Extending the Clinical and Economic Evaluations of a Randomised Controlled Trial: The IONA Study

Neil James Kerr Henderson

A Dissertation Submitted to the
Faculty of Information and Mathematical Sciences
at the University of Glasgow
for the degree of
Doctor of Philosophy

Department of Statistics

October 2008

Abstract

In modern society people are concerned about their state of health and if they do unfortunately become ill they want the best possible treatment to be made available to them. In order to satisfy these demands new treatments have to be developed. This can be a long and expensive process. Before any new treatment can come to market it has to be proved to be both clinically effective and economically cost-effective. With limited health care resources the cost-effectiveness of treatments is becoming ever more relevant.

In order to show whether a treatment is clinically effective a clinical trial is carried out and this is now usually accompanied by an economic evaluation, so that the cost-effectiveness of the treatment can be assessed. When a clinical trial aimed at preventing clinical events is analysed, a time-to-first event analysis is often performed together with a cost-effectiveness analysis. These analyses do not always make the best use of the large amounts of patient information recorded during the clinical trial. Using the randomised controlled trial (RCT) the Impact Of Nicorandil in Angina (IONA) as an exemplar, ways in which the clinical and economic evaluations of clinical trials can be expanded are explored.

There are three main parts of this thesis. Firstly, following a more detailed introduction in Chapter 1, in Chapters 2 and 3 the IONA Study is introduced

and the main clinical results of the study are given. Secondly, in Chapters 4, 5 and 6 the fact that patients could suffer more than one clinical endpoint is considered. The models that can be used to incorporate the recurrent events are introduced and then applied to the data from the IONA Study. Following on from this, through the simulation of recurrent event data, the performance of the models under different known conditions is assessed. Thirdly, in Chapters 7 and 8 an introduction to health economics is given and following this the main results of the economic evaluation of the IONA Study are presented. Areas in which the results of the economic evaluation can be expanded are then investigated. Finally, in Chapter 9 there is a discussion of the work as a whole and areas where there would be the possibility of further work.

Acknowledgements

I would firstly like to thank my two supervisors Prof. Ian Ford and Dr Andrew Walker. I would especially like to thank Prof. Ford and the Robertson Centre for Biostatistics for funding my research. Prof. Ford has provided me with invaluable help and support throughout the process of completing my thesis. He has always found time to see me and offer his wealth of knowledge. Having never done any economics before let alone health economics I would like to thank Dr Walker for being very patient when I did not understand things and being there if I needed someone to talk to.

Secondly, I would like to thank all the staff in the RCB for being so welcoming and friendly. I would also like to thank all the staff in the Statistics Department. Thanks must also go to Harper Gilmour for being my third supervisor. I would also like to thank all my friends, especially Helen and her lilac pen, for always taking an interest in my work and how it was going. I would also like to say a big thank you to Lindsay for having to endure sharing an office with me for so long. The spelling help was also greatly appreciated!

I would like to thank my Dad and sister for supporting me through the whole process. Without them I could not have done it.

Finally, this is for my Mum.

Declaration

This thesis has been composed by myself and it has not been submitted in any previous application for a degree. The work reported within was executed by myself, unless otherwise stated.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Clinical trials play a critical role in the development of new types of treatments and are important to many different groups including the pharmaceutical industry, who develop the new products, the National Health Service (NHS), who go on to implement the use of such products in the UK, and patients, who are the ones who will ultimately benefit from the development of new treatments by using them. During clinical trials large amounts of data are collected and recorded on the patients who are taking part. In clinical trials where the time until a specific event is important survival analysis is used to analyse the data. This is usually based on a Cox proportional hazards model (Cox, 1972), with the only variable included in the model being the treatment that is being investigated. As a result a large amount of information which has been collected on patients is unused.

In recent times the importance of carrying out of an economic evaluation alongside clinical trials has increased due to the budget constraints that healthcare providers like the NHS are under and the high cost of some of the new

treatments which are coming to market, with the resulting high net cost of prescribing them to patients. When an economic evaluation is undertaken commonly only the benefits a treatment gives and the costs both saved and accrued are considered and, as with a time-to-first event analysis when the clinical outcome is being investigated, any additional information that is available on patients is not considered or used in the analysis.

As during the course of a clinical trial large amounts of information will be collected this information could be included in the analysis of the trial, by looking for variables that may be predictive of whether patients will suffer events and not solely concentrating any analysis on the first event that patients suffer and the time to that first event but including any further events that patients go onto suffer. As well as making best use of all the information collected on patients during the investigation of the clinical outcome the recorded patient data could also be incorporated into an extended economic evaluation of the trial.

The basis of this work is the randomised controlled trial (RCT): the Impact Of Nicorandil in Angina (IONA) Study and the resultant clinical and economic evaluations that were undertaken on the data from the study and have been published. (The IONA Study Group, 2002; Walker et al., 2006) Using the published analyses as a base both the clinical and economic evaluations are expanded on to incorporate all the data that was recorded on patients before and during the study. Hence, the analysis is no longer concentrated solely on the standard time-to-first event analysis and the simple economic evaluation that have already been undertaken.

In the paper by The IONA Study Group (2002) the main clinical results of the study were presented in the form of a time-to-first event analysis with the only

explanatory variable fitted in the model being whether patients were receiving nicorandil, the study drug, or a placebo. The study design and implementation as well as the main findings of the IONA Study are given in Chapter 2. The additional demographic and clinical variables recorded on patients were not included in the analysis nor were any further clinical endpoints that patients went onto suffer after they had experienced their initial event.

In a follow up paper by The IONA Study Group (2005) the univariate significance of all the baseline variables recorded on patients who took part in the IONA Study, including whether they were in the nicorandil or placebo group, was explored and from these findings multivariable models predicting the risk of patients suffering a primary as well as a secondary endpoint of the study were built. In Chapter 3, following on from and expanding the results shown in Chapter 2, the univariate significance of the recorded baseline variables is considered for both the primary endpoint of the IONA Study as well as the gastrointestinal (GI) events that patients suffered, GI events are a potential side-effect of being treated with nicorandil. Multivariable predictive models for both the primary endpoint as well as GI events patients suffered are derived. The set of baseline variables used in Chapter 3 is marginally different from the set used in the paper by The IONA Study Group (2005) resulting in slightly different models. The GI events patients suffered were not considered in the paper by The IONA Study Group (2005). Further models are built from sub-groups of the baseline variables so that a full picture of the variables and different combinations of variables that are prognostic for both types of events can be seen.

Following on from expanding the clinical evaluation of the study the implications of incorporating into the analysis the fact that patients could suffer clinical

endpoints more than once will be investigated. As previously stated the main results of the IONA Study were based on a time-to-first event analysis. After patients had experienced a first primary endpoint they still continued to be monitored and attended study visits until study closedown or they withdrew from the study for any other reason, unless the endpoint had been death when, for obvious reasons, the study period automatically ended for the patient.

During the period after patients suffered a first event and the end of the study they were at risk of suffering further events and many patients did go on to suffer a second event and then further subsequent events. All the serious adverse events that patients suffered during the study were documented and therefore all the serious GI events that patients suffered were recorded. During the course of the study some patients suffered multiple GI events. In the time-to-first event analysis none of the additional events that patients went on to suffer were considered. Recurrent event models introduced in Chapter 4, can be fitted to the data from the IONA Study so that the multiple events that patients suffered can be included in the analysis. The results are given in Chapter 5.

Having introduced and used the recurrent event models on real life data from the IONA Study it was seen that the different models have different underlying assumptions and can give different results. Although in the case of the IONA Study the results were broadly similar. In Chapter 6 through the simulation of recurrent event data it is investigated how under known conditions recurrent event models perform. If the results and generalisability of a clinical trial are to be improved by the inclusion of any recurrent events that patients suffered in the analysis it is important that the most appropriate model is used in each situation to ensure the robustness of the findings.

An introduction to health economics and some of the general terms and techniques used in this area is given in Chapter 7. In Chapter 8 the main results of the economic evaluation carried out on the IONA Study are presented. In the paper by Walker et al. (2006) the economic evaluation of the IONA Study was presented as a Cost-Effectiveness Analysis (CEA) with the results being reported in the form of an Incremental Cost-Effectiveness Ratio (ICER). To take into account the changing costs of procedures or equipment and hospital costs univariate sensitivity analysis of the results were undertaken. In the economic evaluation the baseline variables recorded on patients were not incorporated into the analysis, similarly to the case of the main clinical results. In order to evaluate the ICER the only information that was required were the numbers of primary endpoints suffered in both the nicorandil and placebo groups as well as the costs associated with each treatment group. The results of the cost-effectiveness analysis of the IONA Study showed that the cost per primary endpoint prevented was low and in comparison to similarly evaluated treatments it was cost-effective to treat a population of patients similar to those included in the IONA Study with nicorandil. (Walker et al., 2006). However, even within the IONA population it is likely that for some sub-groups of patients treatment would be more cost-effective than for others. It might be possible to improve the cost-effectiveness of a treatment by targeting it to a sub-group of the potential patient population where it would be most beneficial in terms of resource use and hence maximise the benefits that can be gained from limited healthcare budgets. Importantly, IONA recruited particularly high risk angina patients. Evaluation of the ICER in the lowest risk sub-group of the high risk patients recruited in IONA might be more representative of the type of results that might be achieved in the type of

patients excluded from the study. The cost-effectiveness of a treatment, in this case nicorandil, could possibly be improved by targeting it to those patients who are at the highest risk of suffering an event so therefore more likely to benefit from treatment, assuming that treatment benefit is proportional to risk and than harmful effects are not also increased.

All patients who were recruited in the IONA Study were at a high risk of suffering CHD events but in order to identify those patients who were at the highest risk of suffering events the baseline variables recorded on patients, which related to coronary heart disease (CHD) in the case of the IONA Study, are used. The risk of patients suffering events needs to be calculated and in the case of the IONA Study this can be done by means of the multivariable predictive model for the primary endpoint. This is investigated in Chapter 3. The already high risk patients recruited in IONA can be split into sub-groups by their calculated level of risk of suffering a CHD event and separate cost-effectiveness analyses then carried out. The same principle economic techniques are being used as when the overall cost-effectiveness of a treatment is being investigated but this time the additional information recorded on patients is being incorporated into the analysis. This enables those bodies that make recommendations about how and to whom treatments should be prescribed to see the overall cost-effectiveness of the treatment as well as how the cost-effectiveness changes when only those patients who are at the highest risk of suffering events, and possibly being the most likely to benefit from treatment, are treated compared to the situation where only those patients at a low risk of suffering an event are treated.

Treatments should benefit patients that are being prescribed them but they also have the potential to cause patients to experience side-effects, for example

the GI events that patients suffered during the IONA Study. These side-effects may have implications on the cost-effectiveness of a treatment as well as health implications. If there was a link between the benefits, the reduction in the number of CHD events suffered, and the side-effects of nicorandil, the GI events suffered, is investigated by looking at whether there was any relationship between the risks of these events. This is done using the multivariable predictive model for the primary endpoint found in Chapter 3 as well as the multivariable predictive model for the GI events that patients suffered. The balance between harm, benefits and risk in a clinical trial is an important factor to consider when prescribing recommendations are laid out for the treatment evaluated in a clinical trial.

Finally, in Chapter 9 a summary of the main results and findings is given. The implications of these results and findings and the impact they could have on the future analyses of clinical trials are discussed. In addition, a brief introduction is given to areas where there would be the possibility of further work on the evaluations of clinical trials.

Throughout this work it will be shown how both the clinical and economic evaluations of clinical trials can be extended and expanded upon. This will be achieved by making the best use of information recorded on patients during a clinical trial by incorporating it into both the clinical and economic evaluations. Thus, allowing more robust conclusions to be drawn from clinical trials.

Chapter 2

The Impact Of Nicorandil in Angina Study

In the UK 25% of all deaths are attributable to coronary heart disease and, of patients presenting with a first myocardial infarction, 25% have a history of stable angina. (Gandhi, 1997) If a treatment could be shown to have cardioprotective effects in a large scale randomised controlled trial of patients suffering from stable angina the benefits could be potentially significant. It was thought that nicorandil, an existing antianginal drug, might have cardioprotective effects in patients suffering from stable angina and the IONA Study was carried out to test this hypothesis. In this Chapter background information on nicorandil will be given as well as the design and implementation of the study and details of adverse events that patients suffered. Results for both the primary and secondary endpoints of the study will be described. Additionally, the side-effects of nicorandil will be investigated.

2.1 Nicorandil

The drug nicorandil has been marketed in Japan since 1984 and in the UK it was licensed for the long term symptomatic treatment of chronic angina pectoris. Nicorandil is a nicotinamide ester that works in two different ways, firstly it opens ATP-sensitive potassium channels (K_{ATP}) and secondly it has similar properties to a nitrate. Both these properties enable it to be used for the treatment of angina. (The IONA Study Group, 2002) Nicorandil has been shown to have similar antianginal efficacy to oral nitrates, β-blockers and calcium antagonists. (Döring, 1992; Di Somma et al., 1993; The SWAN Study Group, 1999)

2.2 The IONA Study Design

The Impact Of Nicorandil in Angina (IONA) Study was a randomised double-blind placebo controlled clinical trial in patients who were suffering from stable angina of effort. The aim of the study was to test the hypothesis that nicorandil would reduce the incidents of coronary events in both men and women. The IONA Study had a composite primary endpoint of coronary heart disease (CHD) death, non-fatal myocardial infarction (MI) or unplanned hospital admission for cardiac chest pain. There was also a secondary endpoint of coronary heart disease death or non-fatal myocardial infarction. (The IONA Study Group, 2002) The full definitions of the component parts of the primary and secondary endpoints are given in Appendix A. Randomisation for the IONA Study started in May 1998 and ended in August 2000 with patients being recruited from 226 centres in the UK. These were both primary care and hospital centres, with there being

approximately equal numbers of patients recruited from both. The mean follow-up of the study was 1.6 years (standard deviation (SD) 0.5 years), with the range of follow-up being from 1 to 3 years and the last study visit took place in August 2001. In total 5,126 patients were randomised in the IONA Study, of these 2,565 were assigned to 20 mg of nicorandil twice daily and 2,561 to an identical placebo.

There were a number of inclusion and exclusion criteria before patients could be randomised in the IONA Study. Both men and women were recruited for the study, with men having to be aged $\geq$ 45 years old and women $\geq$ 55 years old. Patients could be existing sufferers of chronic angina of effort or could be newly diagnosed. There were no standard antianginal therapies set out in the study protocol and the medication that each individual patient received was judged to be the optimum for them by the study investigator. As the patients were suffering from angina they were all expected to be receiving at least one symptom relieving oral antianginal drug: a long acting nitrate formulation, a β-blocker or a calcium channel blocker. All patients who were randomised should also have experienced one of the following:

1. previous myocardial infarction;
2. previous coronary artery bypass graft;
3. coronary heart disease proven by angiography or a documented positive exercise test ($\geq$ 1 mm ST depression) in the previous two years.

The last of the three inclusion criteria was required to be accompanied by at least one of the following: left ventricular hypertrophy on ECG (tall R in aVL, SV1 + RV6 $>$ 35 mm, lateral T inversion); evidence of left ventricular dysfunction

(ejection fraction $\leq$ 45% or end diastolic dimension $>$ 5.5 cm); age $\geq$ 65 years; diabetes (types I or II); hypertension (treated, and/or systolic blood pressure $>$ 160 mm Hg or diastolic blood pressure $>$ 95 mm Hg); documented evidence of other vascular disease (stroke, transient ischaemic attack requiring hospital admission, peripheral arterial disease). (The IONA Study Group, 2001) The rationale behind recruiting patients with additional risk factors was so that the patients would be at high risk of suffering a primary endpoint during the period of the study. Patients with any of the following were excluded from taking part in the study:

1. uncontrolled cardiac failure or arrhythmias;
2. unstable angina;
3. coronary artery bypass graft or myocardial infarction in the previous three months
4. percutaneous transluminal coronary angioplasty in the previous six months;
5. uncontrolled hypertension (systolic blood pressure $>$ 180 mm Hg or diastolic blood pressure $>$ 110 mm Hg);
6. the presence of other diseases that in the investigator's opinion would reduce life expectancy or influence significantly the patient's cardiovascular condition;
7. current treatment with nicorandil;
8. current treatment with sulfonylureas (this group of antidiabetic drugs blocks potassium channel opening);

9. pregnancy or lactation;

10. legal incapacity or limited legal capacity;

11. participation in another clinical study within the previous 30 days;

12. presence of contraindications to the study medication;

13. known drug or alcohol abuse.

All subjects provided written informed consent.

Before patients were randomised they had the following baseline characteristics recorded: age, height, weight, body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate, recorded as beats per minute (bpm), sex, whether they were diabetic, hypertensive, a current smoker, whether they had suffered a previous myocardial infarction, coronary artery bypass graft (CABG), percutaneous transluminal coronary angioplasty (PTCA), angiogram, stroke, whether they had been admitted to hospital due to a transient ischemic attack (TIA), whether they had a history of left ventricular hypertrophy (LVH), of left ventricular dysfunction (LVD) or of peripheral vascular disease (PVD). In addition patients had the severity of their angina assessed using the Canadian Cardiovascular Society Functional classification of angina (CCSF), this is a four point scale with patients classified as level I having the mildest level of angina and patients classified as level IV having the most severe level of angina. The baseline characteristics for the patients who took part in the IONA Study can be seen in Table 2.1.

When patients were randomised they were assigned to nicorandil, 10 mg twice daily, or identical placebo in a double-blind manner for the first two weeks of

Table 2.1: Baseline characteristics for patients in the IONA Study shown with mean (SD) for continuous variables and number of patients (%) for categorical variables

	Nicorandil (n = 2,565)	Placebo (n = 2,561)
Continuous Risk Factors, Mean (SD)		
Age (years)	67 (8)	67 (9)
Height (cm)	169 (9)	169 (9)
Weight (kg)	79 (15)	80 (15)
BMI (kg/m^2)	28 (5)	28 (4)
SBP (mm Hg)	138 (19)	138 (19)
DBP (mm Hg)	79 (10)	79 (10)
Heart rate (bpm)	66 (12)	67 (12)
Categorical Risk Factors, n (%)		
Male	1,962 (76.5%)	1,948 (76.1%)
Diabetic	197 (7.7%)	232 (9.1%)
Hypertensive	1,197 (46.7%)	1,178 (46.0%)
Current Smoker	417 (16.3%)	425 (16.6%)
History of Vascular Disease, n (%)		
Previous MI	1,696 (66.1%)	1,682 (65.7%)
Previous CABG	572 (22.3%)	590 (23.0%)
Previous PTCA	360 (14.0%)	392 (15.3%)
Previous Angiogram	1,508 (58.8%)	1,525 (59.6%)
Previous Stroke	134 (5.2%)	116 (4.5%)
Hospital Admission for TIA	47 (1.8%)	55 (2.1%)
History of LVH	259 (10.1%)	260 (10.2%)
History of LVD	230 (9.0%)	206 (8.1%)
History of PVD	289 (11.3%)	335 (13.1%)
CCSF Classification for Angina, n (%)		
Level I	671 (26.2%)	692 (27.0%)
Level II	1,695 (62.6%)	1,583 (61.9%)
Level III	272 (10.6%)	275 (10.7%)
Level IV	15 (0.6%)	9 (0.4%)

the study. After two weeks a study visit was scheduled where the tolerance to nicorandil or the placebo was assessed. For those patients who were tolerating 10 mg twice daily of nicorandil the dose of nicorandil was increased to 20 mg twice daily or matching placebo. Eight weeks after randomisation there was a second study visit where again tolerance to the drug as well as compliance to the treatment regime were assessed. The compliance of patients to their treatment regime was assessed by counting the number of pills, nicorandil or placebo, they returned at study visits. If required, the dosage of nicorandil could be down titrated at this visit. After the study visit at eight weeks further visits took place every sixteen weeks until the study follow-up came to an end. Patients were still followed until study closedown even if they had ceased to comply with their study treatment and all patients were expected to have a study closedown visit. At all study visits compliance was assessed as well as all serious adverse events and adverse events leading to discontinuation of study drug. In addition the angina status of patients was assessed at each visit by the CCSF classification of angina and also the concomitant drug treatments patients were being treated with were recorded. Common types of cardiovascular medications that patients who took part in IONA Study were being treated with were: ACE inhibitors, antiplatelets (and aspirin), β-blockers, long acting nitrates, diuretics, statins and calcium channel blockers (CCBs). The diuretics could be further divided into ordinary and loop diuretics. In addition the three most prescribed CCBs were amlodipine, diltiazem and nifedipine. The numbers of patients being treated with the different cardiovascular medications at baseline can be seen in Table 2.2.

The IONA Study had a target samples size of 5,000 and with this sample size the study had 80% power, with a 5% significance level, to detect a 20% reduction

Table 2.2: The numbers (%) of patients being prescribed the different cardiovascular medications at baseline

Cardiovascular Medication	Nicorandil (n = 2,565)	Placebo (n = 2,561)
ACE Inhibitors	739 (28.8%)	759 (29.6%)
Antiplatelets (and Aspirin)	2283 (89.0%)	2238 (87.4%)
β-Blockers	1469 (57.3%)	1433 (56.0%)
Long Acting Nitrates	1359 (53.0%)	1358 (53.0%)
Ordinary Diuretics	273 (10.6%)	271 (10.6%)
Loop Diuretics	542 (21.3%)	528 (20.6%)
Statins	1449 (56.5%)	1486 (58.0%)
Calcium Channel Blockers	1411 (55.0%)	1397 (54.6%)
Most commonly prescribed Calcium Channel Blockers		
Amlodipine	501 (19.5%)	472 (18.4%)
Diltiazem	613 (23.9%)	638 (24.9%)
Nifedipine	204 (8.0%)	199 (7.8%)

in the rate of the primary endpoint, assuming the placebo event rate would be 13%. There was 80% power, again with 5% significance level, to detect a 25% reduction in the rate of the secondary endpoint assuming a placebo event rate of 8%. (The IONA Study Group, 2002) The assumption regarding the placebo event rate for the secondary endpoint turned out to be incorrect. Instead of the assumed 8% event rate the actual event rate in the placebo group was substantially lower at 5.2%. As a result the IONA Study was underpowered to show statistical significance when analysing the secondary endpoint.

2.3 Adverse Events and Reasons for Study Non-Completion

As was stated in Section 2.2, at all study visits serious adverse events and adverse events leading to discontinuation of study drug that patients suffered were recorded. The serious adverse events were coded using a proprietary dictionary provided by Merck KGaA, Merck KGaA were one of the sponsors of the IONA Study along with Aventis Pharma and Chugai Pharmaceutical Company. There were a number of different types of adverse events recorded which led to patients discontinuing their study drug, the most common of which was headaches. It is known that nicorandil when first taken can cause patients to suffer from headaches. In total 445 patients stopped their study drug due to suffering from headaches, of which 364 were in the nicorandil group (14.2% of the group) and 81 in the placebo group (3.2% of the group). This reinforces the fact that headaches are a side-effect of nicorandil. The reasons for patient non-completion, excluding death, are shown in Table 2.3. Also shown are the numbers of patients who were lost to follow-up and who withdrew informed consent.

Table 2.3: Reasons and patient numbers (%) for non-completion in the IONA Study

Reason for Non-Completion	Nicorandil (n = 2,565)	Placebo (n = 2,561)
Discontinued intervention	1,003 (39.1%)	809 (31.6%)
Reason for discontinuation		
Headache	364 (14.2%)	81 (3.2%)
Other non-fatal adverse events	342 (13.3%)	375 (14.6%)
Non-adverse event reason	297 (11.6%)	353 (13.8%)
Lost to follow-up	2 (0.04%)	2 (0.04%)
Withdrew informed consent	44 (1.7%)	41 (1.6%)

Potentially more severe side-effects were the GI events and resultant hospitalisations that nicorandil may cause. The events that were defined as GI events were serious adverse events that were coded as occurring within the gastrointestinal body system using the coding dictionary provided by Merck KGaA. The lengths of any resultant hospitalisations due to serious adverse events were also recorded. Therefore, even though GI events were not a predefined study outcome the numbers and types of all GI hospitalisations that patients suffered during their participation in the IONA Study were recorded. The statistical power of finding a difference in the rate of GI events that patients suffered between the nicorandil and placebo groups was not considered before the start of the study. The data that were collected will be used to investigate whether nicorandil does, in this instance, increase the likelihood of suffering GI events. It should be noted that all serious adverse events were recorded regardless of whether or not they could definitely be attributed to the study drug that patients were taking.

2.4 Results of the IONA Study

To analyse the primary and secondary endpoints of the IONA Study as well as the GI events a Cox proportional hazards model was used to estimate the hazard ratio (HR) associated with random treatment allocation with accompanying 95% confidence intervals (CI). (Cox, 1972) In all three of these analyses the only covariate that was fitted in the model was the randomised treatment patients were receiving. When fitting a Cox model there is an underlying assumption of proportional hazards between the treatment groups. The assumption of proportional hazards was checked for the three types of event analysed and found to be valid

in each case. The analyses were carried out on an intention-to-treat basis. For the patients who were lost to follow-up, withdrew informed consent for follow-up or died from none CHD reasons their observation times were censored at the time of their last study visit. For all other patients clinical outcomes were sought until one of the following occurred: death or study closedown. The results shown in Sections 2.4.1 and 2.4.2 can be found in the original results paper. In the paper only the total number of GI events suffered by patients during the study was given and no formal analysis was conducted.

The majority of patients suffered an unplanned hospital admission as their first primary endpoint but patients could go on to either die due to CHD or suffer a non-fatal MI. When the secondary endpoint was being investigated the

Table 2.4: The breakdown and numbers (%) of the component parts for the first primary and secondary endpoints that patients suffered during the IONA Study

Endpoint Component Part	Nicorandil (n = 2,565)	Placebo (n = 2,561)
Primary Endpoint		
CHD death	36 (1.4%)	54 (2.1%)
Non-fatal MI	47 (1.8%)	60 (2.3%)
Unplanned hospital admission for cardiac chest pain	254 (9.9%)	284 (11.1%)
Secondary Endpoint		
CHD death	51 (2.0%)	62 (2.4%)
Non-fatal MI	56 (2.2%)	72 (2.8%)

unplanned hospital admission that patients suffered were not considered. Therefore, for the first events that patients suffered, included in the analysis contained in Sections 2.4.1 and 2.4.2, the numbers of patients who suffered the component parts common to both the primary and secondary endpoints were greater for the

secondary endpoint. The breakdown of the component parts and the numbers of patients who suffered them for the first primary and secondary endpoints are shown in Table 2.4.

2.4.1 Primary Endpoint

The Kaplan-Meier estimates for the primary endpoint split by treatment group can be seen in Figure 2.1. As the time from randomisation increases there is

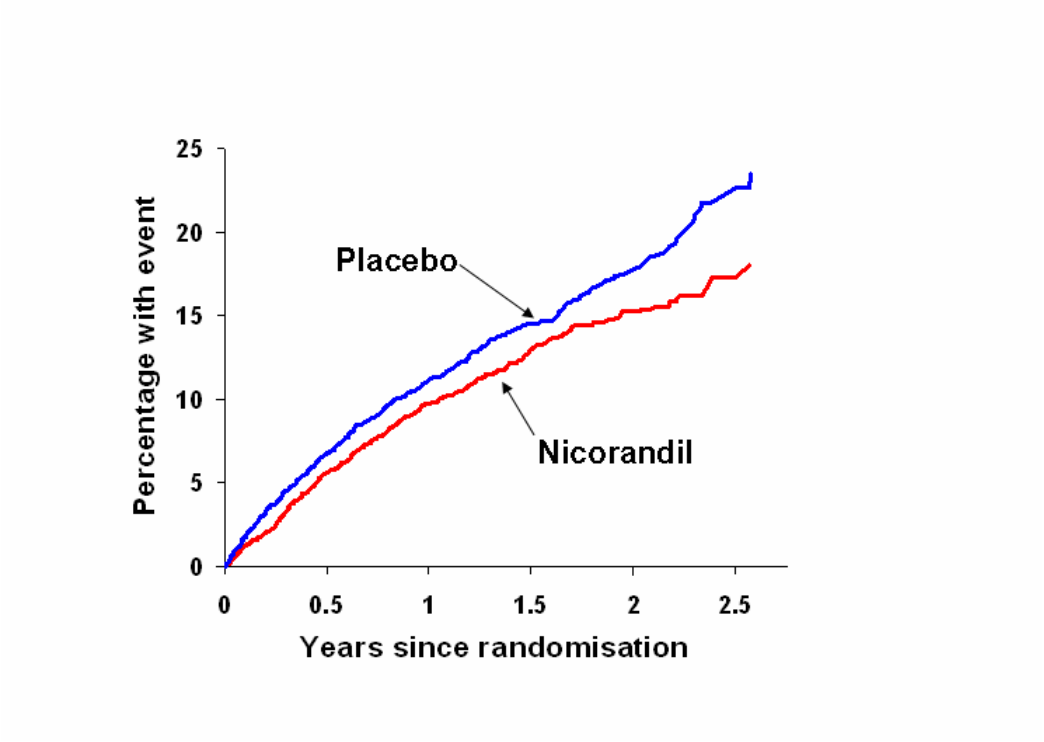


Figure 2.1: Kaplan-Meier estimates for the primary endpoint of the IONA Study

an apparent separation in the curves for the nicorandil and placebo groups with the percentage of patients in the nicorandil group who have suffered a primary

endpoint being lower than in the placebo group. If nicorandil does have cardio-protective properties then this is the pattern that would have been expected.

To see if the reduction in the number of primary endpoints suffered by patients treated with nicorandil was statistically significant a Cox proportional hazard model was fitted to the data, the results can be seen in Table 2.5 along with the numbers of patients who suffered each endpoint. (The IONA Study Group,

Table 2.5: Results of the IONA Study. The table contains numbers of events (%), estimated hazard ratios, 95% CIs and p-values

Endpoint	Nicorandil (n = 2,565)	Placebo (n = 2,561)	Hazard Ratio	95% Confidence Interval	P-Value
Primary	337 (13.1%)	398 (15.5%)	0.83	(0.72 – 0.97)	0.014
Secondary	107 (4.2%)	134 (5.2%)	0.79	(0.61 – 1.02)	0.068

2002) The HR for treatment with nicorandil relative to placebo for the primary endpoint, the composite endpoint of CHD death, non-fatal MI or unplanned hospital admission for cardiac chest pain, was 0.83 (95% CI: 0.72 - 0.97) and the p-value = 0.014. This implies that nicorandil reduced the risk of a primary endpoint, by 17% with this reduction likely to be in the range of 3% to 28%.

2.4.2 Secondary Endpoint

The Kaplan-Meier estimates split by treatment group for the secondary endpoint can be seen in Figure 2.2. As was seen for the primary endpoint as the time since randomisation increases there is an apparent separation in the curves for the nicorandil and placebo groups. The percentage of patients in the nicorandil group who suffered secondary endpoints was lower than the percentage in the placebo group. Subjectively again the Kaplan-Meier estimates indicate that nicorandil

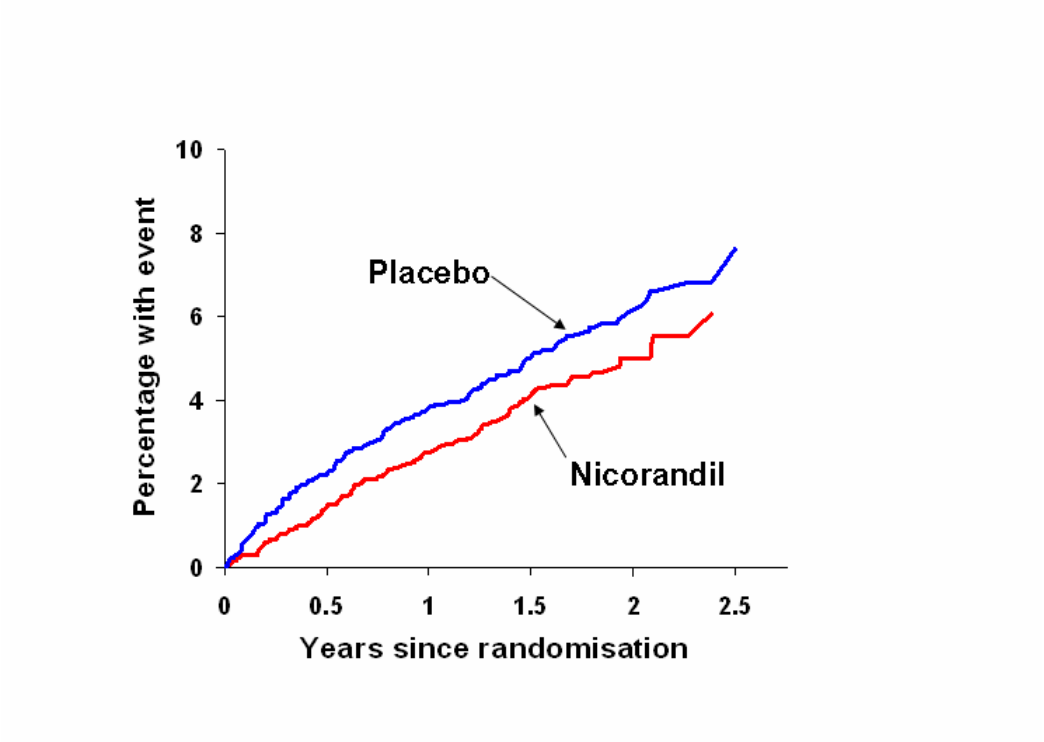


Figure 2.2: Kaplan-Meier estimates for the secondary endpoint of the IONA Study

does have cardioprotective effects.

Looking at the results for the secondary endpoint, of CHD death or non-fatal MI, in Table 2.5, the HR was 0.79 (95% CI: 0.61 - 1.02) and the p-value = 0.068, just failing to achieve statistical significance. The point estimate for the reduction in risk of suffering a secondary endpoint was 21% which is greater than the point estimate of 17% for the reduction in risk of suffering a primary endpoint. Due to the event rate for the secondary endpoint being lower than expected the results fail to reach statistical significance but they are consistent with the results for the primary endpoint.

2.4.3 Gastrointestinal Events

As was mentioned in Section 2.3, the incidence of GI events was not a predefined outcome of the IONA Study, although an increase in the number of GI events patients in the nicorandil group suffered was observed compared to patients in

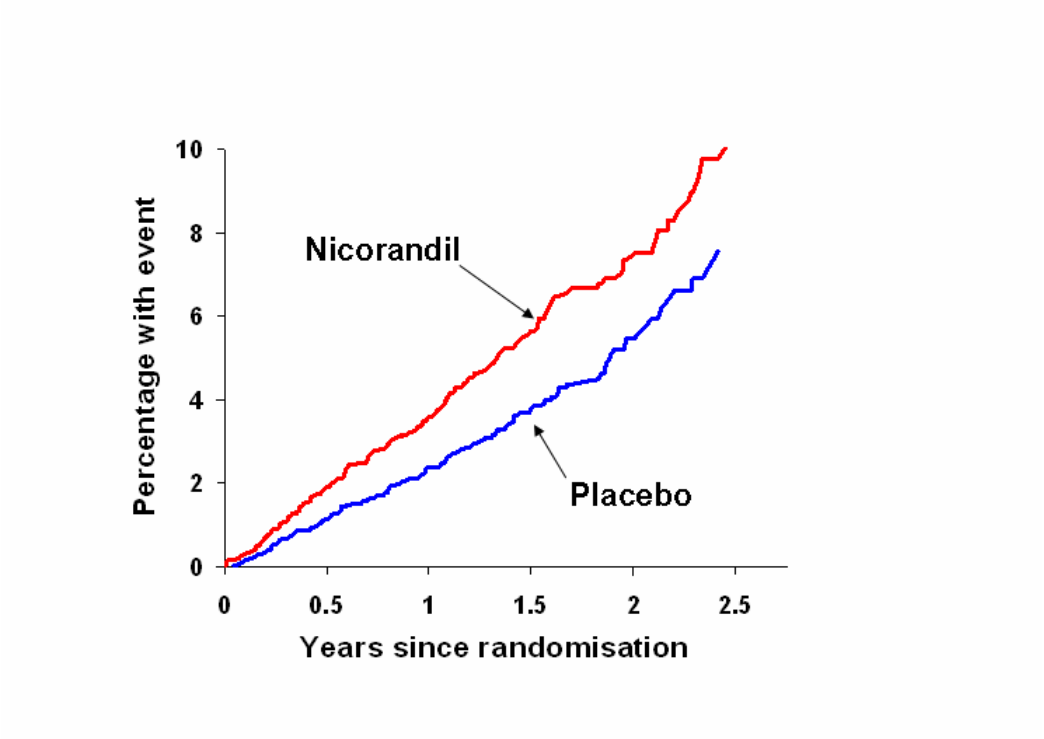


Figure 2.3: Kaplan-Meier estimates for the GI events suffered by patients during the IONA Study

the placebo group. Whether this increase in GI events was statistically significant will be explored. The Kaplan-Meier estimates split by treatment group for the GI events that patients suffered can be seen in Figure 2.3.

As with both the primary and secondary endpoints as the time since patients

were randomised increases the curves for the nicorandil and placebo groups separate but on this occasion it was the nicorandil group that had the higher percentage of patients who suffered events. The Kaplan-Meier estimates suggest that nicorandil does increase the risk of patients suffering GI events. The number of patients in the nicorandil group of the study who suffered a first GI event was 157 (6.1% of the group) and in the placebo group it was 108 (4.2% of the group). The HR for suffering a GI event, nicorandil relative to placebo, was 1.46 (95% CI: 1.14 - 1.86) and the p-value = 0.0027. Patients treated with nicorandil had an increased risk of suffering a GI event of 46%, with this increase likely to be in the range of 14% to 86%.

Chapter 3

Further Modelling of the IONA Study

In Section 2.4 the results of the time-to-first event analysis for the primary endpoint of the IONA Study were reported as well as the findings for the GI events that patients suffered. In both of these analyses the only baseline variable fitted was the randomised treatment patients were receiving. Other variables were also recorded at baseline on all patients who took part in the IONA Study, see Tables 2.1 and 2.2. In this Chapter, in addition to randomised treatment, the other baseline variables will be used to build prognostic models for the primary endpoint of the study and for the GI events patients suffered. These will be based on Cox proportional hazards models. (Cox, 1972) These models will be then used in Chapter 8, when the economic impact of the IONA Study is being assessed.

3.1 The Model Fitting Process

When it came to the model fitting there were two alterations to the baseline variables that are shown in Table 2.1. Firstly due to the close clinical relationship between suffering a stroke and a TIA and the relatively small numbers of patients who had been admitted to hospital due to a TIA, 47 (1.8%) in the nicorandil group and 55 (2.1%) in the placebo group, a stroke and hospitalisation due to a TIA were combined. The numbers of patients with the composite outcome were 172 (6.7%) in the nicorandil group and 162 (6.3%) in the placebo group. Secondly due to the small number of patients who had their angina classified as level IV by the CCSF classification, 15 (0.6%) patients in the nicorandil group and 9 (0.4%) patients in the placebo group, levels III and IV of the CCSF classification were combined together to form a new level III/IV. The new combined level contained 287 (11.2%) patients from the nicorandil group and 284 (11.1%) patients from the placebo group. All other variables were unaltered.

In order to select the variables that were to be included in the final models as explanatory variables a stepwise variable selection procedure was implemented. The significance level to enter the model was set at 0.05 and the significance level to leave the model was also set at 0.05. The rationale behind using a variable selection procedure to build the models was the large number of potential explanatory variables that were available for inclusion. In addition to using a stepwise variable selection procedure both forward and backward variable selection procedures were also used throughout to see if any differences in the variables selected in the final models by the different selection methods were apparent and if any of these differences would have implications for the interpretation of the

final models. In all cases the models shown in Sections 3.4 and 3.6 were those generated by the stepwise variable selection procedure. As the change in risk caused by an increase or decrease of one unit for some of the continuous variables has little clinical significance, especially if the range of the variable is large, the following increases in the continuous variables were considered: an increase of 10 years in age, an increase of 10 cm in height, an increase of 5 kg in weight, an increase of 2 kg/m^2 in BMI, an increase of 10 mm Hg in SBP, an increase of 5 mm Hg in DBP and an increase of 10 bpm in heart rate. All other variables were categorical with two categories, apart from the CCSF classification, which had three. In all of the analyses CCSF level I was used as the reference level to which levels II an III/IV were compared. For the categorical variable sex, females were coded as zero and males were coded as one.

3.2 Missing Baseline Values

When the variable selection procedures were implemented fewer than 5,126 patients were included in the analysis data set due missing values for one or more of the baseline variables. There were no missing values for the cardiovascular medications. Details of the missingness of the data are given in Table 3.1. It can be seen that a maximum of 50 (1.9%) patients were removed from the nicorandil group and 30 (1.2%) from the placebo group leaving 2,515 patients in the nicorandil group and 2,531 patients in the placebo group due to missingness of baseline data and a new total of 5,046 patients. As all the data is baseline data it can be assumed that it is missing at random. Values could have been imputed

Table 3.1: The numbers (%) of patients with missing baseline characteristics

Variable	Nicorandil (n = 2,565)	Placebo (n= 2,561)
Height	20 (0.8%)	15 (0.6%)
Weight	14 (0.5%)	11 (0.4%)
BMI	29 (1.1%)	20 (0.8%)
SBP	2 (0.08%)	0 (0.0%)
DBP	2 (0.08%)	0 (0.0%)
Heart Rate	7 (0.3%)	2 (0.08%)
Current Smoker	0 (0.0%)	1 (0.04%)
Previous PTCA	0 (0.0%)	1 (0.04%)
Previous Stroke or TIA	0 (0.0%)	1 (0.04%)
History of LVH	10 (0.4%)	2 (0.08%)
History of LVD	3 (0.1%)	1 (0.04%)
CCSF Classification	2 (0.08%)	2 (0.08%)
Patients With Missing Values	50 (1.9%)	30 (1.2%)

for the missing data but as the level of missingness was so low this was not attempted. Most of the missingness was caused by patients not having their height and or weight measured at baseline. To minimise any problems associated with missing data models identified by the variable selection procedure were refitted using patients who had data for all of the selected variables.

3.3 Model Building

Two different methods for producing the multivariable predictive models for both the primary endpoint of the study and the GI events patients suffered were used. This was again due to the large number of potential explanatory variables available for inclusion. By using two different model building approaches, in addition to using a stepwise variable selection procedure to actually select the variables for inclusion in the models, the final models produced should model the data as

accurately as possible given the available data. The first method used the set of baseline variables as a whole. It would be possible to use a stepwise variable selection procedure on all the available baseline variables. For each type of event this procedure could be carried out twice once with the CCBs included as a whole and secondly with the CCBs separated in to the three most commonly prescribed ones. Instead the univariate significance of the variables was firstly investigated for both types of event and then a stepwise variable selection procedure performed on the sub-group of variables found to be univariately significant for each type of event.

The second method used involved splitting the baseline variables into three categories: traditional risk factors for CHD, clinical indicators for CHD and background cardiovascular medication. Prognostic models were produced for each of the categories of baseline variables, using a stepwise variable selection procedure, and from the variables selected in these models a further model was then produced, again using a stepwise variable selection procedure, and this was carried out for both the primary endpoint and the GI events. The randomised treatment that patients were receiving was included in each of the three categories. As some of the traditional risk factors for CHD, such as the age and sex of patients, are also clinical indicators for suffering from CHD they were included in both categories. There were two variations for the baseline cardiovascular medication category, firstly the CCBs were included as a whole and secondly amlodipine, diltiazem and nifedipine were included separately. The variables that were included in the three categories are shown in Table 3.2.

Table 3.2: The baseline variables included in the different categories

Traditional Risk Factors for CHD	Clinical Indicators for CHD	Background Cardiovasular Medications
Randomised Treatment	Randomised Treatment	Randomised Treatment
Age	Age	ACE Inhibitors
Height	Sex	Antiplatelets (and Aspirin)
Weight	Hypertensive	β-Blockers
BMI	Previous MI	Long Acting Nitrates
SBP	Previous CABG	Ordinary Diuretic
DBP	Previous PTCA	Loop Diuretic
Heart Rate	Previous Angiogram	Statins
Sex	Previous Stroke or TIA	Calcium Channel Blockers*
Diabetic	History of LVH	Amlodipine†
Current Smoker	History of LVD	Diltiazem†
	History of PVD	Nifedipine†
	CCSF Classification	

* Included in the background cardiovascular medication category used in Model A
† Included in the background cardiovascular medication category used in Model B

3.4 The Primary Endpoint Model

The two methods of model building were implemented to produce models for the primary endpoint of the IONA Study.

