THE UNIVERSITY *of* EDINBURGH

Efficient analysis of ordinal data from clinical trials in head injury

Gillian S McHugh

PhD – The University of Edinburgh – 2012

Preface

I, Gillian McHugh, declare that the following thesis has been composed by me and has not been submitted for any other degree or professional qualification.

This thesis was part of a larger project. Grant support for the IMPACT study was provided by NIH NS 42691. This thesis is my own work, however as it is part of a larger project, some results provided by others have been incorporated. For example, I personally only extracted data from three of the eleven studies in the IMPACT database. Where this has been done it is acknowledged within the thesis.

Acknowledgements

I would like to thank Professor Gordon Murray for his excellent supervision, advice and guidance. I am very grateful for having this opportunity.

This project was part of a collaboration and could not have been done without my IMPACT colleagues in Richmond, Rotterdam and Antwerp. It certainly could not have been done without the Edinburgh 'team' and I thank them for their support and friendship.

Thank you to my mother for her enthusiasm. To John, thank you for all of your encouragement and support, as well as doing all the things I would have been doing had I not been doing this. Finally, to Natalie, thank you for making me smile as I've been doing my "peachy - dee".

Abstract

Many promising Phase II trials have been carried out in head injury however to date there has been no successful translation of the positive results from these explanatory trials into improved patient outcomes in Phase III trials. Many reasons have been hypothesised for this failure. Outcomes in head injury trials are usually measured using the five point Glasgow Outcome Scale. Traditionally the ordinality of this scale is disregarded and it is dichotomised into two groups, favourable and unfavourable outcome. This thesis explores whether suboptimal statistical analysis techniques, including the dichotomisation of outcomes could have contributed to the reasons why Phase III trials have been unsuccessful.

Based on eleven completed head injury studies, simulation modelling is used to compare outcome as assessed by the conventional dichotomy with both modelling that takes into account the ordered nature of the outcome (proportional odds modelling) and modelling which individualises a patient's risk of a good or poor outcome (the 'sliding dichotomy'). The results of this modelling show that both analyses which use the full outcome scale and those which individualise risk show great efficiency gains (as measured by reduction in required sample sizes) over the conventional analysis of the binary outcome. These results are consistent both when the simulated treatment effects followed a proportional odds model and when they did not. Consistent results were also observed when targeting or restricting improvement to groups of subjects based on clinical characteristics or prognosis. Although proportional odds modelling shows consistently greater sample size reductions the choice of whether to use proportional odds modelling or the sliding dichotomy depends on the question of interest.

Contents

Tables

Figures

Abbreviations

AIDS	Acquired Immune Deficiency Syndrome
AUC	Area Under Curve
CRASH	Corticosteroid Randomisation After Significant Head injury
CRF	Case Record Form
CT	Computerised Tomography
EBIC	European Brain Injury Consortium
ECASS	European Cooperative Acute Stroke Studies
EDH	Epidural Haematoma
GAIN	Glycine Antagonist In Neuroprotection
GCS	Glasgow Coma Scale
GEE	Generalised Estimating Equations
GOS	Glasgow Outcome Scale
GOSE	Glasgow Outcome Scale Extended
ICP	Intracranial Pressure
IMPACT	International Mission on Prognosis and Analysis of Clinical Trials in TBI
MABP	Mean Arterial Blood Pressure
MCAR	Missing Completely At Random
MI	Myocardial Infarction
MRI	Magnetic Resonance Imaging
mRS	modified Rankin Scale
NIH	National Institutes of Health
NIHSS	National Institutes of Health Stroke Scale
NINDS	National Institute of Neurological Disorders and Stroke

NMDA	N-Methyl-D-aspartate
OAST	Optimising Analysis of Stroke Trials
PAIS	Paracetamol (Acetaminophen) In Stroke
PEGSOD	Polyethylene glycol-conjugated bovine superoxide dismutase
PRoFESS	Prevention Regimen For Effectively avoiding Second Strokes
RCT	Randomised Controlled Trial
ROC	Receiver Operating Characteristic
RSS	Reduction in Sample Size
SAINT	Stroke Acute Ischemic NXY Treatment
SBP	Systolic Blood Pressure
SCAST	Scandinavian Candesartan Acute Stroke Trial
SDH	Subdural Haematoma
SLIN	Selfotel International
STICH	Surgical Trial in Intracerebral Haemorrhage
TBI	Traumatic Brain Injury
TCDB	Traumatic Coma Data Bank
TINT	Tirilazad International
TIUS	Tirilazad United States
TOAST	Trial of Org 10172 in Acute Stroke Treatment
t-PA	Tissue Plasminogen Activator
TPDS	Top Priority Data Set
tSAH	Traumatic Subarachnoid Haemorrhage
UK4	UK Four Centres study
VISTA	Virtual International Stroke Trials Archive
WHO	World Health Organisation

1 Chapter 1 Background and Introduction

1.1 Introduction

Given the difficulties often experienced in recruiting patients into randomised controlled trials (RCTs) we need more sensitive methods of analysis, so that fewer subjects are required. For example, a review of 114 trials (McDonald et al. 2006) reported that less than a third of these trials achieved their original recruitment target. We also need more efficient trial designs to allow us to maximise the chance of detecting any treatment benefits that may exist.

Qureshi et al (Qureshi et al. 2004) wrote an excellent article on design methods for evaluating treatments for stroke covering randomisation, intention to treat analysis etc. However, even although he does discuss the ascertainment of outcomes, analysing outcomes scales as they were originally recorded rather than, as a typically done, collapsing into fewer categories is not discussed.

It has been postulated that large trials based on synthesised evidence from previous studies are needed to find moderate but worthwhile effects (Yusuf, Collins, & Peto 1984). Choi and Bullock (Choi & Bullock 2001) subsequently agreed with the philosophy: "There is currently a statistical ground swell of opinion calling for much larger trials in head injury, directed at much smaller magnitudes of effect, as has been the case with successful stroke trials." However, as McDonald et al (McDonald et al. 2006) have shown, few trials, especially those in head injury, achieve their original recruitment target. There has been one large 'mega-trial' in head injury in the last decade, Corticosteroid Randomisation After Significant Head injury (CRASH) (Edwards et al. 2005), which followed the "keep it simple" philosophy with the aim to recruit as many patients as possible with minimal data recording. This trial did demonstrate a statistically significant treatment effect, but unfortunately the effect of the experimental intervention was to ***increase*** mortality. Perhaps

suggesting that the large simple trial is not the answer to the problems facing trials in head injury.

1.2 Head injury - epidemiology

Traumatic Brain Injury (TBI) is the leading cause of death and disability among young adults in developed countries and the incidence in the elderly population is increasing (Kannus, Palvanen, & Niemi 2001;Luukinen et al. 1999;Vink & Bullock 2010). Approximately 1.5 million people die each year world wide from traumatic brain injury with most of the deaths occurring in the developing world (Menon 2009). In the United States approximately 2 million head injuries occur each year. The direct and indirect costs of these injuries are estimated to be $56 billion per annum in the United States. The World Health Organisation (WHO) has projected that by 2020, road traffic accidents, a major cause of traumatic brain injury, will rank third as a cause of the global burden of disease and disablement, behind only ischemic heart disease and unipolar major depression (Finfer & Cohen 2001).

1.3 Head injury - classification

Head injury is heterogeneous with different pathological processes and causes (Guha 2004). Injury severity can be classified in several fundamentally different ways: using the degree of impairment of consciousness - typically measured by the Glasgow Coma Scale (GCS); measuring evidence of structural brain damage - observed on a scan; and measuring at the cellular level the underlying cytotoxic cascade which follows brain injury. Saatman et al (Saatman et al. 2008) set out a pathoanatomic classification, this describes the anatomical features, location or the abnormality to be treated. The Marshall computerised tomography (CT) classification is one such classification (Marshall et al. 1991). It categorises patients based on their CT scan with categories going from Diffuse Injury I -No visible intracranial pathology seen on CT scan to having a mass lesion. More detail is given

in Table 9-3 in the Appendix. However this approach does not take into account the cytotoxic cascade which has been shown to occur in human head injury, validating the previous results observed in animal models (Hovda et al. 1995). A cellular approach rather than a CT approach may be the way forward for developing new drug treatments.

1.4 Head injury - outcome measures

The two most commonly used measures to assess outcome after head injury are the Glasgow Coma Scale (GCS), used to asses a patient's degree of impairment of consciousness soon after the initial insult, and the Glasgow Outcome Scale (GOS), used to assess longer term functional outcome.

The Glasgow Coma Scale (Teasdale & Jennett 1974) comprises three components: an eye response score ranging from 'one-no eye opening' to 'four-spontaneous eye opening'; a verbal response score ranging from 'one-no response' to 'five – orientated' and a motor response score ranging from 'one-no response' to 'six-obeys'. These components are then summed to give a total Glasgow Coma Scale score of 3 to 15. Fuller details of the GCS are given in the Appendix in Table 9-2.

The Glasgow Outcome Scale, (Jennett & Bond 1975), is a five point ordinal scale with categories of good recovery, moderate disability, severe disability, vegetative state and death, conventionally assessed six months after injury. Fuller details of the scoring of this scale are given in the Appendix in Table 9-1.

The components of the Glasgow Coma Scale and the Glasgow Outcome Scale are all ordinal scales. Ordinal data occur when data are classified into a small number of

ordered categories and the ordering has no numerical value. For example, the Glasgow Outcome Scale with its categories of good recovery, moderate disability, severe disability, vegetative state and death is clearly ordered although no numerical value can be assigned to the ordering.

Traditionally, the GOS has been dichotomised into a favourable outcome group, combining the good recovery and moderate disability categories, and an unfavourable outcome group, combining the severe disability, vegetative state and death categories. Dichotomising the GOS does reflect clinical judgement as to what constitutes a 'desirable' outcome. It also has the advantage of allowing relatively unsophisticated techniques to be used to analyse the data which give an easily understood outcome. However, dichotomising the scale in this way does not exploit its underlying ordinal nature. The dichotomisation may also be insensitive in patients with more favourable outcomes. The effects of varying the threshold between "good" and "poor" outcome for different prognosis groups will be examined in this thesis. This is the concept of the "sliding dichotomy" as first proposed by Barer, (Barer 1999); more detail on this is given in subsequent chapters. Analysis across the distribution of scores maximises the use of the scale (Kasner 2006). Dichotomising into favourable and unfavourable does not reflect the clinical benefit that moving up or down the scale has (Menon 2009). Also, analysing a binary outcome forces trialists to discard large quantities of outcome information which can lead to underestimation of treatment benefit or harm or both (Saver 2007).

A limitation of current analysis of the GOS and indeed the Disability Rating Scale (Rappaport et al. 1982), also used to measure outcome after head injury; and the Barthel Index (Mahoney & Barthel 1965) and the modified Rankin Scale (mRS) (van Swieten et al. 1988), used to measure outcome after stroke, is that the criteria for assessment do not specify how pre-existing physical or psychological problems should be taken into account (Pettigrew, Wilson, & Teasdale 1998). The sliding

dichotomy, although it does not take into account pre-injury morbidities, individualises the prognosis for each patient rather than seeing if they achieve an arbitrary "good" outcome. These outcome scales are also often inappropriately dichotomised. It is remarked, (Lai & Duncan 2001), that the definition of favourable outcomes should include the transition in the modified Rankin Scale rather than the difference either side of a cutpoint. Using the full range of a scale to determine an outcome that is clinically meaningful is much more appropriate than an arbitrary dichotomisation. Treatment effect will also more likely be captured by using a shift analysis rather than a dichotomisation (Schabitz & Fisher 2006). As with all analyses, selection of the primary endpoint is crucial in detecting differences between interventions (Kasner 2006).

This thesis will primarily focus on moderate to severe head injury and explore whether using a more sophisticated analysis 'adds value' and provides an intuitive and useful outcome or whether by using a more sophisticated analysis, unnecessary complexity is added without efficiency gains and a simple dichotomous analysis provides the most useful summary.

1.5 Head injury – failure of previous studies to show an effect

Many trials in TBI have been carried out over the last three decades and the reasons given for them failing to show efficacy are varied. A series of articles have been written citing the failure of neuroprotective drugs to work in confirmatory (Phase III) trials in head injury (Dickinson et al. 2000;Maas et al. 1999). At an initial clinical level, the data available from Phase I and Phase II trials may not be reliable; the results obtained in explanatory trials may not be extrapolated to pragmatic trials or the drug may simply not work (Maas, Roozenbeek, & Manley 2010). Also at the clinical level, there is the juxtaposition between the desire for early treatment and the need to obtain consent and fulfil inclusion criteria prior to commencing treatment. At a mechanistic level, the TBI patient population is very heterogeneous with very

differing mechanisms of injury and differing prognostic risk (Bullock, Lyeth, & Muizelaar 1999;Maas 2000;Narayan et al. 2002). Doppenberg et al (Doppenberg, Choi, & Bullock 1997) hypothesised that poor brain penetration of drug was the reason for the failure of a large number of Phase III trials. There is also the requirement to test new drugs on healthy volunteers, a requirement not needed for other diseases with high morbidity and mortality, such as cancer chemotherapy or Acquired Immune Deficiency Syndrome (AIDS) therapy (Doppenberg, Choi, & Bullock 1997). A review of over one thousand neuroprotectants found little relationship in experimental effectiveness between drugs that were effective clinically and those tested only in animal models (O'Collins et al. 2006). It is still not known if the mechanisms that cause secondary injury in humans are the same as those in animals and, even if they are, whether these findings can be translated from bench to bedside (Bullock, Lyeth, & Muizelaar 1999;Ioannidis 2006).

It has been hypothesised that the failure of head injury studies to translate from exploratory Phase II trials to confirmatory Phase III may be to do with the population being studied but could equally and even more so be to do with the design and analysis of the trials carried out so far (Teasdale et al. 1999). One of the key requirements for any trial is that the outcome must be sufficiently sensitive to detect a clinically meaningful effect of the intervention (Bullock, Lyeth, & Muizelaar 1999). With all trials, but fundamentally with head injury trials, the design has to match the question being asked (Murray & Teasdale 2000). At the analysis stage, there are problems with the use of surrogate endpoints and the dichotomisation of the outcome measure, usually the GOS which leads to a loss of sensitivity. There are a plethora of factors which affect outcome after head injury, therefore head injury lends itself to statistical modelling (Helmy, Timofeev, & Hutchinson 2010). The different prognostic risk factors and the analysis of the outcome measures are explored within this thesis.

A partial explanation of the failure of Phase III trials may also be observer variation. The Glasgow Outcome Scale and the extended Glasgow Outcome Scale (GOSE) (Teasdale et al. 1998) have both been found to have significant observer variation (Choi et al. 2002;Lu et al. 2008;Lu et al. 2010;Wilson et al. 2007). The GOSE, an extension of the GOS, is an eight point scale with the categories of good recovery, moderate disability and severe disability being split into upper and lower bands. Wilson et al (Wilson et al. 2007), in examining misclassification of the GOSE, found that the greatest levels of disagreement were found between the categories of upper severe disability and lower moderate disability traditionally where the dichotomisation is made between good and poor outcomes. Whitehead (Whitehead 1993) showed that for a four category outcome, misclassification of 20% could lead to the same loss of efficiency as having a binary outcome.

Other factors to consider in terms of the efficiency of a confirmatory trial in traumatic brain injury are the entry criteria and the use of covariate adjustment. Machado et al (Machado, Murray, & Teasdale 1999) showed that by only targeting the intermediate prognosis patients sample size savings of up to 30% could be achieved. However this was when using a binary rather than an ordinal outcome analysis. Roozenbeek et al (Roozenbeek et al. 2009b) showed that by adopting more strict enrolment criteria, as expected, a more homogeneous population is recruited which can lead to a more efficient trial but recruitment is slowed to such an extent that the manoeuvre becomes self defeating. Others have shown the benefits of adding covariates into the model (Hauck, Anderson, & Marcus 1998;Hernandez, Steyerberg, & Habbema 2004;Roozenbeek et al. 2009a).

1.6 Literature review

1.6.1 Ordinal methods of analysis in the literature

Four methods of ordinal analysis have been focused on in the literature to varying degrees: the proportional odds model (McCullagh 1980); the continuation ratio model (Fienberg 1980); the stereotype (alternatively named canonical regression) model (Anderson 1984) and the latent normal model (Poon 2004). Both the proportional odds and continuation ratio approaches can be regarded as grouped continuous. That is they use discrete versions of an underlying continuous outcome variable (Anderson & Philips 1981;Feldmann & Steudel 2000). The proportional odds and the latent normal model both assume that there is an underlying latent variable. However, the proportional odds model assumes that the distribution of the variables are logistic whereas the latent normal model assumes that the underlying distribution is normal (Poon 2004). The stereotype model assumes that several multidimensional factors contribute to the scale. More details on these models and the extensions to clustered or repeated ordinal data are reported (Liu & Agresti 2005). The proportional odds model has most commonly been used in applied rather than theoretical settings.

1.6.2 Search strategy – head injury and stoke trials and ordinal methods in the literature

Literature searches were undertaken to identify: 1) all Phase III trials in head injury and 2) all Phase III stroke studies that had used an ordinal method of analysis.

Stroke studies were included in the search as head injury and stroke have many similar features. The most universally used description of a stroke is " a clinical syndrome characterised by rapidly developing clinical signs of focal, at times global (applied to patients in deep coma and those with subarachnoid haemorrhage), loss of

cerebral function, with symptoms lasting more than 24h or leading to death, with no apparent cause other than that of vascular origin" (Hatano 1976).

Further literature searching was done to identify studies which had adjusted for baseline severity. Studies that had cited either Barer's (Barer 1999) original paper that mentioned the concept of the sliding dichotomy or Murray et al's (Murray et al. 2005) paper which expanded on the topic in more detail were identified. The term "sliding dichotomy" was also searched for. Alternative methods of adjusting for baseline severity such as patient-specific and responder analysis were also searched for. The term "baseline severity adjusted end point analysis" was also searched for. Also papers that had cited Saver's (Saver 2007) paper on this topic were searched for.

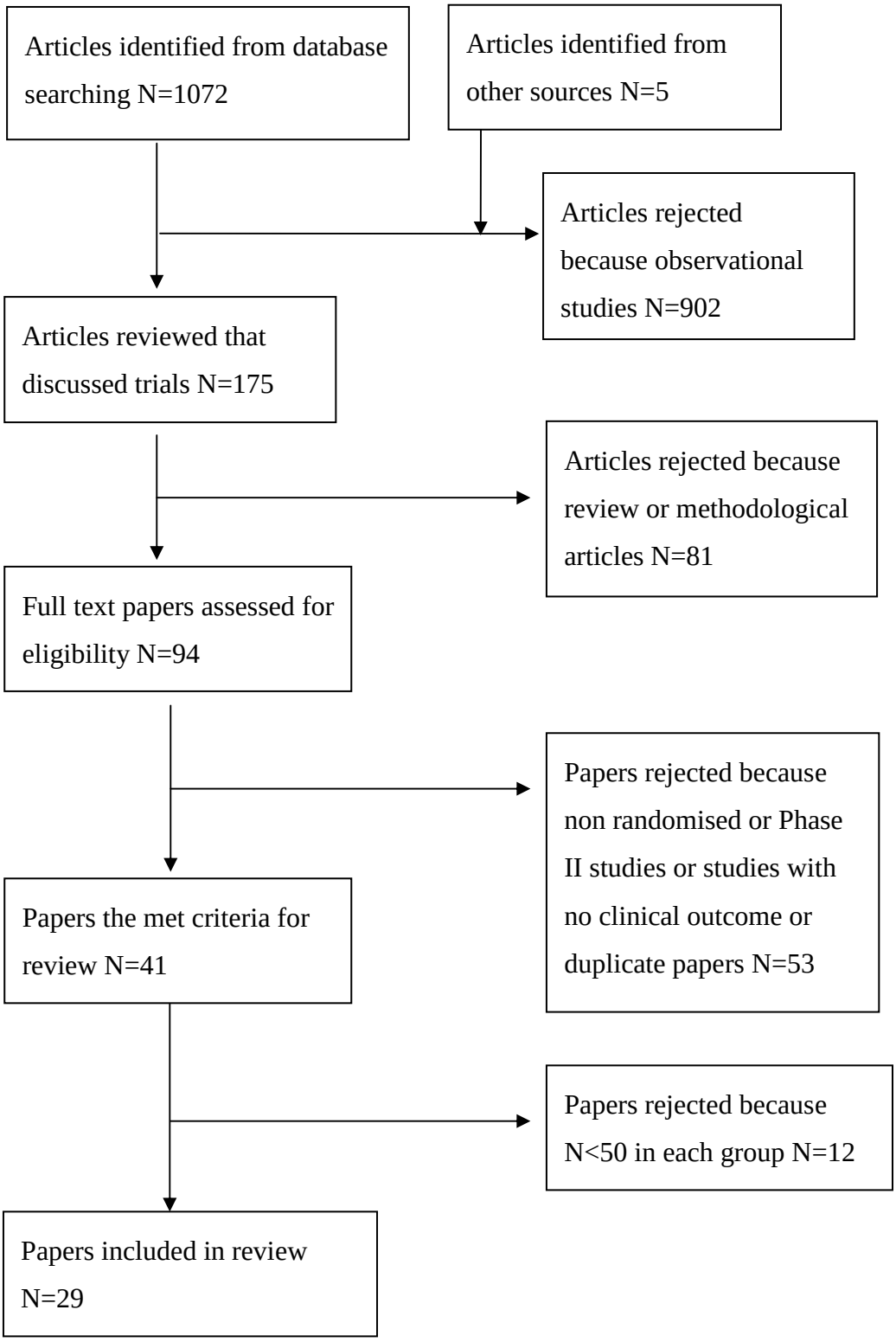
Papers discussing responder analysis were searched for and more detailed searches were done on responder analysis and stroke, head injury or brain injury.

1.6.3 Phase III trials in head injury

In order to establish what outcome measures are currently used in Phase III trials in head injury and how these outcome measures are analysed a systematic review was carried out. Particular attention was placed on whether the outcome was dichotomised or not. PubMed, Medline, the Cochrane Library and Web of Knowledge were searched for head or brain or cranial injur* and Phase III or Phase 3 or trial* or random*. An asterisk indicating that the search term was truncated on the right. The search was last updated in September 2011.

The results of the search strategy are summarised in the flow chart below, Figure 1-1

Figure 1-1 Flow chart showing references considered for Phase III trials in head injury



The first stage was to examine the abstracts of the papers found in the literature search. Although the word random* was used as one of the search terms when some of the studies were examined in more detail they were found to be observational studies. The full paper was then obtained for any relevant study and examined carefully to see if a 'clinical' outcome was measured e.g. GOS or mortality. For this review any studies with laboratory measures as the outcome measure e.g. increase in intracranial Pressure (ICP) were not included. As a final stage in the review process, only studies which included at least 50 subjects in each arm were included.

As the flow chart shows, 29 references were found which related to Phase III trials in head injury. These were all studies with a clinical outcome measure and with greater than or equal to 50 subjects in each arm. Details of these trials are given in Table 1-1. Most of the trials used the Glasgow Outcome Scale to assess outcome. Three trials used the extended GOS (Cooper et al. 2004;Maas et al. 2006;Temkin et al. 2007) and one used a modification of the GOS with six points, distinguishing between severe disability requiring constant care and severe disability without the need for constant care (Gaab et al. 1994). Twenty of the trials recorded GOS at six months. GOS was reported at three months only by Xiao and Young (Xiao et al. 2008;Young et al. 1996); twelve months only by Gaab (Gaab et al. 1994) and at multiple time points up to one year for the remaining six studies.

The GOS was dichotomised before being analysed in 18 of the trials. Fifteen of these trials used the conventional dichotomy and combined the good recovery and moderate disability categories into a favourable outcome group and the severe disability, vegetative state and death categories into an unfavourable group. One trial (Saul et al. 1981) considered good recovery, moderate disability or severe disability as a favourable outcome and vegetative state or death as an unfavourable outcome. Maas et al (Maas et al. 2006) used a prognosis based dichotomy (sliding dichotomy)

and it is unclear where Gaab and colleagues (Gaab et al. 1994) dichotomised their modification of the GOS.

Table 1-1 Phase III trials in head injury[1]

Trial	Year	Population	Outcome	N
Saul	1981	GCS<=7. US, single centre	GOS at 6 months	100
Braakman	1983	Comatose patients. Europe, 2 centres	GOS at 6 months	161
Dearden	1986	Severe head injury. UK, single centre	GOS at 6 months	130
Bailey (HIT I)	1991	Not obeying commands. Europe, 6 centres	GOS at 6 months	351
Muizelaar	1991	GCS<=8. US, single centre	GOS at 3, 6 and 12 months	113
Rockswold	1992	GCS<=9. US, single centre	GOS at 6, 12 and 18 months	168
Wolf	1993	GCS<=8. US, 2 centres	GOS at 3, 6 and 12 months	149
Gaab	1994	GCS<=13. Europe, 10 centres	modified GOS at twelve months	300
ESGN (HIT II)	1994	Not obeying commands. Europe, 13 centres	GOS at 6 months	852
Grumme	1995	Severe head injury. Europe, 9 centres	GOS at discharge and twelve months	396
Harders (HIT III)	1996	tSAH Germany, 21 centres	GOS at 6 months	123

[1] Abbreviations in table: GCS - Glasgow Coma Scale; GOS – Glasgow Outcome Scale; GOSE – Glasgow Outcome Scale Extended; tSAH – traumatic subarachnoid haemorrhage; ESGN – The European Study Group on Nimodipine in Severe Head Injury

Trial	Year	Population	Outcome	N
Young (PEGSOD)	1996	GCS<=8. US, 29 centres	GOS at 3 months	463[2] (1574)
Marshall (TINT)	1998	85% GCS<=8, 15% GCS 9-12. Worldwide, 50 centres	GOS at 6 months	1131
Marmarou (BRADYCOR)	1999	GCS <=8. US, 31 centres	GOS at 3 and 6 months	139
Morris (SELFOTEL)	1999	GCS 4-8. Worldwide, 99 centres	GOS at 6 months	693
Robertson	1999	GCS<=5. US, single centre	GOS at 3 and 6 months	189
Clifton (NABIS)	2001	GCS 3-8. US, 11 centres	Mortality. GOS at 6 months	392
Cruz	2001	Acute subdural haematoma. Brazil, single centre	GOS at 6 months	178
Cruz	2002	Temporal lobe haemorrhage. Single centre, unknown location	GOS at 6 months	141
Lu	2003	GCS<=8. China, single centre	GOS at 6 months	230
Zhi	2003	GCS<=8. China, single centre	GOS at 6 months	396
Cooper	2004	GCS<=8. Australia, multiple centres	GOSE at 6 months	229
Edwards[3] (CRASH)	2005	GCS<=14 Worldwide, 52 centres	GOS at 6 months	10008
Jiang	2005	GCS<=8 & hypertension. China, five centres	GOS at 6 months	486
Yurkewicz	2005	GCS 4-8. US, 40 centres	GOS at 6 months	404

[2] Paper shows results for 463 subjects however all 1574 subjects included in analysis in thesis.

[3] CRASH trial did not distinguish between outcome measures of vegetative state and severe disability.

Trial	Year	Population	Outcome	N
Jiang	2006	GCS<=8. China, three centres	GOS at 6 months	215
Maas (PHARMOS)	2006	Motor 2-5 & CT abnormalities. Worldwide, 86 centres	GOSE at 6 months	861
Temkin	2007	GCS <=12. US, single centre	GOSE at 6 months	499
Xiao	2008	GCS<=8. China, single centre	GOS at 3 months	159

In the 18 trials which dichotomised the GOS, six (Bailey et al. 1991;Jiang et al. 2006;Morris et al. 1999;Rockswold et al. 1992;Saul et al. 1981;Xiao et al. 2008) specified that a Chi-squared analysis was used. However, Jiang only compared within rather than between study groups. Marmarou et al (Marmarou et al. 1999) did not specify what analysis had been used, however, from examining the figures given in the paper, it seems likely that Chi-squared analyses were performed. Three of the trials (Dearden et al. 1986;Marshall et al. 1998;The European Study Group on Nimodipine in Severe Head Injury 1994) compared the proportions with unfavourable outcome, good recovery and favourable outcome respectively. Three of the trials (Harders et al. 1996;Wolf et al. 1993;Yurkewicz et al. 2005) used logistic regression analysis, with Harders also using Fisher's exact test, a modification of the Chi-squared test, to analyse outcome. Two of the trials (Clifton et al. 2001;Edwards et al. 2005) presented relative risks. Braakman et al (Braakman et al. 1983) used a Mann Whitney test, Gaab et al (Gaab et al. 1994) used a Cochrane Mantel Haenszel test and Maas et al (Maas et al. 2006) the sliding dichotomy analysis.

Seven of the trials trichotomised the GOS. Four of these (Cruz, Minoja, & Okuchi 2001;Cruz, Minoja, & Okuchi 2002;Robertson et al. 1999;Young et al. 1996) grouped good and moderate outcomes together and vegetative and death outcomes

together; whereas the other three trials (Jiang et al. 2005;Lu et al. 2003;Muizelaar et al. 1991) grouped good and moderate outcomes together and severe and vegetative outcomes together. Most of these trials (5/7) performed Chi-squared analysis. Young et al (Young et al. 1996) used a Cochran Mantel Haenszel mean score test assigning scores of 0, 0.5 and 1 to poor, intermediate and good outcomes respectively to each trichotomy, and compared the mean score between groups. Muizelaar et al (Muizelaar et al. 1991) stratified by motor score and carried out a multiple logistic regression analysis; no mention was made in the paper if this was a proportional odds analysis.

The remaining four trials did not group the outcome variable. Two trials used the Mann Whitney test to assess differences in GOS between comparison groups (Cooper et al. 2004;Grumme et al. 1995). Zhi et al (Zhi, Zhang, & Lin 2003) tried to assess differences using a two-sample t-test! It was unclear from the paper how the GOSE measured by Temkin et al (Temkin et al. 2007) was analysed.

Bolland et al (Bolland, Sooriyarachchi, & Whitehead 1998) reported a sample size review of a study comparing eliprodil (a N-Methyl-D-aspartate (NMDA) receptor antagonist) with placebo in head injury. Outcome was measured using the GOS and a proportional odds model was fitted. The clinical results of this study have not been published to date.

Therefore, in conclusion, only five of the trials used an analysis which incorporated all of the outcome scale (Cooper et al. 2004;Grumme et al. 1995;Maas et al. 2006;Temkin et al. 2007;Zhi, Zhang, & Lin 2003). Of these five, one (Zhi, Zhang, & Lin 2003) used an inappropriate t-test to analyse outcome and Temkin et al's (Temkin et al. 2007) analysis was unclear, leaving only three Phase III trials which used the full outcome scale.

1.6.4 Ordinal analysis - stroke

Stroke and head injury can lead to similar functional impairments. The most commonly used scale to assess outcome following stroke, the modified Rankin Scale is an ordinal scale. It has seven categories ranging from '0 - no symptoms' to '6 – death'. It was of interest to explore whether this scale had been analysed as an ordinal outcome, as typically, as with head injury outcomes, it is dichotomised.

The Barthel Index is also commonly used to measure outcome after stroke. It comprises a ten item scale covering items such as mobility and toileting. For each item, subjects are scored 0 if dependent or unable to perform a task and up to 15 if able to perform a task independently. Each item is therefore measured on an ordinal scale although the total score (ranging from 0 to 100) is not ordinal.

In order to find stroke trials that had used an ordinal analysis the following search strategy was used. PubMed, Medline, the Cochrane Library and Web of Knowledge were searched for (stroke or cerebrovascular) and (Phase III or Phase 3 or random or trial) and (ordinal or proportional odds or cumulative logit or continuation ratio). The search was last updated in September 2011.

Twenty studies were identified from the literature. These were either trials or subgroup analysis based on trials. Cohort studies were not included in this search. Almost all (18/20) trials used the modified Rankin Scale to assess outcome. Zingmark et al (Zingmark et al. 2003) measured outcome as the degree of sedation. Mendelow et al (Mendelow et al. 2005) in the Surgical Trial in Intracerebral Haemorrhage (STICH) measured outcome using the GOSE at six months. Nine of the studies used ordinal regression analysis to examine outcome (Bath et al. 2009;Bath et al. 2010;Diener et al. 2008a;Lees et al. 2006b;Mishra et al. 2010;Mishra et al. 2011;Sandset et al. 2011;Shuaib et al. 2007;Zingmark et al. 2003).

Two subgroup analyses of the Prevention Regimen For Effectively avoiding Second Strokes (PRoFESS) trial (Diener et al. 2008b) were published, both of which used ordinal regression. In the first, Bath et al (Bath et al. 2009) examined the effects of telemisartan on recurrence, functional outcome and blood pressure using ordinal logistic regression and the Mann Whitney U test to examine outcome. Both analyses, the ordinal logistic regression and the Mann Whitney U test, showed similar results and the results of the ordinal logistic regression showed the same magnitude as those obtained from a conventional binary analysis. No statistically significant effect of telemisartan was found on functional dependency. In the second published subgroup analysis of the PRoFESS trial, Bath et al (Bath et al. 2010) examined 30 day functional outcome in a post-hoc analysis. This analysis included over 1300 patients and compared those randomised to aspirin and extended release dipyridamole with those on clopidogrel. Using ordinal logistic regression, no difference was found in modified Rankin Scale between groups at 30 days. Outcomes were also compared by examining a shift in the distribution of a composite endpoint measure, where again no statistically significant differences were found.

Two major studies were carried out of drug NXY-059 in acute stroke. The first trial, Stroke Acute Ischemic NXY Treatment (SAINT) I (Lees et al. 2006b), showed a positive outcome in reducing disability at 90 days. A subgroup analysis was also done of SAINT I trial which examined the number of days over the first 90 days of stroke onset that a patient spends at home (Mishra et al. 2011). This was found to be extended in patients who received thrombolytic therapy. The second much larger trial, SAINT II (Shuaib et al. 2007), showed no evidence that the treatment reduced mortality or disability. A pooled analysis was then done of both the SAINT I and SAINT II studies, (Diener et al. 2008a). This confirmed the results of the SAINT II study and showed no benefit of the treatment. In all of the analyses of the SAINT trials, the Cochran-Mantel-Haenszel test adjusted for stratification variables was used

for the primary outcome analysis. Proportional odds regression was used to give an estimate of the treatment effect.

Mishra et al (Mishra et al. 2010) used an archive of individual patient data, the Virtual International Stroke Trials Archive (VISTA) (Ali et al. 2007) to compare outcomes in over five thousand patients between those receiving thrombolysis and those not, adjusted for baseline National Institutes of Health Stroke Scale (NIHSS) (Brott et al. 1989) and age. As a formal test of the interaction between baseline NIHSS and the use of thrombolysis showed a statistically significant result, the authors decided to categorise the baseline NIHSS into seven categories. The authors used a Cochran Mantel Haenszel test to assess the association of treatment (thrombolysis) with outcome and then performed a proportional odds regression, even although they believed that the assumption of proportional odds was violated as the interaction test was statistically significant. For all NIHSS baseline categories, functional outcome improved for those in the thrombolysis group compared to the control group. For all categories of the baseline NIHSS score, except the lowest and highest (which had fewer subjects that the other categories), this was a statistically significant result.

Zingmark et al (Zingmark et al. 2003) used proportional odds modelling to model the probability of observing a particular degree of sedation. Sedation was measured using a six point scale ranging from '1-fully awake' to '6 - does not react to pain'.

The Scandinavian Candesartan Acute Stroke Trial (SCAST) (Sandset et al. 2011) recruited over 2000 stroke patients in a trial comparing candesartan and placebo. There were two co-primary effect variables: a composite endpoint variable and modified Rankin Scale at 6 months. Ordinal logistic regression was used as the primary method of analysis for the mRS with the sliding dichotomy being performed

as a sensitivity analysis. Using ordinal regression a statistically significant shift in favour of placebo was observed at the 5% significance level. However the result was not statistically significant following adjustment for the co-primary outcome variables. The sliding dichotomy method showed a similar magnitude and direction of result to the ordinal regression although it was not statistically significant.

Two stroke trials (den Hertog et al. 2009;Mendelow et al. 2005) have used the sliding dichotomy methodology as their primary analysis. The Paracetamol (Acetaminophen) In Stroke (PAIS) stroke trial examined whether early treatment with paracetamol to reduce body temperature would lead to improved outcomes (den Hertog et al. 2009). More patients receiving high dose paracetamol had improvement beyond what would be expected compared with those receiving placebo although the difference was not statistically significant. STICH was a randomised trial of a policy of early surgery with medical treatment versus a policy of initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas (Mendelow et al. 2005). Just over 1000 subjects were recruited. The primary outcome was death or disability, assessed using the GOSE at six months. Risk stratification was based on baseline prognosis using a score derived from observational studies of non-STICH intracerebral haemorrhage data. The prognostic score was a composite of GCS, age and haematoma volume. The prognostic score was used to divide patients into two groups, above and below the median score. For those with a better score a favourable outcome was deemed to be good recovery or better as classified by the GOSE. For those with a worse score a favourable outcome was upper severe disability or better. A neutral result was observed showing neither a benefit nor a detriment to early surgery or initial conservative strategy in those patients for whom the surgeon was uncertain about operating on.

Eight studies (Adams, Jr. et al. 2004;Bluhmki et al. 2009;Levy et al. 2009;Saver & Yafeh 2007;Solling et al. 2009;Song et al. 2008;Thomassen et al. 2005;Yoo et al. 2008) used an analysis which uses a marker of stroke severity at baseline to distinguish between groups of patients. This analysis uses a binary cut point to distinguish between a favourable and unfavourable outcome. A favourable outcome is however defined differently for each of these baseline groups. This type of analysis is denoted as responder or baseline severity adjusted analysis (Saver 2007).

Adams et al (Adams, Jr. et al. 2004) used responder analysis to analyse the data from three studies: the two European Cooperative Acute Stroke Studies (ECASS) and Trial of Org 10172 in Acute Stroke Treatment (TOAST) (Hacke et al. 1995;Hacke et al. 1998;The Publications Committee for the Trial of ORG 10172 in Acute Stroke Treatment (TOAST) Investigators 1998). Instead of using a prognostic score to determine baseline risk, as the sliding dichotomy uses, Adams used the baseline pre treatment NIHSS as a measure of baseline stroke severity. For those with NIHSS scores 1-7 a good outcome was a Rankin Scale=0; for those with NIHSS scores 8-14 a good outcome was Rankin Scale=0 or 1 and for those with NIHSS scores >14 a good outcome was a Rankin Scale of 0 to 2. For the ECASS studies the outcome was the modified Rankin Scale and for TOAST aggregate scores of the Barthel Index and GOS were substituted for the modified Rankin Scale. For all of these trials the overall result remained the same as the conventional dichotomised analysis. However, as would be expected, the number of less severely affected patients who had a favourable outcome was inflated under the conventional analysis whilst the number of seriously ill patients who had a favourable outcome was reduced.

Five studies (Bluhmki et al. 2009;Saver & Yafeh 2007;Solling et al. 2009;Song et al. 2008;Yoo et al. 2008) performed baseline severity adjusted analysis using the same baseline grouping of NIHSS scale and outcome as Adams et al (Adams, Jr. et al. 2004).

Bluhmki et al (Bluhmki et al. 2009) performed a secondary analysis of the ECASS III trial. Using responder, and more conventional analysis, all additional endpoints studied showed improvement in favour of the group given alteplase. The National Institute of Neurological Disorders and Stroke **(**NINDS) (The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group 1995) trials were reanalysed (Saver & Yafeh 2007). Using baseline severity adjusted analysis confirmed the initial trial results. Solling et al (Solling et al. 2009) examined outcome after acute ischemic stroke. Magnetic Resonance Imaging (MRI) and CT were compared to allow identification of patients who would benefit from recombinant tissue plasminogen activator (rtPA) treatment. Using MRI to identify patients who would benefit from rtPA was found to be safe in the population studied. Song (Song et al. 2008) assessed outcome after stroke in Korean stroke patients. The primary analysis was a comparison of those with and without elevated cardiac serum troponin T levels. Yoo et al (Yoo et al. 2008) in a study of under-nutrition found that one week NIHSS score strongly predicted poorer outcome and one week under-nutrition was a weaker predictor of poorer outcome.

Thomassen et al (Thomassen et al. 2005) compared responder analysis with a uniform assessment of outcome. For the uniform assessment of outcome, using the mRS, excellent outcome was 0-1, favourable 0-2, unfavourable 3-5 and death 6 for all subjects. For the responder analysis excellent outcome varied by original NIHSS score in the same manner as Adams et al (Adams, Jr. et al. 2004). Comparing the responder analysis to the uniform analysis, more subjects with severe strokes had favourable outcomes using the responder analysis whilst fewer subjects with mild strokes had favourable outcomes. Although the overall number of subjects who had a favourable outcome was very similar using both methods of analysis, the actual subjects were different.

Levy et al (Levy et al. 2009) also used a responder analysis in the "Ancrod in Acute Ischemic Stroke" trial. Responders were defined based on both pre-stroke mRS and pre-treatment NIHSS scores. For those with a pre-stroke mRS of 0-1 and an NIHSS score of 5-15 a favourable outcome was a mRS of 0-1. For those with a NIHSS score >=16 and a pre-stroke mRS of 0-1 a favourable outcome was mRS of 0-2. For subjects with a pre-stroke mRS of >=2 a favourable outcome was defined as returning to the pre-stroke mRS or better. This trial was halted as a planned interim analysis showed that the study should be stopped early for lack of efficacy.

Diener et al (Diener et al. 2008b), in the primary analysis of the PRoFESS trial examined outcome at 3 months using the modified Rankin Scale (using all points on the scale without dichotomy). This was a trial of over 20000 subjects who had previously had a stroke. Subjects were randomised in a 2x2 factorial design to aspirin (ASA) + extended release dipyridamole (ER-DP) and either telmisartan or placebo; or clopidogrel and telemisartan or placebo. Only subjects with an initial modified Rankin Scale score of less than or equal to four were included. Using one way ANOVA no difference was found in modified Rankin Scale score between the two groups for analyses comparing ASA+ER-DP versus clopidogrel or telmisartan versus placebo. These analyses would assume that the mRS scores were normally distributed – an assumption which seems unlikely looking at the figures in the paper. Although this was not an ordinal analysis it did include all the points on the mRS scale in the analysis, hence its inclusion here.

1.6.5 Discussion

This systematic review has shown that the Glasgow Outcome Scale is the most commonly used outcome measure for Phase III head injury trials. However, only three trials exploited the ordinal nature of the scale appropriately when assessing outcome. None of the 29 head injury trials identified carried out a proportional odds analysis of the data. Examining stroke trials with an ordinal outcome found ordinal

logistic regression being used appropriately in eight studies and the sliding dichotomy in three, one of which also used ordinal logistic regression. This thesis will use the data from eleven published studies in TBI to explore the strengths and limitations of using an ordinal analysis in place of the conventional binary analysis.

Note – the published papers from the IMPACT (International Mission on Prognosis and Analysis of Clinical Trials in TBI) project, details given in section 1.7, which show results of ordinal modelling in head injury, have not been included in the results of the searches above as these results form part of this thesis and are expanded on in detail throughout the thesis.

1.7 IMPACT project

This thesis describes part of the IMPACT project, a National Institutes of Health (NIH) funded project, looking at the application of conventional and innovative methods for the design and analysis of studies in traumatic brain injury. The primary aim of the project is to optimise chances for demonstrating benefit of a new therapeutic agent in the field of TBI. The rationale for this project is that to date no Phase III randomised clinical trials in TBI have convincingly shown benefit.

Within the NIH project four specific areas were to be examined: 1) the heterogeneity of the TBI patient population; 2) the Glasgow Outcome Scale (GOS) assessment; 3) the development and validation of prognostic models and 4) design issues such as recruitment, sequential designs and early endpoints. Each of these four specific aims has been further split into two areas of interest. The project was run over three centres: Edinburgh, Richmond and Rotterdam with each centre having lead responsibility for different parts of the project, although all centres shared responsibility for the development and validation of prognostic models in head injury. In Edinburgh there were two major areas of responsibility: specific aims 1a)

prognostic and mechanistic targeting - within the investigation of the heterogeneity of the TBI population and 2b) the "sliding dichotomy" - within the examination of the GOS.

Specific aim 1a) - prognostic and mechanistic targeting - studied the potential for increasing statistical power by targeting therapy to specific risk groups, (Machado, Murray, & Teasdale 1999). As the population of traumatic brain injury is very heterogeneous it is reasonable to target therapies at patients that can potentially show benefit. Those patients with either a very good prognosis or those with a very poor prognosis contribute little to the statistical power of a conventional study. Intermediate risk groups can be defined according to prognosis (prognostic targeting) or type of head injury (mechanistic targeting).

Specific aim 2b) – the application of the sliding dichotomy – addressed problems in assessing outcome after head injury. The GOS was explored as an ordinal scale with the point of dichotomisation varying between different prognostic risk groups.

1.8 Thesis outline

The first aim of this thesis is to examine ordinal methods of analysing Phase III data, such as the sliding dichotomy and proportional odds modelling and compare these to more conventional methods of analysing outcomes. The second aim is to determine which method of analysis is best under which circumstances. That is the merits of a simple easily understood outcome, for example a clinical audience may prefer the simplicity of an outcome which is e.g. favourable or not versus the merits of a model which is statistically more efficient (such as the proportional odds model). The primary approach of this thesis is to use data from real trials in order to determine how trials with ordinal outcomes may best be analysed.

In order to do this, this thesis will describe methods currently used in the literature for trials in head injury and stroke studies – Chapter 1. The IMPACT data used in the modelling with the preliminary stages prior to analysis (data extraction, imputation of missing data etc.) are then described in Chapter 2. A description of the methodology used for the binary analysis, ordinal analysis and meta analysis is given in Chapter 3. Binary analysis and ordinal analyses – both unadjusted and adjusted for covariates are described in Chapter 4. Chapter 5 compares different sliding dichotomy modelling strategies. Then in Chapter 6 the optimised sliding dichotomy is compared with other modelling strategies. A discussion of the results shown in the thesis, placing them in context and recommendations for future work are shown in Chapter 7.

2 Chapter 2 Subjects and Methods

2.1 Data collection and datasets

2.1.1 Introduction

As part of the IMPACT project we were able to obtain the original subject-level data from 11 completed head injury studies, eight Phase III trials and three observational studies. For the eight trials, no re-analysis by original treatment group was permitted as a condition of obtaining the data. Staff at each centre (Edinburgh, Richmond and Rotterdam) took responsibility for extracting variables from different studies. Summary characteristics of the studies are given below in Table 2-1.

After reviewing the literature and providing an outline of the project in Chapter 1, this chapter, Chapter 2, will describe the IMPACT datasets. The data collection, extraction and synthesis will all be described. This chapter will also discuss the handling of missing data, the pre-screening of continuous variables exploring their distribution, relationship with outcome and the codings and categorisations used in subsequent analysis. Binary analysis of outcomes, with illustration, to allow data checking, is then shown.

Table 2-1 Studies in the IMPACT database

Study	Publication	Population	N
Traumatic Coma Data Bank TCDB	(Foulkes et al. 1991)	GCS<=8 following non surgical resuscitation or GCS deteriorating to <8	677
UK Four Centre Study UK4	(Murray et al. 1999b)	Admitted to centre within 3 days of severe head injury resulting in coma	988
Nimodipine Trial HIT I	(Bailey et al. 1991)	Not obeying commands within 24h of injury	351

Study	Publication	Population	N
Nimodipine Trial HIT II	(The European Study Group on Nimodipine in Severe Head Injury 1994)	Not obeying commands within 24h of injury	852
Tirilazad US study TIUS	Not published	70% severe head injury (GCS 4-8) 30% moderate head injury (GCS 9-12)	1155
Tirilazad International study TINT	(Marshall et al. 1998)	85% severe head injury (GCS 4-8) 15% moderate head injury (GCS 9-12)	1131
PEGSOD[4]	(Young et al. 1996)	Severe head injury, GCS<=8 after resuscitation	1574
European Brain Injury Consortium EBIC	(Murray et al. 1999a)	GCS <=12	1005
Selfotel International SLIN	(Morris et al. 1999)	GCS<=8 post resuscitation	409
Bradycor SKB	(Marmarou et al. 1999)	GCS<=8	139
SAPHIR	Not published	Not obeying commands, >=1 reacting pupil	924

The main initial aim was to extract data which would be available at the point of entry into a trial (i.e. randomisation) with outcome. At a later stage potential early/surrogate endpoints were extracted. In order to combine data from these 11 studies, a common format for the extraction of variables was specified. A detailed template was developed giving variable names and formats in order that a permanent record could be kept of the derivation of variables, the link between the extracted variables from each dataset and the merged dataset, and how the extracted variables would be derived if they did not have a one-to-one correspondence. In preparation

[4] polyethylene glycol-conjugated bovine superoxide dismutase

for extracting these variables literature searching was performed to find the key publications for each of the studies in order that the results obtained using the raw data could be compared with published results. Problems with the data extraction or invalid or inconsistent data values were documented. This extraction of the data was not a one off process performed in isolation but rather an iterative process which evolved over time. We wished to not only assess traditional predictors of outcome, such as pupil reactivity, but also to explore novel predictors of outcome such as glucose. We also wished to convert the data we had into a useable, well-documented format. The database within which the extracted data were held was denoted the Top Priority Data Set (TPDS).

It was agreed that the data should be extracted 'as is' with no cleaning to give a record for posterity of what was recorded in each dataset, albeit in a common format. These data were denoted source one data. Subsequently the data were cleaned and denoted source two data. These source two data were used in all of the analyses.

2.1.2 Description of datasets used in analysis

2.1.2.1 TCDB - Traumatic Coma Data Bank

The TCDB (Foulkes et al. 1991) was the oldest dataset included in the database. This was an unselected series of 677 severely head injured patients admitted to four US centres (California, Texas, Virginia University and Virginia Medical College) between 1984 and 1987. It recorded details on initial severity, demographics and mechanisms of injury. Outcome was assessed at 6 months post injury. These were also the data on which the Marshall CT (Marshall et al. 1991) classification was based. These data were extracted by staff in Richmond.

2.1.2.2 UK4 – UK Four Centres study

As with the TCDB study this too was an unselected series. It consisted of 988 patients admitted to four neurosurgical centres within the UK (Edinburgh, Liverpool, Glasgow and Southampton) (Murray et al. 1999b). The study ran from 1986 to 1988 and subjects were recruited within three days of their head injury. Demographics, CT classification and neurosurgical unit management were measured. Outcome was assessed six months post injury using the Glasgow Outcome Scale. These data were extracted by staff in Edinburgh.

2.1.2.3 HIT I

This was a prospective randomised trial to study the effects of the calcium antagonist nimodipine (Bailey et al. 1991). Subjects with a head injury and not obeying commands were recruited within 24 hours of injury. The study recruited 351 subjects between 1987 and 1989. Outcome was assessed at 6 months using the GOS. These data were extracted by staff in Edinburgh.

2.1.2.4 HIT II

This was, as with HIT I, a prospective randomised study to compare nimodipine with placebo although in a larger number of subjects, 852 (The European Study Group on Nimodipine in Severe Head Injury 1994). The study ran from January 1989 to June 1991 and collected data on demographics and CT characteristics as well as outcome as measured by the GOS at six months. These data were extracted by staff in Edinburgh.

