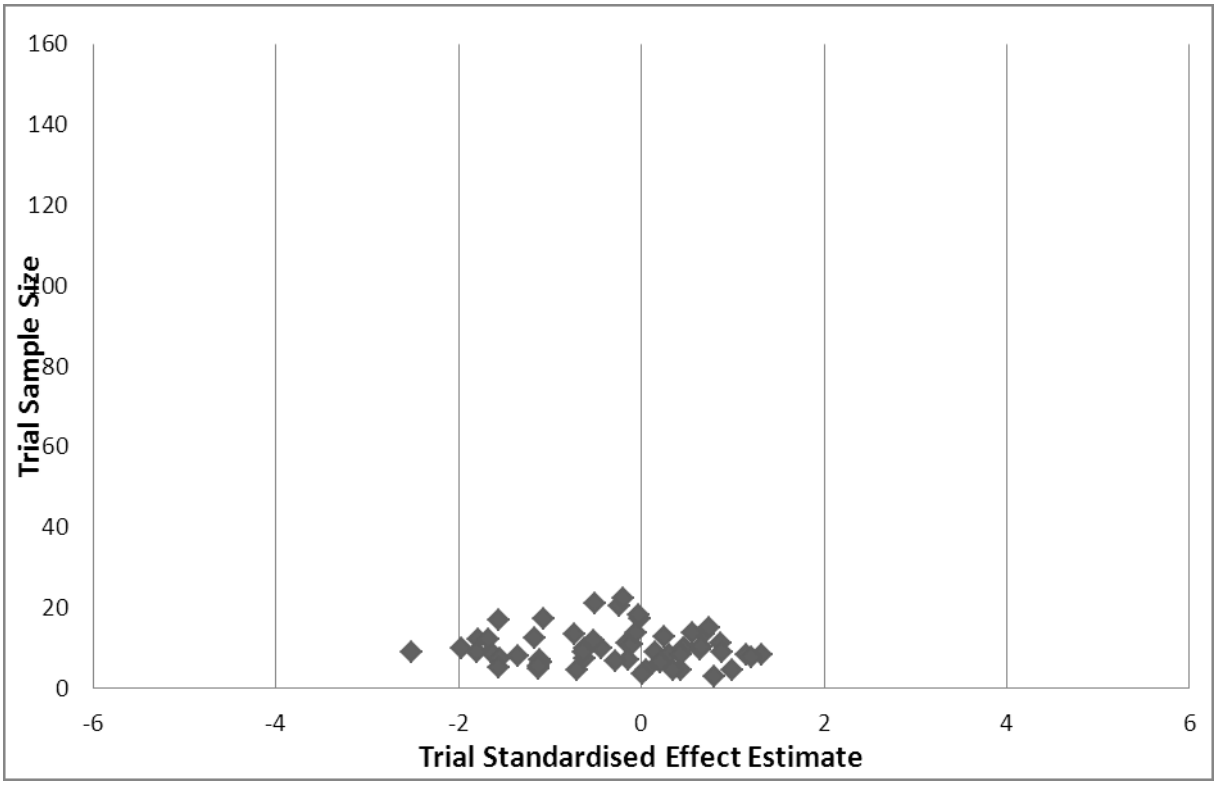


FIGURE 6.8: THE SPREAD OF ASIAN TRIAL TREATMENT EFFECT ESTIMATES FOR NON-FATAL ENDPOINTS



Figures 6.6-6.8 were then used to observe the spread of each continent's individual trial results in relation to non-fatal endpoints. This proved again to provide interesting findings regarding inter-continental differences in treatment effectiveness since it was observed, when comparing these three figures, that Asian trials produced more positive trial results compared to Europe and North America.

Table 6.33 shows the summary statistics calculated for each continent: the mean effect estimate and its standard error, each continent's z-score and its 2-sided p-value. It also provides the

number of positive trials found in each continent and the 2-sided p-value for the binomial sign test that was conducted.

Table 6.33: Summary Statistics for each Continent for Non-Fatal Endpoints.

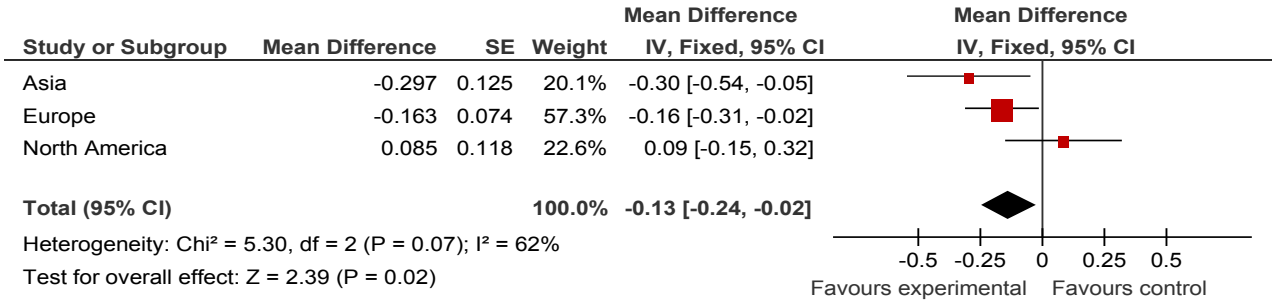
	Europe	North America	Asia
Mean Effect Estimate	-0.163	0.085	-0.297
Standard Error	0.074	0.118	0.125
Z-Score	-2.22	0.72	-2.38
P-value	0.026	0.472	0.017
Number of Trials (% of Positive Trials)	281 (55%)	174 (49%)	56 (59%)
Sign Test P-Value	0.121	0.940	0.229

The mean effect sizes between continents did differ. In North America, the mean treatment effect estimate (Mean = 0.085, SE = 0.118) favoured the control, whereas in Europe (Mean = -0.163, SE = 0.074) and Asia (Mean = -0.297, SE = 0.125) the intervention was more effective. This was confirmed by the 2-sided p-value calculated for the z-scores in Europe (2-sided p-value = 0.026) and Asia (p-value = 0.017). In addition, Asia provided the largest mean effect estimate, which suggested that interventions, relative to controls, performed better in Asia than in Europe and North America. However, this was not supported by any of the binomial signs tests that were conducted for each continent, as none was statistically significant at the 5% level.

A visual examination of Figure 6.9 provided further evidence that the effectiveness of interventions differed between the three continents investigated since the controls appeared to be more effective in North America, while the interventions appeared more effective in Europe and Asia. This was further confirmed by the I^2 statistic, which showed that the percentage of total

variation due to heterogeneity rather than chance was 62%. However, further examination of the forest plot suggested that the level of heterogeneity between the three continents was moderate as there was some degree of overlap in confidence intervals.

FIGURE 6.9; A FOREST PLOT OF THE EFFECTIVENESS OF INTERVENTIONS IN EUROPE, NORTH AMERICA AND ASIA FOR NON-FATAL ENDPOINTS.



Two independent samples t-tests were conducted between Europe and Asia and North America and Asia to confirm or refute these findings. These compared the effect of interventions, relative to controls, in Europe and Asia and North America and Asia. The t-test for Europe versus Asia revealed the effectiveness of interventions, relative to controls, did not differ significantly between Europe (Mean = -0.163, SE=0.074) and Asia (Mean = -0.297, SE = 0.125) (t (97) = 0.9225, 2-sided p-value = 0.359). This contrasted with the independent samples t-test results for North America versus Asia where interventions, relative to controls, performed better in Asia (Mean = -0.297, SE = 0.125) than in North America (Mean = 0.085, SE = 0.118) with the difference being statistically significant (t(157) = 2.222, 2-sided p-value = 0.028). It was concluded, therefore, that there were statistically significant differences in the effect of

interventions between North America and Asia for non-fatal endpoints, with interventions performing better in Asia than in North America.

6.2.3: Types of Intervention Prone to Inter-Continental Differences in Treatment Effectiveness - Non-Fatal Endpoints

As there was some evidence that differences in the clinical effectiveness of interventions existed, it was important to investigate whether any of the intervention types were prone to showing inter-continental differences in treatment effect. This was carried out using the same procedure as for fatal endpoint data (See section 6.1.3). The types of intervention were Drug, Device, Surgery, and Lifestyle. Management interventions were not included as there were no Asian trials involving these types of intervention with non-fatal data and so it was not possible to make comparisons.

6.2.3.1: Europe vs. Asia

First, the types of intervention prone to differences in treatment effect were examined over and within the meta-analyses that provided non-fatal data for the Europe versus Asia comparison. The meta-analyses were grouped according to the type of intervention investigated and placed into tables. Table 6.34 shows the meta-analyses included in the Europe versus Asia pair-wise comparison that investigated drug interventions.

Table 6.34: Meta-analyses Investigating Drugs with their Z-statistic for the Difference in Treatment Effect between Europe and Asia for Non-Fatal Endpoints.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (Europe/ Asia)	Pooled Estimate for Europe *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
226 [P125]	Glycoprotein IIb/IIIa Antagonists	RR	4/1	0.86 *(0.58, 1.28)*	0.13 *(0.01, 2.38)*	1.890 *(-1.04, 4.82)*	1.264	0.206
247 [C40]	Stem Cell Treatment	RR	5/1	1.15 *(0.71, 1.87)*	0.20 *(0.01, 3.97)*	1.753 *(-1.27, 4.78)*	1.136	0.256
51 [C11]	Platelet Glycoprotein IIb/IIIa Blockers	RR	5/1	0.56 *(0.35, 0.88)*	2.74 *(0.12, 63.63)*	-1.591 *(-4.77, 1.59)*	-0.981	0.327
188 [P25]	GCSF Therapy	RR	4/1	0.97 *(0.61, 1.53)*	2.75 *(0.24, 31.72)*	-1.045 *(-3.53, 1.44)*	-0.823	0.410
237 [P19]	Intracoronary Cell Therapy	RR	6/1	0.99 *(0.70, 1.43)*	0.36 *(0.02, 8.41)*	1.027 *(-2.15, 4.21)*	0.633	0.527
18 [C12]	Phosphodiesterase III Inhibitors	RR	3/2	1.44 *(1.00, 2.06)*	0.52 *(0.27, 0.98)*	1.022 *(0.29, 1.75)*	2.746	0.006
153 [P149]	Amiodarone prophylaxis	RR	6/1	0.70 *(0.55, 0.90)*	0.36 *(0.18, 0.71)*	0.684 *(-0.04, 1.41)*	1.836	0.066
25 [P94]	Vasopressin	RR	2/1	1.03 *(0.97, 1.11)*	0.59 *(0.38, 0.91)*	0.556 *(0.19, 0.99)*	2.488	0.013
388 [P101]	Steroid Use	RR	3/1	0.78 *(0.62, 0.98)*	1.25 *(0.49, 3.16)*	-0.467 *(-1.42, 0.49)*	-0.958	0.338
176 [C39]	Calcium Channel Antagonists	RR	2/1	1.54 *(0.58, 4.06)*	0.97 *(0.39, 2.41)*	0.462 *(-0.87, 1.80)*	0.680	0.497
200 [P155]	Angiotensin-converting Enzymes or Receptor Blockers	RR	6/1	0.93 *(0.68, 1.27)*	0.60 *(0.10, 3.76)*	0.434 *(-1.42, 2.29)*	0.458	0.647
165 [P130]	Angiotensin-converting Enzymes or Receptor Blockers	RR	6/1	0.93 *(0.86, 1.00)*	0.60 *(0.37, 0.97)*	0.432 *(-0.05, 0.92)*	1.751	0.080
4 [P2]	GCSF Therapy	RR	4/1	0.91 *(0.58, 1.41)*	1.22 *(0.08, 18.20)*	-0.300 *(-3.04, 2.44)*	-0.215	0.830

400 [P5]	Stem Cell Treatment	RR	4/2	1.00 *(0.63, 1.59)*	0/83 *(0.40, 1.74)*	0.187 *(-0.68, 1.05)*	0.423	0.672
232 [C10]	Magnesium	RR	4/1	0.79 *(0.65, 0.97)*	0.66 *(0.39, 1.10)*	0.185 *(-0.37, 0.74)*	0.656	0.512
227 [C4]	Anti-arrhythmics	RR	11/2	0.75 *(0.71, 0.79)*	0.85 *(0.70, 1.03)*	-0.124 *(-0.32, 0.07)*	-1.225	0.220
27 [P156]	Amiodarone	OR	5/1	0.47 *(0.10, 2.12)*	0.27 *(0.01, 7.34)*	0.548 (-3.09, 4.18)	0.296	0.767
6 [P141]	ACE Inhibitors	MD	6/1	-2.37 *(-15.95, 11.21)*	-18 *(-88.43, 52.53)*	15.632 *(-56.09, 87.36)*	0.427	0.669
184 [P51]	Bone Marrow Cell Therapy	MD	2/1	3.41 *(-11.33, 18.17)*	13 *(4.96, 21.04)*	-9.583 *(-26.38, 7.22)*	-1.118	0.264
4 [P116]	Bone Marrow Cell Therapy	MD	7/5	2.10 *(-2.26, 6.45)*	8.15 *(-0.19,16.48)*	-6.050 *(-15.45, 3.35)*	-1.261	0.207
207 [P146]	Bone Marrow Cell Therapy	MD	5/1	2.60 *(-2.25, 7.45)*	6.7 *(-11.61, 25.01)*	-4.100 *(-23.04, 14.84)*	-0.424	0.671
208 [P17]	GCSF Therapy	MD	4/3	0.47 *(-7.24, 8.18)*	3.33 *(-9.17, 15.83)*	-2.857 *(-17.55, 11.83)*	-0.381	0.703
107 [P95]	Beta-Blocker Therapy	MD	1/1	-0.28 *(-1.69, 1.13)*	-0.36 *(-4.19, 3.47)*	0.080 *(-4.00, 4.16)*	0.038	0.853

As can be seen, two meta-analyses investigating drugs found a statistically significant difference in treatment effect between Europe and Asia. The first meta-analysis (C12 (18)) compared Phosphodiesterase III Inhibitors (the intervention) to a placebo (the control), with Phosphodiesterase III Inhibitors performing better, with respect to a worsening in heart failure, in Asia than in Europe. The 2-sided p-value for this Z-statistic was 0.006. The second meta-analysis was P94 (25), in which Vasopressin (the intervention) was compared with Epinephrine (the control). In this meta-analysis, Vasopressin was found to perform better, with respect to the

failure of return of spontaneous circulation, in Asia than in Europe. The p-value for this intervention was 0.013.

However, when looking over all of the meta-analyses investigating drug interventions, no clear pattern of statistically significant differences was found to suggest that this type of intervention was prone to inter-continental differences in treatment effect.

Table 6.35 provides the details of the meta-analyses that investigated medical devices.

Table 6.35: Meta-analyses Investigating Medical Devices with their Z-statistic for the Difference in Treatment Effect between Europe and Asia for Non-Fatal Endpoints.

Meta-analysis Ref. No.	**Intervention**	**Treatment Effect Measure**	**No. of Trials (Europe/ Asia)**	**Pooled Estimate for Europe *(95% CIs)***	**Pooled Estimate for Asia *(95% CIs)***	**Difference in Treatment Effect *(95% CIs)***	**Z-Statistic**	**2-sided p-value of difference**
96 [P90]	Adjunctive Mechanical Devices	RR	4/2	0.65 *(0.37, 1.13)*	0.40 *(0.21, 0.75)*	0.485 *(-0.35, 1.32)*	1.138	0.255
383 [C41]	Non-Invasive Positive Pressure Ventilation	RR	4/1	0.86 *(0.58, 1.28)*	0.13 *(0.01, 2.38)*	-0.600 *(-0.13, 1.33)*	1.619	0.206
33 [P97]	Thrombectomy or Embolic Protection Devices	RR	6/2	0.94 *(0.30, 2.95)*	0.66 *(0.02, 21.39)*	0.345 *(-3.32, 4.01)*	0.185	0.853

As can be seen, none of the meta-analyses investigating medical devices found a statistically significant difference in treatment effect between Europe and Asia. Therefore, no clear pattern of

statistically significant differences was found to suggest that medical devices were prone to differences in treatment effect.

Table 6.36 provides the details of the meta-analyses that investigated surgical interventions.

Table 6.36: Meta-analyses Investigating Surgical Interventions with their Z-statistic for the Difference in Treatment Effect between Europe and Asia for Non-Fatal Endpoints.

Meta-analysis Ref. No.	**Intervention**	**Treatment Effect Measure**	**No. of Trials (Europe/ Asia)**	**Pooled Estimate for Europe** *(95% CIs)*	**Pooled Estimate for Asia** *(95% CIs)*	**Difference in Treatment Effect** *(95% CIs)*	**Z-Statistic**	**2-sided p-value of difference**
191 [P15]	PCI	RR	2/1	1.72 *(0.52, 5.71)*	0.38 *(0.11, 1.37)*	1.512 *(-0.24, 3.27*	1.687	0.092
259 [P102]	Off-Pump Coronary Artery Bypass Surgery	RR	15/1	0.74 *(0.46, 1.19)*	1.06 *(0.15, 7.36)*	-0.364 *(-2.36, 1.63)*	-0.357	0.721
2 [P7]	PCI	RR	5/1	0.49 *(0.31, 0.77)*	0.38 *(0.11, 1.37)*	0.253 *(-1.10, 1.61)*	0.366	0.715
328 [P108]	Off-Pump Coronary Artery Bypass Surgery	RR	9/1	0.41 *(0.18, 0.95)*	0.35 *(0.01, 8.56)*	0.153 *(-3.14, 3.45)*	0.091	0.928
23 [P14]	Late PCI	MD	1/1	4 *(-1.55, 9.55)*	4 *(-7.41, 15.41)*	0.000 *(-12.69, 12.69)*	0.00	1

Again, it can be seen that no meta-analyses found the difference in treatment effect between Europe and Asia to be statistically significant at the 5% level. As such, no clear pattern of statistically significant differences was found to suggest that this type of intervention was prone to inter-continental differences in treatment effect.

Table 6.37 provides the details of the meta-analysis that investigated a lifestyle intervention.

Table 6.37: Meta-analyses Investigating Lifestyle Interventions with their Z-statistic for the Difference in Treatment Effect between Europe and Asia for Non-Fatal Endpoints.

Meta-analysis Ref. No.	**Intervention**	**Treatment Effect Measure**	**No. of Trials (Europe/ Asia)**	**Pooled Estimate for Europe** ***(95% CIs)***	**Pooled Estimate for Asia** ***(95% CIs)***	**Difference in Treatment Effect** ***(95% CIs)***	**Z-Statistic**	**2-sided p-value of difference**
48 [P166]	Vitamin-mineral Supplementation	RR	2/1	1.05 *(0.43, 2.61)*	0.56 *(0.06, 5.31)*	0.633 *(-1.79, 3.06)*	0.512	0.609

It shows that there was no statistically significant difference between Europe and Asia in treatment effect for vitamin mineral supplementation when non-fatal endpoints were considered. As such, there was no evidence that this intervention was prone to between-continental differences in treatment effect.

As no group of interventions was found to show a clear pattern of statistically significant differences, the types of intervention prone to differences between Europe and Asia could not be identified. Furthermore, as there were 29 meta-analyses overall that provided non-fatal data, the finding that there were two meta-analyses among these that were statistically significant at the 5% level is not more than would be expected by random chance. Therefore, there was insufficient evidence to conclude which types of intervention were likely to show differences in treatment effect between Europe and Asia for non-fatal endpoints.

To investigate further, the trials from each meta-analysis were analysed with independent samples t-tests conducted for each type of intervention: Drug, Device, Surgery and Lifestyle. Table 6.38 shows the number of trials from each continent that investigated each intervention type and provided non-fatal data. It also provides the mean effect estimate and its standard error per continent for each intervention type and the 2-sided p-value for the t-tests that were conducted.

Table 6.38: Summary Statistics for Intervention Types for Non-Fatal Endpoints – Europe versus Asia.

Intervention Type	Number of European Trials	European Mean Effect Estimate	Standard Error	Number of Asian Trials	Asian Mean Effect Estimate	Standard Error	T-Test p-value
Drug	122	-0.119	0.119	26	-0.287	0.231	0.523
Device	38	-0.344	0.205	15	-0.433	0.162	0.735
Surgery	48	-0.023	0.120	12	-0.199	0.225	0.499
Lifestyle	38	-0.116	0.154	3	-0.093	0.332	0.956

The table shows that the number of trials investigating each intervention type differed between continents, as did the mean effect estimate for each continent. The 2-sided p-values for the t-tests also show that the difference in treatment effect between Europe and Asia was not statistically significant at the 5% level for any of the intervention types. The results of the t-tests for each intervention type are discussed further below.

6.2.3.1.1: Drug Interventions

The mean difference between continents for drug therapies was 0.168, showing a small effect size (d = 0.21), with 95% CIs of -0.358 to 0694. An independent samples t-test revealed that

there was not a statistically significant difference in the effectiveness of drug therapies, relative to controls, between Europe (Mean = -0.119, SE= 0.119) and Asia (Mean = -0.287, SE = 0.231) (t (39) = 0.6465, 2-sided p-value = 0.523). Therefore, there was insufficient evidence for the existence of a difference in the effect of drug interventions between Europe and Asia.

6.2.3.1.2: Device Interventions

The mean difference for medical devices was found to be -0.089, representing a small effect size (*d*=0.10), with 95% CIs of -0.436 to 0.5614. An independent samples t-test revealed that there was not a statistically significant difference in the effect of medical devices between Europe (Mean = -0.344, SE=0.205) and Asia (Mean = -0.433, SE = 0.162), (*t* (48) = 0.3406, 2-sided p-value =0.735). As such, it was again concluded that there was not sufficient evidence for the existence of a difference in the effect of medical devices between Europe and Asia.

6.2.3.1.3: Surgical Interventions

The mean difference between continents for surgical interventions was 0.176, which is a small-medium effect size (*d*=0.33), with 95% CIs of -0.362 to 0.714. An independent samples t-test revealed that there was not a statistically significant difference in the effect of surgical interventions, relative to controls, between Europe (Mean= -0.023, SE = 0.120) and Asia (Mean = -0.199, SE = 0.225) (*t*(17)= 0.6902, 2-sided p-value =0.499). Therefore, there was a lack of evidence for the existence of a difference in the effect of surgical interventions, relative to controls, between Europe and Asia.

6.2.3.1.4: Lifestyle Interventions

The mean difference between continents for lifestyle interventions was found to be -0.023, a small effect size (d=0.13), with 95% CIs of -1.598 to 1.552. An independent samples t-test revealed that there was not a statistically significant difference in the effect of lifestyle interventions, relative to controls, between Europe (Mean = -0.116, SE = 0.154) and Asia (Mean = -0.093, SE = 0.332), (t(2) = 0.0628, 2-sided p=0.956). Therefore, there was insufficient evidence for the existence of a difference in the effectiveness of lifestyle interventions between these two continents.

6.2.3.2: North America vs. Asia

As with the previous comparison, the types of intervention prone to between-continental differences in treatment effect were first examined over and within meta-analyses. Table 6.39 shows the meta-analyses that investigated drug therapies and had non-fatal data for both North America and Asia.

Table 6.39: Meta-analyses Investigating Drugs with their Z-statistic for the Difference in Treatment Effectiveness between North America and Asia for Non-Fatal Data.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (N. Amer./ Asia)	Pooled Estimate for North America *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
51 [C11]	Glycoprotein IIb/IIIa Antagonists	RR	9/1	0.54 *(0.44, 0.67)*	2.74 *(0.12, 63,63)*	-1.6224 *(-4.77,1.53)*	-1.009	0.313
226 [P125]	Glycoprotein IIb/IIIa Antagonists	RR	7/1	0.63 *(0.56, 0.72)*	0.13 *(0.01, 2.38)*	1.5784 *(-1.33, 4.48)*	1.064	0.287
200 [P155]	ACEI's or ARB's	RR	1/1	0.22 *(0.02, 2.68)*	0.60 *(0.10, 3.76)*	-0.9859 *(-4.07, 2.09)*	-0.627	0.530
228 [P6]	ACEI's plus ARB	RR	1/1	3.05 *(0.39, 24.15)*	8.14 *(1.04, 63.9)*	-0.9805 *(-3.90, 1.94)*	-0.658	0.510
312 [P124]	Thiazolidinedione Therapy	RR	1/3	0.67 *(0.15, 2.98)*	0.28 *(0.14, 0.55)*	0.8654 *(-0.78, 2.51)*	1.033	0.302
388 [P101]	Steroid Use	RR	7/1	0.58 *(0.45, 0.75)*	1.25 *(0.49, 3.16)*	-0.7625 *(-1.73,0.20)*	-1.549	0.121
153 [P149]	Amiodarone prophylaxis	RR	5/1	0.64 *(0.54, 0.77)*	0.36 *(0.18, 0.71)*	0.5936 *(-0.12, 1.30)*	1.638	0.101
18 [C12]	Phosphodiesterase III Inhibitors	RR	6/2	0.92 *(0.83, 1.02)*	0.52 *(0.27, 0.98)*	0.5746 *(-0.07,1.22)*	1.752	0.080
25 [P94]	Vasopressin	RR	1/1	0.99 *(0.71, 1.39)*	0.59 *(0.38, 0.91)*	0.5165 *(-0.03, 1.06)*	1.848	0.065
251 [P167]	ARB	RR	1/1	1.49 *(0.06, 35.96)*	1.03 *(0.04, 24.70)*	0.3662 *(-4.13, 4.86)*	0.160	0.774
55 [P131]	Clopidogrel plus Aspirin	RR	1/1	0.74 *(0.57, 0.95)*	0.92 *(0.87, 0.97)*	-0.2175 *(-0.48, 0.05)*	-1.606	0.108
227 [C4]	Anti-arrhythmics	RR	6/2	0.86 *(0.82, 0.89)*	0.85 *(0.70, 1.03)*	0.0087 *(-0.19, 0.20)*	0.087	0.930
27 [P156]	Amiodarone	OR	7/1	0.51 *(0.21, 1.21)*	0.27 *(0.01, 7.34)*	0.63(-2.78, 4.05)	0.365	0.715
6 [P141]	ACE Inhibitors	MD	1/1	-13 *(-52.93, 26.93)*	-1.8 *(-88.43, 52.43)*	5 *(-75.96, 85.96)*	0.121	0.904
107 [P95]	Beta-blocker Therapy	MD	5/1	-0.05 *(-0.46, 0.36)*	-0.36 *(-4.19, 3.47)*	0.310 *(-3.54, 4.16)*	0.158	0.875

As can be seen, none of the meta-analyses investigating drugs found a statistically significant difference in treatment effect. Therefore, no clear pattern of statistically significant differences was found to suggest that drug therapies were prone to between-continental differences in treatment effect.

Table 6.40 shows the meta-analyses that investigated medical devices.

Table 6.40: Meta-analyses Investigating Medical Devices with their Z-statistic for the Difference in Treatment Effectiveness between North America and Asia for Non-Fatal Data.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (N. Amer./ Asia)	Pooled Estimate for North America *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
33 [P97]	Thrombectomy or Embolic Protection Devices	RR	1/2	5.00 *(0.00, 21379.8)*	0.66 *(0.02, 21.39)*	2.0211 *(-7.03, 11.08)*	0.437	0.662
96 [P90]	Adjunctive Mechanical Devices	RR	1/2	0.84 *(0.37, 1.90)*	0.40 *(0.21, 0.75)*	0.7379 *(-0.29, 1.77)*	1.406	0.160
383 [C41]	Non-Invasive Positive Pressure Ventilation	RR	1/4	0.71 *(0.27, 1.89)*	0.37 *(0.23, 0.60)*	0.6649 *(-0.42, 1.75)*	1.202	0.229

As Table 6.40 shows, none of the meta-analyses that investigated medical devices found the difference in treatment effect between North America and Asia to be statistically significant at the 5% level. As such, no clear pattern was identified to show these types of intervention as being prone to between-continental differences.

Table 6.41 shows the meta-analyses that investigated surgical interventions.

Table 6.41: Meta-analysis Investigating Surgical Interventions with the Z-statistic for the Difference in Treatment Effectiveness between North America and Asia for Non-Fatal Data.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (N. Amer./ Asia)	Pooled Estimate for North America *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
328 [P108]	Off-pump Surgery	RR	3/1	1.34 *(0.36, 5.05)*	0.35 *(0.01, 8.56)*	1.3323 *(-2.12, 4.78)*	0.757	0.449
259 [P102]	Off-Pump Surgery	RR	4/1	1.34 *(0.47, 3.82)*	1.06 *(0.15, 7.36)*	0.2323 *(-1.97, 2.43)*	0.207	0.836
23 [P14]	Late PCI	MD	1/1	1.2 *(-11.85, 14.25)*	4 *(-7.41, 15.41)*	-2.800 *(-20.13, 14.53)*	-0.317	0.752

Again, it was found that the difference in treatment effect was not statistically significant in any of the meta-analyses. As such, no clear pattern of statistically significant differences was found to suggest that surgical interventions were prone to differences in treatment effect between North America and Asia.