3.4.1 Method One

The univariate significance of the baseline variables for the primary endpoint was investigated using Cox proportional hazards models and the results can be seen in Table 3.3. The p-values for those variables found to be univariately significant for either the primary endpoint or GI events are highlighted. The following variables were found to be individually significant prognostic factors for whether patients would suffer a primary endpoint: treatment with placebo, older age, lower weight,

Table 3.3: Univariate predictors for the IONA primary endpoint and for GI events. The table contains estimated hazard ratios, 95% CIs and p-values

	Primary Endpoint			GI Events		
Variable	Hazard Ratio	95% Confidence Interval	P-Value	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	0.84	0.72 – 0.97	**0.014**	1.46	1.14 – 1.86	**0.0027**
Age (increase of 10 years)	1.12	1.02 – 1.22	**0.013**	1.24	1.07 – 1.44	**0.0038**
Height (increase of 10 cm)	0.96	0.88 – 1.04	0.27	0.89	0.78 – 1.02	0.080
Weight (increase of 5 kg)	0.97	0.94 – 0.99	**0.011**	0.94	0.90 – 0.98	**0.0056**
BMI (increase of 2 kg/m^2)	0.96	0.93 – 0.99	**0.019**	0.94	0.89 – 0.99	**0.045**
SBP (increase of 10 mm Hg)	0.99	0.95 – 1.03	0.55	0.95	0.89 – 1.01	0.086
DBP (increase of 5 mm Hg)	0.97	0.94 – 1.01	0.12	0.97	0.91 – 1.03	0.27
Heart Rate (increase of 10 bpm)	1.09	1.02 – 1.16	**0.0074**	1.08	0.98 – 1.20	0.14
Male	1.01	0.85 – 1.20	0.91	0.71	0.55 – 0.92	**0.011**
Diabetic	1.15	0.90 – 1.48	0.27	1.39	0.94 – 2.05	0.10
Hypertensive	0.99	0.86 – 1.15	0.90	1.07	0.84 – 1.36	0.61
Current Smoker	1.36	1.14 – 1.62	**0.0007**	1.01	0.73 – 1.39	0.97
Previous MI	1.62	1.37 – 1.62	**<0.0001**	1.33	1.02 – 1.74	**0.035**
Previous CABG	1.09	0.92 – 1.29	0.32	1.14	0.86 – 1.50	0.36
Previous PTCA	0.97	0.79 – 1.19	0.74	1.13	0.82 – 1.56	0.45
Previous Angiogram	0.92	0.80 – 1.07	0.29	0.95	0.75 – 1.23	0.70
Previous Stroke or TIA	1.58	1.24 – 2.03	**0.0003**	0.96	0.58 – 1.59	0.87
History of LVH	1.53	1.24 – 1.89	**<0.0001**	1.25	0.86 – 1.82	0.24
History of LVD	1.34	1.06 – 1.70	**0.014**	1.59	1.11 – 2.28	**0.011**
History of PVD	1.34	1.10 – 1.63	**0.0043**	1.16	0.83 – 1.64	0.39
CCSF level II vs. level I	1.27	1.04 – 1.51	**0.016**	1.54	1.12 – 2.13	**0.0084**
CCSF level III/IV vs. level I	2.45	1.95 – 3.08	**<0.0001**	2.12	1.40 – 3.19	**0.0004**
ACE Inhibitors	1.34	1.15 – 1.55	**0.0002**	1.38	1.07 – 1.78	**0.012**
Antiplatelets (and Aspirin)	0.90	0.73 – 1.12	0.36	0.92	0.64 – 1.33	0.66
β-Blockers	0.83	0.72 – 0.96	**0.0099**	0.75	0.59 – 0.95	**0.016**
Long Acting Nitrates	1.68	1.55 – 1.95	**<0.0001**	1.70	1.32 – 2.19	**<0.0001**
Ordinary Diuretics	0.79	0.61 – 1.02	**0.073**	0.92	0.61 – 1.38	0.68
Loop Diuretics	1.82	1.56 – 2.13	**<0.0001**	1.69	1.30 – 2.20	**<0.0001**
Statins	0.84	0.72 – 0.97	**0.016**	0.65	0.51 – 0.83	**0.0005**
Calcium Channel Blockers	1.18	1.02 – 1.36	**0.031**	1.14	0.89 – 1.46	0.30
Amlodipine	1.15	0.97 – 1.38	0.11	0.93	0.68 – 1.27	0.65
Diltiazem	1.16	0.98 – 1.36	0.081	1.37	1.06 – 1.77	**0.018**
Nifedipine	1.00	0.77 – 1.31	1.00	0.75	0.45 – 1.24	0.26

lower BMI, higher heart rate, current smoking, previous MI, previous stroke or TIA, history of LVH, history of LVD, history of PVD, higher angina status as classified by the CCSF classification, treatment with an ACE inhibitor, non-treatment with a β-blocker, treatment with a long acting nitrate, treatment with a loop diuretic, non-treatment with a statin and treatment with a CCB. This sub-group of variables was then used to build a model for the primary endpoint.

From the variables that were univariately significant the following nine variables were selected in the model by the variable selection procedure: treatment with nicorandil, lower BMI, current smoking, previous MI, previous stroke or TIA, history of LVH, increased angina status as assessed by the CCSF classification, treatment with a long acting nitrate and treatment with a loop diuretic. This model was then refitted to maximise the number of patients without missing data and the results can be seen in Table 3.4. The results were qualitatively

Table 3.4: Multivariable predictive model for the primary endpoint of the IONA Study found from the univariately significant baseline variables. For this model n = 5,059. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	0.83	0.71 – 0.96	0.010
BMI (increase of 2 kg/m^2)	0.96	0.92 – 0.99	0.0065
Current Smoker	1.27	1.06 – 1.52	0.0095
Previous MI	1.43	1.22 – 1.71	<0.0001
Previous Stroke or TIA	1.38	1.07 – 1.77	0.012
History of LVH	1.42	1.14 – 1.75	0.0014
CCSF level II vs. level I	1.20	0.99 – 1.45	0.060
CCSF level III/IV vs. level I	2.00	1.58 – 2.53	<0.0001
Long Acting Nitrates	1.41	1.20 – 1.64	<0.0001
Loop Diuretics	1.53	1.30 – 1.80	<0.0001

unchanged.

3.4.2 Method Two

Separate sub-models for the three categories of baseline variables were produced and from these models a predictive model for the primary endpoint was produced.

3.4.2.1 Traditional Risk Factors for CHD Category

From the traditional risk factors for CHD category the variables that were selected by the variable selection procedure were: treatment with nicorandil, older age, higher heart and current smoking. This sub-model was then refitted to maximise the number of patients without missing data and the results can be seen in Table 3.5.

Table 3.5: Traditional risk factors for CHD sub-model for the primary endpoint of the IONA Study. For this model n = 5,114. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	0.84	0.73 – 0.97	0.018
Age (increase of 10 years)	1.16	1.06 – 1.27	0.013
Heart Rate (increase of 10 bpm)	1.07	1.01 – 1.41	0.027
Current Smoker	1.43	1.19 – 1.72	0.0001

3.4.2.2 Clinical Indicators for CHD Category

The variables that were selected by the variable selection procedure in the clinical indicators for CHD category sub-model were: treatment with nicorandil, previous MI, previous stroke or TIA, history of LVH, history of LVD and increased angina

status as assessed by the CCSF classification. The results of fitting this sub-model outwith the variable selection procedure to maximise the number of patients without missing data can be seen in Table 3.6.

Table 3.6: Clinical indicator variables for CHD sub-model for the primary endpoint of the IONA Study. For this model n = 5,109. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	0.84	0.72 – 0.97	0.016
Previous MI	1.58	1.34 – 1.87	<0.0001
Previous Stroke or TIA	1.42	1.10 – 1.82	0.0062
History of LVH	1.46	1.18 – 1.80	0.0005
History of PVD	1.25	1.02 – 1.53	0.031
CCSF level II vs. level I	1.27	1.05 – 1.53	0.012
CCSF level III/IV vs. level I	2.33	1.85 – 2.92	<0.0001

3.4.2.3 Baseline Cardiovascular Medication Category

For the primary endpoint the same variables were selected by the stepwise variables selection procedure from both variations of the baseline cardiovascular medication category, those selected were: treatment with nicorandil, treatment with a long acting nitrate and treatment with a loop diuretic. The models produced by the variable selection procedure and outwith the procedure were the same as there were no missing values for the baseline cardiovascular medication that patients were receiving. The results for the sub-model can be seen in Table 3.7.

3.4.2.4 Multivariable Predictive Model

From the three categories of baseline variables the following variables were selected as being prognostic in at least one of the models: treatment with nicorandil,

Table 3.7: Baseline cardiovascular medications sub-model for the primary endpoint of the IONA Study. For this model n = 5,126. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	0.83	0.72 – 0.96	0.013
Long Acting Nitrates	1.56	1.34 – 1.82	<0.0001
Loop Diuretics	1.69	1.44 – 1.98	<0.0001

older age, higher heart rate, current smoking, previous MI, previous stroke or TIA, history of LVH, history of PVD, higher angina status of patients as assessed by the CCSF classification, treatment with a long acting nitrate and treatment with a loop diuretic. Using a stepwise variable selection procedure a model was built from these variables. This generated a prognostic model containing eight of the variables with age, heart rate and a history of PVD the variables omitted. This model was then refitted outwith the variable selection procedure, to maximise the number of patients without missing data, and the results can be seen in Table 3.8.

3.5 Interpretation of the Primary Endpoint Models

The multivariable predictive models for the primary endpoint produced by the different methods used were the same apart from the inclusion of the BMI of patients in the first model and exclusion of it from the second model. The differences in HRs, 95% CIs and p-values between the two versions of the model were minor and caused by the inclusion and exclusion of the variable BMI. From

Table 3.8: Multivariable predictive model for the primary endpoint of the IONA Study based on the sub-models produced from the baseline variable categories. For this model n = 5,108. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	0.83	0.72 – 0.96	0.012
Current Smoker	1.29	1.08 – 1.54	0.0058
Previous MI	1.46	1.23 – 1.73	<0.0001
Previous Stroke or TIA	1.38	1.07 – 1.76	0.012
History of LVH	1.46	1.18 – 1.80	0.0004
CCSF level II vs. level I	1.18	0.98 – 1.42	0.091
CCSF level III/IV vs. level I	1.94	1.54 – 2.45	<0.0001
Long Acting Nitrates	1.42	1.21 – 1.67	<0.0001
Loop Diuretics	1.50	1.28 – 1.76	<0.0001

the prognostic models for the primary endpoint it was reinforced that the risk of patients suffering CHD events was reduced for those who were treated with nicorandil. It can be seen from version one of the model that patients who were treated with nicorandil had a reduced risk of 17% (95% CI: 4%, 29%) of suffering a primary endpoint. The results were similar for the second version of the model as well as to those seen previously in Section 2.4.1. The HR point estimate for the reduction in risk of suffering a primary endpoint associated with treatment with nicorandil was the same in both versions of the multivariable predictive model as it had been for the original time-to event analysis shown in Section 2.4.1. This suggests that balance was achieved between the two treatment arms of the study.

The age of patients is a known risk factor for CHD, with the risk of patients suffering CHD events increasing as their age increases but the variable age was not included in the either of the final models for the primary endpoint. From Table 3.3 it can be seen that age was a significant univariate predictor. However

the BMI of patients at baseline was included in the first version of the model, with the risk of a primary endpoint decreased with a higher BMI. In contrast to this it has been shown elsewhere that the risk of CHD events is actually increased with higher BMI. (Kim et al., 2006) Through further investigation it became apparent that there was a relationship between age and BMI. The general trend was that with increasing age the BMI of patients decreased. There was a statistically significant negative Pearson correlation of -0.20 between age and BMI, p-value <0.0001. Older patients are more likely to suffer from poorer health and weight loss, which leads to a decrease in BMI, is a common sign that the health of patients is deteriorating. This in turn could imply that these patients are more likely to suffer CHD events. Decreasing BMI was therefore acting as a marker for increasing age and deteriorating health in patients. Due to this fact and confounding factors with the other variables included in the model age was not selected in the final model. For the sub-model produced from the traditional risk factors for CHD, see Table 3.5, age was included whereas BMI was not. BMI could then not be included in the final prognostic model found via method two. Even with the removal of BMI from possible inclusion the confounding factors between age and the other variables that were included in the model meant that age was not included in this model either.

The other baseline variables included in the final model were all known indicators of patients who will have an increased risk of suffering CHD events. These variables included smoking history, previous MI, previous stroke or TIA, history of LVH and more severe symptoms of angina. As the CCSF classification measures the severity of angina it is likely that those patients who were classified higher on the scale would be at a higher risk of suffering from CHD events. With

level I, the mildest level of angina, used as the reference level for the other two levels to be compared against it can be see from Tables 3.4 and 3.8 that there was a non-statistically significant trend to increased risk between levels I and II. The comparison between levels I and III/IV, the mildest and the most severe levels, is highly significant in both models, with the p-value being <0.0001. In this case the risk between those patients classified as level III/IV compared to level I was increased by 100% (95% CI: 58%, 153%), according to the first version of the model. The results for the second version were comparable. In additional analysis that was carried out levels II and III/IV were compared. The HR of patients classified as level III/IV compared to level II of the CCSF classification was 1.67 (95% CI: 1.38 - 2.02) and the p-value was <0.0001. This was for version one of the model with similar results seen for version two. The sub-groups of patients classified as level III/IV had a significantly increased risk of suffering a primary endpoint compared to those classified as level II.

The results of the models for the primary endpoint show that those patients who were being treated with a long acting nitrate or a loop diuretic at baseline had an increased risk of suffering a CHD event. As both of these are classes of cardiovascular medications these findings may appear to be strange at first glance. This is likely to be due to confounding by indication. Patients with angina are prescribed long acting nitrates in order to help with the pain caused by angina but nitrates are not the front line treatment. In normal circumstances patients are firstly prescribed β-blockers, then calcium antagonists and probably nitrates third, thus identifying that those patients who end up being treated with nitrates as a more difficult sub-group of patients and hence possibly at a higher risk of suffering CHD events. Loop diuretics are given to patients who suffer from

heart failure to relieve one of the potential symptoms: fluid accumulation in the body. Patients with congestive heart failure (CHF) are at particularly high risk. Hence these cardiovascular medications were markers of patients at high risk of clinical endpoints. Even though the medications should help the patients who were being treated with them with symptomatic relief they were at an increased risk of suffering a CHD event because of the underlying history that led to the medications being prescribed to the them in the first place.

Looking at the three sub-models produced from the different categories of baseline variables of the variables selected in at least one of the sub-models three were not selected in the final version of the model for the primary endpoint produced by method two: age, which has previously been discussed, heart rate and history of PVD. There is supporting evidence of an association between a higher resting heart rate and the likelihood of suffering a CHD event (Shaper et al., 1993; Hjalmarson, 2007) and having a history of PVD is a known risk indicator for being at an increased risk of suffering a CHD event. In the presence of the other variables selected in the final model these variables did not provide any additional significant prognostic information.

As the Cox model was used to model the data the underlying assumption of proportional hazards should be checked. All of the variables included in the multivariable predictive models for the primary endpoint produced by the two model building methods met this assumption with the exception of whether patients had suffered a previous stroke or TIA. As both models contained a large number of variables it was not unexpected that one of the variables would violate the assumption of proportional hazards. As a result the models were refitted with the variable whether patients had suffered a previous stroke or TIA omitted. To

ascertain what effect the removal of the variable whether patients had suffered a previous stroke or TIA had on the models fitting the data Akaike's Information Criterion (AIC) (Akaike, 1974) was compared for the two versions of each model. For both models the values of AIC were lower for the model containing the variable whether patients had suffered a previous stroke or TIA. As a result the variable whether patients had suffered a previous stroke or TIA was not removed from either model.

Having produced two multivariable predictive models for the primary endpoint a decision had to be made on which to use in later analyses. The value of AIC was compared for the two versions of the model for the primary endpoint and the model produced by method one had a lower value indicating it to be the better fitting model. However, to gain an understanding of the predictive performance of the models their C-statistic values were calculated. (Harrell et al., 1982) The C-statistic can be though of as being equivalent to calculating the area under a receiver operating characteristic (ROC) curve (Hanley and McNeil, 1982) in that the C-statistic can have values between 0 and 1, with a value close to 1 indicating that the predictive performance of the model is good. The value of the C-statistic for the model produced by method one was 0.6421 (95% CI: 0.6211 - 0.6630) and for the model produced by method two it was 0.6409 (95% CI: 0.6201 - 0.6617). The point estimates for the C-statistic are very close together and the CIs overlap considerably but as the value of the C-statistic was marginally larger for the model produced by method one this is the model that will be used in later analyses.

3.6 The Gastrointestinal Event Model

The two methods of model building were implemented to produce models for the first GI events patients suffered during the IONA Study. For both of the methods for producing predictive models for the GI events two models were explored. This was due to whether the CCBs were included as a whole or whether amlodipine, diltiazem and nifedipine were included separately. Model A incorporates the CCBs as a whole and Model B includes amlodipine, diltiazem and nifedipine separately

3.6.1 Method One

The univariate significance of the baseline variables for the GI events was investigated using Cox proportional hazards models and the results can be seen in Table 3.3. For the GI events the following variables were found to be individually significant predictors: treatment with nicorandil, older age, lower weight, lower BMI, being female, previous MI, history of LVD, higher angina status as classified by the CCSF classification, treatment with an ACE inhibitor, non-treatment with a β-blocker, treatment with a long acting nitrate, treatment with a loop diuretic, non-treatment with a statin and, although the combined variable of whether patients were being prescribed a CCB was non significant, treatment with diltiazem. This sub-group of variables was then used to build models for the GI events.

For Model A1 the following six variables were included in the model by the variable selection procedure: treatment with nicorandil, lower weight, higher angina status as assessed by the CCSF classification, treatment with a long acting

Table 3.9: Multivariable predictive model, Model A1, for the GI events patients suffered during the IONA Study found from the univariately significant baseline variables. For this model n = 5,097. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	1.44	1.13 – 1.84	0.0038
Weight (increase of 5 kg)	0.95	0.91 – 0.99	0.014
CCSF level II vs. level I	1.39	1.01 – 1.92	0.046
CCSF level III/IV vs. level I	1.74	1.14 – 2.64	0.010
Long Acting Nitrates	1.50	1.15 – 1.95	0.0024
Loop Diuretics	1.43	1.09 – 1.88	0.0093
Statins	0.71	0.55 – 0.90	0.0053

nitrate, treatment with a loop diuretic and non-treatment with a statin. This model was refitted to maximise the numbers of patients without missing values and the results can be seen in Table 3.9.

For Model B1 the following seven variables were included in the model by the variable selection procedure: treatment with nicorandil, lower weight, higher

Table 3.10: Multivariable predictive model, Model B1, for the GI events patients suffered during the IONA Study found from the univariately significant baseline variables. For this model n = 5,097. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	1.44	1.13 – 1.84	0.0036
Weight (increase of 5 kg)	0.95	0.91 – 0.99	0.012
CCSF level II vs. level I	1.37	0.99 – 1.89	0.060
CCSF level III/IV vs. level I	1.70	1.11 – 2.58	0.014
Long Acting Nitrates	1.48	1.14 – 1.93	0.0032
Loop Diuretics	1.44	1.10 – 1.89	0.0088
Statins	0.71	0.55 – 0.90	0.0018
Diltiazem	1.31	1.01 – 1.70	0.046

angina status as assessed by the CCSF classification, treatment with a long acting nitrate, treatment with a loop diuretic, non-treatment with a statin and treatment with a diltiazem. The result of re-fitting the model, Model B1, can be seen in Table 3.10. It can be seen that the common explanatory variables between Models A1 and B1 had the same implications for the risk of patients suffering GI events. The additional variable included in Model B1 was treatment with diltiazem.

3.6.2 Method Two

Separate sub-models for the three categories of baseline variables were produced and from these models predictive models for the GI events were produced.

3.6.2.1 Traditional Risk Factors for CHD Category

From the traditional risk factors for CHD the following variables were selected in the sub-model for the GI events: treatment with nicorandil, older age, decreasing SBP and being female. This model was then fitted outwith the variable selection procedure to minimise the numbers of patients with missing values and the results can be seen in Table 3.11.

Table 3.11: Traditional risk factors for CHD sub-model for the GI events patients suffered during the IONA Study. For this model n = 5,124. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	1.46	1.14 – 1.87	0.0025
Age (increase of 10 years)	1.24	1.07 – 1.45	0.0049
SBP (increase of 10 mm Hg)	0.92	0.87 – 0.98	0.013
Male	0.74	0.56 – 0.96	0.025

3.6.2.2 Clinical Indicators for CHD Category

The following variables were selected in the sub-model for the GI events from the category of clinical indicators for CHD: treatment with nicorandil, older age, being female, a history of LVD and higher angina status as assessed by the CCSF classification. When the sub-model was fitted outwith the variable selection procedure, so that the number of patients without missing data was maximised, the sex of patients was no longer a significant predictor. This was due to the differing sizes of data set used to produce the models. For the variable selection procedure the data set contained 5,104 patients and sex was significant, but it was on the borderline of significance as the p-value = 0.049, highlighted in Table 3.12. Outwith the selection procedure the data set contained 5,118 patients, a difference of only 14 patients, and sex was no longer significant. Shown in the left half of Table 3.12 are the results produced by the variable selection procedure and shown in the right half are the results when the model was refitted with sex removed from the model.

Table 3.12: Clinical indicator variables for CHD sub-model for the GI events patients suffered during the IONA Study. For both models n = 5,118. The table contains estimated hazard ratios, 95% CIs and p-values

	Model as Selected by the Variable Selection Procedure			Model Fitted Outwith the Variable Selection Procedure with Sex Removed		
Variable	Hazard Ratio	95% Confidence Interval	P-Value	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	1.45	1.14 – 1.86	0.0029	1.44	1.13 – 1.85	0.0033
Age (increase of 10 years)	1.20	1.03 – 1.40	0.017	1.24	1.07 – 1.44	0.0048
Male	0.76	0.58 – 0.99	**0.049**	–	–	–
History of LVD	1.57	1.09 – 2.26	0.015	1.51	1.05 – 2.17	0.025
CCSF level II vs. level I	1.51	1.10 – 2.08	0.013	1.53	1.11 – 2.11	0.0096
CCSF level III/IV vs. level I	1.90	1.25 – 2.88	0.0026	1.99	1.32 – 3.01	0.0011

3.6.2.3 Baseline Cardiovascular Medication Category

Firstly, from the category of cardiovascular medications that will be used to produce Model A2, the following variables were selected in the sub-model for GI events: treatment with nicorandil, treatment with a long acting nitrate, treatment with a loop diuretic and non-treatment with a statin. Secondly, the following variables were included in the sub-model that will be used to produce Model B2: treatment with nicorandil, treatment with a long acting nitrate, treatment with a loop diuretic, non-treatment with a statin and additionally treatment with the CCB diltiazem. The results for both models are shown in Table 3.13.

Table 3.13: Baseline cardiovascular medications sub-models for the GI events patients suffered during the IONA Study. For both models n = 5,126. The table contains estimated hazard ratios, 95% CIs and p-values

	Model A2			Model B2		
Variable	Hazard Ratio	95% Confidence Interval	P-Value	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	1.45	1.14 – 1.85	0.0029	1.46	1.14 – 1.86	0.0026
Long Acting Nitrates	1.58	1.22 – 2.05	0.0005	1.56	1.21 – 2.02	0.0007
Loop Diuretics	1.51	1.16 – 1.97	0.0025	1.50	1.16 – 1.97	0.0026
Statins	0.68	0.53 – 0.87	0.0019	0.68	0.53 – 0.87	0.0018
Diltiazem	–	–	–	1.33	1.03 – 1.73	0.031

3.6.2.4 Multivariable Predictive Model

From the categories of baseline variables the following variables were selected as being prognostic in at least one of the sub-models for the GI events: treatment with nicorandil, older age, decreasing SBP, being female, history of LVD, higher angina status of patients as assessed by the CCSF classification, treatment with a long acting nitrate, treatment with a loop diuretic and non-treatment with a statin. In addition treatment with diltiazem was also included in the sub-model for the category of cardiovascular medications that will be used to produce Model B2. Using a stepwise variable selection procedure two further models were then built using these variables. Prognostic models containing six and seven variables respectively were produced. The only variable not included in either model was the SBP of patients. Models A2 and B2 were fitted outwith the variable selection procedure to maximise the number of patients without missing data and the results can be seen in Table 3.14.

Table 3.14: Multivariable predictive models for the GI events patients suffered during the IONA Study based on the sub-models produced from the baseline variable categories. For both models n = 5,122. The table contains estimated hazard ratios, 95% CIs and p-values

	Model A2			Model B2		
Variable	Hazard Ratio	95% Confidence Interval	P-Value	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	1.45	1.14 – 1.85	0.0029	1.46	1.14 – 1.86	0.0026
Male	0.75	0.58 – 0.98	0.032	0.73	0.56 – 0.95	0.018
History of LVD	–	–	–	1.47	1.01 – 2.13	0.043
CCSF level II vs. level I	1.41	1.02 – 1.94	0.040	–	–	–
CCSF level III/IV vs. level I	1.68	1.10 – 2.55	0.016	–	–	–
Long Acting Nitrates	1.50	1.15 – 1.94	0.0025	1.54	1.19 – 2.00	0.0010
Loop Diuretics	1.43	1.09 – 1.87	0.010	1.40	1.06 – 1.83	0.019
Statins	0.68	0.53 – 0.86	0.0016	0.67	0.52 – 0.85	0.0010
Diltiazem	–	–	–	1.34	1.03 – 1.74	0.027

3.7 Interpretation of the Gastrointestinal Event Models

Models A1 and B1, produced by the first method of model building, were similar with the only difference being the inclusion of treatment with diltiazem in Model B1. Models A2 and B2, produced by the second method, were also similar. The differences were that the angina status of patients as assessed by the CCSF classification was included in Model A2 whereas in Model B2 a history of LVD and treatment with diltiazem were included. Comparing the two version of Model A they were the same except the first included weight and the second included sex, the same difference was seen for the two versions of Model B. The angina status of patients was included in Model B1 but not in Model B2, where a history of LVD was included. The weight of patients was not available for inclusion in the final models produced by method two as weight was not selected in the traditional risk factors for CHD sub-model. In all of the models, even in the presence of other explanatory variables, treatment with nicorandil was shown to be associated with an increased risk of GI complications. In Model A1 the increase in risk of patients suffering a GI event who were in the nicorandil group was 44% (95% CI: 12%, 86%). The results for the other models were similar.

The results from both Models A1 and B1 show that the risk of suffering a GI event decreased by 5% (95% CI: 1%, 9%) for an increase in weight of 5 kg. The point raised earlier in Section 3.5 is noteworthy. Age in the IONA Study was negatively correlated with both BMI and weight. The Pearson correlation between age and weight was -0.31, p-value <0.0001. As people get older they are more susceptible to illness and this fact may be being reflected in the GI event

models with decreasing weight being a surrogate for increasing age in the model. In younger individuals an overweight status is associated with high risk. However in older diseased populations maintenance of weight is important as evidence of weight loss is often a sign of a deterioration in health. As weight was not available for inclusion in either Models A2 or B2 the sex of patients may be acting as a maker for weight in these models. Being male was associated with a lower risk of GI events and the mean (SD) weight of males was 83 kg (14) and for females it was 70 kg (13).

With patients who have the mildest level of angina being used as the reference level it was shown that in Models A1, B1 and A2 that those patients with severe levels of angina were at a higher risk of suffering GI events. Those patients who were classified a level II had an increased risk of 39% (95% CI: 1%, 93%) and those patients classified as level III/IV had an increased risk of 74% (95% CI: 18%, 171%), according to Model A1, compared to those classified as level I. In Model B1 the change in risk between levels I and II was non-significant but the trend was of increased risk for patients classified as level II, otherwise the result for Models B1 and A2 were similar to those for Model A1. There was no significant difference in the risk of patients suffering GI events between those classified as levels II and III/IV. Patients who had more severe levels of angina may in general be in poorer health so more susceptible to other forms of illness such as GI events. The same reasoning could be applied to why having a history of LVD, included in the clinical indicators for CHD sub-model, would lead to an increased risk of GI events. A further possible reason to why increasing severity of angina may lead to a higher risk of GI events is that patients with worse levels of angina symptoms may well have been prescribed a higher number

of cardiovascular medications and these medications could have been causing the increase in the risk of suffering GI events.

Many different types of medication have GI events as documented side-effects, including: long acting nitrates, loop diuretics and diltiazem as well as other CCBs. Those patients who were treated with a long acting nitrate, a loop diuretic or diltiazem had an increased risk of suffering GI events. As in the IONA Study it was only treatment with nicorandil that was randomised and due to confounding with other factors it cannot be claimed that these types of medications caused an increase in the risk of suffering GI events based on these findings alone. The findings suggest that there is an association between the treatment with these types of medication and increased risk of GI events. According to the GI models treatment with a statin was associated with a reduced risk of suffering GI events even though GI events are documented side-effects of statins. If statins do reduce the risk of patients treated with them from suffering GI events there is potentially some confirmatory evidence that statins do reduce the risk of patients suffering certain types of GI events. (Atar et al., 2006) However, this finding may just be a manifestation on the IONA Study itself or be related to the types of patients that were being treated with a statin. The issue of statins will be returned to in Section 3.8

It should be noted that it was likely that patients recruited in the IONA Study were being prescribed other forms of medication in addition to cardiovascular medications. It was not readily known what other types of medications patients were been prescribed. However, it is likely that some of the patients who participated in the IONA Study would have been prescribed non-steroidal anti-inflammatory drugs (NSAIDs). NSAIDs are associated with an increased

risk of suffering GI events. (Richy et al., 2004) It was unclear whether there was a difference in the percentage of patients in the two treatment groups who were being prescribed NSAIDs. Therefore, it was unknown what effect, if any, treatment with NSAIDs had on the risk of patients suffering GI events. Although, the number of patients who were being prescribed antiplatelets (and aspirin) was recorded and this class of medications has similar properties to NSAIDs, in that they increase the risk of patients suffering GI events. Whether patients were being prescribed antiplatelets (and aspirin) was not found to be univariately significant of whether patients would suffer a GI event. However, this was likely due to the fact that such a high proportion of patients who took part in the IONA Study were being prescribed antiplatelets (and aspirin), see Table 2.2, that any effect on the risk of suffering GI events was being masked. In addition if patients had previously experienced GI events as a result of being prescribed antiplatelets (and aspirin) they would have most likely stopped taking them, so the increased risk of suffering GI events they had caused would not be seen.

Looking at the sub-models based on the different categories of baseline variables the only variable selected in one of the sub-models not included in either Models A2 or B2 was the SBP of patients. Patients who have higher than normal BP and are not able to lower it by lifestyle changes are often prescribed medication to lower it. If the BP lowering medications, such as ACE inhibitors, β-blockers, loop diuretics or CCBs, were associated with an increase in the risk of GI events then it would appear that those patients with a higher SBP, who were unlikely to be taking any medication to lower it, were at a reduced risk of suffering GI events. With the cardiovascular medications available for inclusion in Models A2 and B2 SBP no longer provides any significant prognostic information

over and above that provided by the cardiovascular medications.

Again as had been the case for the primary endpoint multivariable predictive model as the Cox model was used to model the data the underlying assumption of proportional hazards should be checked. All of the variables included in the multivariable predictive models for the GI events produced by the two model building methods met this assumption. AIC was again used to compare how the different models fitted the data. Model A1 had a lower value of AIC than Model A2 and Model B1 and a lower value of AIC than Model B2. When the values of AIC were compared for Models A1 and B1 it was found that Model B1 had a lower value indicating that Model B1 was the best fit to the data.

As for the primary endpoint a decision had to be made on which of the GI event multivariable predictive models should be used in later analyses. Thus was again done by calculating the C-statistic for the four models to ascertain their predictive performance. The point estimates and accompanying 95% CI for the C-statistic for the GI event multivariable predictive models are shown in Table 3.15. The point estimates for the C-statistic were all close together and the four sets

Table 3.15: C-statistic point estimates and accompanying 95% CIs for the GI event multivariable predictive models

GI Event Model	Point Estimate	95% Confidence Interval
Model A1	0.6307	0.5971 – 0.6644
Model B1	0.6333	0.5992 – 0.6675
Model A2	0.6268	0.5941 – 0.6615
Model B2	0.6297	0.5956 – 0.6638

of CIs overlap one another but Model B1 has the highest point estimate for the C-statistic and will therefore be used in later analyses. It can also been see that the predictive performance of the GI event models were marginally poorer than

for the primary endpoint models. Although, for both the primary endpoint and the GI events the predictive performance of the multivariable models were better than for the analyses shown in Sections 2.4.1 and 2.4.3, where the only covariate fitted in the model was the randomised treatment patients were receiving.

3.8 Interactions Between the Cardiovascular Medications

Having produced multivariable predictive models for the primary endpoint and GI events it was found that all the models contained at least two of the different classes of cardiovascular medications in addition to the randomised treatment patients were receiving. It was investigated whether any interactions would be found between these different classes of cardiovascular medications and nicorandil. This was again done by using a stepwise variable selection procedure. At the start of the stepwise variable selection procedure the variables that had already been included in the models were forced into the model and the only variables that were under consideration for inclusion were the interaction terms between nicorandil and the different classes of cardiovascular medications.

3.8.1 Primary Endpoint Model Interactions

When the stepwise variable selection procedure was carried out on the models for the primary endpoint neither the interaction between treatment with nicorandil and a long acting nitrate nor the interaction between treatment with nicorandil

and a loop diuretic were added to the models. In addition there was no interaction found between treatment with a long acting nitrate and a loop diuretic. It was also investigated whether there were any interactions between treatment with nicorandil and the traditional risk factors for CHD variables or the clinical indicators for CHD variables included in the models and none were found.

3.8.2 Gastrointestinal Event Model Interactions

When the stepwise variable selection procedure was run on both versions of Models A and B in all cases the only interaction term added was the one between treatment with nicorandil and a statin. No significant interactions were found between the different classes of cardiovascular medications or individual drug, in the case of diltiazem. In addition no interactions were found between treatment with nicorandil and the traditional risk factors for CHD variables or the clinical indicators for CHD variables included in the models. Both versions of Models A and B were then refitted with the combination of whether patients were being treated with neither nicorandil nor a statin, only nicorandil, only a statin or both nicorandil and a statin included as a categorical variable in addition to the other variables in the model. Through investigation and by altering the group used as the reference level the only comparisons that were found to be statistically significant were the three comparisons involving the patients who were only treated with a statin. Therefore, the patients who were treated with placebo and a statin were used as the reference level to compare the other three groups with. The results produced for all the models were similar and as this was an exploratory analysis only one of the models will be shown here, Model B1. The results are

shown in Table 3.16.

Table 3.16: Multivariable predictive model, Model B1, for the GI events patients suffered during the IONA Study including the nicorandil statin interaction terms with those patients treated with placebo and a statin used as the reference level. For this model n = 5,097. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Weight (increase of 5 kg)	0.95	0.91 – 0.99	0.012
CCSF level II vs. level I	1.36	0.98 – 1.88	0.064
CCSF level III/IV vs. level I	1.69	1.11 – 2.58	0.014
Long Acting Nitrates	1.49	1.15 – 1.94	0.0028
Loop Diuretics	1.43	1.09 – 1.88	0.0095
Diltiazem	1.31	1.01 – 1.70	0.043
Neither vs. Placebo and Statins	1.96	1.33 – 2.89	0.0007
Nicorandil vs. Placebo and Statins	2.21	1.52 – 3.23	<0.0001
Nicorandil and Statins vs. Placebo and Statins	1.94	1.33 – 2.81	0.0005

It can be seen that compared to the other three groups the patients who were only being treated with a statin had a decreased risk of suffering GI events. The increase in risk of suffering a GI event was smallest for the group being treated with nicorandil and a statin, followed by the group treated with either and the increase was greatest for the group only being treated with nicorandil. The apparent reduction in the risk of suffering GI events linked to treatment with a statin, seen previously in Section 3.6, is reinforced by these results. It may not actually have been treatment with a statin that was causing patients to be at a reduced risk of GI events, it may have been the type of patients who were treated with a statin and the state of their health. Treatment with nicorandil was still

shown to be associated with an increased risk of GI events.

A possible explanation to why those patients who were treated with a statin were at a reduced risk of GI events is that those patients who were being treated with a statin were naturally resistant to GI events. Patients who had previously experienced GI events during treatment with a statin may have stopped treatment. Those patients who remained on a statin may have had a naturally lower risk of suffering GI events. This was reflected in the reduction in risk of GI events for patients who were treated with a statin. Patients who were treated with a statin had a lower underlying risk of GI events so although treatment with nicorandil increased the risk in conjunction with treatment with a statin those patients were still at a lower risk than patients who were not treated with a statin. Those patients treated with nicorandil but not a statin did not have a lower underlying risk of GI events and this group of patients were at the highest risk of suffering from GI events.

A second possible explanation is these finding may be due to confounding by indication. It is believed that patients with heart failure have an increased risk of GI events, this may be as a result of: the heart failure itself (Weil et al., 2000), the medication that patients are treated with (Verhamme et al., 2006) or that patients with heart failure are in a poorer state of health so are more susceptible to other forms of illness. Typically patients with heart failure are treated with ACE inhibitors and loop diuretics in addition to other types of medication. Treatment with a loop diuretic was included in all of the GI event models. If treatment with a loop diuretic was removed from the variable set then treatment with an ACE inhibitor was included in all the models for GI events. Those patients who were treated with an ACE inhibitor were at an increased risk of GI events. That

patients were treated with either of these classes of cardiovascular medication may indicate that these patients suffered from heart failure. It is unclear whether treatment with a statin is effective in patients with heart failure, although their use is becoming more prevalent. (Krum and McMurray, 2002) During the IONA Study those patients who were suffering from heart failure may not have been treated with a statin hence the apparent reduced risk of suffering GI events for those patients. It is not known which patients were suffering from heart failure so the validity of this argument cannot be assessed. The interaction analysis was exploratory in nature so no definitive conclusions are able to be drawn from these findings.

Chapter 4

Recurrent Event Models

In the majority of cases when survival analysis is being used the analysis that is carried out is based on only the first event that a patient suffers and the time to that event. However, once patients who are taking part in a clinical trial have suffered a first event it is likely that they will continue to be monitored until the end of the trial. During this time patients may experience further events. In traditional survival analysis this additional information would not be considered. In order that all the available information on patients can be taken into consideration in any analysis, models that incorporate the recurrent events have been developed and in this Chapter the reasoning behind three of the most familiar types of recurrent event models will be explained.

4.1 Cox Proportional Hazards Model

The proportional hazards model was first presented in the paper by Cox (1972) and since then it has became the model of choice for censored survival analysis.

The notation used in the proportional hazards model is as follows. For each patient a number of variables, p, will be recorded and these variables can be time dependent. For the ith patient these variables can be denoted as $\underline{z}_i(t) = (z_{1i}(t), ., z_{pi}(t))$ and the proportional hazards model is defined as:

$$\lambda(t, \underline{z}(t)) = \lambda_0(t) exp(\underline{z}(t)\underline{\beta}) \tag{4.1}$$

where $\underline{\beta}$ is a $p \times 1$ vector of unknown parameters and $\lambda_0(t)$ is an unknown baseline hazard function for the standard set of conditions when $\underline{z}(t) = \underline{0}$. (Cox, 1972)

4.2 Introduction to the Recurrent Event Models

It should be noted that as multiple events for each patient are being modelled it is likely that there will be correlation among the observations for individual patients and as a result Lin and Wei (1989) developed a robust sandwich estimator for the covariance matrix of the estimated parameters. When analysing recurrent event data it is advisable that this robust sandwich estimator for the covariance matrix be used leading to a robust standard error for the parameter estimates calculated in the modelling process. Secondly, a cut-off point is needed for the number of recurrent events to be included so that there are still sufficient numbers of patients at risk to make the model fitting effective. If there are small numbers of recurrent events the benefit of using a recurrent event model to analyse the data may be lost. A counting process model, a marginal model and two variations of a conditional model for analysing recurrent events will be described.

4.3 Anderson-Gill Model

The Anderson-Gill model, which will be referred to as the AG model, was introduced in the paper by Andersen and Gill (1982). The AG model is based on a counting process. A counting process, defined by $N(t)$ where t is time, is a process which is constant between events and only moves one unit at each event time. In the recurrent event setting a counting process can be considered as a stochastic process where the occurrence of a number of types of disjoint discrete events in time is recorded. (Andersen et al., 1993) The AG model is defined as:

$$\lambda(t, \underline{z}(t)) = Y(t)\lambda_0(t)exp(\underline{z}(t)\underline{\beta}) \tag{4.2}$$

Equation 4.2 is the same as the Cox proportional hazards model seen in Equation 4.1 apart from the inclusion on the additional $Y(t)$ term. The term $Y(t)$ is known as the risk indicator and in theory it could also be included in Equation 4.1. The reason that it can be omitted from the Cox proportional hazards model is due to the different definition of the risk indicator between the two models. For patient i, when the Cox proportional hazards model is being used, once patient i has suffered the event of interest they are no longer at risk of suffering that event so the risk indicator, $Y_i(t)$, becomes zero. Therefore, for the Cox model once patient i has suffered the event of interest they can no longer provide any more information for the modelling process. In the AG model the risk indicator for patient i, $Y_i(t)$, remains at one for as long as patient i is being monitored no matter how many or few events they suffer during that period of time. For the AG model even when patient i has suffered the event of interest for the first time they can still provide information to use in the modelling process.

In the AG model all events that patients suffer are assumed to be independent. There is no distinction made between whether an event patient i suffers is a first, second or third event as all the events are treated the same. The time scale used in the AG model is continuous with it starting at time zero, when the patient begins the period of observation, and running until time T, when they leave the study. The clock is not reset to zero after an event occurs. As indicated previously, patient i contributes to the risk set for time T_i, the total time they are under observation, no matter what the event number that is being modelled for and whether patient i has or will go onto suffer that number of events. Patient i contributes the event defining the risk set at times when they suffer events. As events are treated as being independent and the number of events suffered is not differentiated between the events patient i suffers could in effect be treated as single events experienced by n different patients, with n depending on the number of events patient i suffered, with the time scale starting for each patient at the time of the previous event.

4.4 Wei, Lin and Weissfeld Model

The Wei, Lin and Weissfeld model, which will be referred to as the WLW model, was introduced in the paper by Wei et al. (1989). The WLW model is a marginal model as the correlations among events are not modelled. The model is defined as:

$$\lambda_k(t, \underline{z}_k(t)) = \lambda_{0k}(t) exp(\underline{z}_k(t)\underline{\beta}_k) \tag{4.3}$$

Equation 4.3 is the same as Equation 4.1 with the exception on the inclusion on the subscript k, where k refers to the k^{th} event that patients suffer. Implying that each event has a different unknown underlying hazard function, $\lambda_{0k}(t)$, and set of parameter estimates, $\underline{\beta}_k$. The time scale for the WLW model starts at zero, when patients are first under observation, and this is the case for all events being modelled for. In the WLW model only the time since the start of the observation period for patients is important and no consideration is given to the time between events or when the previous event occurred. When using the WLW model if k events are being modelled for each patient is thought to be a risk for those k events. Therefore in the WLW model each patient is artificially considered to be at risk of suffering an event whether or not they have suffered the preceding event.

4.5 Prentice, Williams and Peterson Model

The Prentice, Williams and Peterson model, which will be referred to as the PWP model, was introduced in the paper by Prentice et al. (1981). The PWP model is a conditional model and there are two variations of the model. The PWPa model uses the total time and is defined as:

$$\lambda(t|N(t), \underline{z}(t)) = \lambda_{0s}(t)exp(\underline{z}(t)\underline{\beta}_s) \tag{4.4}$$

Whereas the PWPb model uses the gap time and is defined as:

$$\lambda(t|N(t), \underline{z}(t)) = \lambda_{0s}(t - t_{n(t)})exp(\underline{z}(t)\underline{\beta}_s) \tag{4.5}$$

The subscript s represents a stratification variable where $s = s(N(t), z(t), t)$ which may change as a function of time for a given patient. $N(t)$ is the number of events patients have suffered prior to time t. (Prentice et al., 1981) Both models PWPa and PWPb are conditional due to the fact that patients cannot be at risk of suffering an event until they have suffered the previous event. The order in which patients suffer events is therefore important. In the PWPa model the time scale is the same as that used in the AG model in that the time starts at zero and is continuous throughout the period of observation for patients, with the total time being of interest in the PWPa model. Whereas in the PWPb model the gap time between the successive events patients suffer is of interest. The time scale for the first event starts at zero, when the observation period for patients begins, and after every subsequent event patients suffer it is then reset back to zero. This is shown in Equation 4.5 by the underlying hazard term, $\lambda_{0s}(t - t_{n(t)})$, where t is the time till the current event of interest and $t_{n(t)}$ was the time until the previous event. In the PWPa model the underlying baseline hazard is taken as a function of the time since the start of the period of observation but in the PWPb model the baseline hazard is taken as a function of the time since the previous event.

4.6 Recurrent Event Data Example

To aid the understanding of the different recurrent models and the differences between them an example of recurrent data will be given. The example data used is from patients who were enrolled in the IONA Study. (The IONA Study Group, 2002) In this example only the first four primary endpoints that patients

could suffer will be considered and the recurrent event data for three patients will be presented. Patient one, as denoted in Table 4.1, suffered the maximum four events that are being considered at 313 days, 315 days, 691 days and 863 days after randomisation. Patient two suffered an event 402 days after randomisation and then 958 days after randomisation the observation period ended for patient two without them suffering a further event, this may have been as the study follow-up had come to an end or patient two may have been censored for another reason. Patient three suffered no events during their period of observation and that ended 966 days after randomisation. The example recurrent event data is shown in Table 4.1. In Table 4.1 if the event indicator equals zero then a patient did not suffer an event and if it equals 1 then the patient did suffer an event.

Table 4.1: The different ways the data is set up for the four models for analysing recurrent event data using data from the IONA Study as an example

		Patient One			Patient Two			Patient Three		
Model	Event Number	Time Interval	Event Indicator	Stratum Indicator	Time Interval	Event Indicator	Stratum Indicator	Time Interval	Event Indicator	Stratum Indicator
AG	1	(0, 313]	1	1	(0, 402]	1	1	(0, 966]	0	1
	2	(313, 315]	1	1	(402, 958]	0	1			
	3	(315, 691]	1	1						
	4	(691, 863]	1	1						
WLW	1	(0, 313]	1	1	(0, 451]	1	1	(0, 966]	0	1
	2	(0, 315]	1	2	(0, 958]	0	2	(0, 966]	0	2
	3	(0, 691]	1	3	(0, 958]	0	3	(0, 966]	0	3
	4	(0, 863]	1	4	(0, 958]	0	4	(0, 966]	0	4
PWPa	1	(0, 313]	1	1	(0, 402]	1	1	(0, 966]	0	1
	2	(313, 315]	1	2	(402, 958]	0	2			
	3	(315, 691]	1	3						
	4	(691, 863]	1	4						
PWPb	1	(0, 313]	1	1	(0, 402]	1	1	(0, 966]	0	1
	2	(0, 2]	1	2	(0, 556]	0	2			
	3	(0, 376]	1	3						
	4	(0, 172]	1	4						

Firstly, looking at the data set up for the AG model it can be seen that patient one is in the risk set for any event from 0 days until 863 days and has an event which defines the risk set at four time points: 313 days, 315 days, 691 days and 863 days. In the AG model events are assumed to be independent and the order in which they take place is not considered. As a result the stratum indicator is set to 1 for all events for all patients. Patient two is in the risk set from 0 days until 958 days and has only one event that defines the risk set and that is at 402 days. Patient three suffered no events so has not events which define the risk set but is contained within the risk set from 0 days until 966 days.

Secondly, looking at the data set up for the WLW model the stratum indicator is important for this model and increases as the event number increases. For the WLW model the time interval each patient is at risk for always starts at the time patients were enrolled in the study. As patient one experiences four events the time intervals in this case are straightforward with them running from: 0 days until 313 days, 0 days until 315 days, 0 days until 691 days and 0 days until 863. In the WLW model each patient is thought to be at risk of suffering the total number of events being modelled. In this example this is four events. As a result of this even though patients two and three did not suffer four events they still have to have four time intervals as they are still thought to be at risk of suffering all four events. As a consequence of this the dataset used in the modelling process for the WLW model is very large. Patient two suffered an event 402 days after randomisation so the time interval for the first event or stratum is from 0 days until 402 days. Patient two continued to be monitored for a further 556 days without experiencing a second event so the time interval for the second stratum is from 0 days until 958 days, with the event indicator set to 0. For both the

third and fourth strata a time interval is still required. The time interval used for the missing events is just the last available time interval repeated. As a result the time intervals for both the third and fourth strata are from 0 days until 958 days. As patient two did not suffer a third or fourth event the event indicators for both of these strata are set to 0. Patient three did not suffer any events and the inclusion of them in Table 4.1 is mainly to illustrate a drawback of the WLW model. The time intervals for all four strata for patient three are from 0 days until 966 days. Even though this patient suffered no events they are still artificially considered to be at risk of suffering the four events. The information known about patients and the number of events that they suffer and in turn are at risk of suffering from is not utilised in the fitting of the WLW model.