2.1.2.5 TIUS – Tirilazad United States study

This unpublished randomised trial studied the effects of tirilazad mesylate in 1155 patients with a head injury in the US and Israel. Patients were recruited from 1991 to 1994. Information on demographics, CT classification and neurosurgical management was recorded. Outcome was assessed at 6 months using the GOS. These data were extracted by staff in Rotterdam.

2.1.2.6 TINT – Tirilazad International study

This trial ran in parallel to the US study and studied the effects of tirilazad mesylate in countries excluding the US and Israel (Marshall et al. 1998). It recruited 1131 patients with a head injury with the same variables and outcomes being collected as the US study. These data were extracted by staff in Rotterdam.

2.1.2.7 PEGSOD – polyethylene glycol-conjugated bovine superoxide dismutase

This was a series of randomised, multicentre parallel group trials studying the effects of pegorgotein, a scavenger of oxygen-derived free radicals (Young et al. 1996). Between 1993 and 1995 the trials recruited 1574 subjects. The data from all of these trials were obtained by IMPACT even although only the analysis of one of the series of trials, on 463 subjects, was published. Subjects were recruited from 29 American centres. Outcome was assessed using the GOS at three months. These data were extracted by staff in Richmond.

2.1.2.8 EBIC – European Brain Injury Consortium

This was a survey of 1005 moderately and severely injured adult head injured patients admitted to 67 centres in 12 European countries (Murray et al. 1999a). The survey was conducted in 1995. Demography, clinical features, management and investigations were recorded. Outcome was assessed at six months using the GOS. These data were extracted by staff in Edinburgh.

2.1.2.9 SLIN – Selfotel International study

This was the international arm of a trial of the N-methyl-D-aspartate receptor antagonist Selfotel. It recruited 409 subjects between 1994 and 1996. Subjects were severely injured having a post-resuscitation Glasgow Coma Scale of four to eight. Information on demographics, clinical features, and treatment management was recorded. We were unable to obtain data for the American arm of the trial. This trial

was stopped early because of possible safety concerns. However, six month outcome was assessed by the GOS. Staff in Rotterdam extracted these data.

2.1.2.10 SKB – the Bradycor study

This was a randomised, placebo controlled trial studying the effects of the bradykinin antagonist Bradycor in 31 North American centres (Marmarou et al. 1999). One hundred and thirty one severely brain injured patients with GCS three to eight and at least one reacting pupil were recruited in 1996. The primary objective was to assess the efficacy of a continuous infusion of Bradycor in preventing elevation of intracranial pressure (ICP). Demographic details, clinical details and ICP were all recorded. Outcome was assessed at three and six months by the GOS. These data were extracted by staff in Richmond.

2.1.2.11 SAPHIR

This was an unpublished double blind placebo controlled trial of the drug CPP-ene, a competitive NMDA receptor. Subjects not obeying commands and with at least one reactive pupil were recruited from 51 European Centres within 12 hours of injury. In total, 924 head injured subjects were recruited between 1995 and 1997. Demographics, clinical features and treatment management were recorded. Outcome was assessed at three and six months by the GOS. Staff in Edinburgh extracted these data.

2.2 Data extraction

The data extraction was not a trivial task as the historical data were in numerous formats with many files relating to the one study and often little, incomplete or even incorrect and inconsistent documentation. An overview of the extraction of the three datasets for which I was explicitly responsible, EBIC, HIT I and SAPHIR is given below.

2.2.1 EBIC

The EBIC data consisted of one file with information on 1005 subjects from 67 centres. An annotated Case Record Form (CRF) was available for the data making the process of extracting information relatively straightforward. The extracted variables were renamed and reformatted and then checked against the main EBIC publication (Murray et al. 1999a). All extracted variables were found to agree with the published paper.

2.2.2 HIT 1

The HIT I data consisted of 25 data files and one comments file with records on 352 subjects from six centres. One subject was found to have been randomised twice, having had two separate head injuries within the recruitment period of the study. Only information from the first head injury within the recruitment period was used, giving data on 351 subjects. The data were originally stored as character files in column format with no text headings or identifiers on the files. Each dataset had a corresponding descriptive file which specified the column names, formats and position in the dataset. The extraction of the data into a readable format was very challenging. The completed paper case record forms were available and had to be compared with the dataset in order to decipher the placement, format and names of columns for all 25 data files as only one of the data files was formatted as the descriptive file had specified. An un-annotated CRF was available for the HIT I data. As with the EBIC data, the extracted HIT I data were compared with the study publication, (Bailey et al. 1991), and found to agree.

2.2.3 SAPHIR

The SAPHIR data consisted of thirty-nine files with records on 924 subjects. The data were originally stored in zipped SAS files with embedded formats. The data

were unzipped and formatted in SAS. The structure of the data files was very complex making the selection of the correct variables and records problematic. An incomplete partially annotated CRF was available for the SAPHIR data. The SAPHIR data have not been published however an unpublished study report and draft paper were available. There were many major and minor discrepancies between the raw data and the study report and indeed between the study report and the draft paper. For example: there was no one unique patient identifier for subjects; dates of death varied for some patients in the three and six month outcome files; many subjects had a start date of intracranial pressure monitoring before their date of injury and, for approximately 5% of subjects, their date of baseline examination varied in every file in which it was recorded.

Monitoring data, such as blood pressure and intracranial pressure, were particularly difficult to clean and convert to a useable format. These data were recorded over multiple files, notionally being recorded every hour for up to seven days of monitoring. As these data had never been cleaned there were many discrepancies and errors within the file, such as multiple different records at the same time point for the same subject. It took many days of work to extract the data into a suitable format with each subject only having one record for each time point. These discrepancies were resolved in discussion with members of the three IMPACT centres.

2.3 Data synthesis

The five Edinburgh datasets, EBIC, UK4 Centres, HIT I, HIT II and SAPHIR were sent to Richmond where they were entered into a common database along with the other extracted studies. Templates produced in Edinburgh, specifying how variables were coded in our datasets, were circulated so that the other centres could copy this model when sending their own data.

Once nine of the eleven datasets had been received by Richmond they circulated a SAS file which contained all of the extracted data with a consistent coding system where possible. For example, in the Edinburgh datasets death was coded as 0, 1 and Richmond recoded this to 1, 2. Upon checking the extracted data it was apparent that firstly not all variables sent to Richmond had been included, and secondly, for those that were included, some errors had been made in recoding. These problems were notified to Richmond and after many iterations the Edinburgh datasets were entered into the common database and had been recoded correctly and consistently. As part of the data extraction and checking process preliminary results were circulated between the centres. These identified various problems, such as the extracted data for TINT showing 85% of subjects with bilateral non-reacting pupils at first hospital, changing to only 11% on admission to study hospital and then changing again to 93% at post resuscitation – obviously an impossible sequence of events. TIUS also had similar problems. This led to more iterations of the datasets. This process, although vital, was very time consuming. However it did allow the progression of the analysis with confidence in the results.

2.4 Missing data

The datasets used in IMPACT had two kinds of missing data. The first kind was where a specific study did not record a particular variable, for example in HIT 1 place and cause of injury were not recorded. Using Little and Rubin's definition this would be denoted as Missing Completely At Random (MCAR) where the probability of data being missing does not depend on the observed or unobserved data (Little & Rubin 1987). The second kind of missing data occurred when a variable was recorded within a study although some individuals did not have a value recorded. For example, traumatic subarachnoid haemorrhage was recorded in all studies except UK4 however its completeness ranged from 100% in SLIN to 73% in HIT I. For the univariate analyses shown in Chapter 4 a complete case analysis was performed in that analysis was restricted to subjects with data available on the covariate of interest and the GOS.

Imputation techniques were used to minimise the problems associated with the analysis of incomplete data due to missing values in the multivariate analysis, shown in Chapter 4. Using complete case analysis would have been inappropriate as the missingness of a single covariate would have led to the subject's entire record being discarded and not included in the analysis. Complete case analysis has also been shown to lead to bias (Little & Rubin 1987). Multiple Imputation (Rubin 1987) uses the distribution of the observed data, including the outcome variable, to estimate the missing data (van Buuren, Boshuizen, & Knook 1999;White, Royston, & Wood 2011). In the IMPACT database, staff in Rotterdam used Multiple Imputation to create 10 datasets based on all of the data. The first dataset created using the multiple imputation procedure was then used in the multivariate analyses (McHugh et al. 2007a).

2.5 Pre-screening of continuous variables

Previous studies in head injury have typically categorised continuous variables. This approach leads to loss of sensitivity and, as with categorising continuous outcome variables, can lead to a substantial loss of information. Prior to conducting any formal analysis the distribution of the continuous variables and their relationship with outcome was explored. This exploration was performed on all of the studies merged together.

Boxplots and histograms were examined for each continuous variable. These and the results from univariate analysis showed that in a very few instances, less than 1% of all data, 'extreme' values for continuous variables had been recorded. It would have been wrong for a few extreme values to influence the results of prognostic modelling. It was therefore decided to truncate values that were outside the range (defined in Table 2-2 below) to the lower and upper points of the range. This process keeps all of the original data and simply shifts extreme values to the ends of the range. There are no guidelines as to what values the range should take. This was

therefore determined by examining the plots and the results from the univariate analysis.

Only two studies had included children younger than 14 years of age, TCDB and UK4. It was therefore decided to restrict analyses to those aged 14 or older. These were the only data excluded from the analysis. For all of the other continuous variables, more extreme values were shifted to the ranges shown in Table 2-2. For example all haemoglobin values less than 6 g/dL were set to 6 g/dL and all values greater than 17 g/dL were set to 17 g/dL after truncation.

Splines

After truncating each variable, splines were used to visualise the shape of the relationship with outcome for each variable and thus guide the statistical modelling.

Spline functions are piecewise polynomials used in curve fitting which can be fitted using regression programs. They have been used primarily in the physical sciences to approximate a wide variety of functions (Harrell, Jr. 2001). The most commonly fitted splines are linear and cubic. In order to fit the spline function knots (i.e. the points at which the spline bends) have to be chosen. Cubic spline functions can be used to fit highly curved shapes, however, they can be poorly behaved in the tails (Stone & Koo 1985). It is therefore recommended that restricted cubic splines are fitted which constrain the tails of the function to be linear and also result in fewer parameters being specified (Harrell, Jr. 2001).

For the continuous variables modelled in the IMPACT data restricted cubic splines were fitted with five prespecified knots at the 5, 27.5, 50, 72.5, and 95 percentiles of

the data for each continuous prognostic factor (McHugh et al. 2007a). It has been shown that the number of knots is more important than the exact location (Harrell, Jr. 2001). Five knots (giving four degrees of freedom) were chosen as it has been shown in practice that five knots are sufficient to model most non linear functions (Stone 1986).

Fitting splines although technically intricate can often lead to much simpler forms of the model between variables and outcome as was shown with the IMPACT data. A linear spline function was observed to fit approximately the relationship observed between seven of the covariates and GOS. The relationship between each of Systolic Blood Pressure, Mean Arterial Blood Pressure and Sodium with GOS showed a U shaped effect (McHugh et al. 2007a;Royston 2000).

Coding/categorisation

For those variables which had a linear relationship with outcome, the variables were scaled by the interquartile range 25th percentile to 75th percentile, p75-p25. For example age was divided by 24. Scaling creates estimates of odds ratios which are comparable between different prognostic factors both categorical and continuous (Harrell, Jr. & Shih 2001). For those variables with a U shaped relationship with outcome a range was defined for the central category p75-p25 as previously. So, for example, Mean Arterial Blood Pressure would be defined as low if below 85mmHg, middle if between 85 and 110mmHg inclusive and high if greater than 110 mmHg.

Table 2-2 below produced by staff in Rotterdam shows the ranges chosen to truncate at, whether or not the variable had a linear relationship with GOS, and what coding or categorisation was used in the analysis.

Table 2-2 Handling of continuous variables

Variable	Range	Linear	Coding /Categorisation
Age (years)	14+	Yes	/24
Systolic Blood pressure (mmHg)	60-230	No	120-150
Mean Arterial Blood Pressure (mmHg)	40-160	No	85-110
pH	7.0-7.7	Yes	/0.15
Haemoglobin (g/dL)	6-17	Yes	/3.3
Glucose (mmol/L)	3-20	Yes	/3.7
Sodium (mmol/L)	125-155	No	137-142
Haematocrit (%)	18-50	Yes	/10
Platelets $x10^9$/L	Lower Range-450	Yes	/100
Prothrombin Time (seconds)	Lower Range-20	Yes	/2

2.6 Data checking introduction

Prior to proceeding with any analysis, each variable in the merged database was summarised in tabular form and graphically to provide not only an indication for future analyses but also to assist with the checking of the data. The relationship between each variable and outcome was also examined. Outcome was measured using the Glasgow Outcome Scale at six months for all studies except PEGSOD, where it was only recorded at three months. In all analysis outcome was taken at six months where available and three months if not available at six. In order to develop prognostic models the first stage was to carry out a conventional analysis of all covariates with the GOS as a binary outcome.

2.7 Summary methods for 2x2 tables

Using a similar approach to that postulated by Chesnut et al (Chesnut et al. 1993) and illustrated by McHugh et al (McHugh et al. 2007a) the relationship between dichotomous predictors and outcome can be thought of in a similar way as the methods used to evaluate diagnostic tests. The dichotomous outcome measure can be thought of as the reference measure against which a diagnostic test is assessed with, in effect, the prognostic factor of interest replacing the diagnostic test. Dichotomous predictors can be categorised as present or absent, for example pupils reactive or unreactive. Using this notation the relationship between a prognostic factor and outcome can be shown in a 2 x 2 table, Table 2-3 below. Many summary measures can be obtained from such a table: odds ratio ((ad)/(bc)); relative risk ((a/g)/(b/h)); sensitivity (a/g); specificity (d/h); positive predictive value (a/e) and negative predictive value (d/f).

Table 2-3 Illustration for the calculation of summary measures of association in a 2 x 2 table

	Unfavourable outcome	Favourable outcome	Total
Factor present	a	b	e
Factor absent	c	d	f
Total	g	h	N

For the data in the merged database the odds of an unfavourable outcome were modelled. Reference categories were picked for each of the variables, usually the category that made sense clinically or the category with the largest number of subjects. For example, for the variable place of injury, street/highway was used as the reference category and the odds of a poor outcome for each of the other categories relative to the odds of a poor outcome of street/highway were compared. For some categorical variables, e.g. cause of injury, there were many possible categorisations over all of the studies. A decision was made to group the levels of

these categorical variables in order to have reasonable numbers within each level whilst still keeping the distinct causes separate. For example, car occupant and motor vehicle occupant were combined however these categories were not combined with, for example, pedestrian. For the continuous variables modelled with a linear relationship with outcome, the odds were scaled by the interquartile range. For continuous variables with a U shaped relationship with outcome the interquartile range was taken as the reference category, as detailed previously. Each level of each variable was compared back to a reference category and the odds tabulated and graphed.

2.8 Illustration of binary outcome analysis - each covariate by study

Three outcomes were modelled based on the GOS classification: (i) mortality - death versus everything else; (ii) an unfavourable outcome - grouping the GOS into death, vegetative state and severe disability versus moderate disability and good recovery; (iii) a bad outcome - grouping death, vegetative state, severe disability and moderate disability versus a good recovery. For each of the eleven studies separately univariate odds ratios, 95% confidence intervals for the odds ratios and areas under the Receiver Operating Characteristic (ROC) curve, were calculated for each of the comparisons. A SAS macro was written by Dr Butcher (Edinburgh centre) to calculate these values as to look at each level of each variable relative to a reference category involved over 100 comparisons. This number of comparisons was then multiplied three fold for the three different outcomes. This was a crude preliminary analysis in that all of the data from the eleven studies were analysed together without any weighting.

The table below, Table 2-4, examines gender and shows the number of subjects in each category as well as odds of mortality for each of the eleven studies and overall. For each study the category ‘male’ was used as the reference category and the odds shown are the odds of mortality for females relative to the odds of mortality for males. The ‘ALL’ category is an amalgamation of the numbers from all of the

studies and does not represent any weighted estimate; this was done to simplify the first examination of the data. Table 2-4, shows that no significant difference was found between the odds of mortality of females compared with males for any of the studies as all of the confidence intervals overlap one. The AUCs (area under the curve) are of a similar magnitude (0.50 to 0.54) for all studies.

Table 2-4 Odds Ratios and 95% Confidence Intervals for mortality by gender

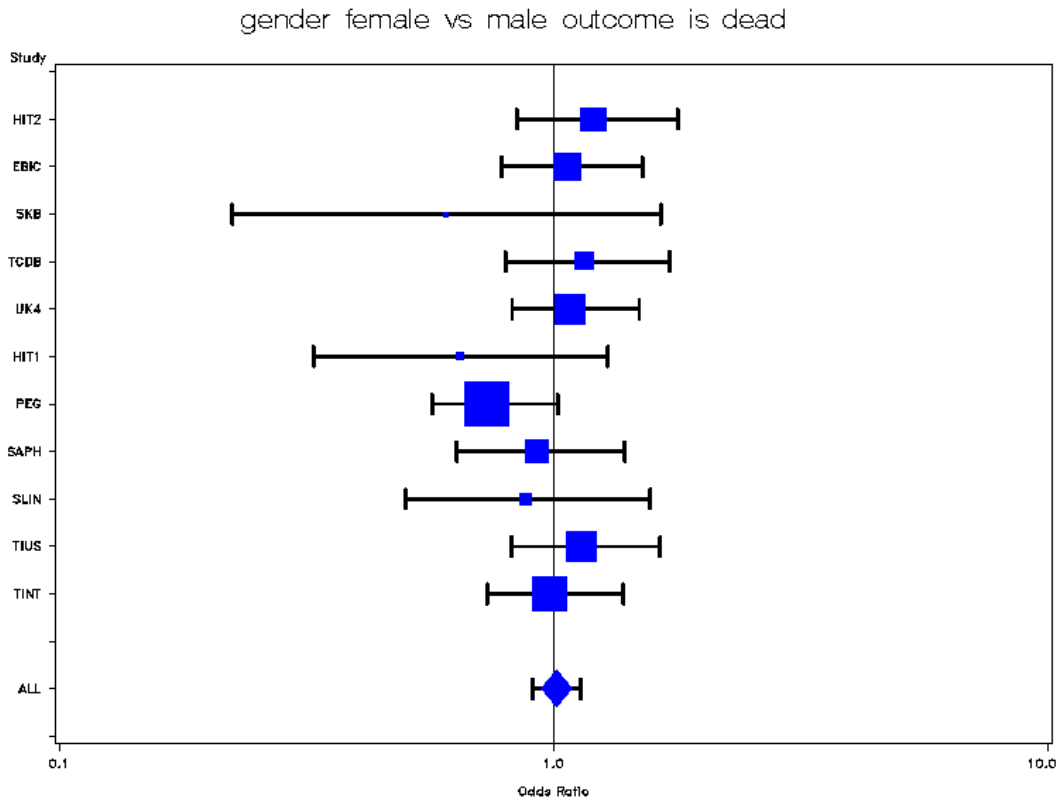
Study Name	Class Level	N	Event (Death)	%	Odds Ratio	95% CI	AUC
HIT2	female	190	49	25.79	1.23	(0.84 to 1.78)	0.519
	male	629	139	22.10			
EBIC	female	212	75	35.38	1.09	(0.79 to 1.51)	0.508
	male	622	208	33.44			
SKB	female	30	6	20.00	0.61	(0.22 to 1.65)	0.542
	male	96	28	29.17			
TCDB	female	137	64	46.72	1.17	(0.80 to 1.72)	0.514
	male	467	200	42.83			
UK4	female	242	99	40.91	1.11	(0.82 to 1.49)	0.510
	male	744	286	38.44			
HIT1	female	56	12	21.43	0.65	(0.33 to 1.29)	0.527
	male	294	87	29.59			
PEG	female	350	71	20.29	0.76	(0.57 to 1.02)	0.523
	male	1160	291	25.09			
SAPH	female	180	40	22.22	0.94	(0.64 to 1.39)	0.505
	male	739	172	23.37			
SLIN	female	89	19	21.35	0.89	(0.50 to 1.57)	0.510
	male	320	75	23.44			
TIUS	female	233	55	23.61	1.16	(0.82 to 1.64)	0.513
	male	809	170	21.01			
TINT	female	268	67	25.00	1.01	(0.73 to 1.38)	0.501
	male	853	212	24.85			

Study Name	Class Level	N	Event (Death)	%	Odds Ratio	95% CI	AUC
ALL	female	1987	557	28.03	1.01	(0.91 to 1.13)	0.501
	male	6733	1868	27.74			

In order to represent visually these odds ratios and confidence intervals, forest plots were constructed for each of the three outcomes: mortality, unfavourable and bad. Again, SAS macros were written to produce these as no standard SAS procedure is able to construct these types of graph. These plots showed for each of the variables the odds ratio and corresponding 95% confidence interval. The odds ratios are shown on a log scale with the size of the square, representing the point estimate of the odds ratio, proportionate to the size of the study. As with the tables, the diamond, representing the estimated odds for the 'ALL' group is based on an amalgamation of the numbers in each study without weighting. The study name is given on the left.

The forest plot for gender, measured in all 11 studies, is shown in Figure 2-1 below. This reflects the information in the table showing no evidence of a difference in the odds of a poorer outcome between males and females for any of the eleven studies.

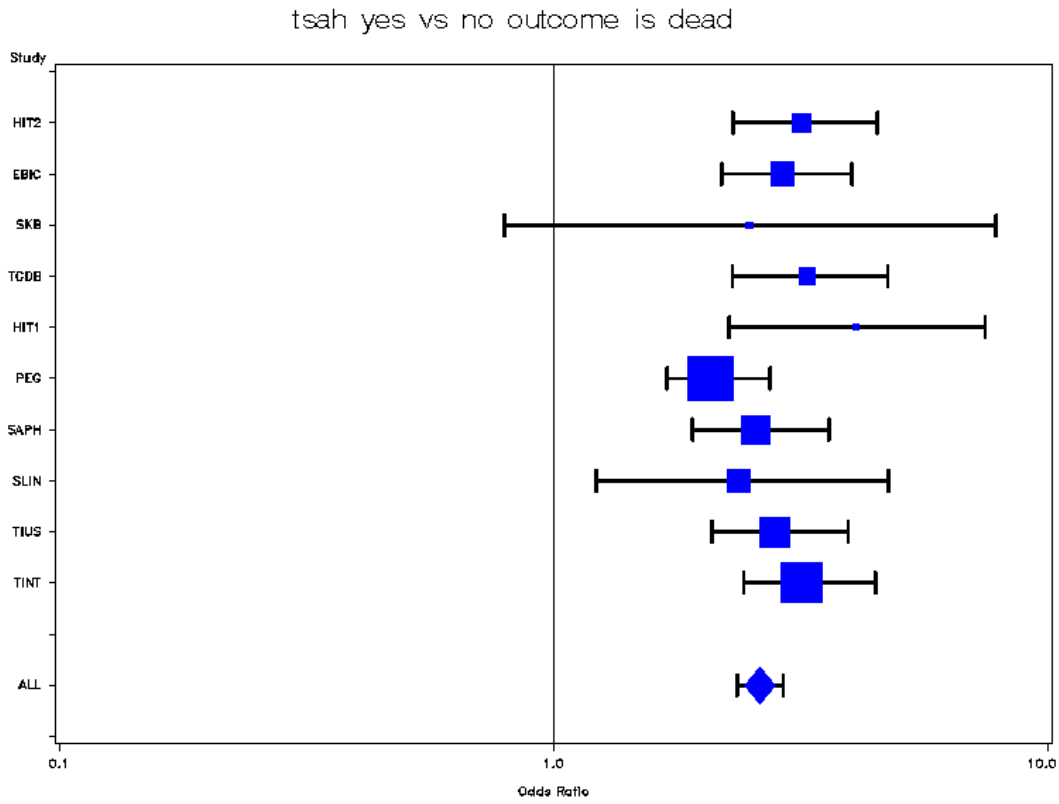
Figure 2-1 Forest plot showing odds of mortality by gender



More interesting patterns however emerge when examining the plot for traumatic subarachnoid haemorrhage (tSAH).

Figure 2-2 below shows the corresponding forest plot for this variable. It can be seen that subjects who have had a subarachnoid haemorrhage have a greater odds of dying than those subjects who have not. For all studies except SKB the 95% confidence intervals do not overlap one. However, the SKB study does have very few subjects. The UK4 centres study did not record information on subarachnoid haemorrhage therefore it is not shown in this figure.

Figure 2-2 Forest plot showing odds of mortality by traumatic subarachnoid haemorrhage



2.9 Software

All analyses were performed using SAS 9.1 (SAS v9. 2004. Cary, NC, USA) with the exception of the generation of the odds ratios chosen to increase the proportion of favourable outcomes in the simulation modelling shown in Chapter 5 which were calculated using the Solver add on in Excel (Microsoft Office Excel 2003).

2.10 Discussion

This chapter has detailed the eleven IMPACT datasets and the careful extraction and synthesis of these data. The handling of continuous variables and summary methods for 2x2 tables has also been discussed. This first stage allowed assessment of odds ratios prior to carrying out any more formal analysis. It was helpful in screening for

any coding errors in data extraction. It was also useful to inspect visually the variables for heterogeneity. A more formal test of heterogeneity was performed on the estimates from the proportional odds modelling, as detailed in Chapter 3. Chapter 3 also discusses more generally the methodology used in the thesis.

3 Chapter 3 Methodology

3.1 Introduction

After the data checking and visual inspection of the variables, the next stage was to model more formally the 'treatment' effects over all eleven studies. Modelling was performed both by taking the outcome (GOS) as a binary outcome, over all four possible splits, and as a four point ordinal outcome. A four point ordinal outcome was modelled as the categories of death and vegetative state were pooled on both statistical and ethical grounds. The statistical grounds being that vegetative state is a relatively rare outcome (only 4% of patients had a 6 month GOS classification of vegetative state); and the ethical grounds being that a vegetative state should never be regarded as being a favourable outcome, irrespective of baseline prognosis. For the binary outcomes logistic regression models were fitted. For the ordinal outcome cumulative logit models with proportional odds, hereafter referred to as proportional odds models, were fitted. The first part of this chapter discusses both binary analysis and proportional odds theory. The proportional odds model is then considered in more depth paying particular attention to the addition of covariates, the application of the model and testing goodness of fit. The second part of the chapter discusses meta analysis demonstrating how it can be applied using fixed and random effects methods.

3.2 Binary analysis theory

For a binary outcome variable using standard notation let

Y be the response variable, X be the explanatory variable and π be the probability of success.

Then:

$\pi(x)$ =P(Y=1|X=x) =1-P(Y=0|X=x)

The logistic regression model is

$$\pi(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)}$$

where α is the intercept and β the parameter estimate from the explanatory variable.

Equivalently the logit (log odds) is

$$\text{logit}\,[\pi(x)] = \log \frac{\pi(x)}{1 - \pi(x)} = \alpha + \beta x \quad \text{(Agresti 2002)}$$

These formulae can be extended to incorporate multiple covariates.

3.3 Proportional odds theory

The proportional odds model has no distributional assumptions about the ordered categorical response. However it does have a parametric assumption about the relationships between the response distributions of any two individuals. If the assumption of proportional odds holds, using proportional odds analysis will exploit the ordinal nature of the GOS without losing any information.

The proportional odds model is palindromic invariant in that the model will accept the ordering of the outcome categories (1 to k) in increasing order 1,…,k or decreasing order k, k-1,…,1 but reject all other orderings (McCullagh 1978). Changing the order of the outcome categories from increasing to decreasing or vice versa will change only the sign of the parameter estimates on the logit scale.

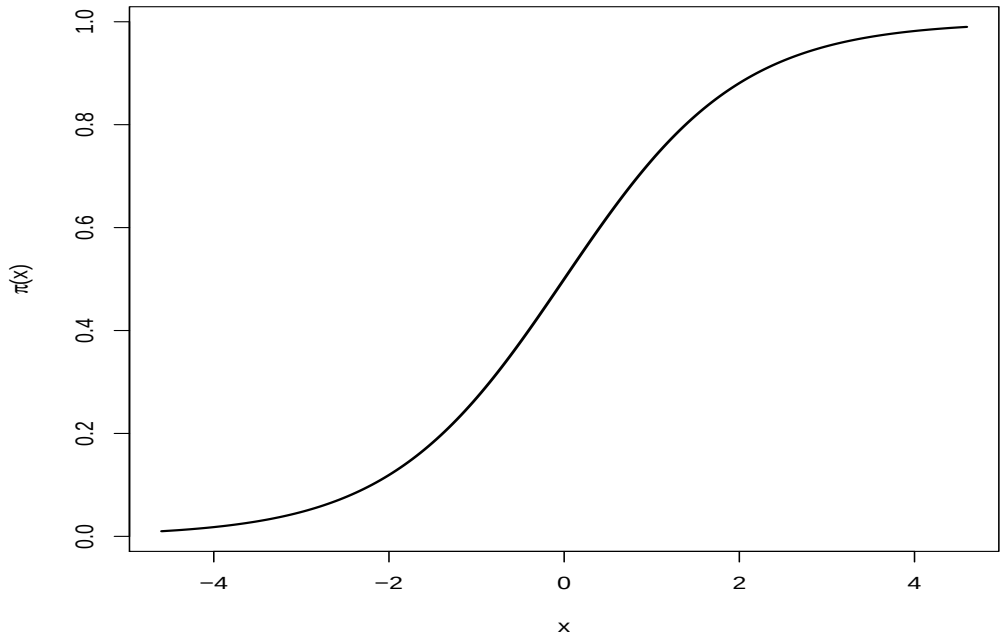
Using similar notation to the binary logistic model

$$\text{logit}[P(Y \le k|x)] = \log \frac{P(Y \le k \mid x)}{1 - P(Y \le k \mid x)} = \alpha + \beta x \quad \text{(Agresti 2002)}$$

3.4 Proportional odds model – addition of covariates

The logit distribution, unlike the Normal distribution is a sigmoidal shape, as shown Figure 3-1 below. With $\pi(x)$ as defined above.

Figure 3-1 Logistic curve



As illustrated by the example below, this non-linearity means that the addition of even a completely balanced covariate inflates the estimate of the odds ratio (Pocock et al. 2002;Robinson & Jewell 1991). This is not seen with the addition of covariates to a normal linear model where the addition of a completely balanced covariate does not change the treatment effect estimate.

With linear regression the addition of covariates to a model gives more precise treatment effect estimates(Gail, Wieand, & Piantadosi 1984). This improvement in

precision can be explained due to a reduction in residual variance (Robinson & Jewell 1991). In a linear model if the covariates are completely balanced then unadjusted and adjusted estimates are the same. This improvement in precision however does not occur with logistic regression and almost counter-intuitively the addition of even completely balanced covariates to a non-linear model leads to a loss of precision (Ford & Norrie 2002;Robinson & Jewell 1991).

This is illustrated using the hypothetical example below in a similar manner to that shown by Steyerberg and Eijkemans (Steyerberg & Eijkemans 2004). Table 3-1 below shows treatments A and B and overall survival. Table 3-2 shows stratification by tSAH. The odds ratio for survival changes from 0.44 overall, to 0.38 when stratified by having a tSAH. This covariate although completely balanced between treatment groups is highly predictive of outcome with far greater numbers of those without a tSAH surviving compared to those with a tSAH.

Table 3-1 Hypothetical example showing overall survival by treatment

Overall			
Treatment	Dead	Survive	Total
A	40	60	100
B	60	40	100
		50%	

Table 3-2 Hypothetical example showing survival stratified by tSAH

tSAH no			tSAH yes		
Dead	Survive	Total	Dead	Survive	Total
10	40	50	30	20	50
20	30	50	40	10	50
	70%			30%	

This hypothetical result has been shown in a published trial. Steyerberg et al (Steyerberg, Bossuyt, & Lee 2000) illustrated the addition of covariates with the GUSTO I trial (GUSTO investigators 1993). The GUSTO I trial was a comparison of tissue plasminogen activator (t-PA) and streptokinase in patients with a myocardial infarction (MI). The primary outcome was death by 30 days. The unadjusted OR for mortality was 0.853, changing to 0.829 when adjusted for age. Corresponding to logistic regression coefficients and standard errors of -0.1586 (0.09) and -0.1878 (0.050) respectively. This gave an increase of 18% on the logistic scale for the treatment effect and an increase of 3% for the standard errors. However, this increase was less than the increase due to stratification. Robinson and Jewell (Robinson & Jewell 1991) showed that adjusting for covariates is always more statistically efficient when using logistic regression despite the increase in standard error.

Covariate adjustment provides individualised (conditional) effect estimates in logistic regression in comparison to an unadjusted analysis which gives a population effect (Hernandez, Steyerberg, & Habbema 2004). The unadjusted (unconditional) estimate is a biased estimate; therefore all of the simulation modelling in this thesis uses covariate adjusted multinomial regression, details of which are given in Chapter 5.

3.5 Proportional odds application

As discussed by McHugh et al (McHugh et al. 2007a) a binary covariate such as hypotension (no/yes) can be summarised as a 2x4 table for each study. Table 3-3 below shows the data for the EBIC study. A far greater percentage of those with hypotension died or were in a vegetative state compared to those without hypotension.

Table 3-3 Relationship between hypotension on admission and six month GOS in EBIC

hypotension / GOS N (row %)	Death/ Vegetative	Severe Disability	Moderate Disability	Good Recovery	Total
Yes	116 (57.7)	30 (14.9)	28 (13.9)	27 (13.4)	201
No	182 (29.0)	94 (15.0)	131 (20.9)	221 (35.2)	628

Using the conventional approach to prognostic analysis with the GOS dichotomised into favourable and unfavourable the table would be as shown in Table 3-4.

Table 3-4 Relationship between hypotension on admission and favourable six month GOS in EBIC

hypotension / GOS N (row %)	Unfavourable outcome	Favourable outcome	Total
Yes	146 (72.6)	55 (27.4)	201
No	276 (43.9)	352 (56.1)	628

The corresponding odds ratio for this table is (146/55)/(276/352)=3.39. Table 3-3 can also be dichotomised as dead/vegetative versus better outcome or as not good versus a good outcome, giving odds ratios of 3.34 and 3.50 respectively.

The three estimated odds ratios (3.39, 3.34 and 3.50) are very similar. The proportional odds model assumes that no matter where the outcome is dichotomised the true underlying odds ratio is the same. The estimate of this common odds ratio is 3.38 95%CI (2.49 to 4.60). That is having hypotension shifts the distribution of outcome in a negative direction. This estimate of 3.38 is very close to the estimates from the three dichotomous splits.

Goodness of fit of the proportional odds model can be examined in a variety of ways, using Generalised Estimating Equations (GEE) (Liang & Zeger 1986), comparing score and Wald tests (Stiger, Barnhart, & Williamson 1999) and comparing observed and expected values (Ashby, Pocock, & Shaper 1986). Here the score test for goodness of fit of the proportion odds model gives p=0.976; showing that the proportional odds model fits well for these data.

In other IMPACT studies however the goodness of fit test showed that the proportional odds assumption did not hold. Table 3-5 below shows hypotension in the UK4 centres study.

Table 3-5 Relationship between hypotension on admission and six month GOS in UK4

hypotension / GOS N (row %)	Death/ Vegetative	Severe Disability	Moderate Disability	Good Recovery	Total
Yes	161 (67.6)	24 (10.1)	22 (9.2)	31 (13.0)	238
No	233 (33.2)	138 (19.7)	133 (18.9)	198 (28.2)	702

The odds ratio for comparing unfavourable versus favourable is 3.11; death and vegetative state versus better is 4.21; and not good recovery versus good recovery is 2.62. The formal goodness of fit test gives p=0.016 showing there is evidence that the underlying odds ratios are not identical. However the three estimates are in the same direction and of similar magnitude suggesting that the odds ratio obtained from the proportional odds model 3.74 95%CI (2.78 to 5.04) is still a sensible summary of the three odds ratios from the dichotomous splits even though the formal assumption of proportional odds does not hold.

This method can be extended to covariates with more than two levels. A reference category is chosen, usually the most clinically meaningful and quite common. The other categories are then individually compared back to this reference category. For example, with place of injury, street/highway was taken as the reference category and the other categories e.g. home or recreational etc. were compared back to the reference category. Occasionally, within each study, some categorical variables had no observations in one cell of the table of GOS by variable. Where this occurred, one extra observation was added to each cell of the table to avoid dividing by zero. This is typically done in the analysis of categorical data (Agresti 2002).

Proportional odds analysis can also be used in the analysis of continuous variables. As shown above the log odds of, say, an unfavourable outcome can be written as a linear function. As with the analysis of binary or categorical variables, the regression

modelling can be performed for different dichotomies of the GOS. A common slope, assuming all of the individual slopes are similar, can be modelled to allow fitting of the proportional odds model for continuous data.

Each of the levels of all variables was tested to see if the assumption of proportional odds held. The assumption of proportional odds was not violated in most cases. This test has however been shown to perform poorly. When modelling data no model is, of-course, 'correct' - "Essentially, all models are wrong, but some are useful" (Box & Draper 1987). As with any model fitting the choice of model can be quite arbitrary and there may be other models which fit the data equally well (Copas & Eguchi 2010). One of the most important things is to develop a prognostic model which fits the data well (Harrell, Jr., Lee, & Mark 1996). This is something that a goodness of fit test can perform poorly at. Generally goodness of fit tests are irrelevant: with small numbers they will not pick up deviations; and with large samples, as studied here, the test is potentially over-sensitive. This is particularly illustrated in the goodness of fit test for proportional odds. For sparse data or where an explanatory variable is continuous it has been shown to perform poorly (Peterson & Harrell 1990). Also, where the data are not sparse it has been shown to be too liberal in that it may show a significant result with only small departures from proportional odds (Peterson & Harrell 1990). It has also been shown that the conventional Pearson's Chi Squared test used to assess goodness of fit may not be valid when the model includes covariates (Pulkstenis & Robinson 2004). Fitting the proportional odds model can be thought of as being analogous to meta analysis. The final estimate, as with that from a proportional odds model, gives an overall summary measure of the data. Even when the assumption of proportional odds is violated the pooled estimate might still be judged to give a useful summary measure for the variable of interest.

3.6 Meta analysis introduction

Meta analysis allows results to be combined from different studies and can provide a quantitative measure of 'treatment' effect allowing conclusions to be drawn about therapeutic effectiveness (L'Abbe, Detsky, & O'Rourke 1987;Whitehead & Whitehead 1991). Thompson and Pocock stated that "meta analysis has the potential to remove idiosyncrasy, …, however it is unrealistic to imagine that it will produce simple statistical answers to complex clinical problems" (Thompson & Pocock 1991). There are two main ways of combining data in a meta analysis using either a fixed or random effects analysis. A fixed effects analysis assumes that the studies observed are from the same population and that the underlying effect is the same in each trial i.e. the studies are homogeneous (DeMets 1987;Thompson & Pocock 1991). A random effects analysis assumes that the studies observed are from a random sample of possible studies and that the true underlying effect varies between studies i.e. it takes into account the heterogeneity between studies (DeMets 1987).

Cochran's Q test is used to test for heterogeneity. The test statistic, Q, is a weighted sum of squares of deviations. When treatment effects are homogeneous it follows a Chi-squared distribution with (k-1) degrees of freedom, where k is the total number of studies under consideration. The test is however known to be insensitive for detecting true heterogeneity amongst studies as statistically significant (Higgins et al. 2003). Therefore a cut off of $p<0.1$ is conventionally used to assess heterogeneity although this increases the type I error rate.

Using conventional notation:

$$Q=\sum w_i\left(\hat{\theta}_i-\hat{\theta}\right)^2$$

where $\hat{\theta}_i$ is the *i*th treatment effect estimate (here the log odds ratio for the proportional odds model), $\hat{\theta}$ is the estimate of the overall effect and w_i is the reciprocal of the variance of $\hat{\theta}_i$, $\frac{1}{\left(SE\left\{\hat{\theta}_i\right\}\right)^2}$

3.7 Meta analysis application

Meta analysis was performed on each variable separately. The proportional odds model was fitted for each trial, with each level of covariate analysed separately and compared to a reference category in a similar manner as described in Chapter 2. These resulting proportional odds estimates were pooled over all the studies. The process used for fixed and random effects pooling is described below.

3.7.1 Fixed effects pooling

Under a fixed effects model, using the inverse variance method to combine studies with Whiteheads' notation (Whitehead & Whitehead 1991)

$\hat{\theta} = \frac{\sum \hat{\theta}_i w_i}{\sum w_i}$ with the terms as defined above.

It is assumed that

$\hat{\theta}_i \sim N(\theta_i, w_i^{-1})$ for i=1, …, k.

To test the global null hypothesis H_0: $\theta_1 = \ldots = \theta_k$ the statistic U can be used, defined below. It has a Chi-squared distribution with one degree of freedom under the null hypothesis. This assumes that there is a common 'treatment' effect in all studies.

$U=\frac{\left(\sum \hat{\theta}_i w_i\right)^2}{\sum w_i}$ with corresponding $SE=\frac{1}{\sqrt{\sum w_i}}$

Therefore a 95% CI for $\hat{\theta}$ is $\hat{\theta} \pm 1.96\sqrt{\frac{1}{\sum w_i}}$

with the 95% CI for the odds ratio obtained by exponentiation.

For example, for traumatic subarachnoid haemorrhage (tSAH) the odds ratios from the proportional odds model for the individual studies are shown in the Table 3-6 below.

Table 3-6 tSAH estimates from proportional odds model by study

Study	OR	95% CI	$\hat{\theta}_i$	SE	w_i	$\hat{\theta}_i w_i$	$\hat{\theta}_i^2 w_i$
TINT	2.72	(2.17,3.41)	1.000	0.116	74.517	74.491	74.465
TIUS	2.48	(1.95,3.14)	0.906	0.121	68.717	62.291	56.465
SLIN	2.39	(1.53,3.73)	0.871	0.228	19.290	16.796	14.624
SAP	2.51	(1.97,3.20)	0.921	0.124	64.984	59.849	55.120
PEG	2.30	(1.90,2.78)	0.832	0.096	108.108	89.998	74.922
HIT1	3.90	(2.32,6.57)	1.361	0.266	14.126	19.224	26.163
TCDB	3.18	(2.26,4.46)	1.156	0.173	33.383	38.605	44.645
SKB	2.73	(1.25,5.98)	1.005	0.399	6.279	6.313	6.347
EBIC	3.24	(2.48,4.23)	1.174	0.136	53.834	63.214	74.227
HIT II	2.64	(2.01,3.45)	0.969	0.138	52.718	51.070	49.474
Sum					495.956	481.851	476.452

Using the formula above to obtain the fixed effect pooled estimate and its 95%CI

$$\hat{\theta} = \frac{\sum \hat{\theta}_i w_i}{\sum w_i} = \frac{481.851}{495.956} = 0.97$$

$$\text{SE} = \frac{1}{\sqrt{\sum w_i}} = \frac{1}{\sqrt{495.959}} = 0.045$$

Odds ratio and 95%CI 2.64 (2.42 to 2.89)

Testing the null hypothesis:

$$\text{U} = \frac{\left(\sum \hat{\theta}_i w_i\right)^2}{\sum w_i} = \frac{(481.857)^2}{495.956} = 468.147 \qquad p<0.0001$$

Testing heterogeneity

$$Q = \sum w_i \left(\hat{\theta}_i - \hat{\theta}\right)^2 = \sum \hat{\theta}_i^2 w_i - U = 476.452 - 468.147 = 8.30 \text{ (9df)} \qquad p=0.50$$

Thus there is strong evidence of an association between tSAH and outcome, p<0.0001 with little evidence of heterogeneity.

Other variables however do show statistically significant heterogeneity. For example Table 3-7 below shows estimates for hypoxia in the eight studies that recorded it.

Table 3-7 Hypoxia estimates from proportional odds model by study

Study	$\hat{\theta}_i$	SE	w_i	$\hat{\theta}_i w_i$	$\hat{\theta}_i^2 w_i$
TINT	0.702	0.164	37.164	26.103	18.333
TIUS	0.938	0.135	55.068	51.655	48.454
SLIN	-0.209	0.382	6.837	-1.461	0.299
SAP	0.703	0.186	28.954	20.359	14.316
UK4	0.609	0.146	47.160	28.712	17.481
TCDB	0.638	0.209	22.948	14.649	9.351
SKB	0.536	0.416	5.790	3.104	1.665
EBIC	1.165	0.147	46.043	53.623	62.451
Sum			249.964	196.776	172.350

In a similar manner to above, the pooled fixed effect estimate, 95% Confidence Interval and Cochrane's Q are calculated.

$$\hat{\theta} = \frac{\sum \hat{\theta}_i w_i}{\sum w_i} = \frac{196.776}{249.964} = 0.79$$

$$\text{SE} = \frac{1}{\sqrt{\sum w_i}} = \frac{1}{\sqrt{249.964}} = 0.06$$

Odds ratio and 95%CI 2.20 (1.94 to 2.49)

Testing the null hypothesis:

$$\text{U} = \frac{\left(\sum \hat{\theta}_i w_i\right)^2}{\sum w_i} = \frac{(196.776)^2}{249.964} = 154.905 \qquad p<0.0001$$

Test of heterogeneity:

$Q=\sum w_i\left(\hat{\theta}_i-\hat{\theta}\right)^2$ = 196.776-154.905=17.45 (7df) p=0.01

It can be observed that there is statistically significant heterogeneity between the studies for hypoxia.

3.7.2 Random effects pooling

The fixed effect estimate takes no account of variation between studies. A random effects model relaxes the assumption of a common intervention effect. Using the DerSimonian and Laird (DerSimonian & Laird 1986) random effects model with Whiteheads' notation (Whitehead & Whitehead 1991).

$\hat{\theta}_i \sim N\left(\theta, w_i^{-1}+\hat{\tau}^2\right)$ for i=1, …, k

Letting $w_i^* = \left(w_i^{-1}+\hat{\tau}^2\right)^{-1}$ gives

$\hat{\theta}_i \sim N\left(\theta, \left(w_i^*\right)^{-1}\right)$

where $\hat{\tau}^2$ is a measure of the degree to which the 'treatment' effect varies across studies and the degree to which individual studies give biased estimates of 'treatment' effects.

$$\hat{\tau}^2 = \frac{Q-(k-1)}{\sum w_i - \left(\sum w_i^2\right)/\sum w_i}$$

with $Q=\sum w_i\left(\hat{\theta}_i-\hat{\theta}\right)^2$ as with the fixed effect model.

The test statistic for the random effects model U* is denoted as below. As with the fixed effects model, this follows a Chi-squared distribution with one degree of freedom under the null hypothesis.

$$U^* = \frac{\left(\sum \hat{\theta}_i w_i^*\right)^2}{\sum w_i^*} \text{ with } \hat{\theta}^* = \frac{\sum \hat{\theta}_i w_i^*}{\sum w_i^*}$$

Extending Table 3-7 to show w_i^* and $\hat{\theta}\, w_i^*$ below in Table 3-8.

Table 3-8 Hypoxia estimates from random effects model

Study	$\hat{\theta}_i$	SE	w_i	w_i^*	θw_i^*
TINT	0.702	0.164	37.164	13.012	9.139
TIUS	0.938	0.135	55.068	14.684	13.774
SLIN	-0.209	0.382	6.837	5.097	-1.066
SAP	0.703	0.186	28.954	11.837	8.323
UK4	0.609	0.146	47.160	14.055	8.557
TCDB	0.638	0.209	22.948	10.693	6.826
SKB	0.536	0.416	5.790	4.491	2.408
EBIC	1.165	0.147	46.043	13.954	16.252
Sum			249.964	87.823	64.213

Using the figures for hypoxia in Table 3-8 gives the following.

$$\hat{\tau}^2 = \frac{Q-(k-1)}{\sum w_i - \left(\sum w_i^2\right)/\sum w_i} = \frac{17.446-(8-1)}{249.964-10202.918/\ 249.964} = 0.05$$

$$\hat{\theta}^* = \frac{\sum \hat{\theta}_i w_i^*}{\sum w_i^*} = \frac{64.213}{87.823} = 0.73$$

$$\text{SE}\,\hat{\theta}^* = \frac{1}{\sqrt{\sum w_i^*}} = \frac{1}{\sqrt{87.823}} = 0.107$$

Odds ratio and 95%CI 2.08 (1.68 to 2.56)

Using a random effects model gives a similar estimate, albeit slightly smaller, and a wider confidence interval than the estimates from the fixed effects model. As can be seen in Table 3-8 the weights from the random effects model, w_i^*, are closer to each other than the weights from the fixed effects model, w_i.

3.8 Discussion

This chapter has shown the theory for both the binary and proportional odds models used in subsequent chapters. The methodology for meta analysis, both fixed and random effects has also been shown. For the studies in the IMPACT database both the fixed and random effects estimates were calculated. However, as the assumption of homogeneity between the different studies seemed implausible, the proportional odds modelling shown in subsequent chapters uses the pooled estimates from the random effects analysis.

4 Chapter 4 Analyses using IMPACT data

4.1 Introduction

The next stage was to model the IMPACT data using both unadjusted and adjusted analyses. Known and novel predictors of outcome were examined. Using the methods in Chapter 3, models (both logistic and proportional odds) were fitted separately for each study. The resulting odds ratios were pooled by using random effects models.

In the first part of this chapter univariate analyses are performed on all four possible binary splits of the five point GOS. These show the odds of an unfavourable outcome for each level of each covariate compared back to a reference category. These results were then examined to see if fitting a proportional odds model using the full outcome scale would provide a useful summary measure of the four binary splits. Examining the results this seemed a reasonable assumption. The second part of this chapter shows the unadjusted pooled random effects estimates of the common odds ratios from the proportional odds model. Subsequently, multivariate analyses are then performed, adjusting the results obtained from the proportional odds model for covariates that were considered to be both a-priori important and also those that showed statistically significant associations with outcome in the univariate proportional odds analysis. Four sets of covariate models with three, four, seven and nine covariates were developed.

In the tables below showing the associations, Global is the U test Global test statistic assuming that the difference between categories across studies=0 and Q is the test for heterogeneity as described in Chapter 3. The p-value for the Q test is only shown to two decimal places as $p<0.1$ is considered to be statistically significant.

4.2 Binary analysis on pooled data – four dichotomies

The odds of an unfavourable outcome were estimated for each level of each covariate referred back to a reference category. This was repeated four times for each for the four possible splits of the five point GOS. Table 4-1 below shows each of the four splits labelled by the 'unfavourable' outcome. Items in bold below the double line were the outcome of interest for each of the four modelling strategies. These outcomes were modelled using logistic regression.