Table 6.42 shows the meta-analysis that investigated a lifestyle intervention.

Table 6.42: Meta-analysis Investigating a Lifestyle Intervention with its Z-statistic for the Difference in Treatment Effectiveness between North America and Asia for Non-Fatal Data.

Meta-analysis Ref. No.	Intervention	Treatment Effect Measure	No. of Trials (N. Amer./ Asia)	Pooled Estimate for North America *(95% CIs)*	Pooled Estimate for Asia *(95% CIs)*	Difference in Treatment Effect *(95% CIs)*	Z-Statistic	2-sided p-value of difference
48 [P166]	Vitamin-Mineral Supplementation	RR	2/1	1.00 *(0.33, 3.01)*	0.56 *(0.06, 5.31)*	0.5804 *(-1.92, 3.08)*	0.454	0.424

As can be seen, the difference in treatment effect between North America and Asia for Vitamin-mineral supplementation was not statistically significant at the 5% level. Therefore, there was no evidence that this lifestyle intervention was prone to between-continental differences in treatment effect.

As no group of interventions could be identified as showing a clear pattern of statistically significant differences in treatment effect, the types of intervention prone to differences in treatment effect between North America and Asia could not be identified for non-fatal endpoints.

To explore further, each continent's trial results were examined with independent samples t-tests conducted for each type of intervention: Drug, Device, Surgery and Lifestyle. Table 6.43 shows the number of trials from each continent that investigated each intervention type and provided non-fatal data. It also provides the mean effect estimate and its standard error per continent for each intervention type and the 2-sided p-value for the t-tests that were conducted.

Table 6.43: Summary Statistics for Intervention Types for Non-Fatal Endpoints –North America versus Asia

Intervention Type	Number of North American Trials	North America Mean Effect Estimate	Standard Error	Number of Asian Trials	Asian Mean Effect Estimate	Standard Error	T-Test p-value
Drug	90	-0.015	0.171	26	-0.287	0.231	0.348
Device	13	0.691	0.640	15	-0.433	0.162	0.113
Surgery	9	0.229	0.259	12	-0.199	0.225	0.229
Lifestyle	26	0.385	0.221	3	-0.093	0.332	0.297

The number of trials investigating each intervention type differed between continents, as did the mean effect estimate for each continent. The 2-sided p-values for the t-tests show that the difference in treatment effect between North America and Asia was not statistically significant at the 5% level for any of the intervention types. The results of the t-tests, for each intervention type, are discussed further below.

6.2.3.2.1: Drug Interventions

The mean difference between continents for drug therapies was 0.272, which is a large effect size (d = 3.21), with 95% CIs of -0.304 to 0.848. An independent samples t-test showed that there was no statistically significant difference in the effect of drug interventions, relative to controls, between North America (Mean = -0.015, SE=0.171) and Asia (Mean = -0.287, SE = 0.231): $t(55)$ = 0.9464, 2-sided p-value=0.348. Therefore, there was insufficient evidence for the existence of a difference in treatment effect for drug interventions between North America and Asia.

6.2.3.2.2: Device Interventions

The mean difference for medical devices was found to be 1.124, showing a large effect size (d = 10.16), with 95% CIs of -0.302 to 2.550. An independent samples t-test found this difference between North America (Mean = 0.691, SE = 0.640) and Asia (Mean = -0.433, SE=0.162) to be non-significant (t (13.) =1.703, 2-sided p-value =0.113). Thus, it was concluded that there was insufficient evidence for a difference in treatment effect for medical devices between North America and Asia.

6.2.3.2.3: Surgical Interventions.

The mean difference between continents for surgical interventions was 0.428, which is a large effect size (d = 5.21), with 95% CIs of -0.296 to 1.152. An independent samples t-test revealed that there was no statistically significant difference in the effect of surgical interventions, relative to controls, between North America (Mean = 0.229, SE = 0.259) and Asia (Mean = -0.199, SE = 0.225): t (17) =1.248, 2-sided p-value =0.229. This suggested that there was currently insufficient evidence that the effect of surgical interventions differed between North America and Asia.

6.2.3.2.4: Lifestyle Interventions

The mean difference for lifestyle interventions was 0.478, representing a large effect size (d = 4.40), with 95% CIs of -0.629 to 1.586. An independent samples t-test showed that there was not a statistically significant difference in the effect of these types of intervention between North America (Mean = 0.385, SE = 0.221) and Asia (Mean = -0.093, SE = 0.332): t (4) = 1.1985, 2-sided p-value =0.297. Again, it was concluded that there was insufficient evidence for any differences in the effect of lifestyle interventions between North America and Asia.

6.2.3.3: Summary of the Types of Intervention Prone to Inter-Continental Differences

From the investigation conducted over and within meta-analyses, the types of intervention prone to show between-continental differences in treatment effect could not be identified for either pair-wise comparison. This was because, for both Europe versus Asia and North America versus Asia, no clear pattern of statistically significant differences was identified for any type of intervention. This was supported by the investigation conducted over trials. In this, none of the t-tests

(conducted for each intervention type) showed statistically significant results for Europe versus Asia or North America versus Asia.

6.3: Conclusion.

This chapter provides some evidence to suggest that inter-continental differences in treatment effect exist when both fatal and non-fatal endpoints are considered.

For fatal endpoints, the findings suggest that there is some evidence for the existence of a difference in treatment effect between continents, with interventions performing better in Asia than in Europe and North America. Furthermore, when looking at different Asian countries from which data was available, the observations suggested that Japanese trials produced more positive trial results compared to other Asian countries. From investigating the types of intervention likely to show such differences, the results for fatal endpoints revealed that medical devices may be prone to showing inter-continental differences in treatment effect.

For non-fatal endpoints, the findings suggested that interventions performed better in Asia than in both Europe and North America, with both pair-wise comparisons finding statistically significant results. However, the results for the analyses conducted over meta-analyses between North America and Asia should be interpreted with caution, since these results may have been inflated as a consequence of including overlapping meta-analyses. Furthermore, for both the North America versus Asia and the Europe versus Asia comparisons, the types of intervention prone to differences in treatment effect could not be identified.

To conclude, this analysis has shown the existence of some inter-continental differences in treatment effectiveness; in particular, differences exist with regard to fatal endpoints. Since this result could not have been a consequence of the endpoint's subjectivity, as is the case for non-fatal endpoints, it needs to be investigated further. It is important to try to discover an explanation for this as it is possible that this result has the potential to impact on the extrapolation of clinical trial results between continents.

As the results for fatal endpoints for Japan were unexpected, especially the finding that Japan produced a high proportion of positive trial results, it was important to investigate further in order to determine if there was a bias among Japanese trials that led to the true effect of the interventions being overestimated. Investigation of the influence of trial quality has become a standard procedure in meta-analytic approaches since the conclusions they draw from data cannot be trusted if the raw material is unsound (Juni, Altman & Egger, 2001). As was shown in Figure 6.5, the Japanese trials were small. According to Egger et al (2003), small-scale trials are, in most cases, conducted and analysed with less methodological rigour than larger trials, resulting in an exaggeration of treatment effect. This points to trial quality as a potentially significant factor in producing the fatal endpoint findings. In other words, if all the Japanese trials suffered from low methodological rigour on account of their small size, they would be more likely to produce the positive trial results that were found. An exploratory study of the quality of the Japanese trials was therefore undertaken.

CHAPTER 7.
TRIAL QUALITY INVESTIGATION.

Trial quality was investigated as a potential explanation for the high proportion of positive Japanese trial results for fatal endpoints reported in the previous chapter (Section 6.3.1). Evidence suggests that trial quality is closely related to treatment effect estimates in that it can exaggerate the effect of treatment (Moher et al. 1998). To explore this, the Japanese trials were compared to a sample of trials from Europe and North America. This was done by two assessors using a modified Jadad Scale (Jadad et al. 1996). The Jadad scale is a validated scale for assessing the methodological quality of RCTs. However, to achieve a broader viewing of the quality of these trials, additional questions were included.

This chapter is organised into four main sections. First, it provides the background to the investigation by presenting and discussing other research conducted on the impact of trial quality on treatment effect estimates. Section two provides a thorough description of the methods used in this study and explains how any discrepancies in ratings between investigators were resolved. The results of the investigation are provided in section three. The chapter concludes by summarising the findings of the trial quality investigation and describing their implications.

7.1: The Impact of Trial Quality on Treatment Effect Estimates

Quality is a complicated concept and is not easy to define in relation to clinical trials (Jadad, 1998). Quality evaluation should include not only the design, conduct, and analysis of a trial but also its clinical relevance and the standard of reporting (Berlin & Rennie, 1999; Ioannidis & Lau, 1998). One of the most important dimensions of trial quality is the validity of the findings generated by the trial (Juni, Altman and Egger, 2001).Validity can be separated into two distinct types: internal and external. Internal validity is considered as the extent to which trial results are correct for the circumstances under investigation (Fletcher, Fletcher, & Wagner, 1982). It implies that any differences found between the groups being investigated can, apart from random error, be attributed to the intervention examined. External validity is the extent to which a trial's result can provide an accurate foundation for generalisations to other populations and conditions (Chow & Liu, 2004). There are two types of external validity: population and ecological validity. Population validity is the extent to which the results of a study conducted in one sample can be applied to another set of patients (Houser, 2008) while ecological validity refers to the extent to which the findings of a study can be applied to other settings (Houser, 2008). Clearly, for clinical trials, the assessment of internal validity is essential and is a prerequisite for the assessment of external validity since if the trial results are flawed, and therefore unsound, the issue of whether or not the findings apply elsewhere is irrelevant and does not need to be contemplated (Jimenez-Buedo & Miller, 2010).

In any trial, internal validity can be threatened by bias. Potential biases to a trial's internal validity fall into four main types: selection bias, performance bias, detection bias and attrition

bias. The first of these, selection bias, is the biased allocation of patients and relates to systematic differences in patient characteristics between the groups being compared. Randomisation aims to reduce, if not eliminate, selection bias by creating groups that are comparable with respect to identified or unidentified confounding factors (Altman & Bland, 1999). However, the success of randomisation relies upon two inter-connected procedures. The first of these is the generation of an allocation sequence that is unpredictable. This can be done, for example, by using computer generated random numbers, drawing lots or envelopes, using random number tables or tossing a coin. The second procedure involves ensuring that these sequences are concealed from those who are enrolling the patients and can be done with the use of central randomisation, sequentially numbered and sealed opaque envelopes or by using coded drug containers that have been prepared by an independent pharmacy. This procedure is of the utmost importance, to prevent selective enrolment of patients on prognostic factors.

The second and third types of bias are performance and detection bias. Performance bias is the unequal provision of care between groups, apart from the intervention investigated, that occurs if extra treatments are given to one group in preference to another. Detection bias is a systematic difference between groups in how outcomes are ascertained, diagnosed or verified and occurs if knowledge of patient assignment has influenced the process of outcome assessment. Both of these types of bias can be prevented with the use of blinding. Performance bias can be prevented by blinding both patients and care-givers and detection bias can be prevented by blinding those who are assessing outcomes including patients, clinicians, investigators and endpoint review committees.

The fourth type of bias that can impact upon internal validity is attrition bias. This is the biased occurrence and handling of protocol deviations and loss of participants to follow-up (Juni, Altman, & Egger, 2001) that occurs when there are systematic differences in the loss of patients between comparison groups in the trial. Patients may be excluded after allocation to treatment because they deviated from the trial protocol (e.g. non-adherence to prescribed treatment) or because they were not able to be followed-up (e.g. the patient could no longer be contacted). Patients that are excluded are unlikely to be representative of all the patients enrolled in the trial and so, once they have been excluded, investigators can no longer be sure that important baseline characteristics, which may impact on therapeutic effect, are similar between the comparison groups. To reduce the effects of attrition bias, analysis of outcomes should use the "intention to treat" principle, whereby analysis involves all randomised patients kept in the groups to which they were originally assigned.

In recent years, many studies have provided evidence of the biases occurring and distorting trial results. These studies suggest that low quality trials tend to exaggerate the effect of treatment (Egger et al. 2003; Kjaergard, Villumsen, & Gluud, 2001; Noseworthy et al. 1994; Peduzzi et al. 1993). This is particularly the case in trials that have small sample sizes, as it has been found that such trials are usually conducted and analysed with low methodological rigour resulting in the inflation of treatment effect estimates (Egger et al. 2003).

As there is evidence to suggest an association between treatment effect estimates and trial quality, it was appropriate to investigate whether the high proportion of positive trial results found in this

study from Japanese trials was a consequence of poor trial quality. This was especially important since the trials from Japan were small.

Trial design and execution within the Japanese trials was compared to a sample of European and North America trials, using a modified Jadad scale. It was hypothesised that the high proportion of positive trial results reported in Japanese trials, when compared to trials from both Europe and North America, were a consequence of trials not using "adequate[52]" methodological approaches.

7.2: Methods and Materials

As the samples of European (240) and North American (155) trials were large, not all trials from these continents were quality assessed. Instead, a sample of trials from Europe and North America were selected. These were selected from the 10 meta-analyses that were identified as containing a Japanese trial that had fatal endpoint data. In each meta-analysis containing a Japanese trial, the Japanese trial data was extracted and the European and North American trials listed. A stratified random sample was then taken of one European and one North American trial from each of the 10 meta-analyses. In the cases where the meta-analysis only contained Japanese trials and European trials, the European trials were listed with a random sample of two[53] European trials being taken. This was also the case in meta-analyses that only contained Japanese and North American trials. This ensured that Japanese trials were only being compared to trials answering the same or similar scientific questions. Furthermore, it meant that trials could be

[52] RCTs that were well designed and properly executed
[53] This allowed the quality of the Japanese trial to be compared to a larger sample of trials assessing the same intervention. One trial may not be representative of the quality of trials from the continents to which the Japanese trials are compared.

assessed for quality both within and over meta-analyses. A sample of 33 trials was created containing 13 Japanese trials, 14 European trials and 6 North American trials (see Table 7.1). It should be noted that prior to quality assessments being conducted all trial duplications were removed so that no trial had the opportunity to be assessed for quality more than once.

Table 7.1: Details of the Trials included in the Trial Quality Investigation.

Ref. No (Meta-analysis code)	Trial Author	Region	Methods	Participants	Intervention	Outcomes	Trial Effect Estimate (RR) and Confidence Intervals	Jadad Score
280 [C4]	Okishige, Nishizaki, Azegami et al (2000)	Japan	Randomised, Single-blind, Multi-centred. 12 month follow up	62 Patients with AF lasting > 6 months	Pilsicainide vs. Placebo	Mortality Pro-arrhythmia AF recurrence	0.62 (0.03, 14.30)	4
347 [C4]	Stroobandt, Stiels, Hoebrechts (1997)	Europe	Randomised, double – blind, Multi-centred. Follow up at 6 months	136 patients with persistent Atrial Fibrillation	Propafenane 2mg/kg over 30 mins followed by oral Propafenone 150mg 3 times daily vs. Matching Placebo	Pharmacologic Conversion to Sinus Rhythm Free from Recurrent Symptomatic Arrhythmia Time to Atrial Fibrillation Relapse	0.11 (0.00, 2.64)	3
10 [C4]	AFIB Investigators (1997)	North America	Randomised, Multi-centre, Placebo-controlled. 12 month follow up	1227 Patients with AF and 187 with Paroxysmal Supraventricular Tachycardia	Bidisomide vs. Placebo	Mortality Symptomatic recurrence of AF or Paroxysmal Supraventricular Tachycardia	2.01 (0.55, 7.41)	3
266 [C10]	Nakashima, Katayama, Honda et al (2004)	Japan	Randomised, Single-blind, Single-centre	180 Patients with successful PCI for AMI	Magnesium Sulphate vs. Placebo	Mortality Complications	0.34 (0.04, 3.22)	1
135 [C10]	Feldstedt, Boesgaard, Bouchelouche et al (1991)	Europe	Randomised, double-blind, single-centred. Median Follow up of 245 days	298 Patients with suspected AMI	Intravenous Magnesium Chloride vs. Placebo	In-hospital death Death at end of follow up Incidence of atrioventricular conduction disturbances	1.23 (0.50, 3.04)	4
393 [C10]	Woods, Fletcher, Roffe et al (1992)	Europe	Randomised, double-blind, Single-centre	2316 Patients with suspected AMI	Intravenous Magnesium Sulphate vs. Placebo	Mortality Morbidity	0.76 (0.59, 0.99)	5

3690 [C12]	The EPOCH Study Group (2002)	Japan	Randomised, Double – blind and Multi-centred	306 Patients with CHF and a radionuclide or echocardiographic LVEF	Pimobendon 1.25 or 2.5 mg vs. Placebo	Adverse cardiac events	0.51 (0.05, 5.60)	5
87 [C12]	Cowley and Skene (1994)	Europe	Randomised, Double-blind, Placebo controlled Multi-centred with 1 year minimum follow up	151 patients with severe symptomatic Heart Failure	Enoximore vs. Placebo	Survival Quality of Life Physical Mobility	1.46 (0.84, 2.53)	3
224 [C12]	Kubo, Gollub, Borge et al (1992)	North America	Randomised, double-blind, multi-centred.	198 patients with symptoms of moderate to severe heart failure	Pimobendanvs. Placebo	Drug safety Exercise duration	0.88 (0.24, 3.18)	3
318 [C17]	Saito, Hosokawa, Tanaka et al (1999)	Japan	Randomised, Multi-centre. 12 month follow up	137 Patients with AMI	Stent vs. Angioplasty	Mortality MI Revascularisation	0.42 (0.08, 2.08)	2
215 [C17]	Kawashima, Ueda, Nishida et al (1999)	Japan	Randomised, Multi-centre. 6 month follow up	222 Patients with AMI	Stent vs. Angioplasty	Revascularisation Mortality	0.34 (0.01, 8.24)	3
244 [C17]	Maillard, Haman, Khalife et al (2000)	Europe	Randomised, Multi-centre. 12 month follow up	211 Patients with AMI	Stent vs. Balloon Angioplasty	Mortality MI Revascularisation	3.26 (0.13, 79.24)	4
348 [C17]	Suryapranata ,van't Hof, Hoorntje et al (1998)	Europe	Randomised, Single-Centre. 12 month follow up	227 Patients with AMI	Stent vs. Angioplasty	Mortality MI Revascularisation	0.68 (0.12, 4.02)	5
352 [C41]	Takeda, Takano, Ogawa et al (1997)	Japan	Randomised, unblended, single-centre	30 Patients with severe cardiogenic PE	CPAP vs. Standard Medical Care	Mortality Tracheal intubation rate	0.33 (0.04, 2.85)	1
351 [C41]	Takeda, Neijima, Takano et al (1998)	Japan	Randomised, unblinded, Single-centre.	22 Patients with PE complicating AMI	CPAP vs. Standard Medical Care and Oxygen mask	Mortality Tracheal Intubation Rate	0.14 (0.02, 0.98)	1
89 [C41]	Crane, Elliott, Gilligan et al (2004)	Europe	Randomised, unblinded, Multi-centre	60 Patients with acute Cardiogenic Pulmonary Oedema	CPAP vs. BILEVEL vs. Standard Medical care plus Oxygen mask	Mortality Tracheal intubation rate Incidence of AMI Side Effects	0.42 (0.14, 1.20)	3
231 [C41]	Levitt (2001)	North America	Randomised, unblinded, Single-centre	38 Patients with severe acute Congestive HF	BILEVEL vs. Standard Medical Care and Oxygen Mask	Mortality Tracheal Intubation Rate Incidence of AMI	1.00 (0.23, 4.40)	3
350 [P2]	Takano, Hasegawa, Kuwabara et al (2007)	Japan	Randomised, blinded, Multi-centre. 6 month follow up	40 Patients with AMI related with left anterior descending coronary artery who underwent successful PCI	G-CSF vs. Placebo	LVEF LVEDV LVESV Restenosis Rate Mortality	3.63 (0.16, 84.11)	5
401 [P2]	Zohlnhofer, Ott, Mehilli et al (2006)	Europe	Randomised, double – blind and	114 patients with ST-segment Elevation AMI	Subcutaneous daily dose of 10 microg/kg of G-SF of Placebo for 5	Reduction of left ventricular infarct size	3.11 (0.13,	5

			Multi-centred		days	Improvement of left ventricular ejection fraction Incidence of angiographic restenosis	74.66)	
124 [P2]	Ellis, Penn, Bolwell et al (2006)	North America	Randomised, double-blind, Single-centred. 30 day follow up	18 Patients with AMI	G-CSF vs. Placebo	Rupture free survival Recovery of LVF	0.18 (0.01, 3.85)	5
180 [P7]	Horie, Takahashi, Minai et al (1998)	Japan	Randomised, Single-centre.	83 Patients with initial Q-wave anterior MI > 24hrs after onset	PTCA vs. No PTCA	Cardiac Death, Non-fatal recurrence of MI Development of Congestive HF	0.18 (0.02, 1.45)	2
126 [P7]	Erne, Schoenenberger, Burckhardt et al (2007)	Europe	Randomised, Unblinded, Multi-centred trial with final follow up at 10 years	201 patients with recent MI, Silent Myocardial Ischemia and 1- or 2- vessel Coronary Artery Disease	PCI vs. Intensive Anti-Ischemic Drug Therapy	Survival free of major cardiac events Exercise induced ischemia Resting left ventricular ejection fraction	0.15 (0.05, 0.48)	2
113 [P7]	Dzavik, Beanlands, Davies et al (1994)	North America	Randomised, Multi-centre. 4 month follow up	44 Patients with occluded infarct-related coronary artery	FrCA vs. no PTCA	Mortality LVEF	0.76 (0.05, 11.39)	2
263 [P87]	Muramatsu, Iwasaki, Inou et al (2002)	Japan	Randomised, Multi-centre. 6 month follow up	302 Patients with CAD	Stenting vs. PTCA	Mortality Restenosis rate MACE	0.12 (0.01, 2.13)	1
254 [P87]	Moer, Myreng, Molstad et al (2001)	Europe	Randomised, Multi-centre. 6 month follow up	145 Patients with stable and unstable angina	Elective stenting with Heparin coated beStent vs. PTCA	MLD Restenosis Rate Event free Survival Angina status	0.96 (0.06, 15.05)	3
212 [P87]	Kastrati, Schomig, Dirschinger et al (2000)	Europe	Randomised, Multi-centre. 7 month follow up	404 Patients with symptomatic CAD	Stenting vs. PTCA	Incidence of angiographic restenosis Adverse clinical events	0.65 (0.11, 3.87)	3
187 [P90]	Ikari, Kawano, Sakurada et al (2005)	Japan	Randomised, Multi-Centre. 8 month follow up	368 Patients with STEMI	Primary PCI with thrombus aspiration vs. PCI alone	Success and complication rate of thrombus aspiration therapy Damage of Myocardial Microcirculation MACE	0.97 (0.06, 15.42)	1
147 [P90]	Fujita, Suwa, Koyama et al (2004)	Japan	Randomised, Multi-Centre.	341 Patients with acute MI	Adjunctive embolic protection vs. no adjunctive embolic protection	30 day major cardiac adverse events	0.69 (0.22, 2.14)	1
40 [P90]	Beran, Lang, Schreiber et al (2002)	Europe	Randomised, Single – centre. 30 day follow up	66 Patients with ACS and suspected intracoronary thrombus	Pretreatment with X-sizer with PCI vs. Conventional PCI	MACE	2.07 (0.20, 21.61)	4
14 [P90]	Ali, Cox, Dib et al (2006)	North America	Randomised, Multi-centre. 30 day	480 Patients presenting in 12hr of symptom onset	Rheolytic thrombectomy adjunct to PCI vs. PCI alone	Infarct size MACE	5.50 (1.23,	3

			follow up	of STEMI			24.55)	
248 [P97]	Matsuo, Inoue, Sakurada et al (2007)	Japan	Randomised, Multi-centre. 6 month follow up	154 Patients with AMI in 24hr from onset	PCI with Guardwire vs. PCI	TIMI perfusion grade MACE	0.31 (0.03, 2.90)	1
97 [P97]	DeLuca, Sardella, Davidson et al (2006)	Europe	Randomised, single-centred. 6 month follow up	76 Patients with anterior STEMI	Intracoronary Thrombectomy and stent placement vs. Conventional stenting	Major Adverse Cardiovascular Events (MACE)	0.20 (0.01, 4.03)	2
349 [P97]	Svilaas, Vlaar, van der Horst (2008)	Europe	Randomised, single-centred, blinded. Follow up of 30 days	1071 patients with ST-segment elevation MI	Thrombus aspiration vs. Conventional PCI	Post procedural frequency of a Myocardial blush grade of 0 or 1 Post procedural frequencies of TIMI flow grade of 3. Death, reinfarction or combination of adverse Cardiac events	0.61 (0.38, 0.99)	5

As only a sample of European and North American trials were chosen, it was important to see whether these samples had typical effect sizes. This would ensure that no bias was introduced into this assessment by comparing Japanese trials to an unrepresentative, biased sample of European and/or North American trials. To do this, the distribution of trials from both continents were examined by producing scatter plots in which each trial's standardised effect estimate was plotted against the square root of trial sample size and in which the trials used in the trial quality assessment were highlighted (see Figures 7.1 and 7.2). This showed that the samples of both European and North American trials provided a range of typical effect sizes. As such, they would provide a good sample to compare with the Japanese trials.

FIGURE 7.1: THE SPREAD OF EUROPEAN TRIAL TREATMENT EFFECT ESTIMATES WITH THE TRIALS INCLUDED IN THE TRIAL QUALITY ASSESSMENT HIGHLIGHTED. 0 REPRESENTS A UNIVERSAL TREATMENT EFFECT ESTIMATE.