Thirdly, the data set up for the PWPa model is exactly the same as for the AG model but this time the stratum indicator is used in a meaningful way. For both forms of the PWP model the order in which patients suffer events is important as they can only be a risk of suffering the k^{th} event once they have suffered the $k-1^{th}$ event. For the PWPb model the stratum variable is again meaningful but the time intervals are set up differently from in the AG model. As was stated in Section 4.5 the gap time between events is of importance when the PWPb model is being used so once a patient suffers an event the time interval is reset back to zero. Therefore, for patient one the time intervals for the four events they suffered or strata are: 0 days until 313 days, 0 days until 2 days, 0 days until 376 days and 0 days until 172 days. The two time intervals for patient two are from 0 days until 402 days, when they suffered an event, and from 0 days until 556 days when their period of study observation ended. The time interval for the one and only stratum for patient three is from 0 days until 966 days.

4.7 Summary of the Recurrent Event Models

Having described the different recurrent event models and illustrated how the data is set up for each of the models through an example it will now be discussed how the recurrent events models fit to the reality of a clinical trial and in particular the IONA Study. The AG model is the simplest of the models to use but due to its underlying assumption of independence it is unlikely that the AG model would be able to suitably model what happens during a clinical trial. Especially, in the case of CHD were it is known that once patients have suffered a CHD event they are at a heightened risk of suffering another CHD event. This would then also violate the assumption of constant risk for the AG model. It would therefore appear that the AG model would not be the most appropriate model to use to analyse recurrent event data where the events of interest are CHD events.

In the WLW, PWPa and PWPb models events are not assumed to be independent. It would therefore appear that one of these models would be more appropriate to model recurrent event data from a clinical trial involving CHD events. In the WLW model all patients are required to have a time interval for all the events being modelled for regardless of how many they actually suffered. This could potentially lead to problem in a recurrent event setting, especially if the numbers of later events are low, as earlier time intervals are carried forward and this could potentially bias the results. In addition, another potential problem is that as the time interval for all events in the WLW model start at the time patients are first under observation no consideration is given to the time between events. As a result the WLW model may be better suited to a competing risks (Crowder, 2001) setting as opposed to a recurrent event one.

When using either of the PWP models to analyse recurrent event data the models condition on post-randomisation events. This could potentially lead to bias in the results. However, when the recurrent event models were used to analyse the recurrent event data from the IONA Study, see Section 5.3, the estimates for the treatment effect for the PWP models were similar to those for the AG and WLW models. Neither the AG model nor the WLW model conditions on post-randomisation events. Therefore, if there is bias in the results for the PWP models due to conditioning on post-randomisation events the bias appears to be minor and have minimal effect on the results for the models. In both the PWPa and PWPb models the order in which patients suffer events is considered and patients are not artificially considered to be at risk of suffering the total number of events being modelled for if they did not actually suffer that number of events. Of the two versions of the PWP model the most appropriate model to analyse data from a clinical trial involving CHD events would appear to be the PWPb model. This is as a result of the fact that the time between events that patients suffer is used in the modelling process as opposed to just the total time, as is the case in the PWPa model. The PWPb model makes the best use of the available information to model the recurrent events.

Having introduced the four recurrent events in this Chapter and discussed the merits of using them to analyse recurrent event data from clinical trials the models will be used in Chapter 5 to model the recurrent event data recorded during the IONA Study. It should be noted that in Equations 4.3, 4.4 and 4.5 the full definitions of the WLW, PWPa and PWPb models are shown. When using with the full versions of the models each of the different event numbers has a separate treatment effect fitted. This is indicated in Equations 4.3, 4.4 and 4.5

by the fact that the vector of unknown parameters, $\underline{\beta}$, has a subscript attached to it. As there is no distinction made between event numbers in the AG model separate treatment effects for the different event numbers cannot be fitted. When the recurrent event data from the IONA Study was analysed of interest was the overall treatment effect as opposed to the treatment effect for the individual event numbers. As a result the actual versions of the WLW, PWPa and PWPb models fitted to the data were as follows:

1. The WLW Model

$$\lambda_k(t, \underline{z}_k(t)) = \lambda_{0k}(t) exp(\underline{z}(t)\underline{\beta}) \tag{4.6}$$

2. The PWPa Model

$$\lambda(t|N(t), \underline{z}(t)) = \lambda_{0s}(t) exp(\underline{z}(t)\underline{\beta}) \tag{4.7}$$

3. The PWPb model

$$\lambda(t|N(t), \underline{z}(t)) = \lambda_{0s}(t - t_{n(t)}) exp(\underline{z}(t)\underline{\beta}) \tag{4.8}$$

Chapter 5

Recurrent Event Modelling of the IONA Study

In Section 2.4 survival analysis was carried out on the time-to-first event for the primary and secondary endpoints of the IONA Study as well as the GI events patients suffered. Patients were at risk of suffering more than one of each type of event and in this Chapter the recurrent event models introduced in Chapter 4 will be applied to the three different types of event.

5.1 The Number of Recurrent Events Patients Suffered

After patients had suffered a first primary or secondary endpoint or a GI event their period of follow-up did not end. They continued to be monitored until the study ended or they were lost to follow-up for other reasons. Any subsequent CHD and GI events that patients suffered were recorded.

5.1.1 Numbers of Primary Endpoints

The distribution of patients by the number of primary endpoints suffered can be seen in Table 5.1. The numbers of patients who suffered a particular number of events and the cumulative numbers are given. From Table 5.1 it can be seen that 337 (13.1%) patients suffered at least one primary endpoint in the nicorandil

Table 5.1: Distribution of the number (%) of patients by the number of recurrent primary endpoints suffered

	Number of Patients with $\geq k$ Events		Number of Patients with k Events	
k	Nicorandil (n = 2,565)	Placebo (n = 2,561)	Nicorandil (n = 2,565)	Placebo (n = 2,561)
0	2,228 (86.9%)	2,163 (84.5%)	2,228 (86.9%)	2,163 (84.5%)
1	337 (13.1%)	398 (15.5%)	247 (9.6%)	291 (11.4%)
2	90 (3.5%)	107 (4.2%)	67 (2.6%)	66 (2.6%)
3	23 (0.9%)	41 (1.6%)	14 (0.5%)	17 (0.7%)
4	9 (0.3%)	24 (0.9%)	6 (0.2%)	11 (0.4%)
5	3 (0.1%)	13 (0.5%)	1 (0.04%)	3 (0.1%)
6	2 (0.08%)	10 (0.4%)	1 (0.04%)	6 (0.2%)
7	1 (0.04%)	4 (0.2%)	0 (0.0%)	3 (0.1%)
8	1 (0.04%)	1 (0.04%)	0 (0.0%)	0 (0.0%)
9	1 (0.04%)	1 (0.04%)	1 (0.04%)	0 (0.0%)
10	0 (0.0%)	1 (0.04%)	0 (0.0%)	0 (0.0%)
11	0 (0.0%)	1 (0.04%)	0 (0.0%)	0 (0.0%)
12	0 (0.0%)	1 (0.04%)	0 (0.0%)	0 (0.0%)
13	0 (0.0%)	1 (0.04%)	0 (0.0%)	0 (0.0%)
14	0 (0.0%)	1 (0.04%)	0 (0.0%)	1 (0.04%)

group and of those 337 patients 247 (9.6%) suffered only one primary endpoint in total and 24 (0.9%) patients in the placebo group suffered at least four primary endpoints and of those 24 patients 11 (0.4%) suffered exactly four primary endpoints. The greatest number of primary endpoints a patient suffered was fourteen and this patient was in the placebo group of the study. All fourteen of the events

this patient suffered were unplanned hospital admissions for cardiac chest pain. The maximum number of events suffered by a patient in the nicorandil group was nine and of these nine events eight were unplanned hospital admissions for cardiac chest pain and the remaining event was a non-fatal MI. The distribution of the number of patients who suffered primary endpoints of the IONA Study is also illustrated graphically in Figure 5.1.

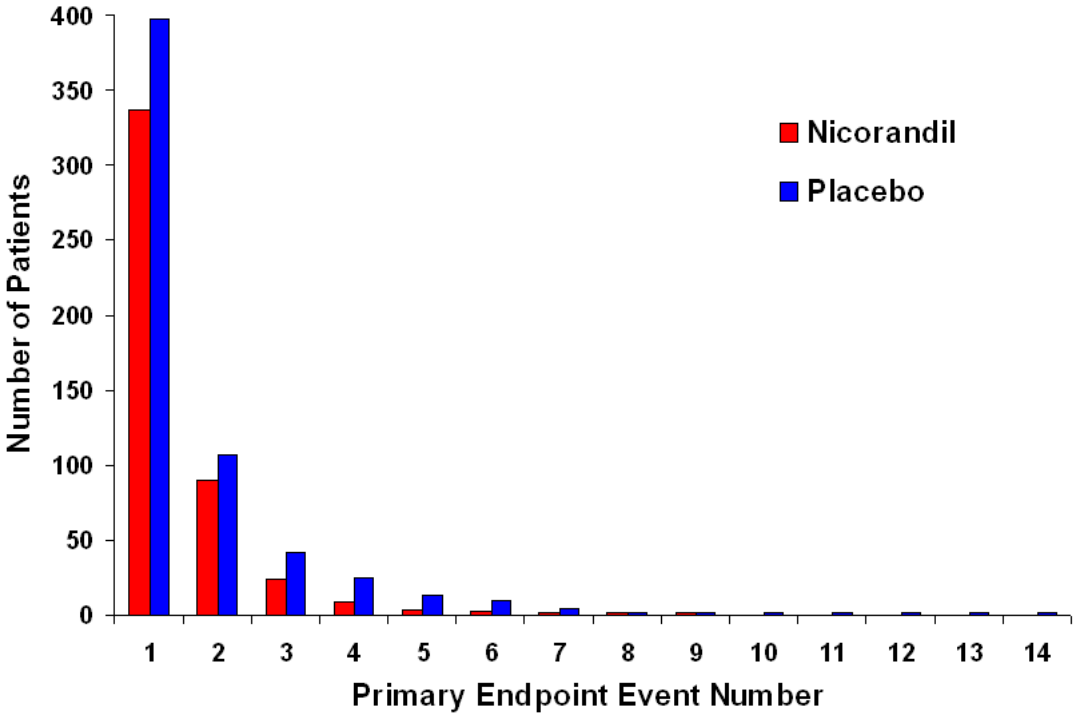


Figure 5.1: Cumulative distribution of the number of primary endpoints patients suffered during the IONA Study

The overall impression is that patients in the nicorandil group suffered fewer recurrent primary endpoints. This could imply that nicorandil not only reduced the risk of first CHD events but also the risk of recurrent CHD events. It can

also be seen that a limited number of patients suffered more than six events, 1 (0.04%) patient in the nicorandil group and 4 (0.2%) patients in the placebo group. Whether the difference in the risk of patients suffering primary endpoints between the nicorandil and placebo groups remained significant with the inclusion of the recurrent events patients suffered, as it was for the time-to-first event, will be examined using the four models discussed in Chapter 4 for analysing recurrent event data.

5.1.2 Numbers of Secondary Endpoints

The numbers of secondary endpoints are presented in an analogous fashion in Table 5.2. The distribution of the number of patients who suffered secondary endpoints of the IONA Study is also illustrated graphically in Figure 5.2. There were fewer secondary endpoints than primary endpoints and far fewer recurrent

Table 5.2: Distribution of the number (%) of patients by the number of recurrent secondary endpoints suffered

	Number of Patients with $\geq k$ Events		Number of Patients with k Events	
k	Nicorandil (n = 2,565)	Placebo (n = 2,561)	Nicorandil (n = 2,565)	Placebo (n = 2,561)
0	2,458 (95.8%)	2,427 (94.8%)	2,458 (95.8%)	2,427 (94.8%)
1	107 (4.2%)	134 (5.2%)	95 (3.7%)	117 (4.6%)
2	12 (0.5%)	17 (0.7%)	12 (0.5%)	15 (0.6%)
3	0 (0.0%)	2 (0.08%)	0 (0.0%)	2 (0.08%)

events. This is mainly due to the large number of primary endpoints that were due to unplanned hospital admissions for cardiac chest pain. In the secondary endpoint only the harder clinical outcomes of CHD death and non-fatal MI were considered. The largest number of secondary endpoints that a patient suffered

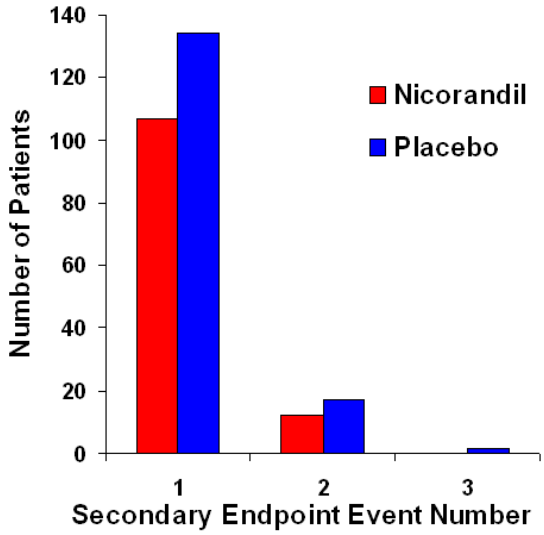


Figure 5.2: Cumulative distribution of the number of secondary endpoints patients suffered during the IONA Study

was three and 2 (0.08%) patients in the placebo group suffered three events. The greatest number of secondary endpoints a patient in the nicorandil group suffered was two and 12 (0.5%) patients suffered two events. Patients in the placebo group suffered more recurrent secondary endpoints that those patients in the nicorandil group but the difference between the two treatment groups appears less marked in comparison to the difference for the primary endpoint.

5.1.3 Numbers of Gastrointestinal Events

The number of GI events that patients suffered can be seen in Table 5.3. The distribution of the number of patients who suffered secondary endpoints of the IONA Study is also illustrated graphically in Figure 5.3. It can be seen that in the nicorandil group 2 (0.08%) patients suffered four GI events and in the placebo

Table 5.3: Distribution of the number (%) of patients by the number of recurrent GI events suffered

	Number of Patients with $\geq k$ Events		Number of Patients with x Events	
k	Nicorandil (n = 2,565)	Placebo (n = 2,561)	Nicorandil (n = 2,565)	Placebo (n = 2,561)
0	2,408 (93.9%)	2,453 (95.8%)	2,408 (93.9%)	2,453 (95.8%)
1	157 (6.1%)	108 (4.2%)	135 (5.3%)	91 (3.6%)
2	22 (0.9%)	17(0.7%)	17 (0.7%)	13 (0.5%)
3	5 (0.2%)	4 (0.2%)	4 (0.2%)	4 (0.2%)
4	2 (0.08%)	0 (0.0%)	2 (0.08%)	0 (0.0%)

group the largest number of GI events was three and 4 (0.2%) patients suffered three GI events. Even though in total the number of GI events that patients suffered in the nicorandil group was larger than the number in the placebo group the difference after the first GI events was not substantial. In the placebo group 15.7% of patients compared to 14.0% in the nicorandil group went on to suffer a subsequent GI event among those suffering a first GI event.

5.2 Changing Risk of Successive Events

Before fitting the recurrent event models the distributions of times to recurrent events for the three types of event were explored. In order to do this Kaplan-Meier (Kaplan and Meier, 1958) estimates were constructed to estimate the time-to-event distributions for each event conditional on the previous event and measuring time from the calendar time of the previous event. This is equivalent to how the PWPb model models recurrent event data. The numbers of successive events studied has been limited based on the number of patients suffering that number of events.

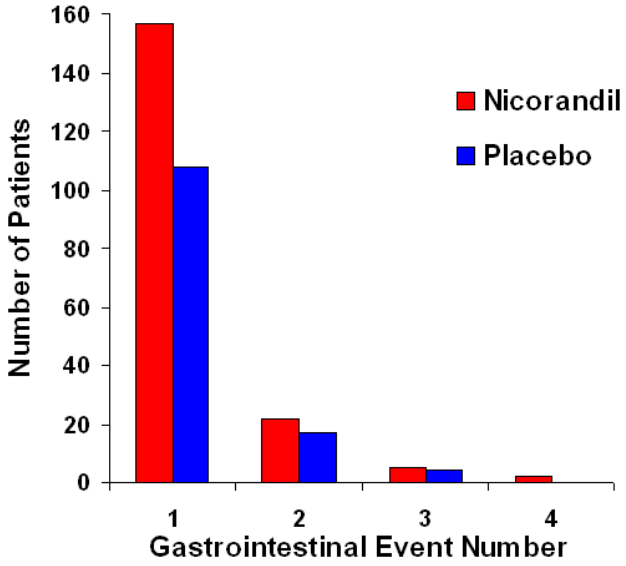
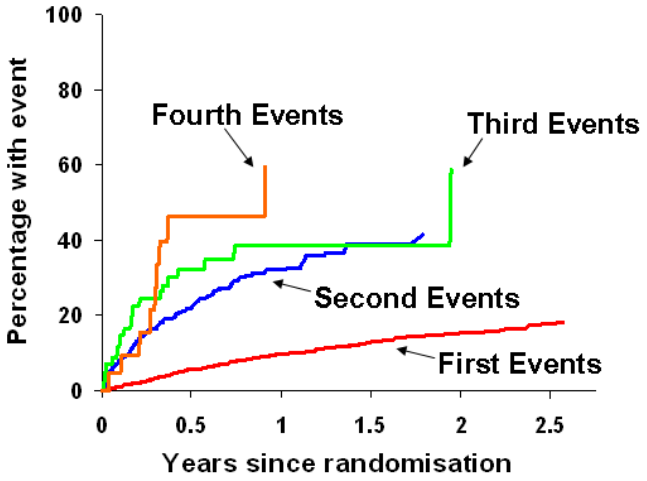


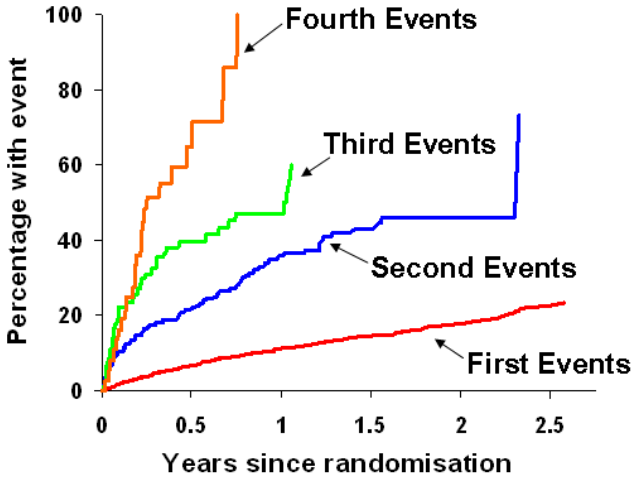
Figure 5.3: Cumulative distribution of the number of GI events patients suffered during the IONA Study

5.2.1 Risk of Successive Primary Endpoints

For the primary endpoint only the first four events were considered due to the limited number of patients who suffered more than four primary endpoints. The Kaplan-Meier estimates for the first four events for the nicorandil and placebo groups can be seen in Figure 5.4. Once a patient had suffered a first primary endpoint their risk of suffering a second primary endpoint was increased and the same was true for the risk of suffering a third and then a fourth event. The primary endpoints were clearly not independent and this will have implications for fitting the AG model. There were two stages to the increased risk. Firstly, there was an acute increase immediately after patients had suffered an event. Secondly, after this initial acute phase was over the increase in risk flattened



(a) The nicorandil group



(b) The placebo group

Figure 5.4: Risk of successive primary endpoints for patients in the nicorandil and placebo groups of the IONA Study

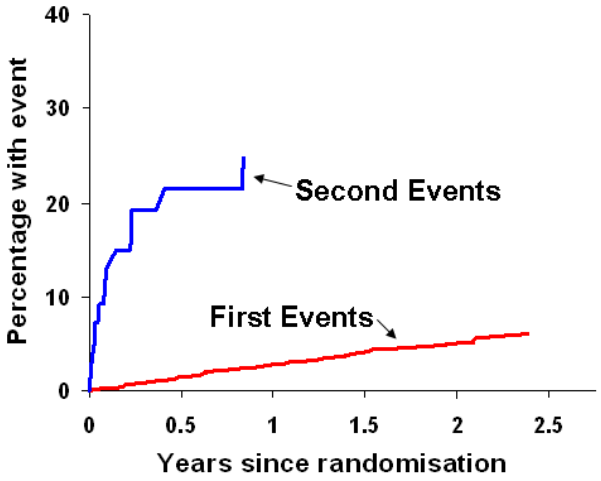
out to a level that was greater that the risk of suffering the previous event. The increasing risk between successive primary endpoints appears to be greater in the placebo group, Figure 5.4b, than in the nicorandil group, Figure 5.4a. Indicating that nicorandil may not only reduce the risk of patients suffering a first primary endpoint but also reduce the risk of patients suffering subsequent CHD events.

5.2.2 Risk of Successive Secondary Endpoints

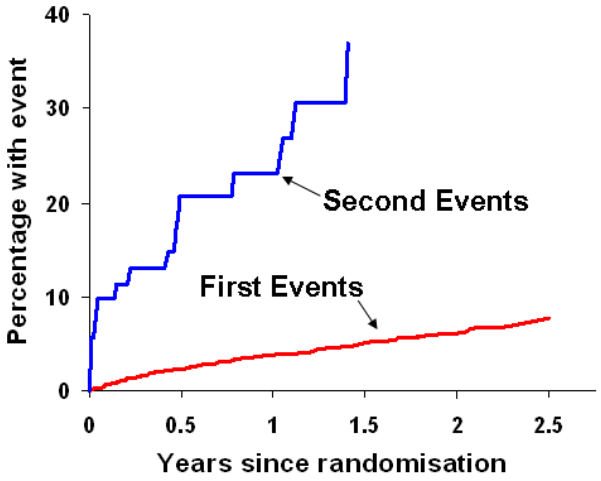
As no patients suffered three secondary endpoints in the nicorandil group and only 2 (0.08%) did in the placebo group, see Table 5.2, only the first two secondary endpoints were considered. The Kaplan-Meier estimates for the first two secondary endpoints that patients suffered in both the nicorandil and placebo groups can be seen in Figure 5.5. As was the case for the primary endpoint in both the nicorandil and placebo groups the risk of suffering a second secondary endpoint was increased compared to the risk of suffering a first secondary endpoint. The increase in risk was greatest immediately after a patient's first event then lessened. As only a small number of patients suffered recurrent events there may well be insufficient patient data available to draw any reliable conclusions from the Kaplan-Meier curves.

5.2.3 Risk of Successive Gastrointestinal Events

The risk of suffering the first three GI events was studied. The Kaplan-Meier estimates for the first three GI events that patients suffered in both the nicorandil and placebo groups can be seen in Figure 5.6. As with the primary and secondary endpoints the risk of suffering a second and then a third event was increased

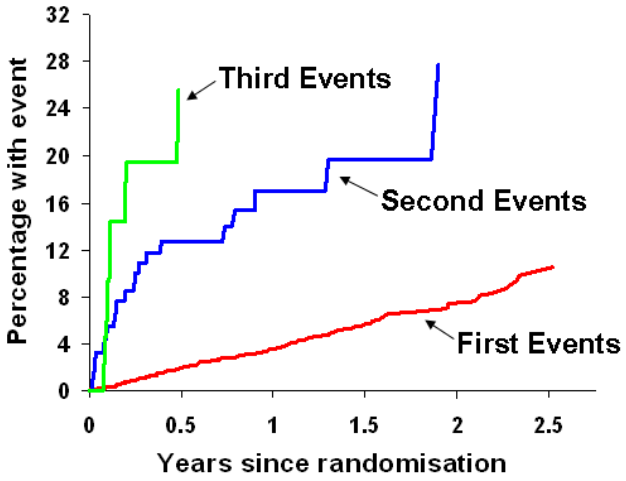


(a) The nicorandil group

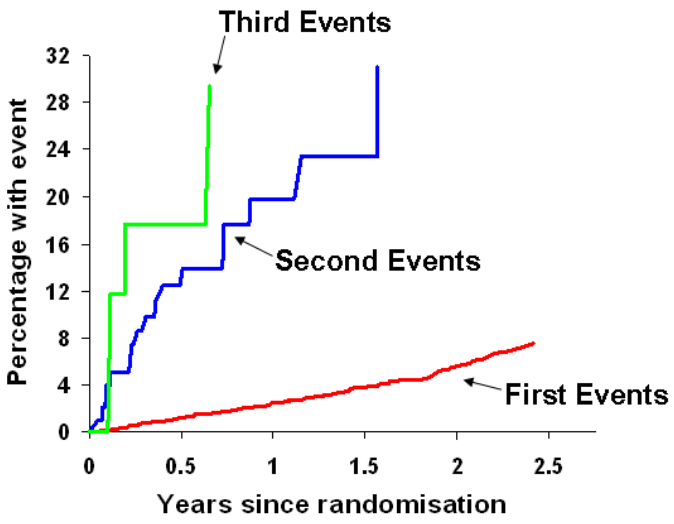


(b) The placebo group

Figure 5.5: Risk of successive secondary endpoints for patients in the nicorandil and placebo groups of the IONA Study



(a) The nicorandil group



(b) The placebo group

Figure 5.6: Risk of successive GI events for patients in the nicorandil and placebo groups of the IONA Study

significantly after each previous event. The difference between the Kaplan-Meier estimates for the first and second GI events in both the nicorandil and placebo groups was substantial. Any difference between the curves for the second and third GI events was less evident.

5.3 Recurrent Event Analysis

The four recurrent event models that were introduced in Chapter 4 were applied to the data from the IONA Study for the recurrent primary and secondary endpoints that patients suffered as well as the recurrent GI events. (Ford et al., 2007) The robust sandwich estimator for the covariance matrix devised by Lin and Wei (1989) and previously introduced in Section 4.2 will be used in all of the recurrent event analyses. Additionally, as was stated in Section 4.2, a cut-off point is needed for the number of recurrent events to be included in the modelling process. The cut-off point was judged individually for each different type of event, based on the number of recurrent events that patients suffered. In the initial recurrent event analyses the only explanatory variable included in the models was the randomised treatment patients were receiving.

5.3.1 Primary Endpoints

It can be seen from Table 5.1 that only 5 patients, 1 (0.04%) in the nicorandil group and 4 (0.2%) patients in the placebo group, suffered more that six primary endpoints and due to the small number of patients and events that they went on to suffer the recurrent event analysis was carried out on only the first six events, resulting in 14 (1.3%) of the 1,071 primary endpoints that patients suffered being

omitted from the analysis. Even using only the first six primary endpoints there were 322 more primary endpoints used in the modelling process and nicorandil prevented an additional 68 events that were not considered in the time-to-first event analysis. The results for the four recurrent event models as well as the original time-to-first event Cox Model are shown in Table 5.4.

Table 5.4: The time-to-first event Cox model and recurrent event analyses based on the first six primary endpoints patients suffered during the IONA Study. The table contains estimated hazard ratios, 95% CIs and p-values

Model	Hazard Ratio	95% Confidence Interval	P-Value
Cox	0.83	0.72 – 0.97	0.014
AG	0.78	0.66 – 0.91	0.0024
WLW	0.77	0.65 – 0.91	0.0028
PWPa	0.82	0.73 – 0.93	0.0015
PWPb	0.82	0.73 – 0.92	0.0012

It can be seen that, based on the findings of the recurrent event models, treatment with nicorandil was not only effective at reducing the risk of patients suffering first primary endpoints but also recurrent primary endpoints. Analysis was carried out using the recurrent event models on the total number of primary endpoints that patients suffered and the findings were similar to those seen in Table 5.4. The HRs and 95% CIs were broadly similar for the four different recurrent event models, the AG and WLW models gave almost identical results, as did the PWPa and PWPb models. Compared to the time-to-first event analysis the HRs and 95% CIs for the AG and WLW models were shifted downward away from the null value of 1. The point estimates for the reduction in risk of suffering recurrent primary endpoints were 22% (95% CI: 9%, 34%) and 23% (95% CI: 9%,

35%), respectively for the AG and WLW models compared to 17% (95% CI: 3%, 28%) for the time-to-first event analysis. The noticeable difference between the four recurrent event models and the time-to-first event analysis for the primary endpoint was in the p-values. The p-values for the four recurrent event models were considerably sharper than for the time-to-first event analysis, the p-values were 0.0024, 0.0028, 0.0015 and 0.0012 for the recurrent event models compared to 0.014 for the original Cox Model.

5.3.2 Secondary Endpoints

As there were only 2 (0.08%) patients in the placebo group and no patients in the nicorandil group who suffered three secondary endpoints the recurrent event analysis was carried out on only the first two secondary endpoints that patients suffered. With the inclusion of the second secondary endpoints there were an additional 29 events included in the modelling process and nicorandil prevented an additional 5 events compared to the time-to-first event analysis. The recurrent event models were used to see whether the inclusion of these additional events made the reduction in risk of suffering secondary endpoints that treatment with nicorandil caused in patients nominally statistically significant. The results for the recurrent event models as well as the Cox Model for the time-to-first analysis can be seen in Table 5.5.

The results for the recurrent events models were almost identical to the result for the time-to-first event analysis in terms of HRs and 95% CIs. The only real difference was in terms of the p-values and that was for the PWPa and PWPb models where the p-values were numerically larger and therefore further away

Table 5.5: The time-to-first event Cox model and recurrent event analyses based on the first two secondary endpoints patients suffered during the IONA Study. The table contains estimated hazard ratios, 95% CIs and p-values

Model	Hazard Ratio	95% Confidence Interval	P-Value
Cox	0.79	0.61 – 1.02	0.068
AG	0.78	0.60 – 1.02	0.066
WLW	0.78	0.60 – 1.02	0.065
PWPa	0.80	0.63 – 1.03	0.079
PWPb	0.80	0.63 – 1.02	0.072

from the statistically significant threshold of 0.05. The point estimates for the reduction in risk of suffering recurrent secondary endpoints were between 20% and 22% for the recurrent event models. If the four models were fitted again but this time with the inclusion of the third secondary endpoints that patients suffered, the results were similar to the findings shown in Table 5.5. The results for the secondary endpoint remain non-significant even with the inclusion of the recurrent events patient suffered and this may be due to the small number of recurrent events that patients did suffer, so the potential benefits of using the recurrent events models to analyse the data were lost.

5.3.3 Gastrointestinal Events

With only 2 (0.08%) patients in the nicorandil group and no patients in the placebo group suffering four GI events the recurrent event analysis was based on the first three GI events that patients suffered. More patients in the nicorandil group suffered a first GI event and this was reflected in the time-to-first event

analysis, shown in Section 2.4.3, showing a significant increase in the risk of patients suffering a GI event in the nicorandil group compared to the placebo group. After the first GI event patients suffered the numbers of patients who suffered further GI events in the nicorandil and placebo groups were similar but the percentages were higher in the placebo group. This may imply that the increase in risk of patients suffering a GI event that treatment with nicorandil caused was restricted to first events only or that due to the small number of recurrent GI events the continued increase in risk caused by treatment with nicorandil was masked. The four recurrent event models were fitted to the GI event data and the results are shown in Table 5.6, along with the results for original time-to-first event Cox Model.

Table 5.6: The time-to-first event Cox model and recurrent event analyses based on the first three GI events patients suffered during the IONA Study. The table contains estimated hazard ratios, 95% CIs and p-values

Model	Hazard Ratio	95% Confidence Interval	P-Value
Cox	1.46	1.14 – 1.86	0.0027
AG	1.41	1.09 – 1.83	0.0082
WLW	1.43	1.10 – 1.86	0.0082
PWPa	1.34	1.07 – 1.69	0.012
PWPb	1.36	1.08 – 1.71	0.0087

It can be seen that the results of the four recurrent event models all indicate that nicorandil did increase the risk of patients suffering recurrent GI events. The results for the AG and WLW models were broadly similar to the results for the time-to-first event analysis with only minimal changes in the point estimates for the HRs and the 95% CIs. For the PWPa and PWPb models the point estimates

for the HRs were shifted downward towards the null value of 1 as were the 95% CIs compared to the time-to-first event analysis. The p-values for the recurrent event models were less sharp compared to the time-to-first event analysis. The p-values for the AG, WLW and PWPb models ranged from 0.0082 to 0.0087, with the p-value for the PWPa model numerically larger still at 0.012 compared to 0.0027 for the time-to-first event analysis. The recurrent event models were run again to include the fourth GI events that patients suffered and only the first two GI events and results all indicated that treatment with nicorandil caused an increase in the risk of patients suffering GI events.

5.4 Interpreting the Performance of the Recurrent Event Models

In Section 5.3 the four recurrent event models were fitted to data for the primary and secondary endpoints of the study as well as the GI events patients suffered. For each of the three types of recurrent event analysed all of the recurrent event models gave broadly similar results but each method has different underlying assumptions and treats the time interval when patients were at risk of suffering events differently. It is important that any method used to analyse the data reflects what actually happened. Due to the limited recurrent event data available for both the secondary endpoint and the GI events the model interpretation will be based on the primary endpoint analysis

Firstly, looking at the AG model it assumes that the recurrent events for each individual patient are independent from each other and that the underlying risk of

suffering an event does not change as patients suffer further events. Once patients have suffered a non-fatal MI or an unplanned hospital admission for cardiac chest pain they are likely to be at an elevated risk of suffering a further CHD event. The elevated risk in both treatment groups is confirmed by looking at Figure 5.4. The increase in risk would also depend on the severity of the CHD they suffered. This implies that the recurrent primary endpoints patients suffered could not be classed as being independent. The assumptions that there is a constant risk of events and that the events are independent, which are fundamental to the AG model, were violated so the use of this model in this situation is not valid.

Secondly, with regard to the WLW model there is an issue with using this model which is due to the underlying theory behind the model, which is that every patient is at risk of suffering the total number of events being considered in the analysis and therefore must have a time interval for each event. As a result when the WLW model is being used each patient is artificially considered to be at risk of recurrent events whether or not they have had a preceding event. As the time intervals used in the WLW model always start at the point when patients are first under observation the time between events that patients suffer is not considered in the model. Only the time between the start of the observation period and the actual events is used in the modelling process. Therefore, the use of the WLW model is not appropriate.

Thirdly, looking at both the PWPa and PWPb models the theory behind the models does not assume that all the events are independent, the order in which events occur is considered and patients are not artificially considered to be at risk of recurrent events whether or not they have had a preceding event. However, there are potential drawbacks to using either of the PWP models. One

of the drawbacks to using either of the PWP models is the potential for bias in the results of the models with increasing rates of recurrent events. Although the recurrent event rate during the IONA Study was not sufficiently high to significantly bias the results of the PWP models. It would therefore appear that the use of one of the PWP models would be most appropriate in this situation.

The choice of which of the PWP models to use was based on the differences between how the time intervals for the recurrent events are set up for the two different versions of the model. In the PWPa model the total time patients have been under observation is considered as the time scale is continuous from the start of the period of observation whereas in the PWPb model the gap times between successive events is of interest as the time scale is reset to zero after every event patients suffer. This results in a difference in the risk for the different events or strata for the PWP models. For the first event both PWP models are the same as well as being identical to the Cox model for a time-to-first event analysis but for subsequent events this is not true. In the PWPa model the subsequent risk is for patients who have suffered k events, where $k \geq 2$, and have been observed for the same length of time since the start of the study. Whereas, for the PWPb model the risk is for patients who have suffered k events, where $k \geq 2$, and have been at risk of suffering the $k + 1^{th}$ event for the same period of time with the total length of time patients have been in the study for not considered. (Prentice et al., 1981) The increased risk patients were at after suffering recurrent events is better reflected in the PWPb model. The use of the PWPb model is recommended in the case of the IONA Study as this model most faithfully reflects the observed patterns in the data.

In other situations the PWPb model may not be the most appropriate model

to use. For example if the events could be considered to be independent and the underlying risk of suffering events did not alter as patients suffered events then the AG model may well be the most efficient model to use. The AG model also has the advantage that it is the simplest of the models to implement. (Hosmer and Lemeshow, 1999) For each different set of circumstances the choice of which model to use should be carefully considered and based on all the available information.

5.5 Multivariable Models for the Recurrent Event Data

In Sections 3.4 and 3.6 multivariable predictive models for the first primary endpoint and GI events that patients suffered were produced. Having fitted recurrent event models to the primary endpoints and GI events, with the only explanatory variable included in the models being the randomised treatment patients were receiving, the multivariable predictive models produced for the time-to-first event analysis were fitted to the recurrent event data. As was discussed in Section 5.4 the PWPb model appears to model the data most accurately so the multivariable model results will only be shown for the PWPb model.

5.5.1 Primary Endpoints

The results for the multivariable recurrent event model for the first six primary endpoints that patients were at risk of suffering using the PWPb model can be seen in Table 5.9. The model for the primary endpoint used was the one produced by the first method of model building. All of the nine variables selected

in the time-to-first event model as being statistically significant remained so in the multivariable recurrent event model. In both models the comparison between CCSF levels I and II were non-significant but were still included in the models to aid their use and interpretation.

Table 5.7: The multivariable recurrent event model for the primary endpoint fitted using the PWPb model based on the first six primary endpoints patients suffered during the IONA Study. For this model n = 5,059. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	0.81	0.72 – 0.92	0.0010
BMI (increase of 2 kg/m^2)	0.97	0.94 – 1.00	0.048
Current Smoker	1.22	1.05 – 1.43	0.0097
Previous MI	1.40	1.21 – 1.63	<0.0001
Previous Stroke or TIA	1.30	1.07 – 1.57	0.0082
History of LVH	1.37	1.15 – 1.64	0.0005
CCSF level II vs. level I	1.14	0.96 – 1.36	0.15
CCSF level III/IV vs. level I	1.57	1.28 – 1.93	<0.0001
Long Acting Nitrates	1.30	1.12 – 1.50	0.0004
Loop Diuretics	1.39	1.22 – 1.59	<0.0001

The HRs were not greatly changed in the recurrent event model from those seen in the time-to-first event model apart from the HR for the comparison between CCSF levels I and III/IV where the increased risk for patients classified as level III/IV compared to level I dropped from 100% to 57%. The 95% CIs were narrower when the recurrent events were included and this may be as a result of more information being utilised in the modelling process. Looking at the p-values there were more apparent differences when the models were compared. As was seen with the original recurrent event analysis for the multivariable recurrent event model the p-value for treatment with nicorandil was significantly sharper

than it was for the time-to-first event analysis, it was 0.0010 compared to 0.010. For the recurrent event model the p-value for the BMI of patients was close to being non-significant at 0.048 as opposed to 0.0065 for the time-to-first event model. Although, the point estimates for the HRs were virtually the same for the two models. If either of the AG or WLW models were fitted BMI no longer remained a significant predictor. In the time-to-first event multivariable model the p-value for the comparison between CCSF levels I and II was bordering on the significant threshold at 0.060. In the recurrent event model the p-value was larger at 0.15. The p-values for the remaining variables included in the model were not greatly altered between the time-to-first event and recurrent event models.

5.5.2 Gastrointestinal Events

Four multivariable time-to-first event models were produced for the GI events patients suffered and recurrent event versions were produced of all models but only Model B1 will be shown. However, when recurrent event versions of Models B1 and B2 were produced in all cases apart from when the PWPa model was used with Model B2 the variable treatment with diltiazem no longer remained significant, although the trend was still to an increased risk of GI events. As a consequence the results for the multivariable recurrent event model, Model B1, using the PWPb model can be seen in the left half of Table 5.8 and in the right half is shown Model B1 with the variable treatment with diltiazem removed. With the variable treatment with diltiazem removed Model B1 becomes Model A1. The interpretation of the model will therefore be based on the results for Model A1. With the inclusion of the first three GI events patients were at risk of suffering

all the variables which were selected in the time-to-first event model remained significant apart from the comparison between CCSF levels I and II, which in the recurrent event version of Model A1 was non-significant. The comparison between CCSF levels I and II was still included in the model, as it had previously been for the primary endpoint models.

Comparing the time-to-first event and recurrent event models there were minimal changes in the HRs and p-values. For the comparison between CCSF levels I and III/IV the increased risk dropped from 76% to 56% and the p-value increased from 0.010 to 0.023. Additionally the p-value for treatment with loop diuretics was sharper. The increased risk of GI events associated with treatment with nicorandil dropped by 9% to 35% with the inclusion of the recurrent events in the analysis. The p-value for treatment with nicorandil was numerically larger and therefore closer to being non-significant. These points further indicate that while treatment with nicorandil increased the risk of patients suffering a first GI event, the effect of treatment with nicorandil on the risk of suffering subsequent GI events was reduced. The significant difference in the risk of suffering first GI events, associated with treatment with nicorandil, may be behind the increased risk of GI events shown in the recurrent event analysis.

Table 5.8: The multivariable recurrent event models, Models B1 and A1, for the GI events fitted using the PWPb model based on the first three GI events patients suffered during the IONA Study. For both models n = 5,097. The table contains estimated hazard ratios, 95% CIs and p-values

	Model B1			Model A1		
Variable	Hazard Ratio	95% Confidence Interval	P-Value	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	1.35	1.07 – 1.70	0.011	1.35	1.07 – 1.69	0.011
Weight (increase of 5 kg)	0.95	0.92 – 0.99	0.0074	0.95	0.92 – 0.99	0.0084
CCSF level II vs. level I	1.31	0.96 – 1.78	0.089	1.33	0.98 – 1.81	0.070
CCSF level III/IV vs. level I	1.54	1.05 – 2.26	0.029	1.56	1.06 – 2.29	0.023
Long Acting Nitrates	1.45	1.14 – 1.85	0.0024	1.47	1.15 – 1.86	0.0019
Loop Diuretics	1.47	1.16 – 1.87	0.0017	1.46	1.15 – 1.86	0.0019
Statins	0.73	0.58 – 0.91	0.0054	0.73	0.58 – 0.91	0.0054
Diltiazem	1.21	0.96 – 1.54	0.11	–	–	–

5.6 Fitting a Shared Frailty Model to the Recurrent Event Data

Having produced multivariable predictive models for the first primary endpoint and GI events in Chapter 3 it was seen that other variables apart from treatment with nicorandil had a significant effect on the risk of patients suffering both types of event. This was also the case when these models were fitted to the recurrent event data in Section 5.5. With the inclusion of additional variables the models should explain much of the distribution for the risk of suffering events, especially for the time-to-first event analysis. However, the heterogeneity of patients is unlikely to be fully explained as other factors, some that were recorded and others that were not, would likely influence the risk of patients suffering events.

In fitting the same multivariable predictive models to the recurrent event data as for the time-to-first event the level of heterogeneity that was explained was likely decreased. This will be for a variety of reasons. Firstly, the models were not specifically developed for the recurrent event data so more appropriate models may exist. Secondly, as the recurrent events were being modelled it is likely other developing risk factors will play a larger roll in determining the risk of patients. Factors which influence second and subsequent events, such as the type of previous event patients suffered, are not included in the modelling process.

In order to try and account for the unknown heterogeneity for the multivariable recurrent event model a shared frailty model was fitted. This involved adding a frailty parameter to the model in addition to the other prognostics variables already included. A Gamma frailty parameter was added to the multivariable

recurrent event model for the primary endpoint. A shared frailty model was fitted to the GI recurrent event data but the frailty parameter was found not to be significant so will not be shown here. As in Section 5.5 the results for the only the PWPb model will be shown. The results of fitting the model including the frailty parameter can be seen in Table 5.9.

Table 5.9: The multivariable recurrent event model for the primary endpoint fitted with a Gamma frailty parameter using the PWPb model based on the first six primary endpoints patients suffered during the IONA Study. For this model n = 5,059. The table contains estimated hazard ratios, 95% CIs and p-values

Variable	Hazard Ratio	95% Confidence Interval	P-Value
Nicorandil	0.77	0.65 – 0.90	0.0011
BMI (increase of 2 kg/m^2)	0.96	0.93 – 0.99	0.018
Current Smoker	1.34	1.09 – 1.65	0.0048
Previous MI	1.56	1.31 – 1.87	<0.0001
Previous Stroke or TIA	1.37	1.02 – 1.84	0.036
History of LVH	1.55	1.22 – 1.98	0.0004
CCSF level II vs. level I	1.19	0.98 – 1.45	0.084
CCSF level III/IV vs. level I	1.97	1.51 – 2.58	<0.0001
Long Acting Nitrates	1.43	1.21 – 1.69	<0.0001
Loop Diuretics	1.57	1.30 – 1.90	<0.0001
Frailty	SD = 1.65		<0.0001

It can be seen that the frailty parameter did add significantly to the model as the p-value for it was <0.0001. The general effect of adding the frailty parameter was to increase the effect of the variables included in the model on the risk of patients suffering CHD events. The CIs were also wider. With the inclusion of the frailty parameter the decrease in the risk of patients who were treated with nicorandil increased from 19% to 23%. The p-values were not greatly altered with the exception of the variable whether patients had suffered a previous stroke or

TIA. The variable remained a significant predictor but the p-value was increased and was therefore much nearer the significance threshold of 0.05. The addition of the frailty parameter did add significantly to the model and altered the point estimates of the HRs and accompanying CIs but it did not change the overall findings. It is noted that the use of a single frailty term suffers from some of the weakness associated with the use of adjustment for baseline covariates in that it cannot deal with evolving frailty as a patient experiences recurrent events.

Chapter 6

Simulation of Recurrent Event Data

In Chapter 4 the various recurrent event models were introduced and it was seen that they have different underlying assumptions, hence their use may not be appropriate for every situation. In this Chapter the performance of the models will be assessed through the simulation of recurrent event data.

6.1 The Simulation Process

The recurrent event models were introduced in Chapter 4 and then applied to the recurrent event data from the IONA Study in Chapter 5. It was seen that the models have different underlying assumptions and that they gave slightly different results when applied to the IONA Study data. The differences were small in the case of the IONA Study. In order to assess the performance of the models under different known conditions, simulations were carried out. The data simulated

were the gap times between events for patients. The gap times were generated from Weibull distributions. The parameters used to generate the gap times were chosen in relation to what had been observed for the primary endpoint of the IONA Study. In addition it is known that patients have different underlying risks of suffering events, such as CHD events. This is due to the heterogeneity of patients. To reflect this heterogeneity a random subject effect was included in some of the simulations.

