A study of response and non-response to postal questionnaire follow-up in clinical trials

by

Rachel Anne Nakash

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Health Care Research

University of Warwick, Warwick Medical School

June 2007

Table of Contents

1 Chapter 1 1

Introduction 1

2 Chapter 2 7

Summary of the survey literature on methods of improving response to postal questionnaires 7

2.1 A background to survey research 7

2.2 Questionnaire design and maximising response rates 9

2.2.1 Questionnaire length 10

2.2.2 Paper colour 11

2.2.3 Print details 12

2.2.4 Page layout 13

2.2.5 Question wording 14

2.2.6 Question order 15

2.2.7 Content of the covering letter 16

2.3 Questionnaire administration and maximising response rates 17

2.3.1 Pre-notification contacts 17

2.3.2 Follow-up contacts 18

2.3.3 Postal Method 19

2.3.4 Saliency 21

2.3.5 Incentives 22

2.4 Behavioural theories of survey response decisions 23

2.4.1 Reactance theory 24
2.4.2 Dissonance theory 26
2.4.3 Functional theory 27
2.4.4 Compliance theory 29
2.4.5 Leverage-Saliency theory 34
2.4.6 Chapter Summary 37
3 Chapter 3 38
Research aims and objectives 38
3.1 The context of the research 38
3.1.1 The importance of response rates in clinical trials 38
3.1.2 The setting for this research 39
3.2 Research Objectives 40
4 Chapter 4 42
The Collaborative Ankle Support Trial (CAST) 42
4.1 Background 42
4.2 Methods 44
4.2.1 Participants 44
4.2.2 Interventions 45
4.2.3 Objectives 46
4.2.4 Trial Procedures 46
4.2.5 Outcomes 47
4.2.6 Sample size 49
4.2.7 Randomisation 50

4.3 Results 50
4.3.1 CAST pilot study 51
4.4 Discussion 62
4.5 Chapter summary 63
5 Chapter 5 65
Systematic review and meta-analysis of methods of improving response to postal questionnaire follow-up in health research 65
5.1 Background 66
5.2 Methods 68
5.2.1 Search Strategy 68
5.2.2 Study Selection 70
5.2.3 Quality assessment 70
5.2.4 Quantitative Data Synthesis 77
5.3 Results 83
5.3.1 Trial Flow 83
5.3.2 Study Characteristics 84
5.3.3 Study Quality 89
5.3.4 Quantitative data synthesis 93
5.3.5 Sensitivity analysis 95
5.4 Discussion 99
5.4.1 Follow-up strategies 99
5.4.2 Questionnaire length 101
5.4.3 Incentives 102

5.4.4 Question order 103
5.4.5 Future research 103
5.5 Conclusions 104
5.6 Chapter summary 105
6 Chapter 6 106
A randomised controlled trial of a method of improving response to postal questionnaire follow-up in a clinical trial 106
6.1 Background 107
6.2 Methods 109
6.2.1 Participants 109
6.2.2 Development and design of the Trial Calendar 110
6.2.3 Objectives 120
6.2.4 Outcome measures 120
6.2.5 Sample size 121
6.2.6 Randomisation 124
6.2.7 Implementation of Trial Calendar Procedures 125
6.2.8 Data cleaning 126
6.2.9 Statistical methods 126
6.3 Results 128
6.3.1 Trial Flow 128
6.3.2 Descriptive statistics 129
6.3.3 Inferential statistics 130
6.3.4 Summary of results 155

6.4 Discussion 156
6.5 Conclusions 162
6.6 Chapter summary 164
7 Chapter 7 165
Issues surrounding response and non-response to postal questionnaires in a clinical trial – the patients perspective 165
7.1 Background 166
7.2 Methods 168
7.2.1 Sample selection 168
7.2.2 Data collection 170
7.2.3 Analytic framework 173
7.3 Results 179
7.3.1 Thematic framework 181
7.3.2 Mapping and Interpretation 183
7.4 Discussion 202
7.4.1 Response 202
7.4.2 Non-response 208
7.4.3 Suggestions to improve response 211
7.5 Conclusions 213
7.6 Rationale for choice of analytic framework 215
7.7 Ensuring rigour 220
7.7.1 Transparency 220
7.7.2 Validity 221

7.7.3 Reliability 223
7.7.4 Comparative 224
7.7.5 Reflexivity 225
7.8 Generalisability 226
7.8.1 Definitions of generalisability in qualitative research 226
7.9 Chapter Summary 230
8 Chapter 8 232
Characteristics of responders and non-responders 232
8.1 The difference between 'refusals' and 'non-contacts' 233
8.2 Socio-demographic characteristics of responders and non-responders to general and health related surveys 234
8.2.1 Socio-economic status 235
8.2.2 Age 236
8.2.3 Sex 238
8.2.4 Marital status 238
8.2.5 Work status 239
8.2.6 Urban versus rural dwelling 240
8.2.7 Symptoms 241
8.2.8 Summary 242
8.3 Socio-demographic characteristics of CAST participants 243
8.3.1 Definitions of CAST non-responders 243
8.3.2 Methods 246
8.3.3 Results 247

8.3.4 Discussion 257
8.3.5 Conclusions 262
8.4 Chapter summary 262
9 Chapter 9 264
Conclusions and recommendations 264
9.1 Summary 264
9.2 Practical conclusions 267
9.3 Theoretical conclusions 269
9.4 Recommendations 272
9.4.1 Recommendations for clinical trialists 272
9.4.2 Recommendations for further research 274
References 277
Appendices 292
Appendix 1 – CAST Background information questionnaire 293
Appendix 2 – CAST Outcome measure questionnaire 302
Appendix 3 – List of studies excluded from systematic review and reasons for exclusion 323
Appendix 4 – The Trial Calendar 340
Appendix 5 – Trial Calendar instructions for use 352
Appendix 6 – Interview schedule for qualitative study 354

Table of Tables

Table 1 *Response rates to the CAST pilot postal questionnaire follow-up* 51

Table 2 *Percentage of CAST pilot participants requiring no reminders* 52

Table 3 *Follow-up procedures used in CAST pilot study* 57

Table 4 *Response categories for CAST participants* 59

Table 5 *Types of responder based on response categories* 60

Table 6 *CAST pilot response categories* 60

Table 7 *Number of CAST pilot participants in each response category at each time point* 61

Table 8 *Electronic bibliographic databases searched and search strategy used* 69

Table 9 *Delphi list for quality assessment[146]* 76

Table 10 *Extracted data of randomised trials of methods of improving response rates to postal questionnaires in health care research* 86

Table 11 *Grading used on quality assessment tool* 91

Table 12 *Quality assessment scores of included studies* 92

Table 13 *Baseline characteristics of participants randomised to receive a Trial Calendar or no Trial Calendar. Values are numbers (%) (except for mean age).* 129

Table 14 *Final response rates at each follow-up point for participants supplied with a Trial Calendar compared to those not supplied with*

a Trial Calendar. Values are numbers responding and non-responding (percentage). 131

Table 15 *Code sheet for variables in Trial Calendar data set*..... 133

Table 16 *Four week follow-up: Summary of univariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)* 135

Table 17 *Four week follow-up: Multivariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)* 136

Table 18 *Four week follow-up: Effect of addition of interaction effects to previous multivariate model*..... 137

Table 19 *Crosstabulation of four week response rates showing age group and Trial Calendar possession. Figures are observed and expected numbers responding and non-responding* 138

Table 20 *Twelve week follow-up – Summary of the univariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)* 139

Table 21 *Twelve week follow-up: Multivariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)* 140

Table 22 *Twelve week follow-up: Refined multivariate model* 141

Table 23 *Twelve week follow-up: Effect of addition of interaction effects to previous main effects model*..... 142

Table 24 *Nine month follow-up: Univariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)* 143

Table 25 *Nine month follow-up: Multivariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)* 144

Table 26 *Nine month follow-up: Refined multivariate model*..... 145

Table 27 *Crosstabulation of nine month response rates showing relationship between age and Trial Calendar possession. Figures are observed and expected numbers responding and non-responding* 146

Table 28 *Nine month follow-up: Effect of addition of interaction effects to previous main effects model*..... 147

Table 29 *Crosstabulation of nine month response rates showing the relationship between sex and Trial Calendar possession. Figures are observed and expected numbers responding and non-responding* 148

Table 30 *Effect of the Trial Calendar on amount of prompting required to return questionnaire at each follow-up point. Values are numbers responding (percentage).* 149

Table 31 *Effect of the Trial Calendar on categorisation of 'keen' or 'reluctant' responder at each follow-up point. Values are numbers responding (percentage).* 150