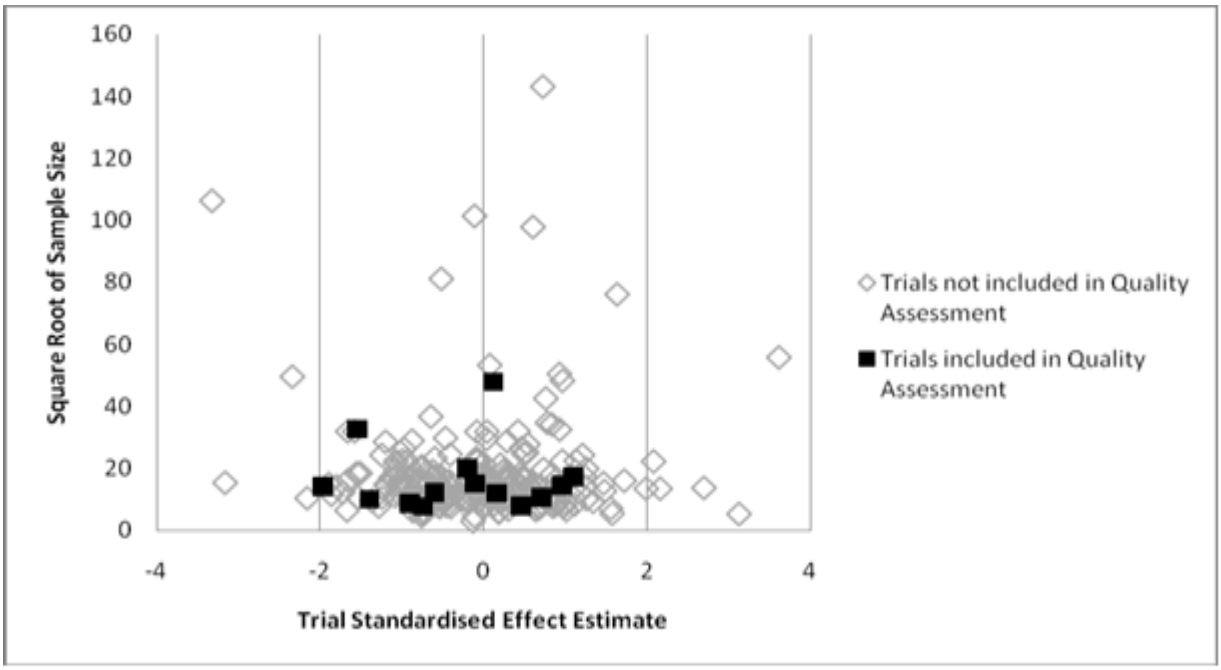
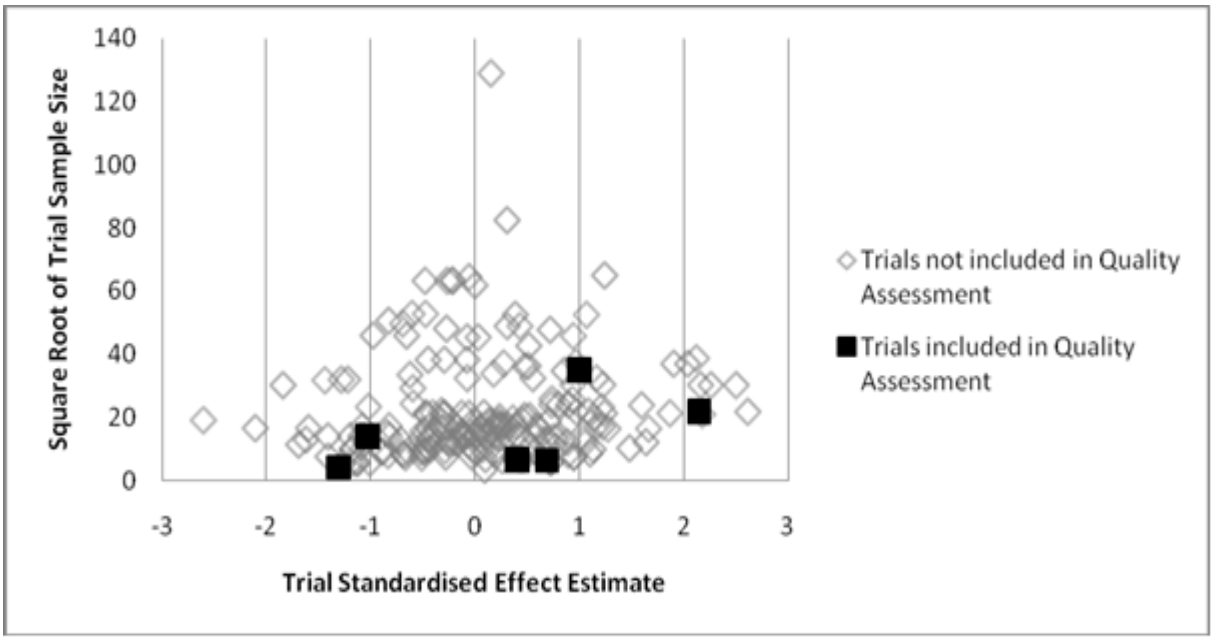


FIGURE 7.2: THE SPREAD OF NORTH AMERICAN TRIAL TREATMENT EFFECT ESTIMATES WITH THE TRIALS INCLUDED IN THE TRIAL QUALITY ASSESSMENT HIGHLIGHTED. 0 REPRESENTS A UNIVERSAL TREATMENT EFFECT ESTIMATE.



7.2.1 The Assessment of Trial Quality

To assess the quality of each of the 33 trials, a trial quality questionnaire was developed (see Appendix 1). The core of this questionnaire was the Jadad Scale. This was used because it is both quick and simple and is the most frequently used scale to assess trial quality in health care (Moja et al. 2005). In addition, this scale is reported to be the most reliable and valid scale in assessing the quality of randomised controlled trials (Olivo et al. 2008).The Jadad scale gives every trial a quality score which lies between one and five. Trials are scored according to the presence of the

three most important methodological aspects of a randomised controlled trial that can impact on the reliability of results: randomisation, blinding and attrition. A maximum of two points are available for appropriate and well reported randomisation and blinding and a maximum of one point for well-reported and explained attrition. If no information is provided for each of these elements of trial quality, the trial scores no points for each category. Trials are recorded as low quality if they score between zero and two, and high quality if they score between three and five.

For this study, the Jadad score was calculated independently by both investigators using the Jadad scoring system mentioned above. Any discrepancies in Jadad score between reviewers were resolved by examining the papers and discussing disagreements with a third reviewer (RL).

As well as using the Jadad scale, some additional trial quality questions about randomisation, blinding and attrition were included on the quality assessment questionnaire in this research. This allowed information not collected by the Jadad Scale to be collated. This enhanced Jadad scale included questions about the following:

1. *Was the trial multi-centred?*
2. *If the trial does not specify blinding, does it specify that a placebo was used?*
3. *Was the patient blinded?*
4. *Was the investigator blinded?*
5. *Was the main outcome considered objective rather than subjective?*
6. *Did the trial follow a Consort type detailed breakdown of both withdrawals and drop-outs following randomisation?*

These additional questions were important as they meant that a more in-depth trial quality assessment could be conducted. The Jadad scale only asks whether there was a description of withdrawals and drop-outs and therefore, does not ask for a detailed breakdown, whereas in the

enhanced Jadad scale used here, question six allowed attrition to be investigated more thoroughly. This type of breakdown, however, is important since knowing the number of patients who were not eligible, who did not receive their allocated treatment or did not complete treatment means that the extent to which the treatment effect may be biased can be assessed. Furthermore, these additional questions allow observations to be made about whether those trials scored as low quality have a biased design as a consequence of the procedures it has used, for instance, its randomisation procedures. This is not possible when only using the Jadad scale, which although it asks whether appropriate randomisation techniques were used, does not take into account that some of these appropriate techniques are more prone to bias than others.

Since sample size is thought to mediate the relationship between trial quality and treatment effect (Singh, Murphy, & Bhandari, 2010), single-centred trials are likely to exaggerate the effect of treatment due to their small sample sizes (Bellomo, Warrillow, & Reade, 2009) . Therefore, the first additional question asking whether the trial being assessed was single or multi centred enabled another factor that may be linked to trial quality and the exaggeration of treatment effect to be investigated. In addition, research suggests that subjective outcomes are specifically vulnerable to bias when a trial has a lack of blinding (Higgins & Altman, 2008) which may result in an exaggeration of treatment effect.[54] As such, the addition of question five, allowed the assessment of whether an objective or subjective outcome had been measured for each trial.

[54] After trial quality assessment had been conducted all included trials were found to have assessed objective outcomes. Therefore, the results of this question are not reported in this thesis.

Once the questionnaire had been developed, it was pre-tested by two investigators (LH and UN). To do this, a selection of the included trials was chosen at random and the two investigators independently assessed them using the questionnaire. The investigators then compared their trial quality assessments and any discrepancies were discussed with a third party (RL). After this discussion, minor changes were made to the enhanced Jadad scale. It was agreed that two additional questions about randomisation processes should be included. The revised list of questions is given below:

1. *Was the trial multi-centred?*
2. ***Was randomisation done by a third party i.e. it was not within the control of a given clinic?***
3. ***Was randomisation done by envelopes in the local setting?***
4. *If the trial does not specify blinding, does it specify that a placebo was used?*
5. *Was the patient blinded?*
6. *Was the investigator blinded?*
7. *Was the main outcome considered objective rather than subjective?*
8. *Did the trial follow a Consort type detailed breakdown of both withdrawals and drop-outs following randomisation?*

Inter-rater reliability and validity of the enhanced questionnaire was checked by comparing the Jadad scores provided by this questionnaire to those given in the quality assessment of Cochrane reviews. This was only done, however, if the trial being assessed had been extracted from a Cochrane review and that review had provided a Jadad score.

The original reports of the all the trials being assessed were obtained and photocopied three times. On two copies, a black marker pen was used to make information regarding affiliations incomprehensible. This was done so that all trials could be assessed fairly. The other copy was

left with this information unmasked for reference after the assessment. Each investigator (LH and UN) was then handed a masked copy of each trial to assess trial quality.

7.2.2 Data abstraction

The two investigators[55] (LH and UN) independently assessed trial quality using the enhanced Jadad questionnaire. Agreement on data abstraction was reached in 82% of cases. However, there were 7 disagreements between the reviewers. Three of these were with regards to blinding, three were concerned with randomisation and one was related to attrition. All of these disagreements were resolved through examining the papers with a third reviewer (RL) and discussing the issue concerned. The discussions resulted in a consensus between all three reviewers about trial quality.

7.3 Results of Trial Quality Assessment

The assessment of trial quality was conducted both over meta-analyses and trials. This section reports the findings of these assessments.

7.3.1: Assessment at meta-analysis level

To examine whether the quality of trials differed between Japan, Europe and North America, trials were first assessed within meta-analyses. This eliminated variation in quality between procedures since trials were only being compared to those using the same procedure. This investigation was important since quality assessment criteria may differ according to procedure

[55] Two investigators were chosen to assess trial quality so that mistakes and the risk of bias during assessment could be reduced.

type and, as such, may impact on the quality score given. For instance, when assessing the quality of trials for surgical interventions blinding is not feasible. Therefore, these trials would score lower than a trial investigating a procedure where blinding was possible and, as such, appear to be of lower quality. Table 7.2 shows the quality of each region's trials within each of their meta-analyses.

As can be seen from Table 7.2, in two meta-analyses, concerning Anti-arrhythmic drugs and Phosphodiesterase III Inhibitors, (C4 (227) and C12 (18)) the Japanese trials appeared to be of better quality than both European and North American trials. In one meta-analysis (P2 (4)) that examined G-CSF therapy, the Japanese trial was of equal quality to its European and North American counterparts. In P7 (2), the Japanese trial was of the same quality as the North American trial but of lower quality compared to the European trial. However, in most cases (n=6) the Japanese trials were of lower quality than trials conducted in Europe and/or North America.

Table 7.2: A Table Showing the Jadad score for each Region's Trials in each Meta-analysis.

Ref. No/ (Meta-analysis)	Intervention	Trial												Highest Quality Region
		1			2			3			4			
		Region	Jadad Score	Effect Estimate (RR and CIs)	Region	Jadad Score	Effect Estimate (RR and CIs)	Region	Jadad Score	Effect Estimate (RR and CIs)	Region	Jadad Score	Effect Estimate (RR and CIs)	
227 [C4]	Anti-arrhythmic	Japan	4	0.62 (0.03, 14.30)	Europe	3	0.11 (0.00, 2.64)	North America	3	2.01 (0.55, 7.41)	N/A	N/A	N/A	Japan
232 [C10]	Intravenous Magnesium	Japan	1	0.34 (0.04, 3.22)	Europe	4	1.23 (0.50, 3.04)	Europe	5	0.76 (0.59, 0.99)	N/A	N/A	N/A	Europe
18 [C12]	Phosphodiesterase III Inhibitors	Japan	5	0.51 (0.05, 5.60)	Europe	3	1.46 (0.84, 2.53)	North America	3	0.88 (0.24, 3.18)	N/A	N/A	N/A	Japan
273 [C17]	Stenting	Japan	2	0.42 (0.08, 2.08)	Japan	3	0.34 (0.01, 8.24)	Europe	4	3.26 (0.13, 79.24)	Europe	5	0.68 (0.12, 4.02)	Europe
383 [C41]	Non-invasive Positive Pressure Ventilation	Japan	1	0.33 (0.04, 2.85)	Japan	1	0.14 (0.02, 0.98)	Europe	3	0.42 (0.14, 1.20)	North America	3	1.00 (0.23, 4.40)	Europe/ North America
4 [P2]	G-CSF Therapy	Japan	5	3.63 (0.16, 84.11)	Europe	5	3.11 (0.13, 74.66)	North America	5	0.18 (0.01, 3.85)	N/A	N/A	N/A	All Continents Equal
2 [P7]	Percutaneous Coronary Intervention	Japan	2	0.18 (0.02, 1.45)	Europe	3	0.15 (0.05, 0.48)	North America	2	0.76 (0.05, 11.39)	N/A	N/A	N/A	Europe
11 [P87]	Stenting	Japan	1	0.12 (0.01, 2.13)	Europe	3	0.96 (0.06, 15.05)	Europe	3	0.65 (0.11, 3.87)	N/A	N/A	N/A	Europe
96 [P90]	Adjunctive Mechanical Devices	Japan	1	0.97 (0.06, 15.42)	Japan	1	0.69 (0.22, 2.14)	Europe	4	2.07 (0.20, 21.61)	North America	3	5.50 (1.23, 24.55)	Europe
33 [P97]	Thrombectomy or Embolic Protection Device	Japan	1	0.31 (0.03, 2.90)	Europe	2	0.20 (0.01, 4.03)	Europe	5	0.61 (0.38, 0.99)	N/A	N/A	N/A	Europe

For instance, in C10 (232) it was found that the Japanese trial scored between 3-4 points lower on the Jadad scale than its European counterparts. This was also the case for both meta-analyses (C17 (273), P87 (11)) concerned with stenting, where the Japanese trials scored lower than the trials conducted in Europe. Indeed, in 60% of the meta-analyses, the Japanese trials appeared to be of lower quality compared to European and North American trials. This can be seen in Figure 7.3 where six of the 10 meta-analyses show the Japanese trials scoring lower on the Jadad scale than their European and North American counterparts. However, a binomial sign test conducted under the null hypothesis that the quality of trials would be the same in Japanese trials compared to Europe and North America indicated that trial quality was not significantly different between the Japanese trials and their European and North American counterparts (2-sided p-value = 0.7539).

FIGURE 7.3: THE TREATMENT EFFECT ESTIMATE AND JADAD SCORE OF EACH REGION'S TRIALS PER META-ANALYSES.

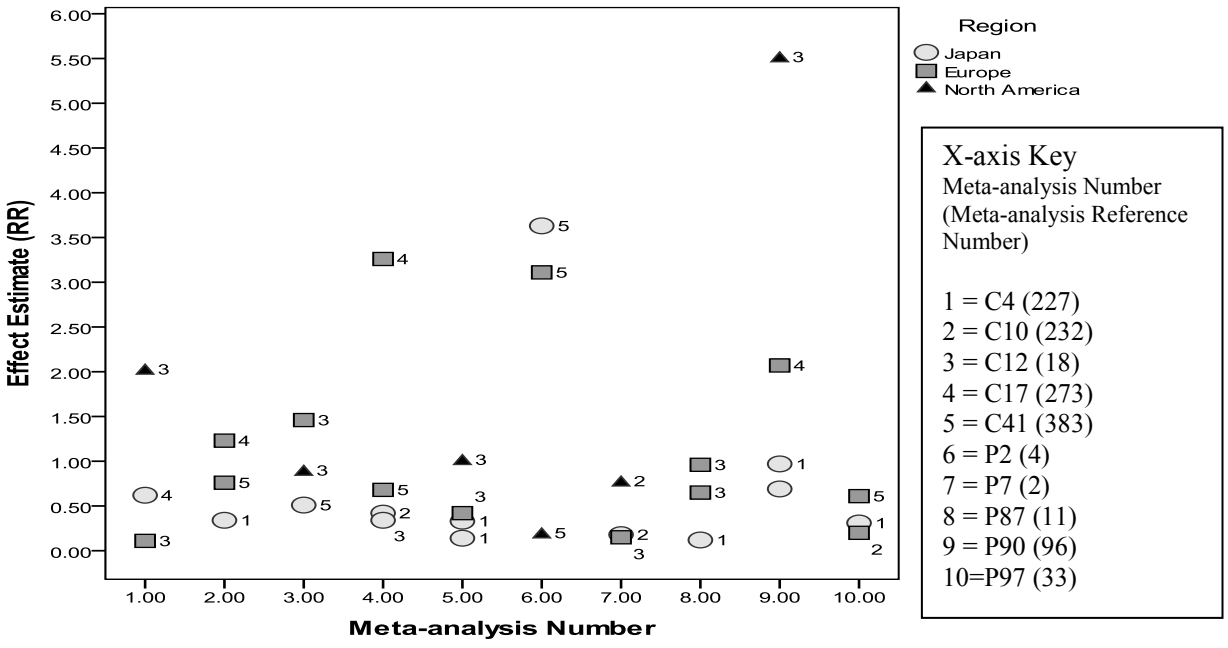


Table 7.2 was also used to see whether the Japanese trials in each meta-analysis provided higher treatment effect estimates than their European and North American counterparts. This was because the high proportion of positive Japanese trial results may have been due to such trials being of lower quality, compared to their European and North American counterparts, and, as such, exaggerating the effect of treatment. It was found that in the majority of meta-analyses the treatment effect estimates provided by the Japanese trials were higher than those provided by their European and North American counterparts. Indeed, in 7 of the 10 meta-analyses the treatment effect estimates provided by the Japanese trials were higher than their European and North American counterparts (see Table 7.2). To see if this was a significant finding, a binomial sign test was conducted under the null hypothesis that the effectiveness of interventions would be the same in the Japanese trials as in their European and North American counterparts. This indicated that treatment effectiveness was not significantly different between trials conducted in Japan and those conducted in Europe and North America (2-sided p-value = 0.344).

7.3.2: Analysis over Trials

An analysis was also conducted over trials. This involved assessing all of the trials included in the trial quality sample together. Trials from Japan, Europe and North America were compared in relation to whether they were single or multi-centred[56] trials and with regards to randomisation, blinding and attrition.

[56] Trials conducted with all centres in one continent.

7.3.2.1: Single or Multi-centre Trials

The percentage of single centre and multi-centre trials was first calculated for each continent. This was to ascertain if trials in Japan were more likely to be single-centred compared to trials conducted within Europe and North America. This was important to investigate, as the positive results in Japanese trials may have been created by chance as a consequence of small sample sizes in single-centred trials. This comparison showed that the proportion of single-centre and multi-centre trials between continents were similar. Table 7.3 shows the percentage of trials in each region that were single or multi-centred.

Table 7.3: The Percentage of Single Centre and Multi-Centre Trials in each Continent.

	Japan	**Europe**	**North America**
Single Centre Trials	4 (31%)	5 (36%)	2 (33%)
Multi-Centre Trials	9 (69%)	9 (64%)	4 (67%)
Total number of Trials	13	14	6

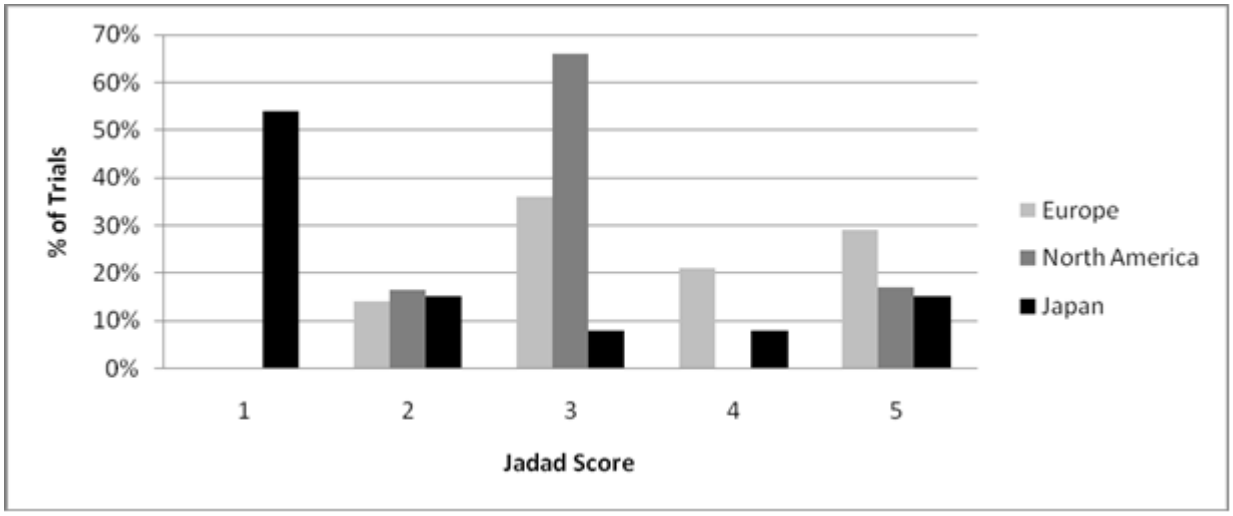
As can be seen from Table 7.3, the percentage of single-centre trials in Japan was 31% (4 out of 13 trials). This is similar to the percentages in both Europe (36% or 5 out of 14 trials) and North America (33% or 2 out of 6 trials). This suggested that the positive trial results found in Japanese trials could not solely be due to such trials being single-centred with small sample sizes.

7.3.2.2: Trial Jadad Scores

Examining the Jadad scores of each region's trials showed that Japanese trials scored lower on the Jadad scale compared to trials from Europe and North America. As can be seen from Figure 7.4, 54% of Japanese trials only scored one on the Jadad scale. Furthermore, only 31% of Japanese trials scored between three and five. This was low compared to Europe and North

America, where no trial scored one and where 86% and 83% of trials respectively scored between three and five.

FIGURE 7.4: A BAR CHART SHOWING THE PERCENTAGE OF TRIALS IN EACH CONTINENT SCORING AT EACH LEVEL OF THE JADAD SCALE.



To investigate whether this difference in trial quality between Japan, Europe and North America was significant, a Kruskal-Wallis one-way ANOVA was performed on the regional groups. As expected, Japanese trials had the lowest quality scores (Median=1) compared to Europe (Median = 3.5) and North America (Median=3). The test was also statistically significant: χ^2 (2df, N=33) = 8.296, p= 0.016. Therefore, it was concluded that there was a statistically significant difference in the quality of trials between Japan, Europe and North America.

To investigate where this significant difference may lie, follow-up Mann-Whitney U Tests were conducted that evaluated pair-wise differences among the three regions, controlling for Type I error across tests by using the Bonferroni Correction. The results of these tests

indicated that Japanese trials scored significantly lower on the Jadad score than European trials (U = 37.5, p = 0.008, N = 27, 2 tailed test).[57] However, the tests also indicated that there was no significant difference in Jadad score between Japan and North America (U = 20.0, p = 0.106, N = 19, 2 tailed test) or between Europe and North America (U = 28.5, p = 0.274, N = 20, 2 tailed test).

The quality of each region's trials was then examined with regards to the three key methodological aspects described earlier in the chapter: randomisation, blinding and attrition. The findings from these comparisons will now be discussed in turn.

7.2.3.3: Randomisation

All trials from Japan, Europe and North America stated that they were randomised. However, the number of trials reporting an appropriate randomisation technique differed between continents, as can be seen from Table 7.4. It should be noted that in this instance an appropriate randomisation technique was considered to be where the allocation sequence was unpredictable and was concealed from those who enrolled the participants.

Table 7.4: The Number of Trial in Each Region That Used an Appropriate Randomisation Technique.

Region	Randomisation Technique Appropriate			Total Number of Trials
	Yes	No	Not Stated	
Japan	3 (23%)	2 (15%)	8 (62%)	13
Europe	9 (64%)	1 (7%)	4 (29%)	14
North America	2 (33%)	1 (17%)	3 (50%)	6

[57]The probability is exact and was found using SPSS version 11.

As Table 7.4 shows, only 23% (3 of 13 trials) of Japanese trials reported the use of an appropriate randomisation technique. This was lower than the percentage of European trials (64% or 9 of 14 trials) and North American trials (33% or 2 of 6 trials). Moreover, 62% (8 of 13 trials) of Japanese trials, 29% (4 of 14 trials) of European trials and 50% (3 of 6 trials) of North American trials did not state how they had randomised patients. Therefore, the low percentage of trials in both Japan and North America reporting an adequate randomisation technique could be the consequence of this non-reporting rather than the randomisation technique not actually being used in the study. Nonetheless, statistical analysis (2-sided Fisher's Exact Test) did not show a statistically significant difference between regions in reporting adequate randomisation techniques ($p = 0.261$). As such, there was not sufficient evidence to claim that the reporting of adequate randomisation techniques was different between the regions investigated.

Additional analysis of randomisation processes were conducted to see whether regions differed in the methods they employed to conceal allocation. This was done by observing the number of trials in each region that used third party randomisation or used envelopes in the local setting. This was important because deciphering allocation schedules is more likely to occur in trials that have used envelopes in the local setting since unsealed envelopes may be opened and translucent envelopes held up to a bright light (Schulz, 1995; Schulz & Grimes, 2002) meaning that allocation may still be in the hands of the researcher. This could result in selection bias and consequently, an exaggeration of treatment effect. Third party randomisation, on the other hand, is considered to be the most desirable approach (Akkerhuis et al. 2000; Hotopf, 2007; Sim & Wright, 2000) because it minimizes selection bias and, therefore, is less likely to lead to over-stating treatment effect.

Table 7.5: Randomisation Techniques by Region

Region	Randomisation Method			Total Number of Trials
	Third Party	Envelopes in Local Setting	Not Stated	
Japan	3 (23%)	0 (0%)	10 (77%)	13
Europe	4 (29%)	3 (21%)	7 (50%)	14
North America	1 (17%)	0 (0%)	5 (83%)	6

As Table 7.5 shows, 23% (3 out of 13) of Japanese trials used third party randomisation compared to 29% (4 out of 14) of European trials and only 17% (1 out of 6) of North America trials. No Japanese or North American trials reported using envelopes in the local setting compared to 21% (3 of 14 trials) of European trials that did this. These results suggested, therefore, that the positive trial results found in Japanese trials may not be explained by the use of less robust "local envelope" randomisation techniques. Furthermore, a 2-sided Fisher's Exact Test showed no evidence of a statistically significant difference between regions in reporting allocation concealment methods (p=0.386). However, it should be noted that 77% (10 of 13) of Japanese trials did not state whether third party randomisation or envelopes in the local setting were used. Therefore, it may be that Japanese trials did use envelopes in the local setting more often but did not report doing so.

7.3.2.2: Blinding

Only 31% (4 of 13) of Japanese trials were found to have reported the use of blinding (see Table 7.6). This was lower than both Europe (64% or 9 of 14) and North America (50% or 3 of 6).

Table 7.6: The Number of Each Regions Trials That Used, Did Not Use or Did Not State if They Had Used Blinding.

Region	Was Blinding Used in the Trial?			Total Number of Trials
	Yes	No	Not Stated	
Japan	4 (31%)	6 (46%)	3 (23%)	13
Europe	9 (64%)	4 (29%)	1 (7%)	14
North America	3 (50%)	1 (17%)	2 (33%)	6

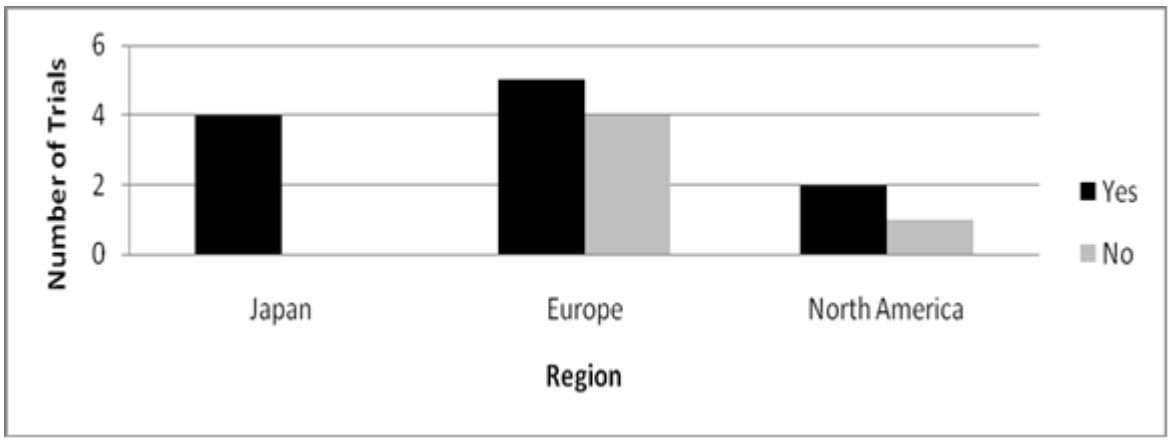
Furthermore, 23% (3 of 13 trials) of trials in Japan did not state whether blinding had been used. As can be seen from Table 7.6, this was higher than the percentage for Europe (7% or 1 of 14 trials) but lower than the percentage for North America (33% or 2 of 6 trials). However, a 2-sided Fisher's Exact Test showed no evidence of a statistically significant difference between regions in reporting the use of blinding ($p = 0.304$). This suggests that the differences in trial quality between regions may not be a consequence of differences in the use of blinding.

The number of blinded trials stating the blinding method was also examined between regions. Table 7.7 shows that all Japanese trials that reported the use of blinding (n=4) stated the method used compared to 56% (5 of 9) of European trials and 67% (2 of 3) of North America trials (see also Figure 7.5).