In each simulation there were two treatment groups indexed by i, with the placebo group = 1 and the treatment group = 2. Each simulation contained $2n$ patients, indexed by j, j = 1,...,n. The recurrent events patients were at risk of suffering from were indexed by k, k = 1,...,K, 1 for the first event, 2 for the second etc. There were two models for the gap times. The first did not include a random subject effect, see Equation 6.1, and the second did, see Equation 6.2.

$$log(y_{ijk}) = \alpha + T_{i_k} + log(\epsilon_j); i = 1, 2, j = 1, ..., n, k = 1, ..., K \qquad (6.1)$$

$$log(y_{ijk}) = \alpha + T_{i_k} + log(\epsilon_j) + log(\eta_j); i = 1, 2, j = 1, ..., n, k = 1, ..., K \qquad (6.2)$$

where y_{ijk} is the gap time for patient j for event k in treatment group i, α is the intercept, T_{i_k} is the treatment effect for group i for event k, ϵ_j is the random error term for patient j, which is an Exponential random variable, and η_j is the random subject effect for patient j, which is a positive stable random variable. In the simulations patients were at risk of suffering four events, $K = 4$, or being followed for a maximum time interval of two years, whichever occurred first. At the two year point patients who were still at risk of suffering further events had their current time interval censored.

Three groups of simulations were performed:

1. The No Treatment Effect Group of Simulations

   In order that the calibration of the simulation process could be assessed a group of null simulations were performed. In this group of simulations the treatment effect was set to zero and $T_{1_k} = T_{2_k}$ for k = 1,2,3,4.

2. The Treatment Effect Group of Simulations

   Simulations were performed where the treatment was effective compared to the placebo for all four events and $T_{1_k} < T_{2_k}$ for k = 1,2,3,4.

3. The Treatment Effect for First Event Group of Simulations

   Simulations were performed were the treatment was effective compared to the placebo for only the first event after which the treatment effect was set to zero. In this group of simulations $T_{1_k} < T_{2_k}$ for $k = 1$ and $T_{1_k} = T_{2_k}$ for k = 2,3,4.

When the treatment was effective it reduced the risk of patients suffering events compared to the placebo.

The risk of suffering a primary endpoint of the IONA Study increased with each successive event patients suffered, see Figure 5.4. This increase in risk was also seen for the secondary endpoint as well as the GI events. Therefore, as well as keeping the risk of events constant over time the risk of patients suffering a second, third and fourth event were increased. This was done using the increased risk of suffering a primary endpoint. The hazard for a second event was increased by a factor of four compared to that for a first event. For the third events the hazard was increased by a further factor of one and a half and finally for the

fourth events by another factor of one and a half. There were then four different sets of conditions for each group of simulations:

1. The First Set of Simulation Conditions

   The risk of patients suffering events did not increase as patients suffered events and a random subject effect was not applied.

2. The Second Set of Simulation Conditions

   The risk of patients suffering events did increase as patients suffered events and a random subject effect was not applied.

3. The Third Set of Simulation Conditions

   The risk of patients suffering events did not increase as patients suffered events and a random subject effect was applied.

4. The Fourth Set of Simulation Conditions

   The risk of patients suffering events did increase as patients suffered events and a random subject effect was applied.

6.2 Generating the Recurrent Event Data

There were 10,000 replicates of each simulation performed. The individual replicates contained 1,000 patients in both the placebo and treatment groups, $n =$ 1,000. The random patient effect, η, was generated from the family of positive stable random variables so that the assumption of proportional hazards was maintained for both the model conditioning on this term and for the unconditional model. (Hougaard, 1986; Crowder, 1989) The patient effect was generated as

a positive stable random variable with index $\frac{1}{2}$. This stable random variable is equivalent to an inverse Gaussian which in turn is equivalent to the inverse of a χ^2 random variable with one degree of freedom (Ford et al., 1995), which was used to generate the subject effect. The effect of increasing the risk of patients suffering events and adding a random subject effect was to increase the event rate. Shown in Table 6.1 are the distributions used to generate the gap times between events for the four sets of simulation conditions.

Table 6.1: Distributions used to generate the gap times between events for patient j for the four sets of simulation conditions

Group	Distribution for Gap Times			
	Event One	Event Two	Event Three	Event Four
The First Set of Simulation Conditions				
Placebo	$\text{We}(\lambda_1, \gamma_p)$	$\text{We}(\lambda_2, \gamma_p)$	$\text{We}(\lambda_3, \gamma_p)$	$\text{We}(\lambda_4, \gamma_p)$
Treatment	$\text{We}(\lambda_1, \gamma_p)$	$\text{We}(\lambda_2, \gamma_p)$	$\text{We}(\lambda_3, \gamma_p)$	$\text{We}(\lambda_4, \gamma_p)$
The Second Set of Simulation Conditions				
Placebo	$a_1\text{We}(\lambda_1, \gamma_p)$	$a_2\text{We}(\lambda_2, \gamma_p)$	$a_3\text{We}(\lambda_3, \gamma_p)$	$a_4\text{We}(\lambda_4, \gamma_p)$
Treatment	$a_1\text{We}(\lambda_1, \gamma_p)$	$a_2\text{We}(\lambda_2, \gamma_p)$	$a_3\text{We}(\lambda_3, \gamma_p)$	$a_4\text{We}(\lambda_4, \gamma_p)$
The Third Set of Simulation Conditions				
Placebo	$\eta\text{We}(\lambda_1, \gamma_p)$	$\eta\text{We}(\lambda_2, \gamma_p)$	$\eta\text{We}(\lambda_3, \gamma_p)$	$\eta\text{We}(\lambda_4, \gamma_p)$
Treatment	$\eta\text{We}(\lambda_1, \gamma_p)$	$\eta\text{We}(\lambda_2, \gamma_p)$	$\eta\text{We}(\lambda_3, \gamma_p)$	$\eta\text{We}(\lambda_4, \gamma_p)$
The Fourth Set of Simulation Conditions				
Placebo	$\eta a_1\text{We}(\lambda_1, \gamma_p)$	$\eta a_2\text{We}(\lambda_2, \gamma_p)$	$\eta a_3\text{We}(\lambda_3, \gamma_p)$	$\eta a_4\text{We}(\lambda_4, \gamma_p)$
Treatment	$\eta a_1\text{We}(\lambda_1, \gamma_p)$	$\eta a_2\text{We}(\lambda_2, \gamma_p)$	$\eta a_3\text{We}(\lambda_3, \gamma_p)$	$\eta a_4\text{We}(\lambda_4, \gamma_p)$

The scale parameters of the Weibull distributions, λ, are indexed by m, m = 1,2,3,4, as the values of λ_m were dependent on which group of simulations were being performed. Show in Table 6.2 are the different values of λ_m used in the simulations. The shape parameters of the Weibull distributions, γ, are indexed by p, $p = 1, 2$, as two different shape parameters were used. Firstly, $\gamma_1 = 1$ and secondly $\gamma_2 = 0.75$. The second value of γ was used to reflect the

Table 6.2: Scale parameters used in the different groups of simulations

Simulation Group	Parameter			
	λ_1	λ_2	λ_3	λ_4
The No Treatment Effect Group				
Placebo	0.1116	0.1116	0.1116	0.1116
Treatment	0.1116	0.1116	0.1116	0.1116
The Treatment Effect Group				
Placebo	0.1116	0.1116	0.1116	0.1116
Treatment	0.0558	0.0558	0.0558	0.0558
The Treatment Effect for First Event Group				
Placebo	0.1116	0.1116	0.1116	0.1116
Treatment	0.0558	0.1116	0.1116	0.1116

change in risk seen for patients after they had suffered a primary endpoint. There was an immediate acute increase in risk after which the increase flattened out to a level above that for the previous event. In order to achieve this effect the shape parameter of the Weibull distribution has to be $0 < \gamma < 1$. (Metalfe and Thompson, 2006) However, patients who enrolled in the IONA Study had not just suffered a CHD event so they would not have been at an initial heightened risk of suffering a further event. Therefore, to reflect the patient population of the IONA Study for the first events $\gamma_p = \gamma_1$ in all cases. The hazard increasing factors, a, indexed by q, q = 1,2,3,4, discussed in Section 6.1 were only applied in the second and fourth sets of simulations conditions. The hazard of a first event was not increased so $a_1 = 1$. As a result the event times for the first and second sets of simulation conditions were the same. This was also true for the third and fourth sets of simulation conditions. The random subject effect η was only applied in the third and fourth sets of simulation conditions.

The gap times between events were of a proportional hazards form. This was

true for the individual events independently from the other events and of censoring. The inclusion of the random subject effect, η, should halve the parameter estimate. (Ford et al., 1995) As a result the expected parameters estimates and the resulting HRs for the individual events were known in advance of the simulations. Therefore, the results for the recurrent event analyses, under perfect conditions, could be inferred for the first and second groups of simulations. For the first group of simulations the parameter estimate should equal 0 and the HR should equal 1. For the second group of simulation the parameter estimate should equal -0.6931 and the HR should equal 0.5 for the first and second sets of simulation conditions. For the third and fourth sets of simulation conditions the parameter estimates should equal -0.3466 and the HR should equal 0.7071. As the treatment effect for the first event group of simulations contained a mixture of events where the treatment was both effective compared to the placebo and the same as the placebo the expected results for this group of simulations were unknown. The simulation program was written and performed in SAS version 9.1.3 (SAS, Cary, North Carolina). The SAS code used to generate the gap time data for the three groups of simulations is shown in Appendix B. Additionally, shown in Appendix B is an example of the SAS code used to incorporate censoring into the gap time data and then analyse it using the different recurrent event models.

6.3 The Simulation Results

The results shown are the mean and standard error (SE) for the parameter estimates and HRs for the 10,000 replicates of each simulation. In addition the

percentage of replicates that showed a significant effect for either the placebo or treatment are shown. A replicate was judged to have shown a significant effect if the p-value was below the significance threshold of 0.05. The Cox model results for the first event only are also shown. The gap times for the first event were the same for the first and second and the third and fourth sets of simulation conditions. As a result for each pair of simulation conditions the results for the time-to-first analysis were the same. With the introduction of the increased risk of patients suffering events this was not the case for subsequent events. The results for the mean HRs are also illustrated graphically. For each set of simulation conditions the mean of the estimated HRs, with 95% CIs, for the different recurrent event models are shown. Due to the narrowness of the 95% CIs and that the results for the models varied the results for the AG and WLW models are shown on different axes from the results for the PWPa and PWPb models. For the no treatment effect and the treatment effect groups of simulations the expected HRs were known. Therefore, where the simulation results allow a reference line indicating the expected HR is shown for these two groups of simulations.

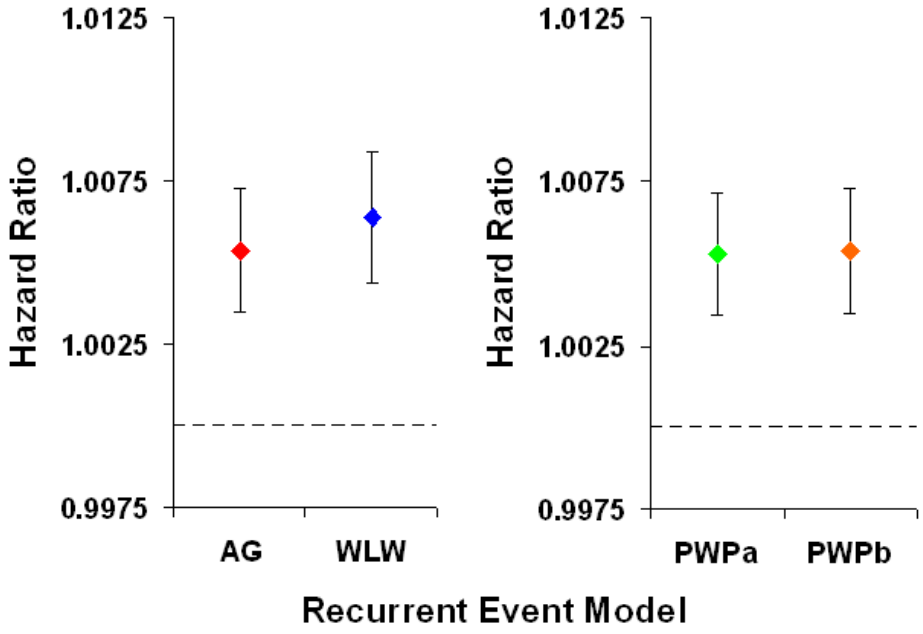
6.3.1 The No Treatment Effect Group of Simulations

The results for the no treatment effect group of simulations where the treatment effect was set to zero are shown first. The simulation results for $\gamma = 1$ are shown in Table 6.3 and the results for the HRs only are shown in Figures 6.1 and 6.2. The simulation results for $\gamma = 0.75$ are shown in Table 6.4 and the results for the HRs only are shown in Figures 6.3 and 6.4.

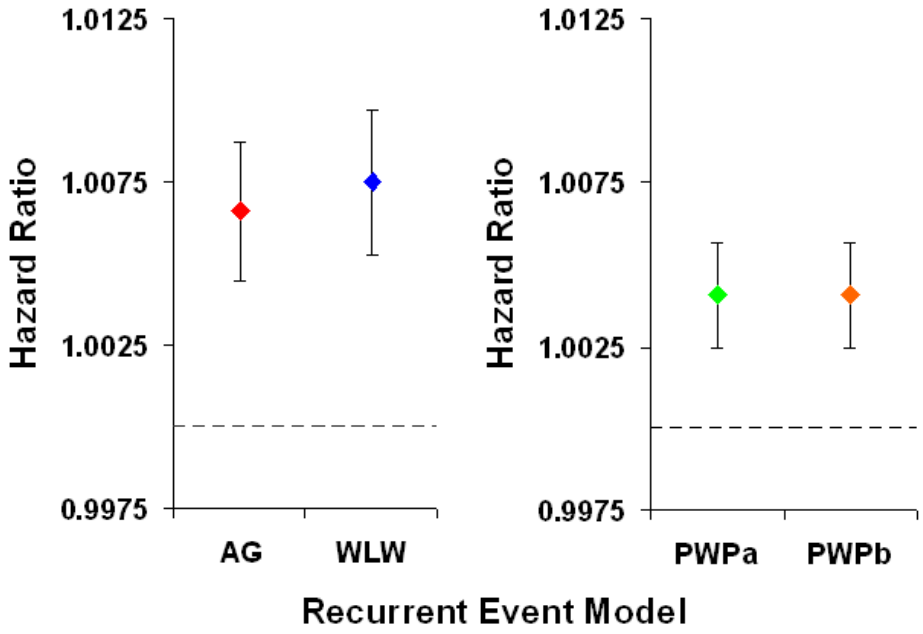
The aim of performing this group of simulations was to ascertain if there was

Table 6.3: Results for the no treatment effect group of simulations where $\gamma = 1$. The table contains the mean of the estimated parameter estimates (SE), hazard ratios (SE) and the percentage of replicates that showed a significant effect for either treatment group

Model	Parameter Estimate (SE)	Hazard Ratio (SE)	Significant Replicates	
			Treatment (%)	Placebo (%)
The First Set of Simulation Conditions				
Cox	0.0017 (0.0010))	1.0066 (0.0010)	2.26%	2.41%
AG	0.0010 (0.0009)	1.0054 (0.0009)	2.22%	2.52%
WLW	0.0011 (0.0010)	1.0064 (0.0010)	2.22%	2.49%
PWPa	0.0009 (0.0009)	1.0053 (0.0009)	2.23%	2.50%
PWPb	0.0009 (0.0009)	1.0054 (0.0009)	2.19%	2.54%
The Second Set of Simulation Conditions				
Cox	0.0017 (0.0010)	1.0066 (0.0010)	2.26%	2.41%
AG	0.0011 (0.0010)	1.0066 (0.0011)	2.45%	2.60%
WLW	0.0012 (0.0011)	1.0075 (0.0011)	2.43%	2.54%
PWPa	0.0008 (0.0008)	1.0041 (0.0008)	2.39%	2.62%
PWPb	0.0008 (0.0008)	1.0041 (0.0008)	2.31%	2.67%
The Third Set of Simulation Conditions				
Cox	<0.0001 (0.0006)	1.0021 (0.0006)	2.68%	2.25%
AG	-0.0001 (0.0007)	1.0021 (0.0007)	2.66%	2.46%
WLW	-0.0002 (0.0007)	1.0024 (0.0007)	2.74%	2.37%
PWPa	-0.0001 (0.0004)	1.0008 (0.0004)	2.74%	2.64%
PWPb	0.0002 (0.0005)	1.0013 (0.0005)	2.61%	2.41%
The Fourth Set of Simulation Conditions				
Cox	<0.0001 (0.0006)	1.0021 (0.0006)	2.68%	2.25%
AG	-0.0003 (0.0007)	1.0018 (0.0006)	2.64%	2.31%
WLW	-0.0003 (0.0007)	1.0020 (0.0007)	2.64%	2.31%
PWPa	-0.0003 (0.0005)	1.0008 (0.0005)	2.66%	2.77%
PWPb	0.00006 (0.0004)	1.0010 (0.0004)	2.50%	2.67%

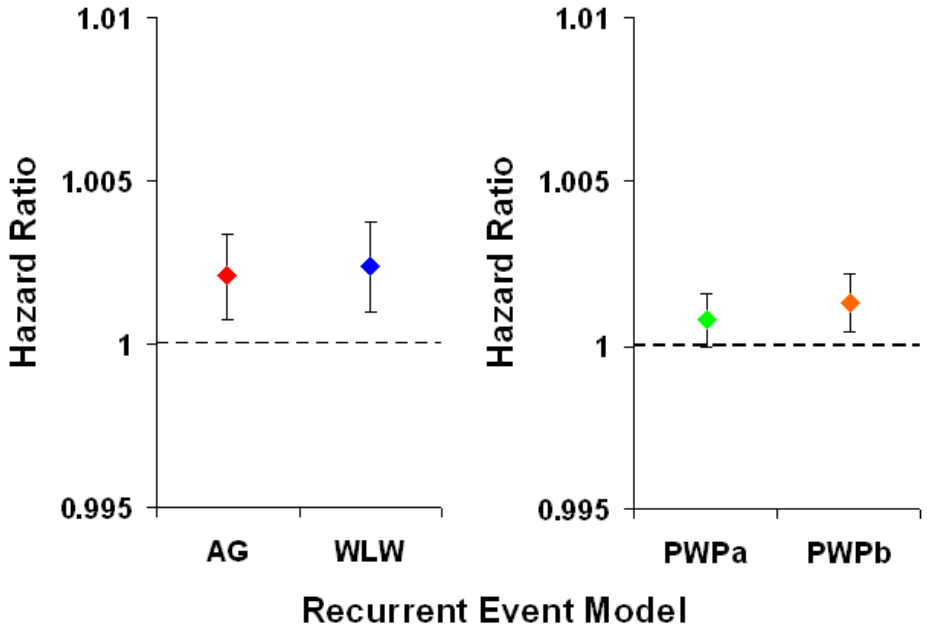


(a) The first set of simulation conditions

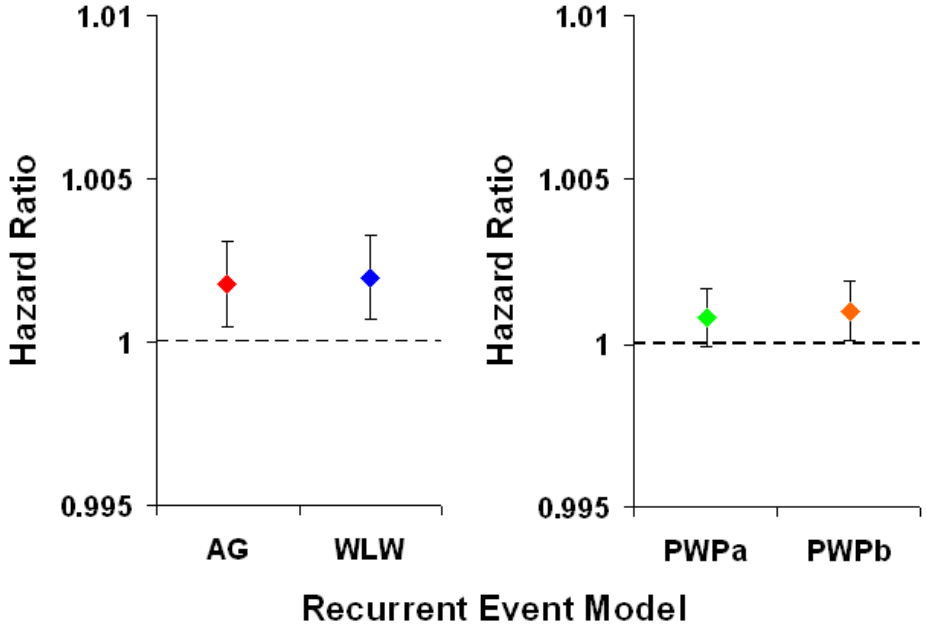


(b) The second set of simulation conditions

Figure 6.1: Hazard ratios for the first and second sets of simulation conditions for the no treatment effect group of simulations where $\gamma = 1$ with accompanying 95% CIs



(a) The third set of simulation conditions

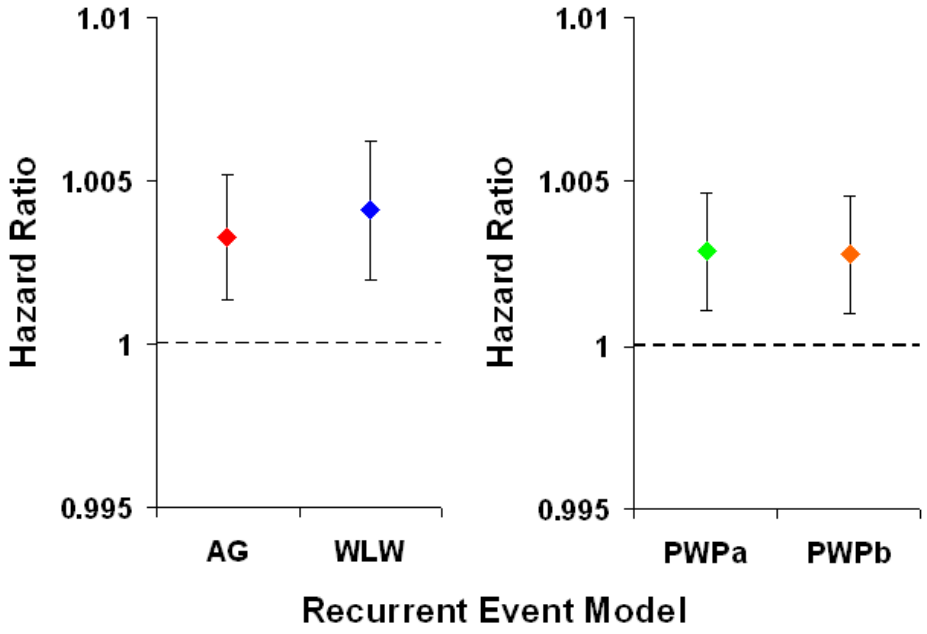


(b) The fourth set of simulation conditions

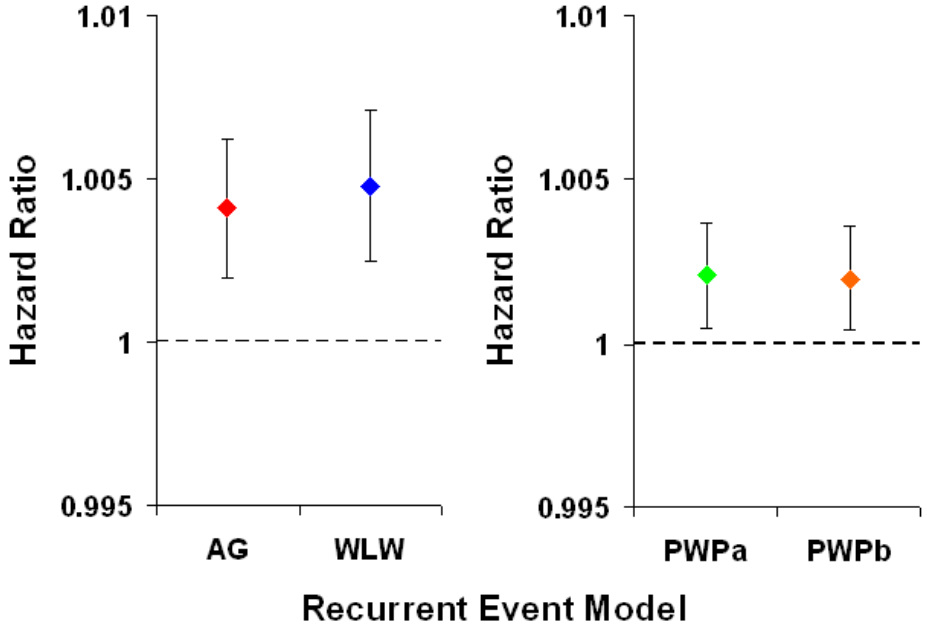
Figure 6.2: Hazard ratios for the third and fourth sets of simulation conditions for the no treatment effect group of simulations where $\gamma = 1$ with accompanying 95% CIs

Table 6.4: Results for the no treatment effect group of simulations where $\gamma = 0.75$. The table contains the mean of the estimated parameter estimates (SE), hazard ratios (SE) and the percentage of replicates that showed a significant effect for either treatment group

Model	Parameter Estimate (SE)	Hazard Ratio (SE)	Significant Replicates	
			Treatment (%)	Placebo (%)
The First Set of Simulation Conditions				
Cox	-0.0013 (0.0010)	1.0037 (0.0010)	2.70%	2.57%
AG	-0.0015 (0.0010)	1.0033 (0.0010)	2.57%	2.58%
WLW	-0.0017 (0.0011)	1.0041 (0.0011)	2.60%	2.54%
PWPa	-0.0014 (0.0009)	1.0029 (0.0009)	2.59%	2.47%
PWPb	-0.0013 (0.0009)	1.0028 (0.0009)	2.52%	2.55%
The Second Set of Simulation Conditions				
Cox	-0.0013 (0.0010)	1.0037 (0.0010)	2.70%	2.57%
AG	-0.0017 (0.0011)	1.0041 (0.0011)	2.57%	2.61%
WLW	-0.0018 (0.0011)	1.0048 (0.0012)	2.62%	2.51%
PWPa	-0.0011 (0.0008)	1.0021 (0.0008)	2.67%	2.66%
PWPb	-0.0010 (0.0008)	1.0020 (0.0008)	2.54%	2.72%
The Third Set of Simulation Conditions				
Cox	-0.0007 (0.00064)	1.0013 (0.0006)	2.49%	2.66%
AG	0.0001 (0.0007)	1.0022 (0.0007)	2.36%	2.46%
WLW	<0.0001 (0.0007)	1.0025 (0.0007)	2.25%	2.56%
PWPa	0.0001 (0.0004)	1.0010 (0.0004)	2.67%	2.36%
PWPb	0.0004 (0.0004)	1.0013 (0.0004)	2.26%	2.59%
The Fourth Set of Simulation Conditions				
Cox	-0.0007 (0.0006)	1.0013 (0.0006)	2.49%	2.66%
AG	-0.0004 (0.0006)	1.0017 (0.0006)	2.37%	2.66%
WLW	-0.0005 (0.0007)	1.0018 (0.0007)	2.41%	2.67%
PWPa	0.0002 (0.0005)	1.0013 (0.0005)	3.06%	2.84%
PWPb	0.0002 (0.0004)	1.0011 (0.0004)	2.17%	2.44%

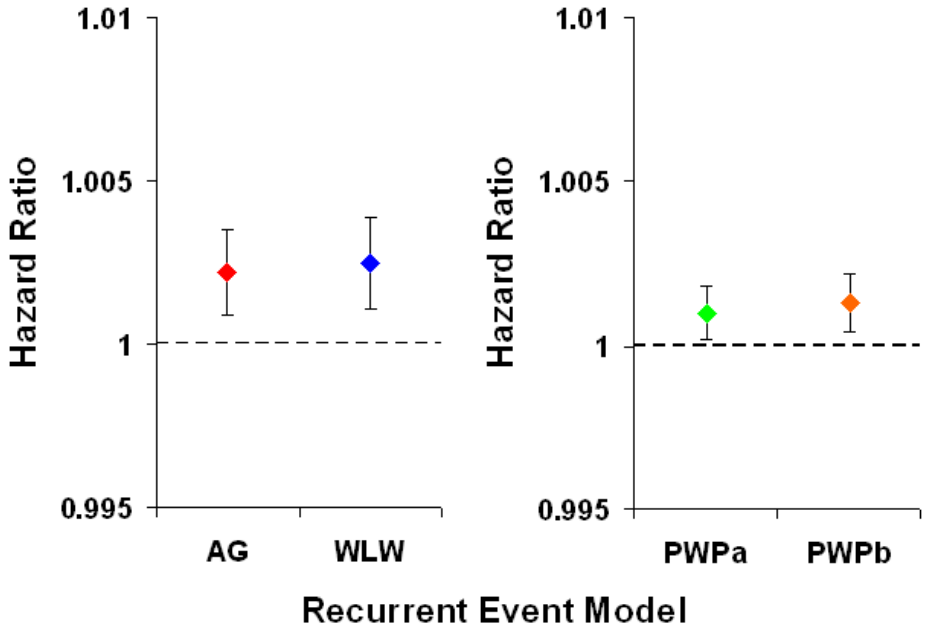


(a) The first set of simulation conditions

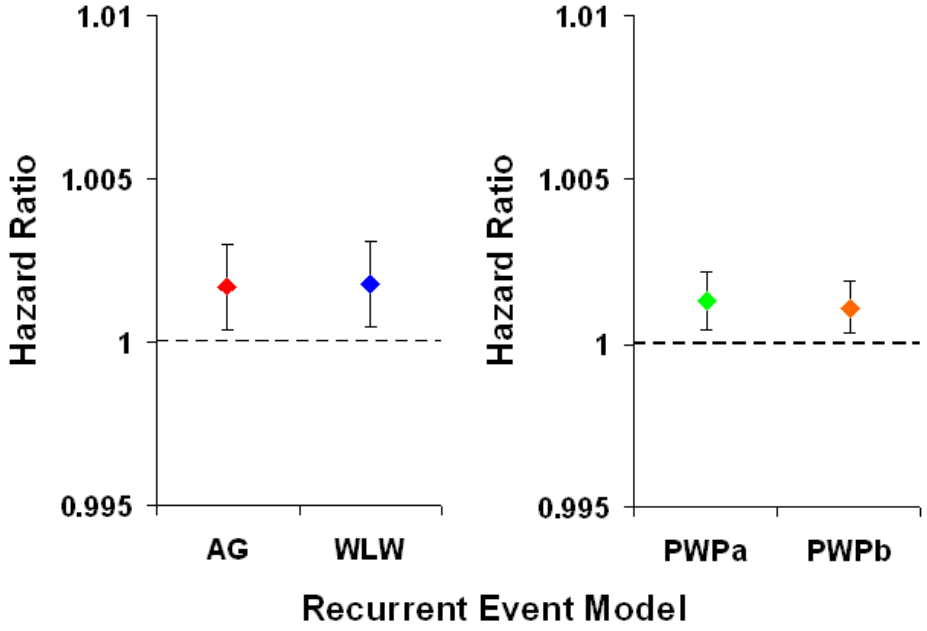


(b) The second set of simulation conditions

Figure 6.3: Hazard ratios for the first and second sets of simulation conditions for the no treatment effect group of simulations where $\gamma = 0.75$ with accompanying 95% CIs



(a) The third set of simulation conditions



(b) The fourth set of simulation conditions

Figure 6.4: Hazard ratios for the third and fourth sets of simulation conditions for the no treatment effect group of simulations where $\gamma = 0.75$ with accompanying 95% CIs

any form of bias in the simulation process. As all the HRs approximately equalled 1 there was no bias. As the treatment effect was set to zero it would be expected that roughly 2.5% of the replicates would show a significant effect for both the placebo and the treatment. This was reflected in the percentage of replicates that were significant. In this group of simulations even though the parameter estimates were close to zero the inclusion of the random subject effect did halve the parameter estimates. For the third and fourth sets of simulations conditions, where the random subject effect was included, the 95% CIs for the HRs were narrower than for the first and second sets of simulation conditions.

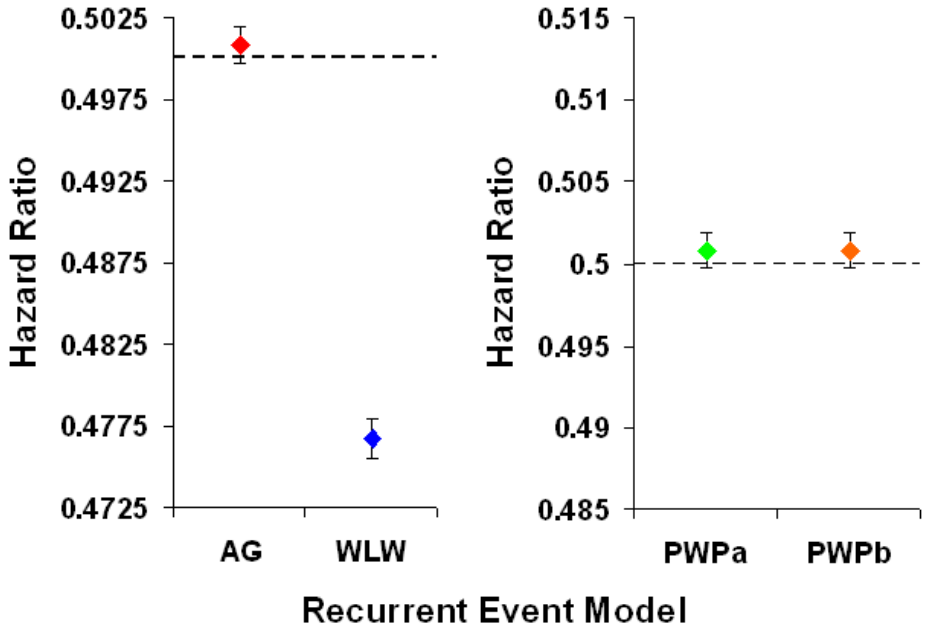
6.3.2 The Treatment Effect Group of Simulations

The results for the treatment effect group of simulations where the treatment was effective compared to the placebo for all four events are now shown. The simulation results for $\gamma = 1$ are shown in Table 6.5 and the results for the HRs only are shown in Figures 6.5 and 6.6. The simulation results for $\gamma = 0.75$ are shown in Table 6.6 and the results for the HRs only are shown in Figures 6.7 and 6.8.

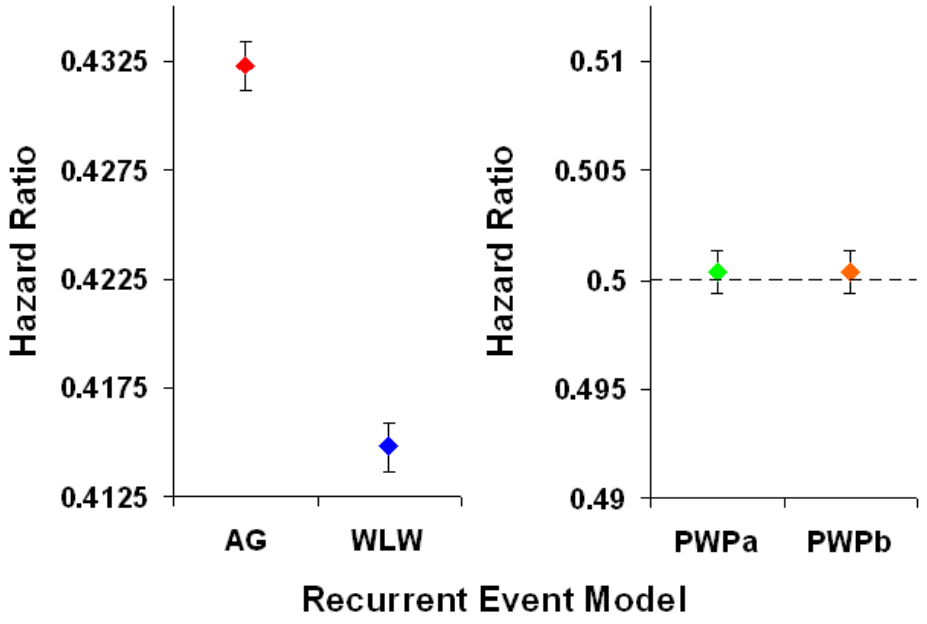
As the parameter estimates were not so close to zero for this group of simulations it is clearer that the random subject effect did halve the parameter estimates. Looking at the HRs, starting with the first set of simulation conditions when $\gamma = 1$ the AG, PWPa and PWPb models gave the expected results whereas the WLW model overestimates the treatment effect. When $\gamma = 0.75$ the AG model gives close to the expected result and the PWPa and PWPb models underestimated the treatment effect. The WLW model again overestimated the

Table 6.5: Results for the treatment effect group of simulations where $\gamma = 1$. The table contains the mean of the estimated parameter estimates (SE), hazard ratios (SE) and the percentage of replicates that showed a significant effect for either treatment group

Model	Parameter Estimate (SE)	Hazard Ratio (SE)	Significant Replicates	
			Treatment (%)	Placebo (%)
The First Set of Simulation Conditions				
Cox	-0.6982 (0.0012)	0.5011 (0.0006)	100%	0.0%
AG	-0.6982 (0.0012)	0.5008 (0.0006)	100%	0.0%
WLW	-0.7485 (0.0012)	0.4767 (0.0006)	100%	0.0%
PWPa	-0.6983 (0.0012)	0.5008 (0.0006)	100%	0.0%
PWPb	-0.6983 (0.0012)	0.5008 (0.0006)	100%	0.0%
The Second Set of Simulation Conditions				
Cox	-0.6982 (0.0012)	0.5011 (0.0006)	100%	0.0%
AG	-0.8465 (0.0013)	0.4323 (0.0005)	100%	0.0%
WLW	-0.8887 (0.0013)	0.4148 (0.0005)	100%	0.0%
PWPa	-0.6979 (0.0011)	0.5004 (0.0005)	100%	0.0%
PWPb	-0.6979 (0.0011)	0.5004 (0.0005)	100%	0.0%
The Third Set of Simulation Conditions				
Cox	-0.3469 (0.0007)	0.7086 (0.0005)	99.90%	0.0%
AG	-0.3343 (0.0007)	0.7178 (0.0005)	99.51%	0.0%
WLW	-0.3641 (0.0008)	0.6969 (0.0005)	99.67%	0.0%
PWPa	-0.1635 (0.0005)	0.8500 (0.0004)	95.81%	0.0%
PWPb	-0.1715 (0.0005)	0.8435 (0.0004)	93.28%	0.0%
The Fourth Set of Simulation Conditions				
Cox	-0.3469 (0.0007)	0.7086 (0.0005)	99.90%	0.0%
AG	-0.3567 (0.0007)	0.7017 (0.0005)	99.54%	0.0%
WLW	-0.3712 (0.0007)	0.6917 (0.0005)	99.95%	0.0%
PWPa	-0.1488 (0.0005)	0.8629 (0.0004)	85.53%	0.0%
PWPb	-0.1546 (0.0005)	0.8577 (0.0004)	90.62%	0.0%

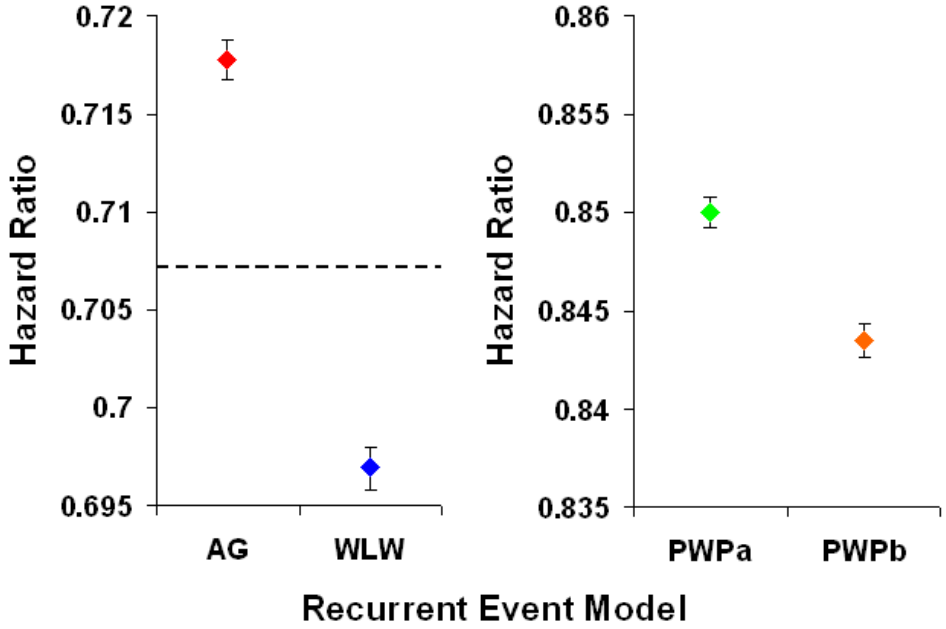


(a) The first set of simulation conditions

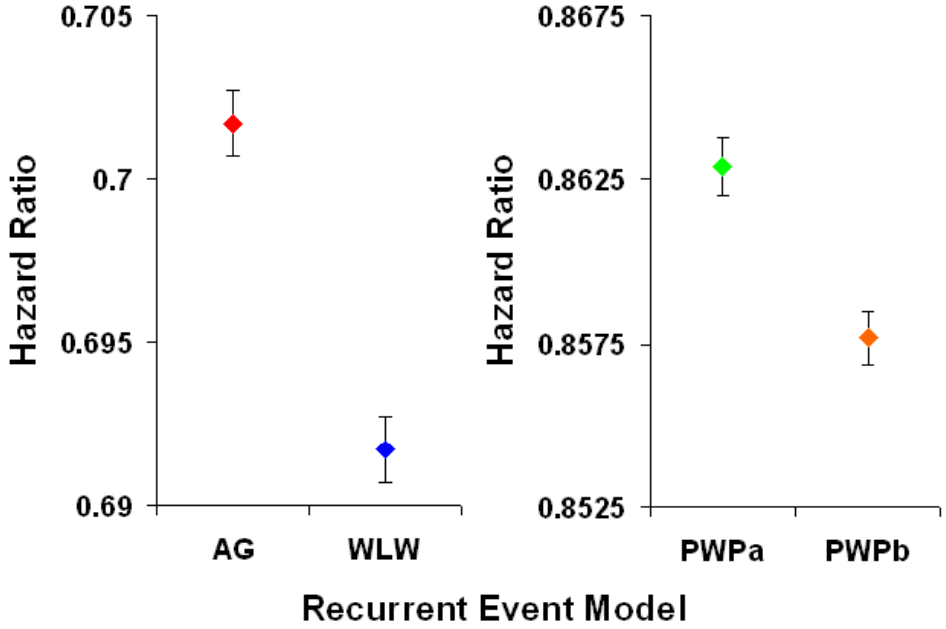


(b) The second set of simulation conditions

Figure 6.5: Hazard ratios for the first and second sets of simulation conditions for the treatment effect group of simulations where $\gamma = 1$ with accompanying 95% CIs



(a) The third set of simulation conditions

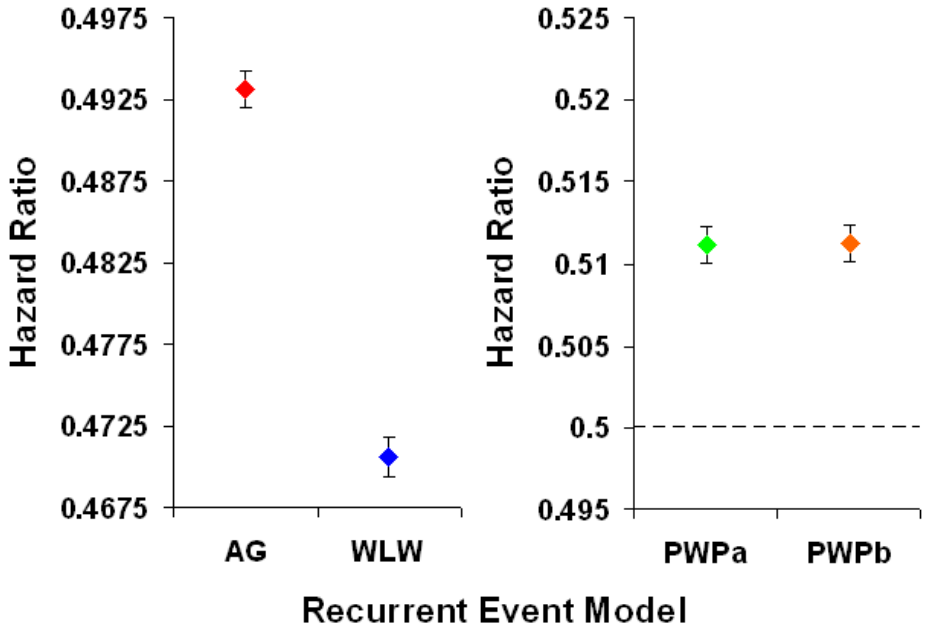


(b) The fourth set of simulation conditions

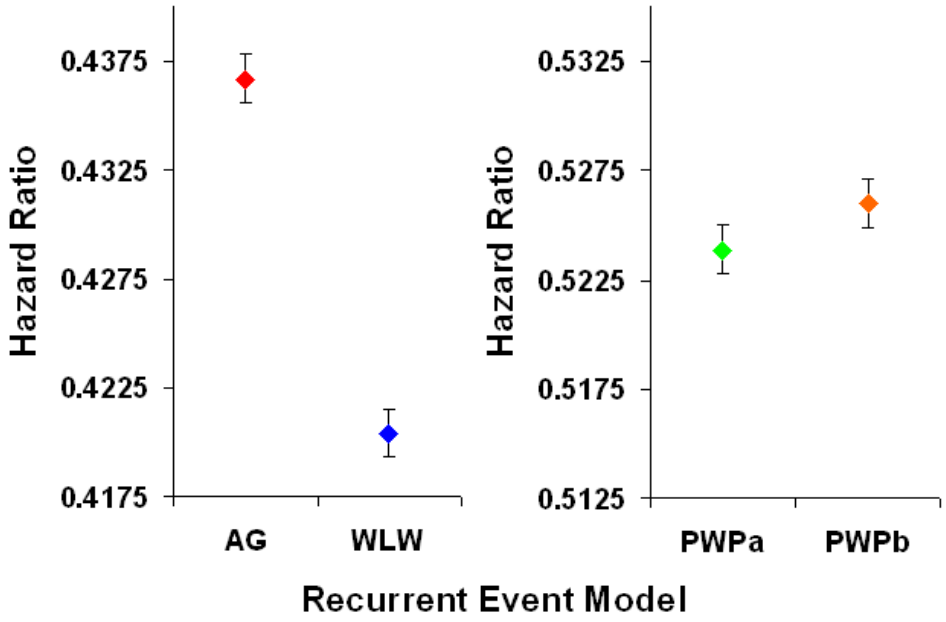
Figure 6.6: Hazard ratios for the third and fourth sets of simulation conditions for the treatment effect group of simulations where $\gamma = 1$ with accompanying 95% CIs