Table 4-1 GOS with four binary splits

Mortality	**Dead/Vegetative State**	**Unfavourable**	**Not Good**
Good Recovery	Good Recovery	Good Recovery	Good Recovery
Moderate Disability	Moderate Disability	Moderate Disability	**Moderate Disability**
Severe Disability	Severe Disability	**Severe Disability**	**Severe Disability**
Vegetative State	**Vegetative State**	**Vegetative State**	**Vegetative State**
Death	**Death**	**Death**	**Death**

4.2.1 Dichotomous analysis – mortality outcome

The first binary outcome was death versus good outcome, moderate disability, severe disability and vegetative status. The results for all of the variables considered are shown below in Table 4-2. The odds ratio for e.g. gender shows the odds of an unfavourable outcome for females compared to males. As can be seen there is very little difference in the odds of dying between men and women p=0.925. The odds ratio is almost one (1.01) with a narrow 95% confidence interval (0.90 to 1.13). These results are homogeneous across the 11 studies, p=0.50. The odds ratios for age, pH, haemoglobin, glucose, haematocrit, platelets and prothrombin time are scaled by their interquartile range over all studies as described in Chapter 2. Little differences were observed in outcome between races, years of education, primary or

secondary referral and contusions with most other variables showing an association with mortality. Having a traumatic subarachnoid haemorrhage, increasing age, CT class of swelling or shift, GCS of 3-5, not having a motor response of localises or obeys and having lower lab values (with the exception of glucose for which a higher value had a stronger association) were very strongly associated with death. Having an epidural haematoma was strongly associated with a better outcome after injury, which may be explained by the potential ease with which these haematomas can be surgically evacuated. With an epidural haematoma, brain function is disturbed because of compression although there is generally little intrinsic brain damage. If compression is relieved promptly, full recovery is more likely to occur (Steyerberg et al. 2008). Little statistically significant heterogeneity between studies was observed with the exceptions of hypotension, cisterns, contusions and some of the GCS components which did show statistically significant heterogeneity between studies.

Table 4-2 Pooled random effects estimates of binary odds ratios for mortality outcome (unadjusted)

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	1.01	(0.90 to 1.13)	0.01	0.925	9.35	0.50
Race	Caucasian						
	Black	1.24	(0.98 to 1.55)	3.30	0.069	4.17	0.53
	Asian	1.42	(0.94 to 2.15)	2.84	0.092	2.69	0.61
	other	1.15	(0.77 to 1.71)	0.46	0.496	5.73	0.22
Education	0-8 years						
	9-12 years	0.92	(0.56 to 1.54)	0.09	0.763	3.08	0.21
	> 12 years	0.87	(0.57 to 1.34)	0.39	0.531	2.30	0.32

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Place Injury	street/highway						
	home	1.82	(1.34 to 2.48)	14.85	<0.001	0.98	0.61
	work/school	1.20	(0.69 to 2.08)	0.40	0.527	5.11	0.08
	recreational	0.50	(0.28 to 0.87)	5.91	0.015	0.06	0.97
	other	1.29	(0.88 to 1.89)	1.70	0.192	1.25	0.54
Cause Injury	domestic/fall						
	road traffic accident	0.54	(0.48 to 0.61)	99.78	<0.001	3.84	0.95
	assault	0.56	(0.44 to 0.71)	21.96	<0.001	9.92	0.45
	work-related	0.84	(0.61 to 1.16)	1.12	0.291	4.49	0.81
	sports/recreation	0.40	(0.23 to 0.71)	9.91	0.002	13.52	0.10
	other	0.78	(0.63 to 0.98)	4.74	0.030	7.21	0.62
Referral	primary						
	secondary	0.99	(0.87 to 1.13)	0.01	0.913	2.00	0.85
Hypoxia	no						
	suspected/definite	2.02	(1.61 to 2.55)	36.18	<0.001	16.18	0.02
Hypotension	no						
	suspected/definite	2.62	(1.99 to 3.47)	46.19	<0.001	29.42	<0.01
Hypothermia	no						
	suspected/definite	2.11	(1.35 to 3.30)	10.61	0.001	14.15	0.01
CT class	diffuse class II						
	no visible pathology	0.51	(0.28 to 0.92)	5.03	0.025	8.82	0.18
	swelling/shift	3.50	(2.67 to 4.60)	81.60	<0.001	12.06	0.06
	mass	2.88	(2.45 to 3.39)	162.52	<0.001	4.73	0.58
Cisterns	present						
	compressed/absent	2.74	(1.98 to 3.78)	37.43	<0.001	19.48	<0.01
Shift	no						
	1-5 mm	1.62	(1.30 to 2.03)	18.26	<0.001	8.17	0.23
	> 5mm	2.66	(2.04 to 3.480	51.84	<0.001	15.13	0.02
tSAH	no						
	yes	2.82	(2.52 to 3.16)	325.88	<0.001	9.19	0.42

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Epidural Haematoma (EDH)	no						
	yes	0.68	(0.57 to 0.80)	19.75	<0.001	8.77	0.36
Subdural Haematoma (SDH)	no						
	yes	2.41	(2.11 to 2.76)	169.04	<0.001	11.26	0.19
Contusion	no						
	yes	1.18	(0.89 to 1.57)	1.32	0.251	31.27	<0.01
Eye	none						
	pain/sound/ spontaneous	0.42	(0.30 to 0.59)	25.90	<0.001	41.13	<0.01
	missing/untestable	0.66	(0.50 to 0.87)	8.53	0.004	9.43	0.22
Verbal	none						
	sounds-orientated	0.38	(0.29 to 0.49)	56.00	<0.001	22.86	0.01
	missing/untestable	0.88	(0.68 to 1.15)	0.87	0.350	20.16	0.01
Motor	localises/obeys						
	none	5.73	(3.72 to 8.82)	62.87	<0.001	40.08	<0.01
	extension	5.46	(3.95 to 7.55)	105.91	<0.001	34.09	<0.01
	abnormal flexion	2.75	(2.12 to 3.56)	58.58	<0.001	19.65	0.03
	normal flexion	1.60	(1.28 to 1.99)	17.67	<0.001	17.80	0.06
	missing/untestable	1.94	(1.50 to 2.51)	25.64	<0.001	5.19	0.52
Glasgow Coma Scale (GCS)	6-8						
	3-5	3.54	(2.72 to 4.61)	88.53	<0.001	31.12	<0.01
	9-15	0.71	(0.43 to 1.16)	1.90	0.168	36.57	<0.01
	missing/untestable	1.98	(1.67 to 2.36)	60.17	<0.001	3.08	0.88
Pupil	both sides +ve						
	one side +ve	2.38	(2.02 to 2.79)	111.39	<0.001	6.69	0.57
	both side -ve	5.93	(4.42 to 7.96)	141.09	<0.001	37.05	<0.01
Systolic Blood Pressure (SBP)	120-150mmHg						
	<120 mmHg	1.65	(1.23 to 2.21)	11.23	0.001	34.39	<0.01
	>150 mmHg	1.54	(1.32 to 1.79)	30.37	<0.001	8.72	0.37

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Mean Arterial Blood Pressure (MABP)	85-110 mmHg						
	<85 mmHg	1.42	(1.12 to 1.80)	8.25	0.004	23.72	<0.01
	>110 mmHg	1.47	(1.24 to 1.73)	19.89	<0.001	7.54	0.48
Sodium	137-142 mmol/L						
	<137 mmol/L	1.35	(1.12 to 1.62)	10.04	0.002	7.61	0.27
	>142 mmol/L	1.21	(0.99 to 1.48)	3.51	0.061	8.51	0.20
Age	/24 years	2.08	(1.91 to 2.26)	288.56	<0.001	12.19	0.27
pH	/0.15	0.73	(0.64 to 0.83)	22.07	<0.001	6.14	0.19
Haemoglobin	/3.3 g/dL	0.69	(0.58 to 0.82)	18.56	<0.001	10.12	0.07
Glucose	/3.7 mmol/L	1.85	(1.70 to 2.01)	196.41	<0.001	5.86	0.32
Haematocrit	/10 %	0.69	(0.55 to 0.87)	9.97	0.002	2.65	0.27
Platelets	/100 $x10^9$/L	0.65	(0.55 to 0.77)	24.82	<0.001	1.97	0.58
Prothrombin time	/2 seconds	1.64	(1.34 to 2.00)	23.75	<0.001	3.21	0.20

4.2.2 Dichotomous analysis – dead/vegetative outcome

Table 4-3 below shows the odds ratios when dead/vegetative is taken as 'unfavourable outcome'. Similar results are observed as with the mortality outcome, with increased odds of an unfavourable outcome for older patients, those having a traumatic subarachnoid haemorrhage, lower GCS scores and non reacting pupils. Lower values of the laboratory parameters, except glucose, are again strongly associated with a poorer outcome. Little statistically significant heterogeneity between studies was observed for most variables however it was observed with hypotension, cisterns and some of the GCS components.

Table 4-3 Pooled random effects estimates of binary odds ratios for dead/vegetative outcome (unadjusted)

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	1.03	(0.93 to 1.15)	0.37	0.542	5.94	0.82
Race	Caucasian						
	Black	1.20	(0.97 to 1.49)	2.72	0.099	2.79	0.73
	Asian	1.20	(0.75 to 1.91)	0.58	0.447	5.02	0.28
	other	1.08	(0.83 to 1.40)	0.36	0.549	3.96	0.41
Education	0-8 years						
	9-12 years	0.79	(0.58 to 1.07)	2.41	0.121	1.20	0.55
	> 12 years	0.81	(0.58 to 1.12)	1.58	0.209	1.07	0.58
Place Injury	street/highway						
	home	1.63	(1.21 to 2.21)	10.10	0.002	0.86	0.65
	work/school	1.39	(0.93 to 2.07)	2.58	0.108	3.07	0.21
	recreational	0.59	(0.35 to 0.99)	4.05	0.044	1.39	0.50
	other	1.35	(0.93 to 1.94)	2.55	0.110	0.15	0.93
Cause Injury	domestic/fall						
	road traffic accident	0.60	(0.53 to 0.68)	72.83	<0.001	4.60	0.92
	assault	0.57	(0.45 to 0.72)	21.68	<0.001	7.11	0.71
	work-related	0.84	(0.62 to 1.15)	1.14	0.286	3.44	0.90
	sports/recreation	0.42	(0.25 to 0.71)	10.42	0.001	12.75	0.12
	other	0.81	(0.66 to 1.01)	3.64	0.056	2.58	0.98
Referral	primary						
	secondary	0.97	(0.85 to 1.10)	0.28	0.595	3.24	0.66
Hypoxia	no						
	suspected/definite	2.08	(1.71 to 2.54)	53.54	<0.001	12.59	0.08
Hypotension	no						
	suspected/definite	2.68	(2.02 to 3.54)	47.42	<0.001	30.61	<0.01
Hypothermia	no						
	suspected/definite	2.13	(1.45 to 3.14)	14.60	<0.001	11.30	0.02

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
CT class	diffuse class II						
	no visible pathology	0.43	(0.24 to 0.78)	7.69	0.006	9.13	0.17
	swelling/shift	3.50	(2.75 to 4.45)	104.55	<0.001	10.45	0.11
	mass	2.86	(2.45 to 3.33)	178.21	<0.001	5.74	0.45
Cisterns	present						
	compressed/absent	2.86	(2.06 to 3.95)	39.95	<0.001	21.76	<0.01
Shift	no						
	1-5 mm	1.61	(1.28 to 2.03)	16.61	<0.001	9.39	0.15
	> 5mm	2.50	(1.93 to 3.24)	48.78	<0.001	14.93	0.02
tSAH	no						
	yes	3.00	(2.66 to 3.40)	307.65	<0.001	11.19	0.26
EDH	no						
	yes	0.63	(0.54 to 0.73)	33.74	<0.001	6.05	0.64
SDH	no						
	yes	2.26	(2.04 to 2.51)	234.13	<0.001	6.86	0.55
Contusion	no						
	yes	1.25	(0.97 to 1.62)	2.94	0.086	29.15	<0.01
Eye	none						
	pain/sound/ spontaneous	0.36	(0.26 to 0.51)	33.72	<0.001	47.16	<0.01
	missing/untestable	0.65	(0.50 to 0.85)	9.80	0.002	9.54	0.22
Verbal	none						
	sounds-orientated	0.35	(0.27 to 0.45)	62.88	<0.001	25.33	<0.01
	missing/untestable	0.92	(0.71 to 1.19)	0.41	0.523	20.19	0.01
Motor	localises/obeys						
	none	6.07	(3.70 to 9.95)	51.02	<0.001	55.23	<0.01
	extension	7.29	(5.00 to 0.63)	106.39	<0.001	47.76	<0.01
	abnormal flexion	3.39	(2.58 to 4.45)	77.33	<0.001	23.30	0.01
	normal flexion	1.74	(1.36 to 2.21)	20.02	<0.001	23.29	0.01
	missing/untestable	2.16	(1.68 to 2.77)	36.25	<0.001	4.33	0.63

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
GCS	6-8						
	3-5	3.89	(3.00 to 5.04)	105.55	<0.001	31.97	<0.01
	9-15	0.68	(0.43 to 1.09)	2.59	0.108	36.10	<0.01
	missing/untestable	2.07	(1.75 to 2.44)	72.13	<0.001	4.14	0.76
Pupil	both sides +ve						
	one side +ve	2.55	(2.19 to 2.97)	142.02	<0.001	7.68	0.46
	both side -ve	7.32	(5.62 to 9.54)	217.55	<0.001	30.05	<0.01
SBP	120-150mmHg						
	<120 mmHg	1.63	(1.27 to 2.10)	14.89	<0.001	27.26	<0.01
	>150 mmHg	1.46	(1.25 to 1.69)	23.73	<0.001	9.26	0.32
MABP	85-110 mmHg						
	<85 mmHg	1.44	(1.18 to 1.76)	12.73	<0.001	18.30	0.02
	>110 mmHg	1.49	(1.22 to 1.83)	15.57	<0.001	10.83	0.21
Sodium	137-142 mmol/L						
	<137 mmol/L	1.51	(1.30 to 1.75)	29.57	<0.001	5.79	0.45
	>142 mmol/L	1.27	(1.02 to 1.59)	4.44	0.035	11.09	0.09
Age	/24 years	2.06	(1.92 to 2.21)	379.72	<0.001	9.57	0.48
pH	/0.15	0.78	(0.71 to 0.87)	21.81	<0.001	4.26	0.37
Haemoglobin	/3.3 g/dL	0.69	(0.60 to 0.81)	22.50	<0.001	9.04	0.11
Glucose	/3.7 mmol/L	1.80	(1.61 to 2.00)	111.82	<0.001	9.00	0.11
Haematocrit	/10 %	0.69	(0.56 to 0.86)	11.16	0.001	2.57	0.28
Platelets	/100 $x10^9$/L	0.69	(0.59 to 0.81)	21.57	<0.001	0.88	0.83
Prothrombin time	/2 seconds	1.66	(1.34 to 2.06)	21.09	<0.001	3.60	0.17

4.2.3 Dichotomous analysis - unfavourable outcome

Table 4-4 shows the odds ratios and 95% confidence intervals for the conventional dichotomous split of the GOS grouping death, vegetative state and severe disability together and moderate recovery and good recovery together. As with the previous

two dichotomies a similar pattern is seen between the variables and outcome. Older age, having a tSAH or a CT classification of swelling, shift or mass were very strongly associated with an unfavourable outcome. Statistically significant heterogeneity was observed between studies for compressed or absent cisterns, contusions, shift>5mm, GCS 9-15 and some categories of the GCS components, and bilateral non reacting pupils.

Table 4-4 Pooled random effects estimates of binary odds ratios for unfavourable outcome (unadjusted)

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	1.04	(0.93 to1.16)	0.40	0.526	11.14	0.35
Race	Caucasian						
	Black	1.32	(1.00 to 1.73)	3.83	0.050	6.63	0.25
	Asian	1.15	(0.79 to 1.69)	0.54	0.461	2.42	0.66
	other	1.03	(0.81 to 1.31)	0.06	0.808	1.97	0.74
Education	0-8 years						
	9-12 years	0.62	(0.37 to 1.03)	3.36	0.067	4.64	0.10
	> 12 years	0.59	(0.39 to 0.91)	5.63	0.018	3.17	0.20
Place Injury	street/highway						
	home	1.53	(1.14 to 2.06)	7.94	0.005	0.70	0.71
	work/school	1.43	(1.05 to 1.93)	5.31	0.021	1.40	0.50
	recreational	0.89	(0.59 to 1.35)	0.30	0.581	0.09	0.96
	other	1.23	(0.86 to 1.75)	1.29	0.256	0.33	0.85
Cause Injury	domestic/fall						
	road traffic accident	0.68	(0.60 to 0.76)	45.81	<0.001	7.39	0.69
	assault	0.65	(0.46 to 0.92)	5.95	0.015	20.66	0.02
	work-related	0.73	(0.54 to 1.00)	3.89	0.048	4.66	0.79
	sports/recreation	0.45	(0.29 to 0.71)	11.57	0.001	13.19	0.11
	other	0.95	(0.77 to 1.16)	0.26	0.611	3.61	0.94
Referral	primary						
	secondary	1.06	(0.91 to 1.23)	0.54	0.462	7.10	0.21

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Hypoxia	no						
	suspected/definite	2.14	(1.72 to 2.67)	45.90	<0.001	15.55	0.03
Hypotension	no						
	suspected/definite	2.67	(2.18 to 3.28)	87.47	<0.001	15.95	0.04
Hypothermia	no						
	suspected/definite	2.27	(1.66 to 3.11)	26.03	<0.001	7.49	0.11
CT class	diffuse class II						
	no visible pathology	0.35	(0.25 to 0.49)	39.90	<0.001	5.58	0.47
	swelling/shift	2.54	(2.06 to 3.14)	76.10	<0.001	9.76	0.14
	mass	2.13	(1.69 to 2.67)	42.01	<0.001	15.88	0.01
Cisterns	present						
	compressed/absent	2.30	(1.71 to 3.11)	30.06	<0.001	22.27	<0.01
Shift	no						
	1-5 mm	1.34	(1.03 to 1.74)	4.75	0.029	13.42	0.04
	> 5mm	2.10	(1.53 to 2.88)	20.86	<0.001	24.49	<0.01
tSAH	no						
	yes	2.70	(2.45 to 2.98)	388.40	<0.001	7.27	0.61
EDH	no						
	yes	0.60	(0.52 to 0.69)	51.30	<0.001	6.71	0.57
SDH	no						
	yes	2.16	(1.88 to 2.47)	122.58	<0.001	13.28	0.10
Contusion	no						
	yes	1.31	(1.06 to 1.60)	6.47	0.011	22.18	<0.01
Eye	none						
	pain/sound/ spontaneous	0.35	(0.27 to 0.44)	75.41	<0.001	32.01	<0.01
	missing/untestable	0.75	(0.54 to 1.04)	3.00	0.083	14.16	0.05
Verbal	none						
	sounds-orientated	0.38	(0.30 to 0.48)	70.10	<0.001	26.00	<0.01
	missing/untestable	1.04	(0.81 to 1.35)	0.10	0.749	20.79	<0.01

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Motor	localises/obeys						
	none	5.35	(3.69 to 7.75)	78.54	<0.001	36.11	<0.01
	extension	8.90	(6.93 to 11.42)	294.68	<0.001	19.89	0.03
	abnormal flexion	4.05	(2.95 to 5.57)	74.32	<0.001	37.83	<0.01
	normal flexion	1.94	(1.64 to 2.30)	58.34	<0.001	16.49	0.09
	missing/untestable	2.26	(1.60 to 3.19)	21.30	<0.001	9.77	0.13
GCS	6-8						
	3-5	3.68	(3.04 to 4.45)	178.56	<0.001	19.38	0.04
	9-15	0.58	(0.42 to 0.81)	10.28	0.001	26.31	<0.01
	missing/untestable	2.06	(1.71 to 2.48)	59.31	<0.001	9.38	0.23
Pupil	both sides +ve						
	one side +ve	2.67	(2.21 to 3.23)	102.32	<0.001	12.61	0.13
	both side -ve	6.92	(5.03 to 9.52)	141.58	<0.001	38.26	<0.01
SBP	120-150mmHg						
	<120 mmHg	1.50	(1.28 to 1.75)	26.44	<0.001	12.13	0.15
	>150 mmHg	1.47	(1.19 to 1.81)	12.65	<0.001	19.27	0.01
MABP	85-110 mmHg						
	<85 mmHg	1.23	(1.04 to 1.45)	6.07	0.014	14.72	0.06
	>110 mmHg	1.45	(1.18 to 1.77)	12.81	<0.001	12.05	0.15
Sodium	137-142 mmol/L						
	<137 mmol/L	1.43	(1.24 to 1.65)	25.14	<0.001	5.77	0.45
	>142 mmol/L	1.14	(0.95 to 1.36)	1.91	0.167	8.50	0.20
Age	/24 years	2.15	(2.00 to 2.31)	428.66	<0.001	9.10	0.52
pH	/0.15	0.83	(0.75 to 0.91)	15.80	<0.001	2.65	0.62
Haemoglobin	/3.3 g/dL	0.66	(0.55 to 0.78)	24.00	<0.001	12.69	0.03
Glucose	/3.7 mmol/L	1.69	(1.52 to 1.89)	93.01	<0.001	9.10	0.11
Haematocrit	/10 %	0.60	(0.50 to 0.74)	25.10	<0.001	2.18	0.34
Platelets	/100 $x10^9$/L	0.67	(0.58 to 0.78)	26.75	<0.001	2.90	0.41

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Prothrombin time	/2 seconds	1.42	(1.04 to 1.95)	4.91	0.027	6.58	0.04

4.2.4 Dichotomous analysis – not good outcome

Table 4-5 below shows the odds ratios and 95% confidence intervals for each covariate for the final binary split taking not a good outcome as 'unfavourable' i.e. comparing a not good recovery outcome with good recovery. A similar relationship between the covariates and outcome is observed as with the previous three dichotomies. This pattern is shown both with the odds ratios and the tests for heterogeneity as would be expected. For example, older age and having a tSAH are very strongly associated with a not good outcome.

Table 4-5 Pooled random effects estimates of binary odds ratios for not good outcome (unadjusted)

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	0.98	(0.88 to 1.09)	0.18	0.671	5.09	0.88
Race	Caucasian						
	Black	1.41	(0.91 to 2.17)	2.38	0.123	10.06	0.07
	Asian	1.01	(0.67 to 1.51)	0.00	0.964	2.07	0.72
	other	1.12	(0.87 to 1.44)	0.74	0.388	1.72	0.79
Education	0-8 years						
	9-12 years	0.88	(0.67 to 1.15)	0.91	0.341	1.64	0.44
	> 12 years	0.72	(0.53 to 0.99)	4.08	0.043	2.17	0.34

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Place Injury	street/highway						
	home	1.35	(0.98 to 1.85)	3.46	0.063	0.17	0.92
	work/school	1.40	(1.01 to 1.92)	4.15	0.042	1.83	0.40
	recreational	0.71	(0.47 to 1.07)	2.66	0.103	0.07	0.96
	other	0.92	(0.63 to 1.33)	0.21	0.651	0.00	1.00
Cause Injury	domestic/fall						
	road traffic accident	0.71	(0.63 to 0.81)	28.87	<0.001	8.32	0.60
	assault	0.71	(0.54 to 0.93)	6.33	0.012	11.83	0.30
	work-related	1.06	(0.71 to 1.60)	0.08	0.772	9.25	0.32
	sports/recreation	0.41	(0.25 to 0.66)	13.81	<0.001	15.02	0.06
	other	0.97	(0.76 to 1.23)	0.07	0.793	9.80	0.37
Referral	primary						
	secondary	1.06	(0.83 to 1.35)	0.22	0.640	17.29	<0.01
Hypoxia	no						
	suspected/definite	2.13	(1.70 to 2.67)	43.55	<0.001	12.31	0.09
Hypotension	no						
	suspected/definite	2.43	(2.00 to 2.95)	80.52	<0.001	11.07	0.20
Hypothermia	no						
	suspected/definite	2.09	(1.43 to 3.06)	14.67	<0.001	7.85	0.10
CT class	diffuse class II						
	no visible pathology	0.42	(0.27 to 0.65)	14.79	0.001	14.75	0.02
	swelling/shift	2.19	(1.84 to 2.60)	79.40	<0.001	6.46	0.37
	mass	1.97	(1.63 to 2.39)	49.27	<0.001	10.90	0.09
Cisterns	present						
	compressed/absent	2.17	(1.83 to 2.57)	80.78	<0.001	6.69	0.24
Shift	no						
	1-5 mm	1.14	(0.96 to 1.34)	2.30	0.130	4.44	0.62
	> 5mm	2.02	(1.44 to 2.84)	16.48	<0.001	22.43	<0.01
tSAH	no						
	yes	2.41	(2.12 to 2.74)	176.26	<0.001	12.14	0.21

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
EDH	no						
	yes	0.66	(0.57 to 0.75)	35.78	<0.001	2.81	0.95
SDH	no						
	yes	1.97	(1.62 to 2.39)	45.36	<0.001	21.60	0.01
Contusion	no						
	yes	1.44	(1.22 to 1.70)	19.18	<0.001	13.22	0.07
Eye	none						
	pain/sound/ spontaneous	0.37	(0.29 to 0.47)	67.96	<0.001	34.50	<0.01
	missing/untestable	0.87	(0.62 to 1.22)	0.64	0.424	12.30	0.09
Verbal	none						
	sounds-orientated	0.42	(0.33 to 0.52)	57.22	<0.001	26.89	<0.01
	missing/untestable	1.08	(0.84 to 1.38)	0.35	0.552	15.54	0.03
Motor	localises/obeys						
	none	4.05	(2.86 to 5.75)	61.69	<0.001	27.08	<0.01
	extension	7.68	(6.02 to 9.80)	270.08	<0.001	12.86	0.23
	abnormal flexion	3.77	(2.68 to 5.29)	58.51	<0.001	34.74	<0.01
	normal flexion	1.69	(1.35 to 2.13)	20.61	<0.001	30.75	<0.01
	missing/untestable	2.29	(1.67 to 3.15)	26.16	<0.001	8.02	0.24
GCS	6-8						
	3-5	3.16	(2.66 to 3.75)	169.95	<0.001	13.24	0.21
	9-15	0.55	(0.39 to 0.79)	10.55	0.001	37.65	<0.01
	missing/untestable	1.83	(1.47 to 2.28)	29.13	<0.001	11.81	0.11
Pupil	both sides +ve						
	one side +ve	3.13	(2.58 to 3.80)	132.43	<0.001	9.44	0.31
	both side -ve	6.96	(4.74 to 0.22)	97.93	<0.001	32.43	<0.01
SBP	120-150mmHg						
	<120 mmHg	1.44	(1.27 to 1.64)	31.15	<0.001	3.27	0.92
	>150 mmHg	1.36	(1.12 to 1.65)	9.45	0.002	14.03	0.08

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
MABP	85-110 mmHg						
	<85 mmHg	1.22	(1.08 to 1.38)	10.48	0.001	5.76	0.67
	>110 mmHg	1.42	(1.09 to 1.86)	6.63	0.010	16.62	0.03
Sodium	137-142 mmol/L						
	<137 mmol/L	1.32	(1.12 to 1.56)	10.58	0.001	6.94	0.33
	>142 mmol/L	1.09	(0.84 to 1.41)	0.42	0.516	14.18	0.03
Age	/24 years	2.15	(1.94 to 2.39)	215.14	<0.001	14.16	0.17
pH	/0.15	0.83	(0.74 to 0.92)	13.02	0.003	3.59	0.46
Haemoglobin	/3.3 g/dL	0.72	(0.63 to 0.81)	28.51	<0.001	6.79	0.24
Glucose	/3.7 mmol/L	1.57	(1.45 to 1.70)	121.67	<0.001	2.51	0.78
Haematocrit	/10 %	0.70	(0.55 to 0.88)	9.40	0.002	2.16	0.34
Platelets	/100 $x10^9$/L	0.74	(0.59 to 0.93)	6.97	0.008	4.90	0.18
Prothrombin time	/2 seconds	1.16	(0.63 to 2.14)	0.22	0.638	15.23	<0.01

4.2.5 Discussion - dichotomies

This is the first step before performing any proportional odds analysis as it allows judgement of whether the odds ratios under each of the four dichotomies are broadly similar and hence if the assumption of proportional odds is likely to be met. For example for gender the odds of an unfavourable outcome for men and women were very similar for all four dichotomies. The four odds ratios and 95% confidence intervals were: 1.01 (0.90 to 1.13); 1.03 (0.93 to 1.15); 1.04 (0.93 to 1.16) and 0.98 (0.88 to 1.09) for the mortality, dead/vegetative, unfavourable and not good dichotomies respectively. Other variables showed a much stronger effect, for example having a tSAH was associated with a much poorer outcome for all four dichotomies. Odds ratios and 95% confidence intervals are: 2.82 (2.52 to 3.16) for the mortality outcome; 3.00 (2.66 to 3.40) for the death/vegetative outcome; 2.70

(2.45 to 2.98) for the unfavourable outcome and 2.41 (2.12 to 2.74) for the not good outcome. Again, all of these show a similar pattern for all four dichotomies.

4.3 Proportional odds analysis on pooled data introduction

As the results from the four univariate analyses showed a similar pattern, the next stage was to fit a proportional odds model to these data. Again the odds of an unfavourable outcome were modelled. For this analysis the outcomes of death and vegetative state were pooled on both statistical and ethical grounds, as discussed previously. Thus, the GOS was reduced to a 4-point ordinal scale for all proportional odds modelling.

4.3.1 Unadjusted estimates – using proportional odds model

Univariate results from the proportional odds regression are shown in Table 4-6 below. As with the univariate analysis on the binary outcomes many variables showed a strong univariate effect. Particularly strong relationships were observed with age, CT class, tSAH, hypoxia and hypotension. No differences in the odds of an unfavourable outcome were observed between genders or referral to neurosurgical unit. Very little statistically significant heterogeneity was observed between studies. Only some of the CT parameters, cisterns, shift, contusion, some of the GCS components and prothrombin time showed statistically significant heterogeneity.

Table 4-6 Pooled random effects estimates of the common odds ratios from proportional odds models (unadjusted)

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	1.01	(0.92 to 1.11)	0.06	0.805	7.17	0.71

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Race	Caucasian						
	Black	1.30	(1.09 to 1.56)	8.15	0.004	4.11	0.53
	Asian	1.09	(0.78 to 1.51)	0.24	0.622	2.52	0.64
	other	1.08	(0.88 to 1.34)	0.56	0.453	2.43	0.66
Education	0-8 years						
	9-12 years	0.78	(0.57 to 1.07)	2.36	0.125	2.70	0.26
	> 12 years	0.70	(0.52 to 0.94)	5.46	0.020	2.32	0.31
Place Injury	street/highway						
	home	1.51	(1.16 to 1.98)	9.27	0.002	1.01	0.60
	work/school	1.37	(0.99 to 1.89)	3.59	0.058	2.65	0.27
	recreational	0.72	(0.49 to 1.05)	2.92	0.088	0.15	0.93
	other	1.13	(0.82 to 1.56)	0.58	0.447	0.08	0.96
Cause Injury	domestic/fall						
	road traffic accident	0.66	(0.60 to 0.73)	65.32	<0.001	6.80	0.74
	assault	0.66	(0.52 to 0.84)	11.84	0.001	13.94	0.18
	work-related	0.88	(0.68 to 1.14)	0.96	0.327	2.63	0.96
	sports/recreation	0.45	(0.28 to 0.71)	11.71	0.001	20.07	0.01
	other	0.91	(0.76 to 1.09)	1.06	0.304	5.94	0.75
Referral	primary						
	secondary	1.02	(0.87 to 1.20)	0.05	0.817	10.35	0.07
Hypoxia	no						
	suspected/definite	2.08	(1.69 to 2.56)	46.95	<0.001	17.45	0.01
Hypotension	no						
	suspected/definite	2.67	(2.09 to 3.41)	62.56	<0.001	28.00	<0.01
Hypothermia	no						
	suspected/definite	2.21	(1.56 to 3.15)	19.71	<0.001	11.71	0.02
CT class	diffuse class II						
	no visible pathology	0.45	(0.31 to 0.67)	16.14	<0.001	13.04	0.04
	swelling/shift	2.62	(2.13 to 3.21)	83.67	<0.001	11.89	0.06
	mass	2.18	(1.83 to 2.61)	73.92	<0.001	12.44	0.05

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Cisterns	present						
	compressed/absent	2.45	(1.88 to 3.20)	43.20	<0.001	22.03	<0.01
Shift	no						
	1-5 mm	1.36	(1.09 to 1.68)	7.68	0.006	11.70	0.07
	> 5mm	2.20	(1.64 to 2.96)	27.63	<0.001	26.30	<0.01
tSAH	no						
	yes	2.64	(2.42 to 2.89)	468.15	<0.001	8.30	0.50
EDH	no						
	yes	0.64	(0.56 to 0.72)	53.00	<0.001	4.43	0.82
SDH	no						
	yes	2.14	(1.87 to 2.45)	123.32	<0.001	15.92	0.04
Contusion	no						
	yes	1.34	(1.10 to 1.63)	8.34	0.004	25.68	<0.01
Eye	none						
	pain/sound/ spontaneous	0.36	(0.28 to 0.46)	63.60	<0.001	47.76	<0.01
	missing/untestable	0.74	(0.55 to 1.00)	3.85	0.050	15.07	0.04
Verbal	none						
	sounds-orientated	0.38	(0.30 to 0.48)	70.49	<0.001	33.65	<0.01
	missing/untestable	0.99	(0.78 to 1.26)	0.00	0.949	23.40	<0.01
Motor	localises/obeys						
	none	5.30	(3.49 to 8.04)	61.42	<0.001	65.45	<0.01
	extension	7.48	(5.60 to 9.98)	186.33	<0.001	38.36	<0.01
	abnormal flexion	3.58	(2.71 to 4.73)	80.44	<0.001	38.79	<0.01
	normal flexion	1.74	(1.44 to 2.11)	32.41	<0.001	27.73	<0.01
	missing/untestable	2.20	(1.66 to 2.92)	30.19	<0.001	9.11	0.17
GCS	6-8						
	3-5	3.67	(2.99 to 4.52)	151.85	<0.001	28.56	<0.01
	9-15	0.60	(0.41 to 0.88)	6.90	0.009	51.51	<0.01
	missing/untestable	1.97	(1.65 to 2.34)	58.37	<0.001	10.45	0.16

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Pupil	both sides +ve						
	one side +ve	2.71	(2.36 to 3.12)	194.30	<0.001	8.96	0.35
	both side -ve	7.31	(5.35 to 9.99)	156.11	<0.001	50.85	<0.01
SBP	120-150mmHg						
	<120 mmHg	1.53	(1.31 to 1.80)	27.72	<0.001	15.75	0.05
	>150 mmHg	1.42	(1.20 to 1.68)	17.23	<0.001	14.96	0.06
MABP	85-110 mmHg						
	<85 mmHg	1.30	(1.12 to 1.51)	11.94	0.001	15.04	0.06
	>110 mmHg	1.45	(1.19 to 1.76)	13.74	<0.001	13.74	0.09
Sodium	137-142 mmol/L						
	<137 mmol/L	1.40	(1.22 to 1.60)	23.11	<0.001	6.80	0.34
	>142 mmol/L	1.14	(0.94 to 1.38)	1.88	0.171	11.34	0.08
Age	/24 years	2.14	(2.00 to 2.28)	546.10	<0.001	10.01	0.44
pH	/0.15	0.80	(0.74 to 0.88)	25.17	<0.001	4.05	0.40
Haemoglobin	/3.3 g/dL	0.69	(0.60 to 0.78)	31.84	<0.001	9.72	0.08
Glucose	/3.7 mmol/L	1.68	(1.54 to 1.83)	131.58	<0.001	8.18	0.15
Haematocrit	/10 %	0.68	(0.55 to 0.83)	13.51	<0.001	2.96	0.23
Platelets	/100 $x10^9$/L	0.70	(0.62 to 0.80)	28.52	<0.001	2.88	0.41
Prothrombin time	/2 seconds	1.41	(0.99 to 1.99)	3.72	0.054	11.23	<0.01

4.3.2 Graphical illustration unadjusted estimates

As an illustration of the results in Table 4-6 above, Figure 4-1 and Figure 4-2 below show forest plots for six selected variables. Figure 4-1 shows referral, hypoxia, hypotension and hypothermia. For all four variables a consistent pattern of effects is seen over all the studies. Although there was statistically significant heterogeneity between studies for hypotension, looking at the figure it can be seen that estimates from all of the studies lie in the same direction. Little difference is observed in the

odds of an unfavourable outcome between those who had a primary or secondary referral. Those who had hypoxia, hypotension or hypothermia had a greater odds of a poorer outcome than those who did not in all studies.

Figure 4-2 below shows forest plots for the CT categories and cisterns. It can be seen that those with no visible pathology on the CT scan have a much more favourable outcome than those who had a diffuse injury. Whereas, those who had a classification of mass or swelling/shift have a greater odds of an unfavourable outcome. Statistically significant heterogeneity was observed with cisterns. However, looking at the figure shows that all of the estimates from the different studies lie in the same direction.

Figure 4-1 Forest plots of referral, hypoxia, hypotension and hypothermia

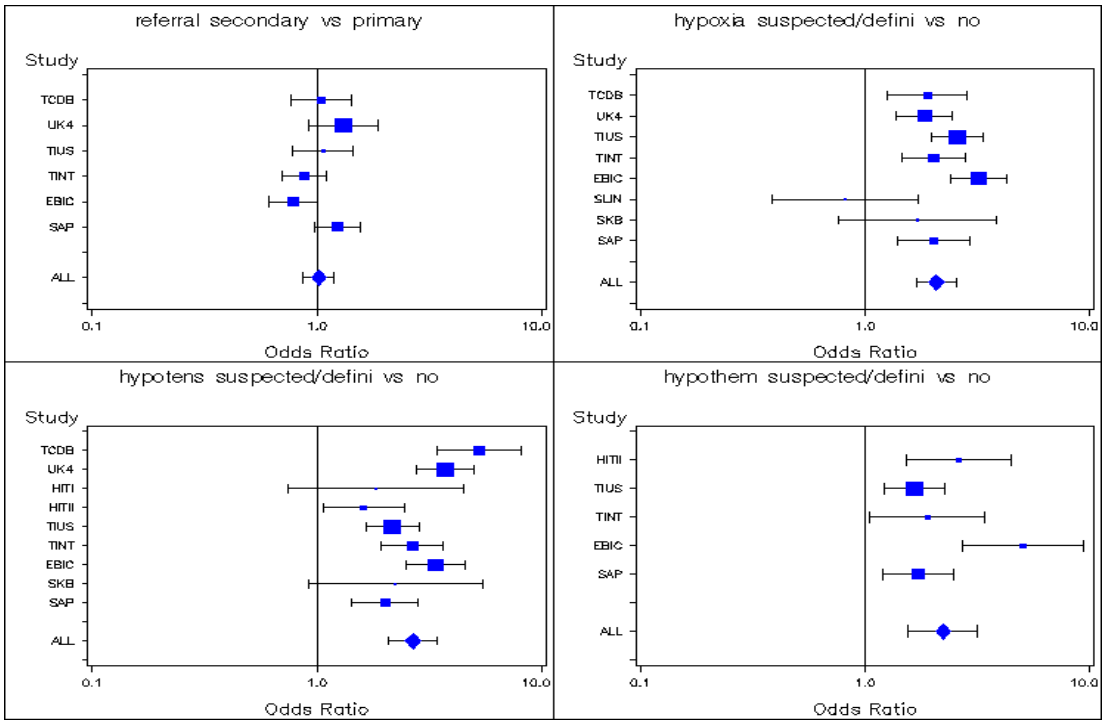
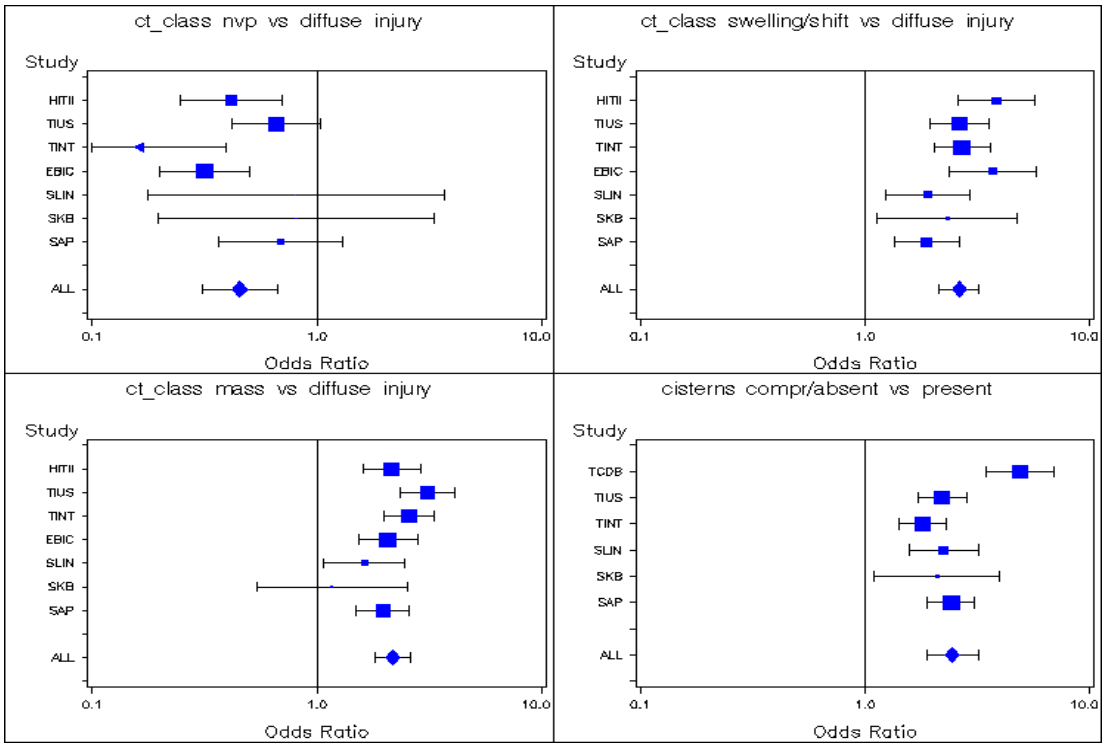


Figure 4-2 Forest plots of CT class and cisterns



4.4 Multivariate analysis using proportional odds modelling

4.4.1 Introduction

In order to perform multivariate analysis staff at the Rotterdam centre developed prognostic models based on the univariate proportional odds regression of the GOS (Steyerberg et al. 2008). Four models were used: a simple three variable model containing age, motor score and pupils (the conventional covariates used in many analyses (Marmarou et al. 2007)); a four variable model containing CT in addition to age, motor and pupils (Maas et al. 2007); a more complex seven variable model which included the addition of tSAH, hypoxia and hypotension to the four variable model (Chesnut et al. 1993;Hukkelhoven et al. 2005;Maas et al. 2007;McHugh et al. 2007b); and a nine variable model containing glucose and haemoglobin in addition to the parameters in the seven variable model (Van Beek et al. 2007). The regression coefficients for a selection of the models have been published (Steyerberg et al. 2011).

Once all initial variables were extracted from each dataset and placed into the database an imputed version of the database was created as detailed in Chapter 2. Values were imputed for any variables which had partially missing values within each study. If a variable was wholly missing within a study it was not imputed. Outcome as measured by the GOS was not imputed. For these multivariate analyses patients were included with complete data on age, baseline motor score and 6 month outcome. In total 82% of the required values were available for a prognostic model with seven predictors (Murray et al. 2007) and 92% were available for a prognostic model with nine predictors (Steyerberg et al. 2008). However, only four studies recorded lab values so the final nine covariate model is based only on these studies.

4.4.2 Results adjusted for three covariates

Table 4-7 below shows the results of the multivariate analysis when each comparison was adjusted for age, motor score and pupil reactivity. Strong associations with an unfavourable outcome were still observed with having a traumatic subarachnoid haemorrhage, higher glucose levels, having hypoxia, having hypotension, CT classes of swelling, shift or mass, lower pH, lower Haemoglobin, fewer platelets and longer prothrombin time. Having an EDH, an eye response of pain/sound/spontaneous or a verbal response of sounds-orientated was strongly associated with a favourable outcome. Statistically significant heterogeneity was observed between studies for referral and some categories of eye movement and GCS score.

Table 4-7 Pooled random effects estimates of the common odds ratios from proportional odds models adjusted for 3 covariates

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	0.94	(0.85 to 1.04)	1.41	0.235	7.44	0.68
Race	Caucasian						
	Black	1.44	(1.08 to 1.93)	6.11	0.013	7.77	0.17
	Asian	1.22	(0.84 to 1.78)	1.08	0.299	4.17	0.38
	other	1.11	(0.89 to 1.40)	0.87	0.350	1.55	0.82
Education	0-8 years						
	9-12 years	0.87	(0.67 to 1.13)	1.08	0.298	0.98	0.61
	> 12 years	0.74	(0.56 to 0.97)	4.57	0.033	1.57	0.46
Place Injury	street/highway						
	home	0.82	(0.59 to 1.14)	1.43	0.232	2.36	0.31
	work/school	1.10	(0.74 to 1.66)	0.23	0.632	3.46	0.18
	recreational	0.75	(0.50 to 1.12)	1.98	0.159	0.31	0.86
	other	0.87	(0.59 to 1.28)	0.51	0.476	2.35	0.31

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Cause Injury	domestic/fall						
	road traffic accident	1.08	(0.96 to 1.21)	1.54	0.214	6.18	0.80
	assault	1.03	(0.76 to 1.40)	0.04	0.834	17.40	0.07
	work-related	1.21	(0.90 to 1.63)	1.67	0.196	5.62	0.69
	sports/recreation	0.74	(0.50 to 1.09)	2.30	0.129	9.60	0.29
	other	1.06	(0.87 to 1.29)	0.33	0.567	8.20	0.51
Referral	primary						
	secondary	1.14	(0.91 to 1.43)	1.37	0.241	16.11	0.01
Hypoxia	no						
	suspected/definite	1.65	(1.37 to 2.00)	26.46	<0.001	12.39	0.09
Hypotension	no						
	suspected/definite	2.06	(1.64 to 2.59)	37.77	<0.001	20.99	0.01
Hypothermia	no						
	suspected/definite	1.63	(1.11 to 2.40)	6.26	0.012	12.32	0.02
CT class	diffuse class II						
	no visible pathology	0.47	(0.32 to 0.70)	13.56	<0.001	11.39	0.08
	swelling/shift	2.23	(1.83 to 2.72)	63.18	<0.001	9.76	0.13
	mass	1.48	(1.27 to 1.71)	26.06	<0.001	7.90	0.25
Cisterns	present						
	compressed/absent	1.83	(1.55 to 2.17)	48.70	<0.001	8.07	0.15
Shift	no						
	1-5 mm	1.31	(1.11 to 1.56)	9.88	0.002	6.97	0.32
	> 5mm	1.38	(1.02 to 1.87)	4.33	0.038	23.91	<0.01
tSAH	no						
	yes	2.01	(1.83 to 2.21)	213.20	<0.001	6.95	0.64
EDH	no						
	yes	0.63	(0.55 to 0.72)	47.04	<0.001	2.09	0.98
SDH	no						
	yes	1.33	(1.21 to 1.47)	32.67	<0.001	7.55	0.48
Contusion	no						
	yes	1.40	(1.18 to 1.67)	14.19	<0.001	17.95	0.01

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Eye	none						
	pain/sound/ spontaneous	0.65	(0.51 to 0.81)	13.74	<0.001	30.59	<0.01
	missing/untestable	0.86	(0.56 to 1.32)	0.48	0.489	9.60	0.21
Verbal	none						
	sounds-orientated	0.66	(0.55 to 0.80)	18.35	<0.001	19.02	0.04
	missing/untestable	1.00	(0.86 to 1.17)	0.00	0.999	4.80	0.68
GCS	6-8						
	3-5	1.37	(1.02 to 1.83)	4.52	0.0340	15.06	0.13
	9-15	0.72	(0.50 to 1.04)	3.00	0.083	39.49	<0.01
	missing/untestable	1.13	(0.95 to 1.34)	1.79	0.181	6.63	0.47
SBP	120-150mmHg						
	<120 mmHg	1.28	(1.12 to 1.45)	14.22	<0.001	9.48	0.30
	>150 mmHg	1.30	(1.08 to 1.57)	7.37	0.007	17.19	0.03
MABP	85-110 mmHg						
	<85 mmHg	1.14	(1.01 to 1.29)	4.80	0.029	9.26	0.32
	>110 mmHg	1.27	(1.02 to 1.58)	4.47	0.035	15.20	0.06
Sodium	137-142 mmol/L						
	<137 mmol/L	1.14	(0.91 to 1.43)	1.35	0.245	14.31	0.03
	>142 mmol/L	1.11	(0.96 to 1.27)	2.07	0.150	5.39	0.49
pH	/0.15	0.84	(0.76 to 0.92)	14.39	<0.001	1.87	0.76
Haemoglobin	/3.3 g/dL	0.76	(0.66 to 0.88)	14.68	<0.001	9.77	0.08
Glucose	/3.7 mmol/L	1.45	(1.36 to 1.55)	118.24	<0.001	4.13	0.53
Haematocrit	/10 %	0.83	(0.65 to 1.07)	2.05	0.152	3.38	0.18
Platelets	/100 $x10^9$/L	0.79	(0.69 to 0.91)	11.06	0.001	2.52	0.47
Prothrombin time	/2 seconds	1.63	(1.40 to 1.89)	40.40	<0.001	1.96	0.38

4.4.3 Results adjusted for four covariates

Table 4-8 shows the results of the multivariate analysis when each level of each covariate was adjusted for age, motor score, pupil reactivity and CT class. Even after adjustment for these four variables, a highly statistically significant relationship was observed between having a subarachnoid haemorrhage, higher glucose levels and longer prothrombin time with an unfavourable outcome. As with adjustment for three covariates, statistically significant heterogeneity was only observed between studies for referral, shift >5mm and some eye and GCS categories.