Table 7.7: Method of Blinding per Region

Region	Blinding Method Described?		Total Number of Trials
	Yes	No	
Japan	4 (100%)	0 (0%)	4
Europe	5 (56%)	4 (44%)	9
North America	2 (67%)	1 (33%)	3

FIGURE 7.5: A GRAPH SHOWING THE PERCENTAGE OF TRIALS IN EACH CONTINENT THAT USED BLINDING AND DESCRIBED THE BLINDING METHOD USED.



However, a 2-sided Fisher's Exact Test again found no evidence of a statistically significant difference between regions in reporting blinding methods ($p = 0.291$).Consequently, there was insufficient evidence to suggest that the reporting of blinding methods differed significantly between regions, suggesting that this factor may not solely explain the between-regional differences in trial quality found.

7.3.2.3: Attrition

Attrition rates refer to the number of patients that withdrew or dropped out of the trial. Table 7.8 shows the number and percentages of each region's trials that did or did not report attrition rates.

Table 7.8: The Number of Trials in Each Region Reporting or Not Reporting Attrition Rates.

Region	Attrition Rates Reported		Total Number of Trials
	Yes	No	
Japan	4 (31%)	9 (69%)	13
Europe	14 (100%)	0 (0%)	14
North America	5 (83%)	1 (17%)	6

It shows that 69% (9 of 13) of Japanese trials did not report attrition rates compared to no European trials and only 1% (1 of 6) of North American trials. To see if this was statistically significant, a 2-sided Fishers Exact Test was conducted. It indicated that there was evidence of a statistically significant difference in the reporting of attrition rates between regions ($p = 0.0001$).

Whether trials provided a CONSORT-type detailed breakdown of attrition rates was also explored. In the CONSORT Statement, trials are required to provide a flow diagram displaying the progress of all participants through the four stages of a trial: enrolment, intervention allocation, follow-up and analysis. In this sample, such a breakdown was only provided in 8% (1 of 13) of Japanese trials compared to 43% (6 of 14) of European trials and 17% (1 of 6) of North American trials (see Table 7.9).

Table 7.9: The Number of Trials in Each Continent Reporting Attrition in a CONSORT-type Detailed Breakdown.

Region	Trials Reporting Attrition Rates in a Consort Type Detailed Breakdown		Total Number of Trials
	Yes	No	
Japan	1 (8%)	12 (92%)	13
Europe	6 (43%)	8 (57%)	14
North America	1 (17%)	5 (83%)	6

It should be pointed out, however, that such a detailed breakdown of attrition rates was not presented in 25 of the 33 trials included in this sample. Moreover, a 2-sided Fisher's Exact Test showed no evidence of a statistically significant difference between regions in reporting a CONSORT type breakdown of attrition rates ($p = 0.107$). The CONSORT Statement was only established in 1996 (Begg et al. 1996), and as a result, trials published before this year could not have implemented its recommendations. What is more, since the CONSORT Statement only makes recommendations (not requirements) about reporting, trials published after this date may still not have put into practice these recommendations about attrition reporting.[58]

7.4: Meta-regression

Random effects meta-regression was conducted in STATA to further explore heterogeneity (for output see Appendix 2). First, a meta-regression was run to see whether treatment effect could be predicted by region. This showed that region could account for 32% of the between-

[58]In this study, trials that could not report attrition using the CONSORT Statement, i.e. those published before the CONSORT Statement existed, were not weighted less favourably in terms of quality. This was because the Jadad scale used to assess quality in this study does not ask a question about the CONSORT Statement, only a question about the reporting of attrition rates. As such, the quality score given to each trial by the Jadad scale does not take into account whether the trial followed the CONSORT Statement or not.

study heterogeneity and showed that there was a marginally significant difference between regions (F (2,30) = 3.29, p = 0.051) in the size of treatment effect. Furthermore, a statistically significant difference in the size of treatment effect was present between Japan and North America (F (1,30) = 6.55, p = 0.016).

A second meta-regression was then conducted to examine the impact of Jadad score after adjusting for region. However, Jadad score was found to not significantly contribute to the differences in the size of treatment effect (p=0.411). The last meta-regression was carried out to observe the association between treatment effect and Jadad score without adjusting for region. This showed that Jadad score did not significantly contribute to the differences in treatment effect (p=0.1157). However, it should be noted that the results of these meta-regressions are purely observational as by containing few trials, they have low power to detect any relationship between the covariates and treatment effect.

7.5: Publication Bias

To further explore the influence of bias on treatment effect estimation, publication bias was investigated in the Japanese, European, and North American sample of trials using the Egger test. The high proportion of positive Japanese trials found may result from publication bias in that trials showing positive results are more likely to be published. Indeed, the Egger test conducted on the Japanese trials showed that there was statistically significant publication bias (Egger statistic = 0.743, p = 0.047). Conversely, the Egger test result for Europe and North America did not show any evidence of publication bias (Europe Egger statistic = -0.714, p = 0.47, North America Egger statistic = -1.214, p = 0.67).

7.6: Conclusion

The investigation into trial quality as an explanation for the high proportion of positive trial results found in Japan provided some interesting findings. Differences in reporting randomisation within trials explored in this chapter were not statistically significant, nor were the differences found in how blinding was reported. The lack of statistical significance for blinding could be explained by the fact that, unless the intervention is drug based, it can be difficult or indeed, impossible, to blind a study. This difference may, therefore, be a consequence of the intervention rather than poorly designed, managed, and reported trials.

Nevertheless, examination over meta-analyses found that trial quality was different between trials conducted in Japan and those conducted in Europe and North America. Trials from Japan scored significantly lower on the Jadad scale of trial quality than trials from Europe and North America ($p = 0.016$). When explored further, it was found that they scored significantly lower than their European counterparts ($p = 0.008$) but not than their North American counterparts ($p = 0.274$). This may, however, be a consequence of the small sample size for North America.

The difference in trial quality could be a result of Japanese trials reporting rates of attrition less often than their European and North America counterparts, a difference that was statistically significant at the 5% level. Furthermore, Japanese trials that did provide attrition information tended not to comply with the CONSORT recommendations for presenting this type of data, even though all Japanese trials involved in this study were published after 1996: only 8% (1 of 13) of Japanese trials, compared with 43% (6 of 14) of European trials and

17% (1 of 6) of North American trials presented attrition data using the CONSORT recommendations. However, this difference was not statistically significant at the 5% level.

As studies about the impact of trial quality on effect estimates suggest, one might expect Japanese trials to provide larger treatment effects than the comparison trials, and this was the case in this study. Many of the included meta-analyses showed Japanese trials to have larger treatment effect estimates than their European and North American counterparts. However, these differences were not statistically significant at the 5% level. Analysis across trials found that the positive results from Japanese trials could not be attributed to these trials being single-centred because the percentage of single-centre trials was similar between Japan, Europe and North America, with the majority of trials from all three regions being multi-centred. However, evidence suggested that the high proportion of positive trials in Japan was due to publication bias.

To conclude the results of this investigation suggest that Japanese trials are of lower quality compared to their European and North American counterparts, especially with regards to the reporting of attrition rates. As might be expected, therefore, Japanese trials provided higher treatment effect estimates. While it is important to bear in mind that studies on trial quality indicate that poor reporting of trial methods and results is a reflection of poor trial methods and administration (Juni, Altman, & Egger, 2001; Liberati, Himel, & Chalmers, 1986; Schulz et al. 1995), as the trials across all three regions provided different types and levels of methodological information, it is possible that the results presented here about trial quality reflect not the trial quality itself but the quality of decisions around reporting of trials.

PART THREE: EVALUATION AND CONCLUSION

CHAPTER 8.
DISCUSSION & CONCLUSION

The purpose of this study was to investigate the existence of international differences in treatment effectiveness in cardiovascular diseases. In addition, it aimed to identify the types of interventions that were likely to show such differences in treatment effect. Panoramic meta-analysis was used to detect and explore differences in treatment effect between Europe, North America and Asia for both fatal and non-fatal endpoints.

This chapter summarises the main findings of this study and explains why the findings are important for research practice and the practice and policy of evidence based medicine. To do this, the findings are contextualised within the literature on inter-continental differences in treatment effectiveness. The limitations of the study are discussed and the areas that warrant further investigation outlined.

8.1: Empirical Findings

This study provides some evidence of the existence of international differences in the effectiveness of cardiovascular interventions between Europe, North America and Asia. These differences were evident for both fatal and non-fatal endpoints.

8.1.1: Fatal Endpoints

For the pair-wise comparison between Europe and North America no evidence of a statistically significant difference in the effectiveness of cardiovascular interventions between Europe and North America was found. This was also the case for the pair-wise comparison

between North America and Asia. For the comparison between Europe and Asia, however, statistically significant differences were found, with interventions performing better, relative to controls, in Asia compared to Europe.

The universal comparison supported this finding, as it indicated that Asian trials provided more positive results favouring the intervention. This was also statistically significant. Furthermore, when Asian trials were grouped by country, it was observed that a high proportion of Japanese trials provided positive trial results, a finding that contrasted with previous research that suggested China would provide the highest proportion of positive study results (Pan et al. 2005; Zhang, Freemantle, & Cheng, 2011).

To explore this unexpected finding, the quality of the included trials was evaluated because poor trial quality has been linked to exaggerated treatment effect estimates (Egger, Juni, Bartlett, Holenstein, & Sterne, 2003; Kjaergard, Villumsen, & Gluud, 2001; Noseworthy, Ebers, Vandervoort, Farquhar, Yetisir, & Roberts, 1994; Peduzzi, Wittes, Detre, & Holford, 1993) . Indeed, Japanese trials were found to be of significantly lower quality compared to their European and North American counterparts and further, the effect estimates from the included Japanese trials were slightly higher, in most cases, compared to Europe and North America.

However, it should be noted that as this evaluation was only able to assess trial quality based on information provided in trial reports, quality assessment may reflect continental differences in the quality of reporting randomised controlled trials rather than indicating the quality of the design and implementation of the trials themselves. Juni et al. (2001), however, suggest that trials that do not report randomisation or blinding methods, do not do so because

the methods for these are inadequate. In short, poorly reported trials indicate poor quality methods. This may, therefore, explain the results found in this study, where Japanese trials reported more positive results and appeared to be of poorer quality.

Nonetheless, publication bias may also explain the high proportion of positive Japanese trial results found, with trials showing positive results being more likely to be published in this country. Indeed, statistical analysis showed that publication bias was present in the Japanese sample of trials but not present in the European or North American sample.

However, it plausible that factors not investigated in this study also contributed towards international differences in treatment effectiveness. Selection bias, for instance, may explain some of the intercontinental differences in treatment effect because more negative trials from Japan may have been indexed in journal databases not used by the authors of the systematic reviews included in this study. Given that Medline and the systematic reviews included in The Cochrane Library were used to identify systematic reviews with meta-analyses for this study, and that they each draw from a broad spectrum of journals, it is unlikely that the results are due to selection bias of this kind. However, it may be the case that these findings are due to Japanese researchers being reluctant to publish negative findings. This would create an abundance of positive Japanese trial results in the public domain that would then be used in meta-analytic studies and therefore included in a study such as this.

Issues around patient compliance with regimens could also offer an explanation. Indeed, Bleyer et al. (1999) suggest that international differences in the rates of compliance could greatly change the effectiveness of treatment, with patients who are more compliant having a more positive outcome from the intervention.

With several factors possibly contributing to these findings, further exploration is required to ascertain how much of the difference in inter-continental treatment effect can be explained by trial quality alone, other possible factors or a combination of these.

8.1.2: Types of Intervention and Fatal Endpoints

From the investigation conducted over and within meta-analyses, the types of intervention prone to inter-continental differences in treatment effect could not be identified for either pair-wise comparison. However, the analysis conducted across the individual trials found that medical devices showed a statistically significant difference in treatment effect between both Europe and Asia and North America and Asia. This suggested that medical devices may be more likely to show international differences in treatment effect.

There may be many reasons for this finding. The difference in treatment effect may be more prominent for medical devices because of regional differences in how the intervention was implemented which, in turn, impact upon effectiveness. Drummond et al. (2009) state that the effectiveness of a medical device depends on how the device works and also on how it is used. This is different to the effectiveness of a drug that solely depends, as long as it is given at the correct dose, route and so on, on the drug itself. Furthermore, continental differences in the training for medical devices may explain the prominent difference in effect found for this intervention type. For instance, in the continent that provides the most training the person implementing the device will have more experience and this may, therefore, impact on the effectiveness of the device and contribute to the large between-continent differences found in this study.

8.1.3: Non-Fatal Endpoints

For non-fatal endpoints, there was some evidence that differences existed in treatment effectiveness between continents. The pair-wise comparison between Europe and North America showed that interventions performed better, relative to controls, in Europe. This being a statistically significant result. The results from the pair-wise comparisons between Europe and Asia and North America and Asia, meanwhile, showed that interventions performed better, relative to controls, in Asia compared to both Europe and North America, with both comparisons being statistically significant. These results were supported by the findings of the universal comparison in which the Independent sample t-test for North America versus Asia showed a statistically significant difference in treatment effect. However, no evidence was found for a difference in treatment effect between Europe and Asia since the Independent sample t-test for this comparison was not statistically significant at the 5% level.

The reasons for the above findings may be complex and chance does not offer a reasonable explanation. However, it is possible that the differences for this type of endpoint are the consequence of measurement bias, whereby an investigator's measurement is affected by their knowledge of treatment assignment. This is particularly important for subjective outcomes (i.e. non-fatal endpoints) as research suggests that knowledge of treatment assignment exaggerates the effect of treatment by 30% when subjective outcomes, compared to objective outcomes, are assessed (Wood et al. 2008). It is possible, therefore, that trial management in the regions explored for this study are different and that investigators involved in Asian trials are more likely to be aware of treatment assignment than their

European and North American counterparts, and therefore, explain the positive treatment effect estimates for Asian trials.

As with the fatal endpoint data, other factors could explain these differences, including publication bias, whereby trials where an intervention is found to be more effective than the control are more likely to be published. However, if this is the case for Asian studies, it is likely that it is also the case for non-Asian trials and as such, the effect on the findings of this study is likely to be minimal.

8.1.4: Types of Intervention and Non-Fatal Endpoints

For non-fatal endpoints the types of intervention prone to inter-continental differences could not be identified for any of the pair-wise comparisons. This was because, from the analysis across and within meta-analyses, no clear pattern of statistically significant differences was identified for any type of intervention. The investigation conducted across the individual trials provided similar results, with none of the t-tests conducted for each intervention type showing statistically significant results. Since the types of CVD intervention likely to show differences in treatment effectiveness could not be identified, this issue will need to be explored through further research.

8.2 Relationship with Existing Literature

Although no study was identified that directly assessed differences in treatment effectiveness between continents or countries, several studies have highlighted an interaction between country and treatment effect (Christensen, Broderick, Vincent, Morris, & Steiner. 2009; O'Shea & DeMets, 2001; Wolfe et al. 1999) which the results of this study can support. Furthermore, the findings presented in this thesis suggested that trials from Asia tended to

produce more positive results (i.e. results that favoured the intervention) when compared to other continents; a finding consistent with recent research conducted by Zhang et al. (2011) and confirmed by the findings of Pan et al. (2005) and Vickers et al. (1998).

However, this thesis has provided additional information, showing that this difference in the proportion of positive trial results can be explained, in part, by studies conducted in Japan reporting more positive results of interventions for cardiovascular disease. A finding that contrasts with that found by Zhang et al. (2011) and Pan et al. (2005), who suggested the differences were attributable to Chinese trials. As such, this country based analysis needs further investigation to ascertain which countries are more likely to produce these positive trial results.

As discussed in Chapter 7, trial quality may explain some of these differences between continents. Studies focusing on trial quality have shown that Asian trials are of low quality and require improvement (Zhang et al, 2011); a finding echoed in this thesis. In addition, Moher et al. (1999) found that low trial quality tends to exaggerate treatment effect and the findings from this thesis could be seen to support this suggestion because Japanese trials were of lower quality and produced more positive results than their European and North America counterparts. However, Chang et al. (2005), in their analysis of inter-country differences in the outcomes of acute coronary syndrome, demonstrated that such differences could mainly be accounted for by patient level factors. Therefore, more in-depth investigation is needed to identify the main factors creating these apparent international differences in treatment effectiveness.

However, as also suggested in Chapter 7, these findings may reflect not just trial quality but rather the quality of reporting trials. Although studies into trial quality have been conducted for the past 30 years, it is disappointing to find that essential trial information, for example, methods of randomisation, is still not necessarily reported adequately. The Zhang et al. (2011) study found that over half of Asian trials failed to report what methods they had used to randomise patients, a finding that is consistent with those found in this panoramic meta-analysis. Although studies suggest that quality of trial reporting is determined by trial quality, (Juni, Altman, & Egger, 2001; Liberati, Himel, & Chalmers, 1986; Schulz et al. 1995), it is not certain that for the trials included in this panoramic meta-analysis, this is the case. Further research is needed to ascertain the degree to which the quality of reporting trials reflects country or region specific norms, rather than trial quality itself.

This thesis has also provided additional information by showing that trial quality is associated with exaggerated treatment effects when objective outcomes are assessed. Previous research had shown this only to be the case when subjective outcomes are considered (Wood, Egger, Gluud et al. 2008), finding no evidence of exaggerated treatment effects resulting from poor quality trials in the assessment of objective outcomes. However, the difference between the results of this thesis and the study by Wood et al. may be due to Wood et al. (2008) not accounting for bias due to attrition which might have confounded assessment of the effects of inadequate allocation concealment and lack of blinding in their study. Furthermore, while this thesis is based on data solely from cardiovascular trials, the work by Wood et al. relates to several conditions and this may account for the difference in the findings between these two studies with regards to objective outcomes.

8.3: Limitations and Strengths

This section will now discuss the strengths and limitations of this study.

8.3.1: Limitations

While all efforts were taken to minimise them, there are still some limitations to this study. The literature search was conducted using a comprehensive search strategy in order to minimise the risk of missing relevant systematic reviews containing meta-analyses. However, this may not have located all published and unpublished systematic reviews and so, some relevant data may have been excluded from analysis.

The search for relevant systematic reviews was only conducted in two databases: Medline and The Cochrane Library and again, may have resulted in relevant systematic reviews not being included in the analysis for this study. However, given the number of systematic reviews that were found to overlap between Medline and The Cochrane Library, and the fact that these databases are commonly viewed as the main listings for this type of study in this subject area, it is unlikely that searching other databases would have identified many additional relevant systematic reviews containing meta-analyses. Furthermore, it is argued that Medline provides an adequately representative sample of the best quality trials and systematic reviews (Vickers et al. 1998) and that, therefore, it is unlikely that the sample analysed in this dataset is unrepresentative of relevant systematic reviews.

Another limitation of this study was that it only included systematic reviews published in English. This may create a language bias since systematic reviews not published in the English language would not have been identified, and some relevant trial data may therefore have been missed and excluded from the analysis. However, those searching and assessing

reviews for inclusion were not competent in other languages and no translator was available to assist the researchers.

The inclusion of overlapping meta-analyses could also be considered a limitation of this study. Some of the systematic reviews included were answering the same clinical question and, therefore, contained several of the same trials. This meant that when meta-analyses were pooled, some trial results were duplicated, and resulted in exaggerated results as was evident in the pair-wise comparisons between Europe and North America for fatal endpoints and North America and Asia for non-fatal endpoints. However, as panoramic meta-analysis is based upon meta-analyses, the inclusion of duplicated trial results is inevitable and, therefore, should be accounted for in any analyses using this approach. Furthermore, the data set for this study would have been too small if overlapping meta-analyses had been excluded from the analysis, meaning that the detection of inter-continental differences in treatment effectiveness would not have been possible.

This study did not attempt to extract patient level information from multinational trials that have arms in Europe, North America and Asia. Analysis involving this type of data could have increased statistical power, enabled a more flexible analysis of patient subgroups and outcomes (Clarke & Stewart, 2001) and may have allowed a more thorough investigation of inter-continental differences in treatment effectiveness. However, while researchers conducting multinational trials might be expected to monitor findings for country differences in treatment effect, these are rarely published. Consequently, obtaining patient level information is a difficult and time consuming task, which is not always possible.

No formal assessment of the quality of the systematic reviews containing meta-analyses was conducted as part of this study because it was initially envisaged that analysis would only pull data from the trials contained in meta-analyses (which were quality assessed see Chapter 7) rather than use the meta-analytic data itself. However, as The Cochrane Collaboration has high quality requirements for systematic reviews, it is likely that any review and meta-analysis linked to a Cochrane review and found on their database will be of high quality. Should further panoramic meta-analyses, using meta-analysis level data, be conducted, and to counter concerns that existing tools are still as yet 'incomplete or inconsistent' (Shea et al. 2000), a comprehensive tool could be developed, bringing together elements of existing tools, to help address all components of quality.

A limitation of this study was the use of the Jadad scale to assess trial quality. This scale places more emphasis on the quality of reporting than it does on the actual methodological quality of the trial. For instance, a trial reporting patient attrition rates will earn one point, irrespective of whether data was analysed using the intention to treat method. This, therefore, means that quality assessment of trials using this scale can only hint at the methodological quality of trials. Furthermore, a Jadad score of three, indicates a trial of high quality (Juni, Altman, & Egger, 2001). However, a trial can score three points even when it does not use random allocation or conceal allocation procedures from patients. Therefore, trial quality scores obtained using this scale may not adequately reflect the methodological quality of trials. In addition, this scale places considerable importance on blinding, even though blinding is not always feasible. Consequently, trials about, for example, surgical procedures, that are methodologically sound, may score low on quality because blinding is not possible.

However, while there is no "gold standard" scale to evaluate methodological quality of trials (Towheed, 2006) and none of the scales at the time of this study could be recommended without reservations (Higgins, 2010; Higgins & Green, 2005), the Jadad scale was the only one that had been constructed according to psychometric principles. As such, the Jadad scale, for this study, had to be the choice instrument for assessing trial quality.

8.3.2: Strengths

This study has a number of strengths. It presents the first empirical, direct, investigation into the existence of inter-continental differences in treatment effect between Europe, North America and Asia and is also the first study to investigate directly the association between these differences and trial quality.

Using the novel panoramic meta-analytic approach, which has its foundations in the meta-analytic approach; it has all of the strengths of this method. First, it followed a structured methodology involving meticulous review and analysis of all trials contained in the included meta-analyses. This overcame the biases that can be associated with basing conclusions on only a few trials and permitted the validity of the studies included to be determined. Second, this method permitted small trials, or those with non-significant effects, to be included in the analysis so that they contributed to the overall results of this study. In doing so, little data was wasted and the impact of bias that could have resulted from not including such trials (that could have distorted the results of this study) was reduced.

The panoramic meta-analytic approach, however, enhances the strengths of meta-analysis in that it is more efficient for data collection. By identifying relevant studies through published meta-analyses and directly extracting data from them, large amounts of data can be collated

in a timely manner. In doing this, a vast amount of data can be included, which adds validity and increases generalisability of the findings.

In addition, the panoramic meta-analytic approach builds on meta-analysis by comparing results across studies and across meta-analyses, thus providing a robust way of testing whether a particular variable can explain between-study heterogeneity. This was important since this study aimed to investigate 'continent' as a component of between-study heterogeneity over different types of interventions for different, but related, medical conditions. This is the first study to used meta-analytic techniques to investigate this issue.

As part of the panoramic meta-analysis, this study conducted a detailed search to determine the countries and continents where each trial originated. This search involved the examination of the original trial reports and entailed contacting authors to obtain country and continent information. Without this, a thorough examination of inter-continental differences in treatment effectiveness could not have been conducted, as details such as these are rarely presented in systematic reviews themselves. This thorough examination about the countries in which trials were conducted, has not been done before.

8.4: Future Research

More work remains to be done in order to provide a comprehensive understanding of international differences in treatment effectiveness. This study focused solely on cardiovascular diseases. The differences found in treatment effectiveness for this group of conditions, however, highlight the need for comparative work on other conditions and groups of conditions to ascertain whether these differences are particular to CVDs or are more widely applicable.

Furthermore, this study only compared trials from Europe, North America and Asia, and, as such, comparative work needs to be conducted to determine whether other continents show similar differences. In particular, as this study did not include data from studies conducted in the Middle East (see Chapter 4), future research could look at whether differences exist between trials conducted in the Middle East and those conducted in the rest of Asia. Including studies from these additional regions would allow international differences to be observed on the world stage and the relative size of international differences in treatment effect to be explored. This would make those wanting to extrapolate overseas clinical data to help make decisions about approving interventions or for making guideline recommendations, more aware of which trial results they can place confidence in.

The types of intervention likely to show inter-continental differences in effectiveness also require further examination. Although, this research has reported some novel findings about how different intervention types perform across the comparison continents, a more in-depth understanding is needed to allow clear decisions to be made about extrapolating data about particular intervention types between continents. For instance, research could be conducted into other types of intervention, such as management strategies.

The original intention of this study was to compare treatment effectiveness between countries, as explained in Chapter 4. However, a lack of country-level data included in meta-analyses meant that this was not possible. As the importance of country-level data on treatment effect has been shown in this research, namely by the fact that studies based in Japan show more positive results than those from Europe and North America, comparisons at the country-level is essential.

To be able to follow-up this research question, an additional research problem needs to be addressed that is both a research question and a research practice recommendation. Future research needs to identify methods by which country-level data can be collated and researchers need to provide more detail when reporting trials. For example, those conducting trials or systematic reviews need to be more explicit about where trials were carried out and those conducting multinational trials need to be prepared to release country-level breakdowns of their results.

Further research is also needed around trial quality, particularly about the impact of trial quality on treatment effect estimates and about methods that allow quality of trials to be assessed directly so that those using the data can be sure this reflects trial quality rather than quality of reporting. While trial quality assessments may be stipulated in the protocols of some RCTs, a wider application of this approach is required and should be used as a matter of course. Future research could also utilise several different trial quality checklists and scales to explore the association between inter-continental differences in treatment effect and trial quality. In particular, the risk of bias tool developed by The Cochrane Collaboration in 2008[59] could be utilised. This tool was developed to distinguish between the actual research methods used in a study and those that the study reports. Therefore, it is a more robust tool and ideal for an investigation into trial quality differences and their impact on treatment effect.

[59] This tool was revised in 2010 after an evaluation project. In 2008 it had not been evaluated for use in systematic reviews and meta-analyses and so was not available for use in this study.

8.5: Empirical Recommendations

The results of this study have provided some evidence that international differences in treatment effectiveness exist but that they may be explained, for fatal endpoints, by continental differences in the quality of trials. Below are the recommendations that can be made on the basis of these findings.

Recommendation 1: Caution should be used when extrapolating clinical data between continents. It is recommended that countries should be cautious when using overseas clinical data as the basis for their local guideline recommendations or in approving new interventions locally. This thesis found clear differences between continents for both fatal and non-fatal endpoints.

Recommendation 2: Trial quality should be taken into account when extrapolating clinical data. It is recommended that when extrapolating and interpreting clinical data for the purpose of approving a new intervention or as the basis of guideline recommendations, trial quality is taken into account. This is because low quality trials can introduce a bias into the purported effectiveness of interventions.

Recommendation 3: Those conducting trials and systematic reviews should be more explicit about where trials are conducted and release country breakdowns of their results. From this study it is recommended that those who are conducting trials and systematic reviews should be explicit about where trials were conducted. It also recommends that those involved in multinational trials should release the country breakdowns of their results. In following these recommendations a more accurate measurement of differences in the trial results of individual countries could be carried out.

8.6: Methodological Recommendations

Several recommendations can be made about the use of panoramic meta-analysis.

Recommendation 1: Those using panoramic meta-analysis should take into account the quality not only of included trials but also of included meta-analyses. In doing this, researchers can ensure that any analysis has its foundations in the best available evidence. Furthermore, the impact of low quality trials and meta-analyses on their findings can be investigated. This could be done by developing a comprehensive tool specifically for panoramic meta-analysis based on the checklists and scales that already exist, for example, the QUORUM statement (Moher et al. 1999).

Recommendation 2: The impact of including trial duplications from overlapping meta-analyses should always be investigated. This is because it has the potential to impact on estimates of treatment effect.

Recommendation 3: Where possible, individual patient data should be obtained. This would give the approach more statistical power and enable more flexible sub-group analyses to be conducted.

8.7: Conclusion

This study has provided some evidence that international differences in the effectiveness of interventions for treating and preventing cardiovascular diseases exist for both fatal and non-fatal endpoints. For fatal endpoints, this study has highlighted that such differences may be explained by continental differences in trial quality. For non-fatal endpoints, on the other hand, such differences may be explained by their subjective nature.