Table 6.6: Results for the treatment effect group of simulations where $\gamma = 0.75$. The table contains the mean of the estimated parameter estimates (SE), hazard ratios (SE) and the percentage of replicates that showed a significant effect for either treatment group

Model	Parameter Estimate (SE)	Hazard Ratio (SE)	Significant Replicates	
			Treatment (%)	Placebo (%)
The First Set of Simulation Conditions				
Cox	-0.6964 (0.0012)	0.5020 (0.0006)	99.99%	0.0%
AG	-0.7138 (0.0012)	0.4932 (0.0006)	100%	0.0%
WLW	-0.7617 (0.0013)	0.4706 (0.0006)	100%	0.0%
PWPa	-0.6775 (0.0011)	0.5112 (0.0006)	100%	0.0%
PWPb	-0.6772 (0.0011)	0.5113 (0.0006)	100%	0.0%
The Second Set of Simulation Conditions				
Cox	-0.6964 (0.0012)	0.5020 (0.0006)	99.99%	0.0%
AG	-0.8366 (0.0013)	0.4367 (0.0006)	100%	0.0%
WLW	-0.8754 (0.0013)	0.4204 (0.0006)	100%	0.0%
PWPa	-0.6519 (0.0010)	0.5239 (0.0005)	100%	0.0%
PWPb	-0.6479 (0.0010)	0.5260 (0.0005)	100%	0.0%
The Third Set of Simulation Conditions				
Cox	-0.3475 (0.0007)	0.7082 (0.0005)	99.93%	0.0%
AG	-0.3331 (0.0007)	0.7186 (0.0005)	99.64%	0.0%
WLW	-0.3636 (0.0008)	0.6973 (0.0005)	99.72%	0.0%
PWPa	-0.1632 (0.0005)	0.8503 (0.0004)	94.90%	0.0%
PWPb	-0.1715 (0.0005)	0.8434 (0.0004)	93.87%	0.0%
The Fourth Set of Simulation Conditions				
Cox	-0.3475 (0.0007)	0.7082 (0.0005)	99.93%	0.0%
AG	-0.3543 (0.0007)	0.7034 (0.0005)	99.88%	0.0%
WLW	-0.3696 (0.0007)	0.6929 (0.0005)	99.90%	0.0%
PWPa	-0.1488 (0.0005)	0.8629 (0.0004)	82.89%	0.0%
PWPb	-0.1540 (0.0005)	0.8582 (0.0004)	91.68%	0.0%

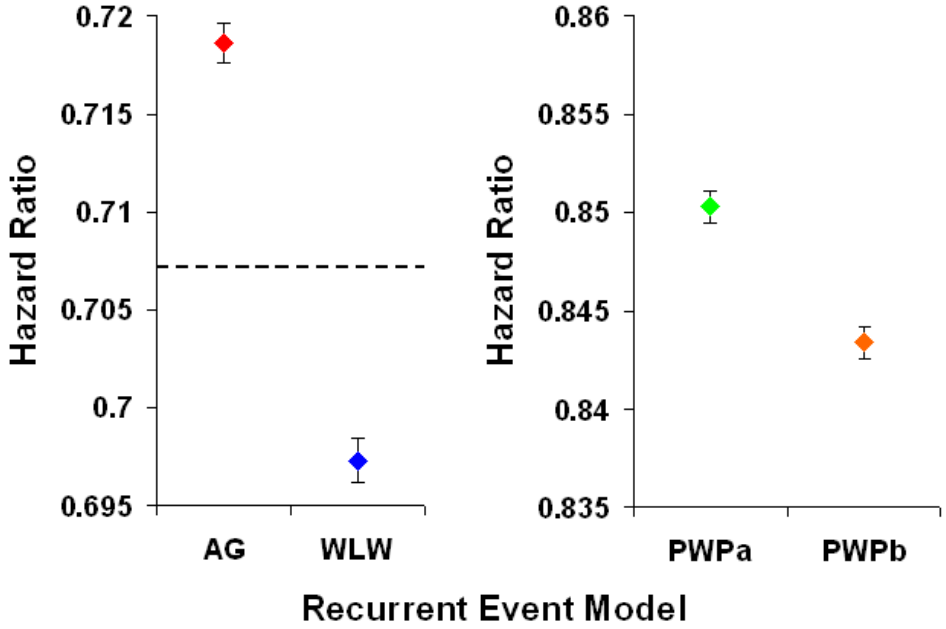


(a) The first set of simulation conditions

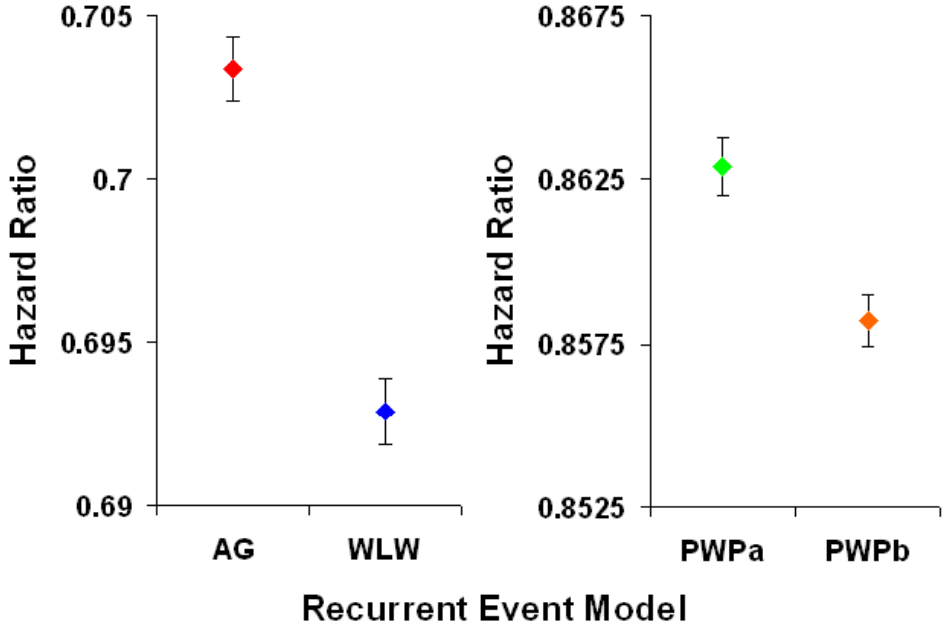


(b) The second set of simulation conditions

Figure 6.7: Hazard ratios for the first and second sets of simulation conditions for the treatment effect group of simulations where $\gamma = 0.75$ with accompanying 95% CIs



(a) The third set of simulation conditions



(b) The fourth set of simulation conditions

Figure 6.8: Hazard ratios for the third and fourth sets of simulation conditions for the treatment effect group of simulations where $\gamma = 0.75$ with accompanying 95% CIs

treatment effect. For the second set of simulation conditions when $\gamma = 1$ the PWPa and PWPb models gave the expected results. However, when $\gamma = 0.75$ they underestimated the treatment effect. For both values of γ the AG and WLW models overestimated the treatment effect. The results for the third and fourth sets of simulation condition were very similar for both values of γ. The results for the AG model were around the expected value and the WLW model slightly overestimated the treatment effect. The PWPa and PWPb models underestimated the expected treatment effect by around 15% for the third set of conditions and 16% for the fourth set of conditions. This percentage underestimation of the treatment effect by the PWPa and PWPb models was an absolute percentage difference as opposed to a relative percentage difference. The PWPa model underestimated the treatment effect marginally more than the PWPb model.

For the first and second sets of simulation conditions 100% of the replicates showed a significant treatment effect. This changed for the third and fourth sets of conditions. For the AG and WLW models there was no real change as the percentage of replicates that showed a significant treatment effect remained above 99%. This was not the case for the PWPa and PWPb models. For the third set of simulation conditions the percentage of replicates that showed a significant treatment effect remained in the nineties but for the fourth set of simulation conditions the percentage decreased further. The decrease in percentage was most apparent for the PWPa model where the percentage dropped to 85.53% when $\gamma = 1$ and 82.89% when $\gamma = 0.75$. No replicates were found to show a significant placebo effect.

6.3.3 The Treatment Effect for First Event Group of Simulations

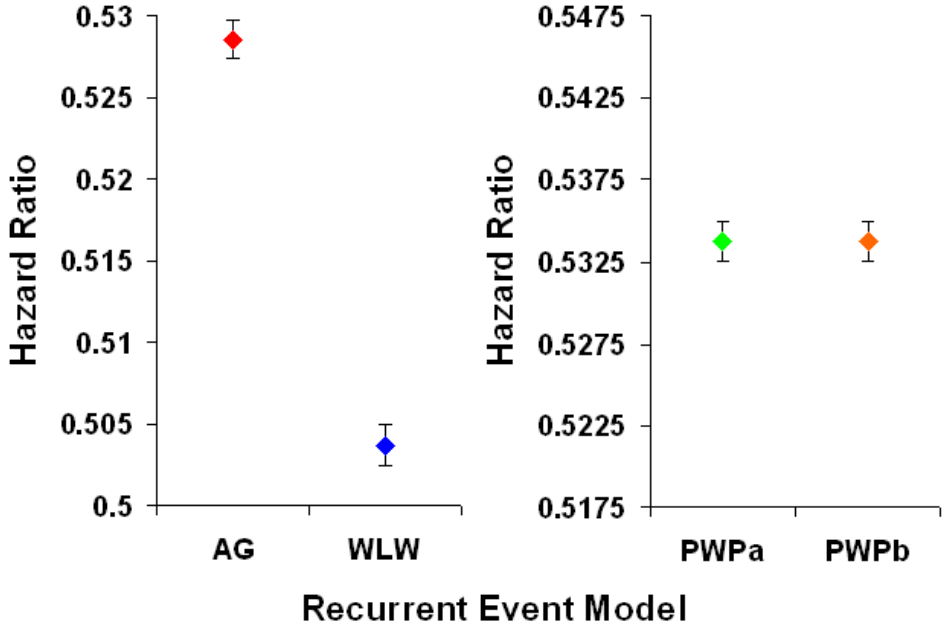
The results for the treatment effect for the first event group of simulations where the treatment was effective compared to the placebo for only the first event after which the treatment effect was set to zero are now shown. The simulation results for $\gamma = 1$ are shown in Table 6.7 and the results for the HRs only are shown in Figures 6.9 and 6.10. The simulation results for $\gamma = 0.75$ are shown in Table 6.8 and the results for the HRs only are shown in Figures 6.11 and 6.12.

For the first and second set of simulation conditions the WLW model gave results very close to the Cox model for the time-to-first event analysis, whereas the results for the AG model were slightly higher. The HRs for the PWPa and PWPb models were all higher than 0.5, which was the HR for the time-to-first event analysis. For the first set of simulation conditions the results were still relatively close to 0.5. However, for the second set of simulation conditions the results were pulled away from 0.5. The effect was more apparent for both sets of simulations conditions when $\gamma = 0.75$. A similar patter as seen in Section 6.3.2 for the third and fourth sets of simulation conditions emerged. The AG and WLW models gave similar results and the PWPa and PWPb models gave similar results. The results for the AG and WLW models were relatively close to the time-to-first event results. They were pulled towards the null value but still showed a large reduction in risk for the treatment group. However, this was not the case for the PWPa and PWPb models where apart from the PWPb model for the third set of simulation conditions the HRs were greater than the null value of 1.

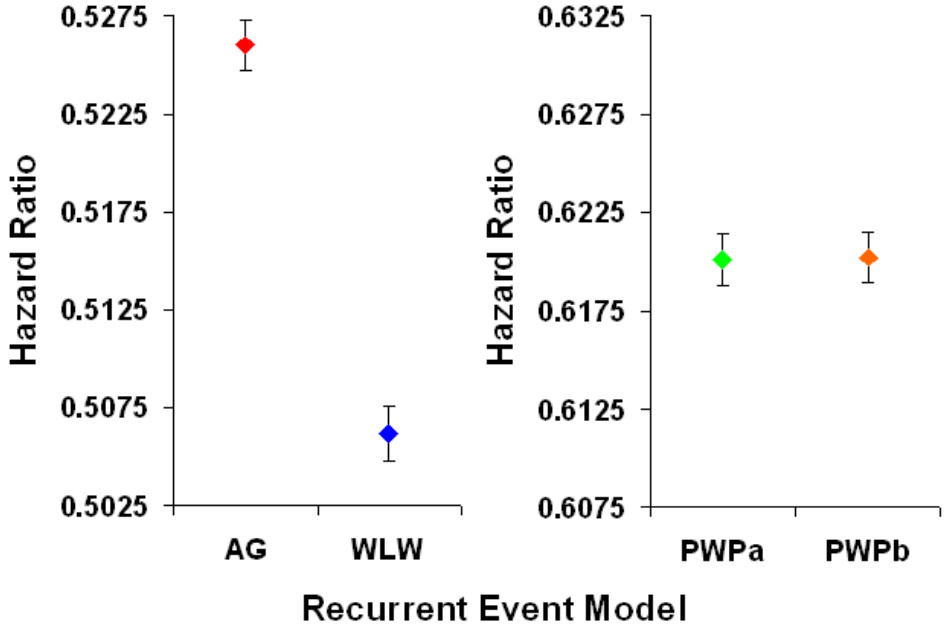
The percentage of replicates that showed a significant treatment effect for the

Table 6.7: Results for the treatment effect for the first event group of simulations where $\gamma = 1$. The table contains the mean of the estimated parameter estimates (SE), hazard ratios (SE) and the percentage of replicates that showed a significant effect for either treatment group

Model	Parameter Estimate (SE)	Hazard Ratio (SE)	Significant Replicates	
			Treatment (%)	Placebo (%)
The First Set of Simulation Conditions				
Cox	-0.6967 (0.0012)	0.5019 (0.0006)	100%	0.0%
AG	-0.6445 (0.0012)	0.5286 (0.0006)	100%	0.0%
WLW	-0.6937 (0.0013)	0.5037 (0.0006)	100%	0.0%
PWPa	-0.6349 (0.0012)	0.5338 (0.0006)	100%	0.0%
PWPb	-0.6348 (0.0012)	0.5338 (0.0006)	99.99%	0.0%
The Second Set of Simulation Conditions				
Cox	-0.6967 (0.0012)	0.5019 (0.0006)	100%	0.0%
AG	-0.6510 (0.0013)	0.5260 (0.0007)	99.94%	0.0%
WLW	-0.6902 (0.0014)	0.5062 (0.0007)	99.94%	0.0%
PWPa	-0.4838 (0.0011)	0.6201 (0.0007)	99.70%	0.0%
PWPb	-0.4835 (0.0011)	0.6202 (0.0007)	99.71%	0.0%
The Third Set of Simulation Conditions				
Cox	-0.3468 (0.0007)	0.7086 (0.0005)	99.90%	0.0%
AG	-0.1938 (0.0007)	0.8259 (0.0006)	77.79%	0.0%
WLW	-0.2299 (0.0008)	0.7969 (0.0006)	85.82%	0.0%
PWPa	0.0226 (0.0005)	1.0239 (0.0005)	0.72%	8.05%
PWPb	-0.0221 (0.0005)	0.9793 (0.0005)	6.08%	0.80%
The Fourth Set of Simulation Conditions				
Cox	-0.3468 (0.0007)	0.7086 (0.0005)	99.90%	0.0%
AG	-0.2535 (0.0007)	0.7780 (0.0005)	95.32%	0.0%
WLW	-0.2758 (0.0007)	0.7609 (0.0005)	96.68%	0.0%
PWPa	0.0916 (0.0005)	1.0975 (0.0006)	0.01%	43.13%
PWPb	0.0206 (0.0005)	1.0220 (0.0005)	0.90%	6.51%

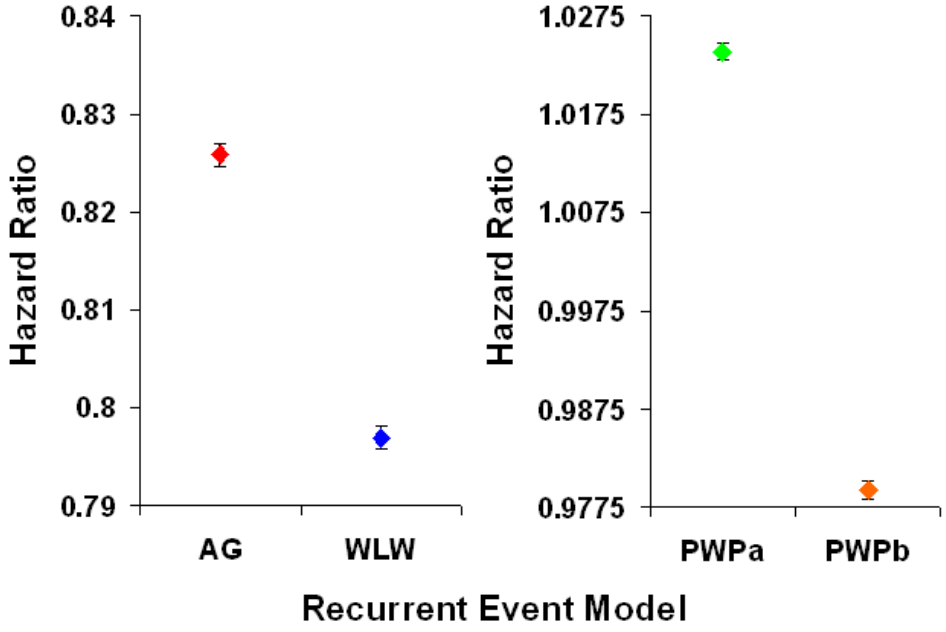


(a) The first set of simulation conditions

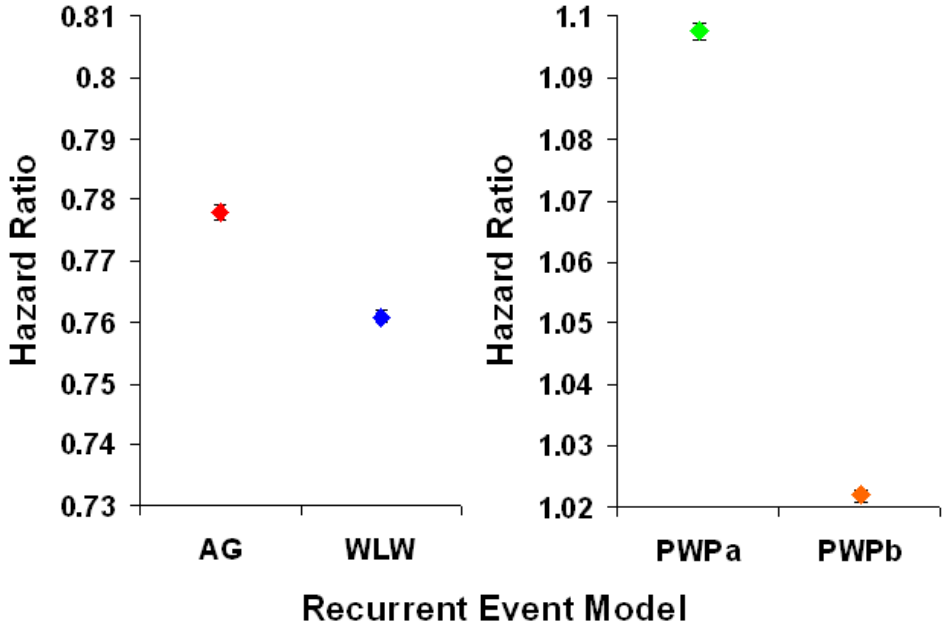


(b) The second set of simulation conditions

Figure 6.9: Hazard ratios for the first and second sets of simulation conditions for the treatment effect for the first event group of simulations where $\gamma = 1$ with accompanying 95% CIs



(a) The third set of simulation conditions

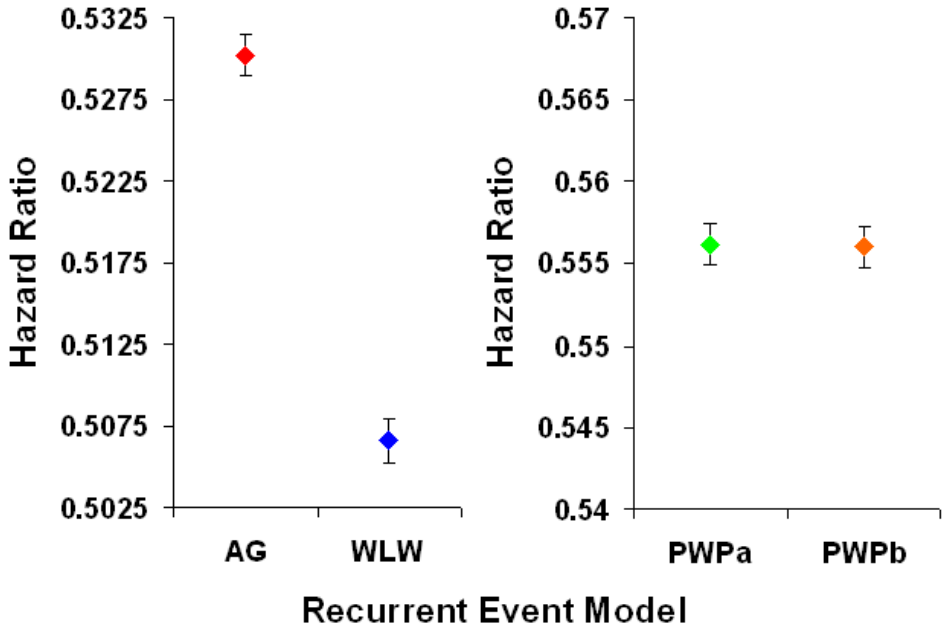


(b) The fourth set of simulation conditions

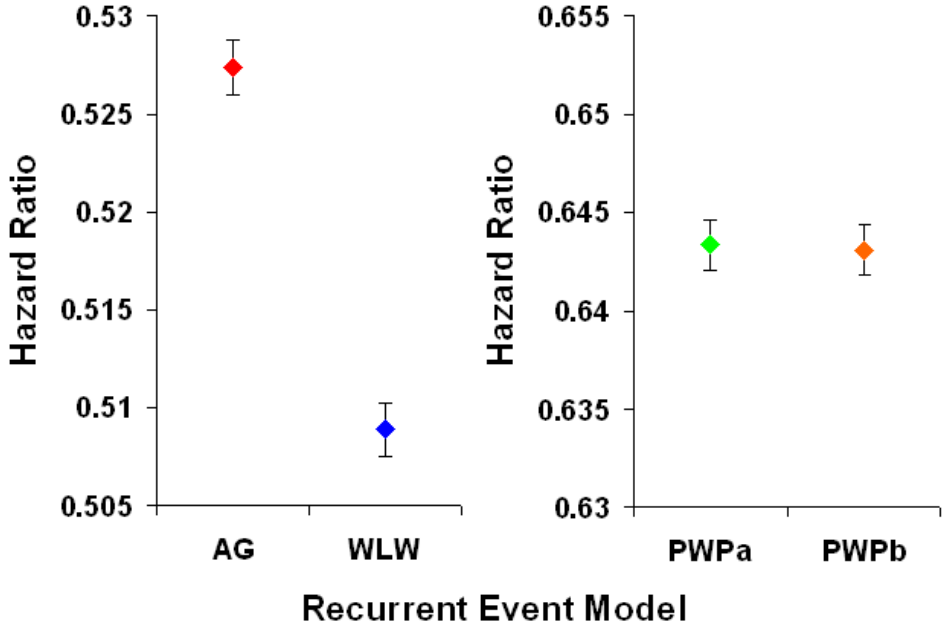
Figure 6.10: Hazard ratios for the third and fourth sets of simulation conditions for the treatment effect for the first event group of simulations where $\gamma = 1$ with accompanying 95% CIs

Table 6.8: Results for the treatment effect for the first event group of simulations where $\gamma = 0.75$. The table contains the mean of the estimated parameter estimates (SE), hazard ratios (SE) and the percentage of replicates that showed a significant effect for either treatment group

Model	Parameter Estimate (SE)	Hazard Ratio (SE)	Significant Replicates	
			Treatment (%)	Placebo (%)
The First Set of Simulation Conditions				
Cox	-0.6933 (0.0012)	0.5035 (0.0006)	100%	0.0%
AG	-0.6417 (0.0012)	0.5302 (0.0006)	100%	0.0%
WLW	-0.6881 (0.0013)	0.5066 (0.0006)	100%	0.0%
PWPa	-0.5934 (0.0012)	0.5562 (0.0006)	100%	0.0%
PWPb	-0.5936 (0.0012)	0.5560 (0.0006)	100%	0.0%
The Second Set of Simulation Conditions				
Cox	-0.6933 (0.0012)	0.5035 (0.0006)	100%	0.0%
AG	-0.6484 (0.0013)	0.5274 (0.0007)	99.93%	0.0%
WLW	-0.6849 (0.0014)	0.5089 (0.0007)	99.94%	0.0%
PWPa	-0.4466 (0.0011)	0.6434 (0.0007)	99.47%	0.0%
PWPb	-0.4469 (0.0010)	0.6431 (0.0007)	99.57%	0.0%
The Third Set of Simulation Conditions				
Cox	-0.3462 (0.0007)	0.7091 (0.0005)	99.88%	0.0%
AG	-0.1986 (0.0007)	0.8220 (0.0006)	79.72%	0.0%
WLW	-0.2354 (0.0008)	0.7927 (0.0006)	87.14%	0.0%
PWPa	0.0208 (0.0005)	1.0222 (0.0005)	0.86%	7.23%
PWPb	-0.0300 (0.0005)	0.9716 (0.0005)	8.99%	0.51%
The Fourth Set of Simulation Conditions				
Cox	-0.3462 (0.0007)	0.7091 (0.0005)	99.88%	0.0%
AG	-0.2529 (0.0007)	0.7785 (0.0005)	95.20%	0.0%
WLW	-0.2759 (0.0007)	0.7609 (0.0006)	96.87%	0.0%
PWPa	0.0879 (0.0006)	1.0936 (0.0006)	0.0%	37.20%
PWPb	0.0085 (0.0005)	1.0096 (0.0005)	1.50%	3.85%

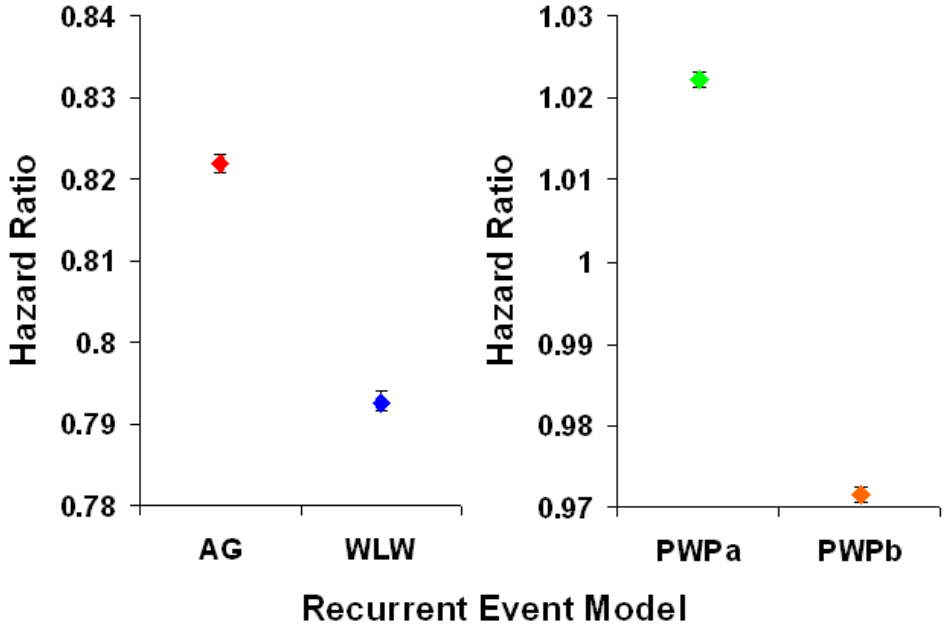


(a) The first set of simulation conditions

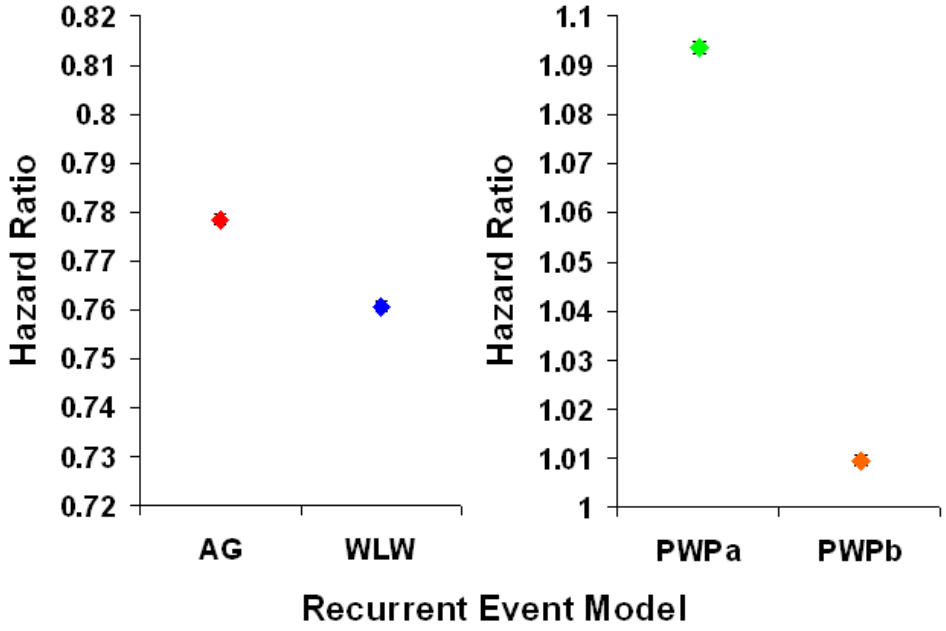


(b) The second set of simulation conditions

Figure 6.11: Hazard ratios for the first and second sets of simulation conditions for the treatment effect for the first event group of simulations where $\gamma = 0.75$ with accompanying 95% CIs



(a) The third set of simulation conditions



(b) The fourth set of simulation conditions

Figure 6.12: Hazard ratios for the third and fourth sets of simulation conditions for the treatment effect for the first event group of simulations where $\gamma = 0.75$ with accompanying 95% CIs

first and second sets of simulation conditions was greater than 99%. This was not the case for the third and fourth sets of simulation conditions. For the AG and WLW models the percentage of replicates that showed a significant treatment effect remained high, especially for the fourth set of simulation conditions. No replicates showed a significant placebo effective. For the PWPa model for the third set of conditions less than 1% of the replicates showed a significant treatment effect whereas 7% or 8%, depending on the value of γ, showed a significant placebo effect. For the PWPb model 6% or 9%, depending on the value of γ, of the replicates showed a significant treatment effect whereas less than 1% showed a significant placebo effect. For the fourth set of simulation conditions no replicates showed a significant treatment effect for the PWPa model. For the PWPb model around 1% of the replicates did show a significant treatment effect. For the PWPb model 4% or 7%, depending on the value of γ, of replicates showed a significant placebo effect but for the PWPa model this jumped to 37% or 43%, depending on the value of γ.

6.4 Interpretation of the Simulation Results

From the results for the no treatment effect group of simulations it was seen that the simulation process was unbiased. The results for the treatment effect and treatment effect for the first event groups of simulations can then be interpreted in the knowledge that there was no form of bias in the simulation process. When examining the results of the simulations the different underlying assumptions of the recurrent event models have to be taken into account. For the first set of simulation conditions very few patients suffered a third or fourth event. With

the increased risk of suffering events under the second set of simulation conditions the number of patients suffering later events did increase. The inclusion of the random subject for the third and fourth sets of simulation conditions dramatically increased the number of patients who suffered third and fourth events. The increased event rate caused by the random subject effect was far greater than purely increasing the risk of patients suffering events. For the third and fourth sets of simulation conditions the event rate was substantially higher than in the IONA Study. The second set of simulation conditions, when $\gamma = 0.75$, most accurately reflected the IONA Study data.

The results for the treatment effect group of simulations are considered first. For the AG model it is assumed that the risk of events remains constant over time and that the events are independent. The gap times were generated independently. However, given patients were followed for a maximum of two years the length of the first gap time would impact on the length of the later gap times. Therefore, the assumption of independence could be questioned. The assumption of constant risk was violated under the second and fourth sets of simulation conditions and to a lesser extent for all simulations where $\gamma = 0.75$. The ideal conditions for the AG model were the first set of simulation conditions. From the results in Section 6.3.2 it is seen that the AG model performed well under these conditions. Even with the inclusion of the random subject effect for the third set of simulation conditions the AG model still performed well. For the second set of simulation conditions where the risk of patients suffering events was increased the AG model overestimated the treatment effect. Under the fourth set of conditions this overestimation of the treatment effect was not seen. The event rate was so high for the third and fourth sets of simulation conditions that

increasing the risk of patients suffering events did not have such an impact on the simulation results as it did for the second set of simulations conditions.

The overall pattern of the results for the WLW model was that the treatment effect was overestimated. This was most obvious for the first and second sets of simulation conditions where the event rate was lower. For the WLW model all patients are required to have risk intervals for all events regardless of how many events they actually suffer. If a treatment is effective less patients who are receiving the treatment will suffer a first event within a given time period or if they do suffer an event it will be later. A long first time interval will impact when patients suffer a second event and the same is true for third and fourth events. In the WLW model if patients do not suffer the number of events being modelled for the last observed time interval for the patient is carried forward for the remaining events. In doing this the reduced risk of suffering events caused by the treatment was artificially replicated once, twice or even three times. This would then lead to the overestimation of the treatment effect. With the inclusion of the random subject effect the event rate was greatly increased and this led to less time intervals being carried forward and as a result the overestimation of the treatment effect was reduced. Even for the third and fourth sets of simulation conditions the results for the WLW model slightly overestimated the treatment effect. In general when the event rate was low the AG and WLW models overestimated the treatment effect. When the event rate was higher the results for the AG and WLW models were near the expected values.

The PWPa and PWPb models gave the expected results for the first and second sets of simulation conditions when $\gamma = 1$. When $\gamma = 0.75$ and as a result the event was slightly higher both models marginally underestimated the

treatment effect. The underestimation of the treatment effect was again seen in the results for the third and fourth sets of simulation conditions. For the third and fourth sets of conditions the PWPa and PWPb models estimated the treatment effect to be only half of what it should have been. Unlike the WLW model and to a lesser extent the AG model the PWPa and PWPb models performed better when the event rate was lower. These finding may be explained by selection bias for patients. As the treatment was effective those patients who suffered events in the treatment group had a higher underlying frailty than patients who did not suffer events. These patients were predisposed to suffer more events and suffer them earlier. Therefore for later events, even though the treatment was effective, these patients were at an elevated risk of suffering events compared to those in the placebo group. This resulted in the HRs being pulled towards the null value of 1. When the event rate was lower the selection bias had a smaller effect. With the inclusion of the random subject effect and the resultant increased event rate the bias was exaggerated.

The results for the treatment effect for the first event group of simulations are now considered. For the first set of simulation conditions, where the event rate was lowest, the WLW model gave almost identical results as the Cox model for the time-to-first event analysis whereas the other models estimated the treatment effect to be smaller. The WLW model again gave almost identical results to the time-to-first event analysis for the second set of simulation conditions. The fact that the treatment effect was set to zero after the first event was not reflected in the results for the WLW model. The results for the other models for the second set of simulation conditions showed a reduced treatment effect. The lessening of effect was more apparent for the PWPa and PWPb models. For the third

and fourth sets of simulation conditions where the event rates were considerably higher the AG and WLW models showed that the treatment did reduce the risk of patients suffering recurrent events. The treatment effect was near to the treatment effect seen in the treatment effect group of simulations. The AG model does not consider the order in which patients suffer events. For the first and second sets of simulation conditions the majority of events suffered were first events, where the treatment was effective. The greater occurrence of first events was reflected in the results for the AG model. For the third and fourth sets of simulation conditions the PWPa and PWPb models indicated that were was a small placebo effect, apart from the PWPb model for the third set of simulation conditions. In this case the results showed the treatment was effective but it only reduced the risk of patients suffering events by 2% or 3%. Here the selection bias of patients was having a larger effect than seen in the treatment effect group of simulations.

The patients in the treatment group who suffered first events had a high level of frailty as they suffered events even though the treatment was effective. For subsequent events these patients still had a high level of frailty but the treatment was no longer effective. The result of which was that for the second, third and fourth events, even though there was no difference between the treatment groups, it appeared that the placebo was effective at reducing the risk of patients suffering events. The results for the PWPa and PWPb models were heavily influenced by the recurrent events over the first events. This was reinforced by looking at the percentage of replicates that showed a significant difference between the treatment groups. For the third and fourth sets of simulation conditions the majority of replicates did not show a significant difference between the groups. Some of

the replicates showed a significant placebo effect. The percentage of replicates that showed a significant placebo effect was higher for the PWPa model than the PWPb model. For the fourth set of simulation conditions the percentage of replicates that showed a significant placebo effect was particularly high at 37% or 43%, depending on the value of γ, for the PWPa model.

The reason why this percentage was so high is not clear but may be as a result of the total time being used in the PWPa model as opposed to the gap time in the PWPb model. The selection bias for the treatment group that the most frail patients suffered first events was applicable to both the PWPa and PWPb models. There may however be an additional form of bias only applicable to the PWPa model. This second form of bias was highlighted under the fourth set of simulation conditions as this was the scenario were patients suffered the highest number of recurrent events. For the PWPa model the time patients enter the risk set for the second event and when they suffer a second event are affected by the time to the first event as the total time is used. For the PWPb model this is not the case as the time is reset to zero after an event. The most frail patients will suffer early first events and then early second events whereas the least frail patients will have longer times to both events. For patients who were at risk of a second event those patients who were in the treatment group would tend to have a higher level of frailty than patients in the placebo group. These patients would then suffer a second event earlier and this would be reflected in the PWPa model as the total time is used. This would then double up the bias in the PWPa model compared to the PWPb model where there was only the selection bias with respect to the first event. However, further simulations would be required to ascertain if this was the reason behind the high percentage of

replicates that showed a significant placebo effect for the PWPa model. These simulations have shown that the recurrent event models do perform differently under known conditions and that they can give measurably different results.

6.5 Summary of the Main Findings of the Simulations

The main findings of the simulations were as follows:

1. Under the ideal conditions for the AG model, the first set of simulation conditions, the model performed well.

2. When the risk of suffering events was not constant, the second set of simulation conditions, the AG model overestimated the treatment effect.

3. The overall pattern of the results for the WLW model was that the treatment effect was overestimated. This was most apparent for the first and second sets of simulation conditions.

4. In general when the event rate was low the AG and WLW models overestimated the treatment effect.

5. When the event rate was higher the results for the AG and WLW models were near the expected values.

6. The results for the AG and WLW models were influenced more by the first events than the recurrent events.

7. For the first and second sets of simulations conditions, where the event rate was low, the PWPa and PWPb models performed well.

8. For the third and fourth sets of conditions, where the event rate was highest, the PWPa and PWPb models estimated the treatment effect to be only half of what it should have been.

9. The results for the PWPa and PWPb models were heavily influenced by the recurrent events over the first events.

10. In the simulations that most accurately reflected what happened in the IONA Study for the primary endpoint, the second set of simulation conditions with $\gamma = 0.75$ for the treatment effect group of simulations, the models that performed best were the PWPa and PWPb models.

Chapter 7

Health Economics

The economic evaluation of the IONA Study was carried out by Walker et al. (2006). Before details of the analysis are given and expanded upon in Chapter 8 an introduction to the topic of health economics will be given. With the increasing costs of providing healthcare for the population and the increased patient awareness of the possible interventions and treatments available to them, decisions made about the potential use of any such interventions cannot be based solely on the results of clinical trials. The economic impact of the intervention has to be considered. As a result the area of health economics has recently seen its profile raised. A good example of this involved the breast cancer drug Herceptin and the debate surrounding whether it should be prescribed or not. (Dent and Clemons, 2005; Anon, 2005; Barrett et al., 2006) Due to the limited amount of money available it is not possible to satisfy everyone. If one treatment is recommended it could well have a direct impact on another treatment not being recommended, as if one is recommended there may well be insufficient money available to fund both. The undertaking of economic evaluations, alongside clinical trials and in

general, is therefore gaining increasing importance. In this Chapter the different types of economic evaluations that are used will be explained and leading on from that Incremental Cost-Effectiveness Ratios and Quality-Adjusted Life Years will be further discussed. The use of sensitivity analysis in economic evaluations will also be covered.

7.1 Economic Evaluations

A product, for example a treatment, that is judged to be cost-effective is one that based on the benefits it produces and all the associated costs is thought economical to use. An economic evaluation involves the comparison of two or more interventions in terms of both their costs and consequences. In health economics the interventions will commonly be drugs or devices and the consequences will be in terms of the clinical benefits they give. When an economic evaluation of a treatment is being carried out an important factor to consider is the overall net cost of that treatment. The net cost of a treatment includes all the cost savings that the treatment makes, for example through the reduction in the number of hospitalisations that patients who are receiving the treatment would have gone on to suffer and any other reductions in costs versus the cost of the actual treatment itself and any accompanying costs. These supplementary costs can include additional out-patient hospital visits and GP appointments for patients or the cost of extra laboratory tests that are required to be undertaken to monitor the patients and how they are responding to the treatment. There are five main types of economic evaluations that are used. They are Cost-Effectiveness Analysis (CEA), Cost-Utility Analysis (CUA), Cost-Minimisation Analysis (CMA),

Cost-Consequence Analysis (CCA) and Cost-Benefit Analysis (CBA). A short explanation of each will be given.

7.1.1 Cost-Effectiveness Analysis

In a Cost-Effectiveness Analysis (Lopert et al., 2003) the costs of two treatments are compared to each other as well as comparing the benefits in terms of one single clinical outcome. The costs are measured in monetary terms and the benefits are measured in natural units pertaining to the given situation, for example the number of deaths prevented or the number of symptom free days. As the costs and benefits are being measured in non-comparable units the comparison takes the form of a ratio known as an Incremental Cost-Effectiveness Ratio (ICER). For comparing two treatments, A and B, the following formula for an ICER would be used:

$$\begin{aligned} \text{ICER} &= \frac{(\text{Total Cost of Treatment A}) - (\text{Total Cost of Treatment B})}{(\text{Total Benefit of Treatment A}) - (\text{Total Benefit of Treatment B})} \\ &= \frac{\text{Difference in Costs Between Treatments A and B}}{\text{Difference in Benefits Between Treatments A and B}} \end{aligned} \tag{7.1}$$

In health economics ICER is a generic term with no standard definition and it can be and is used in a variety of situations whenever cost and benefits are being compared. The topic of ICERs will be returned to in Section 7.2. A limitation of CEA is that as natural units of clinical benefit are being compared the ability to assess the relative benefits of introducing different types of new treatments using this method of economic evaluation is limited to those that have the same natural units of clinical benefit.

7.1.2 Cost-Utility Analysis

In a Cost-Utility Analysis (Brinsmead and Hill, 2003), as in a CEA, the costs of two treatments are compared to each other as well as comparing the benefits in terms of one single clinical outcome. Costs are measured in monetary terms but this time the benefits are measured in a common outcome called a utility. When comparing the two treatments an ICER is again used. A frequently used utility in health economics is the Quality-Adjusted Life Year (QALY). Life-years are a familiar concept to epidemiologists and statisticians as a measure of changes in survival. However, they take no account of the quality of life of the patient, whereas QALYs do. When looking at QALYs states of health are graded on a scale of 0 to 1, with 0 being dead or the equivalent of being dead and 1 being in perfect health. In some cases the scale can have negatives values where patients are still alive but considered to be in a worse state than death, such as patients suffering advanced cancer pain or advanced degenerative disease. The QALY values assigned to different health states are known as health utilities. For example the utility for a person who is an insulin dependent diabetes mellitus sufferer is 0.84 and for someone who is on dialysis it is 0.41. How the utility values are assigned and the use of QALYs will be discussed in Section 7.3.

7.1.3 Cost-Minimisation Analysis

In a Cost-Minimisation Analysis (Newby and Hill, 2003) it is assumed that the two treatments being compared have been shown to offer equivalent clinical benefit. As a consequence unlike CEA and CUA the benefits of the two treatments cannot be compared. In a CMA analysis the only the thing that is compared is

the cost of the two treatments, with the treatment with the overall lower cost being favoured. It is important to remember that when the costs of two treatments are being compared it is not just the cost of the treatments themselves but also the associated costs that are being compared. For example, the cost of one treatment may be lower but patients may have to be closely monitored when they have been prescribed that treatment. The monitoring may include attending GP visits and having laboratory tested carried out. This will then impact on the overall cost of the treatment. The impact on patients also has to be considered. If one treatment involves patients taking a course of tablets whereas another involves them attending hospital outpatient clinics for the duration of their treatment the costs associated with the two treatments for patients will differ greatly. This type of evaluation is often used when a new treatment comes to market that has been shown to be as clinically effective as the current standard but is believed to be more cost-effective to use.

7.1.4 Cost-Consequence Analysis

In a Cost-Consequence Analysis (Mauskopf et al., 1998), as in a CEA and a CUA, the costs of two treatments are compared to each other as well as comparing the benefits. In a CCA the benefits can be presented as more than one outcome and are presented in a disaggregated manner and are also not combined with the costs in the form of a ratio. The different benefits are given in their natural units. When carrying out a CCA all the major costs for both treatments are set out as well as all the benefits and these can then be compared and if investigators wish they can apply their own weighting to the different types of benefits.

7.1.5 Cost-Benefit Analysis

In a Cost-Benefit Analysis (Sugden and Williams, 1979), as in a CCA, the benefits that treatments give to patients can be presented as more than one outcome but they are not measured in their natural units or in utilities but they are assigned monetary values as are the costs associated with the treatments. As a result the net benefit of a treatment can be calculated by comparing the total costs of the treatment with the total gain in monetary terms of the benefits associated with it. As both the cost and benefits are measured using the same units, be it pounds, dollars or euros, CBA can be used to compare treatments for different conditions using the net benefit principle. One way of assigning monetary values to health benefits is based on the idea of willingness to pay (WTP). In WTP people are asked how much money they would be willing to pay for a health benefit, such as controlling high blood pressure, and this value is then assigned to that benefit. As health benefits have to be expressed in monetary terms, even taking into consideration the use of WTP and other such techniques, the use of CBA can be problematic in health economics.