Table 4-8 Pooled random effects estimates of the common odds ratios from proportional odds models adjusted for 4 covariates

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	0.94	(0.85 to 1.03)	1.68	0.194	8.32	0.60
Race	Caucasian						
	Black	1.45	(1.06 to 1.98)	5.31	0.021	8.68	0.12
	Asian	1.19	(0.83 to 1.73)	0.89	0.346	3.73	0.44
	other	1.09	(0.87 to 1.37)	0.60	0.437	2.55	0.64
Education	0-8 years						
	9-12 years	0.86	(0.66 to 1.12)	1.26	0.263	1.54	0.46
	> 12 years	0.72	(0.54 to 0.95)	5.36	0.021	1.93	0.38
Place Injury	street/highway						
	home	0.73	(0.53 to 1.01)	3.71	0.054	2.17	0.34
	work/school	0.97	(0.60 to 1.57)	0.02	0.890	4.71	0.09
	recreational	0.71	(0.47 to 1.06)	2.85	0.091	0.17	0.92
	other	0.80	(0.53 to 1.22)	1.04	0.308	2.73	0.25

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Cause Injury	domestic/fall						
	road traffic accident	1.15	(1.03 to 1.30)	5.61	0.018	6.08	0.81
	assault	1.08	(0.78 to 1.50)	0.24	0.626	20.09	0.03
	work-related	1.25	(0.93 to 1.68)	2.19	0.139	5.71	0.68
	sports/recreation	0.76	(0.51 to 1.15)	1.70	0.193	10.23	0.25
	other	1.11	(0.90 to 1.37)	0.94	0.331	9.61	0.38
Referral	primary						
	secondary	1.12	(0.87 to 1.44)	0.77	0.380	20.52	<0.01
Hypoxia	no						
	suspected/definite	1.65	(1.39 to 1.95)	33.18	<0.001	9.82	0.20
Hypotension	no						
	suspected/definite	2.06	(1.64 to 2.57)	39.81	<0.001	19.45	0.01
Hypothermia	no						
	suspected/definite	1.62	(1.14 to 2.31)	7.25	0.007	10.22	0.04
Cisterns	present						
	compressed/absent	1.68	(1.28 to 2.20)	14.09	<0.001	13.64	0.02
Shift	no						
	1-5 mm	1.09	(0.93 to 1.28)	1.04	0.307	3.98	0.68
	> 5mm	1.14	(0.77 to 1.71)	0.43	0.510	29.12	<0.01
tSAH	no						
	yes	1.90	(1.72 to 2.09)	174.53	<0.001	8.77	0.46
EDH	no						
	yes	0.50	(0.44 to 0.58)	91.64	<0.001	4.10	0.85
SDH	no						
	yes	1.17	(1.05 to 1.30)	7.88	0.005	7.66	0.47
Contusion	no						
	yes	1.34	(1.12 to 1.61)	10.52	0.001	18.03	0.01
Eye	none						
	pain/sound/ spontaneous	0.63	(0.51 to 0.79)	15.78	<0.001	28.76	<0.01
	missing/untestable	0.88	(0.55 to 1.42)	0.27	0.601	11.27	0.13

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Verbal	none						
	sounds-orientated	0.66	(0.54 to 0.80)	17.36	<0.001	20.83	0.02
	missing/untestable	1.01	(0.86 to 1.17)	0.01	0.933	4.13	0.76
GCS	6-8						
	3-5	1.33	(0.97 to 1.82)	3.12	0.077	17.45	0.06
	9-15	0.69	(0.48 to 1.00)	3.80	0.051	37.72	<0.01
	missing/untestable	1.11	(0.93 to 1.33)	1.33	0.249	7.22	0.41
SBP	120-150mmHg						
	<120 mmHg	1.27	(1.13 to 1.44)	14.56	<0.001	9.06	0.34
	>150 mmHg	1.28	(1.06 to 1.54)	6.43	0.011	16.82	0.03
MABP	85-110 mmHg						
	<85 mmHg	1.14	(1.02 to 1.27)	5.28	0.022	7.03	0.53
	>110 mmHg	1.26	(1.01 to 1.59)	4.08	0.043	16.02	0.04
Sodium	137-142 mmol/L						
	<137 mmol/L	1.09	(0.86 to 1.39)	0.50	0.481	16.20	0.01
	>142 mmol/L	1.10	(0.96 to 1.26)	1.85	0.174	5.03	0.54
pH	/0.15	0.83	(0.76 to 0.91)	15.74	<0.001	2.44	0.66
Haemoglobin	/3.3 g/dL	0.76	(0.65 to 0.88)	13.56	<0.001	10.81	0.06
Glucose	/3.7 mmol/L	1.42	(1.33 to 1.52)	103.04	<0.001	3.62	0.60
Haematocrit	/10 %	0.82	(0.63 to 1.05)	2.50	0.114	3.44	0.18
Platelets	/100 $x10^9$/L	0.80	(0.69 to 0.92)	9.13	0.003	3.16	0.37
Prothrombin time	/2 seconds	1.60	(1.38 to 1.86)	37.47	<0.001	1.61	0.45

4.4.4 Results adjusted for seven covariates

Table 4-9 below shows the results of the multivariate analysis when estimates were adjusted for age, motor score, pupil reactivity, CT class, traumatic subarachnoid haemorrhage, hypoxia and hypotension. Having higher glucose levels, lower haemoglobin and longer prothrombin time were strongly associated with an

unfavourable outcome. Having an epidural haematoma, an eye response of pain/sound/spontaneous or a verbal response of sounds-orientated was strongly associated with a favourable outcome. Statistically significant heterogeneity was observed between studies for referral and one category each within shift, GCS and sodium.

Table 4-9 Pooled random effects estimates of the common odds ratios from proportional odds models adjusted for 7 covariates

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	0.94	(0.85 to 1.04)	1.51	0.220	6.78	0.75
Race	Caucasian						
	Black	1.48	(1.03 to 2.13)	4.58	0.032	10.74	0.06
	Asian	1.14	(0.73 to 1.78)	0.34	0.563	5.26	0.26
	other	1.07	(0.85 to 1.35)	0.35	0.555	2.00	0.74
Education	0-8 years						
	9-12 years	0.87	(0.67 to 1.14)	1.03	0.310	1.89	0.39
	> 12 years	0.70	(0.46 to 1.05)	2.94	0.086	3.13	0.21
Place Injury	street/highway						
	home	0.76	(0.56 to 1.03)	3.06	0.080	1.23	0.54
	work/school	0.94	(0.57 to 1.56)	0.06	0.813	4.94	0.08
	recreational	0.72	(0.48 to 1.09)	2.45	0.118	0.12	0.94
	other	0.81	(0.51 to 1.29)	0.81	0.369	3.31	0.19
Cause Injury	domestic/fall						
	road traffic accident	1.16	(1.02 to 1.30)	5.56	0.018	6.83	0.74
	assault	1.17	(0.84 to 1.62)	0.88	0.349	19.48	0.03
	work-related	1.27	(0.94 to 1.72)	2.45	0.117	5.44	0.71
	sports/recreation	0.80	(0.52 to 1.24)	0.98	0.322	11.29	0.19
	other	1.15	(0.94 to 1.41)	1.87	0.172	8.38	0.50
Referral	primary						
	secondary	1.15	(0.87 to 1.50)	0.98	0.323	22.45	<0.01

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Hypothermia	no						
	suspected/definite	1.40	(1.02 to 1.91)	4.38	0.036	7.67	0.10
Cisterns	present						
	compressed/absent	1.64	(1.27 to 2.12)	14.39	<0.001	11.98	0.04
Shift	no						
	1-5 mm	1.10	(0.93 to 1.29)	1.19	0.276	3.80	0.70
	> 5mm	1.18	(0.80 to 1.75)	0.71	0.401	27.59	<0.01
EDH	no						
	yes	0.53	(0.46 to 0.61)	76.70	<0.001	3.19	0.92
SDH	no						
	yes	1.17	(1.05 to 1.31)	7.99	0.005	6.70	0.57
Contusion	no						
	yes	1.26	(1.07 to 1.48)	7.73	0.005	14.30	0.05
Eye	none						
	pain/sound/ spontaneous	0.65	(0.53 to 0.80)	16.17	<0.001	24.35	0.01
	missing/untestable	0.89	(0.53 to 1.51)	0.19	0.663	12.81	0.08
Verbal	none						
	sounds-orientated	0.67	(0.55 to 0.81)	17.37	<0.001	18.95	0.04
	missing/untestable	0.94	(0.80 to 1.10)	0.58	0.446	5.67	0.58
GCS	6-8						
	3-5	1.28	(0.93 to 1.77)	2.33	0.127	17.79	0.06
	9-15	0.71	(0.49 to 1.01)	3.55	0.060	35.44	<0.01
	missing/untestable	1.03	(0.86 to 1.23)	0.11	0.746	6.98	0.43
SBP	120-150mmHg						
	<120 mmHg	1.18	(1.03 to 1.35)	5.83	0.016	9.70	0.29
	>150 mmHg	1.33	(1.10 to 1.61)	8.49	0.004	16.96	0.03
MABP	85-110 mmHg						
	<85 mmHg	1.06	(0.95 to 1.19)	1.01	0.315	6.59	0.58
	>110 mmHg	1.29	(1.03 to 1.61)	4.89	0.027	15.24	0.05

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Sodium	137-142 mmol/L						
	<137 mmol/L	1.07	(0.81 to 1.41)	0.24	0.622	20.25	<0.01
	>142 mmol/L	1.05	(0.91 to 1.20)	0.41	0.522	2.90	0.82
pH	/0.15	0.89	(0.81 to 0.98)	5.89	0.015	2.57	0.63
Haemoglobin	/3.3 g/dL	0.76	(0.67 to 0.85)	20.78	<0.001	6.82	0.23
Glucose	/3.7 mmol/L	1.35	(1.26 to 1.45)	74.10	<0.001	2.89	0.72
Haematocrit	/10 %	0.89	(0.65 to 1.22)	0.50	0.479	4.77	0.09
Platelets	/100 $x10^9$/L	0.81	(0.70 to 0.93)	8.30	0.004	3.09	0.38
Prothrombin time	/2 seconds	1.55	(1.33 to 1.81)	31.63	<0.001	1.07	0.58

4.4.5 Results adjusted for nine covariates

Table 4-10 below shows the results of the multivariate analysis when estimates were adjusted for age, motor score, pupil reactivity, CT class, traumatic subarachnoid haemorrhage, hypoxia, hypotension, haemoglobin and glucose. Here, as mentioned above, these estimates are only based on four studies. Even after adjusting for these nine covariates, highly statistically significant associations with a favourable outcome were observed for those with an epidural haematoma, eye score being pain, sound or spontaneous and verbal score being between sounds to orientated inclusive. A highly statistically significant association between increased prothrombin time and an unfavourable outcome was observed. Statistically significant heterogeneity was observed between the studies for the same variables as had statistically significant heterogeneity when estimates were adjusted for seven covariates.

Table 4-10 Pooled random effects estimates of the common odds ratios from proportional odds models adjusted for 9 covariates

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Gender	male						
	female	0.88	(0.80 to 0.98)	5.50	0.019	9.49	0.49
Race	Caucasian						
	Black	1.50	(1.02 to 2.20)	4.32	0.038	11.82	0.04
	Asian	1.13	(0.72 to 1.76)	0.28	0.594	5.23	0.26
	other	1.07	(0.85 to 1.34)	0.31	0.577	2.61	0.63
Education	0-8 years						
	9-12 years	0.84	(0.65 to 1.10)	1.60	0.206	1.63	0.44
	> 12 years	0.69	(0.47 to 0.99)	3.95	0.047	2.69	0.26
Place Injury	street/highway						
	home	0.78	(0.58 to 1.07)	2.35	0.125	1.16	0.56
	work/school	0.91	(0.53 to 1.57)	0.11	0.739	5.53	0.06
	recreational	0.74	(0.49 to 1.11)	2.14	0.143	0.23	0.89
	other	0.84	(0.52 to 1.37)	0.49	0.485	3.56	0.17
Cause Injury	domestic/fall						
	road traffic accident	1.14	(1.01 to 1.28)	4.18	0.041	7.79	0.65
	assault	1.24	(0.90 to 1.70)	1.71	0.191	18.29	0.05
	work-related	1.26	(0.93 to 1.71)	2.23	0.135	5.10	0.75
	sports/recreation	0.81	(0.53 to 1.25)	0.87	0.350	11.06	0.20
	other	1.15	(0.92 to 1.43)	1.54	0.214	9.67	0.38
Referral	primary						
	secondary	1.14	(0.86 to 1.51)	0.84	0.359	24.34	<0.01
Hypothermia	no						
	suspected/definite	1.36	(0.96 to 1.93)	3.07	0.080	9.16	0.06
Cisterns	present						
	compressed/absent	1.63	(1.25 to 2.14)	12.60	<0.001	13.09	0.02
Shift	no						
	1-5 mm	1.08	(0.92 to 1.28)	0.88	0.348	4.36	0.63
	> 5mm	1.21	(0.82 to 1.77)	0.90	0.344	26.17	<0.01

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
EDH	no						
	yes	0.51	(0.44 to 0.59)	85.03	<0.001	4.12	0.85
SDH	no						
	yes	1.19	(1.07 to 1.33)	9.67	0.002	6.99	0.54
Contusion	no						
	yes	1.25	(1.06 to 1.48)	6.76	0.009	15.52	0.03
Eye	none						
	pain/sound/ spontaneous	0.64	(0.52 to 0.80)	15.90	<0.001	25.64	<0.01
	missing/untestable	0.83	(0.52 to 1.33)	0.58	0.444	10.59	0.16
Verbal	none						
	sounds-orientated	0.66	(0.55 to 0.80)	18.38	<0.001	18.22	0.05
	missing/untestable	0.94	(0.81 to 1.10)	0.53	0.466	5.16	0.64
GCS	6-8						
	3-5	1.33	(0.97 to 1.81)	3.14	0.077	16.59	0.08
	9-15	0.70	(0.49 to 1.00)	3.76	0.052	35.26	<0.01
	missing/untestable	1.03	(0.86 to 1.23)	0.11	0.744	6.26	0.51
SBP	120-150mmHg						
	<120 mmHg	1.09	(0.95 to 1.26)	1.60	0.206	10.20	0.25
	>150 mmHg	1.33	(1.09 to 1.62)	8.23	0.004	17.36	0.03
MABP	85-110 mmHg						
	<85 mmHg	1.00	(0.90 to 1.13)	0.01	0.932	6.39	0.60
	>110 mmHg	1.30	(1.03 to 1.65)	4.83	0.028	16.58	0.03
Sodium	137-142 mmol/L						
	<137 mmol/L	1.03	(0.79 to 1.35)	0.06	0.812	18.77	<0.01
	>142 mmol/L	1.12	(0.97 to 1.29)	2.30	0.129	2.55	0.86
pH	/0.15	0.93	(0.85 to 1.03)	1.97	0.161	2.17	0.70
Haematocrit	/10 %	1.06	(0.70 to 1.59)	0.07	0.795	2.11	0.35
Platelets	/100 $x10^9$/L	0.80	(0.69 to 0.93)	8.18	0.004	3.03	0.39

Variable	Category	Odds Ratio	95%CI	Global Test Statistic	Global p-value	Q Test Statistic	Q p-value
Prothrombin time	/2 seconds	1.46	(1.25 to 1.71)	22.17	<0.001	0.64	0.73

4.5 Discussion

The first part of this chapter showed that using logistic regression to model the four binary outcomes similar odds ratios and overlapping confidence intervals were observed for each of the outcomes for most variables. Using proportional odds modelling therefore seemed an appropriate summary. The second part of the chapter firstly showed unadjusted results. Subsequently four covariate sets (each containing respectively three, four, seven and nine covariates) were adjusted for in multivariate proportional odds analysis. Forest plots were also shown to illustrate the consistency of estimates across studies. Little differences were observed between the three and four covariate models. Therefore in subsequent chapters three covariate sets are used (three, seven and nine). These analyses were all performed using the IMPACT data and using each level of each covariate as though it was a treatment effect. The next stage, shown in Chapter 5 and Chapter 6 is to use the IMPACT data to simulate treatment effects and compare different modelling strategies for ordinal outcomes.

5 Chapter 5 Comparison of different analysis strategies – the sliding dichotomy

5.1 Introduction

Having shown that the proportional odds model was a reasonable summary of the four binary dichotomies using the IMPACT data in Chapter 4, the next stage was to compare different modelling strategies against the conventional dichotomous analysis. This will show which gives the greatest reduction in sample size (RSS) under which circumstances. Various scenarios and treatment effects were modelled using simulated data from the eleven IMPACT studies. The methods for simulating the different treatment effects giving details of all of the scenarios are given in the first part of this chapter. The sample size reduction for the different methods and scenarios studied are compared with recommendations made as to which methods of analysis are optimal under different treatment strategies. In this chapter all simulation analysis strategies, the conventional analysis, the sliding dichotomy and the methodology for all of the simulation studies are described.

The simulation studies were done in two parts: firstly the sliding dichotomy was refined, as shown in the second part of this chapter; then the optimised version of the sliding dichotomy was taken forward to compare with the other modelling strategies. These results are shown in Chapter 6.

Results are presented both by study and as overall median reductions in sample size in tables and graphs. The figures show the unadjusted and adjusted analyses together.

The sections describing the conventional analysis, sliding dichotomy and methodology below have almost all been published and the sections showing the results tables have partly been published (McHugh et al. 2010). This paper discussed the three and seven covariate models with the 5% treatment effect.

5.2 Conventional analysis

In evaluating novel analytical approaches, the conventional unadjusted Chi-squared analysis based on a dichotomisation of the GOS into unfavourable versus favourable was used as a reference technique. All other approaches are compared back to this reference technique in terms of their relative efficiency, expressed as the percentage reduction in sample size, which preserves the statistical power of the novel analysis. This could also be thought of as obtaining narrower confidence limits and hence more precise estimates of the treatment effect for a given sample size. It has previously been shown, using the IMPACT database, that the use of logistic regression to incorporate baseline covariates into the conventional analysis of the dichotomous GOS can yield sample size reductions of the order of 25% (Hernandez et al. 2006).

To extend these earlier findings, and to allow direct comparisons with the results of ordinal analysis, three different sets of covariates were used in adjusted dichotomous analyses. These sets are as follows:

3 covariate - age, motor and pupils

7 covariate - age, motor, pupils, hypoxia, hypotension, tSAH and CT class

9 covariate - age, motor, pupils, hypoxia, hypotension, tSAH, CT class, haemoglobin and glucose

The covariate values were 94.6% complete for the three-covariate model, 81.6% complete for the seven-covariate model and 91.7% complete for the nine-covariate model. As described in detail previously in Chapter 2, the missing covariate values were replaced using a single realisation of a multiple imputation procedure.

5.3 The sliding dichotomy

A fundamental objection to the conventional dichotomous analysis of the GOS is that it gives absolute priority to one transition in the scale, namely, the change from severe disability to moderate disability, and ignores all other transitions, such as the change from vegetative state to severe disability or from moderate disability to good recovery. This goes against clinical practice and also lacks statistical sensitivity. For a patient other than one with an intermediate prognosis, it is unlikely that a therapeutic intervention will alter prognosis sufficiently for the patient's outcome to move from 'unfavourable' to 'favourable'. Therefore, when using the conventional dichotomous analysis, a substantial proportion of patients recruited into a clinical trial do not have the potential to demonstrate the effect of even a highly beneficial intervention (Machado, Murray, & Teasdale 1999). This dilutes the observed effect of any beneficial (or indeed harmful) intervention and so reduces statistical power.

The 'sliding dichotomy' has been proposed to overcome this problem (Murray et al. 2005). The GOS is still dichotomised, but the point of dichotomy is tailored to each individual patient's baseline prognosis. For example, for a patient with an excellent prognosis, their outcome might only be regarded as being 'favourable' if they achieve good recovery, whereas a patient with a very poor prognosis might be regarded as having a 'favourable' outcome if they achieve severe disability or better. In effect, one is defining outcome to be favourable if it is better than would be expected, given the individual's baseline prognosis.

There are a number of operational issues to be decided when implementing the sliding dichotomy (Murray et al. 2005). One needs to use a baseline prognostic model, but does a sophisticated model offer substantial advantages over a simple model? Given the predicted prognostic risk, the patients need to be ordered by risk and then banded into prognostic groups. Should this be done to give roughly equal number of patients in each band, or so that each band is defined by a range of

prognostic risk? How many bands should be used? After the baseline covariates have been incorporated into the prognostic model, can they be used again to undertake an adjusted analysis? Each of these issues is explored in this chapter.

5.4 Simulating treatment effects - methods

5.4.1 Design of the Simulation Study

The simulation study began by fitting an overall multinomial generalised logit regression model for a nominal response to the entire IMPACT database to predict the probability of each possible outcome as a function of a patient's baseline covariates. This was done separately for the three, seven and nine covariate sets described previously, using the SAS procedure PROC LOGISTIC. As with the proportional odds analysis shown in Chapter 4, death and vegetative state were pooled on both statistical and ethical grounds. Thus, the GOS was reduced to a 4-point ordinal scale for all the analyses. The SAS programs for the initial simulations were written jointly by Dr Butcher (Edinburgh) and me. Subsequent programs to model the simulation data were written by me.

The 11 constituent studies within the IMPACT database were used as examples of typically selected head injury populations. For a single simulation with any one of the 11 IMPACT data sets, 400 subjects were sampled at random with replacement. For each individual, their predicted outcome was modelled in terms of their baseline covariates. For example, the modelled probabilities of the outcome for a given individual being death/vegetative state (D/V), severe disability (SD), moderate disability (MD), or good recovery (GR) might be 40%, 20%, 30%, and 10%, respectively. Using these estimated probabilities, an actual outcome was simulated, yielding a random sample of 400 patients reflecting the baseline severity of the study in question, and with simulated outcomes. These represented the placebo group in a trial. Still as part of this single simulation, a further 400 subjects were sampled at

random with replacement, to generate the intervention group. When simulating the outcomes for the intervention group, the predicted probabilities coming from the multinomial model were adjusted to incorporate a treatment effect.

Two treatment effects were fitted, 5% and 8%, both increasing the proportion of favourable outcomes. Eight percent was chosen as this is typically the treatment difference looked for in head injury trials. Five percent was also chosen to reflect a more modest, and perhaps more realistic, treatment benefit.

Two different models were used for the treatment effect, 'uniform' and 'mortality'. For the 'uniform' treatment effect, it was assumed that the impact of the treatment followed precisely a proportional odds model. The common odds ratio that was used was calibrated so that overall, there was a 5% or 8% absolute increase in the proportion of patients with a favourable outcome. For the 'mortality' model, it was assumed that the effect of treatment was to reduce the risk of D/V, but that the relative probability of SD to MD to GR was unaltered. Again, the reduction in mortality was calibrated so that overall, there was a 5% or 8% absolute increase in the proportion of patients with a favourable outcome. The technical details of this procedure are given below.

5.4.1.1 Algorithms used for the simulated treatment effects

Proportional odds model

Step 1: For a simulated patient from the treatment group with given covariate values, the global multinomial model will give initial 'placebo' probabilities p_1, p_2, p_3, p_4 for the four possible outcomes D/V, SD, MD, and GR, respectively (where $p_1+p_2+p_3+p_4=1$).

Step 2: The odds ratios are calculated for each possible dichotomisation of the scale as follows:

$OR_4= p_4/(1-p_4)$

$OR_{34} = (p_3+p_4)/(1-p_3-p_4)$

$OR_{234} = (p_2+p_3+p_4)/(1-p_2-p_3-p_4)$

Step 3: The odds ratios are each multiplied by the treatment effect parameter k.

Step 4: The scaled odds ratios are back transformed to give the new 'treatment' probabilities t_1, t_2, t_3, and t_4 as follows:

$t_4=1/(1+1/(kxOR_4))$

$t_3=1/(1+1/(kxOR_{34}))-t_4$

$t_2=1/(1+1/(kxOR_{234}))-t_3- t_4$

$t_1=1-t_2-t_3-t_4$

The calibration is achieved by taking the overall distribution over D/V, SD, MD, and GR for the study in question and applying the algorithm above. The treatment parameter is obtained by solving $t_3+t_4=p_3+p_4+\delta$ for k, where δ was set to 0.05 or 0.08. This was performed using the Microsoft Excel add on Tool, Solver.

Mortality model

The algorithm is a simplified version of the above. The treated probability of D/V is reduced by scaling and then back transforming the corresponding odds ratio. The probabilities for SD, MD, and GR are scaled linearly so that the four treatment probabilities sum to one. The calibration is achieved as described above.

5.4.1.2 Illustration for the simulated treatment effects

As an illustration, the overall distribution of outcome in the selected study might be 20% D/V, 30% SD, 25% MD, and 25% GR. With a proportional odds model, a

common odds ratio of (55/45)/(50/50), that is, 1.22, leads to a 5% absolute increase in the proportion of patients with an outcome of MD or GR. Under this uniform model, the full outcome distribution in the intervention group would be 17% D/V, 28% SD, 26% MD, and 29% GR. Under the mortality model, the proportion of patients with outcomes of SD, MD, and GR remain in the ratio 30:25:25. To achieve this, with an absolute increase of 5% in the proportion of patients with an outcome of MD or GR requires an overall outcome distribution of 12% D/V, 33% SD, 27.5% MD, and 27.5% GR. With this distribution, the odds ratio for D/V versus better is 1.83, the odds ratio for D/V or SD versus better is 1.22 (as for the proportional odds model), and the odds ratio for D/V or SD or MD versus better is 1.14. Thus, as intended, the mortality model deviates substantially from the proportional odds model. Fitting a proportional odds model to these data under the mortality model gives an estimated common odds ratio of 1.30.

5.4.1.3 Modelling the simulated treatment effects

This process therefore generated a sample of 800 subjects, representing 400 patients from the placebo group and 400 from the intervention group. The total size of 800 was chosen to be representative of a typical Phase III trial in head injury. Once the data were generated, they were analysed by the different approaches, which were to be compared, yielding either a 'significant' or a 'nonsignificant' outcome. For the analyses based on the sliding dichotomy approach, the parameters for the underlying prognostic model were estimated separately for each simulation by fitting a binary logistic regression model with a conventional favourable/unfavourable dichotomy of the GOS as the response variable. The entire process was then repeated 1000 times, and the power of each statistical approach was estimated as the proportion of the 1000 analyses, which yielded a significant result at the 5% level (two-sided). With this sample size, a power of the order of 80% is estimated with a standard error of approximately 1.25%.

For ease of interpretation, the power was then converted into a measure of the efficiency of each analytical approach relative to the reference approach of the conventional dichotomy without covariate adjustment. Thus, each method is reported in terms of the reduction in sample size, which can be achieved while preserving the power of the analysis. "The formula used was: 100-100*[(mean of Z score for reference model)/(mean of Z score for adjusted model)]^2 (Hernandez, Steyerberg, & Habbema 2004) where Z score is equal to the Wald statistic of the treatment effect coefficient" (Hernandez et al. 2006).

In total, the simulations covered 11x3x2x2 scenarios – every combination of the 11 constituent studies with the three sets of covariates (three, seven or nine) with the two treatment effect models (uniform or mortality) and the two treatment effects (5% and 8%).

In total, 16 different analytical approaches were compared:

- The conventional dichotomous analysis, as the reference technique.
- The sliding dichotomy with each combination of three, four, or five prognostic bands; the prognostic bands determined either by having equal numbers of patients per band or by having equal bands for p(fav), the predicted probability of a favourable outcome (e.g., for four equal bands, the bands would be defined by p(fav) lying in the ranges 0–25%, 25–50%, 50–75%, and 75–100%); and with and without covariate adjustment (3x2x2=12 combinations).
- The conventional dichotomous analysis, with covariates.
- The proportional odds model, without covariates.
- The proportional odds model, with covariates.

The different approaches to the sliding dichotomy analysis were compared first, as shown in this chapter, and a single optimised version was taken forward to compare with the other methods, as shown in Chapter 6.

For a limited number of scenarios representing the conventional binary logistic regression analysis, the proportional odds analysis and the sliding dichotomy analysis; the entire process of running 1000 simulations for each of the 11 data sets was repeated 10 times, to allow the standard error of the differences between methods in percentage sample size reduction to be estimated. Similarly, for a number of representative scenarios, a very large number of simulations were run with a null treatment effect to assess whether the actual significance level was close to the nominal value of 5%. These gave the following results. Significance level: the 95% confidence intervals for the significance level achieved with the logistic regression, the sliding dichotomy, and the proportional odds model were, 5.10–5.22%, 4.95–5.06%, and 5.08–5.21% respectively. Precision: the estimated standard error for the difference between methods in terms of the percentage reduction in sample size ranged from 1.5% to 5% with a mean of 3.5%.

5.5 Sliding dichotomy modelling

It was firstly of interest to compare the different sliding dichotomy strategies. The tables below show the sample size reduction achieved when the sliding dichotomy strategy is applied. Each table shows all of the sliding dichotomy scenarios separately for the combinations of the three covariate sets (three, seven and nine covariates), two treatment effects (uniform and mortality) and the two treatment effects (5% and 8%). For all of the tables the table cells are the percentage reductions in sample size, which can be achieved while preserving the power relative to the conventional unadjusted analysis of the dichotomised outcome scale

5.5.1 Sliding dichotomy modelling – uniform treatment effect

Here the treatment effect followed a proportional odds model.

5.5.1.1 Three covariate model

5% treatment effect

Table 5-1 below shows the results from the 5% treatment effect. Greater reductions in sample size are observed, in general, using the bands with equal splits rather than the bands based on p(fav). Using five bands, equal splits with and without covariates performs poorly for all studies. Using five bands equal splits with TINT gives an increase in sample size compared with the conventional dichotomy. The addition of covariates gives a modest increase in efficiency for all of the scenarios.

Table 5-1 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, three covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	16.2	30.2	26.1	25.6	36.5	28.6	28.9	45.3	29.9	38.5	33.3	29.9
4 bands, equal splits	22.6	35.5	29.3	33.7	34.6	27.5	33.8	44.5	30.7	37.0	38.2	33.8
5 bands, equal splits	-4.1	18.2	8.7	17.3	21.4	15.6	26.8	41.0	18.3	29.9	24.6	18.3
3 bands, equal splits +cov	19.8	32.5	29.5	27.9	38.5	31.6	36.6	48.7	30.3	43.9	32.3	32.3

[5] Table cells are the percentage reductions in sample size, which can be achieved while preserving the power relative to the conventional unadjusted analysis of the dichotomised outcome scale. 'Equal splits' means prognostic bands chosen to contain equal numbers of patients. 'p(fav)' means prognostic bands chosen with specific ranges of the predicted probability of a favourable outcome. 'no cov' means that the final analysis was not covariate adjusted. '+cov' means that the final analysis was covariate adjusted.

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
4 bands, equal splits +cov	24.4	35.7	30.3	33.4	36.3	32.3	36.7	46.3	35.0	40.4	39.1	35.7
5 bands, equal splits +cov	-1.5	17.7	12.2	17.6	24.8	16.1	26.6	41.2	20.6	31.3	25.6	20.6
3 bands, p(fav) splits	18.5	29.0	28.9	29.2	33.2	26.4	33.2	42.1	20.9	34.4	32.9	29.2
4 bands, p(fav) splits	14.1	29.2	22.3	27.0	29.7	21.6	31.1	43.8	24.2	31.7	28.1	28.1
5 bands, p(fav) splits	18.0	32.1	21.1	27.0	34.2	23.0	34.2	43.2	28.5	35.0	34.4	32.1
3 bands, p(fav) splits +cov	18.0	29.4	31.8	30.7	35.4	30.2	38.2	47.2	25.3	40.5	35.0	31.8
4 bands, p(fav) splits +cov	15.4	30.0	22.3	28.1	30.3	20.3	37.3	45.0	24.9	35.8	29.2	29.2
5 bands, p(fav) splits +cov	18.5	34.4	21.6	28.6	35.8	22.3	36.9	45.3	27.7	38.9	34.6	34.4

8% treatment effect

All of the sliding dichotomy strategies gave a reduction in sample size compared with the conventional dichotomy for all studies using the 8% treatment effect as shown in Table 5-2 below. Using five bands equal splits with and without covariates again gave the smallest reductions in sample size compared with the other strategies. Similar results were observed both when having equal numbers of patients per band and when banding by the proportion of favourable outcomes.

Table 5-2 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, three covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	26.5	29.8	28.1	27.5	29.8	29.4	34.2	38.2	26.2	35.9	32.7	29.8
4 bands, equal splits	27.6	29.8	26.2	28.4	29.5	21.7	36.5	38.7	21.2	36.7	31.8	29.5
5 bands, equal splits	14.1	12.3	12.6	16.1	19.0	7.0	27.5	30.9	10.0	27.4	19.4	16.1
3 bands, equal splits +cov	27.4	32.8	29.7	28.0	31.4	29.6	38.7	42.4	26.0	39.5	33.9	31.4
4 bands, equal splits +cov	29.7	31.4	29.4	29.5	33.1	25.0	38.2	43.2	24.0	40.8	32.0	31.4
5 bands, equal splits +cov	16.3	14.1	14.7	17.5	21.6	7.0	28.6	34.9	8.9	28.9	20.5	17.5
3 bands, p(fav) splits	28.8	28.4	26.5	28.2	28.7	18.3	36.0	39.2	17.6	35.2	29.9	28.7
4 bands, p(fav) splits	25.6	25.7	21.8	23.1	26.8	14.7	34.2	38.4	13.8	32.8	24.0	25.6
5 bands, p(fav) splits	27.9	28.7	23.7	26.7	30.2	17.6	34.9	40.4	19.2	36.0	27.6	27.9
3 bands, p(fav) splits +cov	32.5	28.7	28.2	29.9	32.8	21.7	41.4	42.9	22.4	41.0	32.3	32.3
4 bands, p(fav) splits +cov	29.4	29.4	23.1	25.2	29.3	16.6	35.3	40.8	16.3	36.0	26.5	29.3
5 bands, p(fav) splits +cov	30.4	30.0	25.8	28.0	33.7	19.9	37.6	42.3	20.2	39.0	28.5	30.0

5.5.1.2 Seven covariate model

For both the 5% and 8% treatment effects, all sliding dichotomy scenarios showed a reduction in sample size compared with the conventional analysis as shown in Table 5-3 and Table 5-4 respectively.

5% treatment effect

Here, typically, using the scenarios based on having equal numbers of patients per band gave slightly greater reductions in sample size than using the scenarios which banded by the probability of a favourable outcome. For all scenarios, using covariates gave a modest increase in efficiency. Using five bands with equal splits gave the smallest reductions in sample size.

Table 5-3 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, seven covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	24.7	33.1	25.7	33.2	39.6	41.5	39.9	42.1	27.6	29.2	30.1	33.1
4 bands, equal splits	29.2	39.8	27.1	33.3	41.6	36.8	41.7	34.2	26.0	28.4	28.1	33.3
5 bands, equal splits	7.3	28.2	10.2	11.6	28.9	25.6	33.0	27.4	17.8	27.5	16.9	25.6
3 bands, equal splits +cov	24.9	39.8	28.1	36.8	43.8	42.8	44.9	50.6	31.9	40.0	34.1	39.8
4 bands, equal splits +cov	32.3	49.1	26.9	42.8	44.6	40.5	43.6	43.0	30.6	37.8	31.7	40.5
5 bands, equal splits +cov	15.8	35.5	11.8	30.8	31.6	23.1	37.7	32.9	13.5	29.9	15.0	29.9

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
3 bands, p(fav) splits	25.6	37.8	23.2	33.3	36.4	38.2	44.0	35.6	20.2	28.7	29.3	33.3
4 bands, p(fav) splits	25.4	35.5	13.2	18.6	34.8	31.6	37.3	30.9	5.5	24.4	24.6	25.4
5 bands, p(fav) splits	29.4	37.6	18.4	26.7	36.0	35.7	40.9	34.2	13.0	26.0	28.1	29.4
3 bands, p(fav) splits +cov	32.3	47.7	23.6	40.8	44.4	36.2	44.0	41.7	26.8	40.7	32.7	40.7
4 bands, p(fav) splits +cov	27.1	44.2	22.3	35.2	35.2	32.6	41.6	41.0	16.6	37.0	27.4	35.2
5 bands, p(fav) splits +cov	31.7	47.0	22.8	36.1	41.3	35.7	44.4	42.9	23.3	39.4	27.4	36.1

8% treatment effect

As with the 5% treatment effect, using five bands with equal splits gave the smallest reductions in sample size although the reductions with the 8% treatment effect were greater than those with the 5% treatment effect. Using banding based on the probability of a favourable outcome gave slightly greater increases in sample size than using equal splits and, as previously, using covariates gave a small additional effect.

Table 5-4 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, seven covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	29.1	38.1	20.4	32.4	35.0	31.2	36.5	42.4	30.3	39.6	34.7	34.7
4 bands, equal splits	44.5	44.3	36.3	40.3	42.4	38.5	38.5	39.7	26.9	43.5	29.8	39.7
5 bands, equal splits	22.9	34.7	1.7	24.4	21.8	17.4	32.1	43.7	30.9	39.6	21.7	24.4
3 bands, equal splits +cov	34.2	41.1	22.4	33.9	38.9	35.2	43.7	48.4	32.6	46.5	39.1	38.9
4 bands, equal splits +cov	46.9	44.2	37.0	34.0	46.1	42.2	37.5	46.1	30.7	40.7	32.7	40.7
5 bands, equal splits +cov	29.8	30.3	22.4	19.6	32.3	23.9	31.6	38.7	15.9	32.8	15.9	29.8
3 bands, p(fav) splits	38.4	45.1	22.2	40.1	33.3	30.1	38.6	50.4	39.0	41.7	42.1	39.0
4 bands, p(fav) splits	33.3	41.2	15.9	34.2	31.7	28.4	35.1	47.2	35.4	43.1	34.3	34.3
5 bands, p(fav) splits	34.3	44.2	22.0	38.7	35.3	29.7	38.3	49.8	37.3	44.1	40.5	38.3
3 bands, p(fav) splits +cov	47.2	42.3	35.1	36.6	45.8	40.1	39.4	46.1	32.0	44.1	30.7	40.1
4 bands, p(fav) splits +cov	41.3	40.6	32.2	30.2	44.3	36.3	36.0	44.7	25.4	40.4	28.4	36.3
5 bands, p(fav) splits +cov	44.3	41.8	35.1	33.9	45.6	38.1	38.9	44.7	28.0	40.1	32.3	38.9

5.5.1.3 Nine covariate model

5% and 8% treatment effects

A similar pattern was observed both for the 5% and 8% treatment effects. Using nine covariates, all sliding dichotomy strategies gave a reduction in sample size of typically 30-40% compared with the conventional dichotomous analysis, as shown in Table 5-5 and Table 5-6 below. Using four bands, when banding by sample size, gave the greatest reductions in sample size with smaller reductions observed with using three and then five bands. Those scenarios that used the proportion of favourable outcomes to group subjects showed a similar picture with the three, four and five bands, with here the three band scenarios typically performing the best.

Table 5-5 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, nine covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	35.1	43.6	20.3	28.3	30.7	35.0	40.4	37.5	35.8	36.2	32.6	35.1
4 bands, equal splits	39.6	49.3	29.8	30.4	29.0	32.0	42.4	37.9	37.9	36.2	22.3	36.2
5 bands, equal splits	16.4	42.0	16.9	8.0	22.4	22.9	32.6	31.8	26.8	37.0	9.9	22.9
3 bands, equal splits +cov	39.8	47.8	23.9	32.1	35.8	38.7	45.2	47.4	37.6	44.7	35.1	38.7
4 bands, equal splits +cov	41.8	50.6	30.5	32.3	32.6	35.6	41.9	47.8	39.2	43.5	30.6	39.2
5 bands, equal splits +cov	28.8	43.3	16.4	12.5	23.6	24.8	30.8	36.9	28.6	36.8	12.3	28.6
3 bands, p(fav) splits	31.7	48.4	28.4	29.3	27.6	31.2	41.1	35.9	27.6	36.0	28.2	31.2

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
4 bands, p(fav) splits	32.5	46.4	20.5	24.5	27.8	27.7	34.8	33.2	23.8	34.5	16.9	27.8
5 bands, p(fav) splits	32.5	47.4	27.8	26.2	27.0	30.4	36.0	32.1	30.0	35.4	23.5	30.4
3 bands, p(fav) splits +cov	42.6	50.7	30.1	32.0	33.1	37.7	40.0	48.1	34.4	43.8	30.4	37.7
4 bands, p(fav) splits +cov	39.6	46.6	24.9	23.6	30.3	30.4	37.8	44.0	30.4	39.6	21.6	30.4
5 bands, p(fav) splits +cov	40.5	50.2	30.3	27.5	36.7	32.2	39.1	45.0	32.8	39.6	25.2	36.7

Table 5-6 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, nine covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	24.0	39.4	21.3	29.8	30.9	32.1	41.5	40.4	34.4	38.9	29.4	32.1
4 bands, equal splits	39.9	42.5	36.2	36.6	46.6	43.6	40.2	42.2	31.6	29.1	28.3	39.9
5 bands, equal splits	24.0	30.7	19.0	14.6	30.1	24.0	34.6	42.3	32.0	30.3	22.5	30.1
3 bands, equal splits +cov	28.5	44.1	23.6	34.9	37.2	35.8	47.8	45.1	37.1	43.7	33.7	37.1
4 bands, equal splits +cov	43.8	53.8	38.3	41.8	48.9	46.4	54.2	48.2	36.8	45.3	32.8	45.3
5 bands, equal splits +cov	27.0	40.0	22.4	26.9	34.2	30.1	42.8	40.1	27.0	40.1	22.7	30.1
3 bands, p(fav) splits	38.6	41.9	36.0	33.1	45.6	40.3	39.6	45.7	40.3	33.6	38.0	39.6

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
4 bands, p(fav) splits	36.2	38.3	30.2	30.4	40.3	34.6	38.4	44.0	37.1	30.1	34.4	36.2
5 bands, p(fav) splits	37.1	41.1	35.8	32.9	43.9	37.6	38.8	44.3	42.1	30.6	36.4	37.6
3 bands, p(fav) splits +cov	44.0	51.5	39.0	43.3	50.0	44.1	54.2	50.1	34.0	47.2	34.6	44.1
4 bands, p(fav) splits +cov	38.2	51.2	30.8	39.3	44.9	43.0	49.7	47.4	32.5	43.9	30.3	43.0
5 bands, p(fav) splits +cov	43.6	52.6	38.1	41.1	46.2	43.8	54.6	48.9	31.3	46.4	33.9	43.8

5.5.1.4 Uniform treatment effect graphical comparison

Figure 5-1, Figure 5-2 and Figure 5-3 below show the results by trial for the Uniform 5% treatment effect when banding by sample size for three, four and five bands respectively. The figures show the three covariate models (three, seven and nine), both unadjusted and adjusted. The studies are ordered with the trials together then the three surveys, UK4, TCDB and EBIC. Similar reductions in sample size are observed using three and four bands, with five bands giving smaller reductions. For most studies the addition of covariates gave additional sample size reductions, the adjusted results. Using models based on seven or nine covariates generally gave greater sample size reductions than using three covariates although this was not consistent over all studies.

Graphs showing the results by trial banded by the probability of a favourable outcome for a Uniform 5% treatment effect are shown in Figure 5-4, Figure 5-5 and Figure 5-6 for the three, four and five bands respectively. For all studies using either three, four or five bands and banding by the probability of a favourable outcome

gives a reduction in sample size compared with the conventional dichotomous analysis. Here, the addition of covariates gives a modest additional gain for all three models. A similar pattern of results is observed over three, four and five bands. This is different to the results observed when banding was done by sample size, where using five bands gave smaller reductions in sample size than using three or four bands.

Figure 5-7, Figure 5-8 and Figure 5-9 below show the results by trial for the sliding dichotomy simulations for the Uniform 8% treatment effect when banding by sample size for three, four and five bands respectively. For all studies banding by three, four or five bands gave a reduction in sample size compared with the conventional dichotomy. Using five bands gave the smallest reductions in sample size. Using three or four bands gave similar reductions. For most studies using the seven or nine covariate models gave greater reductions than the three covariate model.

As with the 5% treatment effect, banding by the probability of a favourable outcome gave a reduction in sample size for all studies modelling the 8% treatment effect, as shown in Figure 5-10, Figure 5-11 and Figure 5-12. The adjusted models typically gave greater reductions than the unadjusted models. However, using the nine covariate model did not consistently give greater reductions than the seven covariate model.

Figure 5-13 and Figure 5-14 show the median reductions in sample size over all trials by prognostic banding group for the Uniform 5% and 8% treatment effects respectively. These show that using either three, four or five bands and banding by sample size or by the probability of a favourable outcome, all covariate models show a reduction in sample size compared with the conventional dichotomous model. When banding by the probability of a favourable outcome a similar pattern is

observed when using three, four or five bands. When banding by sample size using five bands gives smaller reductions than either three or four bands.

Figure 5-1 Sliding Dichotomy: Uniform 5% Equal splits: 3 bands[6]

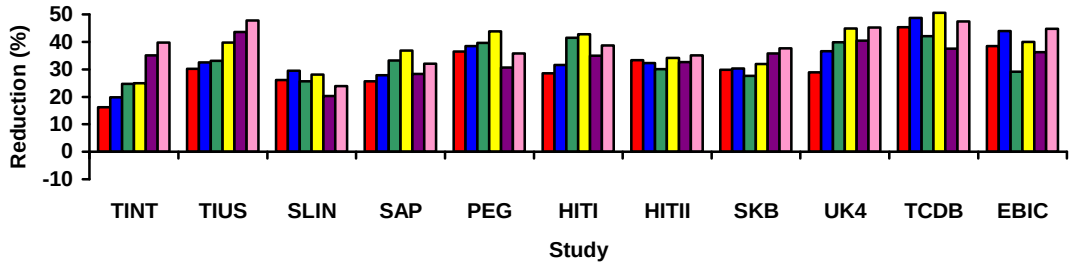


Figure 5-2 Sliding Dichotomy: Uniform 5% Equal splits: 4 bands[6]

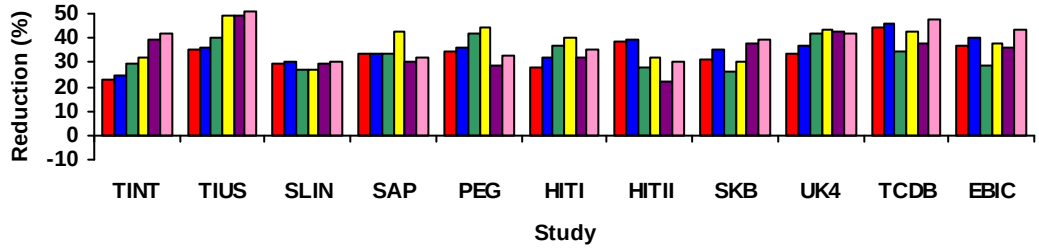
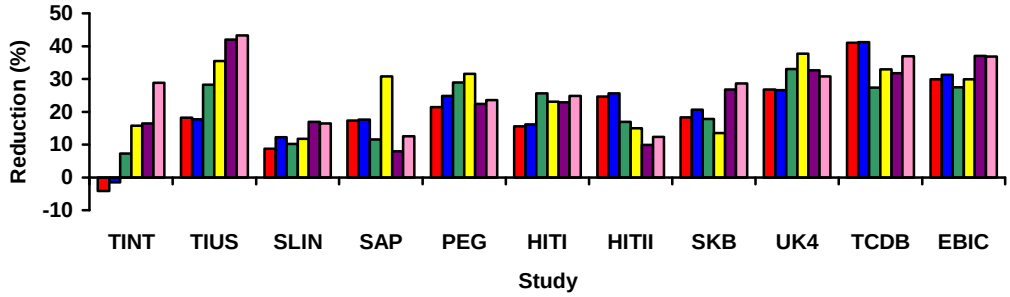


Figure 5-3 Sliding Dichotomy: Uniform 5% Equal splits: 5 bands[6]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

[6] Bars are the percentage reductions in sample size which can be achieved while preserving the power relative to the conventional unadjusted analysis of the dichotomised outcome scale. 'Equal splits' means prognostic bands chosen to contain equal numbers of patients. 'p(fav)' means prognostic bands chosen with specific ranges of the predicted probability of a favourable outcome.