Table 32 *Effect of Trial Calendar on amount of missing data for 'number of days off work' question at four week follow-up. Figures are number of questionnaires with answers missing or filled in (percentage).* 152

Table 33 *Effect of Trial Calendar on amount of missing data for 'time taken to return to sport' question. Figures are number of questionnaires with answers missing or filled in (percentage).* 153

Table 34 *Effect of Trial Calendar on amount of missing data for 'further treatment' question. Figures are number of questionnaires with answers missing or filled in (percentage).* 154

Table 35 *Basic structure of interview questions* 172

Table 36 *Characteristics of interview participants* 180

Table 37 *Thematic framework of response issues of clinical trial participants* 182

Table 38 *Frequency of reported response themes* 190

Table 39 *Understanding of the trial and reason for non-response* 195

Table 40 *Treatment preference, treatment received and response* 196

Table 41 *Features of rigorous qualitative analysis[207]* 220

Table 42 *Definitions of non-responders in CAST* 244

Table 43 *Distribution of socio-demographic characteristics in non-responders compared with responders at the CAST 4 week follow-up point (%)* 250

Table 44 *Distribution of socio-demographic characteristics in non-responders compared with responders at the CAST 12 week follow-up point (%)* 251

Table 45 *Distribution of socio-demographic characteristics in non-responders compared with responders at the CAST 9 month follow-up point (%)* 252

Table 46 *Distribution of socio-demographic characteristics in distinct groups of non-responders compared with responders at the CAST 4 week follow-up point (%)* 254

Table 47 *Distribution of socio-demographic characteristics in distinct groups of non-responders compared with responders at the CAST 12 week follow-up point (%)* 255

Table 48 *Distribution of socio-demographic characteristics in distinct groups of non-responders compared with responders at the CAST 9 month follow-up point (%)*..... 256

Table 49 *Comparison of mean scores (standard deviations) for 'benefit of treatment' question on CAST follow-up questionnaires at 12 weeks and nine months*..... 257

Table of Figures

Figure 1 *Standardised follow-up protocol used in CAST trial* 54

Figure 2 *Response types of CAST pilot participants at each time point* ... 62

Figure 3 *Flow chart of study selection process*..... 84

Figure 4 *Random effects model meta-analyses of methods of improving response rates to postal questionnaires in health care research – response rate after all follow-up efforts* 94

Figure 5 *Fixed effects model meta-analysis of methods of improving response rates to postal questionnaires in health care research* 95

Figure 6 *Meta-analysis of methods of improving response to postal questionnaires excluding studies of unclear quality* 97

Figure 7 *Meta-analysis of methods of improving response to postal questionnaires excluding small studies* 98

Figure 8 *Flow diagram of trial progress* 128

Figure 9 *Diagrammatic representation of response to general surveys* . 184

Figure 10 *Frequency of emergent response themes for 'keen responders' (n=9)*..... 191

Figure 11 *Frequency of emergent response themes for 'reluctant responders' (n=5)* 191

Figure 12 *Diagrammatic summary of reasons for response to clinical trial follow-up*..... 192

Figure 13 *Diagrammatic representation of reasons for non-response in clinical trial follow-up* 199

Publications and presentations

1. Nakash RA, Cooke MW, Marsh J, Lamb SE. Current practice of treatment of severe ankle injuries in A&E departments – a survey to assist the design of a randomised controlled trial. Abstract presented at the Chartered Society of Physiotherapy Annual Congress, ICC Birmingham, October 2003.

2. Nakash RA, Lamb SE, Cooke MW. Conducting clinical trials in accident and emergency medicine - a pilot study highlights recruitment challenges. Poster presented at FAEM conference, London, November 2003.

3. Lamb SE, Nakash RA, Withers EJ, Clark M, Marsh JL, Wilson S, Hutton JL, Szczepura A, Dale JR, Cooke MW, Collabortive Ankle Support Trial research team. Clinical and cost effectiveness of mechanical support for severe ankle sprains: design of a randomised controlled trial in the emergency department [ISRCTN 37807450].*BMC Musculoskelet Disord.* 2005 Jan 13;6(1):1.
[Some material from this publication used in chapter 4]

4. Nakash RA, Hutton JL, Jørstad-Stein EC, Gates S, Lamb SE. Maximising response to postal questionnaires - A systematic review of randomised trials in health research. *BMC Medical Research Methodology* 2006, 6:5 (23 February 2006)
[Condensed version of Chapter 5]

5. Rachel A Nakash, Jane L Hutton, Sarah E Lamb, Simon Gates, Joanne Fisher. Response and non-response to postal questionnaire follow-up in a clinical trial – A qualitative study of the patient's perspective. *Journal of Evaluation in Clinical Practice* 2007 [IN PRESS]
[Condensed version of Chapter 7]

6. Rachel A Nakash. Response to postal questionnaires in clinical trials – the patient's perspective. Poster presented at Warwick University Post Graduate Poster Competition. May 2007

Acknowledgements

I would like to thank the following people for contributing in their own ways to this research endeavour:

Firstly, thanks go to my supervisors; Professors Jane Hutton and Sallie Lamb and Dr Simon Gates for their invaluable advice, for bearing with me throughout this process and for the funding which has enabled me to achieve my ambition.

I would also like to thank my friends and colleagues at Warwick Clinical Trials unit for their support and encouragement - in particular Jo Fisher who was always at hand with a positive word when the chips were down. Thanks also to the recruitment physios on CAST for their invaluable help in implementing the Trial Calendar trial, namely Gill Doughty, Vivien Nichols Cheryl Ritchie, Mark Williams and last, but not least, Esther Williamson who kept the ball rolling whilst I was otherwise engaged.

Thanks to Emma Withers and Lisa Craven for help with the laborious task of transcribing interview data and to Vivien Nichols for re-coding the final transcripts.

Finally, thanks to my most excellent family who have supported and believed in me throughout, especially my husband. Thank you, Adam, for putting up with me, listening to me and understanding me – I owe you one!

This thesis is dedicated to my daughter Maddie:

Even things which you feel are beyond your capabilities are achievable with perseverance, determination and a smidge of self-belief

Declaration

The author declares that the work contained in this thesis is her own work and has not been used before. A condensed version of chapter 5 has been published and a condensed version of chapter 7 is in press. These publications were submitted during the research process of this thesis and are not connected with any prior research activity. A second researcher, Vivien Nichols, re-coded a random sub-sample of the interview transcriptions to enhance the reliability of the qualitative study.

Chapter 4 describes the randomised controlled trial which provided much of the setting for this research. This was a multi-centred trial on which the author was the Trial Manager. The protocol for this trial has been previously published[1], the author being a co-author on that publication.

This thesis has not been submitted for a degree at any other university.

Abstract

Postal questionnaires offer one of the least expensive modes of collecting patient based outcomes in health care research. Many methods of increasing response to questionnaires used in educational and market research surveys have been tested. Behavioural theories have also been applied to survey research to understand response decisions. Little attention, however, has focussed specifically on response issues to postal questionnaires used to collect data in clinical trials. This is the subject of this thesis.

A systematic review of methods of improving response to postal questionnaire follow-up in health care studies was conducted. A method of improving response was then devised and its effectiveness was tested within an existing clinical trial (the Collaborative Ankle Support Trial - CAST). This method was a 'Trial Calendar' which was a prompting and reminder tool to encourage response. Qualitative data were gathered from clinical trial participants to ascertain factors influencing their response decisions. Finally, the socio-demographic characteristics of CAST participants were examined.

The systematic review demonstrated that follow-up reminder systems had the most significant effect on response rates (RR 1.82, CI 95% 1.11 to 2.99). Incorporating such reminders into a tool such as the 'Trial Calendar', however, had no effect on improving response in CAST. The qualitative study revealed aspects of behavioural theories which could be incorporated into trial information and appeals for response. Analysis of the socio-demographic characteristics of CAST participants revealed that the youngest age group (16-24 years) was less likely to respond at every follow-up point.

It is concluded that rather than anticipating low response rates and striving to devise methods of converting non-responders into responders, efforts should be directed at preventing participants becoming non-responders in the first place. This thesis argues for the area of follow-up to postal questionnaires in clinical trials to become a theoretical research issue in its own right.

1 Chapter 1

Introduction

Patient based outcome measures are methods of assessing various aspects of health and health related issues from the patient's perspective.[2] Such methods are being used increasingly as tools for assessing the effectiveness of interventions in health care clinical trials.[2] It is becoming recognised that the patients' concerns and experiences need to be taken into account in evaluations of interventions. Objective outcome measures alone such as clinical and laboratory tests do not allow for the patient's perspective.[3] It is, therefore, argued that clinical trials should incorporate relevant patient based outcome measures.[4]

Patient based outcome measures usually take the format of a questionnaire. There are various ways of collecting the required information: the patient can be contacted and asked questions over the telephone, a face to face interview may be arranged or the questionnaire may be mailed to the patient for self-completion and return.