7.2 Incremental Cost-Effectiveness Ratios

When using Incremental Cost-Effectiveness Ratios the costs will almost always be measured in monetary values and the benefits can be measured in a number of different ways. For example, if a CEA is being carried out, natural units of measurement, such as symptom free days, can be used. Whereas if a CUA is being carried a utility such as QALYs will be used. Regardless of the unit of measurement being used the concept of the ICER is unchanged. Most commonly

an ICER involves the comparison of two treatments. There is no direct three-way ICER equivalent. However individual pairwise comparisons can be used. If one of the three treatments was a placebo or an existing treatment, comparisons could be made with it and the other two, so as to gauge which of the two new treatments is most cost-effective.

Given that differences are being looked at, within the ratio, one or both of the numerator and denominator could have a negatives value. From Equation 7.1 it can be seen that if the difference in costs is negative this would imply that the net cost of treatment A is lower than treatment B. If however the difference in benefits is negative this would imply that treatment B is more beneficial than treatment A. This is assuming that the benefits are measured in a currency that is desirable, such as lives saved or QALYs gained. If however the benefits are being measured in terms of undesirable events, such as the number of heart attacks avoided or in the case of the IONA Study the number of primary endpoints prevented, the converse would be true. If treatment A is a potentially new treatment then ideally the cost difference would be negative and the benefits difference would be positive.

As ICERs are ratios there is a potential drawback with them in that if at anytime the difference in costs is zero the ratio will then automatically be equal to zero regardless of what the difference in benefits is between the two treatments. In addition if the difference in benefits is zero or approaches zero, as would be the case if the benefits were shown to be equivalent in a clinical trial, then the ratio will be undefined or excessively large which will be problematic when it comes to interpretation. In such circumstances a CEA or a CUA, which both use ICERs in the evaluation, would not be used but a CMA would be carried out instead.

7.2.1 The Cost-Effectiveness Plane

A concept that aids the understanding of ICERs and also illustrates some of the problems with them is the cost-effectiveness plane. (Black, 1990) This is illustrated in Figure 7.1. The x-axis measures the effectiveness and the y-axis the costs.

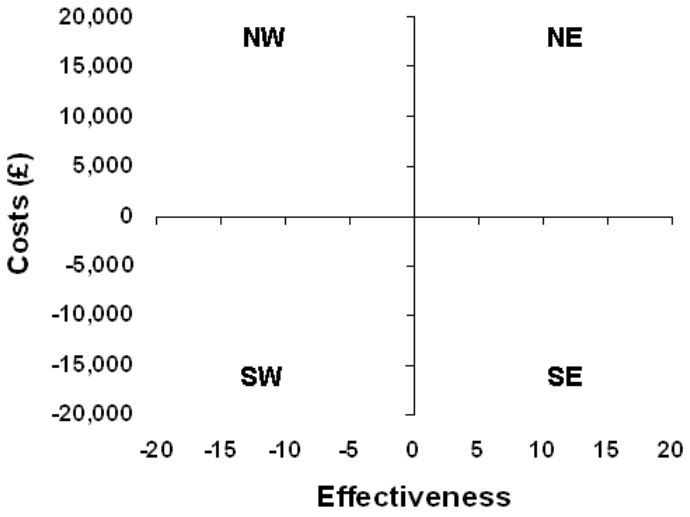


Figure 7.1: The cost-effectiveness plane

From Figure 7.1 it can be seen that the plane can be slit up into four quadrants, north-west (NW), north-east (NE), south-east (SE) and south-west (SW). Here let treatment A be the new treatment and treatment B be the current standard treatment. The origin is the point where both treatments A and B are

equally effective and have the same net cost.

1. North-West

   In the NW quadrant treatment A is totally dominated by treatment B in that it has a higher net cost and is less effective. Treatment B would be favoured over treatment A in all situations. In this quadrant an ICER has a negative value.

2. North-East

   In the NE quadrant there has to be a trade-off between treatment A and treatment B. Treatment A has a higher net cost than treatment B but it is also more effective. This is the most common situation. A judgement has to be about how much more effective does treatment A need to be than treatment B and whether the resultant increase in the overall net cost due to the use of treatment A is economically justifiable. In different situations WTP for increased effectiveness will alter. In this quadrant an ICER has a positive value.

3. South-East

   In the SE quadrant treatment A totally dominates treatment B in that it has a lower net cost and is more effective. Treatment A would be favoured over treatment B in all situations. In this quadrant an ICER has a negative value.

4. South-West

   In the SW quadrant there again has to be a trade-off between treatment A and treatment B. Treatment A has a lower net cost but is also less effective

that treatment B. The decision to be made in this situation is whether the saving in costs is sufficient to outweigh the decrease in effectiveness of treatment A over treatment B. This is the type of situation where a pharmaceutical company will campaign for its new product to be used over the current standard as it is overall cheaper to prescribe and only marginally less effective. How much less effective the new treatment is, is the key issue to be considered. In this quadrant an ICER has a positive value.

Looking at the four quadrants of Figure 7.1 illustrates a problem with ICERs in that if the ratio lies in both the NW or SE quadrant the ratio will have a negative value. If the only information quoted is the ICER value and no further information about the values of the numerator and the denominator are given then interpretation of negative ICERs could be problematic as the quadrants are the polar opposites of each other. The fact that ICERs that lie in both the NE and SW quadrants are positive should not be an issue as if the value does lie in the SW quadrant then treatment A would have both a higher net cost and be less effective that treatment B so an economic evaluation is unlikely to be carried out. There is a second problem with the interpretation of negative values of ICERs and this is illustrated through an example given in the paper by Briggs and Fenn (1998). Two treatments result in the same cost savings and both improve the health of patients, treatment A by 0.5 units and treatment B by 1 unit. As the denominator used in the calculation of the ICER for treatment A is a fraction the ICER will have a larger value in magnitude than the ICER for treatment B. As the ICERs are both negative this implies that treatment A is more cost-effective, but treatment B is more effective and therefore would be more cost-effective to

use. The magnitude of negative ICERs can be misleading. When ICERs are being used in practice it is unlikely that information on the constituent parts of both the numerator and denominator will not be available so any potential problems with interpretation of ICER values will be avoided.

7.3 Quality-Adjusted Life Years

The concept of the Quality-Adjusted Life Year was introduced in Section 7.1 when CUA was being discussed. The main issue with using QALYs is how the utility values are assigned to all the different possible health states.

7.3.1 Assigning Utility Values

There are a number of different methods used for assigning QALY utility values and three of the most widely used are as follows:

1. Time-Trade-Off

   Time-Trade-Off (TTO) (Torrance et al., 1972) is used to calculate QALYs by the following method. People are told that they will remain in a given state of health for ten years, for example, and they are then asked how many of those ten years would they give up to live in perfect health. For example if someone said they would give up four of the ten years in a particular health state to live in perfect health for six years then that health state would have a utility value of 0.6.

2. Standard Gamble

   The Standard Gamble method (von Neumann and Morgenstern, 1944; Torrance, 1976) assigns utilities to given health states as follows. People are given the choice of remaining in a state of health for a given period of time or the alternative of undergoing a treatment or procedure that will either return them to perfect health or it will kill them. Then, given the state of health that they are in, they are asked what probability or chance they would want the treatment to have of success before they would agree to undertake it. For example if a person was suffering from a common cold the probability of success would have to be 0.99 or more likely 1 before anyone would seriously consider undergoing the treatment. On the other hand if they were told that they would be in a coma with little chance of regaining consciousness the probability would most likely be considerably lower.

3. Health Questionnaires

   Health Questionnaires are used to assess the severity of a health state. An example is the EuroQol Group EQ-5D questionnaire, http://www.euroqol.org/. (Brooks, 1996) The EQ-5D questionnaire has five questions focusing on five different areas: mobility, pain and discomfort, self-care, anxiety and depression and usual activities. Each question has three possible answers corresponding to no problem, some problems and major problems. There are 243 possible combinations of answers to the five questions and with the inclusion of being unconscious or being dead they are 245 possible health states in total. Each of theses health states has been assigned a utility value

and when a patient fills in a questionnaire their combination of answers is matched to a look-up table of utilities. A second approach requires patients to grade their state of health on a visual analogue scale of 0 to 100, with 100 being in perfect help. Other variations of patient questionnaires are also available.

There are potential drawbacks with all three of these approaches. If the general public are asked to assess the seriousness of a health state they are likely to have had no direct experience with the condition and they may overestimate the seriousness of the condition. Alternatively patients who have lived with a condition for some time might have adjusted to living with it and as a result might underestimate the seriousness of the condition. A possible way to combat this is to ask doctors to assess the seriousness of the health state. They should have an objective view and have more experience relating to the condition than the general public on which to base their answers and at the same time they will not have become accustomed to living with the condition as patients will have. However, as patients are often likely to see doctors when their symptoms are at their worst doctors may in fact overestimate the seriousness of the given health state. Nurses are also asked and in some cases their answers are regarded as a truer reflection of the seriousness of a disease as they will spend more time with patients than doctors so may have a better insight into the health state and its implications for the patients. For each of the three methods a large sample of people or patients will be asked and the mean value of their answers, scaled to be from 0 to 1, will be taken as the utility value for that particular state of health.

7.3.2 Using Quality-Adjusted Life Years

Given that each state of health has been assigned a utility value the QALY approach can then be used to compare different medical interventions and treatments. For example if there are two treatments that can be used to treat the same medical condition their effectiveness can be measured and compared by looking at QALYs. This means that if treatment C prolongs life for 10 years and treatment D prolongs life for 5 years, both compared to no treatment, on the surface it would appear that treatment C has a clear advantage over treatment D. However, the quality of the life the patient will have is not being taken into consideration at this stage. If treatment D gives the patient a quality of life of 0.9, then treatment D would give the patient a QALY gain of 4.5 years. Whereas if treatment C gives the patient a quality of life of 0.45, then treatment C would also give the patient a QALY gain of 4.5 years. So when the quality of life is considered as well as the quantity of life it appears that both treatments C and D will be equally effective. It would then be down to whether the quality of life or the quantity of life is more important.

In real life situations and when faced with a limited budget the quality and quantity of life gains are not the only things that have to be considered but also the net cost of any such gains has to be taken into consideration and this is done by looking at the net cost per QALY gained. For example let treatment E have a QALY gain of 10 and treatment F have a QALY gain of 5, both compared to standard current best practice. The treatments can be used to treat the same or different conditions. In terms of quality of life treatment E is the better of the two and should be prescribed. This would be true if there was an infinite

amount of money available but this is not the case so economic factors have to be considered as well. If the net cost of prescribing treatment E was £200,000 per patient then the net cost per QALY would be £20,000 per patient and if the net cost of prescribing treatment F was £25,000 per patient then the net cost per QALY would be £5,000 per patient. In this example there is a difference of cost per QALY of £15,000. Treatment E is superior in terms of QALY gains but treatment F is superior in terms of cost per QALY. If the aim was to treat the maximum number of patients then treating patients with treatment F, at a lower net cost per QALY, would benefit more patients and increase the number of QALYs gained from treatment F than treating a few patients with treatment E, at a higher net cost per QALY. It also has to be considered which conditions should be treated if the treatments that treat them have the same net cost per QALY. This is looking at the situation from the viewpoint of the NHS. For individual patients the situation may well be different. For those patients able to afford to pay for private healthcare they would want the treatment that is most effective for their condition and not the most cost-effective treatment. The NHS cannot afford to give each patient the best possible treatment regardless of cost. The NHS has to look at the population of the UK as a whole and not at patients on an individual basis.

The QALY is a useful tool to health economists as it takes into account both the quality of life of a patient as well as the quantity of that life. Any changes in survival of patients as well as any change in the quality of that survival are both considered in the calculation of QALYs but there are other relevant factors that are not measured. If patients feel like they have no autonomy or involvement in their own treatment or decisions being made about their treatment this may

diminish their quality of life. These changes in the quality of life of patients cannot be measured using a questionnaire, such as the EQ-5D questionnaire, but they will be important to the patients themselves. The QALY is an important measure and useful decision making tool especially when the net cost per QALY gained can be calculated and used but other factors also need to be taken into consideration.

7.3.3 Quality-Adjusted Life Years Used in Practice

Recommendations about which new treatments the NHS should advocate for use are made by the Scottish Medicines Consortium (SMC) in Scotland and by the National Institute for Health and Clinical Excellence (NICE) in England and Wales. The net cost per QALY is just one factor that is looked at when a treatment is put forward for use and other things have to be taken into consideration. By looking at the net cost per QALY the cost-effectiveness of the treatment can be judged, but what net cost per QALY is too much? If the NHS budget was infinite this would not be a problem but it is not infinite, so serious consideration has to be taken in the allocation of budgets so that the best use of the money available is made. There is no published fixed cut-off point for the net cost per QALY after which any treatment will not be recommended for use purely on economic grounds, although NICE have issued guidelines on the matter. (NICE, 2004)

> Below a most plausible ICER of £20,000/QALY, judgements about the acceptability of a new technology as an effective use of NHS resources are based primarily on the cost-effectiveness estimate. Above

> a most plausible ICER of £20,000/QALY, judgements about the acceptability of the new technology as an effective use of NHS resources are more likely to make more explicit reference to factors including:
>
> 1. the degree of uncertainty surrounding the calculation of ICERs
> 2. the innovative nature of the technology
> 3. the particular features of the condition and population receiving the technology
> 4. where appropriate, the wider societal costs and benefits.
>
> Above an ICER of £30,000/QALY, the case for supporting the technology on these factors has to be increasingly strong. The reasoning for the Committee's decision will be explained, with reference to the factors that have been taken into account, in the 'Considerations' section of the guidance.

From this statement it would appear that if a treatment has a net cost per QALY of below £20,000 it will be approved, between £20,000 and £30,000 it is likely to be approved as long as there is convincing evidence of its worth and above £30,000 it is unlikely that it will be approved unless the evidence of its benefit is extremely strong. These boundaries, if indeed they are actually the boundaries used, are not set in stone and exceptions to the rules can and will be made. A prime example of this is when an Orphan Medicinal Product (OMP) comes under consideration. (Dear et al., 2006) Due to the nature of an OMP the cut-off points would not be practical so each case has to be judged on its own merits. There is still some debate even then about the highest cost per QALY value that should

be accepted. Money used to fund OMPs to treat a few patients could be used to treat a large number of patients suffering from a more common form of illness. Whether money should be used to treat the few instead of the many is one of the balancing acts that bodies such as the SMC and NICE face.

7.4 Sensitivity Analysis

In health economics and economics in general sensitivity analysis is a useful device as costs of goods and services will not remain constant over time. An example would be when a drug goes off patent. When off patent the manufacturing of a drug is no longer restricted solely to the company that actually devised it. Other pharmaceutical companies are free to produce generic copies of the drug usually at a fraction of the price of the original. Additionally, any parameters that are used in a model will have uncertainty surrounding their values. For example it is unknown whether the results of a clinical trial will be transferable to day-to-day clinical practice and whether utility values that are specified by a small number of doctors are actually valid. Changes in any of these parameters could affect the findings of an economic evaluation. If changing costs and parameter estimates can be incorporated into a cost-effectiveness analysis they can be used to judge how the cost-effectiveness of a treatment will change over time.

When a sensitivity analysis is carried out the cost of one resource use is altered by a fixed quantity for each separate analysis. This type of analysis is an example of a univariate sensitivity analysis. Usually when a univariate sensitivity analysis is being carried out the cost of a number of resource uses can be altered. This could lead to many univariate sensitivity analyses being carried out. The results

of the individual univariate sensitivity analyses can then be presented together in a tornado diagram. (Howard, 1988) The tornado diagram allows the investigator to easily identify the resource uses that have the largest influence on the cost-effectiveness of the treatment.

However, no consideration is given to the distribution of the cost values that are being altered in a univariate sensitivity analysis and the level of increase or decrease in costs is arbitrarily set by the investigator. In some cases a number of quantities are altered by fixed levels and this is then a multi-way sensitivity analysis. Both univariate and multi-way sensitivity analysis have limitations in that only one or a small number of quantities can be altered at a time and only by a fixed level. As the number of quantities that are altered increases the interpretation of the results can become complicated and the analysis itself is computationally time consuming. (Claxton et al., 2005) These issues can be overcome by using Probabilistic Sensitivity Analysis (PSA).

7.4.1 Probabilistic Sensitivity Analysis

Probabilistic Sensitivity Analysis is an alternative to deterministic sensitivity analysis. Instead of the parameters within the cost-effectiveness model being altered individually by fixed amounts, full distributions for the parameters are specified. There are a large number of distributions that can be used but in most cases the choice of which to use will be governed by the parameter itself and what it is measuring. Once the distributions for the model parameters have been decided upon the parameters for the distributions themselves have to be considered. The probability distributions can be derived directly from the data for

the clinical trial. Alternatively the probability distributions could be generated by bootstrap sampling. (Pasta et al., 1999) A Bayesian approach can also be taken to specifying the distributions so that prior beliefs of clinicians and economists can be reflected in the analysis. It is not always possible or sensible to specify whole distributions for some parameters such as drug costs, as these are fixed and unlikely to change in the short term. With the choice of probability distributions having been made, Monte Carlo simulation can be used to investigate how altering the different parameters simultaneously impacts on the cost-effectiveness of the treatment. (Briggs et al., 2002b) In addition this will also allow any uncertainties in other variables, including the costs of services or assigned utility values, to be reflected in the analysis. The results of the PSA can then be presented by the means of a Cost-Effectiveness Acceptability Curve (CEAC).

7.4.2 Cost-Effectiveness Acceptability Curves

The concept of the Cost-Effectiveness Acceptability Curve was first proposed as a possible alternative to calculating confidence intervals for ICERs as a method of assessing the uncertainty surrounding them. (van Hout et al., 1994) The interpretation of CIs for ICERs can be problematic. In the calculations of CEACs the costs and effects per patient are considered to be outcomes from random variables. (O'Brien et al., 1994) Using the same notation as in Section 8.3.3 the cost-effectiveness plane, labelled as the $\Delta C/\Delta E$ plane, is a graphical representation of the joint probability density of ΔC and ΔE. Of interest is the probability that the cost-effectiveness ratio, C/E-ratio, is above or below R, where R is a pre-specified value that would imply that the treatment would be acceptable in

terms of cost-effectiveness. Above R and the treatment would not be recommended on the grounds of cost-effectiveness and below R it would be. The value of R is the altered. The relationship between R and the C/E-ratio is defined by a 'C/E-acceptability curve'. The probability that a treatment will be cost-effective over a range of values of R can then be seen by looking at the curve: the higher the probability for that value of R the more likely that the treatment will be accepted as being cost-effective. A CEAC is shown in Figure 7.2. This is

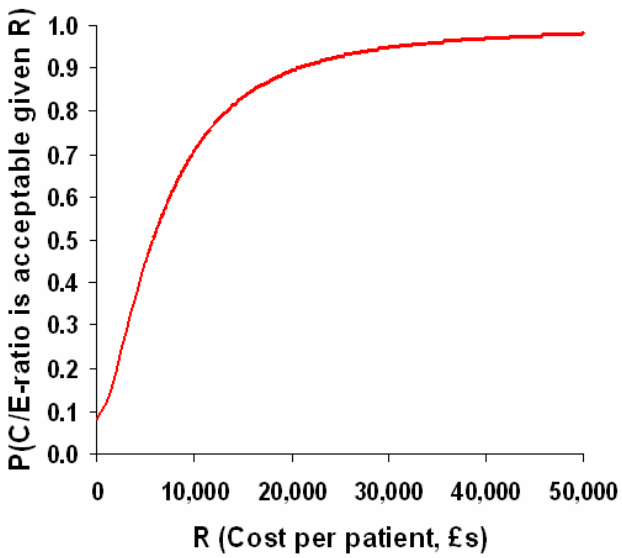


Figure 7.2: An example of a Cost-Effectiveness Acceptability Curve

only an illustrative example and does not imply that all CEACs are of this form. (Fenwick et al., 2004) The structure of CEACs depend on which quadrants of the cost-effectiveness plane the joint density of $(\Delta C, \Delta E)$ lies in. Where the CEAC

cuts the y-axis is the point where the WTP for the increased effectiveness of one treatment over another treatment is zero, the intercept does not automatically occur at the origin. The use of PSA is becoming more widespread in economic evaluations and NICE recently recommended that PSA should be included in submissions to them. (NICE, 2004) Sensitivity analysis is an aid to let health economists see how changing costs will affect the cost-effectiveness of treatments in the future.

Chapter 8

Health Economics of the IONA Study

In addition to the clinical benefits of nicorandil the economic implications of implementing treatment with nicorandil into clinical practice were of great interest when the IONA Study was devised. In this Chapter the rationale behind the economic evaluation will be described. Following on from this the economic evaluation will be expanded upon.

8.1 The Economic Evaluation of the IONA Study

From Section 7.1 it can be seen that there are a number of different types of economic evaluations that can be undertaken. As an active treatment, nicorandil, and a placebo were being compared a CMA could not be undertaken. There was one clear clinical outcome, the number of primary endpoints suffered, so the advantages of using a CCA would not be fully utilised. In a CBA monetary values

would have had to be assigned to preventing death due to CHD, non-fatal MIs and unplanned hospital admissions for cardiac chest pain. This would have been possible but due to the different natures of the component parts of the primary endpoint and the differing severities of non-fatal MIs and unplanned hospital admissions for cardiac chest pain that occurred as well as the difficulties of having to assigning monetary values to clinical outcomes a CBA was not performed. A CEA or a CUA could have been carried out but as there were no utility values, specifically QALYs, readily available a CEA was used. A literature review or a meta-analysis could have been carried out to assign utility values to the component parts of the primary endpoint so that a CUA could have been carried out but this was not attempted by Walker et al. (2006). Using a CEA the net cost per primary endpoint prevented by treatment with nicorandil was calculated and used as the basis for the analysis, with the results being expressed in the form of an ICER.

8.2 Costs Associated with the IONA Study

The costs for the nicorandil and placebo groups had a number of constituent parts, which were: the cost of nicorandil, the cost of cardiovascular and cerebrovascular hospitalisations and the cost of hospital admissions due to adverse events. Costs that were thought to be similar for both treatment groups, such as the cost of background cardiovascular medications patients were being prescribed, would have minimal impact on the cost-effectiveness calculation so were not included. (Walker et al., 2006) The costs were calculated as follows:

8.2.1 The Cost of Nicorandil

There were two parts that made up the cost of nicorandil, the cost of the actual nicorandil tablets themselves and the cost of dispensing them and the administration that went along with that. The cost of one 10 mg tablet of nicorandil was 13.6 pence and the cost of one 20 mg tablet was 25.9 pence. To the cost of each prescription a 10% dispensing fee was added. For the initiation of treatment and the up-titration of nicorandil the cost of two GP visits were included for each patient. The cost per GP visit was £19. The costs associated with nicorandil are shown in Table 8.1. The total cost of nicorandil for each patient was then calculated as the number of 10 mg and 20 mg packs of nicorandil they received plus the cost of the starter pack and the accompanying administration costs.

Table 8.1: The cost of nicorandil

	Cost
Initial starter pack	£28.12
10 mg pack	£27.23
20 mg pack	£51.85
GP visit	£19.00

8.2.2 The Cost of Hospitalisations

There were two types of hospital admissions considered, firstly those for cardiovascular and cerebrovascular reasons and secondly those due to serious adverse events related to treatment with nicorandil. The GI hospital admissions were considered such events by the study investigators. For the cardiovascular and cerebrovascular admissions there were a number of different types of hospital wards that patients were admitted to and these had differing cost associated

with them. (Netten et al., 2001; Information and Statistics Divison, NHS Scotland, 2002) In the IONA Study patients were recruited from throughout the UK. As a result UK costs were used in the analysis. The costs for the different types of ward are shown in Table 8.2. As the hospitalisation records did not distinguish between whether patients had been admitted to intensive care or coronary care units an average of the two costs was used. Patients who were admitted for GI problems were either admitted to an intensive care unit or a general medical ward. For each hospitalisation the costs of two hospital outpatient and two GP visits were also included as follow-up costs after discharge. As records of the number of hospital outpatient and GP visits that patients attended were not recorded the number and resulting costs of these visits were arbitrarily included by Walker et al. (2006) for each hospitalisation patients suffered. These costs are for the year ending 31^{st} March 2002.

Table 8.2: Hospital bed-day costs

Type of Hospital Bed	Cost
General medical ward	£242
Specialist cardiology	£429
Cardiac surgery ward	£666
Intensive care unit	£1,323
Coronary care unit	£610
Intensive/Coronary care unit	£967
Day Cases	
General medicine	£365
Cardiothoracic surgery	£622
Hospital outpatient visit	£72
GP visit	£19

In order to calculate the total hospital costs for both the nicorandil and placebo groups the number and length of hospital admissions for patients who

took part in the study was recorded and this information is shown in Table 8.3. The details of ward admissions and the specific reasons why patients were admitted to hospital are known but not shown here, see Walker et al. (2006). The hospitalisation costs were then calculated as the number of days each patient spent in the different types of hospital ward multiplied by the bed-day cost of that particular ward. The costs of any procedures or surgical operations that were performed during the hospitalisations were not considered.

Table 8.3: Numbers of admissions to hospital and number of days spent in hospital by patients

	Nicorandil (n = 2,565)	Placebo (n = 2,561)
Cardiovascular and Cerebrovascular Admissions		
Number of patients hospitalised	609	683
Number of hospital admissions	968	1,132
Days spent in hopsital	5,230	6,154
Gastrointestinal Admissions		
Number of patients hospitalised	157	108
Number of hospital admissions	186	129
Days spent in hopsital	934	571

8.2.3 Total Costs for the IONA Study

In the paper by Walker et al. (2006) the results of a primary and secondary analysis were both reported and when the separate ICERs were calculated different total costs for the two treatment groups were used. During the study when patients suffered GI events and resultant hospitalisations the investigator recorded whether they thought the study treatment had caused the GI event. Whether the study investigator thought that the study treatment had caused the GI event was coded as follows:

1. Certain
2. Probable
3. Possible
4. Unrelated
5. Unassessable

In the total costs for the primary analysis only the cost of GI hospitalisations directly attributed to the study treatment that patients were receiving by the study investigator were included. This corresponded to the GI events that had their cause coded as certainly, probably and possibly related to the study treatment patients were receiving. Additionally all the aftercare costs, no matter the reason for the hospitalisation, were not included in the total costs used in the primary analysis. In other studies it has been shown that although aftercare costs will be different between treatment groups, the total costs and the difference between the costs of the two treatment groups will likely be small in comparison between the actual hospitalisation costs themselves. (UK Prospective Diabetes Study Group, 1998) The main costs attributed to both treatment groups were due to hospitalisation of patients. Using the information contained in Tables 8.1, 8.2 and 8.3 the total costs for both the nicorandil and placebo groups, that were used in the primary analysis, can be calculated and these are shown in Table 8.4.

In the secondary analysis the costs of all the GI hospitalisations that patients suffered during the study and not just those attributed to the study medication by the study investigator were included in the total costs. The cost of two follow-up hospital outpatient and two GP visits were also included. Again using the

Table 8.4: Costs associated with the IONA Study - Primary analysis

Costs (£)	Nicorandil	Placebo
Nicorandil	596,304	—
Cardiovascular and Cerebrovascular admissions	3,031,671	3,656,733
Gastrointestinal admissions	44,286	15,843
Total Cost	£3,672,261	£3,672,576

information contained in Tables 8.1, 8.2 and 8.3 the total costs for both the nicorandil and placebo groups, that were used in the secondary analysis, can be calculated and these are shown in Table 8.5.

Table 8.5: Costs associated with the IONA Study - Secondary analysis

Costs (£)	Nicorandil	Placebo
Nicorandil	596,304	—
Cardiovascular and Cerebrovascular admissions	3,031,671	3,656,733
Cardiovascular and Cerebrovascular follow-up	168,168	198,016
Gastrointestinal admissions	236,838	153,316
Gastrointestinal follow-up	32,578	22,932
Total Cost	£4,065,559	£4,030,997

8.3 The Cost-Effectiveness Analysis of the IONA Study

Having all the cost information available, see Tables 8.4 and 8.5, and knowing the number of first primary endpoints that occurred in both groups, 337 (13.1%) in the nicorandil group and 398 (15.5%) in the placebo group, see Table 2.5, the

CEA was then carried out and the ICERs for both the primary and secondary analysis were calculated. In both the primary and secondary analyses only the first primary endpoints patients suffered were considered. In the case of the IONA Study the difference in benefits was measured in terms of an undesirable event, namely the difference in the number of primary endpoints that patients in the nicorandil and placebo groups suffered. As a result the total number of primary endpoints suffered were assigned negative values, with the higher the value numerically the more favourable it was. Of interest was the net cost per primary endpoint prevented that treatment with nicorandil caused.

8.3.1 ICER Calculations

The calculation of the ICER for the primary economic analysis of the IONA Study is shown in Equation 8.1. It can be seen that the Primary IONA ICER was equal to -£5, implying that for every primary endpoint treatment with nicorandil prevented it resulted in a net cost saving of £5. Here the cost saving in terms of reducing the number of hospitalisations more than offsets the cost of nicorandil. This was a very favourable finding for nicorandil in terms of cost-effectiveness.

$$\begin{aligned} \text{Primary IONA ICER} &= \frac{\text{Difference in costs between the two groups}}{\text{Difference in number of primary endpoints suffered}} \\ &= \frac{(3,672,261 - 3,672,576)}{(-337 - (-398))} \\ &= -£5 \end{aligned} \tag{8.1}$$

The calculation of the ICER for the secondary economic analysis of the IONA Study was equivalent to the calculation of the Primary IONA ICER. The calculation of the ICER is shown in Equation 8.2.

$$\begin{aligned} \text{Secondary IONA ICER} &= \frac{(4,065,559 - 4,030,997)}{(-337 - (-398))} \\ &= £567 \end{aligned} \tag{8.2}$$

The Secondary IONA ICER was equal to £567. The net cost of treatment with nicorandil was £567 for every primary endpoint prevented. As there was a higher GI event rate for patients treated with nicorandil when the secondary analysis was carried out, which included the costs of all the GI hospitalisations that patients suffered as well as the costs of follow-up care, treatment with nicorandil no longer resulted in a net cost saving per primary endpoint prevented. However, the overall cost-effectiveness of nicorandil still remained good.

8.3.2 ICER Interpretation

From Equations 8.1 and 8.2 it can be seen that the net cost of nicorandil preventing one primary endpoint in the IONA Study was -£5 for the primary analysis and £567 for the secondary analysis. When compared to other new treatments that were evaluated in a similar fashion these figure are very favourable. (Meads et al., 2000; McDonagh et al., 2000; Robinson et al., 2002) The net saving of £5 for every CHD event that treatment with nicorandil prevented was clearly very good but even the figure of £567 seems a small price to pay for preventing a CHD event. However, other facts have to be considered. Firstly, the IONA Study recruited patients that were perceived to be at a high risk of suffering events. As

a result the reduction in risk of suffering events that nicorandil caused was magnified and any comparison to the population as a whole cannot be made directly. Allowances for the reduced risk in the general population have to be inferred. This may well reduce the overall cost-effectiveness of treatment with nicorandil in the population treated in routine cardiology clinics. It could therefore be said that the external validity, or generalisability, of the results may well be poor. Secondly, as the data comes from a clinical trial the follow-up of patients was for a finite time and it was therefore unknown what would happen to these patients in the future and how this may effect the cost-effectiveness of nicorandil. This is true for the vast majority of economic evaluations that are derived from clinical trials. Any modelling of costs and outcomes beyond the end of a clinical trial can be problematic (Raftery, 1999) and was not attempted in this case as it was felt that the economic evaluation of the IONA Study had already achieved its objectives of showing treatment with nicorandil to be cost-effective. Thirdly, due to the composite nature of the primary endpoint of the IONA Study the interpretation of both ICERs may be awkward in that one event prevented could be CHD death whereas another could be an unplanned admission to hospital for cardiac chest pain. To prevent a death as opposed to an unplanned admission to hospital would be of more benefit but this difference was not taken into consideration.

The recording of whether the GI events patients suffered was due to the study treatment they were receiving was subjectively done by the study investigator and the reliability of this assignment cannot be guaranteed. It was therefore thought more appropriate to include the costs of all the GI events in the CEA. After each hospitalisation patients suffered they will have likely received some form of after care. The follow-up care each patient received would likely depend on the

reason for the hospitalisation and vary between individual patients. The precise care that patients did receive was not known. Applying an arbitrary fixed cost for the follow-up care received after each hospitalisation patients experienced seems the most appropriate method of reflecting the cost of follow-up care in the CEA. As a result the ICER for the secondary analysis was more conservative than for the primary analysis, implying it may be more useful in gauging the true cost-effectiveness of treatment with nicorandil. The costs associated with the secondary analysis ICER will therefore be used as the basis for extending the economic evaluation.

In the primary and secondary IONA ICERs calculated in Section 8.3.1 only the first primary endpoints patients suffered were included in the calculations. In total patients in the nicorandil group suffered 467 primary endpoints and in the placebo group patients suffered 604 primary endpoints, see Table 5.1. Treatment with nicorandil prevented an additional 76 events when all the CHD events patients suffered were considered. The resultant decrease in hospitalisation costs were most likely already factored into the analysis as all the hospitalisations cost for cardiovascular and cerebrovascular reasons were included in the costs used in the calculations of the ICERs. If the risk of recurrent CHD events and not just first events were reduced by treatment with nicorandil, as shown in Section 5.3.1, this would then have further implications on the cost-effectiveness, with treatment with nicorandil likely to be even more cost-effective. If the recurrent CHD events were included in the calculation of the Secondary IONA ICER then the net cost of treatment with nicorandil became £252 per CHD prevented, as shown in Equation 8.3. Compared to the Secondary IONA ICER, shown in Equation 8.2,

the cost of preventing a CHD event was more than halved.

$$\begin{aligned} \text{Adjusted Secondary IONA ICER} &= \frac{(4,065,559 - 4,030,997))}{(-467 - (-604))} \\ &= £252 \end{aligned} \tag{8.3}$$

It is still important to remember that the patients who were enrolled in the IONA Study were at a higher risk of suffering CHD events than the general patient population of CHD sufferers. This will have an affect on the cost-effectiveness of nicorandil. The cost-effectiveness will likely be reduced but as it was favourable at a net cost of £252 per CHD event prevented that even with a decrease in the cost-effectiveness nicorandil would still remain cost-effective to use in the treatment of patients with CHD.

8.3.3 Estimation of 95% Confidence Intervals for the ICERs

The ICERs calculated for both the primary and secondary CEA were informative but they were only point-estimates and the uncertainty surrounding them was unknown. To help ascertain the uncertainty 95% confidence intervals will be estimated. This was done by two different methods, one parametric and one non-parametric method. (Briggs et al., 2002a) Firstly, a parametric method was used and this involved using Fieller's Theorem (Fieller, 1954) to calculate the limits for the CI. The assumption that the difference in effects and costs between treatment groups follows a joint Normal distribution was made. Fieller's Theorem was used as the CIs were being estimated for a ratio. The limits for the CI were

estimated using the following formula:

$$\frac{[\Delta E\Delta C - z^2_{\alpha/2}cov(\Delta E,\Delta C)] \pm \sqrt{[\Delta E\Delta C - z^2_{\alpha/2}cov(\Delta E,\Delta C)]^2 - [\Delta E^2 - z^2_{\alpha/2}var(\Delta E)][\Delta C^2 - z^2_{\alpha/2}var(\Delta C)]}}{\Delta E^2 - z^2_{\alpha/2}var(\Delta E)} \tag{8.4}$$

where E represents the effects of the treatments and C represents the costs of the treatments. The limits are the roots of the following quadratic equation:

$$R^2[\Delta E^2 - z^2_{\alpha/2}var(\Delta E)] - 2R[\Delta E\Delta C - z^2_{\alpha/2}cov(\Delta E, \Delta C)] + [\Delta C^2 - z^2_{\alpha/2}var(\Delta C)] = 0 \tag{8.5}$$

where $R = \Delta C/\Delta E$.

Secondly, a non-parametric method was used and this was by means of bootstrap sampling. (Efron and Tibshirani, 1993) The percentile bootstrap method was used to estimate the limits of the CI. Patients were sampled with replacement from the treatment groups, 2,565 patients from the nicorandil group and 2,561 patients from the placebo group in each sample, and the ICER for this sample of patients was then calculated. The bootstrap sampling procedure was run 10,000 times and the samples were then ranked according to the calculated ICER value. The lower limit of the 95% CI was taken as the 251^{st} ICER value and the upper limit as the $9{,}750^{th}$ ICER value in the ranking. Both methods for estimating CIs were used to produce 95% CIs for both the primary and secondary ICERs of the IONA Study. The intervals along with the point estimates for the ICERs are shown in Table 8.6.

The CIs estimated via the two methods were broadly similar. The bootstrap

Table 8.6: The estimated 95% CIs for the primary and secondary ICERs of the IONA Study

Economic Analysis	ICER Point Estimate (£)	Method of Estimating the 95% Confidence Intervals	
		Fieller's Theorem (£)	Bootstrap Sampling (£)
Primary	-5	-11,116 – 30,736	-11,822 – 25,210
Secondary	567	-10,507 – 34,491	-11,071 – 29,746
Adjusted Analyses to Included the Recurrent Events Patients Suffered			
Primary	-2	-4,419 – 6,747	-4,546 – 9,190
Secondary	252	-4,290 – 7,372	-4,288 – 10,891

method gave narrower intervals for both analyses. In the published NICE guidelines (NICE, 2004) it states that that above an ICER value of £30,000/QALY there has to be very strong evidence to support recommendation of a treatment for use. In the IONA Study QALYs were not used but instead the cost per CHD prevented was considered. Using Fieller's Theorem the estimated upper limit of the 95% CI for the primary analysis was marginally above the £30,000 threshold and for the secondary analysis the upper limit was £34,491. Even taking this into consideration the cost-effectiveness of treatment with nicorandil was still favourable. For the bootstrap sampling intervals both the upper limits of the CIs were below the £30,000 threshold. Both sets of intervals for the primary and secondary analyses are illustrated in Figure 8.1.

Also shown in Table 8.6 are the ICERs and estimated 95% CIs for both the adjusted primary and secondary economic analyses. In the adjusted analysis the recurrent events that patients suffered were considered and not just the first CHD events. As when the recurrent events were considered treatment with nicorandil prevented more CHD events than treatment with placebo the difference in the

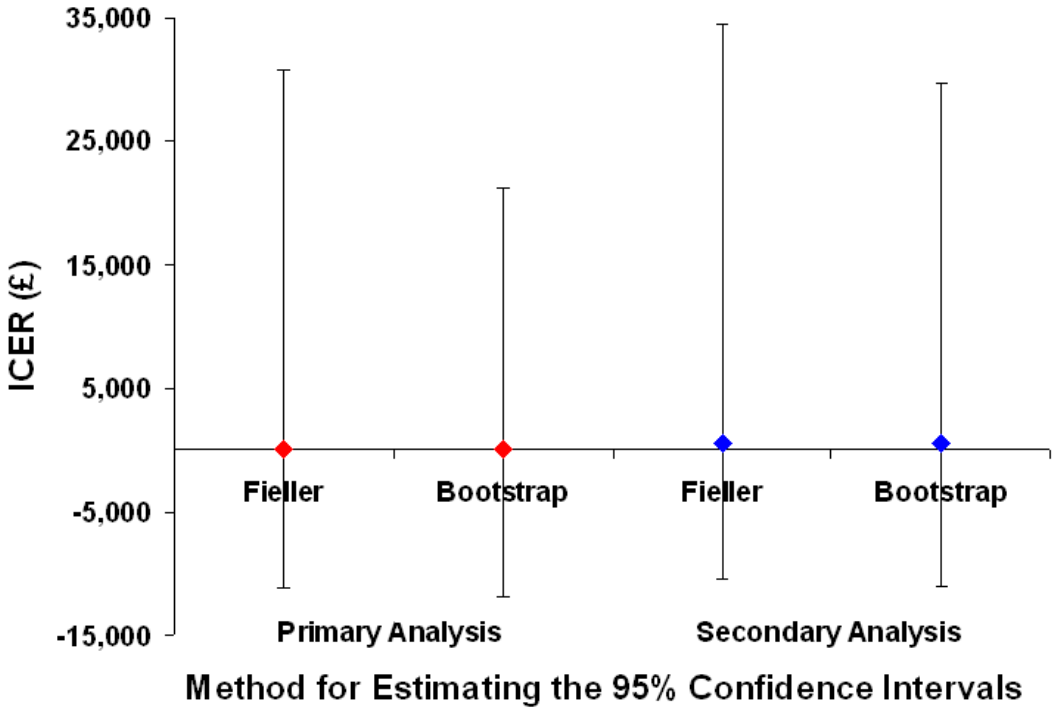


Figure 8.1: Graphical representation of the estimated 95% CIs for the primary and secondary ICERs of the IONA Study

effectiveness of the two treatments was greater. The estimated 95% CIs were narrower as a result. The cost-effectiveness of treatment with nicorandil was shown to be favourable for both analyses and by both methods used to estimate the 95% CIs. The problem of interpretation of negative ICERs discussed in Section 7.2 is illustrated by comparing the original primary analysis with the adjusted primary analysis. The ICER for the primary analysis showed that for every CHD prevented there was a net cost saving of £5. In the adjusted primary analysis, in which the treatment effect was larger, the net cost saving per CHD prevented was £2. If the actual values used in the calculation of the ICER are considered

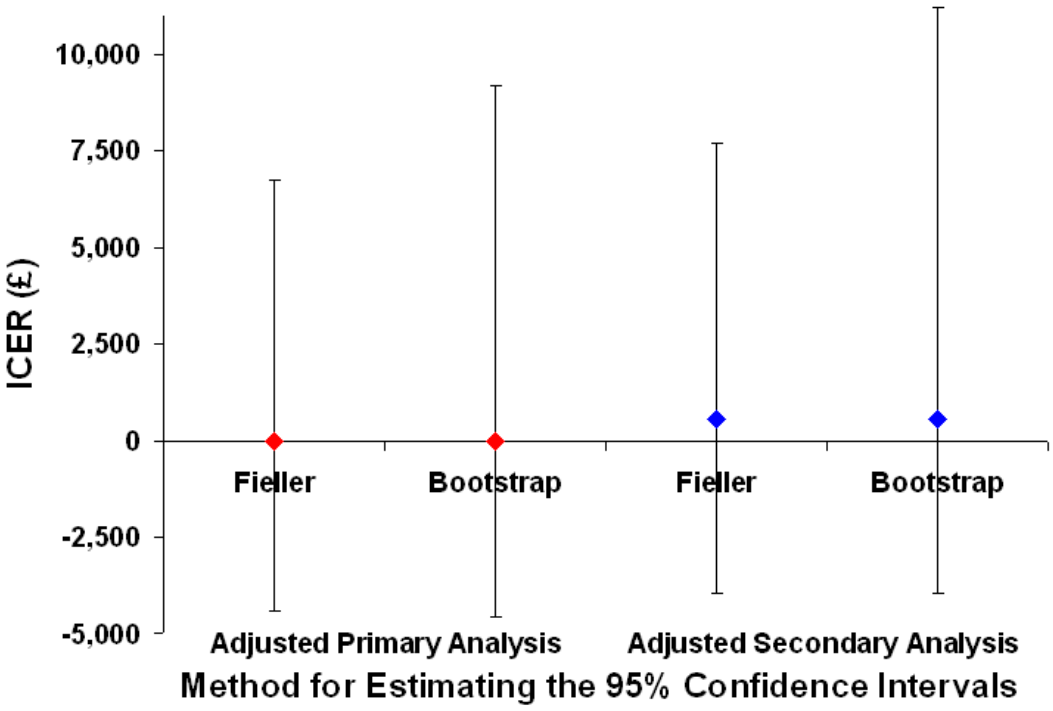


Figure 8.2: Graphical representation of the estimated 95% confidence intervals for the adjusted primary and secondary ICERs of the IONA Study

the adjusted primary analysis should indicate treatment with nicorandil to be more cost-effective but due to the difference in costs being negative this is not the case. The adjusted ICERs along with the estimated 95% CIs are illustrated graphically in Figure 8.2.

In the IONA Study there was a significant difference in the effectiveness of treatment with nicorandil over treatment with placebo implying that the denominator used in the calculation of the primary and secondary ICERs in Equations 8.1 and 8.2 was strictly greater than zero. When this is not the situation there is a potential drawback to using Fieller's Theorem to estimate CIs. In order

to estimate a CI using Fieller's Theorem a quadratic equation has to be solved, Equation 8.5, and doing this gives the roots shown in Equation 8.4. For the roots there are three situations that can occur and these are given in the paper by O'Hagan et al. (2000):

1. The solution may be a proper finite interval, of the form $[a, b]$.
2. The solution may be the complement of an interval, and so consist of 2 distinct infinite intervals of the form $(\text{-}\infty, a] \cup [b, \infty)$.
3. The solution may be the whole line $(\text{-}\infty, \infty)$.

In certain situations the estimates of the CIs are infinite or only exclude a small interval of values. (Heitjan et al., 1999) Using such intervals for inference can therefore be problematic. The benefit of using a parametric method, which will be more powerful if the assumption of a joint Normal distribution is valid, over a non-parametric method like bootstrap sampling will be lost. When there is a significant difference in the effects of the treatments the use of Fieller's Theorem is favoured over bootstrap sampling as it will give the same results each time unlike the bootstrap where random sampling is used and the results generated will alter each time. (Briggs et al., 2002a) The computation required to implement Fieller's Theorem is also significantly less than that required to implement bootstrap sampling. (Severens et al., 1999) In Sections 8.5 and 8.7 where the data set will be split according to the risk of patients suffering CHD events and then the risk of suffering GI and CHD events the sample sizes in each group will be reduced and in most cases there will be a non-significant difference in the effectiveness of treatment with nicorandil over treatment with placebo. Here the use of bootstrap sampling to estimate the 95% CIs for the ICERs will be favoured.

8.4 Sensitivity Analysis of the IONA Study

In the paper of the economic evaluation of IONA Study by Walker et al. (2006) sensitivity analysis was carried out on the result of the primary economic analysis. The sensitivity analysis that was carried out was univariate sensitivity analysis and this was done by firstly increasing the costs of angioplasty due to the increased use of stents by three different monetary values, £100, £200 and £500, and secondly by increasing and decreasing the costs of hospital bed-days by 20%. When the sensitivity to the costs of hospital bed-days was being investigated the costs were inflated and deflated separately for three of the types of hospital wards that patients were admitted to. The three types of ward that had their cost altered were specialist cardiology, cardiac surgery and the intensive care unit and coronary care unit combined group.

From the results of the sensitivity analysis undertaken by Walker et al. (2006), shown in Table 8.7, it was seen that by increasing the cost of an angioplasty procedure, due to the increased use of stents, only altered the cost-effectiveness of nicorandil marginally. The reported reason behind this was the fact that the number of angioplasties that were carried out during the study were similar in the two treatment groups with there only being a difference of eighteen between the groups, 98 (3.8%) in the nicorandil group and 116 (4.5%) in the placebo group. The changing of the costs of hospital bed-days had a larger effect on the cost-effectiveness of nicorandil than the increasing costs of angioplasty procedures. The results were least sensitive to changes in the cost of bed-days on specialist cardiology wards followed by the combined intensive care unit and coronary care unit group and most sensitive to the changes in the bed-day cost in cardiac

surgery wards. Even with the changes in the costs of goods and services within the sensitivity analyses the overall cost-effectiveness of treatment with nicorandil remained favourable.