Although this study naturally has its weaknesses, it is the first, to directly investigate the existence of inter-continental differences in treatment effectiveness. As such, this study's findings, while echoing those of other studies that did not specifically aim to answer this question, provides the groundwork on which future research about inter-country, inter-region, and inter-continental differences in treatment effectiveness, trial quality and quality of trial reporting can be based.

This research has used a new methodological approach – panoramic meta-analysis – an approach to meta-analysis that allows large amounts of data about many different, but related issues, to be collated and analysed. This thesis, therefore, constitutes a foundation on which to build future research using this unique methodology.

Many lessons can be learnt from the findings presented here that can be applied to the development of guidelines and treatment approval processes. It is hoped that by providing a better understanding of inter-continental differences, improvements can be made in the way trial data is extrapolated between continents and in the quality of reporting of randomised controlled trials. Furthermore, it is hoped, on the basis of this study's findings, improvements can be made to the evidence used as the foundation for guideline recommendations and that this will impact on the quality of support, services and interventions patients with CVD receive and the outcomes they are able to achieve.

Appendix 1.

Trial Quality Questionnaire

Trial Authors: **Year:** **Journal:**

		Jadad Questions	Quality Scores
1) Was the trial multi-centred?	Y/N		
Randomisation:			
1) Is the study described as randomised?	Y/N	J	
2) Is the randomisation technique appropriate i.e. not using alternative patients?	Y/N/Not Stated*	J	
3) Was randomisation done by a third party i.e. it was not within the control of a given clinic? (Computer randomised trial centre, Telephone, Web Based)	Y/N/ /Not stated*		
4) Was randomisation done by envelopes within the local setting?	Y/N/Not stated*		
Blinding:			
1) Does the trial specify that blinding was used?	Y/N	J	

2) If not was a placebo used? 3)	Y/N/Not Stated*		
4) Was the patient blinded?	Y/N/Not Stated*		
5) Was the investigator blinded?	Y/N/Not Stated* / NA		
6) If the answer to 3) and 4) was "Yes" then was the blinding method described?	Y/N	J	
7) Was the main outcome considered objective (such as death) rather than subjective (such as pain)?	Y/N		
Withdrawals and Drop-outs:			
1) Was the number of drop-outs between randomisation and the assessment of outcome given by group?	Y/N	J	
2) Did the trial follow a CONSORT-type detailed breakdown of both withdrawals and drop-outs following randomisation?	Y/N		
Total Quality Score			
Comments:			

*** = Also means Unclear**

J = Indicates a question that is to be answered for the Jadad Scale.

Instructions:

In order to provide the trials with a Jadad score the following items need to be scored:

1) **If the trial was described as randomised, then 1 point is awarded.**
2) **If the study was described as blinded, then 1 point is awarded.**
3) **If a description of withdrawals and drop-outs was given, then 1 point is awarded.**
4) **If the method used to generate randomisation was described and it was appropriate, then 1 point is awarded.**
5) **If the method of double blinding was described and it was appropriate, then 1 point is awarded.**

The maximum score possible is 5 while the lowest score is 0. A score of 5 indicates that the quality of the trial is high.

Appendix 2.

Meta-regression Output.

Meta-regression 1 – Region

```
Meta-regression                                  Number of obs  =     33
REML estimate of between-study variance          tau2           =  .1248
% residual variation due to heterogeneity        I-squared_res  =  8.57%
Proportion of between-study variance explained   Adj R-squared  = 32.32%
Joint test for all covariates                    Model F(2,30)  =   3.29
With Knapp-Hartung modification                  Prob> F        = 0.0512
------------------------------------------------------------------------------
logTE |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 _Iregion2_2 |  -.5773012   .3674232    -1.57   0.127     -1.32768     .173077
 _Iregion2_3 |   .6694331   .4119743     1.62   0.115    -.1719307    1.510797
       _cons |  -.3150137   .1835637    -1.72   0.096    -.6899007    .0598733
------------------------------------------------------------------------------
```

```
. test _Iregion2_2= _Iregion2_3

( 1)  _Iregion2_2 - _Iregion2_3 = 0

F(  1,    30) =    6.55
Prob> F =    0.0158
```

Meta-regression 2 – Jadad Score after adjusting for Region

```
Meta-regression                                 Number of obs =     33
REML estimate of between-study variance         tau2          =   .1794
% residual variation due to heterogeneity       I-squared_res =  11.58%
Proportion of between-study variance explained  Adj R-squared =   2.65%
Joint test for all covariates                   Model F(3,29) =    2.27
With Knapp-Hartung modification                 Prob> F       =  0.1012
------------------------------------------------------------------------------
logTE |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      _Iregion2_2 |  -.2827954   .5170503    -0.55   0.589    -1.340282    .7746912
      _Iregion2_3 |   .8002819   .4594369     1.74   0.092    -.1393721    1.739936
jadadscore |   .1316464    .157745     0.83   0.411    -.1909782    .4542711
      _cons |  -.8547196   .6639058    -1.29   0.208    -2.212559    .5031202
------------------------------------------------------------------------------
```

Meta-regression 3 – Jadad Score

```
Meta-regression                                 Number of obs =     33
REML estimate of between-study variance         tau2          =   .2446
% residual variation due to heterogeneity       I-squared_res=  22.98%
Proportion of between-study variance explained  Adj R-squared  = -32.68%
With Knapp-Hartung modification
------------------------------------------------------------------------------
logTE |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
jadadscore |   .1812684   .1249541     1.45   0.157    -.0735772    .4361141
      _cons |  -.9520952   .4422656    -2.15   0.039    -1.854102   -.0500886
------------------------------------------------------------------------------
```

Bibliography

Aasbo, J. D., Lawrence, A. T., Krishnan, K., Kim, M. H., & Trohman, R. G. 2005, "Amiodarone Prophylaxis Reduces Major Cardiovascular Morbidity and Length of Stay after Cardiac Surgery: A Meta-Analysis", *Annals of Internal Medicine*, vol. 143, no. 5, pp. 327-336.

Abbate, A., Biondi-Zoccai, G. G. L., Appleton, D. L., Erne, P., Schoenenberger, A. W., Lipinski, M. J., Agostoni, P., Sheiban, I., & Vetrovec, G. W. 2008, "Survival and Cardiac Remodeling Benefits in Patients Undergoing Late Percutaneous Coronary Intervention of the Infarct-Related Artery: Evidence From a Meta-Analysis of Randomized Controlled Trials", *Journal of the American College of Cardiology*, vol. 51, no. 9, pp. 956-964.

Abdel-Latif, A., Bolli, R., Tleyjeh, I. M., Montori, V. M., Perin, E. C., Hornung, C. A., Zuba-Surma, E. K., Al-Mallah, M., & Dawn, B. 2007, "Adult Bone Marrow-Derived Cells for Cardiac Repair: A Systematic Review and Meta-analysis", *Archives of Internal Medicine*, vol. 167, no. 10, pp. 989-997.

Abdel-Latif, A., Bolli, R., Zuba-Surma, E. K., Tleyjeh, I. M., Hornung, C. A., & Dawn, B. 2008, "Granulocyte colony-stimulating factor therapy for cardiac repair after acute myocardial infarction: A systematic review and meta-analysis of randomized controlled trials", *American Heart Journal*, vol. 156, no. 2, pp. 216-226.

Abdulla, J., Haarbo, J., Køber, L., & Torp-Pedersen, C. 2006, "Impact of Implantable Defibrillators and Resynchronization Therapy on Outcome in Patients with Left Ventricular Dysfunction – A Meta-Analysis.", *Cardiology*, vol. 106, pp. 249-255.

Abdulla, J., Barlera, S., Latini, R., Kjoller-Hansen, L., Sogaard, P., Christensen, E., Kober, L., & Torp-Pedersen, C. 2007, "A systematic review: Effect of angiotensin converting enzyme inhibition on left ventricular volumes and ejection fraction in patients with a myocardial infarction and in patients with left ventricular dysfunction", *European Journal of Heart Failure*, vol. 9, no. 2, pp. 129-135.

Abdulla, J., Pogue, J., Abildstrom, S. Z., Kober, L., Christensen, E., Pfeffer, M. A., Yusuf, S., & Torp-Pedersen, C. 2006, "Effect of angiotensin-converting enzyme inhibition on functional class in patients with left ventricular systolic dysfunction - a meta-analysis", *European Journal of Heart Failure*, vol. 8, no. 1, pp. 90-96.

Abramson, J., & Abramson, Z. 1999. *Survey Methods in Community Medicine: Epidemiological research, Programme evaluation, Clinical trials.* London: Churchill Livingstone. pp. 58.

Abubakar, I., Kanka, D., Arch, B., Porter, J., & Weissberg, P. 2004, "Outcome after acute myocardial infarction: a comparison of patients seen by cardiologists and general physicians", *BMC Cardiovascular Disorders*, vol. 4, no. 1, p. 14.

AFIB Investigators 1997, "Treatment of Atrial Fibrillation and Paroxysmal Supraventricular Tachycardia With Bidisomide", *Circulation*, vol. 96, no. 8, pp. 2625-2632.

Agostoni, P., Biondi-Zoccai, G. G. L., Gasparini, G. L., Anselmi, M., Morando, G., Turri, M., Abbate, A., McFadden, E. P., Vassanelli, C., Zardini, P., Colombo, A., & Serruys, P. W. 2005, "Is bare-metal stenting superior to balloon angioplasty for small vessel coronary artery disease? Evidence from a meta-analysis of randomized trials", *European Heart Journal*, vol. 26, no. 9, pp. 881-889.

Akkerhuis, K. M., Deckers, J. W., Boersma, E., Harrington, R. A., Stepinska, J., Mahaffey, K. W., Wilcox, R. G., Lincoff, A. M., Keltai, M., Topol, E. J., Califf, R. M., & Simoons, M. L. 2000, "Geographic variability in outcomes within an international trial of glycoprotein IIb/IIIa inhibition in patients with acute coronary syndromes. Results from PURSUIT", *European Heart Journal*, vol. 21, no. 5, pp. 371-381.

Alexander, R. A., Scozzaro, M. J., & Borodkin, L. J. 1989, "Statistical and empirical examination of the chi-square test for homogeneity of correlations in meta-analysis", *Psychological Bulletin*, vol. 106, no. 2, pp. 329-331.

Ali, A., Cox, D., Dib, N., Brodie, B., Berman, D., Gupta, N., Browne, K., Iwaoka, R., Azrin, M., Stapleton, D., Setum, C., & Popma, J. 2006, "Rheolytic Thrombectomy With Percutaneous Coronary Intervention for Infarct Size Reduction in Acute Myocardial Infarction. 30-Day Results From a Multicenter Randomized Study", *Journal of the American College of Cardiology*, vol. 48, no. 2, pp. 244-252.

Allison, J. J., Kiefe, C. I., Weissman, N. W., Person, S. D., Rousculp, M., Canto, J. G., Bae, S., Williams, O. D., Farmer, R., & Centor, R. M. 2000, "Relationship of Hospital Teaching Status With Quality of Care and Mortality for Medicare Patients With Acute MI", *JAMA: The Journal of the American Medical Association*, vol. 284, no. 10, pp. 1256-1262.

Almqvist, C., Pershagen, G., & Wickman, M. 2005, "Low socioeconomic status as a risk factor for asthma, rhinitis and sensitization at 4 years in a birth cohort", *Clinical & Experimental Allergy*, vol. 35, no. 5, pp. 612-618.

Altman, D. G. & Bland, J. M. 1999, "How to randomise", *BMJ*, vol. 319, no. 7211, pp. 703-704.

Amsallem, E., Kasparian, C., Haddour, G., Boissel, J., & Nony, P. 2005, "Phosphodiesterase III Inhibitors for Heart Failure.", *Cochrane Database of Systematic Reviews* no. 1.

Anderson, J. J., Wells, G., Verhoeven, A. C., & Felson, D. T. 2000, "Factors predicting response to treatment in rheumatoid arthritis: The importance of disease duration", *Arthritis & Rheumatism*, vol. 43, no. 1, pp. 22-29.

Anderson, T. & Finn, J. 1996, "Measures of Variability," in *The New Statistical Analysis of Data*, T. Anderson & J. Finn, eds., Springer-Verlag, New York, p. 113.

Andraws, R., Berger, J. S., & Brown, D. L. 2005, "Effects of Antibiotic Therapy on Outcomes of Patients With Coronary Artery Disease", *JAMA: The Journal of the American Medical Association*, vol. 293, no. 21, pp. 2641-2647.

Andreotti, F., Testa, L., Biondi-Zoccai, G. G. L., & Crea, F. 2006, "Aspirin plus warfarin compared to aspirin alone after acute coronary syndromes: an updated and comprehensive meta-analysis of 25ΓÇà307 patients", *European Heart Journal*, vol. 27, no. 5, pp. 519-526.

Appleton, D. L., Abbate, A., & Biondi-Zoccai, G. G. L. 2008, "Late percutaneous coronary intervention for the totally occluded infarct-related artery: A meta-analysis of the effects on cardiac function and remodeling", *Catheterization and Cardiovascular Interventions*, vol. 71, no. 6, pp. 772-781.

Ashcroft, R. E. 2004, "Current epistemological problems in evidence based medicine", *Journal of Medical Ethics*, vol. 30, no. 2, pp. 131-135.

Aung, K. & Htay, T. 2005, "Vasopressin for Cardiac Arrest: A Systematic Review and Meta-analysis", *Archives of Internal Medicine*, vol. 165, no. 1, pp. 17-24.

Aziz, O., Rao, C., Panesar, S. S., Jones, C., Morris, S., Darzi, A., & Athanasiou, T. 2007, "Meta-analysis of minimally invasive internal thoracic artery bypass versus percutaneous revascularisation for isolated lesions of the left anterior descending artery", *BMJ*, vol. 334, no. 7594, p. 617.

Bagshaw, S. M., Galbraith, P. D., Mitchell, L. B., Sauve, R., Exner, D. V., & Ghali, W. A. 2006, "Prophylactic Amiodarone for Prevention of Atrial Fibrillation After Cardiac Surgery: A Meta-Analysis", *The Annals of Thoracic Surgery*, vol. 82, no. 5, pp. 1927-1937.

Baker, W. L. & Couch, K. A. 2007, "Azithromycin for the secondary prevention of coronary artery disease: A meta-analysis", *American Journal of Health-System Pharmacy*, vol. 64, no. 8, pp. 830-836.

Balk, E. M., Bonis, P. A. L., Moskowitz, H., Schmid, C. H., Ioannidis, J. P. A., Wang, C., & Lau, J. 2002, "Correlation of Quality Measures With Estimates of Treatment Effect in Meta-analyses of Randomized Controlled Trials", *JAMA: The Journal of the American Medical Association*, vol. 287, no. 22, pp. 2973-2982.

Barth, J., Critchley, J., & Bengel, J. 2008, "Psychosocial interventions for smoking cessation in patients with coronary heart disease", *Cochrane Database of Systematic Reviews* no. 1.

Bassler, D., Briel, M., Montori, V. M., Lane, M., Glasziou, P., Zhou, Q., Heels-Ansdell, D., Walter, S. D., Guyatt, G. H., & Study Group 2010, "Stopping Randomized Trials Early for Benefit and Estimation of Treatment Effects", *JAMA: The Journal of the American Medical Association*, vol. 303, no. 12, pp. 1180-1187.

Bastian, H., Glasziou, P., & Chalmers, I. 2010b, "Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up?", *PLoS Med*, vol. 7, no. 9, p. e1000326.

Bavry, A. A., Kumbhani, D. J., & Bhatt, D. L. 2008, "Role of adjunctive thrombectomy and embolic protection devices in acute myocardial infarction: a comprehensive meta-analysis of randomized trials", *European Heart Journal*, vol. 29, no. 24, pp. 2989-3001.

Bavry, A. A., Kumbhani, D. J., Helton, T. J., & Bhatt, D. L. 2005, "Risk of Thrombosis With the Use of Sirolimus-Eluting Stents for Percutaneous Coronary Intervention (from Registry and Clinical Trial Data)", *The American Journal of Cardiology*, vol. 95, no. 12, pp. 1469-1472.

Becker, L. & Oxman, A. 2009, "Overviews of reviews," in *Cochrane Handbook for Systematic Reviews of Interventions*, 5.0.02 edn, J. Higgins & S. Green, eds., The Cochrane Collaboration, Oxford.

Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K. F., Simel, D., & Stroup, D. F. 1996, "Improving the Quality of Reporting of Randomized Controlled Trials", *JAMA: The Journal of the American Medical Association*, vol. 276, no. 8, pp. 637-639.

Begg, C. B. & Mazumdar, M. /12, "Operating Characteristics of a Rank Correlation Test for Publication Bias", *Biometrics*, vol. 50, no. 4, pp. 1088-1101.

Bellomo, R., Warrillow, S. J., & Reade, M. C. 2009, "Why we should be wary of single-center trials", *Critical Care Medicine*, vol. 37, no. 12.

Belsey, J. What is Evidence Based Medicine? Evidence Based Medicine [2nd], 1-10. 2009. What is....? Series. 12-6-2010. Ref Type: Electronic Citation

Beran, G., Lang, I., Schreiber, W., Denk, S., Stefenelli, T., Syeda, B., Maurer, G., Glogar, D., & Siostrzonek, P. 2002, "Intracoronary Thrombectomy With the X-Sizer Catheter System Improves Epicardial Flow and Accelerates ST-Segment Resolution in Patients With Acute Coronary Syndrome: A Prospective, Randomized, Controlled Study", *Circulation*, vol. 105, no. 20, pp. 2355-2360.

Berger, J. S., Stebbins, A., Granger, C. B., Ohman, E. M., Armstrong, P. W., Van de Werf, F., White, H. D., Simes, R. J., Harrington, R. A., Califf, R. M., & Peterson, E. D. 2008, "Initial Aspirin Dose and Outcome Among ST-Elevation Myocardial infarction Patients Treated With Fibrinolytic Therapy", *Circulation*, vol. 117, no. 2, pp. 192-199.

Bergovec, M., Reiner, E., Milicic D., &Vrazic H. 2008, "Differences in risk factors for coronary heart disease in patients from continental and Mediterranean regions of Croatia", *Wiener Klinische Wochenschrift*, vol. 120, no. 21, pp. 684-692.

Berlin, J. A. & Rennie, D. 1999a, "Measuring the Quality of Trials", *JAMA: The Journal of the American Medical Association*, vol. 282, no. 11, pp. 1083-1085.

Berlin, J. A. 1995, "Invited Commentary: Benefits of Heterogeneity in Meta-analysis of Data from Epidemiologic Studies", *American Journal of Epidemiology*, vol. 142, no. 4, pp. 383-387.

Besselink, M. G. H., Verwer, T. J., Schoenmaeckers, E. J. P., Buskens, E., Ridwan, B. U., Visser, M. R., Nieuwenhuijs, V. B., & Gooszen, H. G. 2007, "Timing of Surgical Intervention in Necrotizing Pancreatitis", *Archives of Surgery*, vol. 142, no. 12, pp. 1194-1201.

Bhandari, M. & Joensson, A. 2009, "Meta-analysis," in *Clinical Research for Surgeons*, M. Bhandari & A. Joensson, eds., Thieme, New York, pp. 68-76.

Bleyer, A. J., Hylander, B., Sudo, H., Nomoto, Y., de la Torre, E., Chen, R. A., & Burkart, J. M. 1999b, "An International Study of Patient Compliance With Hemodialysis", *JAMA: The Journal of the American Medical Association*, vol. 281, no. 13, pp. 1211-1213.

Bleys, J., Miller, E. R., Pastor-Barriuso, R., Appel, L. J., & Guallar, E. 2006, "Vitamin-mineral supplementation and the progression of atherosclerosis: a meta-analysis of randomized controlled trials", *The American Journal of Clinical Nutrition*, vol. 84, no. 4, pp. 880-887.

Bonneux, L., Barendregt, J. J., Meeter, K., Bonsel, G. J., & van der Maas, P. J. 1994a, "Estimating clinical morbidity due to ischemic heart disease and congestive heart failure: the future rise of heart failure", *American Journal of Public Health*, vol. 84, no. 1, pp. 20-28.

Borenstein, M., Hedges, L., & Rothstein, H, 2007.*Introduction to Meta-analysis* [online]. Available at: www.Meta-Analysis.com [Accessed 07/06/2009].

Bosch, X., Loma-Osorio, P., & Marrugat, J. 2001, "Platelet glycoprotein IIb/IIIa blockers for percutaneous coronary intervention and the initial treatment of non-ST segment elevation acute coronary syndromes", *Cochrane Database of Systematic Reviews* no. 4.

Bosworth, H., Oddone, E., & Weinberger, M. 2005, *Patient Treatment Adherence: Concepts, Interventions and Measurement* Lawrence Erlbaum Associates Inc, Mahwah, NJ.

Bourgeois, F. T., Murthy, S., & Mandl, K. D. 2010, "Outcome Reporting Among Drug Trials Registered in ClinicalTrials.gov", *Annals of Internal Medicine*, vol. 153, no. 3, pp. 158-166.

Bowater, R. J., Stirling, S. A., & Lilford, R. J. 2009, "Is antibiotic prophylaxis in surgery a generally effective intervention? Testing a generic hypothesis over a set of meta-analyses. ", *Annals of Surgery*, vol. 249, no. 4, pp. 551-556.

Bowry, A. D. K., Brookhart, M. A., & Choudhry, N. K. 2008, "Meta-Analysis of the Efficacy and Safety of Clopidogrel Plus Aspirin as Compared to Antiplatelet Monotherapy for the Prevention of Vascular Events", *The American Journal of Cardiology*, vol. 101, no. 7, pp. 960-966.

Bravata, D. M., Gienger, A. L., McDonald, K. M., Sundaram, V., Perez, M. V., Varghese, R., Kapoor, J. R., Ardehali, R., Owens, D. K., & Hlatky, M. A. 2007, "Systematic Review: The Comparative Effectiveness of Percutaneous Coronary Interventions and Coronary Artery Bypass Graft Surgery", *Annals of Internal Medicine*, vol. 147, no. 10, pp. 703-716.

Brechner, R. J., Cowie, C. C., Howie, L. J., Herman, W. H., Will, J. C., & Harris, M. I. 1993, "Ophthalmic Examination Among Adults with Diagnosed Diabetes-Mellitus", *Jama-Journal of the American Medical Association*, vol. 270, no. 14, pp. 1714-1718.

Brevetti, G., Oliva, G., Giugliano, G., Schiano, V., De Maio, J., & Chiariello, M. 2007a, "Mortality in Peripheral Arterial Disease: A Comparison of Patients Managed by Vascular Specialists and General Practitioners", *Journal of General Internal Medicine*, vol. 22, no. 5, pp. 639-644.

British Heart Foundation, 2008. *Total costs of CVD in Europe.* British Heart Foundation Statistics Website "Heartstats" [online]. Available at: http://www.bhf.org.uk/heart-health/statistics.aspx [Accessed 09/11/09].

British Heart Foundation. 2011. *Death rates from CHD, stroke and all other diseases of the circulatory system, people aged under 75, 1970–2002, England.*British Heart Foundation Statistical Website "Heartstats" [online]. Available at: http://www.bhf.org.uk/heart-health/statistics/mortality.aspx [Accessed 04/02/11].

British Medical Journal, 2011. *Clinical Evidence Glossary.* BMJ Publishing Group Limited [online]. Available at: http://clinicalevidence.bmj.com/x/index.html [Accessed 04/02/11]

Broderick, J. P., Brott, T., Tomsick, T., Huster, G., & Miller, R. 2010, "The Risk of Subarachnoid and Intracerebral Hemorrhages in Blacks as Compared with Whites", *New England Journal of Medicine*, vol. 326, no. 11, pp. 733-736.

Burgers, J. S., Bailey, J. V., Klazinga, N. S., Van der Bij, A. K., Grol, R., & Feder, G. 2002e, "Inside Guidelines", *Diabetes Care*, vol. 25, no. 11, pp. 1933-1939.

Butkiewicz, E. K., Leibson, C. L., O'Brien, P. C., Palumbo, P. J., & Rizza, R. A. 1995, "Insulin therapy for diabetic ketoacidosis. Bolus insulin injection versus continuous insulin infusion", *Diabetes Care*, vol. 18, no. 8, pp. 1187-1190.

Cantor, W. J., Brunet, F., Ziegler, C. P., Kiss, A., & Morrison, L. J. 2005, "Immediate angioplasty after thrombolysis: a systematic review", *Canadian Medical Association Journal*, vol. 173, no. 12, pp. 1473-1481.

Chang, W. C., Midodzi, W. K., Westerhout, C. M., Boersma, E., Cooper, J., Barnathan, E. S., Simoons, M. L., Wallentin, L., Ohman, E. M., & Armstrong, P. W. 2005, "Are international differences in the outcomes of acute coronary syndromes apparent or real? A multilevel analysis", *Journal of Epidemiology and Community Health*, vol. 59, no. 5, pp. 427-433.

Chaudhry, R., Goel, V., & Sawka, C. 2001, "Breast cancer survival by teaching status of the initial treating hospital", *Canadian Medical Association Journal*, vol. 164, no. 2, pp. 183-188.

Cho, L. & Mukherjee, D. 2005, "Hormone Replacement Therapy and Secondary Cardiovascular Prevention: A Meta-Analysis of Randomized Trials.", *Cardiology*, vol. 104, no. 3, pp. 143-147.

Choudhry, N. K., Singh, J. M., Barolet, A., Tomlinson, G. A., & Detsky, A. S. 2005, "How should patients with unstable angina and non-ST-segment elevation myocardial infarction be managed? A meta-analysis of randomized trials", *The American Journal of Medicine*, vol. 118, no. 5, pp. 465-474.

Chow, S. & Liu, J. 2004, *Design and Analysis of Clinical Trials: Concepts and Methodologies.*, 2nd edn, Wiley, New Jersey.

Christensen, M. C., Broderick, J., Vincent, C., Morris, S., & Steiner, T. 2009, "Global Differences in Patient Characteristics, Case Management and Outcomes in Intracerebral Hemorrhage: The Factor Seven for Acute Hemorrhagic Stroke (FAST) Trial", *Cerebrovascular Diseases*, vol. 28, no. 1, pp. 55-64.

Christiaens, T., De Backer, D., Burgers, J., & Baerheim, A. 2004, "Guidelines, evidence, and cultural factors Comparison of four European guidelines on uncomplicated cystitis", *Scandinavian Journal of Primary Health Care*, vol. 22, no. 3, pp. 141-145.

Claridge, J. A. & Fabian, T. C. 2005, "History and Development of Evidence-based Medicine", *World Journal of Surgery*, vol. 29, no. 5, pp. 547-553.

Clark, A. M., Hartling, L., Vandermeer, B., Lissel, S. L., & McAlister, F. A. 2007, "Secondary prevention programmes for coronary heart disease: a meta-regression showing the merits of shorter, generalist, primary care-based interventions.", *European Journal of Cardiovascular Prevention & Rehabilitation.*, vol. 14, no. 4, pp. 538-546.

Clark, A. M., Hartling, L., Vandermeer, B., & McAlister, F. A. 2005, "Meta-Analysis: Secondary Prevention Programs for Patients with Coronary Artery Disease", *Annals of Internal Medicine*, vol. 143, no. 9, pp. 659-672.

Clark, R. A., Inglis, S. C., McAlister, F. A., Cleland, J. G. F., & Stewart, S. 2007, "Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis", *BMJ*, vol. 334, no. 7600, p. 942.

Clarke, M. & Stewart, L. 2001, "Obtaining Individual Patient Data From Randomised Controlled Trials," in *Systematic Reviews in Health Care: Meta-analysis in Context*, 2nd edn, M. Egger, G. Smith, & D. G. Altman, eds., BMJ, London, pp. 109-121.

Cohn, L. D., & Becker, B. J. 2003. "How Meta-Analysis Increases Statistical Power." *Psychological Methods*, vol.8, pp. 243-253.

Cohen, E. A., Lincoff, M., Sapp, S., Burton, J. R., Tcheng, J. E., & Topol, E. 2001a, "Variations in event rates between Canada and United States in the EPISTENT trial", *Journal of the American College of Cardiology*, vol. 37, no. 2, p. 81A.