Good quality clinical trials often recruit many hundreds or even thousands of patients in their attempts to detect small but clinically relevant treatment effects.[5] The least expensive mode of collecting patient based outcomes in such large target groups is postal self-completion questionnaires.[6] Other advantages of this method are that it prevents interviewer bias (a

subconscious or conscious attempt by the interviewer to bias the respondents answer in some way) and allows the patient to respond in their own time without any pressure. There are disadvantages in that there is limited ability to explain questions or prompt the respondent for answers and also there is no control over who actually completes the questionnaire.[6] The biggest disadvantage with postal questionnaires, however, is non-response which reduces the effective sample size and can introduce bias.[7] Such 'losses to follow up' in a clinical trial are frustrating and potentially damaging. Lost patients could have had outcomes that would have affected the study conclusions. Investigators, however, have little choice but to exclude them from the analysis[8] or, if the context allows, make assumptions about their outcomes. There is some debate surrounding what is an acceptable rate of loss to follow up in a clinical trial. Sackett et al[9] suggest a 'five and twenty' rule of thumb with less than 5% loss to follow up probably leading to little bias and more than 20% loss giving the potential for serious threats to validity. Although there is no empirical evidence for this, some journals are now refusing to publish trials with losses greater than 20%.[10]

Obtaining losses to follow up of less than 20% in a clinical trial using postal questionnaires as a means of data collection takes considerable effort. Much research has been carried out on investigating methods of improving response rates to postal questionnaires but this is largely in the area of survey research rather than follow up in clinical trials.[11, 12] Furthermore,

such research is often conducted in the fields of social, educational and market research on surveys with no connection to health care.[13, 14] A recent Cochrane review carried out by Edwards et al[15] identified 292 eligible randomised controlled trials of methods of improving response rates to postal questionnaires. They found that methods such as unconditional incentives, shorter questionnaires and 'user-friendly' questionnaires can substantially improve response rates. They conclude that health researchers using postal questionnaires can improve their response rates by using these strategies. On closer inspection of the review, however, it is apparent that only about 20% of the included trials appeared to have a health care connection (i.e were published in a health related or medical journal). The majority of the trials were published in market research or educational research journals. This fact has been noted by other authors who suggest some caution in interpreting the findings of the review and question the generalisability of the results into the health care setting.[16] McColl et al[17] recommend, as a high priority, future research to investigate whether methods of improving response rates in non-health related surveys are effective in health surveys. O'Cathian[18] suggests a sub-analysis of the Edwards et al review focussing just on the health literature and also raises the issue of differences in response rates from different target populations. Factors affecting the response rates to a survey of health professionals, for example, may well differ from those that affect surveys of patient populations or the general public.[18] This is an important

issue that has received little attention in the current literature. The motivation of a patient to respond to a follow up questionnaire in a clinical trial is likely to be different to that of a health professional or a member of the general public selected to receive a survey questionnaire.[19] Tactics to encourage response may therefore differ.

There are many ways in which clinical trial participants may differ from survey participants. Saliency of the questionnaire has been shown to be one of the strongest predictors of response rate (a salient topic is defined as 'one which deals with important behaviour or interests that are also current').[20] It is assumed that a participant in a clinical trial receiving a questionnaire regarding their response to the treatment intervention would find the subject of the questionnaire highly salient. Also, the motive for the questionnaire in a clinical trial is clear to the participant. In a cold survey, recipients of the questionnaire may be suspicious as to the purpose of the survey and may therefore be reluctant to respond because of this.

Treatment preferences or satisfaction with the treatment received may also be influential factors in the response behaviour of clinical trials participants. Ludemann et al[21] found that patients in a clinical trial of laparoscopic fundoplication were less likely to respond to postal follow up if they had a poor outcome from the surgery. There are clearly issues surrounding the response of clinical trial participants and what makes them different from other populations that warrant further investigation.

Edwards et al have recently refined their review in a paper entitled 'Follow-up by mail in clinical trials: does questionnaire length matter?'.[7] The title suggests that the review is based on data from health care clinical trials. Twenty seven trials were included in the review and, again, closer inspection reveals that only 14 (52%) had studied health related topics and only four (15%) studied patients rather than members of the general public or student populations. Only one of the patient based trials[22] studied response rates in a clinical trial and two of the other patient based trials were based on unpublished data. Again, Edwards et al extrapolate the findings of their review and conclude that shorter questionnaires should be used in clinical trials to improve response. Unfortunately, despite the promise from the title, this review does not help to define issues of response that may be relevant to participants in a clinical trial. In defence of Edwards and his colleagues a review of methods of improving response rates to postal follow up specifically in clinical trials would be impossible since virtually no trials exist. This is surprising since it would seem relatively easy to include a sub-study in a clinical trial to investigate experimentally the effect of the various methods of improving response that have been identified in previous non-health related studies or surveys. Indeed, McColl et al[17] suggest 'piggy backing' such studies on to 'real' research studies rather than creating an artificial situation simply for the purposes of testing one or more hypotheses of methods of improving response.

This research endeavour was undertaken to address the lack of available literature specific to clinical trial follow-up. An aim was to investigate response to postal questionnaire follow-up in a clinical trial from a variety of perspectives. This included a randomised controlled trial of a method of improving response nested within an existing trial. A qualitative study was also conducted to investigate factors influencing response decisions of participants in clinical trials. Comparisons were then made with existing issues of response widely documented in the survey literature. A full explanation of the aims and objectives of this thesis is given in chapter 3.

2 Chapter 2

Summary of the survey literature on methods of improving response to postal questionnaires

Although this thesis is concerned with improving response rates to postal questionnaires specifically in clinical trials, the wealth of survey literature cannot be ignored. This chapter aims to summarise and discuss this literature. It is necessary to summarise the survey literature so that comparisons can be made with findings and issues arising in subsequent chapters. The findings of the recent Cochrane review of methods of improving response rates to postal questionnaires[15] is heavily referenced as the most extensive and up to date work in this area.

2.1 A background to survey research

It is widely considered that the pioneer of scientific social surveys was Charles Booth who, more than a century ago, began an enquiry into the *Labour and Life of the People of London.*[23] This investigation into the poverty and living conditions of the working class drew attention to the extent and severity of poverty at the turn of the 19th century. It is considered to be a pioneering contribution to the science of social study.[24] Over the subsequent thirty years large scale social surveys were being conducted across the country. These all followed the broad survey methodological

pattern established by the early pioneers.[24] In modern times, social surveys are a significant occupation of government organisations and market research has become a large-scale industry.[24] Surveys are now commonly conducted across a vast range of subjects and disciplines to gain both descriptive and explanatory information.

Early surveys were conducted using face to face interviews. It is not clear who carried out the first survey using the postal questionnaire method. It is, however, evident that postal questionnaires were used early on in the life history of survey research. The earliest reference found on how to maximise response to postal questionnaire surveys dates back to 1924.[25] The problem of non-response was clearly an issue from the outset of this survey methodology. The reason for such emphasis on maximising survey response rates is concern over 'non-response bias'. Underlying the inferences which are made from surveys is the assumption that all elements designated for the sample are actually observed and measured.[26] The concern is that if non-responders differ in important ways to the responders it is misleading to generalise the results of the survey across the population. Non-response in household surveys has become an increasing problem in recent years.[26, 27] Such is the current interest in this area that the journal '*Public Opinion Quarterly'* has recently dedicated an entire issue to 'Non-response bias in household surveys'.[28] This journal has also published many articles concerning methods of maximising response to survey questionnaires. An electronic search of this journal using the search terms

'mail' or 'postal' or 'questionnaire' (1936 to present) returned 106 articles relating to either questionnaire design or methods of maximising response. The body of literature on how to maximise response rates to postal questionnaire surveys can be broadly categorised into two groups. Firstly, many studies investigate design issues of the questionnaire itself and ways this can be manipulated to enhance the likelihood of response. Secondly, the mode of questionnaire administration and follow-up of reluctant responders has been widely investigated. The literature relating to these two categories will be summarised and discussed in the following sections. This chapter concludes with a discussion of some of some of the psychological behavioural theories that survey researchers have adopted in attempts to explain the response decisions of survey participants.