Figure 5-4 Sliding Dichotomy: Uniform 5% p(fav) splits: 3 bands[6]

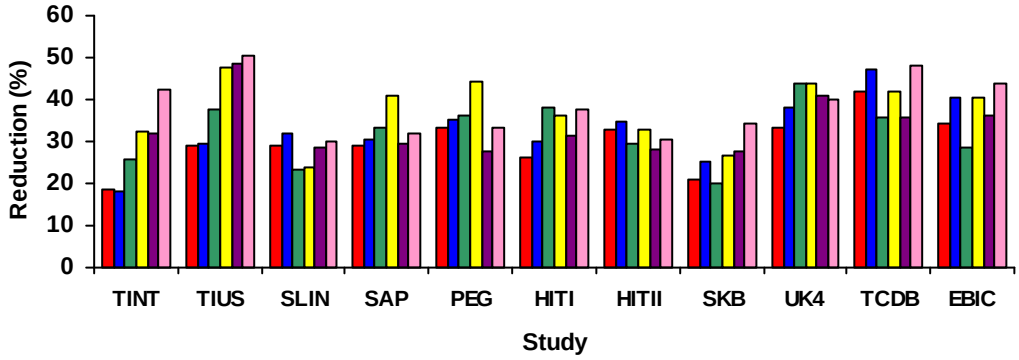


Figure 5-5 Sliding Dichotomy: Uniform 5% p(fav) splits: 4 bands[6]

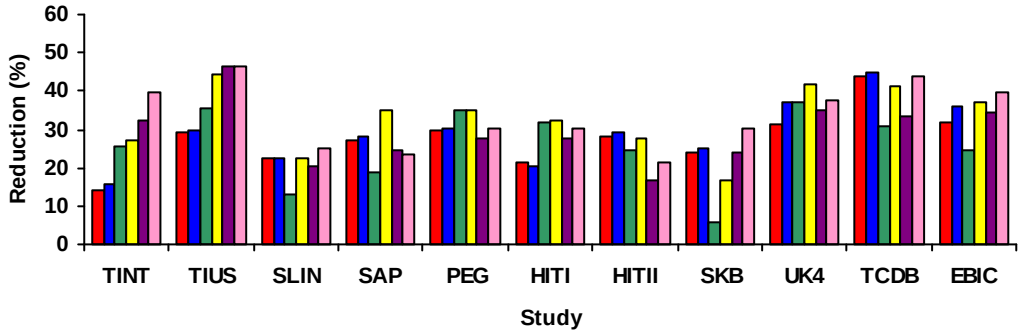
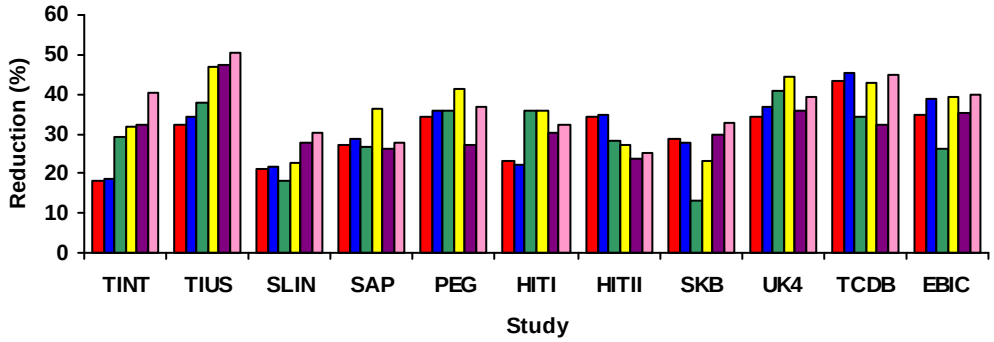


Figure 5-6 Sliding Dichotomy: Uniform 5% p(fav) splits: 5 bands[6]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

Figure 5-7 Sliding Dichotomy: Uniform 8% Equal splits: 3 bands[6]

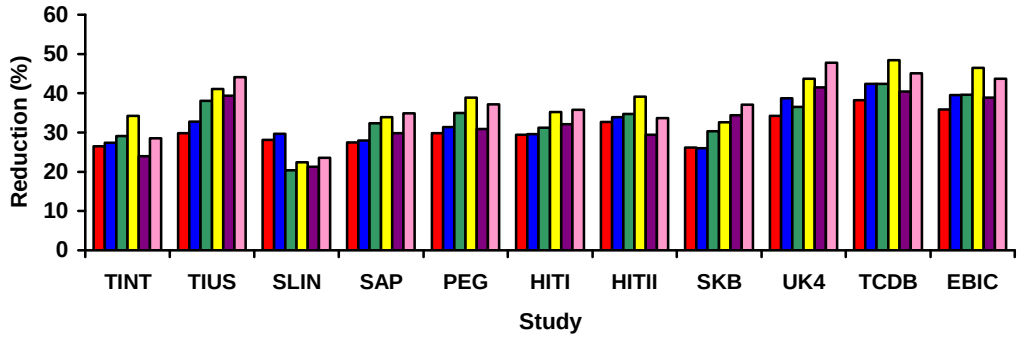


Figure 5-8 Sliding Dichotomy: Uniform 8% Equal splits: 4 bands[6]

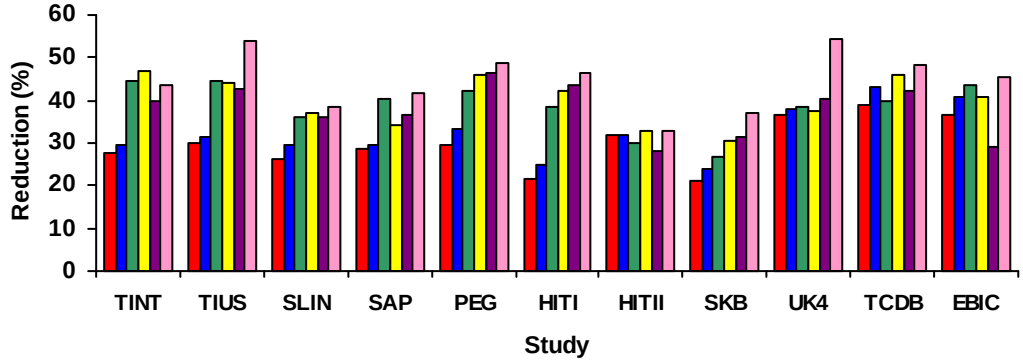
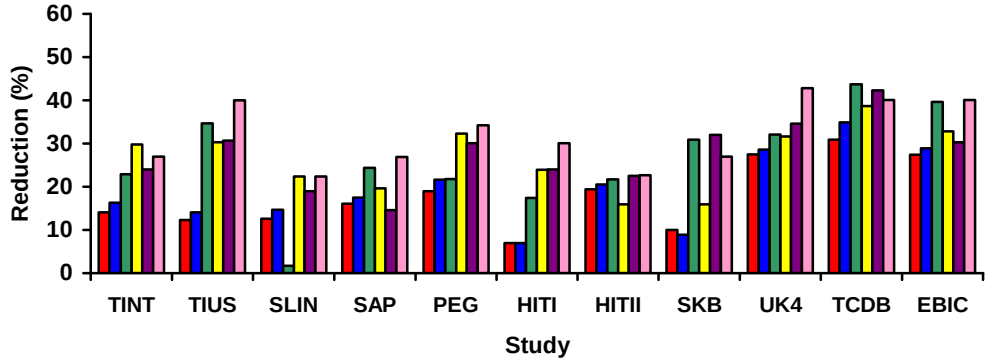


Figure 5-9 Sliding Dichotomy: Uniform 8% Equal splits: 5 bands[6]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

Figure 5-10 Sliding Dichotomy: Uniform 8% p(fav) splits: 3 bands[6]

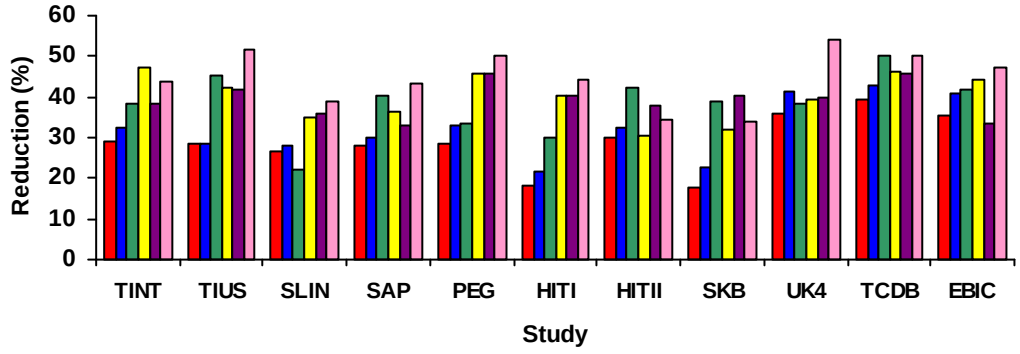


Figure 5-11 Sliding Dichotomy: Uniform 8% p(fav) splits: 4 bands[6]

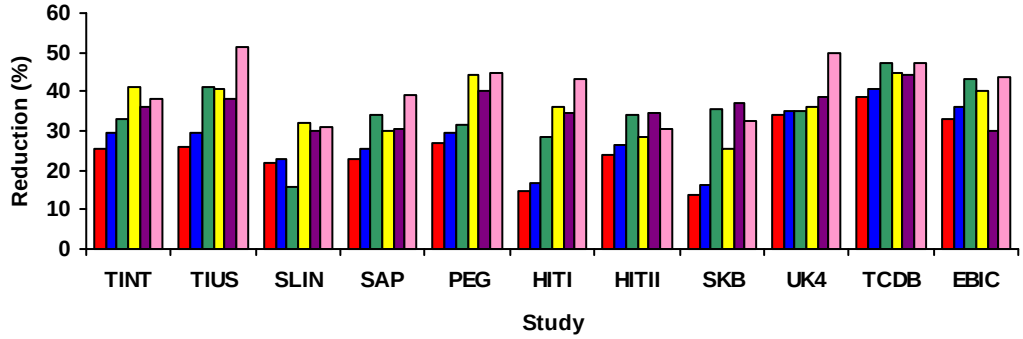
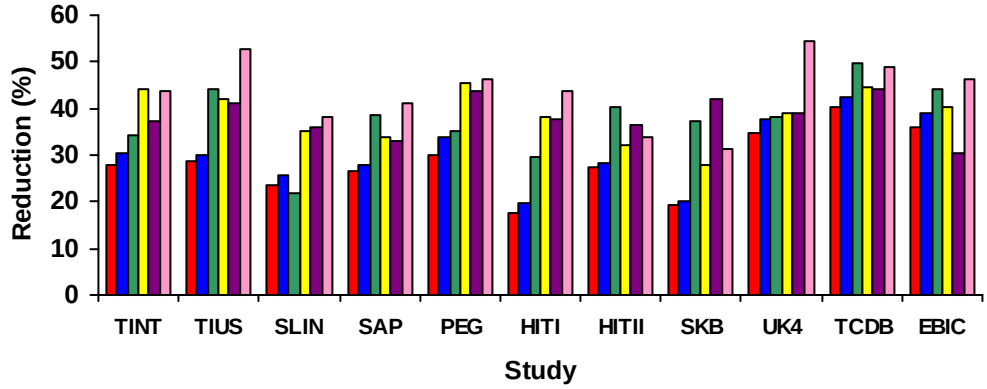


Figure 5-12 Sliding Dichotomy: Uniform 8% p(fav) splits: 5 bands[6]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

Figure 5-13 Sliding Dichotomy: Median reductions in sample size by prognostic banding group. Uniform treatment scenario, 5% treatment effect for all subjects[7]

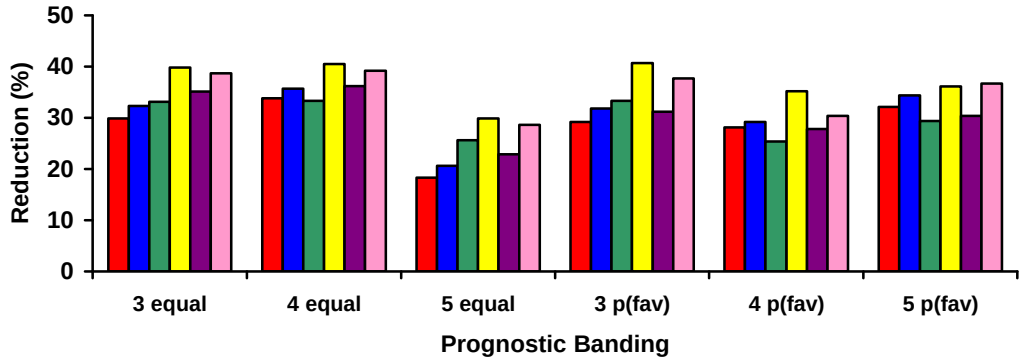
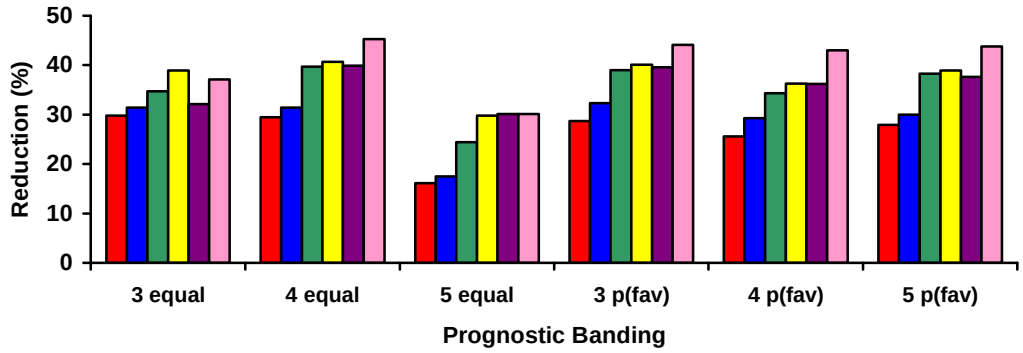


Figure 5-14 Sliding Dichotomy: Median reductions in sample size by prognostic banding group. Uniform treatment scenario, 8% treatment effect for all subjects[7]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

[7] Bars are the percentage reductions in sample size which can be achieved while preserving the power relative to the conventional unadjusted analysis of the dichotomised outcome scale. 'Equal' means prognostic bands chosen to contain equal numbers of patients. 'p(fav)' means prognostic bands chosen with specific ranges of the predicted probability of a favourable outcome. The number refers to the number of bands.

5.5.2 Sliding dichotomy modelling - mortality treatment effect

5% and 8% treatment effects

Here the treatment effect is a reduction in mortality. No sliding dichotomy strategy gave a reduction in sample size for SLIN for both the 5% and 8% treatment effects.

5.5.2.1 Three covariate model

5% treatment effect

Table 5-7 below shows that, with the exception of the three series (UK4, TCDB and EBIC), using five bands equal splits both with and without covariates gave an increase in sample size compared with the conventional dichotomous analysis for all studies. All other scenarios showed a reduction in sample size for all studies except TINT and SLIN. In general greater reductions were observed when equal splits were used as opposed to banding by the probability of a favourable outcome although the reductions were similar overall.

Table 5-7 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, three covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	15.1	21.9	-47.1	34.0	35.1	36.9	42.9	34.6	38.0	37.3	40.3	35.1
4 bands, equal splits	18.5	24.4	-14.8	16.5	32.5	23.9	52.4	49.8	28.8	32.8	23.1	24.4
5 bands, equal splits	-119.6	-87.6	-115.7	-75.8	-9.2	-14.2	15.1	14.4	-27.8	6.3	-42.3	-27.8
3 bands, equal splits +cov	18.8	23.6	-52.4	35.6	38.0	36.6	44.7	37.9	36.6	40.5	40.5	36.6

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
4 bands, equal splits +cov	17.6	25.2	-12.8	20.2	33.2	26.9	54.3	52.9	31.6	38.1	23.3	26.9
5 bands, equal splits +cov	-109.8	-84.4	-115.7	-56.1	-10.0	-14.7	14.8	17.1	-29.4	4.4	-45.5	-29.4
3 bands, p(fav) splits	3.4	9.8	-24.5	10.5	27.7	16.7	44.0	43.1	9.7	26.0	21.5	16.7
4 bands, p(fav) splits	-14.5	14.2	-64.2	11.7	29.9	29.4	47.8	45.7	30.1	34.2	19.4	29.4
5 bands, p(fav) splits	-0.9	19.4	-29.0	22.6	37.2	32.8	49.1	48.1	29.8	39.9	29.5	29.8
3 bands, p(fav) splits +cov	8.5	13.8	-17.5	13.5	28.5	21.7	46.0	46.1	11.4	31.2	23.3	21.7
4 bands, p(fav) splits +cov	-19.9	13.8	-51.5	12.0	32.6	29.2	51.0	47.1	30.1	37.3	21.5	29.2
5 bands, p(fav) splits +cov	3.0	19.4	-25.8	21.4	37.4	33.5	51.4	50.3	30.7	41.5	28.2	30.7

8% treatment effect

In general, using 3 or 4 bands, equal splits with covariates gave the greatest reductions. With TINT the only very modest advantage over the conventional analysis was seen when using 4 bands equal splits. All other scenarios showed a disadvantage of using the sliding dichotomy with increases in the sample size required from 7% up to 148% as shown in Table 5-8.

Table 5-8 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, three covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	-25.5	5.1	-65.6	31.3	42.3	37.2	36.5	38.1	33.5	38.0	36.5	36.5
4 bands, equal splits	5.9	16.6	-41.4	26.4	35.5	25.9	43.9	43.2	19.1	24.2	19.8	24.2
5 bands, equal splits	-148.2	-197.6	-184.6	-70.3	-10.4	-42.3	12.3	5.4	-63.5	-3.5	-74.3	-63.5
3 bands, equal splits +cov	-22.3	5.5	-61.8	32.9	43.0	38.0	37.7	42.5	34.0	39.9	37.5	37.5
4 bands, equal splits +cov	7.2	18.1	-40.7	26.8	38.9	28.0	46.6	46.3	19.1	27.7	21.1	26.8
5 bands, equal splits +cov	-138.1	-180.2	-175.6	-67.9	-7.7	-40.9	13.3	6.0	-60.7	-3.3	-71.3	-60.7
3 bands, p(fav) splits	-12.1	-2.4	-28.0	20.0	32.8	22.6	39.6	35.5	7.3	27.4	16.7	20.0
4 bands, p(fav) splits	-33.6	-5.4	-78.8	13.9	41.4	25.3	46.2	41.6	5.2	30.2	13.1	13.9
5 bands, p(fav) splits	-9.5	6.4	-46.5	24.4	43.0	29.5	48.6	45.5	10.7	36.7	21.7	24.4
3 bands, p(fav) splits +cov	-7.4	0.7	-28.7	21.7	36.0	24.7	42.3	40.1	8.2	31.4	20.8	21.7
4 bands, p(fav) splits +cov	-24.8	-3.6	-75.5	16.3	42.3	27.8	47.7	43.9	7.7	34.0	15.3	16.3
5 bands, p(fav) splits +cov	-6.9	9.4	-46.1	25.0	44.1	29.1	49.5	49.2	12.8	38.2	23.4	25.0

5.5.2.2 Seven covariate model

5% treatment effect

For most studies using three or four bands with equal splits resulted in greater reductions in sample size than banding by the proportion of favourable outcomes as shown in Table 5-9 below. Using five bands with equal splits however gave an increase in sample size compared with the conventional analysis for seven of the eleven studies. Using covariates gave a modest increase in efficiency.

Table 5-9 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, seven covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	29.5	32.9	23.7	36.4	38.1	40.2	38.0	41.7	43.2	39.3	37.2	38.0
4 bands, equal splits	24.9	24.8	12.0	16.6	13.6	23.9	52.9	56.6	56.4	40.4	29.2	24.9
5 bands, equal splits	-68.1	-93.1	-77.9	-79.6	-34.8	-51.4	18.9	19.5	29.4	6.5	-44.1	-44.1
3 bands, equal splits +cov	33.6	34.2	24.4	37.9	41.3	42.9	39.8	47.9	43.9	42.5	38.9	39.8
4 bands, equal splits +cov	25.6	27.0	14.5	7.3	19.7	28.0	53.6	58.7	55.8	47.1	31.6	28.0
5 bands, equal splits +cov	-83.8	-138.2	-126.6	-102.7	-20.0	-89.6	10.8	29.3	29.4	6.1	-29.7	-29.7
3 bands, p(fav) splits	13.6	19.1	14.0	13.0	24.5	24.6	45.0	44.6	41.7	34.7	28.1	24.6
4 bands, p(fav) splits	19.4	20.1	-7.8	12.5	27.4	19.9	49.0	47.8	51.6	38.1	21.8	21.8

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
5 bands, p(fav) splits	23.1	28.0	12.5	24.0	35.5	28.6	53.3	50.3	51.3	42.1	28.8	28.8
3 bands, p(fav) splits +cov	20.6	11.9	-0.4	15.3	20.2	25.6	42.0	54.6	46.4	41.9	32.4	25.6
4 bands, p(fav) splits +cov	18.3	10.0	-30.2	13.0	30.7	25.4	49.3	55.0	53.7	45.8	30.7	30.7
5 bands, p(fav) splits +cov	22.0	20.4	-8.2	23.1	33.8	31.2	52.5	57.6	54.8	45.8	35.7	33.8

8% treatment effect

For all studies using three or four bands with equal splits both with and without covariates gave the greatest reductions in sample size as shown in Table 5-10. As with the 5% treatment effect using five bands with equal splits gave an increase in sample size compared with the conventional dichotomy for seven of the eleven studies. In general using the equal splits gave greater reductions than banding by the probability of a favourable outcome.

Table 5-10 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, seven covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	17.2	28.0	7.3	39.6	40.8	41.9	40.7	45.7	39.2	42.1	37.1	39.6
4 bands, equal splits	21.9	31.7	28.9	19.3	23.7	36.4	43.1	50.7	44.0	35.4	29.8	31.7

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
5 bands, equal splits	-156.7	-130.7	-128.9	-103.1	-32.1	-82.8	2.9	20.8	11.0	18.2	-66.4	-66.4
3 bands, equal splits +cov	20.8	30.9	8.8	40.3	43.8	43.6	40.7	47.2	40.0	46.4	40.0	40.3
4 bands, equal splits +cov	25.5	33.5	28.5	23.0	27.4	37.2	50.9	55.8	41.9	39.8	32.2	33.5
5 bands, equal splits +cov	-154.6	-120.4	-118.1	-78.3	-45.6	-66.0	12.8	23.7	7.7	14.2	-72.9	-66.0
3 bands, p(fav) splits	12.3	23.1	11.7	12.4	32.4	32.1	39.6	50.5	33.0	40.4	25.7	32.1
4 bands, p(fav) splits	-5.0	17.1	-36.4	11.5	37.9	24.8	43.1	52.2	42.4	41.3	13.9	24.8
5 bands, p(fav) splits	8.4	24.5	-10.2	19.2	42.2	34.1	46.8	55.1	42.7	45.2	25.0	34.1
3 bands, p(fav) splits +cov	7.5	21.9	13.1	18.6	32.4	36.0	44.7	51.7	33.7	47.6	26.3	32.4
4 bands, p(fav) splits +cov	-9.3	18.8	-17.4	15.6	37.5	24.6	48.6	56.7	41.4	44.9	18.6	24.6
5 bands, p(fav) splits +cov	-0.9	24.7	-1.9	20.5	40.0	32.4	50.7	57.5	41.7	50.7	26.1	32.4

5.5.2.3 Nine covariate model

5% treatment effect

For the scenarios using three and four bands using banding based on sample size resulted in slightly greater reductions in sample size than using banding based on the probability of a favourable outcome as shown in Table 5-11 below. Five bands equal

splits both with and without covariates performed poorly again resulting in an increase in sample size compared to the conventional dichotomy for most studies. Banding by the probability of a favourable outcome and using four or five bands gave a reduction in sample size compared with the conventional dichotomy for all studies except SLIN.

Table 5-11 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, nine covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	33.5	25.2	20.7	42.0	38.9	40.1	33.2	32.3	40.2	37.7	40.4	37.7
4 bands, equal splits	17.6	26.8	21.5	29.7	33.8	10.2	47.1	45.7	50.3	45.9	24.6	29.7
5 bands, equal splits	-155.6	-113.1	-143.2	-65.0	-22.8	-91.7	4.8	0.0	27.6	17.5	-27.5	-27.5
3 bands, equal splits +cov	37.5	29.1	23.3	42.8	43.3	42.9	36.6	38.1	42.8	43.8	42.5	42.5
4 bands, equal splits +cov	20.2	26.6	27.0	34.4	37.5	16.9	50.3	47.4	51.7	47.1	28.6	34.4
5 bands, equal splits +cov	-129.3	-96.0	-80.1	-48.5	-11.7	-88.9	13.5	9.6	28.4	14.5	-29.3	-29.3
3 bands, p(fav) splits	2.5	18.2	-3.7	22.6	31.8	15.1	38.2	37.4	40.6	38.8	31.0	31.0
4 bands, p(fav) splits	-6.3	12.4	-48.4	27.7	39.2	19.4	42.9	40.7	52.9	42.4	31.4	31.4
5 bands, p(fav) splits	3.7	21.0	-14.4	32.6	45.3	23.7	47.1	44.4	54.1	47.3	35.9	35.9

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
3 bands, p(fav) splits +cov	7.5	21.0	6.7	31.8	38.9	20.2	39.9	39.7	42.2	40.8	28.8	31.8
4 bands, p(fav) splits +cov	3.7	23.9	-15.4	32.2	46.8	15.1	47.6	45.0	50.2	40.8	32.6	32.6
5 bands, p(fav) splits +cov	16.1	26.6	3.1	38.9	49.2	23.5	51.1	49.1	53.6	45.1	33.9	38.9

8% treatment effect

As with the 5% treatment effect using five bands equal splits resulted in greater sample sizes than the conventional dichotomy as shown in Table 5-12 below. Here though greater sample size reductions were observed, in general, for the strategies based on banding the proportion of favourable outcomes rather than having equal splits based on sample size.

Table 5-12 Sliding Dichotomy comparison: Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, nine covariate model for all subjects[5]

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3 bands, equal splits	20.8	25.1	0.5	40.6	43.6	37.0	36.9	39.8	40.9	45.5	38.4	38.4
4 bands, equal splits	26.5	32.4	11.3	26.3	34.7	20.8	44.7	44.6	50.1	37.5	19.3	32.4
5 bands, equal splits	-151.2	-116.8	-120.8	-95.8	-25.4	-101.2	8.4	15.1	30.8	12.5	-58.4	-58.4
3 bands, equal splits +cov	22.8	28.7	2.7	43.2	46.4	38.3	40.6	42.8	42.6	50.2	40.1	40.6

	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	Median
4 bands, equal splits +cov	28.2	34.6	12.9	26.1	37.2	23.9	46.7	46.8	53.0	51.5	24.8	34.6
5 bands, equal splits +cov	-110.8	-126.8	-121.6	-94.0	-17.2	-78.3	9.0	15.9	18.3	18.9	-70.5	-70.5
3 bands, p(fav) splits	6.7	22.5	4.9	18.1	34.8	15.9	39.9	47.0	48.9	41.1	33.7	33.7
4 bands, p(fav) splits	-10.5	13.0	-35.8	7.7	39.7	9.4	45.7	51.5	53.7	39.8	27.9	27.9
5 bands, p(fav) splits	3.7	24.9	-18.0	20.2	44.0	15.3	47.7	53.2	57.6	44.3	33.2	33.2
3 bands, p(fav) splits +cov	14.4	24.8	1.6	16.9	36.7	24.1	41.3	45.4	42.6	53.2	27.4	27.4
4 bands, p(fav) splits +cov	-0.2	15.0	-32.5	13.5	42.4	16.7	49.1	46.1	48.4	53.8	23.1	23.1
5 bands, p(fav) splits +cov	10.5	24.2	-15.8	20.9	45.0	24.3	52.0	50.0	50.1	54.7	29.8	29.8

5.5.2.4 Mortality treatment effect graphical comparison

Figure 5-15, Figure 5-16 and Figure 5-17 below show the results by trial obtained when banding by sample size with three, four and five bands respectively using the 5% treatment effect. Using three or four bands gave reductions in sample size for the three covariate models both unadjusted and adjusted, for all studies except SLIN. Using five bands however gave much larger sample sizes for all of the trials except SKB seven and nine covariate models. For the three and four band models, using covariates gave a small additional benefit. The seven and nine covariate model gave greater sample size reductions than the three covariate model using four bands. However, using three bands little difference was observed between the three

covariate models. A very similar pattern was also observed using the 8% treatment effect as shown in Figure 5-21, Figure 5-22 and Figure 5-23.

Banding by the probability of a favourable outcome did show a reduction in sample size for almost all studies for three, four and five bands, with the 5% treatment effect, as shown in Figure 5-18, Figure 5-19 and Figure 5-20. SLIN showed an increase in sample size for almost all scenarios and TINT also showed an increase in sample size for some of the four band covariate models. Greater reductions in sample size were typically observed for the observational studies rather than the trials. As with the Uniform model, a similar pattern was observed over using three, four or five bands when banding by the proportion of favourable outcomes.

Modelling an 8% treatment effect and banding by the probability of a favourable outcome gave similar results to modelling the 5% treatment effect, as shown in Figure 5-24, Figure 5-25 and Figure 5-26 below. For SLIN and TINT using four bands gave an increase in sample size compared with using the conventional dichotomy for most of the covariate models. Again, the observational studies typically had greater reductions than the trials.

Figure 5-27 and Figure 5-28 below show the median reductions in sample size when banding both by sample size and by the probability of a favourable outcome for the 5% and 8% treatment effects respectively. Banding by the probability of a favourable outcome gave median reductions in sample size of around 40% irrespective of whether three, four or five bands were used. When banding by sample size using three or four bands also gave reductions of around 40% however using five bands consistently gave an increase in sample size of 40 to 60% when compared with the conventional dichotomy.

Figure 5-15 Sliding Dichotomy: Mortality 5% Equal splits: 3 bands[6]

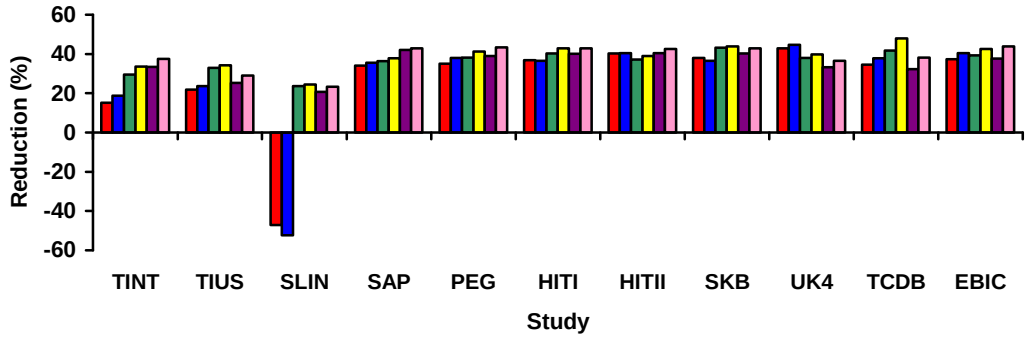


Figure 5-16 Sliding Dichotomy: Mortality 5% Equal splits: 4 bands[6]

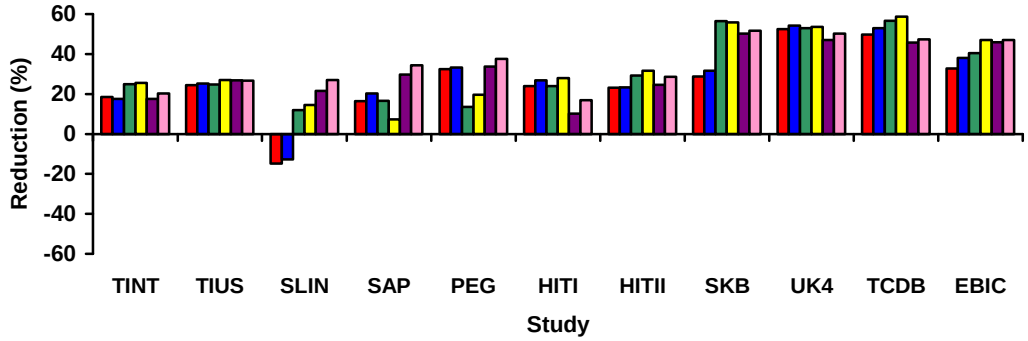
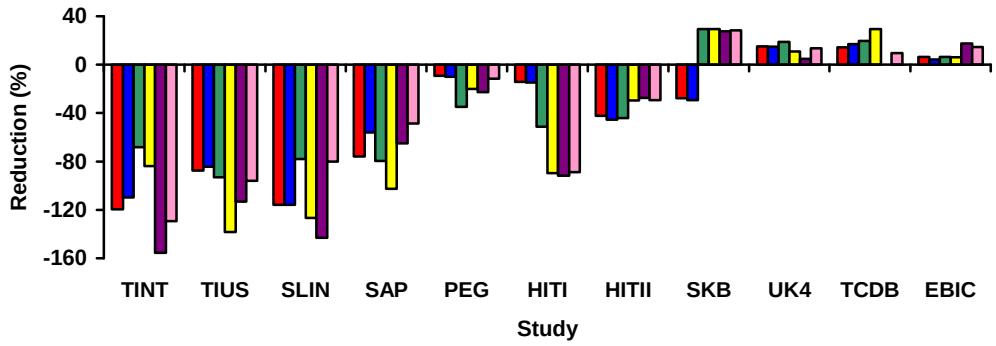


Figure 5-17 Sliding Dichotomy: Mortality 5% Equal splits: 5 bands[6]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

Figure 5-18 Sliding Dichotomy: Mortality 5% p(fav): 3 bands[6]

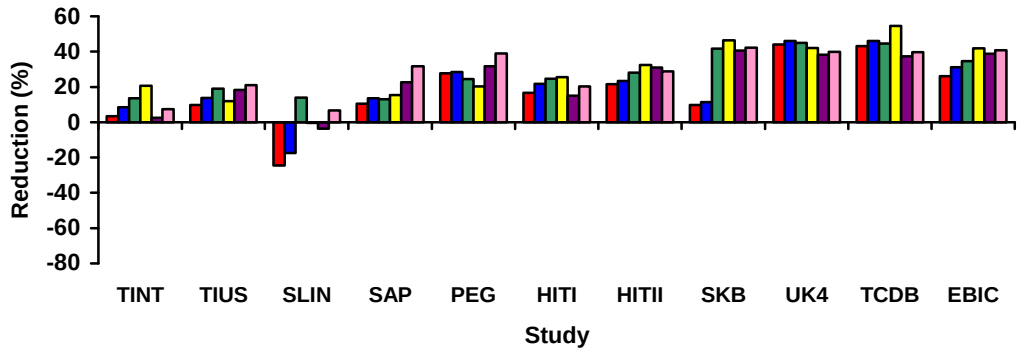


Figure 5-19 Sliding Dichotomy: Mortality 5% p(fav): 4 bands[6]

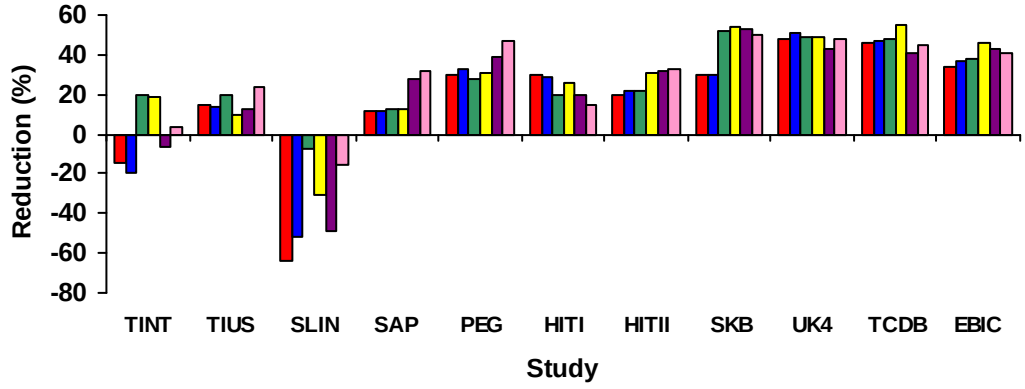
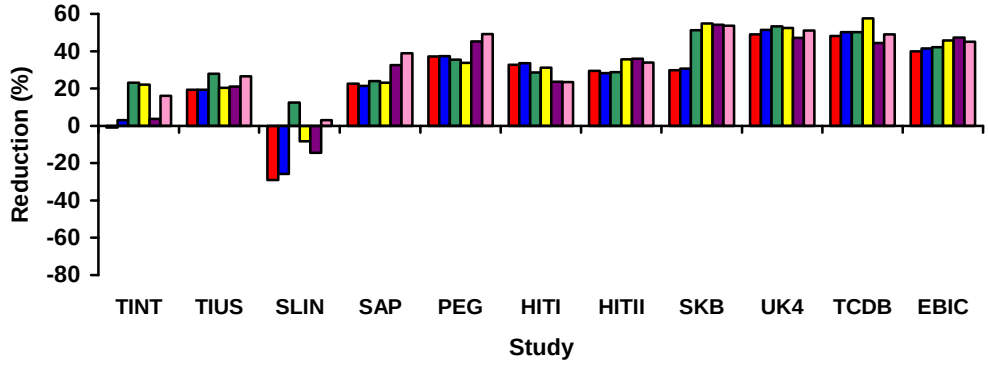


Figure 5-20 Sliding Dichotomy: Mortality 5% p(fav): 5 bands[6]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

Figure 5-21 Sliding Dichotomy: Mortality 8% Equal splits: 3 bands[6]

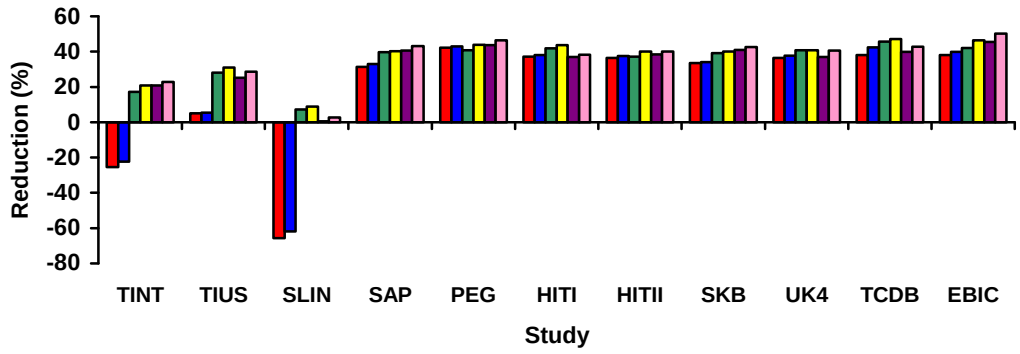


Figure 5-22 Sliding Dichotomy: Mortality 8% Equal splits: 4 bands[6]

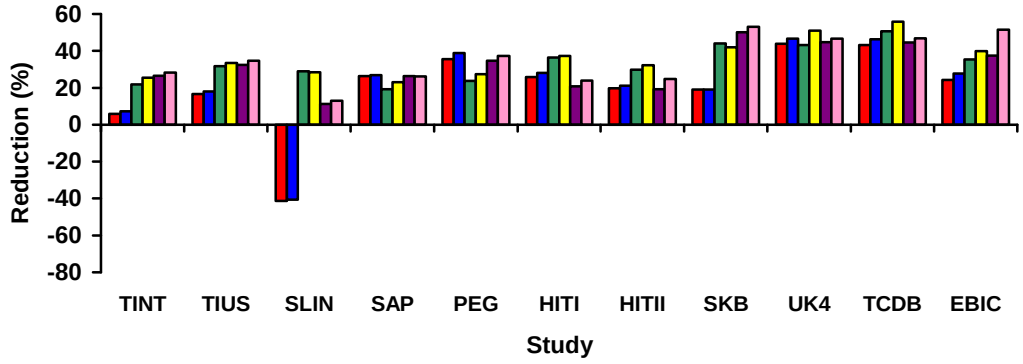
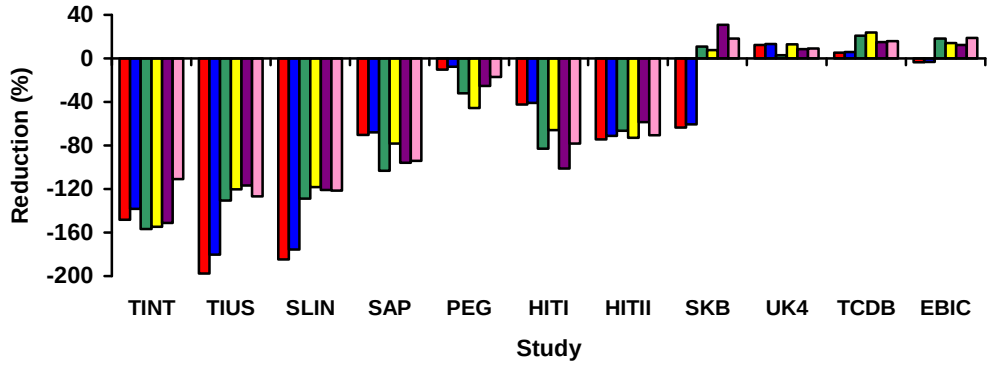


Figure 5-23 Sliding Dichotomy: Mortality 8% Equal splits: 5 bands[6]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

Figure 5-24 Sliding Dichotomy: Mortality 8% p(fav) splits: 3 bands[6]

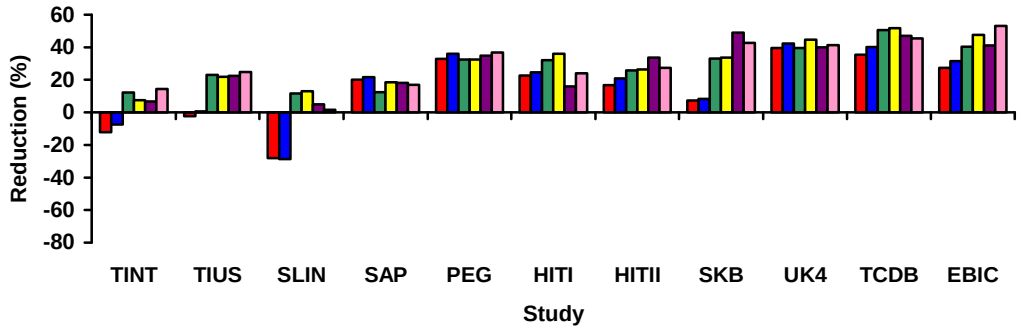


Figure 5-25 Sliding Dichotomy: Mortality 8% p(fav) splits: 4 bands[6]

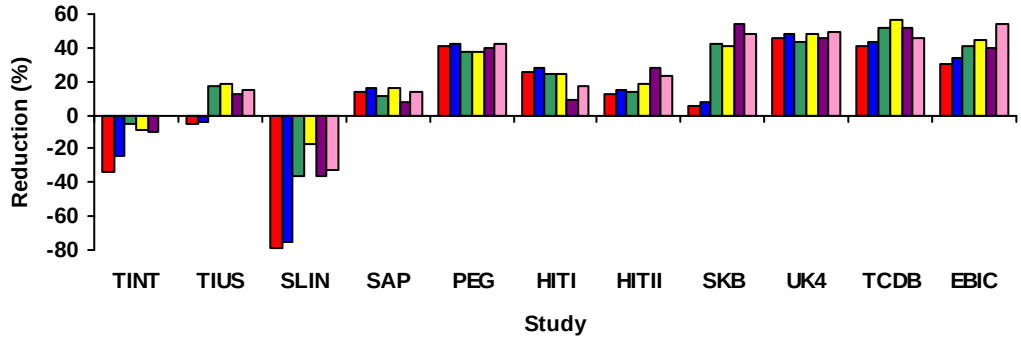
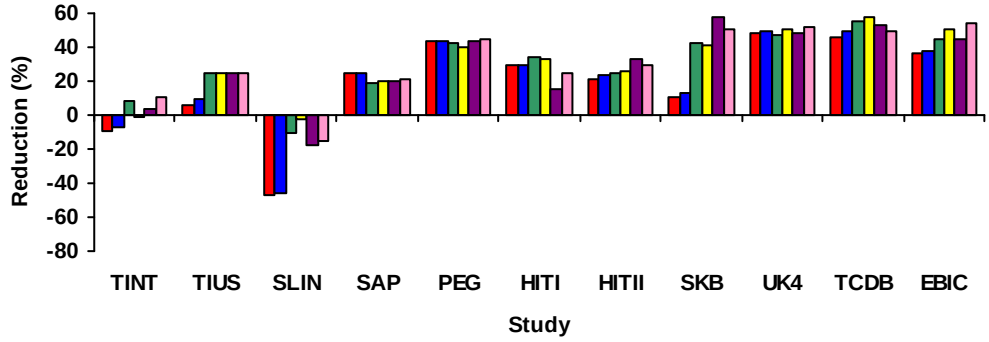


Figure 5-26 Sliding Dichotomy: Mortality 8% p(fav) splits: 5 bands[6]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

Figure 5-27 Sliding Dichotomy: Median reductions in sample size by prognostic banding group. Mortality treatment scenario, 5% treatment effect for all subjects[7]

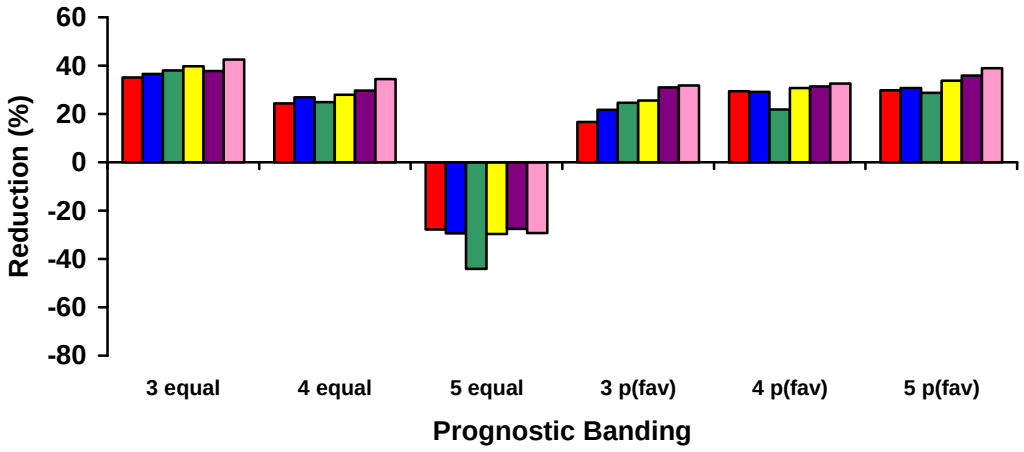
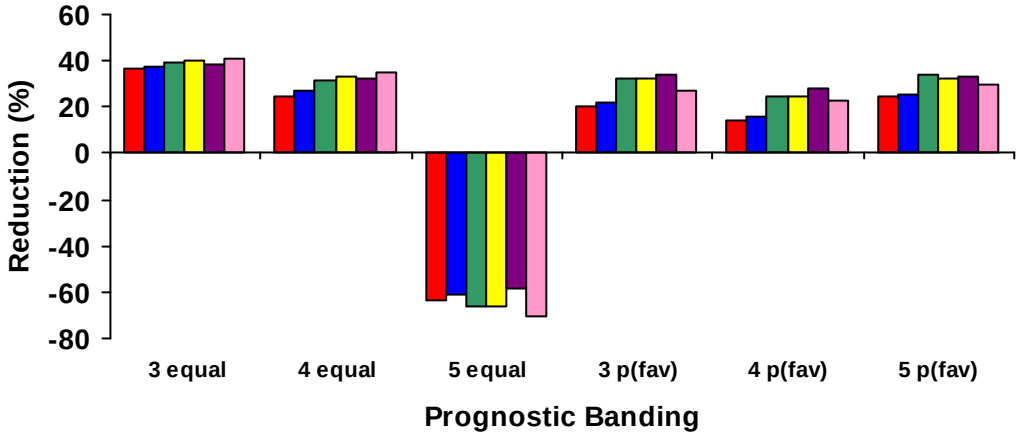


Figure 5-28 Sliding Dichotomy: Median reductions in sample size by prognostic banding group. Mortality treatment scenario, 8% treatment effect for all subjects[7]



Key: 3 covariate model unadjusted (red), 3 covariate model adjusted (dark blue), 7 covariate model unadjusted (green), 7 covariate model adjusted (yellow), 9 covariate model unadjusted (purple), 9 covariate model adjusted (pink)

5.6 Discussion

This chapter has shown that when modelling the sliding dichotomy a consistent pattern is observed. Banding by sample size, either three or four bands, with the addition of covariates gave typically the greatest sample size reductions. The sliding dichotomy performed poorly when using five splits based on sample size as shown both in the tables and in the graphs. Perhaps studies of 800 subjects are too small to show such refinement. Using the sliding dichotomy also depends on the constituent properties of the study sampled. For SLIN in particular, using the sliding dichotomy or the sliding dichotomy plus covariates resulted in an increase in sample size of 30% to 40% compared with using the conventional treatment only strategy. This is because subjects in the 'poor' prognosis band using the sliding dichotomy had a wide range of actual outcomes from poor to good, as illustrated below. Table 5-13 below shows the probability of a favourable outcome, p(fav), for each of the three bands of the sliding dichotomy for SLIN and EBIC. It can be seen for SLIN that for those in the poor prognosis band their probability of a favourable outcome ranges from 0.28 to 0.60. Whereas for EBIC, for those in the poor prognosis band their probability of a favourable outcome ranges from 0.22 to 0.46, a much narrower range.

Table 5-13 Probability of a favourable outcome for 3 bands for SLIN and EBIC

	SLIN	EBIC
	p(fav)	p(fav)
Poor prognosis	0.28 to 0.60	0.22 to 0.46
Mid prognosis	0.60 to 0.72	0.46 to 0.67
Best prognosis	0.72 to 0.82	0.67 to 0.84

Table 5-14 and Table 5-15 below show the percentage of subjects in each of the three sliding dichotomy bands for SLIN and EBIC respectively. These percentages are the numbers with the simulated GOS, simGOS. The line shows where the dichotomy

which gives closest to a 50:50 split is made within each prognosis group. It can be seen with EBIC data, Table 5-15, that the split is optimal for each of the three bands. However for SLIN, Table 5-14, it can be seen that for those in the best and mid prognosis bands the dichotomy is between good recovery versus everything else. For those in the poorest prognostic band the dichotomy is between severe and moderate.

Table 5-14 SLIN, percentage of subjects in each sliding dichotomy band

SLIN

	simGOS	%		simGOS	%		simGOS	%
	D/V	7		D/V	18		D/V	31
Best	S	17	Mid	S	18	Poor	S	23
	M	25		M	28		M	23
	G	51		G	36		G	23

Table 5-15 EBIC, percentage of subjects in each sliding dichotomy band

EBIC

	simGOS	%		simGOS	%		simGOS	%
	D/V	12		D/V	17		D/V	38
Best	S	15	Mid	S	24	Poor	S	27
	M	29		M	27		M	19
	G	43		G	33		G	16

The next stage in modelling, as shown in Chapter 6 is to take forward the optimised version of the sliding dichotomy for each of the covariate sets, treatment effect models and treatment effects and compare with the other strategies for analysing ordinal data.

6 Chapter 6 Comparing different modelling strategies

6.1 Introduction

Having refined the sliding dichotomy, as shown in Chapter 5, the next stage was to compare the sliding dichotomy against different modelling strategies to see which gives the greatest reduction in sample size (RSS) under which circumstances. These will include, amongst others, fitting a proportional odds models with a common odds ratio, modelling a reduction in mortality, modelling a proportional odds model with different odds ratios in different clinical subgroups and also different prognostic subgroups.

As previously three covariate sets (three, seven and nine covariates), two treatment scenarios Uniform and Mortality and two treatment effects 5% and 8% were modelled.

The optimised sliding dichotomy, the conventional analysis with covariates and the proportional odds model with and without covariates are all compared with the conventional dichotomy in the first part of this chapter.

In the second part it was of interest to see if the same pattern of results was observed when improvement was restricted to certain groups (both prognostic and physical) or when groups were targeted for improvement (again based on prognosis or a physical characteristic). As in Chapter 5, results are shown both by trial and as a median reduction in sample size over all trials.

6.1.1 Restricting improvement and targeting - methods

Restricting improvement

Two separate scenarios were explored using all subjects: restricting improvement to patients with an intermediate prognosis and restricting improvement to patients with a mass lesion. That is only applying a treatment effect to these patients.

For the improvement restricted to patients with an intermediate prognosis the treatment effect (5% or 8%) was only applied to those subjects with an individual probability of a favourable outcome between 0.2 and 0.8 inclusive. This was obtained by fitting a conventional binary logistic regression model with the conventional favourable/unfavourable dichotomy of the GOS as the response variable. The modelling of the uniform and mortality treatment effects is as previously described in Chapter 5. For those subjects with either a low (<0.2) or high (>0.8) probability of a favourable outcome then the probabilities remained as the original probabilities from the multinomial model, i.e. no treatment effect was applied.

Restricting improvement to patients with a mass lesion was done in a very similar way. Here, only patients who had a mass lesion had the treatment effect applied with all other subject's probabilities remaining as originally modelled.

Targeting

Here also two separate scenarios were explored: targeting only patients with an intermediate prognosis and targeting only patients with a mass lesion. That is only using these patients in the analysis.

For the first, targeting only patients with an intermediate prognosis, the different strategies were compared only for subjects with an intermediate prognosis of a favourable outcome, between 0.2 and 0.8 inclusive. The numbers available for analysis were therefore: 6279 for the three covariate model; 5636 for the seven covariate model and 5464 for the nine covariate model.

For the second scenario, targeting only patients with a mass lesion, only subjects with this were analysed. As the seven and nine covariate models both contained CT only the three covariate model could be fitted. Analysis here was performed on 2974 subjects.

6.2 All subjects – comparing different modelling strategies

For all tables, the table cells are the percentage reductions in sample size, which can be achieved while preserving the power relative to the conventional unadjusted analysis of the dichotomised outcome scale

6.2.1 Comparing strategies - uniform treatment effect

Here the treatment effect followed a proportional odds model.

6.2.1.1 Three covariate model

5% treatment effect

Across all trials the greatest reduction in sample size was under the proportional odds plus covariates strategy, typically around 40%, although all strategies did show an improvement over the conventional modelling strategy. Using the sliding dichotomy with four bands and covariates gave the next largest reduction in sample size as shown in Table 6-1 below.

Table 6-1 All subjects. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, three covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	21.2	20.1	11.7	18.4	27.2	18.3	26.8	30.9	14.6	28.9	23.9	21.2
Proportional Odds: no cov	21.9	22.7	22.7	23.2	24.3	19.0	18.1	24.0	23.7	16.3	22.9	22.7
Proportional Odds: + cov	40.2	44.7	36.3	41.9	45.6	39.2	48.6	53.5	37.2	50.3	45.3	44.7
SD 4 bands, equal splits +cov	24.4	35.7	30.3	33.4	36.3	32.3	36.7	46.3	35.0	40.4	39.1	35.7

8% treatment effect

Table 6-2 below, shows that, as with the 5% treatment effect, the greatest reduction in sample size occurred using the proportional odds plus covariates strategy with the sliding dichotomy plus covariates strategy again giving the next greatest reduction in sample size. In general, both treatment effects, 5% and 8%, showed similar results with all strategies showing an improvement over the conventional dichotomous treatment strategy.