Table 8.7: Results of the sensitivity analysis performed on the primary cost-effective analysis of the IONA Study

	ICER Value (£)
Primary Analysis	-5
Increase angioplasty cost by £100	-35
Increase angioplasty cost by £200	-64
Increase angioplasty cost by £500	-153
Reduce specialist cardiology bed day cost by 20%	342
Increase specialist cardiology bed day cost by 20%	-353
Reduce cardiac surgery bed day cost by 20%	1,051
Increase cardiac surgery bed day cost by 20%	-1,061
Reduce intensive/coronary care unit bed day cost by 20%	606
Increase intensive/coronary care unit bed day cost by 20%	-617

If the sensitivity analysis that was performed on the results of the economic evaluation of the IONA Study was to be further expanded from the simple univariate sensitivity analyses that were carried out then a multi-way sensitivity analysis could have been performed. If a multi-way sensitivity analysis was to be performed, using the univariate sensitivity analysis that was carried by Walker et al. (2006) as an example, then the costs of angioplasty procedures and hospital bed-days would be altered simultaneously in one analysis as opposed to individually in separate analyses, as was the case when the univariate sensitivity analysis was performed.

8.5 Coronary Heart Disease Risk Levels

The net cost of nicorandil per primary endpoint prevented, as shown by Equations 8.1 and 8.2, indicates that nicorandil was cost-effective when treating the patient population of the IONA Study as a whole. Even taking into consideration the likely smaller reduction in risk of suffering a CHD event that the general population would experience treatment with nicorandil would still appear to be cost-effective. In other circumstances and with other treatments the cost-effectiveness will not be so clear cut. If the risk of patients suffering events could be calculated the patient population could then be split into different levels of risk. Treating only those patients in the higher risk sub-groups of the population may improve the cost-effectiveness of a treatment. The key factor to consider is how the risk of suffering events should be calculated and then using this risk how the population should be split into sub-groups. Using the IONA Study data it was investigated whether treating only the higher risk patients would improve the cost-effectiveness of treatment with nicorandil. (Henderson et al., 2005)

Firstly, any such potential sub-groups of the population had to be identified. In order to do this the risk of a patient suffering a CHD event had to be calculated and this was done using the multivariable predictive model found for the primary endpoint in Section 3.4.1. It should be noted that the multivariable predictive model consisted of variables recorded on patients at baseline. Let R_{CHD} equal the risk of a patient suffering a CHD event. Then R_{CHD} was calculated as follows:

$$R_{CHD} = exp(\underline{x}_i\underline{\hat{\beta}}) \tag{8.6}$$

where $\underline{x}_i$ is the $1 \times p$ vector of prognostic variables for patient i, where p is the

number of variables included in the multivariable predictive model, and $\underline{\hat{\beta}}$ is the $p \times 1$ vector of parameter estimates found from the Cox model for the primary endpoint.

It was of interest to identify sub-groups of the patient population in which the cost-effectiveness of treatment with nicorandil would be improved. As a result when the risk of patients suffering CHD events was calculated all patients were treated as if they were in the placebo group of the study. In addition 67 patients, 41 (1.6%) in the nicorandil group and 26 (1.0%) in the placebo group, could not have their risk of suffering a CHD event calculated due to these patients having missing values for one or more of the baseline variables used to calculate R_{CHD}.

For the remaining 5,059, using Equation 8.6, the risk of them suffering a CHD event was then calculated. The patients were ranked according to their calculated risk. According to this ranking the patients were then split into three groups of approximately equal size: low, medium and high risk of suffering a CHD event. It should be noted that patients who were recruited to the IONA Study were at a high risk of suffering CHD events. Therefore, all three risk groups, including the group defined as the low risk group, would have a higher underlying risk of suffering a CHD event than the general population. As a result when the risk groups are being referred to it is in the context of the IONA Study and not the general population. The grouping of patients into these three groups was based solely on the ranking of their risk and did not take into account which treatment group the patients were in. As a result the balance in numbers of patients between the treatment groups due to randomisation was lost. This resulted in imbalances of patient numbers between the treatment groups for the three levels of CHD risk. Calculating the net cost per event prevented, via an ICER, directly for the

risk groups using the total number of events suffered and the total costs for the groups would likely have be misleading due to the random imbalances in patient numbers between the treatment groups. This would have resulted in an inequality between the number of patients who were at risk of suffering a CHD event and more particularly in the total costs for the nicorandil and placebo groups within the three risk groups. Shown in Table 8.8 are the proportion of patients who suffered a first CHD event and the mean cost per patients for the nicorandil and placebo groups separately for the three levels of risk. As indicated in Section 8.3.2 the total costs used were the ones used in the secondary analysis.

Table 8.8: The proportion of patients who suffered a primary endpoint and the mean cost per patient in the nicorandil and placebo groups for the three levels of CHD risk

Low CHD Risk Group	Nicorandil (n = 821)	Placebo (n = 866)
Proportion of patients who suffered a CHD event	0.07186	0.09238
Mean cost per patient (£)	1,161	1,094
Medium CHD Risk Group	**Nicorandil (n = 852)**	**Placebo (n = 834)**
Proportion of patients who suffered a CHD event	0.1069	0.1389
Mean cost per patient (£)	1,367	1,415
High CHD Risk Group	**Nicorandil (n = 852)**	**Placebo (n = 834)**
Proportion of patients who suffered a CHD event	0.2136	0.2386
Mean cost per patient (£)	2,186	2,233

As expected, starting from the low risk group and moving through the medium risk group to finally the high risk group the proportion of patients who suffered a CHD event increased. In order to calculate the ICERs for the three levels of

CHD risk the following modified version of Equation 7.1 was used:

$$\text{ICER} = \frac{\text{Difference in Proportion of Patients who Suffered a CHD Event}}{\text{Difference in Mean Cost per Patient Between the Treatment Groups}} \tag{8.7}$$

The ICERs for three levels of CHD risk as well as for the overall IONA Study, the secondary analysis, with accompanying estimated 95% CIs can be seen in Table 8.9.

Table 8.9: The cost per CHD event prevented for the three levels of CHD risk as well as for the overall IONA Study with accompanying estimated 95% CIs

Level of CHD Risk	ICER Point Estimate (£)	Method of Estimating the 95% Confidence Intervals	
		Fieller's Theorem (£)	Bootstrap Sampling (£)
Overall	567	-10,507 – 34,491	-11,071 – 29,746
Low	3,255	$-\infty$ – ∞	-73,622 – 81,960
Medium	-1,497	-20,056 – 219,461	-22,111 – 34,488
High	-1,885	$-\infty$ – ∞	-88,195 – 77,403

The ICER for the low CHD risk patients was £3,255, higher than the ICER for the overall data set. Treatment with nicorandil would still reduce the risk of patients in the low risk group from suffering CHD events but not as cost-effectively as treating only those higher risk patients. Both the ICERs for the medium and high risk groups had negative values at -£1,497 and -£1,885 indicating for these two groups treatment with nicorandil resulted in a net cost saving per CHD event prevented. The cost of nicorandil was more than off-set by the reduction in the number of CHD events and resultant hospitalisation costs that these events would have caused. The ICER for the high risk CHD group was numerically lower than for the medium risk CHD group but only marginally.

Looking at the 95% CIs it can be seen that using Fieller's Theorem the intervals for the low and high CHD risk groups were the entire real line. This means that there were no real roots to the solution of Equation 8.5 and that there was no significant difference in the effectiveness of treatment with nicorandil over treatment with placebo. For the medium CHD risk group there were real roots for Equation 8.5 and there was a significant difference in the effectiveness of treatment with nicorandil over treatment with placebo. This difference in effectiveness was only marginal and that resulted in the 95% CI estimated by Fieller's Theorem being so wide. The data set was split up and as a result was not powered to show a significant difference between the treatments in the three levels of CHD risk so these results were not unexpected hence why in this situation the use of bootstrap sampling over Fieller's Theorem is recommend. The ICERs with estimated 95% CIs, found by bootstrap sampling, are illustrated graphically in Figure 8.3. The bootstrap sampling intervals for the low and high CHD risk groups were wide so no meaningful inference can be made. The significance difference in the effectiveness of the treatment for the medium CHD risk group was reflected in that the estimated 95% CI was much narrower than for either the low or high CHD risk group. Comparing the medium CHD risk interval to that for the overall data set indicated that treating only patients who were at a medium risk of suffering a CHD event would improve the cost-effectiveness of nicorandil. Implying the same would be true for only treating those patients who were at a high risk of suffering a CHD. It is likely that treating the medium and high CHD risk groups with nicorandil would result in a net cost saving to the NHS.

In the case of nicorandil and the IONA Study the level of CHD risk for patients was calculated using the model for the primary endpoint found in Section 3.4.1

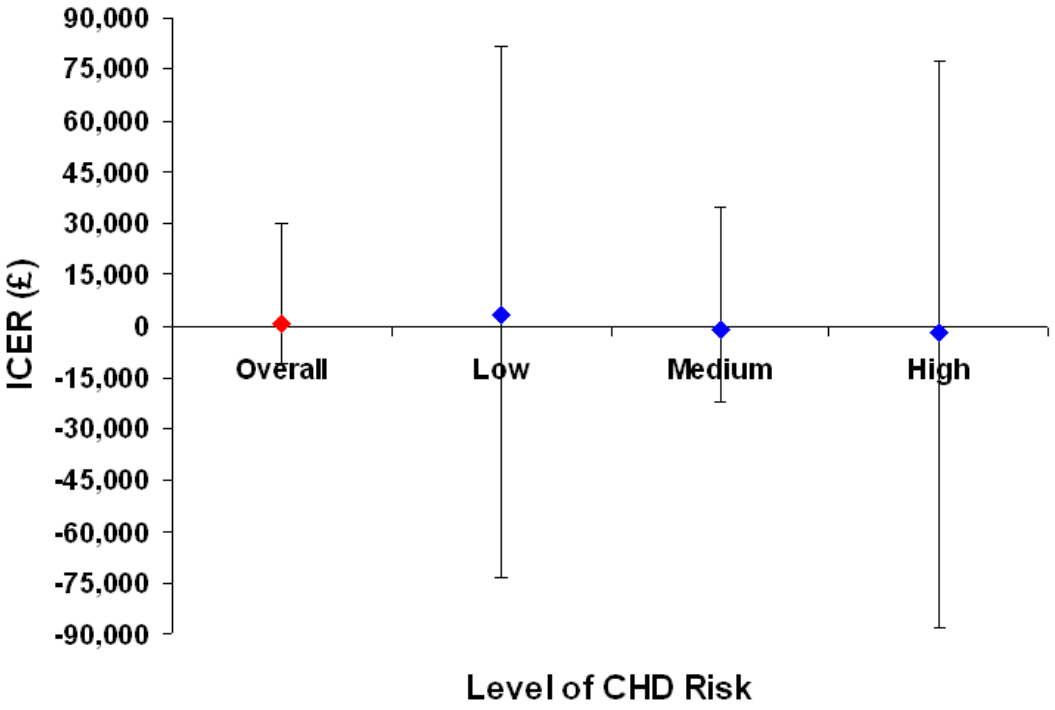


Figure 8.3: The cost per CHD event prevented for the three levels of CHD risk as well as for the overall IONA Study with accompanying 95% CIs estimated by bootstrap sampling

and Equation 8.6. If this model was to be used on the general population of cardiovascular patients it may need to be revised as at present it is only based on a sub-set of the population, namely the patients who took part in the IONA Study, and it may not accurately model the general population. In order to make the risk model more generalisable instead of using Equation 8.6 to calculate the risk of patients suffering CHD events an externally validated CHD risk model could have been used. One such example would have been to use Framingham risk equations (Anderson et al., 1991; D'Agostino et al., 2000) to calculate the risk of patients suffering CHD events. However, there is some evidence that CHD

risk estimation methods derived from the Framingham Study overestimate the risk for patients from the UK. (Brindle et al., 2003) Therefore, using a CHD risk equation derived from European data may be more appropriate, such as the Systematic Coronary Risk Evaluation (SCORE) risk model. (Conroy et al., 2003)

If the procedure of ranking patients by their risk of suffering events was to be used with other treatments, for different health conditions, then risk models would have to be produced for each different condition. In all cases over time the validity of the models would need to be checked and possibly updated and modified. This could be due, for example, to changes in the patient population or modified prescribing recommendations. Using nicorandil as an example if the prescribing recommendations for antianginal drugs for patients with stable angina were altered, since whether not a patient was being treated with a long acting nitrate or a loop diuretic were both variables used in the calculation of R_{CHD}, the risk model would need to be checked to see if it was still valid.

This type of procedure identifying high risk groups of patients would be most applicable where a treatment was known to be effective but the cost-effectiveness of using it to treat the entire patient population was above the £30,000/QALY threshold. If those patients who are at a higher risk of suffering events could be identified using a risk model, then treatment could be targeted at those patients who would most benefit from it due to their high risk of suffering an event. In this sub-group of the patient population the cost-effectiveness of the treatment would be improved.

8.6 The Most Cost-Effective Patients to Treat

Other methods for targeting treatments to the most cost-effective sub-group of patients have been developed and implemented on data from clinical trials. One of these methods was demonstrated in the paper by Jönsson et al. (1999), which is based on the Scandinavian Simvastatin Survival Study (4S). (Scandinavian Simvastatin Survival Study Group, 1994) The patients who were randomised in 4S were all pre-existing sufferers of CHD disease. The method entailed investigating the cost-effectiveness of a sub-group of patients based on one particular clinical characteristic, in the paper by Jönsson et al. (1999) this characteristic was whether patients were diabetic. The cost-effectiveness of simvastatin, the cholesterol lowering drug investigated in 4S, was then investigated for patients who were diabetic and then compared against the non-diabetic patients. The cost-effectiveness of simvastatin for diabetic patients was improved compared to non-diabetic patients but as was the case in the IONA Study with nicorandil it was cost-effective to prescribe simvastatin to the whole patient population. In other circumstances this may not be the case and the idea behind treating or not treating patients based on one clinical or demographic variable, to improve the cost-effectiveness of the treatment, has merit. In the case of CHD the variable used could be, for example, the age of patients, their diabetic status, as was used in the paper by Jönsson et al. (1999) or the smoking status of patients. This method would also be simpler to implement as only information regarding one variable on patients is required unlike in the calculation of R_{CHD} in Section 8.5, where information regarding several variables was required.

A further cost-effectiveness analysis of 4S was undertaken in the paper by

Johannesson et al. (1997) and in this analysis unlike in the one undertaken in the paper by Jönsson et al. (1999) more than one recorded baseline variable was considered when the population of patients were split into sub-groups to look at cost-effectiveness. Unlike the paper by Jönsson et al., 1999 whether patients were diabetic was not considered but the age, sex and cholesterol levels of patients before treatment with simvastatin or placebo were. The cost-effectiveness of the use of simvastatin was then calculated separately for men and women of different ages and baseline levels of cholesterol. As for when diabetic and non-diabetic patients were investigated all the sub-groups of patients were cost-effective to treat with simvastatin. Treating younger males with higher levels cholesterol was the most cost-effective sub-group of patients to treat. This type of procedure could be extended to include additional or different variables depending on the circumstances and the type of treatment being considered. This could lead to a treatment that was not cost-effective to prescribe to all patients being targeted at sub-groups of the potential patient population in which the cost-effectiveness of the treatment would be acceptable.

In the paper by the Heart Protection Study Collaborative Group (2005) the cost-effectiveness of simvastatin was again investigated this time based on the MRC/BHF Heart Protection Study (HPS). (Heart Protection Study Collaborative Group, 2002) As with 4S those patients who were randomised in the HPS had a history of CHD. A five year multivariate risk score similar to R_{CHD}, described in Section 8.5, was calculated using the parameter estimates found from a Cox model. Patients were then ranked according to their five year risk score and split into fifths. The cost-effectiveness of each fifth was then assessed. Prescribing simvastatin to each of the fifths of risk was found to be cost-effective. Those patients

calculated to be at the highest risk were the most cost-effectiveness sub-group of patients to treat. Similar results were seen in Section 8.5 for the IONA Study.

An additional factor that can affect the long term cost-effectiveness and the actual cost of a treatment itself, in the case of both 4S and the HPS the cost-effectiveness of simvastatin, is when a drug goes off patent. Generic copies of the drug can then be manufactured resulting in the price of the drug decreasing. This factor was taken into consideration in a follow up analysis of the HPS, looking at the lifetime cost-effectiveness of simvastatin. (Heart Protection Study Collaborative Group, 2006) As has already been stated simvastatin was cost-effective when treating the whole patient population. For treatments which are on the cusp of being cost-effective to use they could become cost-effective if when the drug goes off patent a generic equivalent can be produced at a fraction of the cost of the original. This could also allow treatments, where cost-effective sub-groups had previously been identified, to be prescribed to an enlarged sub-group of patients as a result of the cost of the treatment having been reduced. Implying that once an economic evaluation of a treatment has been undertaken and the findings implemented the findings should be revisited if and when circumstances change to see if they are still valid.

The discussed procedures of identifying patients who are at a higher risk of suffering events as well as the method introduced in Section 8.5 can all be used to aid the cost-effectiveness evaluations that treatments undergo. In the studies mentioned the treatments being investigated turned out to be within the bounds of cost-effectiveness to treat the whole patient population. Therefore, the full benefit of identifying sub-groups of the patient population where the cost-effectiveness was improved was not seen. Where the overall cost-effectiveness of a treatment

makes it too expensive to be recommended for widespread use these techniques could be implemented to identify sub-groups where the cost-effectiveness was below the cost-effectiveness acceptability threshold. The treatment could then be recommended for use within these sub-groups of patients.

8.7 The Relationship between Beneficial and Adverse Events

Using the IONA Study data it will be explored whether there could be a relationship between the benefits that treatments give to patients and the adverse events that they cause, due to side-effects, and what if any the economic implications of this relationship would be. Patients suffered GI events during the IONA Study and models for the risk of patients suffering GI events were found in Section 3.6. From these models and the analysis contained in Section 2.4.3 it was shown that treatment with nicorandil was associated with an increased risk of suffering a GI event. The nature of the GI events and the increase in the number of events that patients suffered were not serious compared to the benefits treatment with nicorandil provided by preventing CHD events. In other situations and with other treatments this may not be the case, such as chemotherapy treatment for patients suffering from cancer. If a relationship did exist between the beneficial and adverse events a treatment caused this could be taken into consideration on both clinical and economic grounds when the treatment is being prescribed.

In Section 8.5 the risk of patients suffering CHD events was calculated from Equation 8.6 and the patients were then split up into three groups: low, medium

and high risk patients. Kaplan-Meier estimates (Kaplan and Meier, 1958) and the resulting curves for the GI events patients suffered can be plotted for the three levels of CHD risk and these are shown in Figure 8.4.

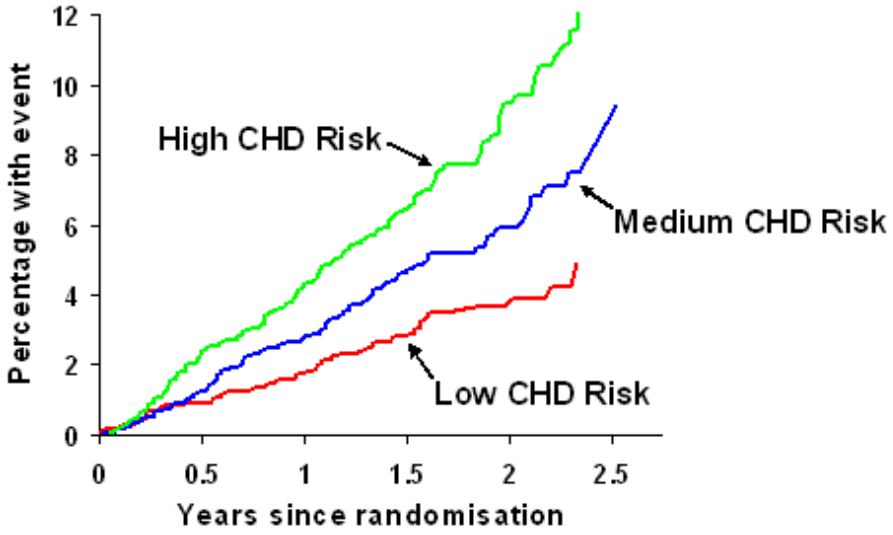


Figure 8.4: Kaplan-Meier curves for suffering a GI event for the three levels of CHD risk

From Figure 8.4 it can be seen that as time progresses there is an apparent separation in the curves for the three levels of CHD risk. There does appear to have been a relationship between the risk of suffering a CHD and GI event: if a patient was at high risk of suffering a CHD event they were also at a higher risk of suffering a GI event. This was true for the IONA Study but it may well not be true in other situations with other treatments. An investigation would have

to be carried out for each separate case and set of circumstances.

To see if the relationship was as strong in the other direction, using Model B2 found in Section 3.6.1, the risk of a patient suffering a GI event was calculated in a similar way to how the risk of suffering CHD events was calculated using Equation 8.6. Let the risk of a patient suffering a GI event equal R_{GI}, then the risk was calculated as follows:

$$R_{GI} = exp(\underline{y}_i \underline{\hat{\beta}}) \tag{8.8}$$

where $\underline{y}_i$ is the $1 \times p$ vector of prognostic variables for patient i, where p is the number of variables included in the model, and $\underline{\hat{\beta}}$ is the $p \times 1$ vector of parameter estimates found from the Cox model for the GI events.

Using Equation 8.8 the risk of patients suffering a GI event was then calculated. As was the case when the risk of patients suffering CHD events was calculated using Equation 8.6 when the risk of patients suffering GI was calculated all patients were treated as if they were in the placebo group of the study. The risk of 16 (0.6%) patients in the nicorandil group and 13 (0.5%) patients in the placebo group could not be calculated as there were missing values for one or more of the baseline variables included in the GI event model. The patients were split according to their risk of suffering GI events into three risk groups: low, medium and high. Kaplan-Meier estimates and the resulting curves were produced for the CHD events patients suffered and plotted for the three levels of GI risk and are shown in Figure 8.5.

There is a similar pattern in Figure 8.5 to the one seen in Figure 8.4 in that those patients who were calculated to be at high risk of GI events were also

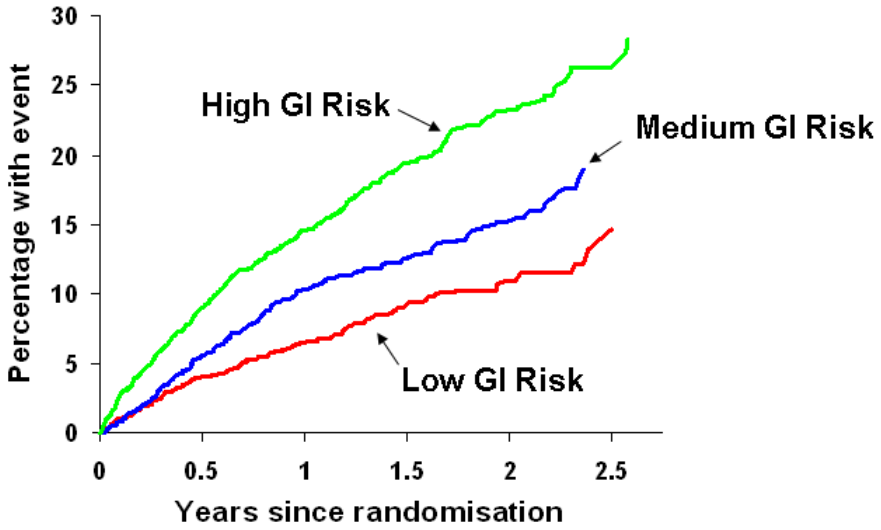


Figure 8.5: Kaplan-Meier curves for suffering a CHD event for the three levels of GI risk

at high risk of suffering CHD events. Given that the multivariable predictive models for the CHD events and GI events share common prognostic variables some relationship between the two types of events would be expected. The model for the CHD events, found in Section 3.4.1, and Model B2 for the GI events, found in Section 3.6.1, contained the following common variables: the angina status of patients as assessed by the CCSF classification, whether patients were being treated with a long acting nitrate and whether patients were being treated with a loop diuretic. In addition as the BMI of a patient was included in the model for the CHD events and the weight of a patient, which is used in the calculation

of BMI, was included in the model for the GI events there was a further link between the models for the two types of event.

Having seen that there was a relationship between the risk of suffering a CHD and GI event how this affects the cost-effectiveness of treatment with nicorandil will be explored. This was done in a similar way to how the cost-effectiveness of the different levels of CHD risk were investigated in Section 8.5. However, before the patients were ranked according to their risk of CHD they were ranked by their risk of suffering a GI event using Equation 8.8. In order that each risk group would contain sufficient numbers of patients from both treatment groups the data set was split in half to give a low and a high GI risk group. Using Equation 8.6 these two groups were further divided into low and high CHD risk groups, giving four risk groups in total. Shown in Table 8.10 are the proportion of patients who suffered a first CHD event and the mean cost per patients for the nicorandil and placebo groups separately for the four levels of combined GI and CHD risk. Using the information contained in Table 8.10 and the modified ICER formula, Equation 8.7, ICERs for the four risk levels were calculated as well as estimated 95% CIs using both Fieller's Theorem and bootstrap sampling. The results are shown in Table 8.11.

From the point estimates for the ICERs for the combined GI and CHD risk groups it was seen that for the low GI risk group it was cost-effective to treat both the low and high CHD risk groups. Although compared to the previously calculated ICERs for the IONA Study the low GI risk group were the least cost-effective group of patients to treat. The ICER for the high CHD risk group was higher than for the low CHD risk group. This may have been as result of the fact that patients in this group were at a higher risk of suffering CHD events and

Table 8.10: The proportion of patients who suffered a primary endpoint and the mean cost per patient in the nicorandil and placebo groups for the combined levels of GI and CHD risk

Low GI Risk Group		
Low CHD Risk Group	Nicorandil (n = 652)	Placebo (n = 742)
Proportion of patients who suffered a CHD event	0.06992	0.09356
Mean cost per patient (£)	1,159	1,057
High CHD Risk Group	Nicorandil (n = 622)	Placebo (n = 645)
Proportion of patients who suffered a CHD event	0.1206	0.1442
Mean cost per patient (£)	1,543	1,396
High GI Risk Group		
Low CHD Risk Group	Nicorandil (n = 670)	Placebo (n = 592)
Proportion of patients who suffered a CHD event	0.1134	0.1233
Mean cost per patient (£)	1,382	1,373
High CHD Risk Group	Nicorandil (n = 617)	Placebo (n = 646)
Proportion of patients who suffered a CHD event	0.2236	0.2601
Mean cost per patient (£)	2,237	2,459

therefore accruing hospitalisation costs than in the low CHD risk group. However, the point where the cost of nicorandil was more than off-set by the reduction in the number of CHD events and resultant hospitalisation costs that these events would have caused had not yet been reached. As a result the cost-effectiveness of treatment with nicorandil was reduced. In addition, although this group of patients was the low GI risk group, patients in the high CHD risk group may have suffered more GI events that in the low CHD risk group due to the relationship seen between the two types of event, see Figures 8.4 and 8.5.

Any increase in the number of GI events and resultant hospitalisations suffered by patients would impact on the cost-effectiveness of treatment with nicorandil.

Table 8.11: The cost per CHD event prevented for the combined levels of GI and CHD risk with accompanying estimated 95% CIs

Level of CHD Risk	ICER Point Estimate (£)	Method of Estimating the 95% Confidence Intervals	
		Fieller's Theorem (£)	Bootstrap Sampling (£)
Low GI Risk Group			
Low	4,324	$-\infty$ – -32,391 $\cup$ -16,071 – ∞	-74,170 – 96,798
High	6,217	$-\infty$ – ∞	-98,599 – 104, 671
High GI Risk Group			
Low	872	$-\infty$ – ∞	-138,063 – 120,148
High	-6,099	$-\infty$ – ∞	-60,373 - 52,290

For the high GI risk group treating both the low and high CHD risk groups was also cost-effective. In the high CHD risk group every CHD prevented resulted in a net cost saving of £6,099, this net cost saving was the largest seen when analysing the IONA Study data. For the low CHD risk group the net cost of ever CHD event prevented was £872 which was again favourable and lower than for either of the low or high CHD risk groups in the low GI risk group.

By ranking patients by their risk of suffering a GI event and then by there risk of suffering a CHD event the cost-effectiveness of nicorandil was improved. There were no significant differences in the effectiveness of treatment with nicorandil over treatment with placebo in the four risk groups and as a result the 95% CIs as estimated by Fieller's Theorem were not very informative. For three of the groups the intervals were the entire real line and for the low risk CHD group in the low risk GI group the interval only excluded the interval (-£32,391, -£16,071). For this group real roots to the solution of Equation 8.5 did exist but as there was no significant difference in the effectiveness of the treatments the interval was of the form $(-\infty, a] \cup [b, \infty)$. The bootstrap sampling estimates were more informative but they were still quite wide due to there being no significant difference between

the effectives of the treatments for any of the risk groups. The bootstrap sampling interval was especially wide for the low CHD risk group in the high GI risk group. As all of the 95% CIs were so wide no definitive conclusions can be made about the cost-effectiveness of treatment with nicorandil. Although the general pattern was that treating the group of patients who were at a higher risk of suffering GI

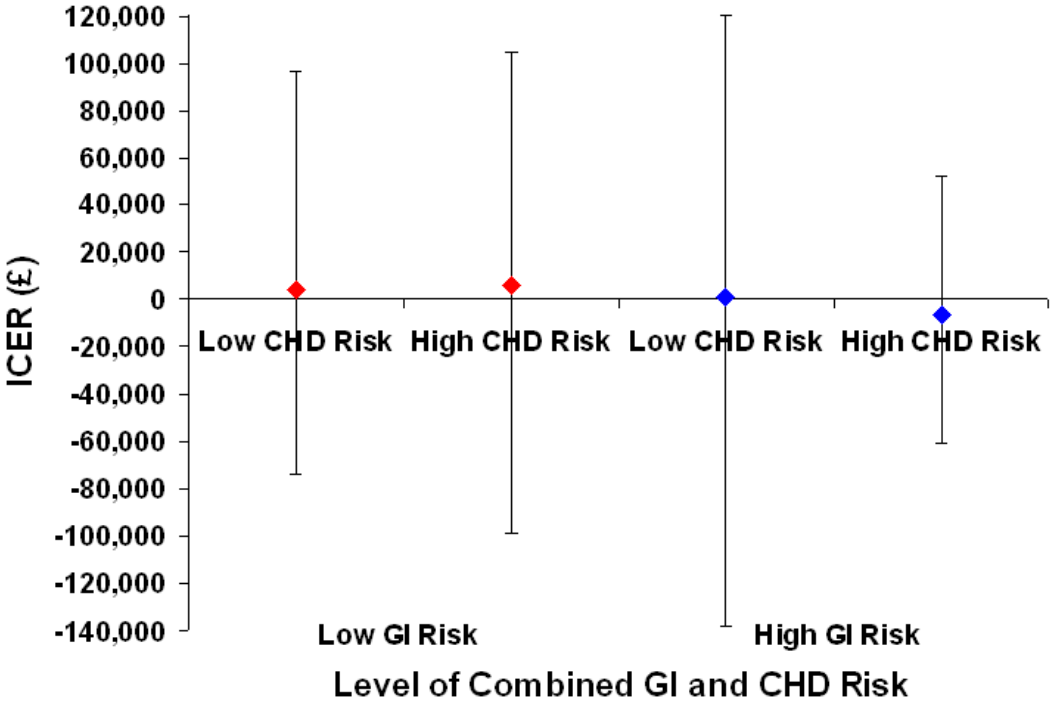


Figure 8.6: The cost per CHD event prevented for the levels of combined GI and CHD risk with accompanying 95% CIs estimated by bootstrap sampling

events was more cost-effective. A graphical representation of the ICERs for the combined levels of GI and CHD risk, along with the 95% CIs found by bootstrap sampling, are shown in Figure 8.6.

It has been shown that treatment with nicorandil reduced the incidence of

CHD events but increased the incidence of GI events and that there was a relationship between the risks of patients suffering both these types of events. If only those patients who were calculated to be at high risk of suffering CHD events were treated with nicorandil this would lead to a higher rate of patients suffering GI events. As the rate of patients suffering GI events was small compared to the rate patients suffered CHD events and the cost of these events were small compared to CHD events the economic impact of treating only patients at high risk of CHD and GI events would be minor. As the cost of the GI events had minimal impact on the cost-effectiveness calculations in this instance the cost-effectiveness of nicorandil was actually improved as the sub-group of patients identified were at the greatest risk of suffering CHD events so therefore would benefit most from treatment with nicorandil. However, if this idea was to be applied in other situations with different treatments where the adverse events they cause are more serious and result in higher costs required to treat them then only treating high risk patients could adversely affect the cost-effectiveness of a treatment. If this process, or a similar one, was to be implemented with other treatments and not just with nicorandil investigation would be needed to see if analogous findings to the ones for the IONA Study could be found. This process would be most applicable to treatments that have high rates of serious side-effects that are costly to treat and therefore have a large impact of the cost-effectiveness of a treatment.

Chapter 9

Conclusions and Further Work

The original clinical and economic evaluations for the IONA Study (The IONA Study Group, 2002; Walker et al., 2006), the details of which are given in Sections 2.4.1, 2.4.2 and 8.3.1, indicated that nicorandil did have cardioprotective effects and that to achieve these clinical benefits treatment with nicorandil was cost-effective. By extending both the clinical and economic evaluations it was seen that the clinical benefits of treatment with nicorandil extended past the first CHD events patients were at risk from and that the cost-effectiveness of treatment with nicorandil could be further improved.

It was found that adjustment for baseline explanatory variables did not substantially affect the estimated risk reduction for CHD events. Nicorandil treatment was also found to be associated with an increased risk of GI events. This increased risk was still present in the multivariable model for the GI events. The benefits of nicorandil treatment would appear to out weigh the potential drawback of increased risk of suffering GI events.

The area where the clinical evaluation could really be extended was by including not just the first event suffered but also the multiple recurrent events that patients suffered. Considering only the primary endpoint where the number of recurrent events was sufficient to make best use of the recurrent event models, when the recurrent event analyses were compared to the time-to-first event analysis the statistical significance of the reduction in risk of patients suffering CHD associated with treatment with nicorandil was improved for all of the methods of analyses investigated. Here the underlying risk reduction associated with treatment with nicorandil is not known. However, as was seen in Chapter 6 for the second set of simulation conditions for the treatment effect group of simulations where the event rate was similar to the event rate for the primary endpoint in IONA, the AG and WLW models estimated the treatment effect to be larger compared to the PWPa and PWPb models. For the simulations the reduction in risk that the treatment should cause was known and the results for the PWPa and PWPb models more accurately reflected this. For the IONA Study the AG and WLW models estimated the reduction in risk caused by treatment with nicorandil to be marginally higher then the PWPa and PWPb models did. It appeared that for the IONA Study data the PWPb model most closely modelled the underlying data.

Having introduced the random subject effect into the data generation process it was seen the PWPb model did not perform well under these conditions. The positive stable variable used to generate the random subject effect had the affect of reducing the parameter estimate for the treatment effect by half. The event rate was also greatly increased. As discussed in Chapter 5 the heterogeneity of patients has to be considered but it is unlikely that in real life situations it

would have such a dramatic effect on the parameter estimate. The advantage of using the positive stable distribution in this case was that the assumption of proportional hazards was preserved. If further simulations were to be performed other data generation processes than used in Chapter 6 should be used to make any findings more generalisable and to possibly reflect more accurately the degree of frailty that might exist in a real population. (Metalfe and Thompson, 2006) For example, a Poisson process, an Autoregressive model or a Gamma frailty could be used to generate the data.

The inclusion of the random subject effect did highlight that the recurrent event models do not always give the same results and in some circumstances the differences between the results for the models can be significant. It is therefore important that the most appropriate recurrent event model is used to analyse data. The particular conditions for the data set such as the event rate and whether the risk of events was constant need to be taken into consideration. The model used to analyse the data should not be based on preconceived ideas but on the actual observed pattern of the data. If a clinical evaluation is extended past a time-to-first event analysis to a recurrent event analysis the findings may gain greater weight. However, if the model that is used to incorporate the recurrent events into the analysis is not appropriate some of the benefits of extending the analysis past a simple time-to-first event analysis may be lost as well as making the results harder to interpret. The same is true if there are insufficient numbers of recurrent events available to include in the analysis.

One obvious approach to remove underlying frailty is to adjust for baseline risk factors. Particularly for cardiovascular disease, many of the risk factors are well understood. It is recommended that all recurrent event analyses should be

adjusted for important baseline predictors of outcome. The incorporation of an explicit frailty parameter in the model is another approach worthy of consideration. However, it is noted that this will not adjust for frailty that develops with the occurrence of recurrent events. In some situations it may be best to use only the traditional time-to-first event analysis, and such analyses should perhaps always be considered as the primary outcome of a study.

An alternative approach that could be used to model recurrent event data that would potentially adjust for frailty that develops with the occurrence of recurrent events would be to use a multi-state model. (Andersen and Keiding, 2002) When using a multi-state model all patients start in the same state at the start of the trial. Patients then transition to other states when they suffer events. The states that patients transition to depend on the type of event they suffer. If the event patients suffer is death or another type of event that means they can no longer participate in the trial they transition to an absorbing state. For patients who transition to non-absorbing states they remain in that state until they suffer a further event. Patients then transition to another state. Again the states that patients transition to depend on the type of event they suffer. This modelling process then continues for as long as patients are at risk of suffering events. The use of multi-state models could be further extended by using competing risks multi-state models. (Andersen et al., 2002; Putter et al., 2007)

It was seen that the cost-effectiveness of treatment with nicorandil could be improved. This was achieved by targeting it at sub-groups of the patient population. The sub-groups being patients who had a higher underlying risk of suffering CHD events. By targeting the prescribing of nicorandil to those patients calculated to be at the highest risk nicorandil would become a very favourable option

for the NHS. The rationale for ongoing treatment was strengthened by the fact that the cardioprotective effects of treatment with nicorandil extended past the first CHD events patients suffered. Patients who had suffered previous CHD events as well as patients newly diagnosed with CHD could benefit from treatment with nicorandil. Once patients have suffered a CHD event treatment with nicorandil could be continued as it would still be clinically effective as well as still being cost-effective. If the long term cost-effectiveness of treatment with nicorandil was to be investigated this could be achieved by using a Markov model. (Briggs and Sculpher, 1997; Lang et al., 2003) Using a Markov model the number of CHD events that patients who took part in the IONA Study would suffer after the end of the study and how frequently they would suffer such events could be modelled. Thus, allowing the long term cost-effectiveness of treatment with nicorandil to be assessed.

A potential drawback of nicorandil is the GI side-effects that nicorandil could cause in patients. The side-effects, if serious and requiring treatment and even hospitalisation, could have implications on the cost-effectiveness of a treatment. These implications were not fully considered when the ICERs for the IONA Study were calculated. Both the primary and secondary IONA ICERs included the cost of the GI hospitalisations but no consideration was given to the extra actual numbers of GI events patients suffered as in some way counteracting the numbers of CHD events prevented. The number of GI events patients suffered could be factored into the calculation of the ICERs as well as the costs that accompanied these events. If the GI events were to be factored into the ICER calculation a weighting factor would need to be applied to the GI events, as the GI events were generally less severe and therefore less costly than the CHD events. They were

also not a predefined outcome of the study. Therefore, if a weighting factor of one was applied to the CHD events then it would be likely that the GI events would be assigned a weighting factor of less than one. If the GI events were included in the denominator of the ICER and given an arbitrary weight of a half the calculation of the secondary IONA ICER would become:

$$\begin{aligned} \text{Secondary IONA ICER} &= \frac{(4,065,559 - 4,030,997)}{(-415.5 - (-452))} \\ &= £947 \end{aligned} \tag{9.1}$$

The adjusted secondary IONA ICER, with the inclusion of the recurrent GI events in addition to the recurrent CHD events, would become:

$$\begin{aligned} \text{Adjusted Secondary IONA ICER} &= \frac{(4,065,559 - 4,030,997))}{(-560 - (-668.5))} \\ &= £319 \end{aligned} \tag{9.2}$$

For the secondary IONA ICER the net cost per 'Event' prevented rises to £947 and for the adjusted secondary IONA ICER it rises to £319. These figures still indicate treatment with nicorandil to be cost-effective. The advantage with these formulations of the ICERs is that the numbers of adverse clinical events as well as the cost associated with them are now factored into the ICER calculation.

This does raise an issue which could be further explored. The primary endpoint was a composite endpoint of three types of CHD event, which greatly differed in their severity. In the CEA the component parts of the primary endpoint were treated as the same and no consideration was given to the difference in their severity and the implications of suffering them had on patients. The

component parts of the primary endpoint could have different weighting factors applied to them. Consideration could also be given to whether there should only be three weighting factors applied to the three component parts of the primary endpoint or whether the non-fatal MIs and unplanned hospitalisations should be further sub-divided based on the severity of the event. Having decided how many weighting factors to apply to the component parts of the primary endpoint, as well as whether there should be multiple weighting factor for the GI events, the values of these weighting factors would have to be estimated. This could be done by looking at QALY values that have been previously assigned to different types and severity of CHD and GI events. The assigned QALY values could then be converted into weighting factors. How introducing these weighting factors would impact the cost-effectiveness of treatment with nicorandil could then be explored. Following on from this, if monetary values could be assigned to the QALY values allocated to the different types of event a net health benefit analysis could be performed. (Stinnett and Mullahy, 1998; Willan, 2001) In performing a net health benefit analysis the difficulties of calculating and interpreting ICERs can be avoided. However, monetary values have to be assigned to QALYs which also has its difficulties.

In Chapter 8 it was seen that there was a positive relationship between the risk of suffering CHD and GI events. By treating those patients who were at a high risk of suffering GI events as well as CHD events the cost-effectiveness of treatment with nicorandil could be improved. This was as a result of the fact that the cost impact of the GI events were small compared to the CHD events. However, if side-effects have a larger impact on the net cost of prescribing treatments these cost-effectiveness improvements would be unlikely. Any relationship between patients

most likely to experience clinical benefit from treatments as well as side-effects could be an important association and the potential links between the two types of event could be investigated for other treatments. If patients will benefit from a treatment but are known to be at a heightened risk of suffering side-effects this could be taken into consideration when treatments are prescribed. Over and above the cost-effectiveness of treatments there always has to be a balance between the benefits and the harm that treatments can cause to patients.

Sensitivity analysis was carried out on the results of the economic evaluation of the IONA Study but this is an area where the analysis could be further extended. To gain a greater understanding of the uncertainty around the cost-effectiveness of treatment with nicorandil instead of performing a univariate sensitivity analysis or extending that to a multi-way sensitivity analysis Probabilistic Sensitivity Analysis could be used. If distributions for the hospitalisation costs and numbers of primary endpoints suffered in both treatment groups could be specified for the IONA Study then PSA could be performed. The fact that high risk patients were enrolled in IONA could then be factored into the analysis. The effectiveness of treatment with nicorandil could be altered so that the cost-effectiveness of nicorandil could be investigated for the general patient population and not just those at an elevated risk of suffering CHD events. Thus, allowing the result of the CEA to be more general.

9.1 Summary

Both the original clinical and economic evaluations have been extended. The results for the extended clinical evaluation have shown that the cardioprotective

effects of nicorandil were not just confined to the first CHD events that patients were at risk of suffering. In the original economic evaluation the cost-effectiveness of treatment with nicorandil was found to be favourable and by extending the analysis this finding has been reinforced and in some cases the cost-effectiveness was further improved. The original clinical and economic findings of the IONA Study have been improved and made more robust as a result of utilising more of the information that was recorded on patients during the study.

In modern society pharmaceutical companies are under great financial pressure to bring their products to market and as a result are trying to shorten the length of clinical trials. For example, by using composite endpoints, as was the case in the IONA Study. If the follow-up of clinical trials are shortened it is likely that modelling of the trial data will become ever more important. In order that this can be best achieved it is important that not only are data recorded on patients during clinical trials but that the data are actually utilised in the clinical and economic evaluations that are undertaken. By incorporating more of the information recorded on patients during a clinical trial and by making the best use of this information both the clinical and economic evaluations of clinical trials can be expanded upon and lead to improved and more robust conclusions being drawn.

A summary of the main findings are as follows:

1. The cardioprotective effects of nicorandil were shown to extend past the first CHD events patients suffered.

2. The recurrent event models all have different underlying assumptions and as a result different strengths and weaknesses.

3. When analysing recurrent event data the choice of which model to use is important and should be based on the particular conditions for the given data set.

4. Not using the most appropriate recurrent event model for the given data set can result in significant levels of bias in the results.

5. The cost-effectiveness of treatment with nicorandil could be further improved by targeting it to those patients who were at the highest risk of CHD events.

6. There has to be a balance between the benefits and the harm that treatments can cause to patients and as a result all such factors should be included in the calculation of cost-effectiveness.

Appendix A

The IONA Study Endpoint Definitions

This Appendix contains the definitions of the components parts of the primary and secondary endpoints of the IONA Study that were used by the critical events committee. (The IONA Study Group, 2001)

A.1 Hospitalisations

Before definitions of the component parts of the primary and secondary endpoints are given a hospital admission will be defined as well as what constituted an unplanned hospital admission.

1. Hospital Admission

   An admission to hospital is defined as any attendance at hospital requiring completion of the hospital admission procedures and usually at least an overnight stay.

2. Unplanned Hospital Admission

   Unplanned admission is defined as an emergency or other urgent, non-elective admission precipitated by general practitioner (GP) referral or self referral to an accident and emergency department, urgent GP referral to hospital in some other way, emergency call to the ambulance service, or urgent admission from a hospital outclinic. The admission must be precipitated by a need for urgent investigation or treatment which cannot be provided on an outpatient basis and which cannot be deferred on an inpatient basis.

A.2 Coronary Heart Disease Death

This consisted of all events satisfying the following definition:

All deaths shall be considered coronary heart disease unless an unequivocal non-coronary heart disease cause can be established. Coronary heart disease deaths will include sudden deaths, death due to myocardial infarction, death due to heart failure, death due to a cardiac investigation/procedure/operation (procedure related death).