Cohen, M. G., Pacchiana, C. M., Corbalβn, R. n., Isea Perez, J. s. E., Ponte, C. I., Oropeza, E. S., Diaz, R., Paolasso, E., Izasa, D., Rodas, M. A., Urrutia, C. E., Harrington, R. A., Topol, E. J., & Califf, R. M. 2001b, "Variation in patient management and outcomes for acute coronary syndromes in Latin America and North America: Results from the Platelet IIb/IIIa in Unstable Angina: Receptor Suppression Using Integrilin Therapy (PURSUIT) trial", *American Heart Journal*, vol. 141, no. 3, pp. 391-401.

Colditz, G. A., Burdick, E., & Mosteller, F. 1995a, "Heterogeneity in Meta-analysis of Data from Epidemiologic Studies: A Commentary", *American Journal of Epidemiology*, vol. 142, no. 4, pp. 371-382.

Collet, J. P., Montalescot, G., Le May, M., Borentain, M., & Gershlick, A. 2006, "Percutaneous Coronary Intervention After Fibrinolysis: A Multiple Meta-Analyses Approach According to the Type of Strategy", *Journal of the American College of Cardiology*, vol. 48, no. 7, pp. 1326-1335.

Cook, D.J., Mulrow, C, Haynes, R.B. 1997,"Systematic reviews: synthesis of best evidence for clinical decisions"*Annals of Internal Medicine*; vol. 126, pp. 376–80.

Cook, D. J., Reeve, B. K., Guyatt, G. H., Heyland, D. K., Griffith, L. E., Buckingham, L., & Tryba, M. 1996, "Stress Ulcer Prophylaxis in Critically Ill Patients", *JAMA: The Journal of the American Medical Association*, vol. 275, no. 4, pp. 308-314.

Cornwell, J. M. 1993, "Monte Carlo Comparisons of Three Tests for Homogeneity of Independent Correlations", *Educational and Psychological Measurement*, vol. 53, no. 3, pp. 605-618.

Cornwell, J. M. & Ladd, R. T. 1993, "Power and Accuracy of the Schmidt and Hunter Meta-Analytic Procedures", *Educational and Psychological Measurement*, vol. 53, no. 4, pp. 877-895.

Cowley, A. J. & Skene, A. M. 1994, "Treatment of severe heart failure: quantity or quality of life? A trial of enoximone. Enoximone Investigators", *British Heart Journal*, vol. 72, no. 3, pp. 226-230.

Cramer, J. A. 2002, "Effect of partial compliance on cardiovascular medication effectiveness", *Heart*, vol. 88, no. 2, pp. 203-206.

Crane, S. D., Elliott, M. W., Gilligan, P., Richards, K., & Gray, A. J. 2004, "Randomised controlled comparison of continuous positive airways pressure, bilevel non-invasive ventilation, and standard treatment in emergency department patients with acute cardiogenic pulmonary oedema", *Emergency Medicine Journal*, vol. 21, no. 2, pp. 155-161.

Critical Reviews Advisory Group, 1996,"Introduction to systematic reviews." School for Health and Related Research [online]. Available at: http://www-user.uni-bremen.de/~henning/Lehre/metaanalysen/Introduction%20_Systematic_Reviews.pdf [Accessed 15/03/2009].

Critical Appraisal Skills Programme. 2002, "10 questions to help you make sense of reviews." Milton Keynes Primary Care Trust [online]. Available at: http://www.phru.nhs.uk. [Accessed 01/03/2005].

Davidson, R. A. 1986, "Source of Funding and Outcome of Clinical-Trials", *Journal of General Internal Medicine*, vol. 1, no. 3, pp. 155-158.

De Luca, G., Gibson, C. M., Bellandi, F., Murphy, S., Maioli, M., Noc, M., Zeymer, U., Dudek, D., Arntz, H. R., Zorman, S., Gabriel, H. M., Emre, A., Cutlip, D., Biondi-Zoccai, G., Rakowski, T., Gyongyosi, M., Marino, P., Huber, K., & vanHof, A. W. J. 2008, "Early glycoprotein IIbΓÇôIIIa inhibitors in primary angioplasty (EGYPT) cooperation: an individual patient data meta-analysis", *Heart*, vol. 94, no. 12, pp. 1548-1558.

De Luca, G. & Marino, P. 2007, "Adjunctive benefits from low-molecular-weight heparins as compared to unfractionated heparin among patients with ST-segment elevation

myocardialinfarction treated with thrombolysis. A meta-analysis of the randomized trials", *American Heart Journal*, vol. 154, no. 6, p. 1085.

De Luca, G., Suryapranata, H., Ottervanger, J. P., & Antman, E. M. 2004, "Time Delay to Treatment and Mortality in Primary Angioplasty for Acute Myocardial Infarction: Every Minute of Delay Counts", *Circulation*, vol. 109, no. 10, pp. 1223-1225.

De Luca, G., Suryapranata, H., Stone, G. W., Antoniucci, D., Neumann, F. J., & Chiariello, M. 2007, "Adjunctive mechanical devices to prevent distal embolization in patients undergoing mechanical revascularization for acute myocardial infarction: A meta-analysis of randomized trials", *American Heart Journal*, vol. 153, no. 3, pp. 343-353.

De Luca, L., Sardella, G., Davidson, C. J., De Persio, G., Beraldi, M., Tommasone, T., Mancone, M., Nguyen, B. L., Agati, L., Gheorghiade, M., & Fedele, F. 2006, "Impact of intracoronary aspiration thrombectomy during primary angioplasty on left ventricular remodelling in patients with anterior ST elevation myocardial infarction", *Heart*, vol. 92, no. 7, pp. 951-957.

Deeks, J. & Altman, D. 2001, "Effect Measures for Meta-anlaysis of Trials with Binary Outcomes," in *Systematic Reviews in Health Care: Meta-anlaysis in Context*, 2nd edn, M. Egger, G. D. Smith, & D. G. Altman, eds., BMJ Publishing Group, London, pp. 313-335.

Deeks, J. J. 2002, "Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes", *Statistics in Medicine*, vol. 21, no. 11, pp. 1575-1600.

Deeks, J. J., Altman, D. G., & Bradburn M 2001, "Statistical Methods for Examining Heterogeneity and Combining Results from Several Studies in Meta-analysis.," in *Systematic Reviews in Health Care: Meta-analysis in Context*, 2nd edn, M. Egger, G. D. Smith, & D. G. Altman, eds., BMJ Publishing Group, London, pp. 285-312.

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. 2008, "Analysing and Presenting Results," in *Cochrane Reviewers Handbook*, 5.0.0 edn, J. P. T. Higgins & S. Green, eds., John Wiley and Sons LTD, Chichester.

DeMaeseneer J & Derese A 1999, "European general practice guidelines: a step too far?", *Eur J Gen Pract*, vol. 5, pp. 86-87.

Dentali, F., Douketis, J. D., Lim, W., & Crowther, M. 2007, "Combined Aspirin-Oral Anticoagulant Therapy Compared With Oral Anticoagulant Therapy Alone Among Patients at Risk for Cardiovascular Disease: A Meta-analysis of Randomized Trials", *Archives of Internal Medicine*, vol. 167, no. 2, pp. 117-124.

Derry, C., Derry, S., McQuay, H., & Moore, R. 2006, "Systematic review of systematic reviews of acupuncture published 1996-2005.", *Clinical Medicine*, vol. 6, no. 4, pp. 381-386.

Dickersin, K. 1990, "The Existence of Publication Bias and Risk Factors for Its Occurrence", *JAMA: The Journal of the American Medical Association*, vol. 263, no. 10, pp. 1385-1389.

Dickersin, K. & Berlin, J. A. 1992, "Meta-analysis: State-of-the-Science", *Epidemiologic Reviews*, vol. 14, no. 1, pp. 154-176.

Dobre, D., van Jaarsveld, C. H. M., deJongste, M. J. L., Haaijer Ruskamp, F. M., & Ranchor, A. V. 2007, "The effect of beta-blocker therapy on quality of life in heart failure patients: a systematic review and meta-analysis", *Pharmacoepidemiology and Drug Safety*, vol. 16, no. 2, pp. 152-159.

Domanski, M., Antman, E. M., McKinlay, S., Varshavsky, S., Platonov, P., Assmann, S. F., & Norman, J. 2004, "Geographic variability in patient characteristics, treatment and outcome in an international trial of magnesium in acute myocardialinfarction", *Controlled Clinical Trials*, vol. 25, no. 6, pp. 553-562.

Dong, B., Hao, Q., Yue, J., Wu, T., & Liu, G. 2006, "Thrombolytic therapy for pulmonary embolism.", *Cochrane Database of Systematic Reviews* no. 2.

Dragano, N., Bobak, M., Wege, N., Peasey, A., Verde, P., Kubinova, R., Weyers, S., Moebus, S., Mohlenkamp, S., Stang, A., Erbel, R., Jockel, K. H., Siegrist, J., & Pikhart, H. 2007, "Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany", *BMC Public Health*, vol. 7, no. 1, p. 255.

Drummond, M., Griffin, A., & Tarricone, R. 2009, "Economic Evaluation for Devices and DrugsΓÇöSame or Different?", *Value in Health*, vol. 12, no. 4, pp. 402-404.

Duval, S. & Tweedie, R. 2000, "Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis", *Biometrics*, vol. 56, no. 2, pp. 455-463.

Dzavik, V., Beanlands, D. S., Davies, R. F., Leddy, D., Marquis, J. F., Teo, K. K., Ruddy, T. D., Burton, J. R., & Humen, D. P. 1994, "Effects of late percutaneous transluminal coronary angioplasty of an occluded infarct-related coronary artery on left ventricular function in patients with a recent (<6 weeks) Q-wave acute myocardial infarction (total occlusion post-myocardial infarction intervention study [TOMIIS] - A pilot study)", *American Journal of Cardiology*, vol. 73, no. 12, pp. 856-861.

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. 1991b, "Publication bias in clinical research", *The Lancet*, vol. 337, no. 8746, pp. 867-872.

Eccles, M. 2003, "Deriving recommendations in clinical practice guidelines", *Quality and Safety in Health Care*, vol. 12, no. 5, pp. 328-329.

Egger, M., Dickersin, K., & Smith, G. D. 2001, "Problems and Limitations in Conducting Systematic Reviews," in *Systematic Reviews in Health Care: Meta-analysis in Context*, 2nd edn, M. Egger, G. D. Smith, & D. G. Altman, eds., BMJ Publishing Group, London, pp. 43-68.

Egger, M., Juni, P., Bartlett, C., Holenstein, F., & Sterne, J. 2003a, "How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study", *Health Technol Assess*, vol. 7, no. 1, pp. 1-76.

Egger, M. & Smith, G. 2001, "Principles and Procedures for Systematic Reviews," in *Systematic Reviews In Health Care: Meta-analysis in Context*, 2nd edn, M. Egger, G. D. Smith, & D. G. Altman, eds., BMJ Publishing Group, London, pp. 23-42.

Egger, M., Smith, G. D., & O'Rourke, K. 2001, *Introduction: Rationale, Potentials, and Promise of Systematic Reviews* BMJ Publishing Group.

Egger, M. & Smith, G. D. 1998e, "meta-analysis bias in location and selection of studies", *BMJ*, vol. 316, no. 7124, pp. 61-66.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. 1997b, "Bias in meta-analysis detected by a simple, graphical test", *BMJ*, vol. 315, no. 7109, pp. 629-634.

Egger, M., Zellweger-Zehner, T., Schneider, M., Junker, C., Lengeler, C., & Antes, G. 1997c, "Language bias in randomised controlled trials published in English and German", *The Lancet*, vol. 350, no. 9074, pp. 326-329.

Eisinger, F., Geller, G., Burke, W., & Holtzman, N. A. 1999a, "Cultural basis for differences between US and French clinical recommendations for women at increased risk of breast and ovarian cancer", *The Lancet*, vol. 353, no. 9156, pp. 919-920.

Ellis, S., Penn, M., & Bolwell, B. e. a. 2006, "Granulocyte colony stimulating factor in patients with large acute myocardialinfarction: results of a pilot dose-escalation randomized trial", *Am Heart J*, vol. 152, p. 1051 e9-1051 e14.

Engels, E. A., Schmid, C. H., Terrin, N., Olkin, I., & Lau, J. 2000, "Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses", *Statistics in Medicine*, vol. 19, no. 13, pp. 1707-1728.

Erne, P., Schoenenberger, A. W., Burckhardt, D., Zuber, M., Kiowski, W., Buser, P. T., Dubach, P., Resink, T. J., & Pfisterer, M. 2007, "Effects of Percutaneous Coronary Interventions in Silent Ischemia After MyocardialInfarction", *JAMA: The Journal of the American Medical Association*, vol. 297, no. 18, pp. 1985-1991.

Ernst, E. 2002, "A systematic review of systematic reviews of homeopathy", *British Journal of Clinical Pharmacology*, vol. 54, no. 6, pp. 577-582.

Etter, J. F., Burri, M., & Stapleton, J. 2007, "The impact of pharmaceutical company funding on results of randomized trials of nicotine replacement therapy for smoking cessation: a meta-analysis*", *Addiction*, vol. 102, no. 5, pp. 815-822.

European Medicines Agency, 2006. "Ethnic Factors in the Acceptability of Foreign Clinical Data."Note for Guidance on Ethnic Factors in the Acceptability of Foreign Clinical Data [online]. Available at: http://www.ema.europa.eu/docs/enGB/documentlibrary/Scientificguideline/2009/09/WC500002842.pdf [Accessed 21/04/09].

Every, N. R., Larson, E. B., Litwin, P. E., Maynard, C., Fihn, S. D., Eisenberg, M. S., Hallstrom, A. P., Martin, J. S., & Weaver, W. D. 1993, "The Association Between On-Site Cardiac-Catheterization Facilities and the Use of Coronary Angiography After Acute Myocardial-Infarction", *New England Journal of Medicine*, vol. 329, no. 8, pp. 546-551.

Ezekowitz, J. A., Rowe, B. H., Dryden, D. M., Hooton, N., Vandermeer, B., Spooner, C., & McAlister, F. A. 2007, "Systematic Review: Implantable Cardioverter Defibrillators for Adults with Left Ventricular Systolic Dysfunction", *Annals of Internal Medicine*, vol. 147, no. 4, pp. 251-262.

Faergeman, O., Kjekshus, J., Cook, T., Pyorala, K., Wilhelmsen, L., Thorgeirsson, G., & Pedersen, T. R. 1998, "Differences in the treatment of coronary heart disease between countries as revealed in the Scandinavian Simvastatin Survival Study (4S)", *European Heart Journal*, vol. 19, no. 10, pp. 1531-1537.

Fahey, T. P. & Peters, T. J. 1996, "What constitutes controlled hypertension? Patient based comparison of hypertension guidelines", *BMJ*, vol. 313, no. 7049, pp. 93-96.

Fedorowicz , Z. & Al-Muharraqi, M. 2009, "Best Evidence from the Cochrane Librar", *Bahrain Medical Bulletin*, vol. 31, no. 1.

Felstedt, M., Boesgaard, S., Bouchelouche, P., Svenningsen, A., Brooks, L., Lech Y., Aldershvile J., skagen, K., Godtfredsen, J. 1991, "Magnesium substitution in acute ischaemic heart syndromes", *European Heart Journal*, vol. 12, no. 11, pp. 1215-1218.

Felson, D. T. 1992b, "Bias in meta-analytic research", *Journal of Clinical Epidemiology*, vol. 45, no. 8, pp. 885-892.

Fervers, B., Burgers, J. S., Haugh, M. C., Latreille, J., Mlika-Cabanne, N., Paquet, L., Coulombe, M., Poirier, M., & Burnand, B. 2006, "Adaptation of clinical guidelines: literature review and proposition for a framework and procedure", *International Journal for Quality in Health Care*, vol. 18, no. 3, pp. 167-176.

Fibrinolytic Therapy Trialists' (FTT) Collaborative Group 1994, "Indications for fibrinolytic therapy in suspected acute myocardialinfarction: collaborative overview of early mortality and major morbidity results from all randomised trials of more than 1000 patients", *The Lancet*, vol. 343, no. 8893, pp. 311-322.

Field, M. & Lohr, K. 1990, *Clinical Practice Guidelines: Directions for a New Program.* National Academy Press, Washington DC.

Fisher LD, Dixon DO, Herson J, Frankowski RK, Hearron MS and Peace KE. (1990). Intention to treat in clinicaltrials. In: Peace, K.E. (ed) *Statistical Issues in Drug Research and Development.* New York: Marcel Dekker, pp. 331-350.

Fisher J, Shahshahani M, & Kitabchi A. Diabetic Ketoacidosis: Low-Dose Insulin Therapy by Various Routes. New England Journal of Medicine 297, 238-241. 1977.

Fletcher, R., Fletcher, S., & Wagner, E. 1982, *Clinical Epidemiology - The Essentials* Williams & Wilkins, Baltimore.

Fox, K. A. A., Goodman, S., Bigonzi, F., LeLouer, V., & Cohen, M. 2000d, "Inter-regional differences and outcome in unstable angina. Analysis of the International ESSENCE trial", *European Heart Journal*, vol. 21, no. 17, pp. 1433-1439.

Fox, K. A. A., Goodman, S. G., Anderson, F. A., Granger, C. B., Moscucci, M., Flather, M. D., Spencer, F., Budaj, A., Dabbous, O. H., & Gore, J. M. 2003, "From guidelines to clinical practice: the impact of hospital and geographical characteristics on temporal trends in the management of acute coronary syndromes", *European Heart Journal*, vol. 24, no. 15, pp. 1414-1424.

Freemantle, N., Tharmanathan, P., Calvert, M. J., Abraham, W. T., Ghosh, J., & Cleland, J. G. F. 2006, "Cardiac resynchronisation for patients with heart failure due to left ventricular systolic dysfunction: a systematic review and meta-analysis", *European Journal of Heart Failure*, vol. 8, no. 4, pp. 433-440.

Fu, Y., Chang, W. C., Mark, D., Califf, R. M., Mackenzie, B., Granger, C. B., Topol, E. J., Hlatky, M., & Armstrong, P. W. 2000b, "Canadian-American Differences in the Management of Acute Coronary Syndromes in the GUSTO IIb Trial : One-Year Follow-Up of Patients Without ST-Segment Elevation", *Circulation*, vol. 102, no. 12, pp. 1375-1381.

Fujita, N., Suwa, S., & Koyama, S. e. a. 2004, "The efficacy of distal embolic protection device during an acute myocardialinfarction: early and long-term results", *Am J Cardiol*, vol. 94, no. Suppl 6A, p. 34.

Gabriel Sanchez, R., Sanchez Gomez, L., Carmona, L., Roqué i Figulsm, M., & Bonfill Cosp, X. 2005, "Hormone replacement therapy for preventing cardiovascular disease in post-menopausal women", *Cochrane Database of Systematic Reviews* no. 2.

Galobardes, B., Lynch, J., & Smith, G. D. 2007, "Measuring socioeconomic position in health research", *British Medical Bulletin*, vol. 81-82, no. 1, pp. 21-37.

Gaziano, T., Reddy, K., Paccaud, F., Horton, S., & Chaturvedi, V. 2006, "Cardiovascular Disease," in *Disease Control Priorites in Developing Countries.*, 2nd edn, D. Jamison et al., eds., Oxford University Press, New York, pp. 645-662.

Gaziano, T. A., Opie, L. H., & Weinstein, M. C. 2006, "Cardiovascular disease prevention with a multidrug regimen in the developing world: a cost-effectiveness analysis", *The Lancet*, vol. 368, no. 9536, pp. 679-686.

Ghersi, D. & Pang, T. 2009, "From Mexico to Mali: four years in the history of clinical trial registration", *Journal of Evidence-Based Medicine*, vol. 2, pp. 1-7.

Gillespie, E. L., Coleman, C. I., Sander, S., Kluger, J., Gryskiewicz, K. A., & White, C. M. 2005, "Effect of Prophylactic Amiodarone on Clinical and Economic Outcomes After Cardiothoracic Surgery: A Meta-Analysis", *The Annals of Pharmacotherapy*, vol. 39, no. 9, pp. 1409-1415.

Giugliano RP, Llevadot J, Wilcox RG, Gurfinkel EP, McCabe CH, Charlesworth A, Thompson SL, Antman EM, & Braunwald E 2001, "Geographic variation in patient and hospital characteristics, management, and clinical outcomes in ST-elevation myocardialinfarctiontreated with fibrinolysis. Results from InTIME-II.", *Eur Heart J*, vol. 22, no. 18, pp. 1702-1715.

Gohler, A., Januzzi, J. L., Worrell, S. S., Osterziel, K. J., Scott Gazelle, G., Dietz, R., & Siebert, U. 2006, "A Systematic Meta-Analysis of the Efficacy and Heterogeneity of Disease Management Programs in Congestive Heart Failure", *Journal of Cardiac Failure*, vol. 12, no. 7, pp. 554-567.

Gornick, M. E., Eggers, P. W., Reilly, T. W., Mentnech, R. M., Fitterman, L. K., Kucken, L. E., & Vladeck, B. C. 2009, "Effects of Race and Income on Mortality and Use of Services among Medicare Beneficiaries", *New England Journal of Medicine*, vol. 335, no. 11, pp. 791-799.

Gray, L. J., Sprigg, N., Bath, P. M. W., Sorensen, P., Lindenstrom, E., Boysen, G., De Deyn, P. P., Friis, P., Leys, D., Marttila, R., Olsson, J. E., O'Neill, D., Ringelstein, B., van der Sande, J. J., & Turpie, A. G. G. 2006, "Significant variation in mortality and functional outcome after acute ischaemic stroke between western countries: data from the tinzaparin in acute ischaemic stroke trial (TAIST)", *Journal of Neurology Neurosurgery and Psychiatry*, vol. 77, no. 3, pp. 327-333.

Gregoire, G., Derderian, F., & Le Lorier, J. 1995, "Selecting the language of the publications included in a meta-analysis: Is there a tower of babel bias?", *Journal of Clinical Epidemiology*, vol. 48, no. 1, pp. 159-163.

Gupta, M., Chang, W. C., Van de Werf, F., Granger, C. B., Midodzi, W., Barbash, G., Pehrson, K., Oto, A., Toutouzas, P., Jansky, P., & Armstrong, P. W. 2003b, "International differences in in-hospital revascularization and outcomes following acute myocardial infarction", *European Heart Journal*, vol. 24, no. 18, pp. 1640-1650.

Haghdoost, A., Sadeghirad, B., Hajarizadeh, B., & Mirzazadeh, A. 2007, "The Application of Systematic Review and Meta-analysis Concepts in Summarizing the Findings of Observational Studies", *Iranian Journal of Psychiatry*, vol. 2, no. 4, pp. 132-136.

Hahn, S., Williamson, P. R., Hutton, J. L., Garner, P., & Flynn, E. V. 2000, "Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies", *Statistics in Medicine*, vol. 19, no. 24, pp. 3325-3336.

Hardy, R. J. & Thompson, S. G. 1998, "Detecting and describing heterogeneity in meta-analysis", *Statistics in Medicine*, vol. 17, no. 8, pp. 841-856.

Harwell, M. 1997, "An empirical study of Hedge's homogeneity test", *Psychological Methods*, vol. 2, no. 2, pp. 219-231.

Hawker, S., Payne, S., Kerr, C., Hardey, M., & Powell, J. 2002, "Appraising the Evidence: Reviewing Disparate Data Systematically", *Qualitative Health Research*, vol. 12, no. 9, pp. 1284-1299.

Healey, J. S., Baranchuk, A., Crystal, E., Morillo, C. A., Garfinkle, M., Yusuf, S., & Connolly, S. J. 2005, "Prevention of Atrial Fibrillation With Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers: A Meta-Analysis", *Journal of the American College of Cardiology*, vol. 45, no. 11, pp. 1832-1839.

Heidenreich, P. A., Trogdon, J. G., Khavjou, O. A., Butler, J., Dracup, K., Ezekowitz, M. D., Finkelstein, E. A., Hong, Y., Johnston, S. C., Khera, A., Lloyd-Jones, D. M., Nelson, S. A., Nichol, G., Orenstein, D., Wilson, P. W. F., Woo, Y. J., on behalf of the American Heart Association Advocacy Coordinating Committee, Stroke, C., Council on Cardiovascular Radiology and Intervention, Council on Clinical Cardiology, Council on Epidemiology and Prevention, Council on Arteriosclerosis, Thrombosis and Vascular Biology, Council on Cardiopulmonary, Critical, C., Perioperative and Resuscitation, Council on Cardiovascular Nursing, Council on the Kidney in Cardiovascular Disease, & Council on Cardiovascular Surgery and Anesthesia, a. I. C. o. Q. o. C. a. O. R. 2011, "Forecasting the Future of Cardiovascular Disease in the United States: A Policy Statement From the American Heart Association", *Circulation*, vol. 123, no. 8, pp. 933-944.

Heijink, R., Noethen, M., Renaud, T., Koopmanschap, M., & Polder, J. 2008, "Cost of illness: An international comparison: Australia, Canada, France, Germany and The Netherlands", *Health Policy*, vol. 88, no. 1, pp. 49-61.

Helton, T. J., Bavry, A. A., Kumbhani, D. J., Duggal, S., Roukoz, H., & Bhatt, D. L. 2007, "Incremental Effect of Clopidogrel on Important Outcomes in Patients with Cardiovascular Disease: A Meta-Analysis of Randomized Trials", *American Journal of Cardiovascular Drugs*, vol. 7, no. 4.

Hemming, K., Lilford, R., & Bowater, R, 2010. A Bayesian Panoramic Meta-analysis for Pooling Systematic Reviews of Systematic Reviews. [Personal Communication].

Higgins, J. P. T. & Thompson, S. G. 2002, "Quantifying heterogeneity in a meta-analysis", *Statistics in Medicine*, vol. 21, no. 11, pp. 1539-1558.

Higgins, J. & Altman, D. 2008, "Assessing risk of bias in included studies," in *Cochrane handbook for systematic reviews of interventions*, Version 5.0.0 edn, J. Higgins & D. G. Altman, eds., www.cochrane-handbook.org/, Cochrane Collaboration.

Higgins, J. & Green, S. 2005, "Assessment of Study Quality," in *Cochrane Handbook for Systematic Reviews of Interventions.*, 4.2.5 edn, J. Higgins & S. Green, eds., The Cochrane Collaboration.

Higgins, J. & Green, S. 2011, *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.0.0 edn, The Cochrane Collaboration, Oxford.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. 2003a, "Measuring inconsistency in meta-analyses", *BMJ*, vol. 327, no. 7414, pp. 557-560.

Ho, K. M., Sheridan, D. J., & Paterson, T. 2007, "Use of intravenous magnesium to treat acute onset atrial fibrillation: a meta-analysis", *Heart*, vol. 93, no. 11, pp. 1433-1440.

Holdgate, A. & Foo, A. 2006, "Adenosine versus intravenous calcium channel antagonists for the treatment of supraventricular tachycardia in adults.", *Cochrane Database of Systematic Reviews 2005* no. 4.

Holland, R., Battersby, J., Harvey, I., Lenaghan, E., Smith, J., & Hay, L. 2005, "Systematic review of multidisciplinary interventions in heart failure", *Heart*, vol. 91, no. 7, pp. 899-906.

Hood, W. B., Dans, A., Guyatt, G. H., Jaeschke, R., & McMurray , J. 2004, "Digitalis for treatment of congestive heart failure in patients in sinus rhythm", *The Cochrane Database of Systematic Reviews*, vol. 1, pp. 1-18.

Hooper, L., Harrison, R., Summerbell, C., Moore, H., Worthington, H., Ness, A., Capps, N., Smith, G., Riemersma, R., & Ebrahim, S. 2004, "Omega 3 fatty acids for prevention and treatment of cardiovascular disease", *Cochrane Database of Systematic Reviews* no. 4.

Horie, H., Takahashi, M., Minai, K., Izumi, M., Takaoka, A., Nozawa, M., Yokohama, H., Fujita, T., Sakamoto, T., Kito, O., Okamura, H., & Kinoshita, M. 1998, "Long-term beneficial effect of late reperfusion for acute anterior myocardial infarction with percutaneous transluminal coronary angioplasty", *Circulation*, vol. 98, no. 22, pp. 2377-2382.

Hotopf, M. 2007, "Clinical Epidemiology," in *Seminars in General Adult Psychiatry*, 2nd edn, G. Wilkinson & G. Stein, eds., Royal College of Psychiatrists, London, p. 705.

Hotopf, M., Lewis, G., & Normand, C. 1997, "Putting trials on trial--the costs and consequences of small trials in depression: a systematic review of methodology", *Journal of Epidemiology and Community Health*, vol. 51, no. 4, pp. 354-358.