2.2 Questionnaire design and maximising response rates

The appearance of the questionnaire is one of many factors which can influence a person's decision on whether or not to respond.[17] Aspects of questionnaire design which have received varying amounts of attention in the survey literature are: questionnaire length, paper colour, print details, page layout, question wording, question order, and content of the covering letter.

2.2.1 Questionnaire length

It is widely reported in survey research that shorter questionnaires have better response rates than longer questionnaires.[14, 20, 29, 30] There is, however, a large amount of variation in the classification of 'long' and 'short' questionnaires. Some studies use number of pages to denote length whilst others use the number of questions. This heterogeneity is evident in a recent Cochrane review of methods of improving response rates to postal questionnaires.[15] Forty five studies were found which investigated the effect of questionnaire length on response rates. The results of the meta-analysis suggested that shorter questionnaires yield the best response (OR 1.73, CI 95% 1.47 to 2.03). There was, however, significant heterogeneity between the trials in the analysis. The variations in length of the questionnaires are identified by the authors as a possible source of this heterogeneity. The same authors have subsequently refined their review to focus just on the effect of questionnaire length on response.[7] In view of the previously identified heterogeneity, the studies in the refined review were stratified into four groups by questionnaire length. The four strata used were; postcard vs one or more pages, one page vs two or three pages, one page vs four or more pages and two or more pages vs longer alternative. In the sub-group where the shortest questionnaire was a postcard, the odds of response were more than halved for each additional page used (OR 0.39 CI 95% 0.34 to 0.45). In the other three strata, however, the effect sizes were much smaller. Further analysis also revealed that the results remained

heterogeneous despite this sub-group analysis. The authors are unable to explain the reason for the heterogeneity but make suggestions such as differences in questionnaire topic and background characteristics of the participants.

It appears that in spite of the claims of the survey literature, the relationship between questionnaire length and response rate is not quite as robust as it may first appear. Due consideration has to be given to the heterogeneity in terms of definitions of questionnaire length. Common sense would suggest that people are more likely to respond to a questionnaire which is concise than to a lengthy one. There is no clarification, however, as to the optimum length either in terms of pages or questions. There is some suggestion that participants respond better when the questions are evenly and well spaced rather than crammed onto one or two pages for the sake of making the questionnaire appear shorter.[29, 31] The issue of making survey questionnaires shorter is also complicated by the fact that the benefits (if any) on response rate needs to be traded off against the disadvantages of a reduction in data gained from using a shorter questionnaire.[15] Some authors suggest that it may be more rewarding to pay attention to details in the mailing procedures and to increase people's motivation than to cut out pages in the questionnaire.[32]

2.2.2 Paper colour

Although several studies have investigated various manipulations of questionnaire colour[33-35] there is no evidence pointing to the single most

effective paper colour. The Cochrane review mentioned above[15] identified ten studies which investigated the effect of questionnaire colour on response rate. No significant differences in response were seen using questionnaires printed on coloured paper. This review, however, grouped all colours together versus white questionnaires. It did not look at the effect of individual colours on response. An earlier review,[36] however, found most references to the effect of questionnaire colour on response related to the comparison of green versus white paper. A meta-analysis of green versus white questionnaires was therefore conducted which showed that green questionnaires lead to a small increase in response rate. Recommendations on the colour of paper to use are made in textbooks of survey administration but are, however, based on expert opinion rather than empirical evidence. For example, Dillman[37] suggests white or off-white paper but gives no explanation for this recommendation. With the lack of good quality evidence in this area it would seem appropriate to agree with Bourque and Fielder[29]: 'When in doubt, use black print on a white background'.

2.2.3 Print details

As well as paper colour, it has been suggested that the way the questionnaire is printed can have an effect on response.[29] Bourque and Fielding[29] propose that factors such as font size and type, ink colour and amount of white space can influence response. Their recommendations are to use a 10-point font which is easy to read, such as `Courier`, in black ink. Again, these recommendations are not referenced to any empirical

evidence and are based on the opinions of the authors. Dillman,[37] however, with a similar lack of evidence recommends a 12-point font. The Cochrane review[15] identified only one study which investigated print details. This study compared a questionnaire 'printed in black and white' with one 'printed in blue and yellow'.[38] They found that the coloured questionnaire significantly improved response rates (OR 1.39 CI 95% 1.16 to 1.67). It is assumed that the 'blue and yellow' questionnaire used blue ink on yellow paper although this is not explicitly described. The Cochrane review appears to have misinterpreted this study as they describe it as comparing 'coloured ink with black or blue ink'.

Until there is further empirical evidence to inform otherwise, common sense would suggest that questionnaires should be printed in a typeface which is large enough and clear enough to avoid strain in reading.[39] There is also no convincing evidence with respect to the provision of 'white space' in the questionnaire.[17]

2.2.4 Page layout

Most textbooks on survey design and administration recommend printing questionnaires in booklet format.[29, 37, 39] Again this is not based on empirical evidence but is justified as it makes the questionnaire look 'more professional',[29] easier to read and reduces the risk of losing pages.[39] The Cochrane review identified two trials comparing booklet format with stapled pages and found no differences in response rates between the two formats.[15] It is also suggested that common sense should be used when

spacing the questions so that a question is not split between two pages.[29, 37] Having to turn a page during the middle of a question is confusing and could introduce response errors.[17]

2.2.5 Question wording

Unlike many of the previous examples of questionnaire design, the wording of the questions is an area that has received 'bewildering' attention in the survey literature.[24] Many investigators have confirmed that changes in the way questions are worded can have a significant impact on the way people respond.[40, 41] Most of the literature in this area, however, is concerned with the way question wording affects the way in which people respond rather than whether they actually respond or not. The recent Cochrane review identified only three studies which looked specifically at the effect of question wording on response rate.[15] These studies compared open-ended questions (those which did not allow 'yes' or 'no' answers) with closed questions (those which did allow 'yes' or 'no' answers). Meta-analysis of the three studies showed that open-ended questions reduced the odds of response by two thirds (OR 0.31 CI 95% 0.09 to 1.04). There was, however, significant heterogeneity between the studies necessitating caution in the interpretation of the results.

An additional consideration in the discussion of the wording of questions is that if the questionnaire is a validated tool, manipulation of the wording of questions will not be possible.

2.2.6 Question order

The way that questions are ordered in the questionnaire has also received much attention in the survey literature. This may be more relevant to maximising response rates than question wording. The general consensus is that placing sensitive, unpleasant or embarrassing questions early on in the questionnaire may increase the likelihood of non-response.[17] Such questions may be of a personal nature or ask about undesirable attitudes or behaviour which respondents may be more reluctant to answer.

Many other variations in question order have been studied, for example; placing more relevant or salient questions at the beginning of the questionnaire,[42] placing demographic information questions first,[43, 44] placing general questions first[45] and placing the easiest to answer questions first.[46, 47] All these variations were included in the recent Cochrane review.[15] The results showed that placing the easiest and most relevant questions at the beginning of the questionnaire improved the odds of response. Placing demographic information questions first had no effect on response and placing the most general questions first lowered the odds of response. Very few trials were included in these analyses, however, making it difficult to generalise the findings. A study which also looked at the placement of demographic questions which was not included in the Cochrane review showed that response rate was improved by placing the demographic questions at the end of the questionnaire.[48]

The optimum question order to maximise response is, therefore, not clear form the current literature. The various effects of question order appear to vary with topic, context and study population.[17]

2.2.7 Content of the covering letter

Textbooks on survey design and administration stress that mailed questionnaires should always be accompanied by a covering letter.[24, 29] The content of the covering letter has received much attention in the survey literature. It is generally accepted that it should explain what the study is about, highlight why the sampled person is important, provide an assurance of confidentiality, offer a summary of the results of the survey, say what to do if questions arise and thank the recipient for their assistance.[37] Although this core structure is accepted, various aspects of the covering letter have been investigated. Such factors include the style of the letter, the characteristics of the signatory, the style of the signature and the nature of the appeal.[17] There is limited evidence that manipulating any of the above factors affects response rates. One study, however, found an increase in response rates if the covering letter was written in a traditional rather than a humorous style.[49] The Cochrane review[15] also found no unequivocal evidence from which to make recommendations for the covering letter content. There is a suggestion that stressing the assurance of confidentiality and asking participants for an explanation if they choose not to respond both increase response rates.[15] These findings, however, come from single studies and have not been verified by further evidence.

Relating the content of the covering letter to theories of individual motivation has been attempted by some authors.[50, 51] This relates to the nature of the appeal. For example, the covering letter could stress the importance of responding for altruistic reasons or for the self-interest reasons of the participant. Understanding the individual motivational factors in response behaviour is central to the theme of this thesis. Theories relating to survey respondent behaviour are therefore explored in detail and are presented in section 2.4.