[8] Table cells are the percentage reductions in sample size, which can be achieved while preserving the power relative to the conventional unadjusted analysis of the dichotomised outcome scale. 'SD' means sliding dichotomy. 'Equal splits' means prognostic bands chosen to contain equal numbers of patients. 'p(fav)' means prognostic bands chosen with specific ranges of the predicted probability of a favourable outcome. 'no cov' means that the final analysis was not covariate adjusted. '+cov' means that the final analysis was covariate adjusted.

Table 6-2 All subjects. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, three covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	15.9	17.3	14.4	18.2	20.9	17.6	22.1	26.1	10.2	24.3	22.2	18.2
Proportional Odds: no cov	22.1	24.9	26.9	22.3	17.5	17.0	22.6	20.6	21.2	20.5	22.9	22.1
Proportional Odds: + cov	40.9	42.7	40.2	42.0	41.4	39.9	47.8	48.0	38.2	46.3	44.8	42.0
SD 3 bands, equal splits +cov	27.4	32.8	29.7	28.0	31.4	29.6	38.7	42.4	26.0	39.5	33.9	31.4

6.2.1.2 Seven covariate model

For both the 5% and 8% treatment effects, shown in Table 6-3 and Table 6-4 respectively below, all modelling strategies showed an improvement over the conventional dichotomous modelling strategy. For both treatment effects the proportional odds plus covariates strategy gave the greatest reduction in sample size, of around 40-50% with again the sliding dichotomy plus covariates giving the next largest decrease. For the 5% treatment effect the conventional plus covariates strategy gave a similar result to the proportional odds treatment only strategy. However, for the 8% treatment effect the conventional plus covariates strategy showed a greater reduction in sample size, for most studies, than the proportional odds treatment only strategy.

Table 6-3 All subjects. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, seven covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	20.0	26.8	19.4	25.3	27.3	26.3	29.3	33.6	13.5	28.9	22.7	26.3
Proportional Odds: no cov	20.5	27.0	25.5	23.2	27.1	28.3	21.7	22.9	15.1	17.8	21.0	22.9
Proportional Odds: + cov	44.6	50.9	41.5	46.5	53.2	50.0	55.6	57.4	41.0	48.7	48.1	48.7
SD 4 bands, equal splits +cov	32.3	49.1	26.9	42.8	44.6	40.5	43.6	43.0	30.6	37.8	31.7	40.5

Table 6-4 All subjects. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, seven covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	23.1	29.1	16.7	24.6	26.5	24.1	29.9	32.6	20.5	33.6	24.4	24.6
Proportional Odds: no cov	25.6	22.6	25.2	22.3	19.9	28.1	22.6	23.3	18.3	22.5	25.8	22.6
Proportional Odds: + cov	47.7	51.3	38.6	46.6	50.0	49.1	50.9	55.1	40.4	52.5	53.3	50.0
SD 3 bands, p(fav) splits +cov	47.2	42.3	35.1	36.6	45.8	40.1	39.4	46.1	32.0	44.1	30.7	40.1

6.2.1.3 Nine covariate model

As with the three and seven covariate models, all modelling strategies showed an improvement over the conventional dichotomous treatment split. This was observed over all studies for both the 5% and 8% treatment effects as shown in Table 6-5 and Table 6-6 respectively below. The proportional odds plus covariates strategy gave the greatest reduction in sample size. The conventional plus covariates strategy gave a similar, or better in some cases, reduction in sample size as the proportional odds treatment only strategy.

Table 6-5 All subjects. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, nine covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	25.8	38.6	17.9	22.2	21.0	27.7	26.9	29.7	22.4	31.0	21.9	25.8
Proportional Odds: no cov	28.4	27.1	22.3	22.0	16.0	26.2	25.4	16.6	21.5	14.4	19.0	22
Proportional Odds: + cov	52.0	55.6	43.8	45.0	44.3	48.9	55.1	55.3	43.9	52.7	46.1	48.9
SD 4 bands, equal splits +cov	41.8	50.6	30.5	32.3	32.6	35.6	41.9	47.8	39.2	43.5	30.6	39.2

Table 6-6 All subjects. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, nine covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	19.6	27.9	17.8	26.7	27.5	21.0	32.8	31.3	22.1	29.8	20.6	26.7
Proportional Odds: no cov	21.9	27.9	24.4	25.8	21.6	26.1	26.3	22.0	22.8	21.0	23.7	23.7
Proportional Odds: + cov	44.5	53.5	42.5	50.0	46.1	48.5	55.6	55.1	47.9	52.6	47.6	48.5
SD 3 bands, p(fav) splits +cov	44.0	51.5	39.0	43.3	50.0	44.1	54.2	50.1	34.0	47.2	34.6	44.1

6.2.2 Comparing strategies – mortality treatment effect

The mortality treatment effect was next explored. This assumes that the effect of treatment is to reduce the risk of death or vegetative state but that the relative probability of severe disability to moderate disability to good recovery is unaltered.

6.2.2.1 Three covariate model

For both the 5% and 8% treatment effects for SLIN and the 8% treatment effect for TINT using the sliding dichotomy strategy resulted in an increase in sample size compared to the conventional treatment only strategy. The proportional odds plus covariates strategy resulted in the largest reduction in sample size for all studies. The sliding dichotomy plus covariates strategy resulted in the next largest reduction in sample size for most studies. The proportional odds treatment only strategy showed a greater reduction in sample size for all studies, except TIUS, than the conventional plus covariates strategy. Using covariates with the conventional dichotomy and with the proportional odds model gave greater reductions than the strategies without covariates as shown in Table 6-7 and Table 6-8 below.

Table 6-7 All subjects. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, three covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	18.5	21.9	13.5	18.6	23.7	20.0	25.6	22.5	19.1	24.3	21.3	21.3
Proportional Odds: no cov	20.6	15.5	15.3	32.7	35.4	35.4	42.1	37.2	38.7	35.4	37.1	35.4
Proportional Odds: + cov	42.0	44.2	31.8	45.6	50.8	50.2	58.9	54.7	49.2	53.2	50.3	50.2
SD 3 bands, equal splits +cov	18.8	23.6	-52.4	35.6	38.0	36.6	44.7	37.9	36.6	40.5	40.5	36.6

Table 6-8 All subjects. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, three covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	13.8	18.1	11.3	14.3	23.3	18.1	19.0	27.0	13.2	23.8	18.8	18.1
Proportional Odds: no cov	18.6	17.0	13.5	28.8	36.7	36.7	37.2	35.2	32.6	35.7	34.2	34.2
Proportional Odds: + cov	38.9	42.2	29.3	44.5	56.0	52.2	52.4	58.1	46.7	52.7	49.6	49.6
SD 3 bands, equal splits +cov	-22.3	5.5	-61.8	32.9	43.0	38.0	37.7	42.5	34.0	39.9	37.5	37.5

6.2.2.2 Seven covariate model

For all studies all strategies showed an improvement over the conventional treatment only strategy. As with the three covariate model the greatest reductions in sample size were observed using the proportional odds plus covariates strategy. Similar results were observed both with the 5% and 8% treatment effects as shown in Table 6-9 and Table 6-10 below.

Table 6-9 All subjects. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, seven covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	24.4	26.5	15.6	17.4	22.0	20.8	25.3	31.8	24.9	30.9	21.5	24.4
Proportional Odds: no cov	21.4	30.0	29.2	28.7	31.4	31.2	40.0	40.3	46.9	35.4	32.0	31.4
Proportional Odds: + cov	48.5	50.1	41.1	47.1	51.6	51.5	58.8	62.9	58.6	57.6	51.3	51.5
SD 3 bands, equal splits +cov	33.6	34.2	24.4	37.9	41.3	42.9	39.8	47.9	43.9	42.5	38.9	39.8

Table 6-10 All subjects. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, seven covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	19.3	22.4	17.7	16.6	19.2	19.9	20.2	25.9	22.2	30.9	22.7	20.2
Proportional Odds: no cov	21.9	24.9	29.4	30.3	35.3	32.1	42.0	41.3	43.7	37.1	34.0	34.0

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Proportional Odds: + cov	44.8	52.2	46.7	51.1	53.1	50.7	57.8	63.6	58.1	61.4	53.7	53.1
SD 4 bands, equal splits +cov	25.5	33.5	28.5	23.0	27.4	37.2	50.9	55.8	41.9	39.8	32.2	33.5

6.2.2.3 Nine covariate model

All strategies showed an improvement over the conventional treatment only strategy for all studies as shown in Table 6-11 and Table 6-12 below. As with the three and seven covariate models, using the proportional odds plus covariates strategy gave the greatest reductions in sample size. The results from all three covariate models showed a similar pattern with, typically, slightly greater magnitudes of reduction going from the three to seven to nine covariate models.

Table 6-11 All subjects. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, nine covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	21.3	30.3	16.1	22.9	21.1	20.7	16.9	22.5	19.8	29.5	22.3	21.3
Proportional Odds: no cov	31.6	28.6	31.3	32.4	34.3	27.9	33.5	34.7	44.6	37.7	34.7	33.5
Proportional Odds: + cov	51.0	52.3	43.2	53.3	56.3	52.5	54.9	57.3	59.6	59.5	53.7	53.7
SD 3 bands, equal splits +cov	37.5	29.1	23.3	42.8	43.3	42.9	36.6	38.1	42.8	43.8	42.5	42.5

Table 6-12 All subjects. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, nine covariate model[8]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	19.9	23.5	15.3	20.4	24.0	21.1	19.9	27.4	22.7	28.6	22.7	22.7
Proportional Odds: no cov	26.0	20.7	30.6	30.9	34.8	31.2	38.0	41.0	49.5	42.8	35.5	34.8
Proportional Odds: + cov	47.1	51.5	43.4	51.2	55.8	50.6	56.5	58.5	58.8	62.6	53.0	53.0
SD 3 bands, equal splits +cov	22.8	28.7	2.7	43.2	46.4	38.3	40.6	42.8	42.6	50.2	40.1	40.6

6.2.3 Graphical comparison – all subjects

The tabulated results showed that the proportional odds model and the sliding dichotomy model with covariates gave the greatest reductions in sample sizes when the treatment effect followed a proportional odds model and when the treatment effect is a reduction in mortality. It is of interest to examine these patterns graphically looking across all of the covariate models together. In the graphs the trials are presented side by side as are the three observational series.

Figure 6-1 and Figure 6-2 below show the results by trial for the 5% and 8% treatment effect respectively when the treatment effect followed a proportional odds model (the Uniform treatment comparison). No clear pattern is observed looking at the sliding dichotomy with the 5% treatment effect. The 8% treatment effect does show, for all studies except HITII, a greater reduction in sample size with the seven and nine covariate models compared with the three covariate models for the sliding dichotomy strategies.

With the proportional odds strategies, using seven or nine covariates almost always shows a greater reduction than with using three covariates both for the 5% and 8% treatment effects. Also, using seven and nine covariates with the proportional odds strategies gave similar or greater reductions in sample size than the sliding dichotomy with three, seven or nine covariates.

Figure 6-3 and Figure 6-4 show the results by trial of the sliding dichotomy and proportional odds models for all three covariate sets when the treatment effect was a reduction in mortality (the Mortality treatment comparison). Again no clear patterns were observed for the sliding dichotomy strategies. Using seven and nine covariates with the proportional odds model typically resulted in greater reductions than using the three covariate model.

Figure 6-5 shows the median reductions in sample size over all trials for the four treatment effects (Uniform 5% and 8% and Mortality 5% and 8%). These show very consistent results both when the treatment effect followed a proportional odds model and when it did not. For almost all models adding covariates gave a small additional benefit. This gave in some cases sample size reductions of just over 50% compared with the conventional dichotomy.

Figure 6-1 Uniform 5%: comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]

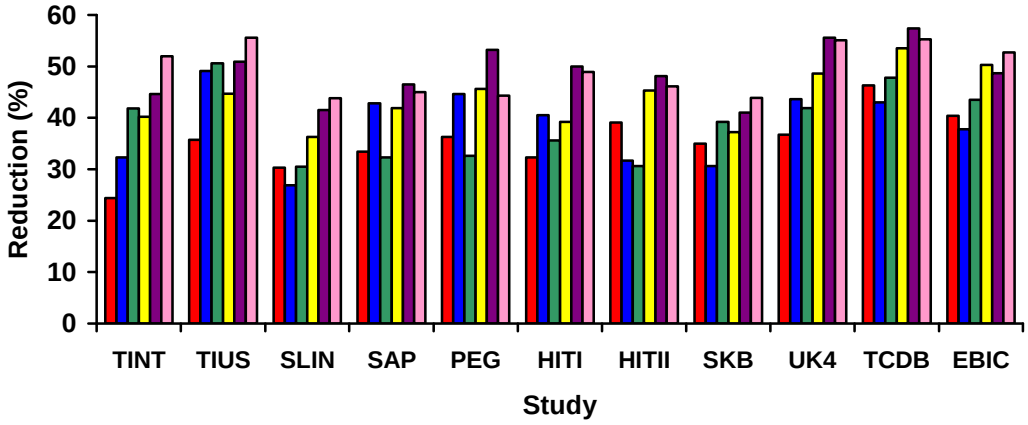
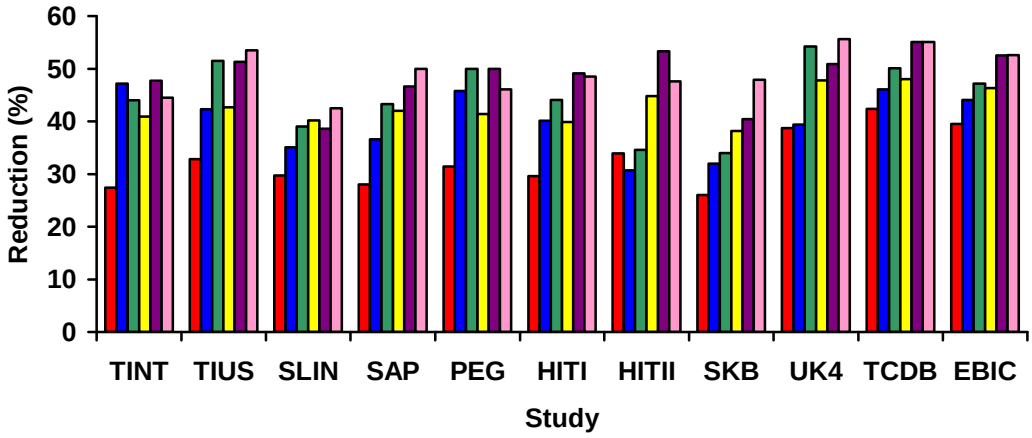


Figure 6-2 Uniform 8%: comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red), Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

[9] Bars are the percentage reductions in sample size which can be achieved while preserving the power relative to the conventional unadjusted analysis of the dichotomised outcome scale.

Figure 6-3 Mortality 5%: comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]

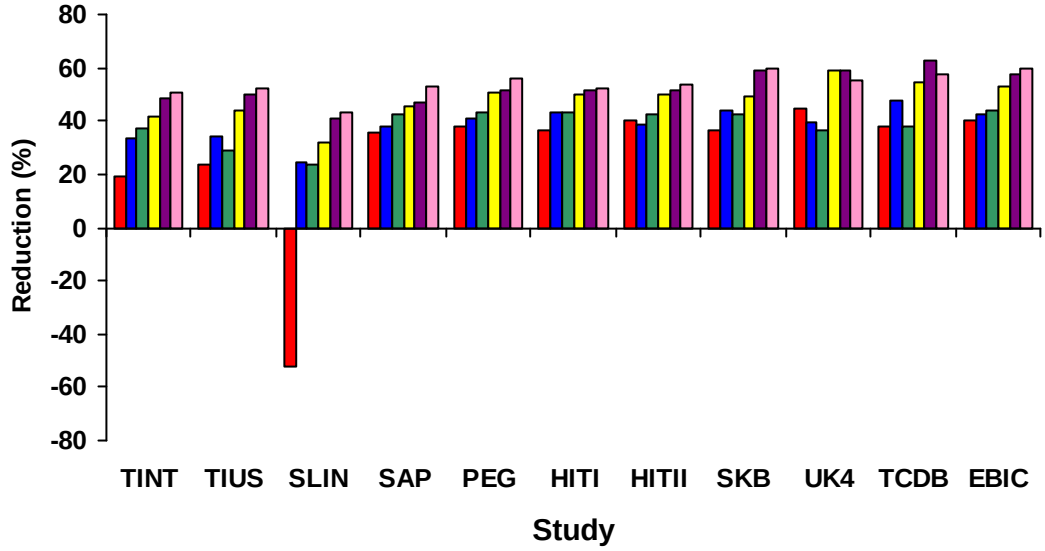
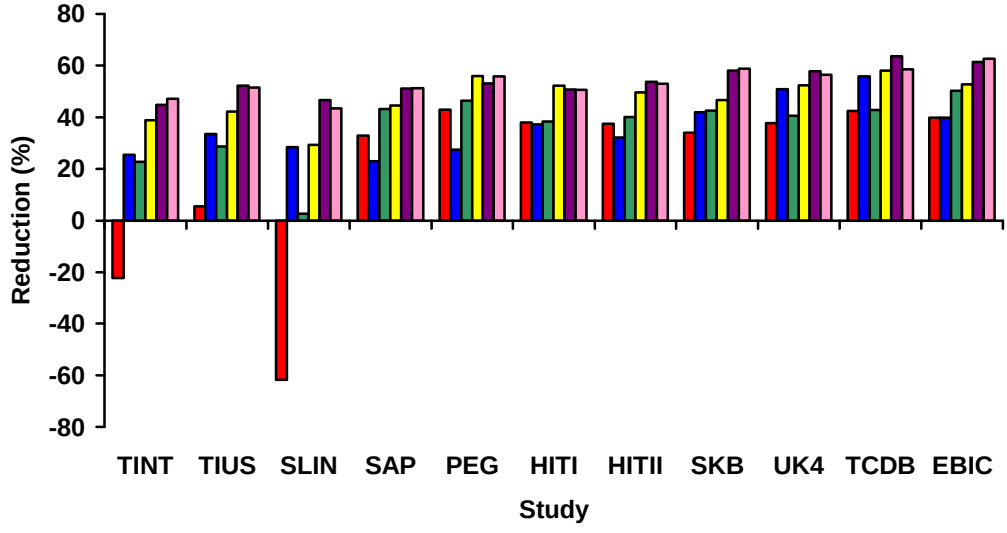
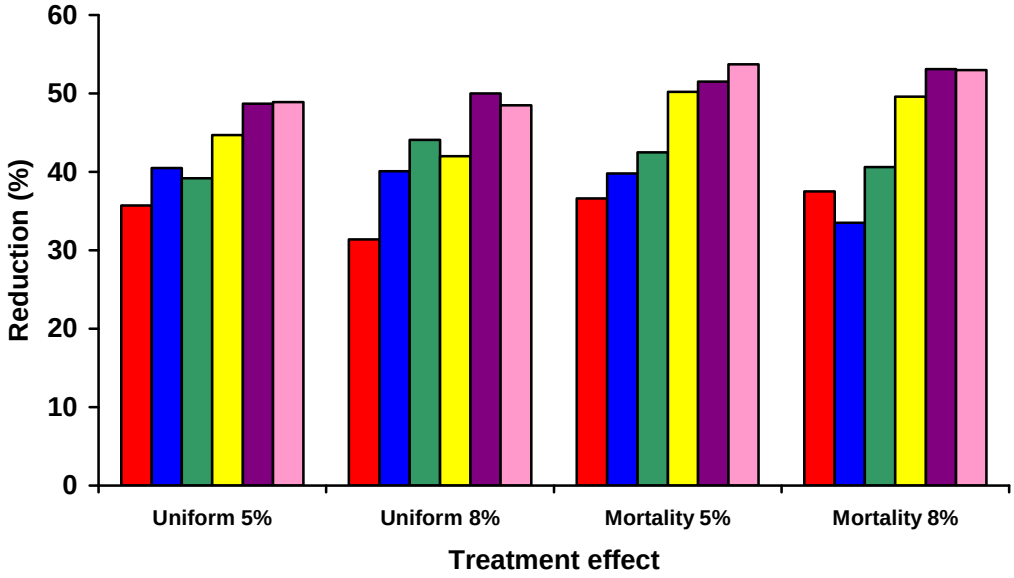


Figure 6-4 Mortality 8%: comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red), Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

Figure 6-5 Median reductions in sample size for all subjects. Comparison of Sliding Dichotomy and Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red), Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

6.2.4 Discussion

For both the 5% and 8% treatment effects and the three, seven and nine covariate models using the proportional odds model with covariates gave the greatest reductions in sample size. This was true both when the data followed a proportional odds model (the Uniform treatment scenario) and when they deviated strongly from a proportional odds model (the Mortality treatment scenario).

6.3 Improvement restricted to patients with an intermediate prognosis

Having shown that using the proportional odds model with covariates gave the greatest sample size reduction for all scenarios, treatment effects and covariates, it is of interest to see if this pattern is repeated when improvement, i.e. the treatment effect, is restricted to subjects with an intermediate prognosis. Here for all comparisons the sliding dichotomy equal splits three bands plus covariates was used as the sliding dichotomy strategy.

6.3.1 Comparing strategies - uniform treatment effect

Here the simulated treatment effect followed a proportional odds model.

6.3.1.1 Three covariate model

For both the 5% and 8% treatment effects, the proportional odds plus covariates strategy gave the greatest reduction in sample size when compared to the conventional treatment only strategy. However, all alternative strategies also showed a reduction in sample size compared with the conventional treatment only strategy.

For the 5% treatment effect, as shown in Table 6-13 below a consistent pattern of results was not observed across all studies. TCDB and EBIC both showed very modest reductions in sample size for the proportional odds treatment only strategy compared with the conventional treatment only strategy. No consistent benefit was observed with the proportional odds treatment only strategy over the conventional plus covariates strategy. Some studies showed a benefit, others did not.

For the 8% treatment effect, for all studies except TINT, the sliding dichotomy strategy gave a greater reduction in sample size for eight of the eleven studies than the conventional plus covariates strategy as shown in Table 6-14 below. In general, the reductions in sample size observed with the 8% treatment effect were greater than with the 5% treatment effect. Using the proportional odds plus covariates gave much greater sample size reductions, typically at least double, than using the proportional odds treatment only strategy.

Table 6-13 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	18.8	13.7	11.0	12.5	22.8	13.7	31.1	19.5	17.6	23.6	25.4	18.8
Proportional Odds: no cov	20.3	18.3	21.6	15.4	15.6	19.0	17.6	7.2	15.7	5.7	18.2	17.6

[10] Table cells are the percentage reductions in sample size, which can be achieved while preserving the power relative to the conventional unadjusted analysis of the dichotomised outcome scale. 'SD' means sliding dichotomy. 'Equal splits' means prognostic bands chosen to contain equal numbers of patients. 'no cov' means that the final analysis was not covariate adjusted. '+cov' means that the final analysis was covariate adjusted.

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Proportional Odds: + cov	37.2	28.6	34.5	32.8	40.8	37.4	45.2	39.8	35.7	38.4	46.9	37.4
SD 3 bands, equal splits +cov	12.5	13.7	15.5	16.8	27.8	24.3	32.5	25.5	27.6	24.6	29.7	24.6

Table 6-14 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	17.5	18.6	12.9	18.6	22.8	20.3	25.8	25.5	15.2	26.5	18.2	18.6
Proportional Odds: no cov	16.6	21.9	18.9	21.2	20.2	24.9	14.1	13.7	19.2	10.2	20.0	19.2
Proportional Odds: + cov	34.6	40.4	33.7	39.7	42.0	42.1	46.9	42.2	34.1	43.3	39.4	40.4
SD 3 bands, equal splits +cov	14.8	23.4	22.8	21.8	27.6	28.1	31.4	32.5	22.8	26.9	23.7	23.7

6.3.1.2 Seven covariate model

For all studies for both the 5% and 8% treatment effects the proportional odds model plus covariates strategy gave the greatest reduction in sample size compared with the conventional treatment only strategy.

5% treatment effect

Almost all alternative strategies showed an improvement over the conventional treatment only strategy for the 5% treatment effect as shown in Table 6-15 below.

For EBIC using the proportional odds treatment only strategy was no better than using the conventional treatment only strategy. For nine of the eleven studies the conventional plus covariates strategy resulted in a greater reduction in sample size compared with the proportional odds treatment only strategy. For seven of the studies the sliding dichotomy strategy gave a smaller reduction in sample size than the conventional plus covariates strategy.

Table 6-15 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, seven covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	15.7	28.1	17.1	23.7	25.4	20.8	29.3	27.7	22.3	30.7	23.6	23.7
Proportional Odds: no cov	20.4	21.1	22.3	17.6	11.9	16.9	14.9	12.3	11.1	0.0	15.2	15.2
Proportional Odds: + cov	32.4	46.3	37.8	36.9	41.6	40.9	45.1	45.9	36.5	45.0	38.5	40.9
SD 3 bands, equal splits +cov	3.7	23.9	17.8	19.1	24.9	24.2	31.1	22.9	25.0	29.9	17.9	23.9

8% treatment effect

For all studies, with the 8% treatment effect all alternative strategies showed an improvement over the conventional treatment only strategy as shown in Table 6-16 below. The conventional plus covariates strategy gave a greater reduction in sample size than the sliding dichotomy strategy for eight of the eleven studies. As with the 5% treatment effect, the results were not consistent from study to study. The proportional odds treatment only strategy gave a smaller reduction in sample size than the conventional plus covariates strategy for all studies except SLIN. A similar

magnitude of reductions in sample size was observed with the 8% treatment effect as with the 5% treatment effect.

Table 6-16 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, seven covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	18.0	29.8	16.2	23.0	23.3	25.6	28.0	30.3	23.2	31.4	19.3	23.3
Proportional Odds: no cov	13.5	20.9	24.1	17.8	14.1	19.0	13.5	11.5	12.5	16.4	17.7	16.4
Proportional Odds: + cov	33.8	45.6	38.5	43.4	42.2	42.7	44.0	44.3	41.7	46.3	39.9	42.7
SD 3 bands, equal splits +cov	10.2	28.2	14.6	17.6	24.6	19.0	29.3	26.7	26.0	29.8	14.6	24.6

6.3.1.3 Nine covariate model

5% and 8% treatment effect

As with the seven covariate model, all alternative strategies showed an improvement over the conventional treatment only strategy for all studies. Using the proportional odds plus covariates gave the greatest reduction in sample size.

5% treatment effect

For all studies, except SKB, a greater reduction in sample size was observed using the conventional plus covariates strategy compared with the sliding dichotomy strategy as shown in Table 6-17 below. Using the covariates with the proportional

odds model typically at least doubled the reduction in sample size compared with using the proportional odds model alone.

Table 6-17 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, nine covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	22.3	29.7	15.3	18.2	26.5	23.1	27.5	29.2	16.5	30.4	24.0	24.0
Proportional Odds: no cov	18.1	17.7	23.9	15.8	6.0	15.0	19.9	5.1	18.9	10.7	14.5	15.8
Proportional Odds: + cov	39.5	42.2	38.6	31.5	41.4	39.9	46.5	39.1	37.3	44.4	41.1	39.9
SD 3 bands, equal splits +cov	17.2	18.0	14.6	9.4	21.0	12.8	21.4	24.7	26.4	23.2	16.7	18.0

8% treatment effect

A similar magnitude of reductions in sample size was observed with the 8% treatment effect as with the 5% treatment effect as shown in Table 6-18 below. The conventional plus covariates strategy gave a greater reduction in sample size than the sliding dichotomy strategy for all studies. The proportional odds treatment only strategy gave a smaller reduction in sample size than the conventional plus covariates strategy for nine of the studies; only for TINT and SLIN did this strategy give a greater reduction in sample size.

Table 6-18 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, nine covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	19.3	27.9	19.6	25.8	23.5	22.6	26.5	31.9	23.9	31.3	22.3	23.9
Proportional Odds: no cov	19.7	17.0	23.6	19.3	11.3	16.1	9.6	10.2	11.3	9.0	14.1	14.1
Proportional Odds: + cov	37.6	44.6	40.9	42.6	38.3	39.9	40.6	45.1	38.0	43.7	39.7	40.6
SD 3 bands, equal splits +cov	8.2	20.7	12.0	18.1	19.5	15.9	17.7	30.8	17.6	23.2	15.9	17.7

All covariate models

For most studies the reduction in sample size using the proportional odds plus covariates strategy was similar for the three, seven and nine covariate models. The exception to this being EBIC where the addition of more covariates did increase the reduction in sample size. The proportional odds treatment only strategy gave a modest reduction in sample size which was usually less than with the conventional plus covariates strategy.

6.3.2 Comparing strategies – mortality treatment effect

Here the simulated treatment effect was a reduction in mortality.

6.3.2.1 Three covariate model

5% and 8% treatment effects

The proportional odds plus covariates strategy gave the greatest reduction in sample size for all studies. Greater reductions in sample size were observed with the three covariate mortality scenario than with the three covariate uniform scenario.

5% treatment effect

Here consistent results were not observed over all studies as shown in Table 6-19 below. The proportional odds treatment only strategy gave a greater reduction in sample size than either the conventional plus covariates or the sliding dichotomy strategy for most studies.

Table 6-19 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	14.9	24.7	16.5	14.4	24.1	18.6	19.7	22.2	16.0	29.8	19.4	19.4
Proportional Odds: no cov	25.2	23.8	12.6	26.4	30.7	32.3	35.7	35.0	30.2	32.3	30.9	30.7
Proportional Odds: + cov	44.4	46.0	36.4	45.2	47.4	47.6	49.0	50.0	45.7	50.0	47.0	47
SD 3 bands, equal splits +cov	9.0	17.0	-39.1	33.5	25.0	29.3	23.2	28.3	28.0	25.0	29.0	25

8% treatment effect

Both TINT and SLIN showed an increase in sample size using the sliding dichotomy strategy compared to the conventional treatment only strategy as shown in Table 6-20 below. Slightly greater reductions in sample size were observed with the 8% compared to the 5% treatment effect.

Table 6-20 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	12.7	17.1	8.3	21.1	19.5	17.0	24.8	25.2	17.2	27.0	23.8	19.5
Proportional Odds: no cov	13.1	16.4	12.6	31.2	32.5	34.3	36.5	35.7	34.5	30.7	32.9	32.5
Proportional Odds: + cov	33.5	39.9	29.8	45.8	49.1	48.8	54.1	50.3	50.5	53.7	48.4	48.8
SD 3 bands, equal splits +cov	-25.5	4.6	-62.6	31.2	28.1	29.9	25.7	27.4	36.4	25.5	29.3	27.4

6.3.2.2 Seven covariate model

For both the 5% and 8% treatment effects the greatest reduction in sample size is observed using the proportional odds plus covariates strategy for all studies.

5% treatment effect

All alternative strategies showed an improvement over using the conventional treatment only strategy as shown in Table 6-21 below. For six of the eleven studies

the sliding dichotomy strategy gave a greater reduction in sample size than the conventional plus covariates strategy. The addition of the covariates to either the conventional or the proportional odds models had the greatest effect for the three observational studies UK4, TCDB and EBIC.

Table 6-21 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, seven covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	25.2	28.5	19.5	16.6	24.5	27.7	30.2	34.3	14.5	31.2	27.2	27.2
Proportional Odds: no cov	25.4	26.0	30.0	29.2	32.4	27.9	35.1	36.0	37.3	27.1	26.6	29.2
Proportional Odds: + cov	47.4	49.1	44.2	46.4	51.2	49.7	52.5	54.5	50.1	53.3	50.6	50.1
SD 3 bands, equal splits +cov	31.4	16.9	20.6	31.1	29.9	35.8	19.8	16.0	16.1	20.4	33.1	20.6

8% treatment effect

As with the three covariate model, using the sliding dichotomy strategy with SLIN gave an increase in sample size compared with using the conventional treatment only strategy. Almost all of the studies showed a decrease in reduction in sample size using the sliding dichotomy compared with the 5% treatment effect. For all studies, except TIUS, the proportional odds treatment only strategy showed the second greatest reduction in sample sizes compared with the proportional odds plus covariates strategy as shown in Table 6-22 below.

Table 6-22 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, seven covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	16.5	27.1	14.9	20.2	24.0	23.5	28.4	27.5	18.6	30.0	23.6	23.6
Proportional Odds: no cov	20.3	21.5	25.1	30.9	34.8	29.5	36.3	34.7	39.1	32.7	31.9	31.9
Proportional Odds: + cov	42.5	49.0	43.0	48.0	52.1	47.8	55.2	53.2	50.2	52.8	49.1	49.1
SD 3 bands, equal splits +cov	7.4	18.3	-6.6	30.3	33.5	26.8	18.4	13.4	18.4	23.6	29.0	18.4

6.3.2.3 Nine covariate model

For both the 5% and 8% treatment effects the proportional odds plus covariates strategy gave the greatest reduction in sample size compared with the conventional treatment only strategy for all studies.

5% treatment effect

In general, the proportional odds treatment only strategy gave the second greatest reduction in sample size as shown in Table 6-23 below. For all studies, except TIUS, a greater reduction in sample size was observed using the proportional odds treatment only strategy than the conventional plus covariates strategy. For TIUS the reductions in sample size with both the proportional odds treatment only and the conventional plus covariates strategies were very similar although the conventional plus covariates gave a slightly greater reduction. For both SAPHIR and PEGSOD

the sliding dichotomy strategy gave a greater reduction in sample size than the proportional odds treatment only strategy.

Table 6-23 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, nine covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	27.8	22.3	14.2	21.5	19.7	23.9	25.2	30.3	28.1	21.9	24.1	23.9
Proportional Odds: no cov	30.4	21.1	27.0	29.9	25.5	27.9	37.5	32.1	44.8	32.6	32.1	30.4
Proportional Odds: + cov	50.0	46.0	37.8	53.3	47.8	46.6	53.9	52.9	57.5	51.1	48.8	50.0
SD 3 bands, equal splits +cov	28.1	17.3	11.6	34.4	25.7	33.2	18.0	15.4	28.1	13.1	30.1	25.7

8% treatment effect

Similar reductions in sample size were observed using the proportional odds plus covariates strategy with the 8% treatment effect as with the 5% treatment effect, as shown in Table 6-24 below. For SLIN, using the sliding dichotomy strategy increased the sample size compared with using the conventional treatment only strategy. For most of the studies, the proportional odds treatment only strategy gave the second greatest reductions in sample size compared with the proportional odds plus covariates strategy.

Table 6-24 Improvement restricted to patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, nine covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	20.4	26.8	18.0	22.2	27.0	28.1	25.7	33.2	22.1	26.3	24.3	25.7
Proportional Odds: no cov	22.0	21.4	24.5	30.1	32.8	31.5	34.8	34.7	42.5	29.3	33.8	31.5
Proportional Odds: + cov	46.1	48.7	42.6	50.9	53.2	52.2	53.8	56.1	54.3	50.6	51.1	51.1
SD 3 bands, equal splits +cov	3.8	15.5	-12.8	27.8	34.6	29.5	23.7	18.7	23.8	17.2	30.4	23.7

Discussion

In general, there was a greater reduction in sample size when using a greater number of covariates. For SLIN, the sliding dichotomy was a poor strategy to use as detailed previously.

General comments for both uniform and mortality scenarios

The addition of the covariates has a much larger effect under the uniform scenario compared with the mortality scenario.

The studies which are unselected series show the greatest reductions in sample size although very large reductions in sample size are observed for all studies.

The studies which are small trials show the least benefit from the alternative strategies however the studies which are small unselected series, e.g. TCDB, do show a larger benefit from using the alternative modelling strategies.

6.3.3 Graphical comparison – improvement restricted to patients with an intermediate prognosis

When the treatment effect followed a proportional odds model using both the sliding dichotomy and proportional odds strategies gave a reduction in sample size for all studies as shown in Figure 6-6 and Figure 6-7. Similar effects were observed with the 5% and 8% treatment effects. For both the 5% and the 8% treatment effects using the proportional odds strategies consistently gave greater reductions in sample size than the sliding dichotomy strategies for all studies. No consistent pattern was observed with increasing numbers of covariates. For example, using a model with nine covariates rarely gives a greater sample size reduction than using a model with seven covariates.

For all studies the proportional odds strategies gave a reduction in sample size even when the treatment effect modelled was a reduction in mortality as shown in Figure 6-8 and Figure 6-9 below. Using the sliding dichotomy strategies gave reductions for all studies except SLIN and TINT (although here only for three covariates with the 8% treatment effect). For both the 5% and 8% treatment effects using the proportional odds strategies gave greater reductions than the sliding dichotomy strategies. Slightly greater reductions were observed with the 8% treatment effect compared to the 5% treatment effect, although, for both, reductions were still around 50% compared with using the conventional dichotomy.

These results are summarised in Figure 6-10. This shows the median reduction in sample size when the treatment effect followed both a proportional odds model and was a reduction in mortality. A consistent pattern is observed over all treatment effects with the three covariate model giving reductions of around 25% and the seven

and nine covariate models giving reductions of 40 to 50% when improvement is restricted to those with an intermediate prognosis.

Figure 6-6 Uniform 5%: Restricting improvement to those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]

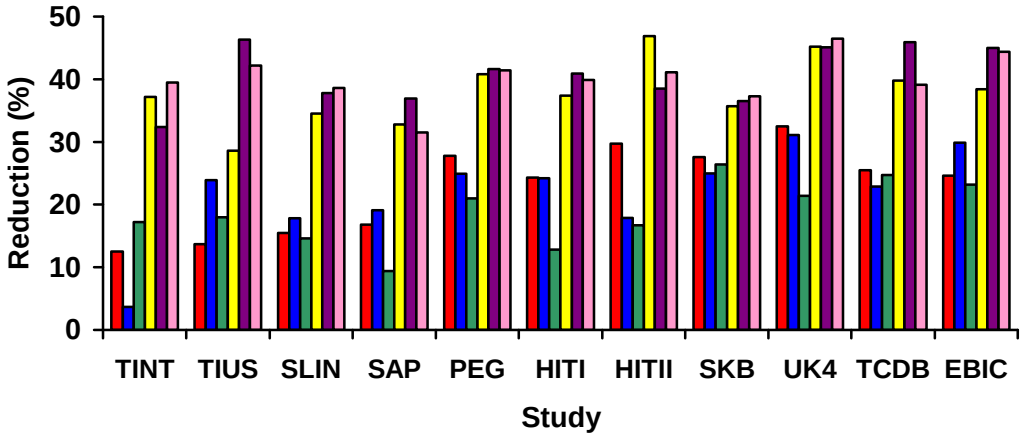
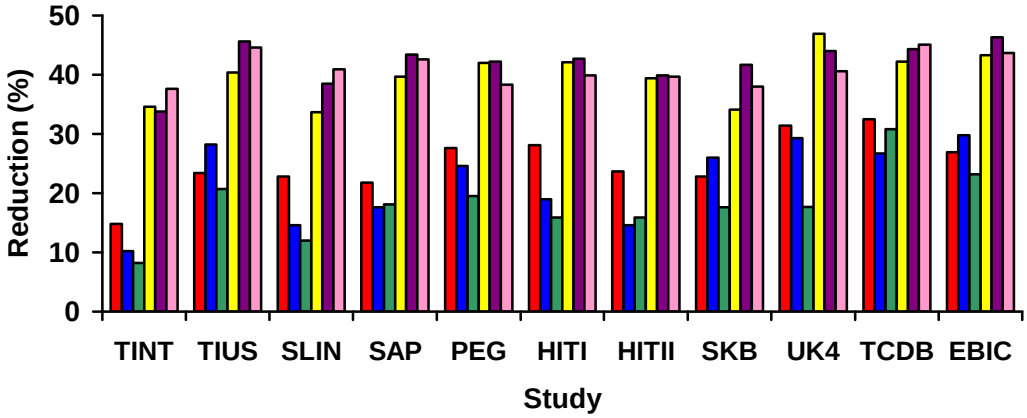


Figure 6-7 Uniform 8%: Restricting improvement to those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red) Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

Figure 6-8 Mortality 5%: Restricting improvement to those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]

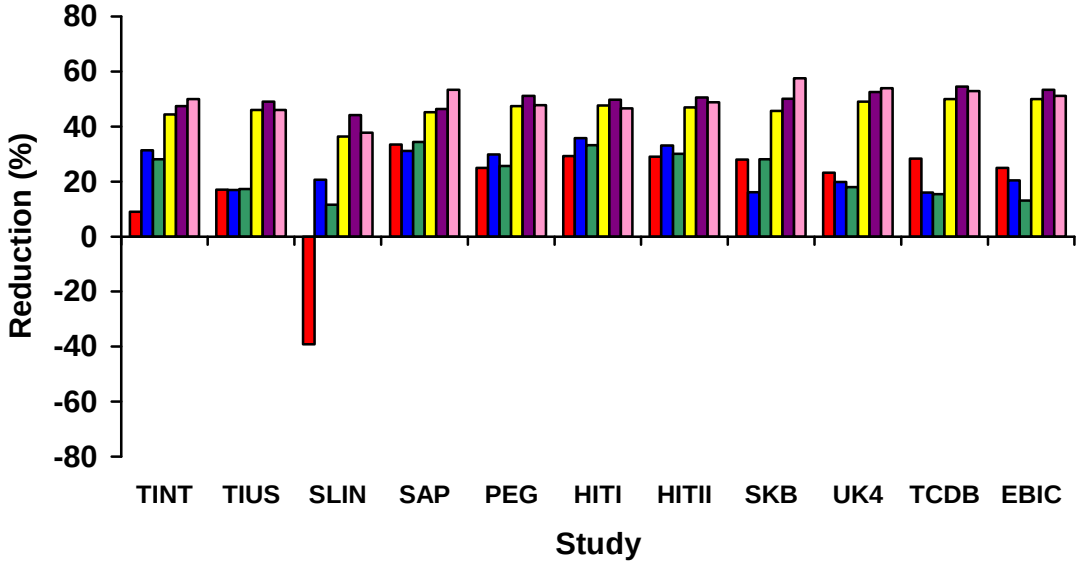
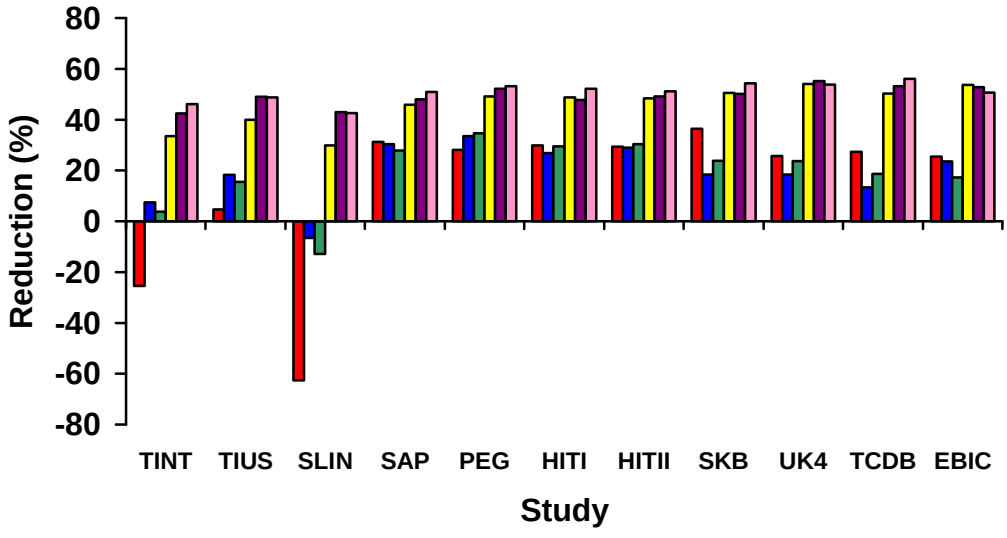
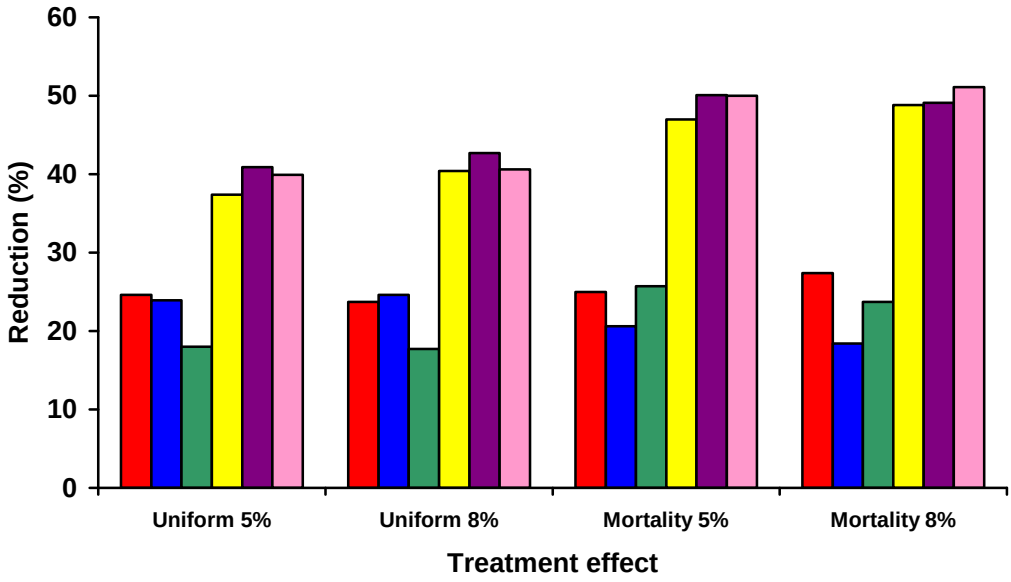


Figure 6-9 Mortality 8%: Restricting improvement to those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red) Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

Figure 6-10 Median reductions in sample size. Restricting improvement to those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red) Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

6.4 Improvement restricted to patients with a mass lesion

With the scenario of restricting improvement to those with a mass lesion, only the three covariate model was used as both the seven and nine covariate models contain CT.

6.4.1 Comparing strategies - uniform treatment effect

Here strategies were compared assuming that the treatment effect followed a proportional odds model.

5% and 8% treatment effects

For all studies, all alternative strategies show a reduction in sample size compared with the conventional treatment only strategy. The greatest reduction in sample size was observed using the proportional odds plus covariates strategy. The reductions seen with the sliding dichotomy were greatest under this scenario, when improvement was restricted to patients with a mass lesion, compared with both analyses restricted to those with a favourable outcome and analysis on all subjects.

5% treatment effect, shown in Table 6-25 below.

The sliding dichotomy strategy gave the second largest reduction in sample size, compared with the proportional odds plus covariates strategy for eight studies, TIUS, SLIN, PEG, HIT I, UK4, TCDB, SKB and EBIC. The unselected series showed the greatest benefit from using the sliding dichotomy strategy.

Table 6-25 Improvement restricted to patients with a mass lesion. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
Conventional + cov	9.5	20.2	18.3	8.3	21.1	13.8	25.0	20.1	23.0	13.9	9.5	18.3
Proportional Odds: no cov	22.0	26.4	30.0	23.1	27.0	23.0	23.4	24.0	18.4	15.8	22.0	23.1
Proportional Odds: + cov	37.3	48.0	43.0	33.6	44.1	40.2	46.9	49.4	46.9	36.3	37.3	43.0
SD 3 bands, equal splits +cov	18.5	35.2	35.7	15.8	39.0	35.1	41.8	46.1	37.9	23.7	18.5	35.2

8% treatment effect

For some studies the sliding dichotomy strategy with the 5% treatment effect gave a greater reduction in sample size than the sliding dichotomy with the 8% treatment effect as shown in Table 6-26 below. In general, the sliding dichotomy strategy gave similar or greater reductions in sample size to the proportional odds treatment only strategy.

Table 6-26 Improvement restricted to patients with a mass lesion. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	15.7	20.6	14.7	15.0	17.8	14.5	19.9	19.4	17.1	13.9	15.7	15.7

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Proportional Odds: no cov	27.0	33.4	35.2	23.6	23.0	17.7	21.8	21.6	20.9	22.6	27.0	23.0
Proportional Odds: + cov	38.5	48.8	46.0	40.5	42.2	38.1	42.6	45.3	48.4	37.7	38.5	42.2
SD 3 bands, equal splits +cov	24.5	34.6	37.6	21.4	26.2	27.8	34.2	36.7	37.3	23.2	24.5	27.8

6.4.2 Comparing strategies – mortality treatment effect

Here the treatment effect was a reduction in mortality.

5% treatment effect

In general, the sliding dichotomy strategy gave a greater reduction in sample size than the proportional odds treatment only strategy as shown in Table 6-27 below. As with the uniform three covariate model, using the sliding dichotomy with SLIN resulted in an increase in sample size compared with the conventional treatment only strategy.

Table 6-27 Improvement restricted to patients with a mass lesion. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	11.8	24.0	13.7	17.5	14.7	12.9	18.8	26.6	17.1	19.8	11.8	17.1
Proportional Odds: no cov	27.4	24.0	15.5	30.8	46.7	41.4	41.9	44.4	39.7	33.8	27.4	33.8

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Proportional Odds: + cov	44.2	51.1	35.2	50.9	60.0	59.1	57.1	63.7	56.0	55.8	44.2	55.8
SD 3 bands, equal splits +cov	25.2	39.6	-41.0	48.6	57.7	57.7	51.9	57.8	48.7	51.1	25.2	48.7

8% treatment effect

Using the sliding dichotomy strategy with TINT, SLIN or HITII resulted in an increase in sample size compared with the conventional treatment only strategy as shown in Table 6-28 below. Both HIT I and TCDB show a slightly greater reduction in sample size using the sliding dichotomy strategy compared with the proportional odds plus covariates strategy. All other studies show the greatest reduction in sample size using the proportional odds plus covariates strategy. The reduction in sample size increased from the 5% to 8% treatment effect using the sliding dichotomy strategy. The reductions in sample size have also very slightly increased for the proportional odds strategy. For both the 5% and 8% treatment effects the addition of the covariates to the proportional odds model gave a large effect.

Table 6-28 Improvement restricted to patients with a mass lesion. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	13.2	16.8	8.9	10.7	21.0	12.1	14.8	14.3	17.4	15.3	13.2	14.3
Proportional Odds: no cov	21.4	23.5	15.0	35.2	43.2	43.7	41.8	44.3	36.2	35.2	21.4	35.2

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Proportional Odds: + cov	45.6	51.2	36.6	52.9	61.1	59.4	59.6	59.9	54.6	52.4	45.6	52.9
SD 3 bands, equal splits +cov	-20.0	26.5	-53.5	48.1	60.1	60.2	57.1	62.1	.	48.6	-20.0	48.4

6.5 Targeting only patients with an intermediate prognosis

Here only subjects who had an intermediate prognosis of a favourable outcome were analysed.

6.5.1 Comparing strategies - uniform treatment effect

Here the treatment effect follows a proportional odds model.