Houser, J. 2008, "The Sampling Strategy," in *Nursing Research: Reading, Using and Creating Evidence.*, J. Houser, ed., Jones and Bartlett, Massachusetts.

Hristov, M., Heussen, N., Schober, A., & Weber, C. 2006, "Intracoronary infusion of autologous bone marrow cells and left ventricular function after acute myocardialinfarction: a meta-analysis", *Journal of Cellular and Molecular Medicine*, vol. 10, no. 3, pp. 727-733.

Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. 2006, "Assessing heterogeneity in meta-analysis: Q statistic or I-2 index?", *Psychological Methods*, vol. 11, no. 2, pp. 193-206.

Huston, P. & Moher, D. 1996, "Redundancy, disaggregation, and the integrity of medical research", *The Lancet*, vol. 347, no. 9007, pp. 1024-1026.

Ikari, Y., Kawano, S., & Sakurada, M. e. a. 2005, "Thrombus aspiration prior to coronary intervention improves myocardial microcirculation in patients with ST elevation acute myocardial infarction: the VAMPIRE study", *Circulation*, vol. 112, no. Suppl 2, p. 659.

Ince, H., Valgimigli, M., Petzsch, M., de Lezo, J. S., Kuethe, F., Dunkelmann, S., Biondi-Zoccai, G., & Nienaber, C. A. 2008, "Cardiovascular events and re-stenosis following administration of G-CSF in acute myocardial infarction: systematic review and meta-analysis", *Heart*, vol. 94, no. 5, pp. 610-616.

Institute for Work and Health, 2011. *Systematic reviews.* [online]. Available at: http://www.iwh.on.ca/systematic-reviews [Accessed 23 March 2010].

Ioannidis, J. P. A. 1998, "Effect of the Statistical Significance of Results on the Time to Completion and Publication of Randomized Efficacy Trials", *JAMA: The Journal of the American Medical Association*, vol. 279, no. 4, pp. 281-286.

Ioannidis, J. P. A. & Katritsis, D. G. 2007, "Percutaneous coronary intervention for late reperfusion after myocardial infarction in stable patients", *American Heart Journal*, vol. 154, no. 6, pp. 1065-1071.

Ioannidis, J. P. & Lau, J. 1998, "Can quality of clinical trials and meta-analyses be quantified?", *The Lancet*, vol. 352, no. 9128, pp. 590-591.

Ioannidis, J. P. A. & Lau, J. 1998, "Heterogeneity of the Baseline Risk within Patient Populations of Clinical Trials: A Proposed Evaluation Algorithm", *American Journal of Epidemiology*, vol. 148, no. 11, pp. 1117-1126.

Jadad, A. 1998, "Assessing the quality of RCTs: why, what, how, and by whom?," in *Randomised controlled trials: a user's guide*, A. R. Jadad, ed., BMJ Books, London.

Jadad, A. R., Cook, D. J., Jones, A., Klassen, T. P., Tugwell, P., Moher, M., & Moher, D. 1998, "Methodology and Reports of Systematic Reviews and Meta-analyses", *JAMA: The Journal of the American Medical Association*, vol. 280, no. 3, pp. 278-280.

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., & McQuay, H. J. 1996, "Assessing the quality of reports of randomized clinical trials: Is blinding necessary?", *Controlled Clinical Trials*, vol. 17, no. 1, pp. 1-12.

Jaffery, Z., Kowalski, M., Weaver, W. D., & Khanal, S. 2007, "A meta-analysis of randomized control trials comparing minimally invasive direct coronary bypass grafting versus percutaneous coronary intervention for stenosis of the proximal left anterior descending artery", *European Journal of Cardio-Thoracic Surgery*, vol. 31, no. 4, pp. 691-697.

Jensen, P. H., Webster, E., & Witt, J. 2009, "Hospital type and patient outcomes: an empirical examination using AMI readmission and mortality records", *Health Economics*, vol. 18, no. 12, pp. 1440-1460.

Jha, A. K., Fisher, E. S., Li, Z., Orav, E. J., & Epstein, A. M. Racial Trends in the Use of Major Procedures Among the Elderly. 14[12], 48. 1-12-2005. Ref Type: Abstract

Jibrini, M., Molnar, J., & Arora, R. 2008, "Prevention of atrial fibrillation by way of abrogation of the renin-angiotensin system: a systematic review and meta-analysis.", *Am J Ther*, vol. 15, no. 1, pp. 36-43.

Jimenez-Buedo, M. & Miller, L. M. 2010, "Why a Trade-Off? The Relationship between the External and Internal Validity of Experiments", *THEORIA.An International Journal for Theory, History and Foundations of Science; Vol 25, No 3 (2010)*.

Jollis, J. G., DeLong, E. R., Peterson, E. D., Muhlbaier, L. H., Fortin, D. F., Califf, R. M., & Mark, D. B. 2009, "Outcome of Acute MyocardialInfarction According to the Specialty of the Admitting Physician", *New England Journal of Medicine*, vol. 335, no. 25, pp. 1880-1887.

Jovicic, A., Holroyd-Leduc, J., & Straus, S. 2006, "Effects of self-management intervention on health outcomes of patients with heart failure: a systematic review of randomized controlled trials", *BMC Cardiovascular Disorders*, vol. 6, no. 1, p. 43.

Jude-Eze, E.N., 2011. *Quality of Meta-analysis in terms of actual statistical analysis.* MSc(R), University of Glasgow.

Juni, P., Altman, D. G., & Egger, M. 2001f, "Assessing the quality of controlled clinical trials", *BMJ*, vol. 323, no. 7303, pp. 42-46.

Juni, P., Holenstein, F., Sterne, J., Bartlett, C., & Egger, M. 2002, "Direction and impact of language bias in meta-analyses of controlled trials: empirical study", *International Journal of Epidemiology*, vol. 31, no. 1, pp. 115-123.

Kang, S., Yang, Y., Li, C., & Gao, R. 2008, "Effects of intracoronary autologous bone marrow cells on left ventricular function in acute myocardial infarction: a systematic review and meta-analysis for randomized controlled trials.", *Coron Artery Dis*, vol. 19, no. 5, pp. 327-335.

Kang, S., Yang, Y., Li, C. J., & Gao, R. 2007, "Effectiveness and tolerability of administration of granulocyte colony-stimulating factor on left ventricular function in patients with myocardial infarction: A meta-analysis of randomized controlled trials", *Clinical Therapeutics*, vol. 29, no. 11, pp. 2406-2418.

Kang, Y., Ohtsu, A., Van Cutsem, E., Rha, S., Sawaki, A., Park, S., Lim, H., Wu, J., Langer, B., & Shah, M. (2010) ."AVAGAST: A randomized, double-blind, placebo-controlled, phase III study of first-line Capecitabine and Cisplatin plus bevacizumab or placebo in patients with Adavanced Gastric Cancer", in *ASCO Annual Meeting Proceedings*, 18 edn.

Kapral, M. K., Wang, H., Mamdani, M., Tu, J. V., Boden-Albala, B., & Sacco, R. L. 2002, "Effect of Socioeconomic Status on Treatment and Mortality After Stroke * Editorial Comment", *Stroke*, vol. 33, no. 1, pp. 268-275.

Kastrati, A., Mehilli, J., Pache, J., Kaiser, C., Valgimigli, M., Kelbaek, H., Menichelli, M., Sabate, M., Suttorp, M. J., Baumgart, D., Seyfarth, M., Pfisterer, M. E., & Schomig, A. 2007, "Analysis of 14 Trials Comparing Sirolimus-Eluting Stents with Bare-Metal Stents", *New England Journal of Medicine*, vol. 356, no. 10, pp. 1030-1039.

Kastrati, A., Schomig, A., Dirschinger, J., Mehilli, J., Dotzer, F., von Welser, N., & Neumann, F. J. 2000, "A Randomized Trial Comparing Stenting With Balloon Angioplasty in Small Vessels in Patients With Symptomatic Coronary Artery Disease", *Circulation*, vol. 102, no. 21, pp. 2593-2598.

Katerndahl, D. A. & Lawler, W. R. 1999, "Variability in Meta-analytic Results Concerning the Value of Cholesterol Reduction in Coronary Heart Disease: A Meta-Meta-Analysis", *American Journal of Epidemiology*, vol. 149, no. 5, pp. 429-441.

Kaul, P., Newby, L. K., Fu, Y., Mark, D. B., Califf, R. M., Topol, E. J., Aylward, P., Granger, C. B., Van de Werf, P. F., & Armstrong, P. W. 2004a, "International differences in evolution of early discharge after acute myocardial infarction", *The Lancet*, vol. 363, no. 9408, pp. 511-517.

Kawashima, A., Ueda, K., Nishida, Y., Inoue, N., Tanaka, S., & Kato O, e. a. 1999, "Quantitative angiographic analysis of restenosis of primary stenting using wiktor stent for acute myocardial infarction: results form a multicenter randomized PRISAM study.", *Circulation*, vol. 100, no. Suppl 1, pp. 1-856.

Keeley, E. C., Boura, J. A., & Grines, C. L. 2006, "Comparison of primary and facilitated percutaneous coronary interventions for ST-elevation myocardial infarction: quantitative review of randomised trials", *The Lancet*, vol. 367, no. 9510, pp. 579-588.

Kelly, S., Frost, G., Whittaker, V., & Summerbell, C. 2004, "Low Glycaemic Index Diets for Coronary Heart Disease.", *Cochrane Database of Systematic Reviews* no. 4.

Khan, K. S., Daya, S., & Jadad, A. R. 1996a, "The Importance of Quality of Primary Studies in Producing Unbiased Systematic Reviews", *Archives of Internal Medicine*, vol. 156, no. 6, pp. 661-666.

Kjaergard, L. L. & Gluud, C. 2002a, "Citation bias of hepato-biliary randomized clinical trials", *Journal of Clinical Epidemiology*, vol. 55, no. 4, pp. 407-410.

Kjaergard, L. L., Villumsen, J., & Gluud, C. 2001a, "Reported Methodologic Quality and Discrepancies between Large and Small Randomized Trials in Meta-Analyses", *Annals of Internal Medicine*, vol. 135, no. 11, pp. 982-989.

Kondo, A., Zierler, B.K., & Hagino, H., 2010,"Relationship between the length of hospital stay after hip fracture surgery and ambulatory ability or mortality after discharge in Japan", *Japan Journal of Nursing Science*, vol. 7, no. 1, pp. 96-107.

Kromhout, D., Menotti, A., Kesteloot, H., & Sans, S. 2002, "Prevention of Coronary Heart Disease by Diet and Lifestyle: Evidence From Prospective Cross-Cultural, Cohort, and Intervention Studies", *Circulation*, vol. 105, no. 7, pp. 893-898.

Krum, H., Haas, S. J., Eichhorn, E., Ghali, J., Gilbert, E., Lechat, P., Packer, M., Roecker, E., Verkenne, P., Wedel, H., & Wikstrand, J. 2005, "Prognostic benefit of beta-blockers in patients not receiving ACE-Inhibitors", *European Heart Journal*, vol. 26, no. 20, pp. 2154-2158.

Kubo, S. H., Gollub, S., Bourge, R., Rahko, P., Cobb, F., Jessup, M., Brozena, S., Brodsky, M., Kirlin, P., & Shanes, J. 1992, "Beneficial effects of pimobendan on exercise tolerance and quality of life in patients with heart failure. Results of a multicenter trial. The Pimobendan Multicenter Research Group", *Circulation*, vol. 85, no. 3, pp. 942-949.

Kuulasmaa, K., Tunstall-Pedoe, H., Dobson, A., Fortmann, S., Sans, S., Tolonen, H., Evans, A., & Ferrario, M. 2000, "Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations", *The Lancet*, vol. 355, no. 9205, pp. 675-687.

Labinaz, M., Ho, C., Banerjee, S., Martin, J., Chen, S., & Mensinkai, S. 2007, "Meta-analysis of clinical efficacy and bleeding risk with intravenous glycoprotein IIb/IIIa antagonists for percutaneous coronary intervention", *The Canadian journal of cardiology*, vol. 23, no. 12, pp. 963-970.

Lafuente-Lafuente, C., Mouly, S., Longas-Tejero, M., & Bergmann, J. 2007, "Antiarrhythmics for maintaining sinus rhythm after cardioversion of atrial fibrillation.", *Cochrane Database of Systematic Reviews* no. 4.

Lakhdar, R., Al-Mallah, M. H., & Lanfear, D. E. 2008, "Safety and Tolerability of Angiotensin-Converting Enzyme Inhibitor Versus the Combination of Angiotensin-Converting Enzyme Inhibitor and Angiotensin Receptor Blocker in Patients With Left Ventricular Dysfunction: A Systematic Review andáMeta-Analysis of Randomized Controlled Trials", *Journal of Cardiac Failure*, vol. 14, no. 3, pp. 181-188.

Lau, J., Ioannidis, J. P. A., & Schmid, C. H. 1997a, "Quantitative Synthesis in Systematic Reviews", *Annals of Internal Medicine*, vol. 127, no. 9, pp. 820-826.

Leizorovicz, A., Haugh, M., & Boissel, J. P. 1992, "Meta-analysis and multiple publication of clinical trial reports", *The Lancet*, vol. 340, no. 8827, pp. 1102-1103.

Levitt, M. A. 2001, "A prospective, randomized trial of BiPAP in severe acute congestive heart failure", *Journal of Emergency Medicine*, vol. 21, no. 4, pp. 363-369.

Li, J., Zhang, Q., Zhang, M., & Egger, M. 2007, "Intravenous magnesium for acute myocardial infarction.", *Cochrane Database of Systematic Reviews 2005* no. 2.

Liberati, A., Himel, H. N., & Chalmers, T. C. 1986, "A quality assessment of randomized control trials of primary treatment of breast cancer", *Journal of Clinical Oncology*, vol. 4, no. 6, pp. 942-951.

Lim, W., Arnold, D. M., Bachanova, V., Haspel, R. L., Rosovsky, R. P., Shustov, A. R., & Crowther, M. A. 2008, "Evidence-Based Guidelines--An Introduction", *Hematology*, vol. 2008, no. 1, pp. 26-30.

Linde, K., Scholz, M., Ramirez, G., Clausius, N., Melchart, D., & Jonas, W. B. 1999, "Impact of Study Quality on Outcome in Placebo-Controlled Trials of Homeopathy", *Journal of Clinical Epidemiology*, vol. 52, no. 7, pp. 631-636.

Lindsay, H. S. J., Zaman, A. G., & Cowan, J. C. 1995, "Ace-Inhibitors After Myocardial-Infarction - Patient Selection Or Treatment for All", *British Heart Journal*, vol. 73, no. 5, pp. 397-400.

Lipinski, M. J., Biondi-Zoccai, G. G. L., Abbate, A., Khianey, R., Sheiban, I., Bartunek, J., Vanderheyden, M., Kim, H. S., Kang, H. J., Strauer, B. E., & Vetrovec, G. W. 2007, "Impact of Intracoronary Cell Therapy on Left Ventricular Function in the Setting of Acute Myocardial Infarction: A Collaborative Systematic Review and Meta-Analysis of Controlled Clinical Trials", *Journal of the American College of Cardiology*, vol. 50, no. 18, pp. 1761-1767.

Littlejohns, P., Cluzeau, F., Bale, R., Grimshaw, J., Feder, G., & Moran, S. 1999, "The quantity and quality of clinical practice guidelines for the management of depression in primary care in the UK", *Br J Gen Pract*, vol. 49, pp. 205-210.

Lyratzopoulos, G. 2006, "The cost of cardiovascular disease: rising, declining or staying still?", *Heart*, vol. 92, no. 10, pp. 1361-1362.

Madan, M., Labinaz, M., Cohen, E. A., Buller, C. E., Cantor, W. J., Seidelin, P., Ducas, J., Carere, R. G., Natarajan, M. K., Pieper, K. S., Hafley, G. E., O'Shea, J. C., Kitt, M. M., Califf, R. M., & Tcheng, J. E. 2004, "A comparison of clinical outcomes between Canadian and American patients after nonurgent coronary stenting", *Canadian Journal of Cardiology*, vol. 20, no. 13, pp. 1343-1349.

Madhukar, P. Data Analysis in Systematic Reviews. 2009. Ref Type: Slide

Magill, A & Shlim, D. 2011, Perspectives: Why Guidelines Differ. [online] Available at: http://wwwnc.cdc.gov/travel/yellowbook/2012/chapter-1-introduction/perspectives-why-guidelines-differ.htm [Accessed 14/2/12].

Mahaffey, K. W., Harrington, R. A., Akkerhuis, M., Kleiman, N. S., Berdan, L. G., Crenshaw, B. S., Tardiff, B. E., Granger, C. B., DeJong, I., Bhapkar, M., Widimsky, P., Corbalon, R., Lee, K. L., Deckers, J. W., Simoons, M. L., Topol, E. J., & Califf, R. M. 2001, "Systematic adjudication of myocardial infarction end-points in an international clinical trial", *Current Controlled Trials in Cardiovascular Medicine*, vol. 2, no. 4, pp. 180-186.

Maillard, L., Hamon, M., Khalife, K., Steg, P. G., Beygui, F., Guermonprez, J. L., Spaulding, C. M., Boulenc, J. M., Lipiecki, J., Lafont, A., Brunel, P., Grollier, G., Koning, R., Coste, P., Favereau, X., Lancelin, B., Van Belle, E., Serruys, P., Monassier, J. P., & Raynaud, P. 2000, "A comparison of systematic stenting and conventional balloon angioplasty during primary percutaneous transluminal coronary angioplasty for acute myocardial infarction", *Journal of the American College of Cardiology*, vol. 35, no. 7, pp. 1729-1736.

Marmot, M. G. 2003, "Understanding social inequalities in health", *Perspectives in Biology and Medicine*, vol. 46, no. 3, p. S9-S23.

Marmot, M. G., Shipley, M. J., & Rose, G. 1984, "Inequalities in death - specific explanations of a general pattern?", *The Lancet*, vol. 323, no. 8384, pp. 1003-1006.

Martin-Rendon, E., Brunskill, S., Doree, C., Hyde, C., Mathur, A., Stanworth, S., & Watt, S. 2008, "Stem cell treatment for acute myocardial infarction.", *Cochrane Database of Systematic Reviews 2005* no. 4.

Matsuo, A., Inoue, N., Suzuki, K., Nakamura, R., Fujita, H., Miki, S., & Yokoi, Y. 2007, "Limitations of using a GuardWire temporary occlusion and aspiration system in

patients with acute myocardial infarction: multicenter investigation of coronary artery protection with a distal occlusion device in acute myocardial infarction (MICADO)", *Journal of Invasive Cardiology*, vol. 19, pp. 132-138.

Matthy, J., DeMeyere, M., van Driel, M., & DeSutter, A. 2007, "Differences Among International Pharyngitis Guidelines: Not Just Academic", *Annals of Family Medicine*, vol. 5, no. 5, pp. 436-443.

McAlister, F. A., Ezekowitz, J., Hooton, N., Vandermeer, B., Spooner, C., Dryden, D. M., Page, R. L., Hlatky, M. A., & Rowe, B. H. 2007, "Cardiac Resynchronization Therapy for Patients With Left Ventricular Systolic Dysfunction", *The Journal of the American Medical Association*, vol. 297, no. 22, pp. 2502-2514.

McDonald, M. A., Simpson, S. H., Ezekowitz, J. A., Gyenes, G., & Tsuyuki, R. T. 2005, "Angiotensin receptor blockers and risk of myocardial infarction: systematic review", *British Medical Journal*, vol. 331, no. 7521, p. 873.

Mehta, S. R., Cannon, C. P., Fox, K. A. A., Wallentin, L., Boden, W. E., Spacek, R., Widimsky, P., McCullough, P. A., Hunt, D., Braunwald, E., & Yusuf, S. 2005, "Routine vs Selective Invasive Strategies in Patients With Acute Coronary Syndromes", *The Journal of the American Medical Association*, vol. 293, no. 23, pp. 2908-2917.

Millar, W. J. & Wigle, D. T. 1986, "Socioeconomic Disparities in Risk-Factors for Cardiovascular-Disease", *Canadian Medical Association Journal*, vol. 134, no. 2, pp. 127-132.

Moer, R., Myreng, Y., Molstad, P., Albertsson, P., Gunnes, P. a., Lindvall, B., Wiseth, R., Ytre-Arne, K., Kjekshus, J., & Golf, S. 2001, "Stenting in small coronary arteries (SISCA) trial: A randomized comparison between balloon angioplasty and the heparin-coated beStent", *Journal of the American College of Cardiology*, vol. 38, no. 6, pp. 1598-1603.

Moher, D., Cook, D., Jadad, A. R., Tugwell, P., Moher, M., Jones, A., Pham, B., & Klassen, T. 1999a, "Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses", *Health Technol Assess*, vol. 3, no. 12.

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. 1999b, "Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement", *The Lancet*, vol. 354, no. 9193, pp. 1896-1900.

Moher, D., Pham, B., Jones, A., Cook, D. J., Jadad, A. R., Moher, M., Tugwell, P., & Klassen, T. P. 1998c, "Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses?", *The Lancet*, vol. 352, no. 9128, pp. 609-613.

Moja, L. P., Telaro, E., D'Amico, R., Moschetti, I., Coe, L., & Liberati, A. 2005, "Assessment of methodological quality of primary studies by systematic reviews: results of the metaquality cross sectional study", *BMJ*, vol. 330, no. 7499, p. 1053.

Moller, C. H., Penninga, L., Wetterslev, J, Steinbruchel, D. A., & Gluud, C. 2008, "Clinical outcomes in randomized trials of off- vs. on-pump coronary artery bypass surgery: systematic review with meta-analyses and trial sequential analyses", *European Heart Journal*, vol. 29, no. 21, pp. 2601-2616.

Montori, V. & Guyatt, G. 2001, "Intention to Treat Principle", *Canadian Medical Association Journal*, vol, 1655, no.10, pp.1339-1341.

Morris, S., Ruel, M., Cohen, R., Dewey, K., de la Briere, B., & Hassan, M. 1999, "Precision, Accuracy and Reliability of Hemoglobin Assessment with use of Capillary Blood", *American Journal of Clinical Nutrition*, vol, 69, no.6, pp. 1243-1248.

Mosteller, F., Gilbert, J. P., & McPeek, B. 1980, "Reporting standards and research strategies for controlled trials: Agenda for the editor", *Controlled Clinical Trials*, vol. 1, no. 1, pp. 37-58.

Muramatsu, T., Iwasaki, K., Inou, N., Horita, Y., Tanaka, T., & Fujita, N. Efficacy of coronary stenting versus balloon angioplasty in vessels of diameters less than 3.0 mm and less than 2.5 mm: CHIVAS investigators. Am J Cardiol 90[96H]. 2002.

Murray, C. J. & Lopez, A. D. 1997, "Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study", *The Lancet*, vol. 349, no. 9064, pp. 1498-1504.

Myers, J. 2003, "Exercise and Cardiovascular Health", *Circulation*, vol. 107, no. 1, p. e2-e5.

Nakashima, H., Katayama, T., Honda, Y., Suzuki, S., & Yano, K. 2004, "Cardioprotective Effects of Magnesium Sulfate in Patients Undergoing Primary Coronary Angioplasty for Acute Myocardial Infarction", *Circulation Journal*, vol. 68, no. 1, pp. 23-28.

Nash, I.S., Nash, D.B., & Fuster, V. Do cardiologists do it better? Journal of the American College of Cardiology. 29[3], 475-478.1-3-1997.

National Institute of Health and Clinical Excellence (NICE), 2011. NICE guidedance. [online]. Available at: http://guidance.nice.org.uk/[Accessed 25/03/10].

Neal, B., Chapman, N., & Patel, A. 2002b, "Managing the global burden of cardiovascular disease", *European Heart Journal Supplements*, vol. 4, no. suppl F, p. F2-F6.

Newell D J. 1992. "Intention-to-treat analysis: implications for quantitative and qualitative research", *International Journal of Epidemiology*. vol. 21, pp 837–841.

NHS Centre for Reviews and Dissemination, 2011. Centre for Reviews and Dissemination. [online]. Available at:http://www.york.ac.uk/inst/crd/index_databases.htm[Accessed 25/03/10].

Nolan, R., Jong, P., Barry-Bianchi, S., Tanaka, T., & Floras, J. 2008, "Effects of drug, biobehavioral and exercise therapies on heart rate variability in coronary artery disease: a systematic review.", *Eur J Cardiovasc Prev Rehabil.*, vol. 15, no. 4, pp. 386-396.

Nordmann, A., Bucher, H., Hengstler, P., Harr, T., & Young , J. 2005, "Primary Stenting versus Balloon Angioplasty for Treating Acute Myocardial Infarction.", *Cochrane Database of Systematic Reviews 2005* no. 2.

Nordmann, A. J., Briel, M., & Bucher, H. C. 2006, "Mortality in randomized controlled trials comparing drug-eluting vs. bare metal stents in coronary artery disease: a meta-analysis", *European Heart Journal*, vol. 27, no. 23, pp. 2784-2814.

Nordmann, A., Kasenda, B., & Briel, M. 2012, "Meta-analyses: What they can and cannot do.", *Swiss Medical Weekly.*, vol.142, pp.w13518.

Noseworthy, J., Ebers, G., Vandervoort, M., Farquhar, R., Yetisir, E., & Roberts, R. 1994, "The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial.", *Neurology.*, vol. 44, no. 1, pp. 16-20.

O'Shea, J. & DeMets, D. 2001b, "Statistical issues relating to international differences in clinical trials", *American Heart Journal*, vol. 142, no. 1, pp. 21-28.

O'Shea, J. C. & Califf, R. M. 2001a, "International differences in treatment effects in cardiovascular clinical trials", *American Heart Journal*, vol. 141, no. 5, pp. 875-880.

O'Shea, J. C., Fu, Y., Chang, W. C., & Armstrong, P. W. 2001, "A tale of two countries: Insights from the differences in Canadian/American patterns of care for patients with acute coronary syndromes", *American Heart Journal*, vol. 142, no. 1, pp. 14-20.

Okishige, K., Nishizaki, M., Azegami, K., Igawa, M., Yamawaki, N., Aonuma, K., & the Kanagawa Arrhythmia Task Force investigators 2000, "Pilsicainide for conversion and maintenance of sinus rhythm in chronic atrial fibrillation: A placebo-controlled, multicenter study", *American Heart Journal*, vol. 140, no. 3, pp. 437-444.

Olivo, S. A., Macedo, L. G., Gadotti, I. C., Fuentes, J., Stanton, T., & Magee, D. J. 2008, "Scales to Assess the Quality of Randomized Controlled Trials: A Systematic Review", *Physical Therapy*, vol. 88, no. 2, pp. 156-175.

Pan, Z., Trikalinos, T. A., Kavvoura, F. K., Lau, J., & Ioannidis, J. P. A. 2005b, "Local Literature Bias in Genetic Epidemiology: An Empirical Evaluation of the Chinese Literature", *PLoS Med*, vol. 2, no. 12, p. e334.

Pan, Z., Trikalinos, T. A., Kavvoura, F. K., Lau, J., & Ioannidis, J. P. A. 2005a, "Local Literature Bias in Genetic Epidemiology: An Empirical Evaluation of the Chinese Literature", *PLoS Med*, vol. 2, no. 12, p. e334.

Papanikolaou, P. N., Christidi, G. D., & Ioannidis, J. P. A. 2006, "Patient Outcomes with Teaching Versus Nonteaching Healthcare: A Systematic Review", *PLoS Med*, vol. 3, no. 9, p. e341.

Park D, Desai P, Aiyengar J, & Balladur A 1998, "Geographic differences in the characteristics of coronary artery disease in India.", *International Journal of Cardiology*, vol. 67, no. 2, pp. 187-189.

Patel, A. A., White, C. M., Gillespie, E. L., Kluger, J., & Coleman, C. I. 2006, "Safety of amiodarone in the prevention of postoperative atrial fibrillation: A meta-analysis", *American Journal of Health-System Pharmacy*, vol. 63, no. 9, pp. 829-837.

Peduzzi, P., Wittes, J., Detre, K., & Holford, T. 1993, "Analysis as-randomized and the problem of non-adherence: An example from the veterans affairs randomized trial of coronary artery bypass surgery", *Statistics in Medicine*, vol. 12, no. 11, pp. 1185-1195.

Pena, V, 1994. OVID MEDLINE: A metamorphosis. *Penn Printout,* vol. 11, no. 2, pp.6-8.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. 2007, "Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity", *Statistics in Medicine*, vol. 26, no. 25, pp. 4544-4562.