2.3 Questionnaire administration and maximising response rates

The previous section looked at the way the design of the questionnaire can be manipulated to improve response rates. This section looks at ways the administration of the questionnaire affects response rates. Many methods of postal questionnaire administration have been investigated in efforts to make the process of response as easy as possible. These methods include pre-notification contacts, follow-up contacts, postal methods, saliency and incentives.

2.3.1 Pre-notification contacts

Contacting the survey participant prior to sending out the questionnaire has received much attention in the survey literature. The Cochrane review[15] identified 39 studies that investigated pre-notification compared to no pre-

notification. The meta-analysis showed that the odds of response were increased by one half if participants were pre-notified (OR 1.50 CI 95% 1.29 to 1.74). There was, however, significant heterogeneity among the trial results. Aspects of pre-notification that have been studied include comparing different modes of pre-notification and comparing the content of different pre-notification messages.[17] The available evidence comparing modes of pre-notification is limited but suggests that the mode (either telephone or letter) does not lead to differences in response rates.[52, 53] The content of the pre-notification message is often referred to in the survey literature as a 'foot in the door'. This technique involves gaining the participants cooperation with a small request with the aim of gaining cooperation with a subsequent larger request.[17] There is no available evidence to suggest that such techniques are more effective than simple pre-notification techniques.[17]

2.3.2 Follow-up contacts

The methods, timing, content and intensity of follow-up contacts of reluctant responders to surveys have also received much attention in the survey literature. The general opinion is that follow-up contacts are effective in improving overall response rates.[29, 37] It has been suggested, however, that this opinion is largely derived from analyses of the differences between initial and final response rates within studies rather than as a result of evidence from randomised controlled trials.[17] The Cochrane review,[15] however, identified 17 randomised trials that investigated follow-up contact.

The meta-analysis showed that the odds of response were almost one half higher when follow-up contacts were used compared to no follow-up (OR 1.44 CI 95% 1.25 to 1.65). Again, there was evidence of significant heterogeneity between the study results. This is probably due to the wide variations in mode and timing of follow-up across the included studies. A recent review of many aspects of questionnaire design and administration does sub-classify follow-up contact.[17] This review investigates the number, content and mode of follow-up contact as well as the effect of including a duplicate questionnaire. The authors conclude that there is no evidence that special mailing procedures, including a 'threat' of further follow-ups or including a duplicate questionnaire with the first reminder improve response rates. However, sending a duplicate questionnaire with the second reminder appears to be effective. The same authors also suggest matching the appeal in the reminder letter to the perceived motivations of the study population. This relates to theories of respondent behaviour as outlined in section 2.4. Heberlein and Baumgartner[20] suggest using such theories to explain those factors which influence a participant to respond to a single or initial mailing. They comment that most of the work carried out on follow-up contacts focuses on the final response rate achieved rather than ways to improve response to the initial mailing.

2.3.3 Postal Method

The survey literature contains many studies investigating the manipulation of postage rates and types and the effect this has on questionnaire

response. Studies tend to fall into two categories: those that investigate the effect of postage on the outgoing questionnaire and those that investigate the effect on the enclosed return envelope. The Cochrane review[15] found six trials which investigated the effect of stamped versus franked envelopes on the outgoing mail. There was no significant difference on response rates between the two conditions. An early study not included in the Cochrane review also showed no difference on response between stamped and franked outgoing mail.[54] The effect of postage on the return envelope has been subjected to much more investigation. The Cochrane review[15] identified 21 relevant studies and concluded that the odds of response were over a quarter higher when a stamped return envelope was used rather than a business reply envelope (OR 1.29 CI 95% 1.18 to 1.42). There was, however, significant heterogeneity between the trial results. Although Linsky (1975)[55] found a similar effect it was suggested that business reply envelopes offer cost advantage since the postage cost is only incurred if the envelope is returned. The cost effectiveness of business reply envelopes has also been noted by other authors.[56] In the absence of any strong evidence to advocate the use of stamps it is recommended that prepaid envelopes are used on both the outgoing and enclosed return envelopes.[17] Prepaid envelopes are more cost effective and less time consuming for the survey administration team. It is also recommended that a return address is added to the outside of the outgoing envelope to

facilitate the return of undeliverable mail. This recommendation, however, is based on expert opinion rather than hard evidence.[17]

2.3.4 Saliency

The use of the word 'salient' has been given slightly different meanings in the survey research literature. Heberlein and Baumgartner (1978) describe a salient topic as "one which deals with important behaviour or interests that are also current".[20] Groves et al, in presenting their 'Leverage-Saliency' theory of survey response, use the term 'salient' to describe the amount of attention drawn to certain attributes of a survey at the time of the request for participation.[57] Most studies of 'saliency' use the term to indicate the level of interest or relevance the questionnaire has to the recipient. The Cochrane review[15] found only two studies which compared the effect on response of 'high interest' and 'low interest' questionnaires. Meta-analysis showed that the odds of response were more than doubled using a questionnaire of 'high interest'. It is generally agreed in the survey literature that questionnaires perceived to be more interesting to the recipient will achieve higher response rates. Moser and Kalton[24] acknowledge that a questionnaire must include all the questions which are essential to answer the research question. They also suggest, however, that if these essential questions are likely to lack appeal for the target population, interesting 'throw-away' questions can be added to act as an incentive for completion.

2.3.5 Incentives

The use of incentives is, perhaps, the most extensively investigated area of methods of enhancing response rates to mailed survey questionnaires.

Some authors have investigated straightforward comparisons between the use of an incentive versus no incentive.[58-60] Others have studied the size of the incentive,[61, 62] compared financial versus non-monetary incentives[63, 64] and studied the effect of enclosed versus promised incentives.[61, 65, 66] The search for trials which used incentives to maximise response returned the most studies in the Cochrane review.[15] Seventy-two trials were found which evaluated the effect of non-monetary incentives and 69 trials investigated the effect of monetary incentives. Both types of incentive improved response but the effect size was bigger for the monetary incentive. There was, however, significant heterogeneity between the trials in both these conditions. The review also found that larger monetary incentives and unconditional incentives (i.e. those given with the questionnaire rather than dependant on its return) also improved response rates. The results of the Cochrane review are very similar to the results of earlier reviews which focussed just on the effect of incentives on mail survey response rates.[67-69] There is some overlap in the included studies of these reviews but the Cochrane review is by far the most extensive. Church (1993)[67] concludes that both monetary and non-monetary incentives are effective in improving response and that conditional incentives should be avoided as they are 'simply not worth the energy involved'.[67] The

evidence for the positive effect of incentives on response rates appears to be unequivocal. There are other issues related to the use of incentives, however, which deserve a mention but which are beyond the scope of this chapter for a detailed discussion. The effect of including an incentive on the quality of the returned data has been investigated.[14, 63] The size of a monetary incentive is also a debated issue with some having the view that the amount is immaterial and that it is the symbolic value which is important.[70] Other authors suggest a direct relationship between the size of the incentive and the increase in response rate.[67] Whether this relationship is linear is another debated point.[20] Pragmatically, the ethical and budgetary constraints of a survey will be important considerations in the use of an incentive to encourage response.[17]

2.4 Behavioural theories of survey response decisions

Many survey researchers have collaborated with social psychologists to gain insights into the psychological processes involved in survey participation. A fundamental question that survey researchers are driven to seek the answer to is: Why do some people answer surveys and others not?[71] A detailed discussion of all the psychological concepts which have been applied to survey research is beyond the scope of this thesis. Five of the major theories, however, will be discussed. These are: Dissonance theory, Functional theory, Reactance theory, Compliance theory and Leverage-Saliency theory. This section outlines the basic principles of these

theories and their possible application in the context of response to follow-up in a clinical trial.

2.4.1 Reactance theory

A 'Theory of psychological reactance' was first introduced by Brehm in 1965.[72] The basis of this theory is that an individual has a set of behaviours which he is free to engage in either at the moment or at some time in the future. If any of these 'free behaviours' are eliminated or threatened with elimination the individual will experience 'reactance'. Reactance is the motivational state directed towards the re-establishment of the free behaviours.[72]

Reactance theory has been applied to survey research by Biner[73] and Biner and Barton.[74] They conducted experiments manipulating both the size of a monetary incentive and the content of the cover letter sent with a postal questionnaire. In the first experiment[73] the cover letter emphasised that response to the survey was either 'essential' or a matter of 'personal choice'. The hypothesis was that the 'essential response' version would induce reactance due to the perceived elimination of the freedom of choice. The manifestation of this reactance would be a refusal to respond resulting in overall lower response rates. This was indeed the case. In a subsequent study, the cover letter explained the incentive as either an 'obligation to respond' or as a 'token of appreciation'.[74] The former explanation of the incentive was expected to induce the lowest response rates if the reactance theory was used as a prediction. The results showed, however, that the

'obligatory' cover letter group actually had a significantly higher response rate than the 'appreciative' cover letter group but only when the incentive was $1 as opposed to $0.25. The authors explain this by suggesting that rather than inducing reactance, by linking the appeal in the cover letter to the incentive, 'equity theory' was induced. Rather than feeling coerced into responding, participants felt truly obligated to return the questionnaire. There appears, therefore, to be a fine line in survey research between inducing reactance and fostering equity. An equity theory explanation of how enclosed monetary incentives enhance response rests on the assumption that the enclosed money will induce a sense of obligation to return the questionnaire. Reactance arousal is induced if participants feel coerced to respond due to the enclosed incentive.