1. Sudden death

   Deaths fulfilling any one of the following criteria:

   (a) witnessed and instantaneous, without new or worsening symptoms;

   (b) witnessed and preceded or accompanied by symptoms attributable to myocardial ischaemia but without other new or worsening symptoms;

(c) witnessed and preceded by symptoms attributable to a cardiac arrhythmia for example, syncope or near syncope;

(d) patients resuscitated from cardiac arrest in the absence of worsening heart failure or other causes of death, including acute myocardial infarction, and who die within 24 hours or without regaining consciousness; similar patients who die despite attempted resuscitation;

(e) unwitnessed death in the absence of worsening heart failure or other causes of death.

2. Death from heart failure

Death occurring when at least one of the following is present in the 48 hours before death:

(a) new or increasing symptoms and/or signs (including worsening renal function) of heart failure requiring the initiation of, or an increase in, treatment directed at heart failure or occurring in a patient already receiving maximum treatment for heart failure;

(b) heart failure symptoms or signs requiring continuous intravenous treatment or oxygen administration;

(c) confinement to bed but only if confinement is for heart failure symptoms;

(d) pulmonary oedema sufficient to cause tachypnoea and distress not occurring in the context of an acute myocardial infarct or as the consequence of an arrhythmia.

3. Death from myocardial infarction

   Death occurring up to 28 days after a documented myocardial infarct. Deaths from a myocardial infarct occurring as a direct result of an investigation, procedure, or operation will be classified as a death caused by myocardial infarction and a procedure related death.

4. Coronary heart disease procedure related

   Deaths deemed to be directly related to that investigation/procedure/operation.

5. Presumed coronary heart disease death

   Death not fulfilling any of the above coronary heart disease categories and any definite cardiovascular or non-cardiovascular definition below.

   (a) Death from stroke

   Death occurring up to 28 days after a documented stroke. Deaths from stroke occurring as a direct consequence of an investigation/procedure/operation will be classified as a death caused by a stroke and a procedure related death.

   (b) Cardiovascular procedure related deaths

   Death occurring within seven days of a cardiovascular investigation, procedure or operation and deemed to be directly related to that investigation/procedure/operation.

   (c) Death from other cardiovascular causes

   Death must be caused by a fully documented cardiovascular cause not included above - for example, pulmonary embolism, ruptured aortic aneurysm, and so on.

A.3 Non-Fatal Myocardial Infarction

This consisted of all events satisfying the following definition:

1. Silent myocardial infarction

   An ECG, at an annual or at an unscheduled visit, that is diagnostic of myocardial infarction (new Q waves $\geq$ 0.04 ms in duration in at least two consecutive leads) and which was not evident on the previous ECG.

2. Acute myocardial infarction

   All definite myocardial infarcts will be counted as events whether they occurred spontaneously or as the direct consequence of an investigational procedure or operation. A diagnosis of myocardial infarction will be made if two of the following three criteria are met:

   (a) At least one of the following:

      i. cardiac ischaemic type pain lasting at least 30 minutes;

      ii. pulmonary oedema;

      iii. cardiogenic shock not otherwise explained.

   (b) Development of new abnormal Q waves ($\geq$ 0.04 ms in duration) in at least two consecutive ECG leads not present on an ECG recorded before the current event, or transient elevation of ST segment followed by T wave inversion in at least two consecutive leads, or new left bundle branch block, or transient elevation of ST segment, new bundle branch block, or other typical ECG changes leading to emergency angiography during which a complete acute occlusion of at least one coronary

artery is demonstrated and following which successful emergency percutaneous revascularisation (recannulation) is performed.

(c) An elevation of cardiac enzymes defined as a transient increase in at least one set of enzymes. (Elevation to at least twice the upper limit of the local normal reference range, or creatine kinase (CK) MB fraction greater equal here 10% of total CK)

A.4 Unplanned Hospital Admission for Cardiac Chest Pain

This consisted of all events satisfying the following definition:

Cardiovascular, involving chest pain.

1. Acute myocardial infarction

   All definite myocardial infarcts will be counted as events whether they occurred spontaneously or as the direct consequence of an investigational procedure or operation. A diagnosis of myocardial infarction will be made if two of the following three criteria are met:

   (a) At least one of the following:

      i. cardiac ischaemic type pain lasting at least 30 minutes;

      ii. pulmonary oedema;

      iii. cardiogenic shock not otherwise explained.

   (b) Development of new abnormal Q waves ($\geq$ 0.04 ms in duration) in at

least two consecutive ECG leads not present on an ECG recorded before the current event, or transient elevation of ST segment followed by T wave inversion in at least two consecutive leads, or new left bundle branch block, or transient elevation of ST segment, new bundle branch block, or other typical ECG changes leading to emergency angiography during which a complete acute occlusion of at least one coronary artery is demonstrated and following which successful emergency percutaneous revascularisation (recannulation) is performed.

(c) An elevation of cardiac enzymes defined as a transient increase in at least one set of enzymes. (Elevation to at least twice the upper limit of the local normal reference range, or creatine kinase (CK) MB fraction greater equal here 10% of total CK)

2. Chest pain which is not associated with a myocardial infarct but requires unplanned admission to hospital:

   (a) Unstable angina

   Typical cardiac ischaemic type chest pain requiring hospital admission for treatment but not meeting the definition of myocardial infarction. The patient must also:

      i. develop new or evolving ST segment/T wave changes on the ECG and

      ii. have treatment with parenteral (buccal other than with short acting preparations, intravenous, transcutaneous or subcutaneous) heparin and/or glyceryl trinitrate, isosorbide dinitrate, or other nitrate.

(b) Definite angina

Typical cardiac ischaemic type chest pain requiring hospital admission for treatment but not meeting the definition of myocardial infarction or definite unstable angina and requiring additional antianginal treatment (new antianginal drugs and/or increased dose of current treatment and/or referral for "revascularisation").

(c) Presumed angina

Admission with chest pain not fulfilling any of the above criteria and where there is no other recorded cause for the pain. That is, any admission with chest pain that is not caused by myocardial infarction, unstable angina, or definite angina is "presumed angina" unless a firm diagnosis to the contrary is recorded by the treating physician.

Appendix B

SAS Simulation Code

This Appendix contains the SAS code used to simulate and then analyse the recurrent event data. Shown is the SAS code used to generate the gap times for the different groups of simulations. Additionally, then shown is the SAS code used to incorporate censoring into the gap time data and then analyse it using the different recurrent event models for the no treatment effect group of simulations with $\gamma = 1$.

B.1 Gap Time Generation Code

No Treatment Effect Group of Simulations with $\gamma = 1$

```
data null_exp (drop = seed seed2);
do k=1 to 10000;
seed = -1; seed2 = -1;
null = 0;
do i=1 to 1000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = ranexp(seed)*(1/0.1116);
d3 = ranexp(seed)*(1/0.1116);
d4 = ranexp(seed)*(1/0.1116);
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end ;
null = 1;
do i=1001 to 2000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = ranexp(seed)*(1/0.1116);
d3 = ranexp(seed)*(1/0.1116);
d4 = ranexp(seed)*(1/0.1116);
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
end;
run;
```

No Treatment Effect Group of Simulations with $\gamma = 0.75$

```
data null_weibull (drop = seed seed2);
do k=1 to 10000;
seed = -1; seed2 = -1;
null = 0;
do i=1 to 1000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = rand('weibull',0.75,(1/0.1116));
d3 = rand('weibull',0.75,(1/0.1116));
d4 = rand('weibull',0.75,(1/0.1116));
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
null = 1;
do i=1001 to 2000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = rand('weibull',0.75,(1/0.1116));
d3 = rand('weibull',0.75,(1/0.1116));
d4 = rand('weibull',0.75,(1/0.1116));
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
end;
run;
```

Treatment Effect Group of Simulations with $\gamma = 1$

```
data treatment_exp (drop = seed seed2);
do k=1 to 10000;
seed = -1; seed2 = -1;
treatment = 0;
do i=1 to 1000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = ranexp(seed)*(1/0.1116);
d3 = ranexp(seed)*(1/0.1116);
d4 = ranexp(seed)*(1/0.1116);
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
treatment = 1;
do i=1001 to 2000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.0558);
d2 = ranexp(seed)*(1/0.0558);
d3 = ranexp(seed)*(1/0.0558);
d4 = ranexp(seed)*(1/0.0558);
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
end;
run;
```

Treatment Effect Group of Simulations with $\gamma = 0.75$

```
data treatment_weibull (drop = seed seed2);
do k=1 to 10000;
seed = -1; seed2 = -1;
treatment = 0;
do i=1 to 1000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = rand(‘weibull’,0.75,(1/0.1116));
d3 = rand(‘weibull’,0.75,(1/0.1116));
d4 = rand(‘weibull’,0.75,(1/0.1116));
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
treatment = 1;
do i=1001 to 2000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = rand(‘weibull’,0.75,(1/0.0558));
d3 = rand(‘weibull’,0.75,(1/0.0558));
d4 = rand(‘weibull’,0.75,(1/0.0558));
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
end;
run;
```

Treatment Effect for First Event
Group of Simulations with $\gamma = 1$

```
data first_exp (drop = seed seed2);
do k=1 to 10000;
seed = -1; seed2 = -1;
first = 0;
do i=1 to 1000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = ranexp(seed)*(1/0.1116);
d3 = ranexp(seed)*(1/0.1116);
d4 = ranexp(seed)*(1/0.1116);
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
first = 1;
do i=1001 to 2000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.0558);
d2 = ranexp(seed)*(1/0.1116);
d3 = ranexp(seed)*(1/0.1116);
d4 = ranexp(seed)*(1/0.1116);
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
end;
run;
```

Treatment Effect for First Event
Group of Simulations with $\gamma = 0.75$

```
data first_weibull (drop = seed seed2);
do k=1 to 10000;
seed = -1; seed2 = -1;
first = 0;
do i=1 to 1000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.1116);
d2 = rand(‘weibull’,0.75,(1/0.1116));
d3 = rand(‘weibull’,0.75,(1/0.1116));
d4 = rand(‘weibull’,0.75,(1/0.1116));
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
first = 1;
do i=1001 to 2000;
subject = cinv(ranuni(seed2),1);
d1 = ranexp(seed)*(1/0.0558);
d2 = rand(‘weibull’,0.75,(1/0.1116));
d3 = rand(‘weibull’,0.75,(1/0.1116));
d4 = rand(‘weibull’,0.75,(1/0.1116));
t1 = d1;
t2 = d2*(1/4);
t3 = d3*(1/4)*(1/1.5);
t4 = d4*(1/4)*(1/1.5)*(1/1.5);
s1 = d1*subject;
s2 = d2*subject;
s3 = d3*subject;
s4 = d4*subject;
a1 = t1*subject;
a2 = t2*subject;
a3 = t3*subject;
a4 = t4*subject;
output; end;
end;
run;
```

B.2 Incorporating Censoring and Analysing the Recurrent Event Data

```
data censoring_null_hr1e (drop = d1 d2 d3 d4);
set null_exp (drop = t1 t2 t3 t4 a1 a2 a3 a4 s1 s2 s3 s4);
by k;
total_time = 2;
if d1 >= 2 then do;
event1 = 0; event2 = .;  event3 = .; event4 = .;
tt1 = 2; tt2 = .; tt3 = .; tt4 = .;
end;
else if . < d1 < 2 then do;
event1 = 1; tt1 = d1; time1 = tt1;
if sum(d1 , d2) >= 2 then do;
event2 = 0; event3 = .; event4 = .;
tt2 = (2 - d1); tt3 = .; tt4 = .;
end;
else if . < sum(d1 , d2) < 2 then do;
event2 = 1; tt2 = d2; time2 = sum(tt1 , tt2);
if sum(d1 , d2 , d3) >= 2 then do;
event3 = 0; event4 = .;
tt3 = (2 - sum(d1 , d2)); tt4 = .;
end;
else if . < sum(d1 , d2 , d3) < 2 then do;
event3 = 1; tt3 = d3; time3 = sum(tt1 , tt2 , tt3);
if sum(d1 , d2 , d3 , d4) >= 2 then do;
event4 = 0;
tt4 = (2 - sum(d1 , d2 , d3));
end;
else if . < sum(d1 , d2 , d3 , d4) < 2 then do;
event4 = 1;
tt4 = d4; time4 = sum(tt1 , tt2 , tt3 , tt4);
total_time = sum(d1 , d2 , d3 , d4);
end;
end;
end;
end;
run;

data censoring_null_hr2e (drop = t1 t2 t3 t4);
set null_exp (drop = a1 a2 a3 a4 d1 d2 d3 d4 s1 s2 s3 s4);
by k;
total_time = 2;
if t1 >= 2 then do;
event1 = 0; event2 = .;  event3 = .; event4 = .;
tt1 = 2; tt2 = .; tt3 = .; tt4 = .;
end;
else if . < t1 < 2 then do;
event1 = 1; tt1 = t1; time1 = tt1;
if sum(t1 , t2) >= 2 then do;
event2 = 0; event3 = .; event4 = .;
tt2 = (2 - t1); tt3 = .; tt4 = .;
end;
else if . < sum(t1 , t2) < 2 then do;
event2 = 1; tt2 = t2; time2 = sum(tt1 , tt2);
if sum(t1 , t2 , t3) >= 2 then do;
event3 = 0; event4 = .;
tt3 = (2 - sum(t1 , t2)); tt4 = .;
end;
else if . < sum(t1 , t2 , t3) < 2 then do;
event3 = 1; tt3 = t3; time3 = sum(tt1 , tt2 , tt3);
if sum(t1 , t2 , t3 , t4) >= 2 then do;
event4 = 0;
tt4 = (2 - sum(t1 , t2 , t3));
end;
else if . < sum(t1 , t2 , t3 , t4) < 2 then do;
event4 = 1;
tt4 = t4; time4 = sum(tt1 , tt2 , tt3 , tt4);
total_time = sum(t1 , t2 , t3 , t4);
end;
end;
end;
end;
run;

data censoring_null_hr3e (drop = s1 s2 s3 s4);
set null_exp (drop = t1 t2 t3 t4 a1 a2 a3 a4 d1 d2 d3 d4);
by k;
atotal_time = 2;
if s1 >= 2 then do;
aevent1 = 0; aevent2 = .;  aevent3 = .; aevent4 = .;
aa1 = 2; aa2 = .; aa3 = .; aa4 = .;
end;
else if . < s1 < 2 then do;
aevent1 = 1; aa1 = s1; atime1 = aa1;
if sum(s1 , s2) >= 2 then do;
aevent2 = 0; aevent3 = .; aevent4 = .;
aa2 = (2 - s1); aa3 = .; aa4 = .;
end;
```

```
else if . < sum(s1 , s2) < 2 then do;
aevent2 = 1; aa2 = s2; atime2 = sum(aa1 , aa2);
if sum(s1 , s2 , s3) >= 2 then do;
aevent3 = 0; aevent4 = .;
aa3 = (2 - sum(s1 , s2)); aa4 = .;
end;
else if . < sum(s1 , s2 , s3) < 2 then do;
aevent3 = 1; aa3 = s3; atime3 = sum(aa1 , aa2 , aa3);
if sum(s1 , s2 , s3 , s4) >= 2 then do;
aevent4 = 0;
aa4 = (2 - sum(s1 , s2 , s3));
end;
else if . < sum(s1 , s2 , s3 , s4) < 2 then do;
aevent4 = 1;
aa4 = s4; atime4 = sum(aa1 , aa2 , aa3 , aa4);
atotal_time = sum(s1 , s2 , s3 , s4);
end;
end;
end;
end;
run;

data censoring_null_hr4e (drop = a1 a2 a3 a4);
set null_exp (drop = t1 t2 t3 t4 d1 d2 d3 d4 s1 s2 s3 s4);
by k;
atotal_time = 2;
if a1 >= 2 then do;
aevent1 = 0; aevent2 = .;  aevent3 = .; aevent4 = .;
aa1 = 2; aa2 = .; aa3 = .; aa4 = .;
end;
else if . < a1 < 2 then do;
aevent1 = 1; aa1 = a1; atime1 = aa1;
if sum(a1 , a2) >= 2 then do;
aevent2 = 0; aevent3 = .; aevent4 = .;
aa2 = (2 - a1); aa3 = .; aa4 = .;
end;
else if . < sum(a1 , a2) < 2 then do;
aevent2 = 1; aa2 = a2; atime2 = sum(aa1 , aa2);
if sum(a1 , a2 , a3) >= 2 then do;
aevent3 = 0; aevent4 = .;
aa3 = (2 - sum(a1 , a2)); aa4 = .;
end;
else if . < sum(a1 , a2 , a3) < 2 then do;
aevent3 = 1; aa3 = a3; atime3 = sum(aa1 , aa2 , aa3);
if sum(a1 , a2 , a3 , a4) >= 2 then do;
aevent4 = 0;
aa4 = (2 - sum(a1 , a2 , a3));
end;
else if . < sum(a1 , a2 , a3 , a4) < 2 then do;
aevent4 = 1;
aa4 = a4; atime4 = sum(aa1 , aa2 , aa3 , aa4);
atotal_time = sum(a1 , a2 , a3 , a4);
end;
end;
end;
end;
run;

proc freq data = censoring_null_hr1e noprint;
by k null;
tables event1 / out = event1_null_hr1e;
tables event2 / out = event2_null_hr1e;
tables event3 / out = event3_null_hr1e;
tables event4 / out = event4_null_hr1e;
run;
proc freq data = censoring_null_hr2e noprint;
by k null;
tables event1 / out = event1_null_hr2e;
tables event2 / out = event2_null_hr2e;
tables event3 / out = event3_null_hr2e;
tables event4 / out = event4_null_hr2e;
run;
proc freq data = censoring_null_hr3e noprint;
by k null;
tables event1 / out = event1_null_hr3e;
tables event2 / out = event2_null_hr3e;
tables event3 / out = event3_null_hr3e;
tables event4 / out = event4_null_hr3e;
run;
proc freq data = censoring_null_hr4e noprint;
by k null;
tables event1 / out = event1_null_hr4e;
tables event2 / out = event2_null_hr4e;
tables event3 / out = event3_null_hr4e;
tables event4 / out = event4_null_hr4e;
run;

proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\IE_hr1e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\IE_hr1e.txt'
new;
run;
ods output parameterestimates =  first_events_null_hr1e;
proc phreg data = censoring_null_hr1e;
model tt1*event1(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = second_events_null_hr1e;
proc phreg data = censoring_null_hr1e;
model tt2*event3(0) = null / rl ties=exact;
by k;
```

```
run;
ods noresults;
ods output parameterestimates = third_events_null_hr1e;
proc phreg data = censoring_null_hr1e;
model tt3*event3(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = fourth_events_null_hr1e;
proc phreg data = censoring_null_hr1e;
model tt4*event4(0) = null / rl ties=exact;
by k;
run;
ods noresults;
proc printto print = print log = log;
run;

proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\IE_hr2e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\IE_hr2e.txt'
new;
run;
ods output parameterestimates = first_events_null_hr2e;
proc phreg data = censoring_null_hr2e;
model tt1*event1(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = second_events_null_hr2e;
proc phreg data = censoring_null_hr2e;
model tt2*event3(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = third_events_null_hr2e;
proc phreg data = censoring_null_hr2e;
model tt3*event3(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = fourth_events_null_hr2e;
proc phreg data = censoring_null_hr2e;
model tt4*event4(0) = null / rl ties=exact;
by k;
run;
ods noresults;
proc printto print = print log = log;
run;

proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\IE_hr3e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\IE_hr3e.txt'
new;
run;
ods output parameterestimates = first_events_null_hr3e;
proc phreg data = censoring_null_hr3e;
model aa1*aevent1(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = second_events_null_hr3e;
proc phreg data = censoring_null_hr3e;
model aa2*aevent3(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = third_events_null_hr3e;
proc phreg data = censoring_null_hr3e;
model aa3*aevent3(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = fourth_events_null_hr3e;
proc phreg data = censoring_null_hr3e;
model aa4*aevent4(0) = null / rl ties=exact;
by k;
run;
ods noresults;
proc printto print = print log = log;
run;

proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\IE_hr4e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\IE_hr4e.txt'
new;
run;
ods output parameterestimates =  first_events_null_hr4e;
proc phreg data = censoring_null_hr4e;
model aa1*aevent1(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = second_events_null_hr4e;
proc phreg data = censoring_null_hr4e;
model aa2*aevent3(0) = null / rl ties=exact;
by k;
run;
ods noresults;
ods output parameterestimates = third_events_null_hr4e;
proc phreg data = censoring_null_hr4e;
model aa3*aevent3(0) = null / rl ties=exact;
by k;
```

```
run;
ods noresults;
ods output parameterestimates = fourth_events_null_hr4e;
proc phreg data = censoring_null_hr4e;
model aa4*aevent4(0) = null / rl ties=exact;
by k;
run;
ods noresults;
proc printto print = print log = log;
run;

/*--------------------*/
/*-- No Increase No Subject --*/
/*--------------------*/
data model_null_hr1ea
(drop = time1-time4 total_time tt1 tt2 tt3 tt4 event1-event4);
set censoring_null_hr1e;
by k;
retain tstart tstop;
array tt time1-time4;
id + 1;
tstart = 0;
ev = 1;
do over tt;
ev = _i_;
if tt = . then do;
tstop = total_time;
status = 0;
end;
else do;
tstop = tt;
status = 1;
end;
output;
tstart= tstop;
end;
if (tstart < total_time) then do;
tstop = total_time;
status = 0;
ev = 4;
output;
end;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\AG_hr1e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\AG_hr1e.txt'
new;
run;
ods output parameterestimates = AG_null_hr1e;
proc phreg data = model_null_hr1ea covs(aggregate) covm;
model (tstart, tstop)*status(0) = null / rl ties=exact;
by k;
id i;
run;
proc printto print = print log = log;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\WLW_hr1e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\WLW_hr1e.txt'
new;
run;
ods output parameterestimates = WLW_null_hr1e;
proc phreg data = model_null_hr1ea covs(aggregate) covm;
strata ev;
by k; model tstop*status(0) = null / rl ties=exact;
id i;
run;
proc printto print=print log=log;
run;
data model_null_hr1eb (drop = lstatus);
retain lstatus;
set model_null_hr1ea;
by k i;
if first.i then lstatus = 1;
if (status = 0 and lstatus = 0) then delete;
lstatus = status;
gaptime = tstop - tstart;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\PWPa_hr1e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\PWPa_hr1e.txt'
new;
run;
ods output parameterestimates = PWPa_null_hr1e;
proc phreg data = model_null_hr1eb covs(aggregate) covm;
model (tstart, tstop)*status(0) = null / rl ties=exact;
by k;
strata ev;
id i;
run;
proc printto print = print log = log;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\PWPb_hr1e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\PWPb_hr1e.txt'
new;
run;
ods output parameterestimates = PWPb_null_hr1e;
proc phreg data=model_null_hr1eb covs(aggregate) covm;
model gaptime*status(0) = null / rl ties=exact;
by k;
strata ev;
```

```
id i;
run;
ods noresults;
proc printto print = print log = log;
run;

/*--------------------*/
/*-- Increasing Risk No Subject --*/
/*--------------------*/
data model_null_hr2ea
(drop = time1-time4 total_time tt1 tt2 tt3 tt4 event1-event4);
set censoring_null_hr2e;
by k;
retain tstart tstop;
array tt time1-time4;
id + 1;
tstart = 0;
ev = 1;
do over tt;
ev = _i_;
if tt = . then do;
tstop = total_time;
status = 0;
end;
else do;
tstop = tt;
status = 1;
end;
output;
tstart= tstop;
end;
if (tstart < total_time) then do;
tstop = total_time;
status = 0;
ev = 4;
output;
end;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\AG_hr2e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\AG_hr2e.txt'
new;
run;
ods output parameterestimates = AG_null_hr2e;
proc phreg data=model_null_hr2ea covs(aggregate) covm;
model (tstart, tstop)*status(0) = null / rl ties=exact;
by k;
id i;
run;
proc printto print=print log=log;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\WLW_hr2e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\WLW_hr2e.txt'
new;
run;
ods output parameterestimates = WLW_null_hr2e;
proc phreg data=model_null_hr2ea covs(aggregate) covm;
strata ev;
by k; model tstop*status(0) = null / rl ties=exact;
id i;
run;
proc printto print=print log=log;
run;
data model_null_hr2eb (drop = lstatus);
retain lstatus;
set model_null_hr2ea;
by k i;
if first.i then lstatus = 1;
if (status = 0 and lstatus = 0) then delete;
lstatus = status;
gaptime = tstop - tstart;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\PWPa_hr2e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\PWPa_hr2e.txt'
new;
run;
ods output parameterestimates = PWPa_null_hr2e;
proc phreg data=model_null_hr2eb covs(aggregate) covm;
model (tstart, tstop)*status(0) = null / rl ties=exact;
by k;
strata ev;
id i;
run;
proc printto print=print log=log;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\PWPb_hr2e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\PWPb_hr2e.txt'
new;
run;
ods output parameterestimates = PWPb_null_hr2e;
proc phreg data=model_null_hr2eb covs(aggregate) covm;
model gaptime*status(0) = null / rl ties=exact;
by k;
strata ev;
id i;
run;
ods noresults;
proc printto print=print log=log;
run;
```

```
/*--------------------*/
/*-- No Increase Subject --*/
/*--------------------*/
data model_null_hr3ea
(drop = atime1-atime4 atotal_time aa1 aa2 aa3 aa4 aevent1-aevent4);
set censoring_null_hr3e;
by k;
retain tstart tstop;
array tt atime1-atime4;
id + 1;
tstart = 0;
ev = 1;
do over tt;
ev = _i_;
if tt = . then do;
tstop = atotal_time;
status = 0;
end;
else do;
tstop = tt;
status = 1;
end;
output;
tstart= tstop;
end;
if (tstart < atotal_time) then do;
tstop = atotal_time;
status = 0;
ev = 4;
output;
end;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\AG_hr3e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\AG_hr3e.txt'
new;
run;
ods output parameterestimates = AG_null_hr3e;
proc phreg data=model_null_hr3ea covs(aggregate) covm;
model (tstart, tstop)*status(0) = null / rl ties=exact;
by k;
id i;
run;
proc printto print=print log=log;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\WLW_hr3e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\WLW_hr3e.txt'
new;
run;
ods output parameterestimates = WLW_null_hr3e;
proc phreg data=model_null_hr3ea covs(aggregate) covm;
strata ev;
by k; model tstop*status(0) = null / rl ties=exact;
id i;run;
proc printto print=print log=log;
run;
data model_null_hr3eb (drop = lstatus);
retain lstatus;
set model_null_hr3ea;
by k i;
if first.i then lstatus = 1;
if (status = 0 and lstatus = 0) then delete;
lstatus = status;
gaptime = tstop - tstart;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\PWPa_hr3e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\PWPa_hr3e.txt'
new;
run;
ods output parameterestimates = PWPa_null_hr3e;
proc phreg data=model_null_hr3eb covs(aggregate) covm;
model (tstart, tstop)*status(0) = null / rl ties=exact;
by k;
strata ev;
id i;
run;
proc printto print=print log=log;
run;
proc printto
print = 'C:\Documents and Settings\nhenderso\Desktop\PWPb_hr3e.txt'
log = 'C:\Documents and Settings\nhenderso\Desktop\PWPb_hr3e.txt'
new;
run;
ods output parameterestimates = PWPb_null_hr3e;
proc phreg data=model_null_hr3eb covs(aggregate) covm;
model gaptime*status(0) = null / rl ties=exact;
by k;
strata ev;
id i;
run;
ods noresults;
proc printto print=print log=log;
run;

/*--------------------*/
/*-- Increasing Risk Subject Effect --*/
/*--------------------*/
data model_null_hr4ea
```

```
(drop = atime1-atime4 atotal_time aa1 aa2 aa3 aa4 aevent1-aevent4);
 set censoring_null_hr4e;
 by k;
 retain tstart tstop;
 array tt atime1-atime4;
 id + 1;
 tstart = 0;
 ev = 1;
 do over tt;
 ev = _i_;
 if tt = . then do;
 tstop = atotal_time;
 status = 0;
 end;
 else do;
 tstop = tt;
 status = 1;
 end;
 output;
 tstart= tstop;
 end;
 if (tstart < atotal_time) then do;
 tstop = atotal_time;
 status = 0;
 ev = 4;
 output;
 end;
 run;
 proc printto
 print = 'C:\Documents and Settings\nhenderso\Desktop\AG_hr4e.txt'
 log = 'C:\Documents and Settings\nhenderso\Desktop\AG_hr4e.txt'
 new;
 run;
 ods output parameterestimates = AG_null_hr4e;
 proc phreg data=model_null_hr4ea covs(aggregate) covm;
 model (tstart, tstop)*status(0) = null / rl ties=exact;
 by k;
 id i;
 run;
 proc printto print=print log=log;
 run;
 proc printto
 print = 'C:\Documents and Settings\nhenderso\Desktop\WLW_hr4e.txt'
 log = 'C:\Documents and Settings\nhenderso\Desktop\WLW_hr4e.txt'
 new;
 run;
ods output parameterestimates = WLW_null_hr4e;
 proc phreg data=model_null_hr4ea covs(aggregate) covm;
 strata ev;
 by k; model tstop*status(0) = null / rl ties=exact;
 id i;
 run;
 proc printto print=print log=log;
 run;
 data model_null_hr4eb (drop = lstatus);
 retain lstatus;
 set model_null_hr4ea;
 by k i;
 if first.i then lstatus = 1;
 if (status = 0 and lstatus = 0) then delete;
 lstatus = status;
 gaptime = tstop - tstart;
 run;
 proc printto
 print = 'C:\Documents and Settings\nhenderso\Desktop\PWPa_hr4e.txt'
 log = 'C:\Documents and Settings\nhenderso\Desktop\PWPa_hr4e.txt'
 new;
 run;
 ods output parameterestimates = PWPa_null_hr4e;
 proc phreg data=model_null_hr4eb covs(aggregate) covm;
 model (tstart, tstop)*status(0) = null / rl ties=exact;
 by k;
 strata ev;
 id i;
 run;
 proc printto print=print log=log;
 run;
 proc printto
 print = 'C:\Documents and Settings\nhenderso\Desktop\PWPb_hr4e.txt'
 log = 'C:\Documents and Settings\nhenderso\Desktop\PWPb_hr4e.txt'
 new;
 run;
 ods output parameterestimates = PWPb_null_hr4e;
 proc phreg data=model_null_hr4eb covs(aggregate) covm;
 model gaptime*status(0) = null / rl ties=exact;
 by k;
 strata ev;
 id i;
run;
 ods noresults;
 proc printto print=print log=log;
 run;
```

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control 19*, 716–723.

Andersen, P. K., S. Z. Abildstrom, and S. Rosthøj (2002). Competing risks as a multi-state model. *Statistical Methods in Medical Research 11*, 203–215.

Andersen, P. K., O. Borgan, R. D. Gill, and N. Keiding (1993). *Statistical Models Based on Counting Processes*. New York: Springer-Verlag.

Andersen, P. K. and R. D. Gill (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics 10*, 1100–1120.

Andersen, P. K. and N. Keiding (2002). Multi-state models for event history analysis. *Statistical Methods in Medical Research 11*, 91–115.

Anderson, K. M., P. W. Wilson, P. M. Odell, and W. B. Kannel (1991). An Updated Coronary Risk Profile. A Statement for Health Professionals. *Circulation 83*, 356–362.

Anon (2005). Herceptin and early breast cancer: a moment of caution. *Lancet 366*, 1673.

Atar, S., C. P. Cannon, S. A. Murphy, S. Rosanio, B. F. Uretsky, and Y. Birnbaum (2006). Statins are associated with lower risk of gastrointestinal bleeding in patients with unstable coronary syndromes: Analysis of the Orbofiban in Patients with Unstable coronary Syndromes-Thrombolysis In Myocardial Infarction 16 (OPUS-TIMI 16) trial. *American Heart Journal 151*, 976.e1–976.e6.

Barrett, A., T. Roques, M. Small, and R. D. Smith (2006). How much will Herceptin really cost? *British Medical Journal 333*, 1118–1120.

Black, W. C. (1990). The CE Plane: A Graphic Representation of Cost-Effectiveness. *Medical Decision Making 10*, 212–214.

Briggs, A. and P. Fenn (1998). Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Economics 7*, 723–740.

Briggs, A. and M. Sculpher (1997). Commentary: Markov models of medical prognosis . *British Medical Journal 314*, 354.

Briggs, A. H., R. Goeree, G. Blackhouse, and B. J. O'Brien (2002b). Probabilistic Analysis of Cost-Effectiveness Models: Choosing between Treatment Strategies for Gastroesophageal Reflux Disease. *Medical Decision Making 22*, 290–308.

Briggs, A. H., B. J. O'Brien, and G. Blackhouse (2002a). Thinking Outside the Box: Recent Advances in the Analysis and Presentation of Uncertainty in Cost-Effectiveness Studies. *Annual Review of Public Health 23*, 377–401.

Brindle, P., J. Emberson, F. Lampe, M. Walker, P. Whincup, T. Fahey, and S. Ebrahim (2003). Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. *British Medical Journal 327*, 1267.

Brinsmead, R. and S. Hill (2003). Use of pharmacoeconomics in prescribing research. Part 4: is cost-utility analysis a useful tool? *Journal of Clinical Pharmacy and Therapeutics 28*, 339–346.

Brooks, R. (1996). EuroQol: the current state of play. *Health Policy 37*, 53–72.

Claxton, K., M. Sculpher, C. McCabe, A. Briggs, R. Akehurst, M. Buxton, J. Brazier, and T. O'Hagan (2005). Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Economics 14*, 339–347.

Conroy, R. M., K. Pyorala, A. P. Fitzgerald, S. Sans, A. Menotti, G. De Backer, D. De Bacquer, P. Ducimetiere, P. Jousilahti, U. Keil, I. Njolstad, R. G. Oganov, T. Thomsen, H. Tunstall-Pedoe, A. Tverdal, H. Wedel, P. Whincup, L. Wilhelmsen, and I. M. Graham on behalf of the SCORE project group (2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *European Heart Journal 24*, 987–1003.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society Series. Series B (Methodological) 34*, 187–220.

Crowder, M. (1989). A Multivariate Distribution with Weibull Connections. *Journal of the Royal Statistical Society. Series B (Methodological) 51*, 93–107.

Crowder, M. J. (2001). *Classical Competing Risks*. Boca Raton: Chapman & Hall/CRC.

D'Agostino, R. B., M. W. Russell, D. M. Huse, R. C. Ellison, H. Silbershatz, P. W. F. Wilson, and S. C. Hartz (2000). Primary and subsequent coronary

risk appraisal: New results from The Framingham Study. *American Heart Journal 139*, 272–281.

Dear, J. W., P. Lilitkarntakul, and D. J. Webb (2006). Are rare diseases still orphans or happily adopted? the challenges of developing and using orphan medicinal products. *British Journal of Clinical Pharmacology 62*, 264–271.

Dent, R. and M. Clemons (2005). Adjuvant trastuzumab for breast cancer. *British Medical Journal 331*, 1035–1036.

Di Somma, S., V. Liguori, M. Petitto, A. Carotenuto, D. Bokor, O. de Divitiis, and M. de Divitiis (1993). A double-blind comparison of nicorandil and metoprolol in stable angina pectoris. *Cardiovascular Drugs and Therapy 7*, 119–123.

Döring, G. (1992). Antianginal and antiischemic efficacy of niocrandil in comparison with isosorbide-5-mononitrate and isosorbide dinitrae - results from 2 multicenter, double-blind, randomized studies with stable coronary heart disease patients. *Journal of Cardiovascular Pharmacology 20(suppl 3)*, S74–S81.

Efron, B. and R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall/CRC.

Fenwick, E., B. J. O'Brien, and A. Briggs (2004). Cost-effectiveness acceptability curves - facts, fallacies and frequently asked questions. *Health Economics 13*, 405–405.

Fieller, E. C. (1954). Some Problems in Interval Estimation. *Journal of the Royal Statistical Society. Series B (Methodological) 16*, 175–185.

Ford, I., N. J. K. Henderson, and H. J. Dargie (2007). Nicorandil Prevents Initial and Recurrent Events in Patients With Stable Angina Pectoris. *Journal of the American College of Cardiology 49(suppl A)*, 238A.

Ford, I., J. Norrie, and S. Ahmadi (1995). Model Inconsistency, Illustrated by the Cox Proportional Hazards Model. *Statistics in Medicine 14*, 735–746.

Gandhi, M. M. (1997). Clinical epidemiology of coronary heart disease in the UK. *British Journal of Hospital Medicine 58*, 23–27.

Hanley, J. A. and B. J. McNeil (1982). The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology 143*, 29–36.

Harrell, F. E., R. M. Califf, D. B. Pryor, K. L. Lee, and R. A. Rosati (1982). Evaluating the yield of medical tests. *Journal of the American Medical Association 247*, 2543–2546.

Heart Protection Study Collaborative Group (2002). MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20 536 high-risk individuals: a randomised placebo-controlled trial. *Lancet 360*, 7–22.

Heart Protection Study Collaborative Group (2005). Cost-effectiveness of simvastatin in people at different levels of vascular disease risk: economic analysis of a randomised trial in 20 536 individuals. *Lancet 365*, 1779–1785.

Heart Protection Study Collaborative Group (2006). Lifetime cost effectiveness of simvastatin in a range of risk groups and age groups derived from a randomised trial of 20 536 people. *British Medical Journal 333*, 1145 doi: 10.1136/bmj.38993.731725.BE.

Heitjan, D. F., A. J. Moskowitz, and W. Whang (1999). Problems with Interval Estimates of the Incremental Cost-Effectiveness Ratio. *Medical Decision Making 19*, 9–15.

Henderson, N. J. K., I. Ford, and A. Walker (2005). Balancing Harm, Benefit and Risk in Clinical Trials. *Clinical Trials 2(suppl 1)*, S58.

Hjalmarson, A. (2007). Heart rate: an independent risk factor in cardiovascular disease. *European Heart Journal 9(suppl F)*, F3–F7.

Hosmer, D. W. and S. Lemeshow (1999). *Applied Survival Analysis: Regression Modeling of Time to Event Data*, Chapter 9: Other Models and Topics, pp. 308–317. New York: John Wiley & Sons.

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika 73*, 387–396.

Howard, R. A. (1988). Decision Analysis: Practice and Promise. *Management Science 34*, 679–695.

Information and Statistics Divison, NHS Scotland (2002). Scottish Health Service Costs 2001/2. http://www.isdscotland.org/isd/info3.jsp?pContentID=2686&p_applic=CCC&p_service=Content.show&.

Johannesson, M., B. Jönsson, J. Kjekshus, A. G. Olsson, T. R. Pedersen, and H. Wedel for the The Scandinavian Simvastatin Survival Study Group (1997). Cost Effectiveness of Simvastatin Treatment to Lower Cholesterol Levels in Patients with Coronary Heart Disease. *New England Journal of Medicine 336*, 332–336.

Jönsson, B., J. R. Cook, and T. R. Pedersen (1999). The cost-effectiveness of lipid lowering in patients with diabetes: results from the 4S trial. *Diabetologia 42*, 1293–1301.

Kaplan, E. L. and P. Meier (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association 53*, 457–481.

Kim, J., T. Meade, and A. Haines (2006). Skinfold thickness, body mass index, and fatal coronary heart disease: 30 year follow up of the Northwick Park heart study. *Journal of Epidemiology and Community Health 60*, 275–279.

Krum, K. and J. J. McMurray (2002). Statins and Chronic Heart Failure: Do We Need a Large-Scale Outcome Trial? *Journal of the American College of Cardiology 39*, 1567–1573.

Lang, D. L., R. Lopert, and S. R. Hill (2003). Use of pharmacoeconomics in prescribing research. Part 5: modelling - beyond clinical trials. *Journal of Clinical Pharmacy and Therapeutics 28*, 433–439.

Lin, D. Y. and L. J. Wei (1989). The Robust Inference for the Cox Proportional Hazards Model. *Journal of the American Statistical Association 84*, 1074–1078.

Lopert, R., D. L. Lang, and S. R. Hill (2003). Use of pharmacoeconomics in prescribing research. Part 3: cost-effectiveness analysis - a technique for decision-making at the margin. *Journal of Clinical Pharmacy and Therapeutics 28*, 243–249.

Mauskopf, J. A., J. E. Paul, D. M. Grant, and A. Stergachis (1998). The Role

of Cost-Consequence Analysis in Healthcare Decision-Making. *PharmacoEconomics 13*, 277–288.

McDonagh, M. S., L. M. Bachmann, S. Golder, J. Kleijnen, and G. ter Riet (2000). A rapid and systematic review of the clinical effectiveness and cost-effectiveness of glycoprotein iib/iiia antagonists in the medical management of unstable angina. *Health Technology Assessment 4*, 1–109.

Meads, C., C. Cummins, K. Jolly, A. Stevens, A. Burls, and C. Hyde (2000). Coronary artery stents in the treatment of ischaemic heart disease: a rapid and systematic review. *Health Technology Assessment 4*, 1–165.

Metalfe, C. and S. G. Thompson (2006). The importance of varying the event generation process in simulation studies of statistical methods for recurrent events. *Statistics in Medicine 25*, 165–179.

Netten, A., T. Rees, and G. Harrison (2001). Unit Costs of Health and Social Care 2001. http://www.pssru.ac.uk/pdf/UC2001/UnitCosts2001ALL.pdf.

Newby, D. and S. Hill (2003). Use of pharmacoeconomics in prescribing research. Part 2: cost-minimization analysis - when are two therapies equal? *Journal of Clinical Pharmacy and Therapeutics 28*, 145–150.

NICE (2004). Guide to the Methods of Technology Appraisal. http://www.nice.org.uk/page.aspx?o=201973.

O'Brien, B. J., M. F. Drummond, R. J. Labelle, and A. Willan (1994). In Search of Power and Significance: Issues in the Design and Analysis of Stochastic Cost-Effectiveness Studies in Health Care. *Medical Care 32*, 150–163.

O'Hagan, A., J. W. Stevens, and J. Montmartin (2000). Inference for the Cost-Effectiveness Acceptability Curve and Cost-Effectiveness Ratio. *PharmacoEconomics 17*, 339–349.

Pasta, D. J., J. L. Taylor, and J. M. Henning (1999). Probabilistic Sensitivity Analysis Incorporating the Bootstrap: An Example Comparing Treatments for the Eradication of Helicobacter pylori. *Medical Decision Making 19*, 353–363.

Prentice, R. L., B. J. Williams, and A. V. Peterson (1981). On the regression analysis of multivariate failure time data. *Biometrika 68*, 373–379.

Putter, H., M. Fiocco, and R. B. Geskus (2007). Tutorial in biostatistics: Competing risks and multi-state. *Statistics in Medicine 26*, 2389–2430.

Raftery, J. (1999). Methodological limitations of cost-effectiveness analysis in health care: implications for decision making and service provision. *Journal of Evaluation in Clinical Practice 5*, 361–366.

Richy, F., O. Bruyere, O. Ethgen, V. Rabenda, G. Bouvenot, M. Audran, G. Herrero-Beaumont, A. Moore, R. Eliakim, M. Haim, and J.-Y. Reginster (2004). Time dependent risk of gastrointestinal complications induced by non-steroidal anti-inflammatory drug use: a consensus statement using a meta-analytic approach. *Annals of the Rheumatic Diseases 63*, 759–766.

Robinson, M., L. Ginnelly, M. Sculpher, L. Jones, R. Riemsma, S. Palmer, Z. Philips, and J. Glanville (2002). A systematic review update of the clinical effectiveness and cost-effectiveness of glycoprotein iib/iiia antagonists. *Health Technology Assessment 6*, 1–171.

Scandinavian Simvastatin Survival Study Group (1994). Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study Group. *Lanet 344*, 1383–1389.

Severens, J. L., T. M. De Boo, and E. M. Konst (1999). Uncertainty of incremental cost-effectiveness ratios. A comparison of Fieller and bootstrap confidence intervals. *International Journal of Technology Assessment in Health Care 15*, 608–614.

Shaper, A. G., G. Wannamethee, P. W. Macfarlane, and M. Walker (1993). Heart rate, ischaemic heart disease, and sudden cardiac death in middle-aged British men. *British Heart Journal 70*, 49–55.

Stinnett, A. A. and J. Mullahy (1998). Net Health Benefits. *Medical Decision Making 18*, S68–S80.

Sugden, R. and A. Williams (1979). *The Principles of Practical Cost-Benefit Analysis*. Oxford: Oxford University Press.

The IONA Study Group (2001). Trial to show the impact of nicorandil in angina (IONA): design, methodology, and management. *Heart 85*, e9.

The IONA Study Group (2002). Effect of nicorandil on coronary events in patients with stable angina: Impact of nicorandil in angina (IONA) randomised trial. *Lancet 359*, 1269–1275.

The IONA Study Group (2005). Determinants of coronary events in patients with stable angina: Results from the Impact of Nicorandil in Angina Study. *American Heart Journal 150*, 689.e1–689.e9.

The SWAN Study Group (1999). Comparison of the antiischaemic and antianginal effects of nicorandil and amlodipine in patients with symptomatic stable angina pectoris: the SWAN study. *Journal of Clinical and Basic Cardiology 2*, 213–217.

Torrance, G. W. (1976). Social preferences for health states: An empirical evaluation of three measurement techniques. *Socio-Economic Planning Sciences 10*, 129–136.

Torrance, G. W., W. H. Thomas, and D. L. Sackett (1972). A utility maximization model for evaluation of health care programmes. *Health Services Research 7*, 118–133.

UK Prospective Diabetes Study Group (1998). Cost effectiveness analysis of improved blood pressure control in hypertensive patients with type 2 diabetes: UKPDS 40. *British Medical Journal 317*, 720–726.

van Hout, B. A., M. J. Al, G. S. Gordon, and F. F. H. Rutten (1994). Costs, Effects and C/E-Ratios Alongside a clinical Trial. *Health Economics 3*, 309–319.

Verhamme, K., G. Mosis, J. Dieleman, B. Stricker, and M. Sturkenboom (2006). Spironolactone and risk of upper gastrointestinal events: population based case-control study. *British Medical Journal 333*, 330.

von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Walker, A., J. McMurray, S. Stewart, W. Berger, A. D. McMahon, H. Dargie,

K. Fox, S. Hillis, N. J. K. Henderson, and I. Ford (2006). Economic evaluation of the impact of nicorandil in angina (IONA) trial. *Heart 92*, 619–624.

Wei, L. J., D. Y. Lin, and L. Weissfeld (1989). Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. *Journal of the American Statistical Association 84*, 1065–1073.

Weil, J., M. J. S. Langman, P. Wainwright, D. H. Lawson, M. Rawlins, R. F. A. Logan, T. P. Brown, M. P. Vessey, M. Murphy, and D. G. Colin-Jones (2000). Peptic ulcer bleeding: accessory risk factors and interactions with non-steroidal anti-inflammatory drugs. *Gut 46*, 27–31.

Willan, A. R. (2001). Analysis, Sample Size, and Power for Estimating Incremental Net Health Benefit from Clinical Trial Data. *Controlled Clinical Trials 22*, 228–237.