Petrie, A. & Sabin, C. 2009, *Medical Statistics at a Glance*, First edn, John Wiley & Sons LTD, Oxford.

Poole, C. & Greenland, S. 1999, "Random-Effects Meta-Analyses Are Not Always Conservative", *American Journal of Epidemiology*, vol. 150, no. 5, pp. 469-475.

Poole-Wilson, P. 2005, "The prevention of cardiovascular disease worldwide: whose task and WHO's task?", *Clinical Medicine, Journal of the Royal College of Physicians*, vol. 5, no. 4, pp. 379-384.

Potosky, A. L., Breen, N., Graubard, B. I., & Parsons, P. E. 1998, "The association between health care coverage and the use of cancer screening tests - Results from the 1992 National Health Interview Survey", *Medical Care*, vol. 36, no. 3, pp. 257-270.

Psaty, B. M., Koepsell, T. D., Wagner, E. H., LoGerfo, J. P., & Inui, T. S. 1990, "The Relative Risk of Incident Coronary Heart Disease Associated With Recently Stopping the Use of {beta}-Blockers", *The Journal of the American Medical Association*, vol. 263, no. 12, pp. 1653-1657.

Pullar, T., Kumar, S., & Feely, M. 1989, "Compliance in clinical trials", *Annals of the Rheumatic Diseases*, vol. 48, no. 10, pp. 871-875.

Qayyum, R., Khalid, M. R., Adomaityte, J., Papadakos, S. P., & Messineo, F. C. 2008, "Systematic Review: Comparing Routine and Selective Invasive Strategies for the Acute Coronary Syndrome", *Annals of Internal Medicine*, vol. 148, no. 3, pp. 186-196.

Rai, A. 2006, "Searching the Medical Literatures through Internet: Review of MEDLINE Database", *TULSSAA Journal*, vol. 4, no. 3, pp. 7-13.

Ramakrishnan, N. 2007, "Thrombolysis is not warranted in submassive pulmonary embolism: a systematic review and meta-analysis", *Crit Care Resusc*, vol. 9, no. 4, pp. 357-363.

Ravago, T. S., Mosniam, J., & Alem, F. 2000, "Evaluation of Community Acquired Pneumonia Guidelines", *Journal of Medical Systems*, vol. 24, no. 5, pp. 289-296.

Ravnskov, U. 1992, "Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome", *British Medical Journal*, vol. 305, no. 6844, pp. 15-19.

Rayner, M., Petersen, S., Moher, M., & et al 2001, *Coronary Heart Disease Statistics: Morbidity Supplement 2001 Edition*, 2001 edn, British Heart Foundation, London.

Reddy, K. S. & Yusuf, S. 1998, "Emerging Epidemic of Cardiovascular Disease in Developing Countries", *Circulation*, vol. 97, no. 6, pp. 596-601.

Reed, S. D., McMurray, J. J. V., Velazquez, E. J., Schulman, K. A., Califf, R. M., Kober, L., Maggioni, A. P., Van de Werf, F., White, H. D., Diaz, R., Mareev, V., & Murin, J. 2006, "Geographic variation in the treatment of acute myocardial infarction in the VALsartan In Acute myocardial iNfarcTion (VALIANT) trial", *American Heart Journal*, vol. 152, no. 3, pp. 500-508.

Rees, K., Bennett, P., West, R., Smith, G., & Ebrahim, S. 2004, "Psychological Interventions for Coronary Heart Disease.", *Cochrane Database of Systematic Reviews* no. 2.

Reinhardt, U. E., Hussey, P. S., & Anderson, G. F. 2004, "U.S. Health Care Spending In An International Context", *Health Affairs*, vol. 23, no. 3, pp. 10-25.

Rimm, E. B., Williams, P., Fosher, K., Criqui, M., & Stampfer, M. J. 1999, "Moderate alcohol intake and lower risk of coronary heart disease: meta-analysis of effects on lipids and haemostatic factors", *BMJ*, vol. 319, no. 7224, pp. 1523-1528.

Roberts, C. & Torgerson, J. 1999, "Understanding controlled trials: Baseline imbalance in randomised controlled trials", *BMJ*, vol. 319, no. 7203, p. 185.

Roberts, C., Guallar, E., & Rodriguez, A. 2007, "Efficacy and safety of statin monotherapy in older adults: a meta-analysis.", *J Gerontol A Biol Sci Med Sci.*, vol. 62, no. 8, pp. 879-887.

Roccaforte, R., Demers, C., Baldassarre, F., Teo, K., & Yusuf, S. 2005, "Effectiveness of comprehensive disease management programmes in improving clinical outcomes in heart failure patients. A meta-analysis", *European Journal of Heart Failure*, vol. 7, no. 7, pp. 1133-1144.

Roiron, C., Sanchez, P., Bouzamondo, A., Lechat, P., & Montalescot, G. 2006, "Drug eluting stents: an updated meta-analysis of randomised controlled trials", *Heart*, vol. 92, no. 5, pp. 641-649.

Rosenberg, W. & Donald, A. 1995, "Evidence based medicine: an approach to clinical problem-solving", *BMJ*, vol. 310, no. 6987, pp. 1122-1126.

Rosmarakis, E. S. & Falagas, M. E. 2007, "Effect of thiazolidinedione therapy on restenosis after coronary stent implantation: A meta-analysis of randomized controlled trials", *American Heart Journal*, vol. 154, no. 1, pp. 144-150.

Rothberg, M. B., Celestin, C., Fiore, L. D., Lawler, E., & Cook, J. R. 2005, "Warfarin plus Aspirin after Myocardial Infarction or the Acute Coronary Syndrome: Meta-Analysis with Estimates of Risk and Benefit", *Annals of Internal Medicine*, vol. 143, no. 4, pp. 241-250.

Rouleau, J. L., Moye, L. A., Pfeffer, M. A., Arnold, J. M., Bernstein, V., Cuddy, T. E., Dagenais, G. R., Geltman, E. M., Goldman, S., Gordon, D., Hamm, P., Klein, M., Lamas, G. A., McCans, J., McEwan, P., Menapace, F. J., Parker, J. O., Sestier, F., Sussex, B., & Braunwald, E. 1993, "A Comparison of Management Patterns after Acute Myocardial Infarction in Canada and the United States", *New England Journal of Medicine*, vol. 328, no. 11, pp. 779-784.

Sackett DL. 2000, "The fall of "clinical research" and the rise of "clinical-practice research".", *Clin Invest Med*, vol. 23, no. 6, pp. 379-381.

Saha, S., Freeman, M., Toure, J., Tippens, K., Weeks, C., & Ibrahim, S. 2008, "Racial and Ethnic Disparities in the VA Health Care System: A Systematic Review", *Journal of General Internal Medicine*, vol. 23, no. 5, pp. 654-671.

Saijo, N. Multinational Trials Reveal Striking Regional Differences. The ASCO Post 1[5]. 2010.
Ref Type: Magazine Article

Saito, S., Hosokawa, G., Tanaka, S., & Nakamura, S. 1999, "Primary stent implantation is superior to balloon angioplasty in acute myocardial infarction: final results of the primary angioplasty versus stent implantation in acute myocardial infarction (PASTA)

trial. PASTA Trial Investigators.", *Catheter Cardiovasc Interv.*, vol. 48, no. 3, pp. 262-268.

Schomig, A., Dibra, A., Windecker, S., Mehilli, J., Suarez de Lezo, J., Kaiser, C., Park, S. J., Goy, J. J., Lee, J. H., Di Lorenzo, E., Wu, J., Juni, P., Pfisterer, M. E., Meier, B., & Kastrati, A. 2007, "A Meta-Analysis of 16 Randomized Trials of Sirolimus-Eluting Stents Versus Paclitaxel-Eluting Stents in Patients With Coronary Artery Disease", *Journal of the American College of Cardiology*, vol. 50, no. 14, pp. 1373-1380.

Schomig, A., Mehilli, J., de Waha, A., Seyfarth, M., Pache, J., & Kastrati, A. 2008, "A Meta-Analysis of 17 Randomized Trials of a Percutaneous Coronary Intervention-Based Strategy in Patients With Stable Coronary Artery Disease", *Journal of the American College of Cardiology*, vol. 52, no. 11, pp. 894-904.

Schulz, K. F. 1995, "Subverting Randomization in Controlled Trials", *JAMA: The Journal of the American Medical Association*, vol. 274, no. 18, pp. 1456-1458.

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. 1995c, "Empirical Evidence of Bias: Dimensions of Methodological Quality Associated With Estimates of Treatment Effects in Controlled Trials", *The Journal of the American Medical Association*, vol. 273, no. 5, pp. 408-412.

Schulz, K. F. & Grimes, D. A. 2002, "Allocation concealment in randomised trials: defending against deciphering", *The Lancet*, vol. 359, no. 9306, pp. 614-618.

Schulz, K. F., Grimes, D. A., Altman, D. G., & Hayes, R. J. 1996, "Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology", *British Medical Journal*, vol. 312, no. 7033, pp. 742-744.

Schwartz, P. J. 1995a, "How reliable are clinical trials?", *European Heart Journal*, vol. 16, no. suppl G, pp. 37-45.

Schwartzman, S. & Morgan, G. J. 2004, "Does route of administration affect the outcome of TNF antagonist therapy?", *Arthritis Res Ther*, vol. 6, no. Suppl 2, p. S19-S23.

Searle, A. 1999, *Introducing Research and Data in Psychology: A Guide to Methods and Analysis* Routledge, London.

Sedrakyan, A., Wu, A. W., Parashar, A., Bass, E. B., & Treasure, T. 2006, "Off-Pump Surgery Is Associated With Reduced Occurrence of Stroke and Other Morbidity as Compared With Traditional Coronary Artery Bypass Grafting: A Meta-Analysis of Systematically Reviewed Trials * Supplemental Appendix I", *Stroke*, vol. 37, no. 11, pp. 2759-2769.

Senn, S. & Harrell, F. 1997, "On Wisdom after the Event", *Journal of Clinical Epidemiology*, vol. 50, no. 7, pp. 749-751.

Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z , et al. 2007. External validation of a measurement tool to assess systematic reviews (AMSTAR). *Public Library of Science One.* vol. 2, no. 12, pp. e1350.

Shen, A. Y.-J., Yao, J. F., Brar, S. S., Jorgensen, M. B., & Chen, W. 2007, "Racial/Ethnic Differences in the Risk of Intracranial Hemorrhage Among Patients With Atrial Fibrillation", *Journal of the American College of Cardiology*, vol. 50, no. 4, pp. 309-315.

Sim, J. & Wright, C. 2000, "Explanatory and Pragmatic Approaches to Clinical Research," in *Research in Health Care: Concepts, Designs and Methods*, J. Sim & C. Wright, eds., Stanley Thomes LTD, Cheltenham, p. 92.

Simes, R. J. 1986, "Publication bias: the case for an international registry of clinical trials", *Journal of Clinical Oncology*, vol. 4, no. 10, pp. 1529-1541.

Sinclair, J. C. & Bracken, M. B. 1994, "Clinically useful measures of effect in binary analyses of randomized trials", *Journal of Clinical Epidemiology*, vol. 47, no. 8, pp. 881-889.

Singh, J. A., Murphy, S., & Bhandari, M. 2010, "Trial sample size, but not trial quality, is associated with positive study outcome", *Journal of Clinical Epidemiology*, vol. 63, no. 2, pp. 154-162.

Singh, R. & Singh, S. 1994, "Research and doctors", *The Lancet*, vol. 344, no. 8921, p. 546.

Sinnaeve, P. R. & Van de Werf, F. J. 2004, "Global patterns of health care for acute coronary syndromes", *Current Opinion in Cardiology*, vol. 19, no. 6, pp. 625-630.

Smith, V., Devane, D., Begley, C., & Clarke, M., 2011, "Methodology in conducting a systematic review of systematic reviews of healthcare interventions", *BioMed Central Medical Research Methodology*, vol. 11, no. 1, pp.1-6.

Song, F., Easterwood, A., Gilbody, S., Duley, L., & Sutton, A. J. 2000, "Publication and Other Selection Biases in Systematic Reviews.", *Health Technology Assessment*, vol. 4, no. 10.

Song, F., Sheldon, T. A., Sutton, A. J., Abrams, K. R., & Jones, D. R. 2001, "Methods for Exploring Heterogeneity in Meta-Analysis", *Evaluation & the Health Professions*, vol. 24, no. 2, pp. 126-151.

Souhami, R. L., Spiro, S. G., & Cullen, M. 1991, "Chemotherapy and radiation therapy as compared with radiation therapy in stage III non-small-cell cancer", *The New England journal of medicine*, vol. 324, no. 16, pp. 1136-1137.

Sterling, T. D., Rosenbaum, W. L., & Weinkam, J. J. 1995, "Publication Decisions Revisited: The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa", *The American Statistician*, vol. 49, no. 1, pp. 108-112.

Stern, J. M. & Simes, R. J. 1997, "Publication bias: evidence of delayed publication in a cohort study of clinical research projects", *BMJ*, vol. 315, no. 7109, pp. 640-645.

Sterne, J., Egger, M., & Moher, D. 2008, "Addressing reporting biases," in *Cochrane Handbook for Systematic Reviews of Interventions*, 5.0.1 edn, J. P. T. Higgins & S. Green, eds., John Wiley & Sons LTD, Chichester.

Sterne, J. & Harbord, R. 2008, "Funnel Plots in Meta-analysis", *The Stata Journal*, vol. 4, no. 2, pp. 127-141.

Straus, S. E. & Jones, G. 2004, "What has evidence based medicine done for us?", *BMJ*, vol. 329, no. 7473, pp. 987-988.

Stroobandt, M. D., Stiels, M. D., & Hoebrechts, R. N. 1997, "Propafenone for Conversion and Prophylaxis of Atrial Fibrillation", *The American Journal of Cardiology*, vol. 79, no. 4, pp. 418-423.

Suryapranata, H., van't Hof, A. W. J., Hoorntje, J. C. A., de Boer, M. J., & Zijlstra, F. 1998, "Randomized Comparison of Coronary Stenting With Balloon Angioplasty in Selected Patients With Acute Myocardial Infarction", *Circulation*, vol. 97, no. 25, pp. 2502-2505.

Svilaas, T., Vlaar, P. J., van der Horst, I. C., Diercks, G. F. H., de Smet, B. J. G. L., van den Heuvel, A. F. M., Anthonio, R. L., Jessurun, G. A., Tan, E. S., Suurmeijer, A. J. H., & Zijlstra, F. 2008, "Thrombus Aspiration during Primary Percutaneous Coronary Intervention", *New England Journal of Medicine*, vol. 358, no. 6, pp. 557-567.

Takano, H., Hasegawa, H., Kuwabara, Y., Nakayama, T., Matsuno, K., Miyazaki, Y., Yamamoto, M., Fujimoto, Y., Okada, H., Okubo, S., Fujita, M., Shindo, S., Kobayashi, Y., Komiyama, N., Takekoshi, N., Imai, K., Himi, T., Ishibashi, I., & Komuro, I. 2007, "Feasibility and safety of granulocyte colony-stimulating factor treatment in patients with acute myocardial infarction", *International Journal of Cardiology*, vol. 122, no. 1, pp. 41-47.

Takeda, S., Nejima, J., Takano, T., Nakanishi, K., Takayama, M., Sakamoto, A., & Ogawa, R. 1998, "Effect of Nasal Continuous Positive Airway Pressure on Pulmonary Edema Complicating Acute Myocardial Infarction", *JAPANESE CIRCULATION JOURNAL*, vol. 62, no. 8, pp. 553-558.

Takeda, S. M. P., Takano, T. M. P., & Ogawa, R. M. P. 1997, "The Effect of Nasal Continuous Positive Airway Pressure on Plasma Endothelin-1 Concentrations in Patients with Severe Cardiogenic Pulmonary Edema.", *Anesthesia & Analgesia*, vol. 84, no. 5, pp. 1091-1096.

Tang, J. L. & Liu, J. L. 2000, "Misleading funnel plot for detection of bias in meta-analysis", *Journal of Clinical Epidemiology*, vol. 53, no. 5, pp. 477-484.

Taylor, S., Bestall, J., Cotter, S., Falshaw, M., Hood, S., Parsons, S., Wood, L., & Underwood, M. 2005, "Clinical Service Organisation for Heart Failure.", *Cochrane Database of Systematic Reviews* no. 2.

Tennant, R., Goens, C., Barlow, J., Day, C., & Stewart-Brown, S. 2007, "A systematic review of reviews of interventions to promote mental health and prevent mental health problems in children and young people", *Journal of Public Mental Health*, vol. 6, no. 1, pp. 25-32.

Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. 2003, "Adjusting for publication bias in the presence of heterogeneity", *Statistics in Medicine*, vol. 22, no. 13, pp. 2113-2126.

The Cochrane Collaboration, 2002a. *Identifying Statistical Heterogeneity.* The Cochrane Collaboration Open Learning Material [online]. Available at: http://www.cochrane-net.org/openlearning/html/mod13-3.htm [Accessed 23/05/09].

The Cochrane Collaboration, 2002b. *Combining Studies.* The Cochrane Collaboration Open Learning Material [online]. Available at: http://www.cochrane-net.org/openlearning/ html/mod12-2.htm [Accessed 23/05/09].

The Cochrane Collaboration, 2011. *Cochrane Reviews.* The Cochrane Collaboration.[online]. Available at:http://www.cochrane.org/cochrane-reviews [Accessed 2 March 2011].

The EPOCH Study Group 2002, "Effects of Pimobendan on Adverse Cardiac Events and Physical Activities in Patients With Mild to Moderate Chronic Heart Failure The Effects of Pimobendan on Chronic Heart Failure Study (EPOCH Study)", *Circulation Journal*, vol. 66, no. 2, pp. 149-157.

The James Lind Library, 2007. *Up-to-date, systematic reviews of all relevant, reliable evidence* [online]. Available at: http://www.jameslindlibrary.org [Accessed 23/05/09].

Thompson, N.J., Snider, D.E., Jr., & Farer, L. S "Variations in National Health Care Practices and Behaviours and Their Influence on International Research", *International Journal of Epidemiology*, vol. 14, no. 3, pp. 457-462.

Thompson, S. G. 1994b, "Systematic Review: Why sources of heterogeneity in meta-analysis should be investigated", *BMJ*, vol. 309, no. 6965, pp. 1351-1355.

Thompson, S. G. 2001, "Why and How Sources of Heterogeneity should be Investigated," in *Systematic Review in Health Care: Meta-analysis in Context.*, 2nd edn, M. Egger, G. D. Smith, & D. G. Altman, eds., BMJ Publishing Group, London, pp. 157-175.

Thompson, S. G. & Higgins, J. P. T. 2002, "How should meta-regression analyses be undertaken and interpreted?", *Statistics in Medicine*, vol. 21, no. 11, pp. 1559-1573.

Thomson, R., McElroy, H., & Sudlow, M. 1998, "Guidelines on anticoagulant treatment in atrial fibrillation in Great Britain: variation in content and implications for treatment", *British Medical Journal*, vol. 316, no. 7130, p. 509.

Thornton, A. & Lee, P. 2000a, "Publication bias in meta-analysis: its causes and consequences", *Journal of Clinical Epidemiology*, vol. 53, no. 2, pp. 207-216.

Tian, L. 2007, "Inferences on standardized mean difference: the generalized variable approach", *Statistics in Medicine*, vol. 26, no. 5, pp. 945-953.

Torgerson, C. 2003, "Developing a Protocol; Searching and Screening; Data Extraction," in *Systematic reviews*, C. Torgerson, ed., Continuum, London, p. 26.

Towheed, T. Is the Jadad Score the Proper Evaluation of Trials? Dr. Towheed replies. *The Journal of Rheumatology*. vol. 33, no. 8, pp. 1711. 2006.

Tramer, M. R., Reynolds, D. J., Moore, R. A., & McQuay, H. J. 1997, "Impact of covert duplicate publication on meta-analysis: a case study", *British Medical Journal*, vol. 315, no. 7109, pp. 635-640.

Turkelson, C. & Hughes, J. 2006, "The Role of Systematic Reviews in Evidence-Based Practice", *American Academy of Orthopaedic Surgeons Bulletin*, vol. 54, no. 4, pp. 1-5.

Turner, T., Misso, M., Harris, C., & Green, S. 2008, "Development of evidence-based clinical practice guidelines (CPGs): comparing approaches", *Implementation Science*, vol. 3, no. 1, p. 45.

United Nations Statistics Division. Standard Country or Area Codes and Geographical Regions for Statistical Use. United Nations . 2008. United Nations. 2008. Ref Type: Electronic Citation

Utts, J. & Heckard, R. 2006, "Testing Hypotheses about Proportions," in *Statistical Ideas and Methods*, J. Utts & R. Heckard, eds., Thomson Higher Education, California, p. 363.

Van de Werf, F., Topol, E. J., Lee, K. L., Woodlief, L. H., Granger, C. B., Armstrong, P. W., Barbash, G. I., Hampton, J. R., Guerci, A., Simes, R. J., Ross, A. M., Califf, R. M., & the, G., I 1995b, "Variations in Patient Management and Outcomes for Acute Myocardial Infarction in the United States and Other Countries", *The Journal of the American Medical Association*, vol. 273, no. 20, pp. 1586-1591.

Van Den Eeden, S. K., Tanner, C. M., Bernstein, A. L., Fross, R. D., Leimpeter, A., Bloch, D. A., & Nelson, L. M. 2003, "Incidence of Parkinson's Disease: Variation by Age, Gender, and Race/Ethnicity", *American Journal of Epidemiology*, vol. 157, no. 11, pp. 1015-1022.

Van Den Noortgate, W. & Onghena, P. 2003, "Multilevel Meta-analysis: A Comparison with Traditional Meta-analytic Procedures.", *Educational and Psychological Measurement*, vol. 63, no. 5, pp. 765-790.

van der Wouden, J. C., Bueving, H. J., & Poole, P. 2005, "Preventing influenza: An overview of systematic reviews", *Respiratory Medicine*, vol. 99, no. 11, pp. 1341-1349.

van Driel, M. L., De Sutter, A., Deveugele, M., Peersman, W., Butler, C. C., De Meyere, M., De Maeseneer, J., & Christiaens, T. 2006, "Are Sore Throat Patients Who Hope for Antibiotics Actually Asking for Pain Relief?", *Annals of Family Medicine*, vol. 4, no. 6, pp. 494-499.

Vickers, A., Goyal, N., Harland, R., & Rees, R. 1998b, "Do Certain Countries Produce Only Positive Results? A Systematic Review of Controlled Trials", *Controlled Clinical Trials*, vol. 19, no. 2, pp. 159-166.

Vital, F., Saconato, H., Ladeira, M., Sen, A., Hawkes, C., Soares, B., Burns, K., & Atallah, Á. 2008, "Non-invasive positive pressure ventilation (CPAP or bilevel NPPV) for cardiogenic pulmonary edema.", *Cochrane Database of Systematic Reviews 2005* no. 3.

Vogel, N., Burnand, B., Vial, Y., Ruiz, J., Paccaud, F., & Hohlfeld, P. 2000, "Screening for gestational diabetes: variation in guidelines", *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 91, no. 1, pp. 29-36.

Ward, S., Lloyd Jones, M., Pandor, A., Holmes, M., Ara, R., Ryan, A., Yeo, W., & Payne, N. 2007, "A systematic review and economic evaluation of statins for the prevention of coronary events", *Health Technology Assessment.*, vol. 11, no. 14, pp. 1-160.

Welsh, R. C., Chang, W., Goldstein, P., Adgey, J., Granger, C. B., Verheugt, F. W. A., Wallentin, L., Van de Werf, F., & Armstrong, P. W. 2005b, "Time to treatment and the impact of a physician on prehospital management of acute ST elevation myocardial infarction: insights from the ASSENT-3 PLUS trial", *Heart*, vol. 91, no. 11, pp. 1400-1406.

White, H. D. 2000, "International differences: selection, noise, or real?", *European Heart Journal*, vol. 21, no. 5, pp. 339-342.

Whitlock, R. P., Chan, S., Devereaux, P. J., Sun, J., Rubens, F. D., Thorlund, K., & Teoh, K. H. T. 2008, "Clinical benefit of steroid use in patients undergoing cardiopulmonary bypass: a meta-analysis of randomized trials", *European Heart Journal*, vol. 29, no. 21, pp. 2592-2600.

Wijeysundera, H. C., Vijayaraghavan, R., Nallamothu, B. K., Foody, J. M., Krumholz, H. M., Phillips, C. O., Kashani, A., You, J. J., Tu, J. V., & Ko, D. T. 2007, "Rescue Angioplasty or Repeat Fibrinolysis After Failed Fibrinolytic Therapy for ST-Segment

Myocardial Infarction: A Meta-Analysis of Randomized Trials", *Journal of the American College of Cardiology*, vol. 49, no. 4, pp. 422-430.

Wijeysundera, H. C., You, J. J., Nallamothu, B. K., Krumholz, H. M., Cantor, W. J., & Ko, D. T. 2008, "An early invasive strategy versus ischemia-guided management after fibrinolytic therapy for ST-segment elevation myocardial infarction: A meta-analysis of contemporary randomized controlled trials", *American Heart Journal*, vol. 156, no. 3, pp. 564-572.

Wolfe, C. D. A., Tilling, K., Beech, R., & Rudd, A. G. 1999, "Variations in Case Fatality and Dependency From Stroke in Western and Central Europe", *Stroke*, vol. 30, no. 2, pp. 350-356.

Wood, L., Egger, M., Gluud, L. L., Schulz, K. F., Juni, P., Altman, D. G., Gluud, C., Martin, R. M., Wood, A. J., & Sterne, J. A. 2008, "Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study", *British Medical Journal (Clinical research ed.)*, vol. 336, no. 7644, pp. 601-605.

Woods, K. L., Fletcher, S., Roffe, C., & Haider, Y. 1992, "Intravenous magnesium sulphate in suspected acute myocardial infarction: results of the second Leicester Intravenous Magnesium Intervention Trial (LIMIT-2)", *The Lancet*, vol. 339, no. 8809, pp. 1553-1558.

World Health Organisation, 2011.*Cardiovascular Disease (CVD).* World Health Organisation [online]. Available at: http://www.who.int/cardiovascular_diseases/en/ index.html [Accessed 21/1/11].

World Health Organisation, 2010. Cardiovascular Diseases Fact Sheet. World Health Organisation [online]. Available at: http://www.who.int/mediacentre/factsheets/fs317/ en/index.html [Accessed 15/06/10].

Yusuf, S., Flather, M., Pogue, J., Hunt, D., Varigos, J., Piegas, L., Avezum, A., Anderson, J., Keltai, M., Budaj, A., Fox, K., & Ceremuzynski, L. 1998, "Variations between countries in invasive cardiac procedures and outcomes in patients with suspected unstable angina or myocardial infarction without initial ST elevation", *The Lancet*, vol. 352, no. 9127, pp. 507-514.

Zed, P. J., Loewen, P. S., & Robinson, G. 1999, "Medication-induced headache: overview and systematic review of therapeutic approaches", *The Annals of Pharmacotherapy*, vol. 33, no. 1, pp. 61-72.

Zhang, D., Freemantle, N., & Cheng, K. K. 2011, "Are randomized trials conducted in China or India biased? A comparative empirical analysis", *Journal of Clinical Epidemiology*, vol. 64, no. 1, pp. 90-5.

Zhang, J. 2008, The efficacy of peer-mediated interventions for promoting social interactions among young children with autism spectrum disorders: A meta-analysis. *Dissertation.* Tennessee Technological University

Zohlnhofer, D., Dibra, A., Koppara, T., de Waha, A., Ripa, R. S., Kastrup, J., Valgimigli, M., Schomig, A., & Kastrati, A. 2008, "Stem Cell Mobilization by Granulocyte Colony-Stimulating Factor for Myocardial Recovery After Acute Myocardial Infarction: A Meta-Analysis", *Journal of the American College of Cardiology*, vol. 51, no. 15, pp. 1429-1437.

Zohlnhofer, D., Ott, I., Mehilli, J., Schomig, K., Michalk, F., Ibrahim, T., Meisetschlager, G., von Wedel, J., Bollwein, H., Seyfarth, M., Dirschinger, J., Schmitt, C., Schwaiger, M., Kastrati, A., Schomig, A., & for the REVIVAL- 2006, "Stem Cell Mobilization by Granulocyte Colony-Stimulating Factor in Patients With Acute Myocardial Infarction", *The Journal of the American Medical Association*, vol. 295, no. 9, pp. 1003-1010.