There are ethical considerations regarding the use of incentives and coercive language to induce response in a clinical trial.[75] The Reactance theory may be more applicable at the recruitment phase of a clinical trial rather than at follow-up. Potential clinical trial participants make an informed decision regarding whether or not to agree to take part in the trial. It is strictly ethically inappropriate to coerce someone to participate in a clinical trial in a way which threatens their 'free behaviour'. Individuals who decide to participate have therefore done so by their own choice and presumably realise that by taking part they will be required to complete the follow-up questionnaires. The word 'presumably' is used here since if it is not made absolutely clear to the individual that participating in the trial means

completing every follow-up questionnaire, reactance may become evident at future follow-up points.

2.4.2 Dissonance theory

Cognitive dissonance theory was developed by Festinger in 1957.[76] Dissonance is seen as an 'unpleasant drive state' and when two or more cognitive elements are dissonant, there will be pressure to reduce that dissonance.[77] Dissonance theory has been suggested as an explanatory model for survey response behaviour by several authors.[77-79] This largely relates to the effects of enclosing monetary incentives on response rates. This theory suggests that if a person accepts a monetary incentive but decides not to participate in the survey they will experience cognitive dissonance. This aversive state of arousal eventually motivates them to reduce the dissonance by deciding to return the questionnaire.[74] The application of dissonance theory to explain survey response has, however, received some criticism. Biner and Barton[74] argue that dissonance is a 'post decision' phenomenon and consequently should not affect decision making. Once the decision to keep the incentive and not return the questionnaire has been made (and dissonance results) the questionnaire will most probably have been thrown away. To overcome this theoretical 'flaw', Biner and Barton suggest that the application of equity theory is more appropriate. Equity theory was proposed by Adams in 1963 as a special case of cognitive dissonance theory.[80] The critical difference when considering the application of the theory to survey response is that equity

considerations are *part* of the decision making process. This means that an individual only has to *consider* keeping the incentive for feelings of equity maintenance to develop. This should induce behavioural attempts to restore equity by returning the questionnaire.[74]

Dillman[37] has integrated many of the principles of equity and exchange in his 'Total Design Method' . This is a well documented approach in the survey literature which provides a framework for enhancing survey response. The framework is built on inducing cooperation between the researcher and the respondent and establishing a sense of trust in the latter.

Aside for the already debated problem of using incentives in a clinical trial, dissonance may be induced in clinical trial participants who feel that they have received 'better' treatment as part of the trial. These individuals may see returning their questionnaire as 'payment ' for this.

2.4.3 Functional theory

The 'Functional approach to the study of attitudes' was proposed by Katz in 1960[81] as an attempt to understand the reasons people hold the attitudes they do. The theory is based on four functions which attitudes perform for the individual. The four functions are: *Adjustive, Ego-defensive, Value expressive* and *Knowledge*. The Functional theory has been applied, in part, to survey research by McKillip and Lockhart in two studies among student populations.[50] They conducted studies which manipulated the content of the cover letter, sent with a postal questionnaire, to appeal to the different

functions. The effect of the different appeals on response rates was assessed. The *adjustive* function recognises the fact that people strive to maximise the rewards in their external environment and to minimise the penalties. McKillip and Lockhart[50] appealed to this function by emphasising in the cover letter the value of the study to the respondent as an individual. The *ego-defensive* function describes the mechanisms by which the individual protects his ego from his own unacceptable impulses and from the knowledge of threatening forces from without. This has similarities with the theories of dissonance and equity as described above. An individual will modify his behaviour and attitudes to avoid or diminish internal cognitive conflict. This function was not specifically appealed to in the experiments of McKillip and Lockhart. The *value-expressive* function describes the mechanism of an individual giving positive expression to his central values and to the type of person he conceives himself to be. McKillip and Lockhart[50] appealed to this function by emphasising the value of the study to students and the university in general. This gave the students an opportunity to express favourable altruistic tendencies. The *knowledge* function describes how individuals seek knowledge to give meaning to what would otherwise be a disorganised, chaotic universe. This function is appealed to by McKillip and Lockhart[50] by the cover letter focussing on the contribution of the survey to general and personal knowledge bases. Other survey researchers have included various appeals in the cover letter to emphasise such things as 'altruism', 'social utility' and egoism' but without

relating the appeal to Katz's Functional theory.[51, 56, 82] The results of these studies and those of McKillip and Lockhart[50] are inconclusive in recommending a universally appropriate cover letter appeal. It has been suggested that the nature of the appeal made in the cover letter should be based on the perceived motivations of the study population and should be ethically sound.[17]

Clinical trials using postal questionnaire follow-up typically include a covering letter with the questionnaire. This theory may offer useful insights into the wording of this letter. Again, however, the language used needs to be chosen carefully to comply with ethical constraints.

2.4.4 Compliance theory

A notable collaboration between survey researchers and social psychologists is that between Robert Groves and Mick Couper from the Survey Research Centre at the University of Michigan and Robert Cialdini, Professor of Psychology at Arizona State University. Cialdini has written extensively on the psychological concepts of compliance and persuasion. [83-85] Groves and Couper have identified compliance as one of the most relevant psychological concepts to explain the response decisions of survey participants.[86]

The decision about whether to respond to a requested activity is often made on the interest value and personal relevance of the activity to the participant. The cost in terms of time, energy and resources required to perform the activity are also factors in the decision process.[86] Six compliance

principles that are used frequently in deciding whether to yield to a request have been identified by Cialdini following several years of controlled psychological research.[85] In the context of survey research, Cialdini describes the decision to respond as 'heuristic'.[87] In psychology the term 'heuristic' refers to simple, efficient rules of thumb which have been proposed to explain how people make decisions. The concept that the decision to respond to a survey is heuristically based is supported by Groves and Couper. They believe that a potential survey respondent does not usually have a large personal interest in the subject of the survey. Consequently the participant is unlikely to invest large amounts of time or energy in the decision to participate.[86] Groves and Couper suggest that Cialdini's compliance principles guide the shallow and quick heuristic decisions to respond to a survey request.[87]

The six compliance principles are outlined below:

1. Reciprocation: This principle suggests that 'one should be more willing to comply with a request to the extent that compliance constitutes the repayment of a perceived gift, favour or concession'.[86] This relates to survey participation in that participants may have a perceived sense of obligation to the organisation making the request or to the broader society it represents. In a clinical trial, participants often take part due to a perceived personal benefit by doing so.[88] This may invoke the reciprocation principle in that participants may respond to the questionnaire as repayment for

perceived 'good' treatment. On a more superficial level, incentives may be sufficient to invoke reciprocity. There is a large body of literature detailing the positive effects of incentives on survey cooperation.[12, 67, 89] The reciprocation principle could be an underlying factor in these findings.[86] An exception to this principle occurs when the gift, favour or concession received is viewed as a bribe or an undue pressure to comply. In these circumstances compliance is inhibited.[90]

2. Consistency: 'After committing oneself to a position, one should be more willing to comply with requests for behaviours that are consistent with that position'.[86] It has been noted by social psychological theorists that most people have a strong desire to be consistent within their attitudes, beliefs, words and deeds.[91] In the context of a clinical trial this principle could be highly relevant. Participants are recruited onto trials after being given an explanation of the purpose of the trial and the requirements of being involved.[92] Participants are required to sign a consent form to confirm that they have been given this information and that they are willing to participate. The consistency principle implies that, once recruited, the participant should see the trial through to its completion. In a clinical trial using postal questionnaire follow-up this entails completing a questionnaire at every requested time point. Although such efforts are made to inform clinical trial participants of what is involved it can not be assumed that participants have fully understood all that is expected of them. The applicability of the

consistency principle in the clinical trial setting can, therefore, not be guaranteed as it is based on assumptions of a full understanding of the trial.