6.5.1.1 Three covariate model

Both 5% and 8% treatment effects

For all studies, all alternative strategies gave a greater reduction in sample size than the conventional treatment only strategy. These results are shown in Table 6-29 and Table 6-30 below. Using the proportional odds plus covariates strategy gave the greatest reduction in sample size and the conventional plus covariates the least reduction in sample size when compared with the conventional treatment only strategy. A very similar pattern was observed with both the 5% and 8% treatment effects. The reductions in sample size observed when targeting only patients with an intermediate prognosis were smaller than when improvement was restricted to those with an intermediate prognosis above – possibly because of the smaller sample pool from which to sample subjects.

Table 6-29 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	15.3	18.5	7.9	11.1	10.9	12.2	11.9	13.0	8.7	10.4	7.0	11.1
Proportional Odds: no cov	24.7	24.3	25.1	17.6	20.4	27.6	21.5	18.8	19.8	17.7	19.3	20.4
Proportional Odds: + cov	36.3	38.3	34.2	32.9	39.3	35.0	38.8	36.3	29.8	32.7	29.7	35.0
SD 3 bands, equal splits +cov	18.5	19.7	18.0	12.1	28.1	23.2	30.1	23.4	22.1	21.6	10.7	21.6

Table 6-30 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	10.9	10.4	5.9	11.3	10.8	10.1	11.3	15.1	6.0	10.9	9.8	10.8
Proportional Odds: no cov	21.3	25.1	25.0	22.3	20.4	21.8	23.6	20.9	17.3	27.3	24.8	22.3
Proportional Odds: + cov	34.9	35.9	32.9	36.8	34.5	29.3	36.7	40.7	27.9	38.2	34.6	34.9
SD 3 bands, equal splits +cov	19.6	22.2	16.5	19.1	22.6	18.5	27.1	29.7	18.7	25.5	18.1	19.6

6.5.1.2 Seven covariate model

Both 5% and 8% treatment effects

All alternative strategies showed an improvement over the conventional treatment only strategy. For all studies, as with the three covariate model, the proportional odds plus covariates strategy gives the greatest reduction in sample size compared with the conventional treatment only strategy. Using the conventional plus covariate strategy gave the smallest reductions in sample size. For nine of the studies, using the proportional odds treatment only strategy gave a greater reduction in sample size than using the sliding dichotomy strategy. Similar reductions in sample size were observed with both the 5% and the 8% treatment effects as shown in Table 6-31 and Table 6-32 respectively below.

Table 6-31 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, seven covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	10.9	17.0	11.2	9.5	8.1	13.4	14.3	15.4	7.8	10.2	10.6	10.9
Proportional Odds: no cov	28.1	24.2	27.7	22.0	19.1	16.9	20.3	18.5	18.6	17.7	21.6	20.3
Proportional Odds: + cov	38.2	37.4	37.1	31.6	30.0	34.8	36.4	36.3	32.0	32.2	36.2	36.2
SD 3 bands, equal splits +cov	21.9	17.2	19.6	20.8	18.0	8.4	22.4	26.9	22.2	17.2	16.9	19.6

Table 6-32 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, seven covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	10.3	11.7	5.2	13.9	14.2	14.4	8.6	12.2	11.2	11.7	7.7	11.7
Proportional Odds: no cov	27.7	25.2	21.7	25.9	21.6	21.0	19.6	23.8	19.6	22.7	21.3	21.7
Proportional Odds: + cov	37.3	35.8	32.0	37.7	34.6	32.9	34.8	33.7	32.1	34.6	33.0	34.6
SD 3 bands, equal splits +cov	14.5	15.6	14.2	21.7	16.6	14.2	22.6	21.5	24.2	16.8	15.7	16.6

6.5.1.3 Nine covariate model

Both 5% and 8% treatment effects

For all trials, all alternative strategies showed an improvement over the conventional treatment only strategy and using the proportional odds plus covariates strategy gave the greatest reduction in sample size. As with the three and seven covariate models, the smallest reductions in sample size were obtained when the conventional plus covariates strategy was used. For SKB only, using a 5% treatment effect, the sliding dichotomy strategy gave a greater reduction in sample size than the proportional odds treatment only strategy as shown in Table 6-33 below. For all other studies, and SKB using an 8% treatment effect, the proportional odds plus treatment strategy gave a greater reduction in sample size than the sliding dichotomy as shown in Table 6-34 below.

Table 6-33 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, nine covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	14.4	15.4	11.7	7.4	6.3	12.8	8.1	11.3	7.7	11.7	8.3	11.3
Proportional Odds: no cov	22.8	27.8	26.2	22.0	17.2	20.5	20.4	18.8	18.8	19.6	21.7	20.5
Proportional Odds: + cov	33.8	38.4	36.6	31.7	33.6	32.6	37.1	32.3	32.3	32.9	33.9	33.6
SD 3 bands, equal splits +cov	18.0	20.8	20.2	16.8	14.0	12.3	16.8	18.4	21.5	17.5	17.2	17.5

Table 6-34 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, nine covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	10.7	12.3	5.9	10.3	12.9	11.6	11.6	12.3	11.8	13.1	8.6	11.6
Proportional Odds: no cov	29.3	25.3	25.3	21.7	25.0	25.9	19.8	24.9	22.0	27.7	21.4	25.0
Proportional Odds: + cov	35.0	37.2	34.6	34.5	34.5	35.0	34.8	38.2	32.0	37.7	30.8	34.8
SD 3 bands, equal splits +cov	16.7	17.4	19.1	13.7	16.0	12.3	15.3	20.7	18.1	19.8	14.2	16.7

General remarks

Using the seven covariate model gave a greater reduction in sample size than the three covariate model. However, there was little difference in reduction in sample size between the nine and seven covariate models.

6.5.2 Comparing strategies – mortality treatment effect

Here the simulated treatment effect is a reduction in mortality.

6.5.2.1 Three covariate model

For both the 5% and 8% treatment effects the proportional odds plus covariates strategy gives the greatest reduction in sample size compared with the conventional treatment only strategy for all studies.

5% treatment effect

Using the sliding dichotomy strategy gave an increase in sample size compared with the conventional treatment only strategy for TINT and SLIN as shown in Table 6-35 below. For all studies the smallest reduction in sample size was observed when the conventional plus covariates strategy is used.

Table 6-35 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	12.4	15.1	12.1	2.8	12.1	10.0	7.6	16.7	10.4	4.3	6.9	10.4
Proportional Odds: no cov	25.8	28.3	24.1	28.8	37.0	34.3	34.4	34.8	36.8	33.6	30.3	33.6
Proportional Odds: + cov	33.8	41.9	34.4	38.0	47.0	42.4	45.3	49.5	49.0	41.2	40.2	41.9
SD 3 bands, equal splits +cov	-9.6	17.1	-28.2	22.4	30.5	32.7	29.9	35.6	37.1	27.2	18.6	27.2

8% treatment effect

For TINT, TIUS and SLIN, using the sliding dichotomy strategy resulted in an increase in sample size compared with using the conventional treatment only strategy as shown in Table 6-36 below. For SLIN this increase is greater than 50% of the sample size with the conventional treatment only strategy. Using the conventional plus covariates strategy gave a modest reduction in sample size compared with using the conventional treatment only strategy. Using the proportional odds treatment only strategy gave the second greatest reduction in sample size.

Table 6-36 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	8.8	16.1	6.0	9.9	9.1	11.4	11.4	13.1	9.2	9.9	11.1	9.9
Proportional Odds: no cov	28.6	27.7	21.5	28.0	33.8	34.5	37.5	41.1	36.4	32.3	29.4	32.3
Proportional Odds: + cov	40.3	42.1	32.3	41.8	42.0	43.6	45.0	51.6	44.5	42.9	39.6	42.1
SD 3 bands, equal splits +cov	-25.8	-25.3	-52.2	16.3	30.9	24.4	31.7	37.8	24.5	23.8	3.0	23.8

6.5.2.2 Seven covariate model

For both the 5% and the 8% treatment effects using the proportional odds plus covariates strategy gave the greatest reduction in sample size compared with using the conventional treatment only strategy.

5% treatment effect

Using the sliding dichotomy strategy gave a very small increase in sample size compared to the conventional treatment only strategy for SLIN as shown in Table 6-37 below. The smallest reductions in sample size were observed when the conventional plus covariates strategy was used. Using the proportional odds treatment only strategy gave the second largest reductions in sample size, greater in all cases than using the sliding dichotomy.

Table 6-37 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, seven covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	7.2	11.3	8.8	13.6	9.5	5.6	7.6	7.1	8.2	11.3	8.8	8.8
Proportional Odds: no cov	35.5	30.2	35.5	37.7	34.7	32.8	34.4	32.9	39.5	38.0	31.3	34.7
Proportional Odds: + cov	43.7	41.8	43.5	47.9	43.8	40.7	41.0	43.2	49.6	45.7	39.0	43.5
SD 3 bands, equal splits +cov	18.5	15.4	-1.0	31.3	28.7	20.8	25.9	31.6	37.2	33.4	16.3	25.9

8% treatment effect

A similar pattern was observed with the 8% treatment effect as with the 5% treatment effect in that the proportional odds treatment only strategy gave greater reductions in sample size than the sliding dichotomy strategy as shown in Table 6-38 below. Using the sliding dichotomy resulted in an increase in sample size for TINT, TIUS, SLIN and HIT II compared with using the conventional treatment only strategy. No consistent increase in reduction in sample size was observed when using the 8% treatment effect compared with the 5% treatment effect.

Table 6-38 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, seven covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	11.6	8.2	9.7	9.4	8.0	9.6	8.1	11.1	7.2	7.8	8.5	8.5
Proportional Odds: no cov	30.8	32.4	31.9	33.6	40.1	31.7	36.6	38.8	37.4	35.1	32.3	33.6
Proportional Odds: + cov	40.4	42.2	39.9	46.6	47.3	41.6	45.5	49.8	45.4	43.4	40.6	43.4
SD 3 bands, equal splits +cov	-13.9	-13.9	-20.3	8.8	27.0	4.5	30.8	27.5	25.5	20.1	-5.8	8.8

6.5.2.3 Nine covariate model

A similar pattern was observed with the nine covariate model as with the seven covariate model. Using the sliding dichotomy strategy resulted in an increase in sample size for SLIN with the 5% treatment effect and TINT, TIUS, SLIN and HIT II with the 8% treatment effect compared with using the conventional treatment only strategy. Very similar reductions in sample size were observed both with the 5% and 8% treatment effects as shown in Table 6-39 and Table 6-40 below. The conventional plus covariates strategy gave the smallest reductions in sample size of any of the alternative strategies.

Table 6-39 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, nine covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	13.9	7.7	14.3	4.9	11.7	6.6	11.0	10.3	9.2	7.5	9.7	9.7
Proportional Odds: no cov	36.6	32.8	33.0	35.2	34.5	32.5	41.2	40.6	44.8	33.8	35.7	35.2
Proportional Odds: + cov	44.7	41.8	40.7	43.9	45.6	40.4	47.9	49.6	50.9	39.4	43.1	43.9
SD 3 bands, equal splits +cov	13.6	10.7	-8.2	21.7	23.5	6.6	35.1	37.0	36.7	17.9	8.4	17.9

Table 6-40 Targeting only patients with an intermediate prognosis. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, nine covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	10.7	14.2	9.6	12.4	12.8	9.5	6.9	12.1	8.0	8.4	9.5	9.6
Proportional Odds: no cov	32.2	33.5	33.1	34.0	37.0	33.1	40.8	38.0	42.2	33.3	33.6	33.6
Proportional Odds: + cov	42.1	43.5	43.9	46.7	47.3	44.1	48.4	50.7	48.5	43.0	44.2	44.2
SD 3 bands, equal splits +cov	-20.1	-15.4	-25.7	3.9	19.0	-19.3	26.2	23.9	31.5	5.6	-14.6	3.9

In general, similar reductions in sample size were observed when using the three, seven and nine covariate models.

6.5.3 Graphical comparison – targeting only patients with an intermediate prognosis

Figure 6-11 and Figure 6-12 show that, for all studies, using either the sliding dichotomy or proportional odds strategies gives a reduction in sample size when the treatment effect follows a proportional odds model. Greater reductions were observed using the proportional odds strategies than the sliding dichotomy strategies. Using models with greater numbers of covariates did not give corresponding greater reductions in sample size for either the sliding dichotomy or proportional odds strategies.

The proportional odds strategies gave greater reductions in sample size compared with the conventional dichotomy for all studies when simulating a treatment effect that gave a reduction in mortality, as shown in Figure 6-13 and Figure 6-14 below. These reductions were typically around 40% for all three covariate models and both the 5% and 8% treatment effects. Using the sliding dichotomy strategies with the 5% treatment effect showed a reduction in sample size for all studies except SLIN and TINT. However, the sliding dichotomy strategies performed more poorly with the 8% treatment effect with four of the trials showing an increased sample size compared with the conventional dichotomy.

This pattern is confirmed when the median reduction in sample size over all trials is shown, Figure 6-15. All scenarios over all treatment effects showed a reduction in sample size compared with the conventional dichotomy. Using the seven covariate model with covariates or the nine covariate models gave the greatest reductions in sample size of 35 to 45%.

Figure 6-11 Uniform 5%: Targeting only those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]

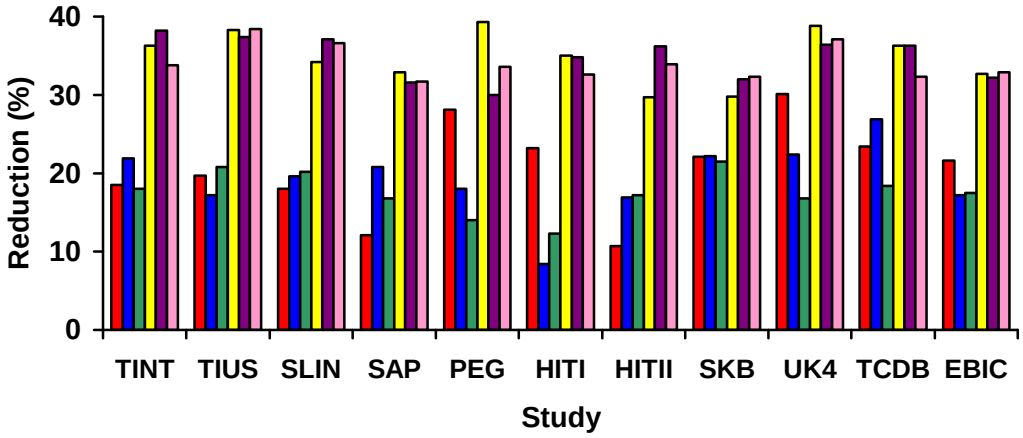
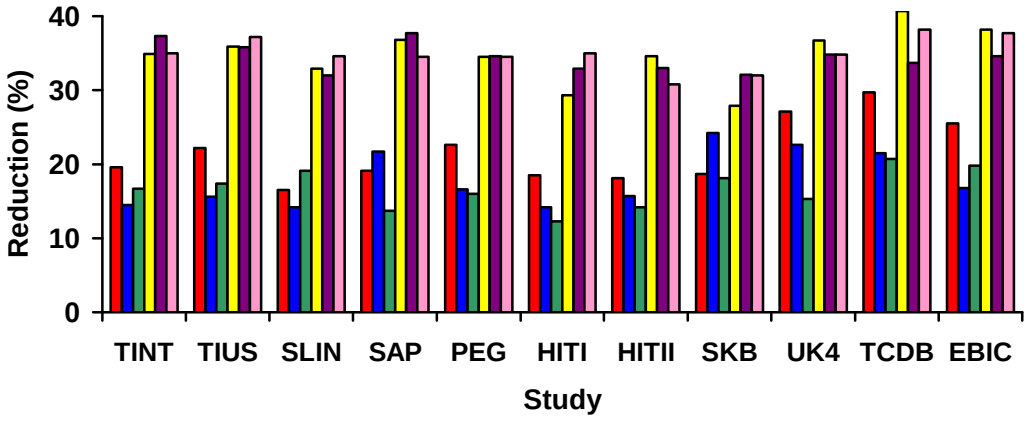


Figure 6-12 Uniform 8%: Targeting only those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red) Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

Figure 6-13 Mortality 5%: Targeting only those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]

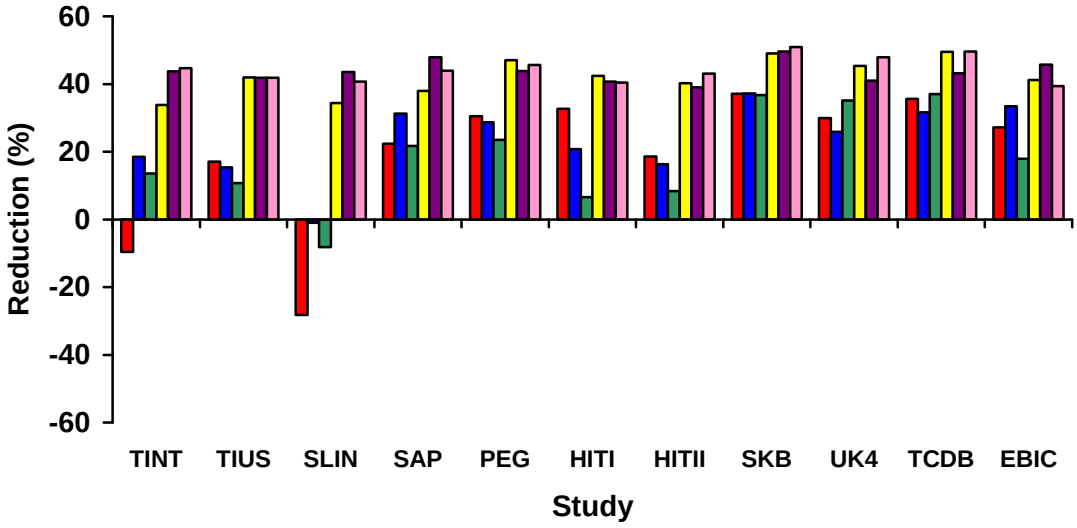
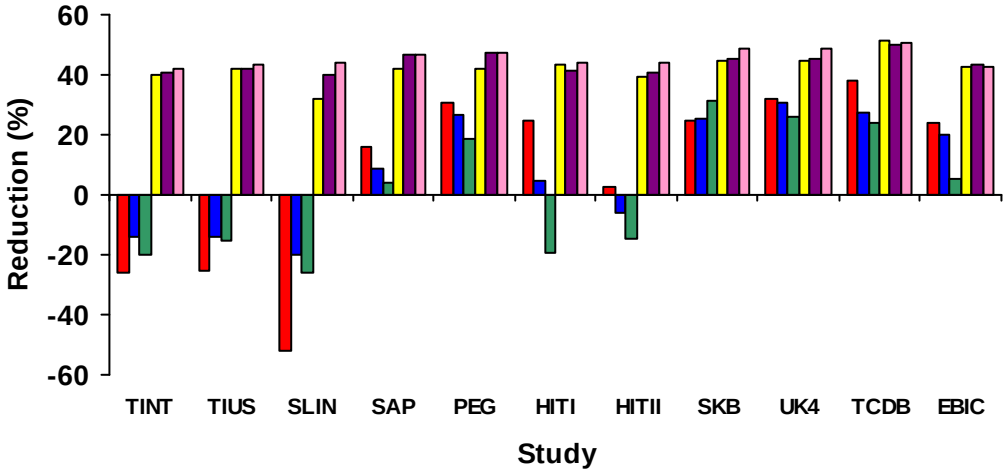
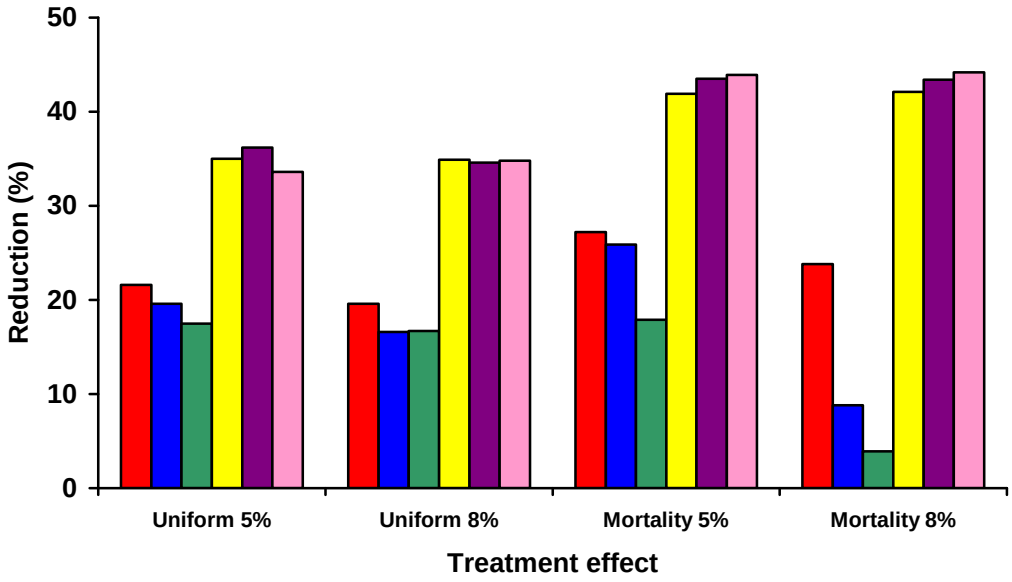


Figure 6-14 Mortality 8%: Targeting only those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red) Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

Figure 6-15 Median reductions in sample size. Targeting only those with an intermediate prognosis. Comparison of Sliding Dichotomy & Proportional Odds models with covariates[9]



Key: Sliding Dichotomy 3 covariates (red) Sliding Dichotomy 7 covariates (dark blue), Sliding Dichotomy 9 covariates (green), Proportional odds 3 covariates (yellow), Proportional odds 7 covariates (purple), Proportional odds 9 covariates (pink)

6.6 Targeting only patients with a mass lesion

Only the results from the three covariate model are shown for both the uniform and mortality scenarios as both the seven and nine covariate models contain CT.

6.6.1 Comparing strategies – uniform treatment effect

For both the 5% and 8% treatment effects all alternative strategies showed an improvement over the conventional treatment only strategy for all studies when the treatment effect followed a proportional odds model as shown in Table 6-41 and Table 6-42 below. Using the proportional odds plus covariates strategy gives the greatest reduction in sample size, again for all studies. Similar reductions are observed with the conventional plus covariates and the proportional odds treatment only strategies. The sliding dichotomy strategy gave a greater reduction in sample size than the conventional plus covariates and the proportional odds treatment only strategies.

Table 6-41 Targeting only patients with a mass lesion. Median and by trial reductions in sample size. Uniform treatment scenario, 5% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	22.3	21.0	16.2	15.0	25.4	26.3	26.5	28.0	34.2	14.8	22.3	22.3
Proportional Odds: no cov	22.5	18.2	24.2	16.9	25.4	23.4	20.7	20.4	27.1	17.3	22.5	22.5
Proportional Odds: + cov	44.1	44.7	39.0	38.1	48.2	48.4	45.5	48.4	53.5	39.5	44.1	44.7

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
SD 3 bands, equal splits +cov	36.0	32.2	27.4	26.2	38.5	42.7	33.4	39.1	47.6	33.0	36.0	36.0

Table 6-42 Targeting only patients with a mass lesion. Median and by trial reductions in sample size. Uniform treatment scenario, 8% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	22.6	25.0	11.8	20.5	21.1	25.1	23.9	31.1	26.2	19.1	22.6	22.6
Proportional Odds: no cov	25.6	22.3	23.8	21.8	19.4	20.3	22.5	25.1	19.6	19.7	25.6	22.3
Proportional Odds: + cov	46.4	46.7	40.0	42.3	43.9	45.9	46.6	52.7	50.2	40.8	46.4	46.4
SD 3 bands, equal splits +cov	33.1	36.7	29.8	31.9	31.1	37.0	34.3	42.4	39.7	28.4	33.1	33.1

6.6.2 Comparing strategies – mortality treatment effect

For both the 5% and 8% treatment effects, for all studies, except SLIN, all alternative strategies show an improvement over the conventional treatment only strategy when the treatment effect was modelled as a reduction in mortality as shown in Table 6-43 and Table 6-44 below. For SLIN, using the sliding dichotomy strategy gave an increase in sample size compared to the conventional treatment only strategy. For all studies the greatest reduction in sample size was observed using the proportional odds plus covariates strategy. The conventional plus covariates strategy gave the smallest reductions in sample size. The sliding dichotomy and the proportional odds treatment only strategies gave similar reductions in sample size. Overall, the greatest

reductions in sample size were observed in the studies which were unselected series. The trials with the smallest numbers of participants had the smallest reductions in sample size.

Table 6-43 Targeting only patients with a mass lesion. Median and by trial reductions in sample size. Mortality treatment scenario, 5% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	24.6	21.6	12.3	17.6	16.0	21.5	22.1	23.9	29.0	14.8	24.6	21.6
Proportional Odds: no cov	32.3	30.9	16.0	42.0	43.9	44.2	45.9	48.1	46.7	39.6	32.3	42.0
Proportional Odds: + cov	49.3	47.6	32.1	51.4	54.1	61.5	55.7	62.1	62.0	53.5	49.3	53.5
SD 3 bands, equal splits +cov	39.0	41.1	-8.9	36.7	44.1	48.8	47.2	60.4	45.5	37.8	39.0	41.1

Table 6-44 Targeting only patients with a mass lesion. Median and by trial reductions in sample size. Mortality treatment scenario, 8% treatment effect, three covariate model[10]

Scenario	Study											Median
	TINT	TIUS	SLIN	SAP	PEG	HITI	UK4	TCDB	SKB	EBIC	HITII	
Conventional no cov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Conventional + cov	17.6	21.2	12.1	19.4	21.8	22.0	19.6	23.5	24.7	17.7	17.6	19.6
Proportional Odds: no cov	30.3	35.0	18.0	38.3	45.6	41.3	45.9	47.5	43.3	37.8	30.3	38.3
Proportional Odds: + cov	46.0	50.2	35.6	52.1	57.7	57.0	58.8	60.7	60.0	53.2	46.0	53.2
SD 3 bands, equal splits +cov	32.0	37.3	-30.3	41.7	41.7	42.2	45.2	58.2	39.4	38.6	32.0	39.4

6.7 Discussion

The first part of this chapter has shown the comparison of the optimised sliding dichotomy with a number of different modelling strategies. The second part has examined whether the results observed for all subjects, as shown in the first part, are replicated when either improvement is restricted to a group of patients or a group of patients were targeted for improvement. When restricting improvement to those with an intermediate prognosis the same pattern was observed for both the Uniform and Mortality treatment effects as with using data from all subjects. For improvement restricted to patients with a mass lesion, for some studies under certain scenarios using the sliding dichotomy actually increased the sample size compared with the conventional analysis. This was possibly because of the smaller subset of patients in this analysis. Targeting only patients with an intermediate prognosis, showed a similar pattern to that observed when improvement was restricted to patients with an intermediate prognosis. However, as expected, the reductions in sample size seen were smaller. When targeting only patients with a mass lesion almost all scenarios for all studies showed a reduction in sample size compared with the conventional dichotomous analysis.

All of the modelling results have shown consistently that adding covariates to the models results in greater sample size reductions for all scenarios. Using covariates with the proportional odds model consistently gave the greatest sample size reductions both when the data followed a proportional odds treatment effect (the Uniform treatment scenario) and when the data did not (the Mortality treatment scenario). Using the sliding dichotomy with covariates typically gave the next greatest reductions in sample size although for some scenarios using the proportional odds model without covariates or the conventional analysis with covariates gave a similar magnitude of reduction. Almost without exception all alternative strategies

performed much better than the conventional analysis leading to reductions in sample size of typically around 40%.

7 Chapter 7 Discussion and Recommendations

7.1 Findings

Some of the text below, in particular the section on the sliding dichotomy has been published (McHugh et al. 2010).

This thesis has shown that by using the ordinal nature of the Glasgow Outcome Scale it is possible to achieve substantial sample size reductions whilst still preserving power. These sample size gains are in addition to those achieved by covariate adjustment alone (Hernandez et al. 2006). Indeed, adding covariates to those strategies which exploited the ordinal nature of the GOS gave an additional benefit, with sample size reductions of over 50% in some cases.

The proportional odds model was the most statistically efficient, consistently giving the greatest sample size reductions. It is perhaps not surprising that it performed well when the simulated treatment effect was structured to follow a proportional odds model - the Uniform model. What is more surprising is that using the proportional odds model still gave the greatest sample size reductions when the treatment effect clearly did not follow the proportional odds assumption – the Mortality model. The proportional odds model was also statistically efficient when the treatment effect was restricted to those with an intermediate prognosis and when it was restricted to those with a mass lesion. Even when targeting groups of patients, those with an intermediate prognosis and those with a mass lesion, it still gave the largest sample size reductions. The sliding dichotomy also performed well. Although it did not give quite the magnitude of sample size reductions observed with using the proportional odds model, it still gave a vastly superior result compared with using the conventional dichotomous analysis. The choice of when to use which approach is a value judgement.

7.2 Related work

Others have found comparable study results in different populations. The Optimising Analysis of Stroke Trials (OAST) collaboration have performed two studies which have shown that statistical approaches which exploited the ordered nature of the outcomes were statistically more efficient than those which collapsed the data into two categories (Bath 2007;The Optimising Analysis of Stroke Trials (OAST) collaboration 2008). Using a combination of individual patient data and summarised data from 55 stroke studies they found that the findings were consistent both over different measures of functional outcome and different studies. In the first study, (Bath et al. 2007), three outcome scales were examined: the modified Rankin Scale; the Barthel Index; and the "3-questions" (taken from the modified Rankin Scale) (Lindley et al. 1994). Within each trial, tests which exploited the ordinal nature of the data, such as ordinal logistic regression and tests which dichotomised the data, such as the Chi-squared test, were compared. Statistical tests which dichotomised the data were assessed multiple times at each of the possible dichotomies. The results of all of these tests were then ranked within each trial. The test which produced the most significant result, i.e. the largest z score was given the lowest rank. Two way analysis of variance was then used to examine which test on average over all of the trials produced the lowest rank. This then enabled ordering of the tests in terms of efficiency of identifying treatment effects. Ordinal logistic regression was consistently given the lowest rank showing it to be the most efficient at identifying treatment effects with the Chi-squared test and the Kolmogorov-Smirnov test consistently showing the least efficiency in identifying treatment effects. In the second study (The Optimising Analysis of Stroke Trials (OAST) collaboration 2008), sample sizes were calculated using comparisons of proportions, means, medians and ordinal data. Ordinal methods, with sample size calculated according to Whitehead's formula (Whitehead 1993) on average reduced sample size by 28%.

The results shown in this thesis, on head injury trials, expand these findings. The OAST collaborators used a mixture of individual patient data where available and summarised data. Only individual patient data made into a common database format

were used in this thesis. It is interesting that our overall results were similar with both groups finding that ordinal methods consistently gave substantial sample size reductions. Interestingly, in the second study, the OAST group found that no method of analysis, except one dichotomy, worked well with the thrombolysis trials. They comment that "ordinal methods may not be optimal for interventions which both improve functional outcome and cause a hazard in a subset of patients, e.g. thrombolysis." This thesis has also shown some situations where an ordinal analysis may not be optimal, for example, when modelling a reduction in mortality for some studies. The OAST work has been expanded here as covariates have been used to develop prognostic models which have been used to discriminate patients on the basis of baseline risk. Here too, many more modelling strategies which are used with ordinal data have been compared, firstly the refinement of the sliding dichotomy then the comparison with the other methods.

Lees et al (Lees et al. 2006a) also examined outcome after stroke, as measured by the NIHSS scale. A number of simulations were performed comparing the power achieved for a fixed sample size when the outcome was analysed, using the Cochran Mantel Haenszel test, both using the full NIHSS scale and when the scale was dichotomised. All simulations showed a decrease in power when using the dichotomous analysis. For example, even with a sample size as large as 4700 subjects, analysing the total NIHSS score gave 75% power and dichotomising the score gave a power of only 55%.

A similar method of analysis has been applied to the Glycine Antagonist In Neuroprotection (GAIN) stroke trial (Lees et al. 2000). Young and colleagues (Young, Lees, & Weir 2003) performed imputation on the stroke data to predict outcome on the Barthel and Rankin (Bonita & Beaglehole 1988) scales. They simulated 24 000 trials exploring various patterns of outcome, although not prognosis based outcomes, and found that using the Rankin Scale instead of the

dichotomised Barthel scale could reduce sample size by up to 84%. In a second examination of the data, Young and colleagues (Young, Lees, & Weir 2005) reported that the use of a prognosis based endpoint could also reduce sample size by up to 49%.

Others have also examined the effects of adjusting for covariates on sample size and power. The OAST collaboration (Gray, Bath, & Collier 2009) performed simulation studies examining this. Using data from over 25 000 patients in 23 stroke trials they compared adjusted (treatment model adjusted for age, sex and baseline severity) and unadjusted (treatment only) ordinal regression models over three levels of treatment effect. They observed similar results to the ones reported in this thesis with reductions of sample size of at least 20% to 30% when covariates were taken into account.

One published trial, SCAST (Sandset et al. 2011), has compared both ordinal logistic regression and the sliding dichotomy on real trial data. Using ordinal regression analysis gave a statistically significant result whereas using the sliding dichotomy did not. This may have been because of the reduced power with this method of analysis compared with ordinal regression.

Two other stroke trials have also used the sliding dichotomy to analyse outcome. The STICH trial (Mendelow et al. 2005) and the PAIS trial (den Hertog et al. 2009) used the sliding dichotomy methodology as their primary analysis. It is interesting that the PAIS trial was originally going to use a dichotomisation of the modified Rankin Scale as its primary outcome. However, as the study recruited patients it was realised that the planned number of patients could not be recruited. The protocol was then modified before the treatment codes were broken (den Hertog et al. 2008) to specify use of the sliding dichotomy as a way of increasing the power of the study to

detect a statistically significant difference. The authors commented that if analysis had been as originally planned the study would have had only 60% power to find the difference specified (den Hertog et al. 2009).

A recent reanalysis of the CRASH (Edwards et al. 2005) trial has compared both the sliding dichotomy and proportional odds approach (Roozenbeek et al. 2011). The published dichotomous analysis showed a non-significant treatment effect 1.09 (0.98 to 1.21) p=0.096. Using both proportional odds modelling and the sliding dichotomy, highly statistically significant treatment effects were observed, showing a deleterious effect of the active treatment. The odds of an unfavourable outcome was 1.15 95%CI (1.06 to 1.25) p=0.0007 using proportional odds modelling and an odds ratio of 1.19, 95%CI (1.08 to 1.30) p=0.0002 for the sliding dichotomy.

7.3 Limitations

All of the simulation studies were based on a sample size of 800 subjects. It would be of interest to see if the findings in the thesis would be repeated with smaller and larger studies. A few limited simulations were run on a larger sample size and these showed a similar pattern to the simulations with 800 subjects. This thesis also only considered two treatment effects, 5% and 8%. Although the results were consistent over both treatment effects, a natural extension would be to model different treatment effects. Only three covariate models were compared. From these findings there was not a direct relationship between number of covariates and increase in sample size. It would however be interesting to explore whether the addition of many more covariates would lead to greater gains.

A potential criticism is that the IMPACT data may not be representative of current head injury trials as the last study was completed in 1995. The same pattern of results were shown when both restricting improvement to certain groups and when

targeting certain groups of patients. As this is currently done in trial recruitment this would seem to show the robustness of these methods.

The analysis of the IMPACT data could be criticised because it was only performed on one large dataset, albeit one which contained eleven composite studies. To have external validity the models would need to be fitted on another dataset different from the ones on which the models were derived (Justice, Covinsky, & Berlin 1999). Two of the prognostic models used in the IMPACT database have been validated in other datasets. Steyerberg at al (Steyerberg et al. 2008) used data from the CRASH trial (Edwards et al. 2005) to validate the models used with the IMPACT data. Subjects with a GCS <=12 and a 6 month GOS were selected from the CRASH data to allow validation on a similar population. As CRASH did not record lab values, two models were validated. The first was the three covariate set (age, motor and pupil). This gave an AUC of 0.776 when modelling survival and an AUC of 0.780 when modelling unfavourable outcome. A five variable model was also fitted to the CRASH data. This added CT and tSAH to the three covariate model, giving AUC of 0.801 and 0.796 for the mortality and unfavourable outcome models respectively. All models had adequate discrimination. Of course, no statistical model is 'correct' so future work could be done on the refinement of these models and their clinical impact (Lingsma et al. 2010). This is, of course, a criticism of prognostic models in general rather than a limitation of the simulation studies described.

Criticisms could be levelled because the assumption of proportional odds did not hold for all analyses. Senn and Julious make the argument that two different statisticians could cut the same continuous scale at two different cut points. If we can accept the results from either dichotomy, then accepting the compromise answer from the proportional odds model seems reasonable (Senn & Julious 2009). Wardlaw et al (Wardlaw et al. 2000) have shown this empirically with stroke data. They performed a meta-analysis of twelve thrombolysis trials which measured

outcome after stroke using the dichotomised modified Rankin Scale. Wardlaw used two different definitions of good and poor outcome. They compared two cut points of the modified Rankin Scale: 0 to 2 (good outcome) versus 3 to 6 (poor outcome) and 0 to 1(good outcome) versus 2 to 6 (poor outcome). Both analyses gave very similar results: poor outcome 3 to 6 OR=0.83; poor outcome 2 to 6 OR=0.79, showing that the treatment effect was robust to using different definitions of good or poor outcome. Changing the definitions did however change the outcomes in several trials from "success" to "failure" and vice versa. This emphasises how changing just one point on a scale can change the outcome from "success" to "failure" when in fact there is no difference in the magnitude of the result.

When combining estimates from the different studies there was statistically significant heterogeneity in some cases. Bailey (Bailey 1987) made the important point that it is not surprising that studies with different designs have heterogeneous outcomes. He proposed three important questions. "(1) which question one is trying to answer; (2) the degree of similarity or dissimilarity of design, and (3) the degree to which heterogeneity of outcomes can be explained." In this thesis the question has always been focussed on what methods are best to analyse ordinal outcome data. The studies that were compared had two distinct designs of being unselected cohorts and trials however all were in patients with moderate to severe head injury and the heterogeneity of the outcomes could usually be explained by the different study designs.

Meta analysis and the pooling of studies can also be criticised. Studies included should have enough in common that the evidence from them can be interpreted for current patients (Thompson & Pocock 1991). The studies pooled for these meta analyses were similar in that they were all in moderately to severely head injured patients. However, three of the studies were observational series and the rest were

trials. However, by using covariates such as gender as a 'treatment effect' this probably offers enough commonalties between studies.

The findings in this thesis that the proportional odds model gives a robust and sensitive analysis even when the formal assumption of proportional odds in not held has been found by others. Tsodikov et al (Tsodikov, Hasenclever, & Loeffler 1998) modelled a U shaped treatment effect and found that the proportional odds model gave the most stable estimates when repeated bootstrap sampling was done, albeit on a much smaller sample than was studied in this thesis. Whitehead et al (Whitehead et al. 2001) also fitted the proportional odds model to a variety of scenarios and found that it performed well.

Other trial designs have been proposed to deal with the problem of recruitment in trials. Relton et al (Relton et al. 2010) proposed a "cohort multiple randomised controlled trial" design. This involves obtaining a large epidemiological cohort who agree to observational measurements being taken. Trial subjects are then randomised from within this cohort. The authors argue that this design will help with patient recruitment issues however no mention is made of the difficulties of recruiting the large epidemiological cohort and why this would be easier that recruiting for a pragmatic trial per se. Using more efficient statistical analysis, giving greater power and hence a reduction in sample size would seem to be a much simpler and more robust strategy for dealing with this problem.

7.4 Future work

More sophisticated approaches to implementing the sliding dichotomy could be used. For example, the thresholds could be constrained to be monotonic, so that moving from a poorer prognosis band to a better prognosis band, the threshold for 'favourable' would need to be at the same point as the previous band or else higher

up the GOS scale. Indeed, an interesting approach would be to use an extremely large number of bands but to enforce monotonicity, giving a very flexible way of assigning an appropriate definition of 'favourable' outcome for each individual patient. Ultimately, it is not necessary to think of bands at all. In this context of analysing a 4-point ordinal scale, the sliding dichotomy will always end up with three groups of patients. Those with poor prognosis, where the threshold will be between D/V and SD; those with intermediate prognosis, where the threshold will be between SD and MD; and those with good prognosis, where the threshold will be between MD and GR. In some situations with extreme outcome distributions, one or even two of these groups might not contain any patients. What is required in general is an efficient way of splitting the ordered list of patients (ordered by baseline prognosis) into these three groups. Another possible approach, for example, might be to work from the worst prognosis patient upwards until a group is obtained with an approximate 50:50 split between D/V and better, and make this the poor prognosis group. Then work from the best prognosis patient downwards until a group is obtained with an approximate 50:50 split between GR versus worse, and make this the good prognosis group. The remaining patients would be the intermediate prognosis group. However, in practice, these groups might overlap or it might not be possible to find a group with the required 50:50 outcome split. Considerable ingenuity would need to be applied to make the approach robust over all possible extreme outcome distributions. Since the sliding dichotomy approach collapses the ordinal scale to a binary one for final analysis, it must still be discarding some relevant information. Thus, from first principles, and as supported by the simulation results, one would expect that, in most situations, the potential efficiency gains from a fully optimised sliding dichotomy analysis would be bounded above by the gains that can be achieved using the proportional odds model.

This thesis could have examined a composite endpoint rather than just GOS. Two studies used this approach on summarised data from trials (Bath et al. 2008;Geeganage et al. 2010). Bath examined vascular prevention trials and Geeganage examined blood pressure reduction and cardiovascular prevention trials.

Instead of categorising outcome as stroke/no stroke and comparing event rates between treatment groups, both Bath and Geeganage expanded this firstly to a three category ordinal outcome i.e. fatal event/non fatal event/no event. This was then expanded to a four category ordinal outcome: fatal stroke/severe non fatal stroke/mild stroke/no stroke and then a five category ordinal outcome: fatal stroke/ severe non fatal stroke/ mild stroke/ transient ischaemic event/no event. Instead of considering outcome as a range of different points on a scale, as with the Glasgow Outcome Scale, different outcomes were thought of as being on the same continuum. In addition to studying stroke Bath and Geeganage similarly categorised MI, vascular events and bleeding (Bath only). Bath compared 10 different statistical tests on each of the 85 trials (four of the tests assumed ordinal data, the rest assumed binary or continuous data). The results within each trial were then ordered and Friedman's test used to assess which test gave the lowest ranks i.e. the most statistically significant values. For all outcomes, using ordinal analysis methods consistently outperformed using binary analysis methods. Geeganage used summarised data from 38 trials and found similar odds ratios were obtained from binary and ordinal outcomes. She showed, with one trial, that using a five level ordinal endpoint could reduce sample size by 56% compared with using a binary endpoint.

The GOS could have been treated as a linear scale. The modified Rankin Scale is conventionally dichotomised between scores 2 and 3 however Hong (Hong & Saver 2009) found that there was little difference between scores 2 and 3 when they are compared using disability weights. This further supports the hypothesis that it is the change in scale that matters rather than the jumping over a, possibly arbitrary, boundary.

This thesis has only looked at the five point GOS. A natural extension to this would be to look at the GOSE to see if the results observed would be observed in this scale. Subsequent work since this thesis has found that the proportional odds model and the sliding dichotomy both still perform excellently when compared to the traditional binary dichotomisation of the GOSE (Weir et al. 2012). The efficiency gains found

are slightly more than was found using the GOS, indicating that more information can be obtained by using the extended scale. Levin et al (Levin et al. 2001) also found that the GOSE was more sensitive to change than the GOS, again indicating that this may be a better outcome scale to use than the GOS.

In conclusion, this thesis has shown that when analysing an ordinal outcome scale using any analysis which exploits this ordering is typically far superior to the binary analysis which is usually performed. When appropriate, using an ordinal approach will either allow far fewer subjects to be recruited for the same power or will allow the same number of subjects to be recruited but a much smaller treatment effect to be detected. Of course, there are some models of treatment effect where using either proportional odds modelling or the sliding dichotomy would be inappropriate, for example a "kill or cure" model. However, with this kind of treatment effect the conventional dichotomous analysis will also be inappropriate.

The concept of the sliding dichotomy has, arguably, more clinical appeal. Bagiella et al (Bagiella et al. 2010) claims that a dichotomised measure is more easily translated to a general population. This may be so but using a fixed dichotomy of the GOSE for all trials, both exploratory and confirmatory does not make clinical sense. An exploratory trial is a mechanistic trial looking for a sensitive outcome whereas a confirmatory trial is looking for something relevant to patients measured using, say, the GOSE.

The choice of whether to use the sliding dichotomy or the proportional odds analysis can depend on the balance between simplicity and clinical appeal (the sliding dichotomy) and statistical efficiency with perhaps a more complex argument (the proportional odds model). This is typically the threshold between, say, a late Phase II trial/early Phase III trial where one wishes to have maximum statistical efficiency, and hence use the proportional odds model, compared with a late Phase III trial where one may sacrifice statistical efficiency for clinical relevance, and hence use the sliding dichotomy. However, no matter which one is used, large gains can potentially be made.

8 References

Adams, H. P., Jr., Leclerc, J. R., Bluhmki, E., Clarke, W., Hansen, M. D., & Hacke, W. 2004, "Measuring outcomes as a function of baseline severity of ischemic stroke", *Cerebrovascular Diseases*, vol. 18, no. 2, pp. 124-129.

Agresti, A. 2002, *Categorical Data Analysis* Wiley, New York.

Ali, M., Bath, P. M. W., Curram, J., Davis, S. M., Diener, H. C., Donnan, G. A., Fisher, M., Gregson, B. A., Grotta, J., Hacke, W., Hennerici, M. G., Hommel, M., Kaste, M., Marler, J. R., Sacco, R. L., Teal, P., Wahlgren, N. G., Warach, S., Weir, C. J., & Lees, K. R. 2007, "The virtual international stroke trials archive", *Stroke*, vol. 38, no. 6, pp. 1905-1910.

Anderson, J. A. 1984, "Regression and Ordered Categorical Variables", *Journal of the Royal Statistical Society Series B-Methodological*, vol. 46, no. 1, pp. 1-30.

Anderson, J. A. & Philips, P. R. 1981, "Regression, Discrimination and Measurement Models for Ordered Categorical Variables", *Applied Statistics-Journal of the Royal Statistical Society Series C*, vol. 30, no. 1, pp. 22-31.

Ashby, D., Pocock, S. J., & Shaper, A. G. 1986, "Ordered Polytomous Regression - An Example Relating Serum Biochemistry and Hematology to Alcohol-Consumption", *Applied Statistics-Journal of the Royal Statistical Society Series C*, vol. 35, no. 3, pp. 289-301.

Bagiella, E., Novack, T. A., Ansel, B., Diaz-Arrastia, R., Dikmen, S., Hart, T., & Temkin, N. 2010, "Measuring Outcome in Traumatic Brain Injury Treatment Trials: Recommendations From the Traumatic Brain Injury Clinical Trials Network", *Journal of Head Trauma Rehabilitation*.

Bailey, I., Bell, A., Gray, J., Gullan, R., Heiskanan, O., Marks, P. V., Marsh, H., Mendelow, D. A., Murray, G., & Ohman, J. 1991, "A trial of the effect of nimodipine on outcome after head injury", *Acta Neurochirurgica*, vol. 110, no. 3-4, pp. 97-105.

Bailey, K. R. 1987, "Inter-study differences: how should they influence the interpretation and analysis of results?", *Statistics in Medicine*, vol. 6, no. 3, pp. 351-360.

Barer, D. 1999, "ECASS II: intravenous alteplase in acute ischaemic stroke. European Co-operative Acute Stroke Study-II", *Lancet*, vol. 353, no. 9146, pp. 66-67.

Bath, P. M., Cotton, D., Martin, R. H., Palesch, Y., Yusuf, S., Sacco, R., Diener, H. C., Estol, C., & Roberts, R. 2010, "Effect of combined aspirin and extended-release dipyridamole versus clopidogrel on functional outcome and recurrence in acute, mild ischemic stroke: PRoFESS subgroup analysis", *Stroke*, vol. 41, no. 4, pp. 732-738.

Bath, P. M., Geeganage, C., Gray, L. J., Collier, T., & Pocock, S. 2007, "Can we improve the statistical analysis of vascular prevention trials? Assessment of ordinal outcomes", *Stroke*, vol. 38, no. 2, pp. 523-524.

Bath, P. M., Geeganage, C., Gray, L. J., Collier, T., & Pocock, S. 2008, "Use of ordinal outcomes in vascular prevention trials: comparison with binary outcomes in published trials", *Stroke*, vol. 39, no. 10, pp. 2817-2823.

Bath, P. M., Martin, R. H., Palesch, Y., Cotton, D., Yusuf, S., Sacco, R., Diener, H. C., Toni, D., Estol, C., & Roberts, R. 2009, "Effect of telmisartan on functional outcome, recurrence, and blood pressure in patients with acute mild ischemic stroke: a PRoFESS subgroup analysis", *Stroke*, vol. 40, no. 11, pp. 3541-3546.

Bath, P. M. W. 2007, "Can we improve the statistical analysis of stroke trials? Statistical reanalysis of functional outcomes in stroke trials", *Stroke*, vol. 38, no. 6, pp. 1911-1915.

Bluhmki, E., Chamorro, A., Davalos, A., Machnig, T., Sauce, C., Wahlgren, N., Wardlaw, J., & Hacke, W. 2009, "Stroke treatment with alteplase given 3.0-4.5 h after onset of acute ischaemic stroke (ECASS III): additional outcomes and subgroup analysis of a randomised controlled trial", *Lancet Neurology*, vol. 8, no. 12, pp. 1095-1102.

Bolland, K., Sooriyarachchi, M. R., & Whitehead, J. 1998, "Sample size review in a head injury trial with ordered categorical responses", *Statistics in Medicine*, vol. 17, no. 24, pp. 2835-2847.

Bonita, R. & Beaglehole, R. 1988, "Recovery of Motor Function After Stroke", *Stroke*, vol. 19, no. 12, pp. 1497-1500.

Box, G. E. P. & Draper, N. R. 1987, *Empirical Model-Building and Response Surfaces* Wiley.

Braakman, R., Schouten, H. J. A., Blaauwvandishoeck, M., & Minderhoud, J. M. 1983, "Megadose Steroids in Severe Head-Injury - Results of A Prospective Double-Blind Clinical-Trial", *Journal of Neurosurgery*, vol. 58, no. 3, pp. 326-330.

Brott, T., Adams, H. P. Jr., Olinger, C. P., Marler, J. R., Barsan, W. G., Biller, J., Spilker, J., Holleran, R., Eberle, R., & Hertzberg, V. 1989, "Measurements of acute cerebral infarction: a clinical examination scale", *Stroke*, vol. 20, no. 7, pp. 864-870.

Bullock, M. R., Lyeth, B. G., & Muizelaar, I. P. 1999, "Current status of neuroprotection trials for traumatic brain injury: Lessons from animal models and clinical studies", *Neurosurgery*, vol. 45, no. 2, pp. 207-217.