3. Social Validation: 'One should be more willing to comply with a request to the degree that one believes that similar others would comply with it'.[86] If a participant believes that most people like themselves agree to participate in surveys they will be more inclined to participate themselves.[86] This principle has the potential to be applicable to postal questionnaire follow-up in clinical trials. Participants may be more likely to respond if they are aware that most other participants have responded. This area, however, has never been investigated.

4. Authority: 'One should be more willing to yield to the requests of someone whom one perceives as a legitimate authority'.[86] Using this principle, survey participation is expected to be greater if the legitimacy of the sponsor (eg government or educational institution) is emphasised. The authority principle, however, may have a negative effect on people who believe they have suffered injustice at the hands of major institutions.[87] This has implications for the clinical trial setting for participants who felt that they did not get the 'best' treatment or that the treatment given did not help them.

5. Scarcity: 'One should be more willing to comply with a request to secure opportunities that are scarce'.[86] This compliance principle is used in survey research where respondents are encouraged to 'make their voice heard'. The survey is perceived as a rare opportunity to participate in an interesting and/or important activity.[87] Clinical trials are usually undertaken

to test a new treatment intervention or to compare treatments to see which one is most effective. The scarcity principle is therefore relevant in this context. Participants in a clinical trial may feel that they are at the 'cutting edge' of medical research and can make a difference to society by taking part. Again, no previous work has been done to verify this.

6. Liking: 'One should be more willing to comply with the requests of liked others'.[86] The liking of strangers has been shown to be influenced by a variety of factors including similarity of attitude, background and dress.[86] This principle is most relevant to surveys involving patient interviews. With postal questionnaires the participant has far less direct contact with the administrator of the questionnaire. In clinical trials, however, the participant is often initially recruited onto the trial by a researcher in a face to face setting. This allows a hypothesis to be made that the liking principle may be applicable in this setting and may encourage the participant to comply with postal follow-up.

The six principles of compliance outlined above have all been shown to affect the decision to 'help'.[93] There are, however, differences in the concepts of help giving and compliance which are used in the decision to respond to a request. A simple request to participate in a survey may be more successful when it includes an appeal to the 'helping norm' (the motivation to help others who are in need).[86] The emotional state of the participant has also been shown to have an affect on requests for help.

Anger and happiness can have negative or positive effects on helping behaviour respectively.[86, 94] The implications of this on survey compliance are that anger would be expected to reduce compliance and happiness would be expected to improve compliance. This has relevance to clinical trials in that the emotional response of the participant to the treatment received could have an influence on their likelihood of responding to postal follow-up. A participant who has negative feelings and has been unhappy with treatment would, according to helping theories, be less likely to continue to help with the research process. The opposite would apply to participants who have been happy with the given treatment. Helping theories can also be applied to clinical trials with respect to the participant's cooperation for altruistic reasons. Studies have shown that altruism is a frequently cited motivational factor for participants to agree to take part in clinical trials.[88, 95] Participants may feel that by being part of a trial they are contributing to a body of research that will potentially help others with similar conditions.

2.4.5 Leverage-Saliency theory

The Compliance theory of survey participation outlined above has been developed by the authors into a theory which includes the relevancy of the topic of the questionnaire to the survey participant. The Leverage-Saliency theory speculates that people vary in the importance they assign to various aspects of a survey request.[96] Survey attributes such as cash incentives or its link to helping the community will be assigned different levels of

importance by different individuals. The propensity of an individual to respond to a survey request is seen as a function of this assigned importance, whether its influence is positive or negative and how salient that aspect of the survey is made during the survey request.[57] The developers of this theory believe that it allows survey researchers to speculate on why there is 'an embarrassing lack of replication'[57] of experimental findings of methods of improving response (as outlined in sections 2.2 and 2.3). The Leverage-Saliency theory is the only theory developed specifically for survey response. Other theories outlined in this section have been drawn form established psychological theories of behaviour and have then been applied to survey research. The Leverage-Saliency theory may, therefore, be more applicable to the clinical trial setting although this has not been tested.

The psychological concepts associated with survey cooperation outlined in this section have many features which appear to be applicable to participants receiving postal questionnaires in a clinical trial. Certain aspects of these theories highlight the potential differences between survey participants and clinical trial participants. The 'compliance principles' (see section 2.4.4), for example, are presented in the survey literature as 'heuristics' or quick, shallow decisions to respond. Clinical trial participants may spend more time considering their decision to respond and demonstrate different motivational factors in this decision. Relevancy of the

questionnaire topic is an important consideration. Even in survey research there will be times when the survey is of interest to the recipient. All the behavioural theories discussed, apart form the Leverage-Saliency theory, were originally identified to explain human behaviour in general and are based on the findings of experimental psychological research. The theories have then been adopted by survey research retrospectively rather than openly questioning people about their attitudes to surveys.

The term 'survey on surveys' has been used to describe evidence which uses survey participants own verbalisations of their response behaviour.[71] This technique has been used by many authors to investigate various aspects of survey response.[97-100] The data which arise from such research suggests that the sponsor of the research and the persistence of the fieldworkers are important factors influencing response.[100] The topic of the survey is also a crucial motive for response and can have either a positive or negative effect. A survey of sexual behaviour which telephoned non-responders to ascertain a reason for their non-response found 36% of non-responders refused due to the nature of the survey.[101] Other reasons for non-response generated from 'surveys on surveys' appear to be quite diverse but are commonly variations of a theme of 'too busy' or 'never got round to it'.[100] DeMaio (1980) however, found that 'invasion of privacy' and 'past experiences' were the most commonly cited reasons for survey refusals.[99] Such personal accounts from non-responders themselves may

offer a better insight into response issues than the application of abstract theories of human behaviour.

2.4.6 Chapter Summary

This chapter gave a brief overview of the literature surrounding methods of improving response to postal questionnaire in survey research. Given the substantial body of research in this area it is surprising that more concrete recommendations have not developed from this literature. Findings appear to be inconsistent for many of the techniques of improving response. Some authors have noted the possibility of complex interactions between particular techniques and the population or situation within which they are used.[55] It has also been suggested that methods that have been identified as improving response rates may be affected by important cross-cultural differences.[30] Little systematic survey research has been conducted which collects information from refusers about their reasons for non-cooperation.[99]

Whether methods which have been shown to improve response in survey research are effective in the context of a clinical trial is the subject of the systematic review presented in chapter 5. A qualitative study of clinical trial participants' own verbalisations of their response decisions is presented in chapter 7. Chapter 7 also discusses the applicability of the behavioural theories of survey response to the clinical trial setting in detail.

3 Chapter 3

Research aims and objectives

3.1 The context of the research

3.1.1 The importance of response rates in clinical trials

The issue of non-response and non-response bias highlighted in the previous chapter relating to survey research is also a major issue in clinical trials. Although this is well recognised in the literature, comparatively little has been written regarding how to maximise response rates to postal questionnaires in this setting. No studies can be found which investigate the deeper theoretical aspects of response in a clinical trial. Poor response rates are a potential source of bias because non-respondents can differ from respondents with respect to important characteristics.[17] It is recognised that elimination of loss to follow-up in a clinical trial is probably impossible. However, it is also suggested that that investigators too frequently claim overwhelming difficulties in reducing these losses and could work harder to obtain higher follow-up rates.[8] Additionally, confidence in the results of a trial is achieved, in part, by the sample size and this is directly related to response rate. The sample size of a clinical trial needs to be sufficient to detect a difference between the groups if one is present (i.e. reducing type 2 error). This relates to what is considered to be the smallest treatment difference that is of such clinical value that it would be undesirable

to fail to detect it. The 'power' of a study is the degree of certainty that the desired treatment difference, if present, will be detected. These factors are used in a standard statistical approach to determine the sample size required.[102] In clinical trials, data is often collected at various points following the treatment intervention to assess long term outcome. Although the desired sample size may be achieved at baseline, if patients fail to respond to follow up questionnaires the sample size effectively gets smaller. This will therefore make subsequent statistical analysis less robust. The anticipation of losing some participants to follow-up can be written into the sample size calculation to maintain the studies power to detect clinically relevant outcomes in spite of these losses. It is, therefore, the effect of loss to follow-up on bias that is of greatest concern to clinical trialists.