Chesnut, R. M., Marshall, L. F., Klauber, M. R., Blunt, B. A., Baldwin, N., Eisenberg, H. M., Jane, J. A., Marmarou, A., & Foulkes, M. A. 1993, "The role of secondary brain injury in determining outcome from severe head injury", *Journal of Trauma*, vol. 34, no. 2, pp. 216-222.

Choi, S. C. & Bullock, R. 2001, "Design and statistical issues in multicenter trials of severe head injury", *Neurological Research*, vol. 23, no. 2-3, pp. 190-192.

Choi, S. C., Clifton, G. L., Marmarou, A., & Miller, E. R. 2002, "Misclassification and treatment effect on primary outcome measures in clinical trials of severe neurotrauma", *Journal of Neurotrauma*, vol. 19, no. 1, pp. 17-22.

Clifton, G. L., Miller, E. R., Choi, S. C., Levin, H. S., McCauley, S., Smith, K. R., Jr., Muizelaar, J. P., Wagner, F. C., Jr., Marion, D. W., Luerssen, T. G., Chesnut, R. M., & Schwartz, M. 2001, "Lack of effect of induction of hypothermia after acute brain injury", *New England Journal of Medicine*, vol. 344, no. 8, pp. 556-563.

Cooper, D. J., Myles, P. S., McDermott, F. T., Murray, L. J., Laidlaw, J., Cooper, G., Tremayne, A. B., Bernard, S. S., & Ponsford, J. 2004, "Prehospital hypertonic saline resuscitation of patients with hypotension and severe traumatic brain injury: a randomized controlled trial", *JAMA*, vol. 291, no. 11, pp. 1350-1357.

Copas, J. & Eguchi, S. 2010, "Likelihood for statistically equivalent models", *Journal of the Royal Statistical Society Series B-Statistical Methodology*, vol. 72, pp. 193-217.

Cruz, J., Minoja, G., & Okuchi, K. 2001, "Improving clinical outcomes from acute subdural hematomas with the emergency preoperative administration of high doses of mannitol: a randomized trial", *Neurosurgery*, vol. 49, no. 4, pp. 864-871.

Cruz, J., Minoja, G., & Okuchi, K. 2002, "Major clinical and physiological benefits of early high doses of mannitol for intraparenchymal temporal lobe hemorrhages with abnormal pupillary widening: A randomized trial", *Neurosurgery*, vol. 51, no. 3, pp. 628-637.

Dearden, N. M., Gibson, J. S., McDowall, D. G., Gibson, R. M., & Cameron, M. M. 1986, "Effect of high-dose dexamethasone on outcome from severe head injury", *Journal of Neurosurgery*, vol. 64, no. 1, pp. 81-88.

DeMets, D. L. 1987, "Methods for combining randomized clinical trials: strengths and limitations", *Statistics in Medicine*, vol. 6, no. 3, pp. 341-350.

den Hertog, H. M., van der Worp, H. B., van Gemert, H. M., Algra, A., Kappelle, L. J., van, G. J., Koudstaal, P. J., & Dippel, D. W. 2008, "Correction: PAIS: paracetamol (acetaminophen) in stroke; protocol for a randomized, double blind clinical trial. [ISCRTN74418480]", *BMC Cardiovascular Disorders*, vol. 8, p. 29.

den Hertog, H. M., van der Worp, H. B., van Gemert, H. M., Algra, A., Kappelle, L. J., van, G. J., Koudstaal, P. J., & Dippel, D. W. 2009, "The Paracetamol (Acetaminophen) In Stroke (PAIS) trial: a multicentre, randomised, placebo-controlled, phase III trial", *Lancet Neurology*, vol. 8, no. 5, pp. 434-440.

DerSimonian, R. & Laird, N. 1986, "Meta-analysis in clinical trials", *Controlled Clinical Trials*, vol. 7, no. 3, pp. 177-188.

Dickinson, K., Bunn, F., Wentz, R., Edwards, P., & Roberts, I. 2000, "Size and quality of randomised controlled trials in head injury: review of published studies", *British Medical Journal*, vol. 320, no. 7245, pp. 1308-1311.

Diener, H. C., Lees, K. R., Lyden, P., Grotta, J., Davalos, A., Davis, S. M., Shuaib, A., Ashwood, T., Wasiewski, W., Alderfer, V., Hardemark, H. G., & Rodichok, L. 2008a, "NXY-059 for the treatment of acute stroke: pooled analysis of the SAINT I and II Trials", *Stroke*, vol. 39, no. 6, pp. 1751-1758.

Diener, H. C., Sacco, R. L., Yusuf, S., Cotton, D., Ounpuu, S., Lawton, W. A., Palesch, Y., Martin, R. H., Albers, G. W., Bath, P., Bornstein, N., Chan, B. P., Chen, S. T., Cunha, L., Dahlof, B., De, K. J., Donnan, G. A., Estol, C., Gorelick, P., Gu, V., Hermansson, K., Hilbrich, L., Kaste, M., Lu, C., Machnig, T., Pais, P., Roberts, R., Skvortsova, V., Teal, P., Toni, D., VanderMaelen, C., Voigt, T., Weber, M., & Yoon, B. W. 2008b, "Effects of aspirin plus extended-release dipyridamole versus clopidogrel and telmisartan on disability and cognitive function after recurrent stroke in patients with ischaemic stroke in the Prevention Regimen for Effectively Avoiding Second Strokes (PRoFESS) trial: a double-blind, active and placebo-controlled study", *Lancet Neurology*, vol. 7, no. 10, pp. 875-884.

Doppenberg, E. M. R., Choi, S. C., & Bullock, R. 1997, "Clinical trials in traumatic brain injury - What can we learn from previous studies?", *Neuroprotective Agents - Third International Conference*, vol. 825, pp. 305-322.

Edwards, P., Arango, M., Balica, L., Cottingham, R., El-Sayed, H., Farrell, B., Fernandes, J., Gogichaisvili, T., Golden, N., Hartzenberg, B., Husain, M., Ulloa, M. I., Jerbi, Z., Khamis, H., Komolafe, E., Laloe, V., Lomas, G., Ludwig, S., Mazairac, G., Munoz Sanchez, M. L., Nasi, L., Olldashi, F., Plunkett, P., Roberts, I., Sandercock, P., Shakur, H., Soler, C., Stocker, R., Svoboda, P., Trenkler, S., Venkataramana, N. K., Wasserberg, J., Yates, D., & Yutthakasemsunt, S. 2005, "Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury-outcomes at 6 months", *Lancet*, vol. 365, no. 9475, pp. 1957-1959.

Feldmann, U. & Steudel, I. 2000, "Methods of ordinal classification applied to medical scoring systems", *Statistics in Medicine*, vol. 19, no. 4, pp. 575-586.

Fienberg, S. E. 1980, *The Analysis of Cross Classified Categorical Data* MIT Press, Cambridge, Mass.

Finfer, S. R. & Cohen, J. 2001, "Severe traumatic brain injury", *Resuscitation*, vol. 48, no. 1, pp. 77-90.

Ford, I. & Norrie, J. 2002, "The role of covariates in estimating treatment effects and risk in long-term clinical trials", *Statistics in Medicine*, vol. 21, no. 19, pp. 2899-2908.

Foulkes, M. A., Eisenberg, H. M., Jane, J. A., Marmarou, A., & Marshall, L. F. 1991, "The Traumatic Coma Data-Bank - Design, Methods, and Base-Line Characteristics", *Journal of Neurosurgery*, vol. 75, p. S8-S13.

Gaab, M. R., Trost, H. A., Alcantara, A., Karimi-Nejad, A., Moskopp, D., Schultheiss, R., Bock, W. J., Piek, J., Klinge, H., & Scheil, F. 1994, ""Ultrahigh" dexamethasone in acute brain injury. Results from a prospective randomized double-blind multicenter trial (GUDHIS). German Ultrahigh Dexamethasone Head Injury Study Group", *Zentralbl.Neurochir.*, vol. 55, no. 3, pp. 135-143.

Gail, M. H., Wieand, S., & Piantadosi, S. 1984, "Biased Estimates of Treatment Effect in Randomized Experiments with Nonlinear Regressions and Omitted Covariates", *Biometrika*, vol. 71, no. 3, pp. 431-444.

Geeganage, C. M., Tracy, M., Bath, M. W., & Bath, P. M. 2010, "Blood pressure reduction and cardiovascular prevention: meta-regression using ordered categorical (ordinal) event data", *Journal of Hypertension*, vol. 28, no. 10, pp. 1995-1999.

Gray, L. J., Bath, P. M., & Collier, T. 2009, "Should stroke trials adjust functional outcome for baseline prognostic factors?", *Stroke*, vol. 40, no. 3, pp. 888-894.

Grumme, T., Baethmann, A., Kolodziejczyk, D., Krimmer, J., Fischer, M., von Eisenhart, R. B., Pelka, R., Bennefeld, H., Pollauer, E., & Kostron, H. 1995, "Treatment of patients with severe head injury by triamcinolone: a prospective, controlled multicenter clinical trial of 396 cases", *Res.Exp.Med.(Berl)*, vol. 195, no. 4, pp. 217-229.

Guha, A. 2004, "Management of traumatic brain injury: some current evidence and applications", *Postgraduate Medical Journal*, vol. 80, no. 949, pp. 650-653.

GUSTO investigators 1993, "An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction", *N.Engl.J.Med.*, vol. 329, no. 10, pp. 673-682.

Hacke, W., Kaste, M., Fieschi, C., Toni, D., Lesaffre, E., von, K. R., Boysen, G., Bluhmki, E., Hoxter, G., & Mahagne, M. H. 1995, "Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke. The European Cooperative Acute Stroke Study (ECASS)", *JAMA*, vol. 274, no. 13, pp. 1017-1025.

Hacke, W., Kaste, M., Fieschi, C., von, K. R., Davalos, A., Meier, D., Larrue, V., Bluhmki, E., Davis, S., Donnan, G., Schneider, D., Diez-Tejedor, E., & Trouillas, P. 1998, "Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Second European-Australasian Acute Stroke Study Investigators", *Lancet*, vol. 352, no. 9136, pp. 1245-1251.

Harders, A., Kakarieka, A., Braakman, R., Hardenack, M., Schmieder, K., Trost, H. A., Hellwig, H., Buchholz, E. M., Klein, T., Peters, R., Zierski, J., Veelken, J., Gilsbach, J. M., Mayfrank, L., Bassiouni, H., Brawanski, A., Holzschuh, M., Hassler, W. E., Rohde, V., Ziebell, P., Emonds, N., Markakis, E., Kolenda, H., Zimmerer, B., Scharphuis, T., Eggert, H. R., Wilkowski, A., May, J. W., Faulhauer, K., Lauer, J., Paulus, J., Schoche, J., Raabe, A., Salger, D., Schibalski, G., Burkert, W., Rainov, N., Heidecke, V., Hamm, K., Grote, E. H., Buchholz, R., Morgalla, M., Meinig, G., Leyendecker, K., Schuerhoff, W., Wassmann, D., Moskopp, D., vonWild, K., Schutze, M., Schonmayr, R., Busch, C., Wallenfang, T., Suadicani, A., Fussler, H., Schakel, E. H., & Beneke, M. 1996, "Traumatic subarachnoid hemorrhage and its treatment with nimodipine", *Journal of Neurosurgery*, vol. 85, no. 1, pp. 82-89.

Harrell, F. E., Jr. 2001, "Regression Modeling Strategies: with applications to linear models, logistic regression, and survival analysis.," Spinger-Verlag, New York.

Harrell, F. E., Jr., Lee, K. L., & Mark, D. B. 1996, "Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors", *Stat.Med.*, vol. 15, no. 4, pp. 361-387.

Harrell, F. E., Jr. & Shih, Y. C. 2001, "Using full probability models to compute probabilities of actual interest to decision makers", *Int.J.Technol.Assess.Health Care*, vol. 17, no. 1, pp. 17-26.

Hatano, S. 1976, "Experience from a multicentre stroke register: a preliminary report", *Bulletin of the World Health Organization*, vol. 54, no. 5, pp. 541-553.

Hauck, W. W., Anderson, S., & Marcus, S. M. 1998, "Should we adjust for covariates in nonlinear regression analyses of randomized trials?", *Controlled Clinical Trials*, vol. 19, no. 3, pp. 249-256.

Helmy, A., Timofeev, I., & Hutchinson, P. J. 2010, "What is the purpose of statistical modelling in traumatic brain injury?", *Acta Neurochirurgica*, vol. 152, no. 11, pp. 2007-2008.

Hernandez, A. V., Steyerberg, E. W., Butcher, I., Mushkudiani, N., Taylor, G. S., Murray, G. D., Marmarou, A., Choi, S. C., Lu, J., Habbema, J. D. F., & Maas, A. I. R. 2006, "Adjustment for strong predictors of outcome in traumatic brain injury trials: 25% reduction in sample size requirements in the IMPACT study", *Journal of Neurotrauma*, vol. 23, no. 9, pp. 1295-1303.

Hernandez, A. V., Steyerberg, E. W., & Habbema, J. D. 2004, "Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements", *Journal of Clinical Epidemiology*, vol. 57, no. 5, pp. 454-460.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. 2003, "Measuring inconsistency in meta-analyses", *British Medical Journal*, vol. 327, no. 7414, pp. 557-560.

Hong, K. S. & Saver, J. L. 2009, "Quantifying the value of stroke disability outcomes: WHO global burden of disease project disability weights for each level of the modified Rankin Scale", *Stroke*, vol. 40, no. 12, pp. 3828-3833.

Hovda, D. A., Lee, S. M., Smith, M. L., Von, S. S., Bergsneider, M., Kelly, D., Shalmon, E., Martin, N., Caron, M., & Mazziotta, J. 1995, "The neurochemical and metabolic cascade following brain injury: moving from animal models to man", *Journal of Neurotrauma*, vol. 12, no. 5, pp. 903-906.

Hukkelhoven, C. W., Steyerberg, E. W., Habbema, J. D., Farace, E., Marmarou, A., Murray, G. D., Marshall, L. F., & Maas, A. I. 2005, "Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics", *Journal of Neurotrauma*, vol. 22, no. 10, pp. 1025-1039.

Ioannidis, J. P. A. 2006, "Evolution and translation of research findings: From to where?", *PLoS Clinical Trials*, vol. 1, no. 7.

Jennett, B. & Bond, M. 1975, "Assessment of outcome after severe brain damage", *Lancet*, vol. i, no. 7905, pp. 480-484.

Jiang, J. Y., Xu, W., Li, W. P., Gao, G. Y., Bao, Y. H., Liang, Y. M., & Luo, Q. Z. 2006, "Effect of long-term mild hypothermia or short-term mild hypothermia on outcome of patients with severe traumatic brain injury", *Journal of Cerebral Blood Flow & Metabolism*, vol. 26, no. 6, pp. 771-776.

Jiang, J. Y., Xu, W., Li, W. P., Xu, W. H., Zhang, J., Bao, Y. H., Ying, Y. H., & Luo, Q. Z. 2005, "Efficacy of standard trauma craniectomy for refractory intracranial hypertension with severe traumatic brain injury: a multicenter, prospective, randomized controlled study", *Journal of Neurotrauma*, vol. 22, no. 6, pp. 623-628.

Justice, A. C., Covinsky, K. E., & Berlin, J. A. 1999, "Assessing the generalizability of prognostic information", *Annals of Internal Medicine*, vol. 130, no. 6, pp. 515-524.

Kannus, P., Palvanen, M., & Niemi, S. 2001, "Time trends in severe head injuries among elderly Finns", *JAMA*, vol. 286, no. 6, pp. 673-674.

Kasner, S. E. 2006, "Clinical interpretation and use of stroke scales", *Lancet Neurology*, vol. 5, no. 7, pp. 603-612.

L'Abbe, K. A., Detsky, A. S., & O'Rourke, K. 1987, "Meta-analysis in clinical research", *Annals of Internal Medicine*, vol. 107, no. 2, pp. 224-233.

Lai, S. M. & Duncan, P. W. 2001, "Stroke recovery profile and the Modified Rankin assessment", *Neuroepidemiology*, vol. 20, no. 1, pp. 26-30.

Lees, K. R., Asplund, K., Carolei, A., Davis, S. M., Diener, H. C., Kaste, M., Orgogozo, J. M., & Whitehead, J. 2000, "Glycine antagonist (gavestinel) in neuroprotection (GAIN International) in patients with acute stroke: a randomised controlled trial. GAIN International Investigators", *Lancet*, vol. 355, no. 9219, pp. 1949-1954.

Lees, K. R., Davalos, A., Davis, S. M., Diener, H. C., Grotta, J., Lyden, P., Shuaib, A., Ashwood, T., Hardemark, H. G., Wasiewski, W., Emeribe, U., & Zivin, J. A. 2006a, "Additional outcomes and subgroup analyses of NXY-059 for acute ischemic stroke in the SAINT I trial", *Stroke*, vol. 37, no. 12, pp. 2970-2978.

Lees, K. R., Zivin, J. A., Ashwood, T., Davalos, A., Davis, S. M., Diener, H. C., Grotta, J., Lyden, P., Shuaib, A., Hardemark, H. G., & Wasiewski, W. W. 2006b, "NXY-059 for acute ischemic stroke", *New England Journal of Medicine*, vol. 354, no. 6, pp. 588-600.

Levin, H. S., Boake, C., Song, J., McCauley, S., Contant, C., Diaz-Marchan, P., Brundage, S., Goodman, H., & Kotrla, K. J. 2001, "Validity and sensitivity to change of the Extended Glasgow Outcome Scale in mild to moderate traumatic brain injury", *Journal of Neurotrauma*, vol. 18, no. 6, pp. 575-584.

Levy, D. E., del Zoppo, G. J., Demaerschalk, B. M., Demchuk, A. M., Diener, H. C., Howard, G., Kaste, M., Pancioli, A. M., Ringelstein, E. B., Spatareanu, C., & Wasiewski, W. W. 2009, "Ancrod in acute ischemic stroke: results of 500 subjects beginning treatment within 6 hours of stroke onset in the ancrod stroke program", *Stroke*, vol. 40, no. 12, pp. 3796-3803.

Liang, K. Y. & Zeger, S. L. 1986, "Longitudinal Data-Analysis Using Generalized Linear-Models", *Biometrika*, vol. 73, no. 1, pp. 13-22.

Lindley, R. I., Waddell, F., Livingstone, M., Sandercock, P., Dennis, M. S., Slattery, J., Smith, B., & Warlow, C. 1994, "Can Simple Questions Assess Outcome After Stroke", *Cerebrovascular Diseases*, vol. 4, no. 4, pp. 314-324.

Lingsma, H. F., Roozenbeek, B., Steyerberg, E. W., Murray, G. D., & Maas, A. I. 2010, "Early prognosis in traumatic brain injury: from prophecies to predictions", *Lancet Neurology*, vol. 9, no. 5, pp. 543-554.

Little, R. J. A. & Rubin, D. B. 1987, *Statistical Analysis with Missing Data* Wiley, New York.

Liu, I. & Agresti, A. 2005, "The analysis of ordered categorical data: An overview and a survey of recent developments", *Test*, vol. 14, no. 1, pp. 1-30.

Lu, J., Marmarou, A., Lapane, K., Turf, E., & Wilson, L. 2010, "A method for reducing misclassification in the extended Glasgow Outcome Score", *Journal of Neurotrauma*, vol. 27, no. 5, pp. 843-852.

Lu, J., Murray, G. D., Steyerberg, E. W., Butcher, I., McHugh, G. S., Lingsma, H., Mushkudiani, N., Choi, S., Maas, A. I., & Marmarou, A. 2008, "Effects of Glasgow Outcome Scale misclassification on traumatic brain injury clinical trials", *Journal of Neurotrauma*, vol. 25, no. 6, pp. 641-651.

Lu, L. Q., Jiang, J. Y., Yu, M. K., Hou, L. J., Chen, Z. G., Zhang, G. J., & Zhu, C. 2003, "Standard large trauma craniotomy for severe traumatic brain injury", *Chinese Journal of Traumatology*, vol. 6, no. 5, pp. 302-304.

Luukinen, H., Herala, M., Koski, K., Kivela, S. L., & Honkanen, R. 1999, "Rapid increase of fall-related severe head injuries with age among older people: a population-based study", *Journal of the American Geriatrics Society*, vol. 47, no. 12, pp. 1451-1452.

Maas, A. I., Murray, G., Henney, H., III, Kassem, N., Legrand, V., Mangelus, M., Muizelaar, J. P., Stocchetti, N., & Knoller, N. 2006, "Efficacy and safety of dexanabinol in severe traumatic brain injury: results of a phase III randomised, placebo-controlled, clinical trial", *Lancet Neurology*, vol. 5, no. 1, pp. 38-45.

Maas, A. I., Roozenbeek, B., & Manley, G. T. 2010, "Clinical trials in traumatic brain injury: past experience and current developments", *Neurotherapeutics*, vol. 7, no. 1, pp. 115-126.

Maas, A. I., Steyerberg, E. W., Murray, G. D., Bullock, R., Baethmann, A., Marshall, L. F., & Teasdale, G. M. 1999, "Why have recent trials of neuroprotective agents in head injury failed to show convincing efficacy? A pragmatic analysis and theoretical considerations", *Neurosurgery*, vol. 44, no. 6, pp. 1286-1298.

Maas, A. I. R. 2000, "Assessment of agents for treatment of head injury", *CNS Drugs*, vol. 13, no. 2, pp. 139-154.

Maas, A. I. R., Steyerberg, E. W., Butcher, I., Dammers, R., Lu, J., Marmarou, A., Mushkudiani, N. A., McHugh, G. S., & Murray, G. D. 2007, "Prognostic value of computerized tomography scan characteristics in traumatic brain injury: Results from the IMPACT study", *Journal of Neurotrauma*, vol. 24, no. 2, pp. 303-314.

Machado, S. G., Murray, G. D., & Teasdale, G. M. 1999, "Evaluation of designs for clinical trials of neuroprotective agents in head injury. European Brain Injury Consortium", *Journal of Neurotrauma*, vol. 16, no. 12, pp. 1131-1138.

Mahoney, F. I. & Barthel, D. W. 1965, "Functional Evaluation: The Barthel Index", *Maryland State Medical Journal*, vol. 14, pp. 61-65.

Marmarou, A., Lu, J., Butcher, I., McHugh, G. S., Murray, G. D., Steyerberg, E. W., Mushkudiani, N. A., Choi, S., & Maas, A. I. R. 2007, "Prognostic value of the Glasgow coma scale and pupil reactivity in traumatic brain injury assessed pre-hospital and on enrollment: An IMPACT analysis", *Journal of Neurotrauma*, vol. 24, no. 2, pp. 270-280.

Marmarou, A., Nichols, J., Burgess, J., Newell, D., Troha, J., Burnham, D., & Pitts, L. 1999, "Effects of the bradykinin antagonist Bradycor (TM) (deltibant, CP-1027) in severe traumatic brain injury: Results of a multi- center, randomized, placebo-controlled trial", *Journal of Neurotrauma*, vol. 16, no. 6, pp. 431-444.

Marshall, L. F., Maas, A. I., Marshall, S. B., Bricolo, A., Fearnside, M., Iannotti, F., Klauber, M. R., Lagarrigue, J., Lobato, R., Persson, L., Pickard, J. D., Piek, J., Servadei, F., Wellis, G. N., Morris, G. F., Means, E. D., & Musch, B. 1998, "A multicenter trial on the efficacy of using tirilazad mesylate in cases of head injury", *Journal of Neurosurgery*, vol. 89, no. 4, pp. 519-525.

Marshall, L. F., Marshall, S. B., Klauber, M. R., Clark, M. V., Eisenberg, H. M., Jane, J. A., Luerssen, T. G., Marmarou, A., & Foulkes, M. A. 1991, "A New Classification of Head-Injury Based on Computerized- Tomography", *Journal of Neurosurgery*, vol. 75, p. S14-S20.

McCullagh, P. 1978, "A Class of Parametric Models for Analysis of Square Contingency-Tables with Ordered Categories", *Biometrika*, vol. 65, no. 2, pp. 413-418.

McCullagh, P. 1980, "Regression-Models for Ordinal Data", *Journal of the Royal Statistical Society Series B- Methodological*, vol. 42, no. 2, pp. 109-142.

McDonald, A. M., Knight, R. C., Campbell, M. K., Entwistle, V. A., Grant, A. M., Cook, J. A., Elbourne, D. R., Francis, D., Garcia, J., Roberts, I., & Snowdon, C. 2006, "What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies", *Trials*, vol. 7, p. 9.

McHugh, G. S., Butcher, I., Steyerberg, E. W., Lu, J., Mushkudiani, N., Marmarou, A., Maas, A. I. R., & Murray, G. D. 2007a, "Statistical approaches to the univariate prognostic analysis of the IMPACT database on traumatic brain injury", *Journal of Neurotrauma*, vol. 24, no. 2, pp. 251-258.

McHugh, G. S., Butcher, I., Steyerberg, E. W., Marmarou, A., Lu, J., Lingsma, H. F., Weir, J., Maas, A. I., & Murray, G. D. 2010, "A simulation study evaluating approaches to the analysis of ordinal outcome data in randomized controlled trials in traumatic brain injury: results from the IMPACT Project", *Clinical Trials*, vol. 7, no. 1, pp. 44-57.

McHugh, G. S., Engel, D. C., Butcher, I., Steyerberg, E. W., Lu, J., Mushkudiani, N., Hernandez, A. V., Marmarou, A., Maas, A. I. R., & Murray, G. D. 2007b, "Prognostic value of secondary insults in traumatic brain injury: results from the IMPACT study", *Journal of Neurotrauma*, vol. 24, no. 2, pp. 287-293.

Mendelow, A. D., Gregson, B. A., Fernandes, H. M., Murray, G. D., Teasdale, G. M., Hope, D. T., Karimi, A., Shaw, M. D., & Barer, D. H. 2005, "Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): a randomised trial", *Lancet*, vol. 365, no. 9457, pp. 387-397.

Menon, D. K. 2009, "Unique challenges in clinical trials in traumatic brain injury", *Critical Care Medicine*, vol. 37, no. 1 Suppl, p. S129-S135.

Mishra, N. K., Lyden, P., Grotta, J. C., & Lees, K. R. 2010, "Thrombolysis is associated with consistent functional improvement across baseline stroke severity: a comparison of outcomes in patients from the Virtual International Stroke Trials Archive (VISTA)", *Stroke*, vol. 41, no. 11, pp. 2612-2617.

Mishra, N. K., Shuaib, A., Lyden, P., Diener, H. C., Grotta, J., Davis, S., Davalos, A., Ashwood, T., Wasiewski, W., Lees, K. R., & Stroke Acute Ischemic NXY Treatment 2011, "Home time is extended in patients with ischemic stroke who receive thrombolytic therapy: a validation study of home time as an outcome measure", *Stroke*, vol. 42, no. 4, pp. 1046-1050.

Morris, G. F., Bullock, R., Marshall, S. B., Marmarou, A., Maas, A., & Marshall, L. F. 1999, "Failure of the competitive N-methyl-D-aspartate antagonist Selfotel (CGS 19755) in the treatment of severe head injury: results of two Phase III clinical trials", *Journal of Neurosurgery*, vol. 91, no. 5, pp. 737-743.

Muizelaar, J. P., Marmarou, A., Ward, J. D., Kontos, H. A., Choi, S. C., Becker, D. P., Gruemer, H., & Young, H. F. 1991, "Adverse-Effects of Prolonged Hyperventilation in Patients with Severe Head-Injury - A Randomized Clinical-Trial", *Journal of Neurosurgery*, vol. 75, no. 5, pp. 731-739.

Murray, G. D., Barer, D., Choi, S., Fernandes, H., Gregson, B., Lees, K. R., Maas, A. I., Marmarou, A., Mendelow, A. D., Steyerberg, E. W., Taylor, G. S., Teasdale, G. M., & Weir, C. J. 2005, "Design and Analysis of Phase III Trials with Ordered Outcome Scales: The Concept of the Sliding Dichotomy", *Journal of Neurotrauma*, vol. 22, no. 5, pp. 511-517.

Murray, G. D., Butcher, I., McHugh, G. S., Lu, J., Mushkudiani, N. A., Maas, A. I. R., Marmarou, A., & Steyerberg, E. W. 2007, "Multivariable prognostic analysis in traumatic brain injury: Results from the IMPACT study", *Journal of Neurotrauma*, vol. 24, no. 2, pp. 329-337.

Murray, G. D. & Teasdale, G. M. 2000, "Quality of randomised controlled trials in head injury - Trials in head injury are more complex than review suggests", *British Medical Journal*, vol. 321, no. 7270, p. 1223.

Murray, G. D., Teasdale, G. M., Braakman, R., Cohadon, F., Dearden, M., Iannotti, F., Karimi, A., Lapierre, F., Maas, A., Ohman, J., Persson, L., Servadei, F., Stocchetti, N., Trojanowski, T., & Unterberg, A. 1999a, "The European Brain Injury Consortium survey of head injuries", *Acta Neurochirurgica*, vol. 141, no. 3, pp. 223-236.

Murray, L. S., Teasdale, G. M., Murray, G. D., Miller, D. J., Pickard, J. D., & Shaw, M. D. 1999b, "Head injuries in four British neurosurgical centres", *British Journal of Neurosurgery*, vol. 13, no. 6, pp. 564-569.

Narayan, R. K., Michel, M. E., Ansell, B., Baethmann, A., Biegon, A., Bracken, M. B., Bullock, M. R., Choi, S. C., Clifton, G. L., Contant, C. F., Coplin, W. M., Dietrich, W. D., Ghajar, J., Grady, S. M., Grossman, R. G., Hall, E. D., Heetderks, W., Hovda, D. A., Jallo, J., Katz, R. L., Knoller, N., Kochanek, P. M., Maas, A. I., Majde, J., Marion, D. W., Marmarou, A., Marshall, L. F., McIntosh, T. K., Miller, E., Mohberg, N., Muizelaar, J. P., Pitts, L. H., Quinn, P., Riesenfeld, G., Robertson, C. S., Strauss, K. I., Teasdale, G., Temkin, N., Tuma, R., Wade, C., Walker, M. D., Weinrich, M., Whyte, J., Wilberger, J., Young, A. B., & Yurkewicz, L. 2002, "Clinical trials in head injury", *Journal of Neurotrauma*, vol. 19, no. 5, pp. 503-557.

O'Collins, V. E., Macleod, M. R., Donnan, G. A., Horky, L. L., van der Worp, B. H., & Howells, D. W. 2006, "1,026 experimental treatments in acute stroke", *Annals of Neurology*, vol. 59, no. 3, pp. 467-477.

Peterson, B. & Harrell, F. E. 1990, "Partial Proportional Odds Models for Ordinal Response Variables", *Applied Statistics-Journal of the Royal Statistical Society Series C*, vol. 39, no. 2, pp. 205-217.

Pettigrew, L. E., Wilson, J. T., & Teasdale, G. M. 1998, "Assessing disability after head injury: improved use of the Glasgow Outcome Scale", *Journal of Neurosurgery*, vol. 89, no. 6, pp. 939-943.

Pocock, S. J., Assmann, S. E., Enos, L. E., & Kasten, L. E. 2002, "Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems", *Statistics in Medicine*, vol. 21, no. 19, pp. 2917-2930.

Poon, W. Y. 2004, "A latent normal distribution model for analysing ordinal responses with applications in meta-analysis", *Statistics in Medicine*, vol. 23, no. 14, pp. 2155-2172.

Pulkstenis, E. & Robinson, T. J. 2004, "Goodness-of-fit tests for ordinal response regression models", *Statistics in Medicine*, vol. 23, no. 6, pp. 999-1014.

Qureshi, A. I., Hutson, A. D., Harbaugh, R. E., Stieg, P. E., & Hopkins, L. N. 2004, "Methods and design considerations for randomized clinical trials evaluating surgical or endovascular treatments for cerebrovascular diseases", *Neurosurgery*, vol. 54, no. 2, pp. 248-264.

Rappaport, M., Hall, K. M., Hopkins, K., Belleza, T., & Cope, D. N. 1982, "Disability rating scale for severe head trauma: coma to community", *Archives of Physical Medicine and Rehabilitation*, vol. 63, no. 3, pp. 118-123.

Relton, C., Torgerson, D., O'Cathain, A., & Nicholl, J. 2010, "Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design", *British Medical Journal*, vol. 340, p. c1066.

Robertson, C. S., Valadka, A. B., Hannay, H. J., Contant, C. F., Gopinath, S. P., Cormio, M., Uzura, M., & Grossman, R. G. 1999, "Prevention of secondary ischemic insults after severe head injury", *Critical Care Medicine*, vol. 27, no. 10, pp. 2086-2095.

Robinson, L. D. & Jewell, N. P. 1991, "Some Surprising Results About Covariate Adjustment in Logistic-Regression Models", *International Statistical Review*, vol. 59, no. 2, pp. 227-240.

Rockswold, G. L., Ford, S. E., Anderson, D. C., Bergman, T. A., & Sherman, R. E. 1992, "Results of a prospective randomized trial for treatment of severely brain-injured patients with hyperbaric oxygen", *Journal of Neurosurgery*, vol. 76, no. 6, pp. 929-934.

Roozenbeek, B., Lingsma, H. F., Perel, P., Edwards, P., Roberts, I., Murray, G. D., Maas, A. I., & Steyerberg, E. W. 2011, "The added value of ordinal analysis in clinical trials: an example in traumatic brain injury", *Critical Care*, vol. 15, no. 3, p. R127.

Roozenbeek, B., Maas, A. I., Lingsma, H. F., Butcher, I., Lu, J., Marmarou, A., McHugh, G. S., Weir, J., Murray, G. D., & Steyerberg, E. W. 2009a, "Baseline characteristics and statistical power in randomized controlled trials: selection, prognostic targeting, or covariate adjustment?", *Critical Care Medicine*, vol. 37, no. 10, pp. 2683-2690.

Roozenbeek, B., Maas, A. I., Marmarou, A., Butcher, I., Lingsma, H. F., Lu, J., McHugh, G. S., Murray, G. D., & Steyerberg, E. W. 2009b, "The influence of enrollment criteria on recruitment and outcome distribution in traumatic brain injury studies: results from the impact study", *Journal of Neurotrauma*, vol. 26, no. 7, pp. 1069-1075.

Royston, P. 2000, "A strategy for modelling the effect of a continuous covariate in medicine and epidemiology", *Statistics in Medicine*, vol. 19, no. 14, pp. 1831-1847.

Rubin, D. B. 1987, *Multiple Imputation for Nonresponse in Surveys* Wiley, New York.

Saatman, K. E., Duhaime, A. C., Bullock, R., Maas, A. I., Valadka, A., & Manley, G. T. 2008, "Classification of traumatic brain injury for targeted therapies", *Journal of Neurotrauma*, vol. 25, no. 7, pp. 719-738.

Sandset, E. C., Bath, P. M., Boysen, G., Jatuzis, D., Korv, J., Luders, S., Murray, G. D., Richter, P. S., Roine, R. O., Terent, A., Thijs, V., & Berge, E. 2011, "The angiotensin-receptor blocker candesartan for treatment of acute stroke (SCAST): a randomised, placebo-controlled, double-blind trial", *Lancet*, vol. 377, no. 9767, pp. 741-750.

Saul, T. G., Ducker, T. B., Salcman, M., & Carro, E. 1981, "Steroids in severe head injury: A prospective randomized clinical trial", *Journal of Neurosurgery*, vol. 54, no. 5, pp. 596-600.

Saver, J. L. 2007, "Novel end point analytic techniques and interpreting shifts across the entire range of outcome scales in acute stroke trials", *Stroke*, vol. 38, no. 11, pp. 3055-3062.

Saver, J. L. & Yafeh, B. 2007, "Confirmation of tPA treatment effect by baseline severity-adjusted end point reanalysis of the NINDS-tPA stroke trials", *Stroke*, vol. 38, no. 2, pp. 414-416.

Schabitz, W. R. & Fisher, M. 2006, "Perspectives on neuroprotective stroke therapy", *Biochemical Society Transactions*, vol. 34, no. Pt 6, pp. 1271-1276.

Senn, S. & Julious, S. 2009, "Measurement in clinical trials: a neglected issue for statisticians?", *Statistics in Medicine*, vol. 28, no. 26, pp. 3189-3209.

Shuaib, A., Lees, K. R., Lyden, P., Grotta, J., Davalos, A., Davis, S. M., Diener, H. C., Ashwood, T., Wasiewski, W. W., & Emeribe, U. 2007, "NXY-059 for the treatment of acute ischemic stroke", *New England Journal of Medicine*, vol. 357, no. 6, pp. 562-571.

Solling, C., Hjort, N., Ashkanian, M., Ostergaard, L., & Andersen, G. 2009, "Safety and efficacy of MRI-based selection for recombinant tissue plasminogen activator treatment: responder analysis of outcome in the 3-hour time window", *Cerebrovascular Diseases*, vol. 27, no. 3, pp. 223-229.

Song, H. S., Back, J. H., Jin, D. K., Chung, P. W., Moon, H. S., Suh, B. C., Kim, Y. B., Kim, B. M., Woo, H. Y., Lee, Y. T., & Park, K. Y. 2008, "Cardiac troponin T elevation after stroke: relationships between elevated serum troponin T, stroke location, and prognosis", *Journal of Clinical Neurology*, vol. 4, no. 2, pp. 75-83.

Steyerberg, E., Mushkudiani, N., Perel, P., Butcher, I., Lu, J., McHugh, G. S., Murray, G. D., Marmarou, A., Roberts, I., Habbema, J. D., & Maas, A. I. Coefficients from model fitting. http://www.tbi-impact.org . 2011.
Ref Type: Electronic Citation

Steyerberg, E. T. W. & Eijkemans, M. U. J. C. 2004, "Heterogeneity bias: The difference between adjusted and unadjusted effects", *Medical Decision Making*, vol. 24, no. 1, pp. 102-104.

Steyerberg, E. W., Bossuyt, P. M., & Lee, K. L. 2000, "Clinical trials in acute myocardial infarction: should we adjust for baseline characteristics?", *American Heart Journal*, vol. 139, no. 5, pp. 745-751.

Steyerberg, E. W., Mushkudiani, N., Perel, P., Butcher, I., Lu, J., McHugh, G. S., Murray, G. D., Marmarou, A., Roberts, I., Habbema, J. D., & Maas, A. I. 2008, "Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics", *PLoS Medicine*, vol. 5, no. 8, p. e165.

Stiger, T. R., Barnhart, H. X., & Williamson, J. M. 1999, "Testing proportionality in the proportional odds model fitted with GEE", *Statistics in Medicine*, vol. 18, no. 11, pp. 1419-1433.

Stone, C. J. 1986, "Generalized additive models", *Statistical Science*, vol. 1, no. 3, pp. 312-314.

Stone, C. J. & Koo, C. Y. 1985, "Additive splines in statistics," in *Proceedings of the Statistical Computing Section ASA*, pp. 45-48.

Teasdale, G. & Jennett, B. 1974, "Assessment of coma and impaired consciousness. A practical scale", *Lancet*, vol. 2, no. 7872, pp. 81-84.

Teasdale, G. M., Maas, A., Iannotti, F., Ohman, J., & Unterberg, A. 1999, "Challenges in translating the efficacy of neuroprotective agents in experimental models into knowledge of clinical benefits in head injured patients", *Acta Neurochirurgica Supplement*, vol. 73, pp. 111-116.

Teasdale, G. M., Pettigrew, L. E. L., Wilson, J. T. L., Murray, G., & Jennett, B. 1998, "Analyzing outcome of treatment of severe head injury: A review and update

on advancing the use of the Glasgow Outcome Scale", *Journal of Neurotrauma*, vol. 15, no. 8, pp. 587-597.

Temkin, N. R., Anderson, G. D., Winn, H. R., Ellenbogen, R. G., Britz, G. W., Schuster, J., Lucas, T., Newell, D. W., Mansfield, P. N., Machamer, J. E., Barber, J., & Dikmen, S. S. 2007, "Magnesium sulfate for neuroprotection after traumatic brain injury: a randomised controlled trial", *Lancet Neurology*, vol. 6, no. 1, pp. 29-38.

The European Study Group on Nimodipine in Severe Head Injury 1994, "A multicenter trial of the efficacy of nimodipine on outcome after severe head injury.", *Journal of Neurosurgery*, vol. 80, no. 5, pp. 797-804.

The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group 1995, "Tissue plasminogen activator for acute ischemic stroke.", *New England Journal of Medicine*, vol. 333, no. 24, pp. 1581-1587.

The Optimising Analysis of Stroke Trials (OAST) collaboration 2008, "Calculation of sample size for stroke trials assessing functional outcome: comparison of binary and ordinal approaches", *International Journal of Stroke*, vol. 3, no. 2, pp. 78-84.

The Publications Committee for the Trial of ORG 10172 in Acute Stroke Treatment (TOAST) Investigators 1998, "Low molecular weight heparinoid, ORG 10172 (danaparoid), and outcome after acute ischemic stroke: a randomized controlled trial.", *JAMA*, vol. 279, no. 16, pp. 1265-1272.

Thomassen, L., Waje-Andreassen, U., Naess, H., Elvik, M. K., & Russell, D. 2005, "Long-term effect of intravenous thrombolytic therapy in acute stroke: responder analysis versus uniform analysis of excellent outcome", *Cerebrovascular Diseases*, vol. 20, no. 6, pp. 470-474.

Thompson, S. G. & Pocock, S. J. 1991, "Can meta-analyses be trusted?", *Lancet*, vol. 338, no. 8775, pp. 1127-1130.

Tsodikov, A., Hasenclever, D., & Loeffler, M. 1998, "Regression with bounded outcome score: evaluation of power by bootstrap and simulation in a chronic myelogenous leukaemia clinical trial", *Statistics in Medicine*, vol. 17, no. 17, pp. 1909-1922.

Van Beek, J. G. M., Mushkudiani, N. A., Steyerberg, E. W., Butcher, I., McHugh, G. S., Lu, J., Marmarou, A., Murray, G. D., & Maas, A. I. R. 2007, "Prognostic value of admission laboratory parameters in traumatic brain injury: Results from the IMPACT study", *Journal of Neurotrauma*, vol. 24, no. 2, pp. 315-328.

van Buuren, S., Boshuizen, H. C., & Knook, D. L. 1999, "Multiple imputation of missing blood pressure covariates in survival analysis", *Statistics in Medicine*, vol. 18, no. 6, pp. 681-694.

van Swieten, J. C., Koudstaal, P. J., Visser, M. C., Schouten, H. J., & van Gijn, J. 1988, "Interobserver agreement for the assessment of handicap in stroke patients", *Stroke*, vol. 19, no. 5, pp. 604-607.

Vink, R. & Bullock, M. R. 2010, "Traumatic brain injury: therapeutic challenges and new directions", *Neurotherapeutics*, vol. 7, no. 1, pp. 1-2.

Wardlaw, J. M., Sandercock, P. A., Warlow, C. P., & Lindley, R. I. 2000, "Trials of thrombolysis in acute ischemic stroke: does the choice of primary outcome measure really matter?", *Stroke*, vol. 31, no. 5, pp. 1133-1135.

Weir, J., Steyerberg, E. W., Butcher, I., Lu, J., Lingsma, H. F., McHugh, G. S., Roozenbeek, B., Maas, A. I., & Murray, G. D. 2012, "Does the Extended Glasgow Outcome Scale add value to the conventional Glasgow Outcome Scale?", *Journal of Neurotrauma*, vol. 29, no. 1, pp. 53-58.

White, I. R., Royston, P., & Wood, A. M. 2011, "Multiple imputation using chained equations: Issues and guidance for practice", *Statistics in Medicine*, vol. 30, no. 4, pp. 377-399.

Whitehead, A., Omar, R. Z., Higgins, J. P. T., Savaluny, E., Turner, R. M., & Thompson, S. G. 2001, "Meta-analysis of ordinal outcomes using individual patient data", *Statistics in Medicine*, vol. 20, no. 15, pp. 2243-2260.

Whitehead, A. & Whitehead, J. 1991, "A general parametric approach to the meta-analysis of randomized clinical trials", *Statistics in Medicine*, vol. 10, no. 11, pp. 1665-1677.

Whitehead, J. 1993, "Sample size calculations for ordered categorical data", *Statistics in Medicine*, vol. 12, no. 24, pp. 2257-2271.

Wilson, J. T., Slieker, F. J., Legrand, V., Murray, G., Stocchetti, N., & Maas, A. I. 2007, "Observer variation in the assessment of outcome in traumatic brain injury: experience from a multicenter, international randomized clinical trial", *Neurosurgery*, vol. 61, no. 1, pp. 123-128.

Wolf, A. L., Levi, L., Marmarou, A., Ward, J. D., Muizelaar, P. J., Choi, S., Young, H., Rigamonti, D., & Robinson, W. L. 1993, "Effect of THAM upon outcome in severe head injury: a randomized prospective clinical trial", *Journal of Neurosurgery*, vol. 78, no. 1, pp. 54-59.

Xiao, G., Wei, J., Yan, W., Wang, W., & Lu, Z. 2008, "Improved outcomes from the administration of progesterone for patients with acute severe traumatic brain injury: a randomized controlled trial", *Critical Care*, vol. 12, no. 2, p. R61.

Yoo, S. H., Kim, J. S., Kwon, S. U., Yun, S. C., Koh, J. Y., & Kang, D. W. 2008, "Undernutrition as a predictor of poor clinical outcomes in acute ischemic stroke patients", *Archives of Neurology*, vol. 65, no. 1, pp. 39-43.

Young, B., Runge, J. W., Waxman, K. S., Harrington, T., Wilberger, J., Muizelaar, J. P., Boddy, A., & Kupiec, J. W. 1996, "Effects of pegorgotein on neurologic outcome of patients with severe head injury - A multicenter, randomized controlled trial", *JAMA*, vol. 276, no. 7, pp. 538-543.

Young, F. B., Lees, K. R., & Weir, C. J. 2003, "Strengthening acute stroke trials through optimal use of disability end points", *Stroke*, vol. 34, no. 11, pp. 2676-2680.

Young, F. B., Lees, K. R., & Weir, C. J. 2005, "Improving trial power through use of prognosis-adjusted end points", *Stroke*, vol. 36, no. 3, pp. 597-601.

Yurkewicz, L., Weaver, J., Bullock, M. R., & Marshall, L. F. 2005, "The effect of the selective NMDA receptor antagonist traxoprodil in the treatment of traumatic brain injury", *Journal of Neurotrauma*, vol. 22, no. 12, pp. 1428-1443.

Yusuf, S., Collins, R., & Peto, R. 1984, "Why do We Need Some Large, Simple Randomized Trials", *Statistics in Medicine*, vol. 3, no. 4, pp. 409-420.

Zhi, D., Zhang, S., & Lin, X. 2003, "Study on therapeutic mechanism and clinical effect of mild hypothermia in patients with severe head injury", *Surgical Neurology*, vol. 59, no. 5, pp. 381-385.

Zingmark, P. H., Ekblom, M., Odergren, T., Ashwood, T., Lyden, P., Karlsson, M. O., & Jonsson, E. N. 2003, "Population pharmacokinetics of clomethiazole and its effect on the natural course of sedation in acute stroke patients", *British Journal of Clinical Pharmacology*, vol. 56, no. 2, pp. 173-183.

9 Appendices

9.1 Appendix A Scales

Table 9-1 Glasgow Outcome Scale

Item	Description
Good recovery	Resumption of normal life despite minor deficits
Moderately disabled	Disabled but independent
Severely disabled	Conscious but disabled. Dependent for daily support
Vegetative state	Minimal responsiveness
Death	Dead

Table 9-2 Glasgow Coma Scale Components

Item	Description
Eye opening	
1	None: no eye opening at all
2	To painful stimulus, should not be applied to the face
3	To speech – not necessarily a request for eye opening
4	Spontaneous – already open with blinking (normal)
Verbal response	
1	None: no vocalisation at all
2	Incomprehensible sounds – grunts and groans but no words
3	Inappropriate words – recognisable words are produced but does not converse or answer
4	Confused – converses and answers questions but is confused
5	Orientated – knows name, age, place date etc.

Item	Description
Motor	
1	None: no movement elicited at all
2	Extensor response – extensor response posture in response to stimulus
3	Abnormal flexion – flexor posture in response to stimulus
4	Withdraws – pulls away from painful stimulus but does not localise
5	Localises - consistent purposeful motion towards a painful stimulus applied at different locations
6	Obeys – moves limb to command and pain is not required

Table 9-3 Marshall CT classification[11]

Classification	Description
Diffuse Injury I (no visible pathology)	No visible intracranial pathology seen on CT scan
Diffuse Injury II	Cisterns present with midline shift 0-5mm and/or no lesion densities present, no-high or mixed density lesion>25cc
Diffuse Injury III (swelling)	Cisterns compressed or absent with midline shift 0-5mm. No high or mixed density lesion >25cc
Diffuse Injury IV (shift)	Midline shift >5mm. No high or mixed density lesion >25cc
Evacuated mass lesion	Any lesion surgically evacuated.
Non-evacuated mass lesion	High or mixed density lesion >25cc not surgically evacuated

[11] Taken from (Marshall et al. 1991)

9.2 Appendix B Publications

The following publications are included both with publishers' and co-authors' permission.

1) A simulation study evaluating approaches to the analysis of ordinal outcome data in randomized controlled trials in traumatic brain injury: results from the IMPACT Project
McHugh GS, Butcher I, Steyerberg EW, Marmarou A, Lu J, Lingsma HF, Weir J, Maas AIR, Murray GD
Clinical Trials 2010; 7: 44-57

2) Statistical approaches to the univariate prognostic analysis of the IMPACT database on traumatic brain injury
McHugh GS, Butcher I, Steyerberg EW, Lu J, Mushkudiani N, Marmarou A, Maas AIR, Murray GD
Journal of Neurotrauma 2007, 24(2): 251-258

3) The Prognostic Value of Secondary Insults in TBI: Results from the IMPACT study
McHugh GS, Engel DC, Butcher I, Steyerberg EW, Lu J, Mushkudiani N, Hernandez A, Marmarou A, Maas AIR, Murray GD
Journal of Neurotrauma 2007, 24(2): 287-293

4) Multivariable prognostic analysis in traumatic brain injury: results from the IMPACT study
Murray GD, Butcher I, **McHugh GS**, Lu J, Mushkudiani N, Maas AIR, Marmarou A, Steyerberg EW
Journal of Neurotrauma 2007, 24(2): 329-337

5) Adjustment for Strong Predictors of Outcome in Traumatic Brain Injury Trials: 25% Reduction in Sample Size Requirements in the IMPACT Study
Hernández AV, Steyerberg EW, Butcher I, Mushkudiani N **Taylor GS**[12], Murray GD, Marmarou A, Choi SC, Lu J, Habbema JDF, Maas AIR
Journal of Neurotrauma 2006; 23(9): 1295-1303.

6) Design and Analysis of Phase III Trials with Ordered Outcome Scales: The Concept of the Sliding Dichotomy
Murray GD, Barer D, Choi C, Fernandes H, Gregson B, Lees KR, Maas AIR, Marmarou A, Mendelow AD, Steyerberg EW, **Taylor GS**, Teasdale GM, Weir CJ
Journal of Neurotrauma 2005; 22: 511-517

[12] My surname was previously Taylor