3.1.2 The setting for this research

The majority of this research was conducted around an existing acute injury clinical trial – the Collaborative Ankle Support Trial (CAST).[1] This trial was conducted by researchers at the University of Warwick in collaboration with eight accident and emergency departments around the West Midlands, Central and South West England. Patients with severe ankle sprains were randomised to receive one of four different types of ankle support. These supports were: tubigrip, plaster cast, Aircast splint and Bledsoe Boot. Patients were followed-up by postal questionnaire at four weeks, 12 weeks and nine months after injury. The pilot phase of this project highlighted the difficulties of maintaining adequate response rates in an acute injury clinical

trial. This led to the development of a standardised sequence of prompts to encourage questionnaire return. To provide detail on the amount of prompting required, participants were coded into a 'response category' according to their response to each follow-up questionnaire.

CAST therefore provided an ideal setting to investigate response issues in clinical trial participants. A full account of CAST with an analysis of the pilot study data is provided in the next chapter.

3.2 Research Objectives

This research has four distinct phases:

i. Systematic Literature Review

Numerous market and educational research studies have been carried out to evaluate strategies of improving response rates to postal questionnaires. Few, however, have been specific to the health care setting, nor to the context in which participants are receiving or being allocated an experimental health care treatment.[12-15, 36, 103]

The aim of this phase of the research was to conduct a systematic review and meta-analysis of methods of improving response to postal questionnaire follow-up in health care studies on patient populations. This is a refinement of other reviews in this area. This is the first review to be conducted which focuses just on the health care literature. The results of the

review will highlight potential differences in response behaviour between participants in health care research and general surveys.

ii. Randomised Controlled Trial

A method of improving questionnaire response and minimising missing data was developed. This was then tested in the context of CAST. By 'nesting' the intervention into an existing clinical trial, a more pragmatic insight into its effectiveness would be gained.[17]

iii. Qualitative Study

CAST participants were interviewed to gain a broader perspective of the response decisions made by participants in acute injury management trials. No previous studies have collected such qualitative data.

iv. Analysis of characteristics of responders and non-responders

Data were also available from the CAST database regarding the characteristics of responders and non-responders. This was then analysed to identify any common socio-demographic variables of responders and non-responders.

This research aims to learn from actual clinical trial participants as well as evaluating current theories of response. The conclusions drawn from this endeavour will help clinical trialists maximise their response rates and deliver more robust research findings.

4 Chapter 4

The Collaborative Ankle Support Trial (CAST)

This chapter provides a contextual background for the rest of this thesis. The CAST methodology is explained in detail to give clarification to subsequent chapters which make reference to specific aspects of CAST.

4.1 Background

Acute ankle sprains are one of the most common conditions seen in accident and emergency departments. It is estimated that ankle sprains account for between 3 and 5 % of all UK A&E attendances[104] with about 5600 injuries each day.[105] The majority of ankle sprains involve the lateral ligament complex[106] (the ligaments on the outside of the ankle) and account for one quarter of all sports injuries.[107] The injury is painful and incapacitating and, unless the injury is minor, weight bearing is difficult to tolerate. Lateral ankle sprains are widely viewed as being uncomplicated and self limiting. Several studies have shown, however, that although the acute symptoms resolve, residual symptoms can linger for months or even years after the initial injury.[108] In a seven year follow-up study of ankle sprains, Konradsen et al[109] found that 32% of subjects experienced residual disability with symptoms of pain, swelling or recurrent sprains. Early

effective treatment is not only crucial to promote a speedy resolution of acute symptoms but is also an important feature in limiting the chronicity of the injury.[108] Clearly defined primary care protocols and a broad knowledge of new methods of rehabilitation are required to restore full activity as soon as possible.[110] No good quality studies can be found that describe the long term outcome of treatments for ankle sprains.

The Collaborative Ankle Support Trial (CAST) was a randomised controlled trial which was conducted by a research team from Warwick Emergency Care and Rehabilitation at the University of Warwick. The trial was commissioned and funded by the NHS Health Technology Assessment programme. It was designed to evaluate the clinical and cost effectiveness of three different methods of mechanical support compared to tubigrip (an elasticated bandage) following severe ankle sprain. The three supports consisted of cast immobilisation and two types of ankle splint. The two types of splint represented fairly new innovations in technology for supporting ankle sprains. Outcomes were assessed in the short term (4 weeks), medium term (12 weeks) and long term (nine months). A national survey of current practice was undertaken prior to the commencement of the trial to inform the trial design.[111]

4.2 Methods

4.2.1 Participants

Trial participants were people attending the accident and emergency departments of eight hospitals across the West Midlands, Central and South West England with a diagnosis of a grade II or grade III ankle sprain. This classification is used to indicate moderate to severe sprains.

The participating hospitals were:

- Coventry and Warwickshire Hospital
- Birmingham Heartlands Hospital
- Frenchay Hospital, Bristol
- John Radcliffe Hospital, Oxford
- Alexandra Hospital, Redditch
- Solihull Hospital
- Hospital of St Cross, Rugby
- Warwick Hospital

Inclusion criteria

All people who attended accident and emergency with a grade II or III sprain of the ankle, aged 16 years and older, who were able to give informed consent. The approach included all ethnic backgrounds.

Exclusion criteria

Age less than 16 years old, ankle fracture, or other fracture sustained in addition to the Grade II and III ankle sprain (such as to the wrist, head etc). Age was used as an exclusion criterion because of the complications involved in the management of epiphyseal injuries (growth plate injuries). Growth plate injuries would not normally be managed using the treatment methods being tested. Patients were also excluded if they had a contra-indication to any of the four arms of the trial. This was most likely to occur in the plaster group for example if a patient had a history of DVT or high risk of DVT or other circulatory disturbance. Other contra-indications included poor skin viability preventing splinting or casting. The decision to exclude on this basis was at the discretion of the attending clinician. Flake fractures of the ankle of less than 2mm were included as these are normally treated as soft tissue injuries.

4.2.2 Interventions

The following three interventions were compared to tubigrip which acted as the control group:

- *Cast immobilisation*

A standard below knee walking cast was used. With reference to current clinical practice and previous research, a time of 10 days in plaster was set.

- *Aircast brace*

This is a removable brace which fits inside the shoe to prevent the ankle twisting but allows normal walking movement.

- *Bledsoe Boot*

This is a larger removable walking brace which fits onto the lower leg. Metal struts completely immobilise the ankle whilst the brace is on.

The duration of use of the two removable braces was according to the manufacturers recommendations. All the interventions were applied by qualified clinicians who were given specific additional training in brace application if necessary. The ankle supports were fitted within three days of injury. A protocol of additional basic treatment and advice for ankle sprains was standardised across all the participating trial centres. This included basic exercises, ice and pain relieving medication.

4.2.3 Objectives

CAST had two main objectives:

1. To estimate the clinical effectiveness of three different methods of ankle support (below knee cast, Aircast brace and Bledsoe Boot) in comparison to Tubigrip.
2. To measure the cost of each strategy, including treatment and subsequent health care costs.

4.2.4 Trial Procedures

Ankle sprain patients were identified as being potentially eligible for trial inclusion at their initial accident and emergency visit. A normal assessment and investigation was carried out by the attending clinician. If the patient

was considered appropriate they were given a trial information pack and invited to return to a 'Trial Clinic' at a later date. To ensure that that their ankle injury was still in the acute phase, the participant had to return to a trial clinic within one week of injury. The trial clinics were staffed by members of the central research team or local collaborators who were trained in the trial procedures. At the trial clinics potential participants were given an explanation of the trial and were invited to take part. If the participant agreed to be involved, informed consent was taken and the baseline questionnaire pack was completed. This pack contained a questionnaire to collect background and socio-demographic information as well as the outcome measure package. The participant was then randomly allocated one of the three ankle supports or the control treatment. The support was applied by the trial clinic clinician (or plaster room staff in the case of the cast) and standardised instructions on the use of the support were given. Patients randomised to receive a cast were given an appointment to return for removal after 10 days.

4.2.5 Outcomes

Primary

- The recovery of mobility.
- The recovery of normal occupation, including return to normal work, study, caring or other activities.

Secondary

- Avoidance of residual symptoms including recurrent instability, lasting limitation of physical activity, and need for further medical, rehabilitation or surgical treatment.

A patient based outcome measure package was compiled to include disease specific and generic quality of life outcomes. It has been suggested that these two factors should be considered in a clinical trial.[2] There are, however, few validated and reliable measures to document recovery after ankle sprain. Furthermore, many existing disease specific outcome measures for ankle sprains require a clinical examination. Such measures were therefore not appropriate for CAST, which used postal questionnaire follow-up. A disease specific measure was finally chosen following a detailed review of the literature. Due to the recognised inadequacies of ankle disease specific measures, two additional measures were included. A further measure was included for the economic analysis. To cover the primary and secondary outcomes and the economic analysis, the finalised outcome measure package therefore consisted of the following measures:

- The Foot and Ankle Outcome Score (FAOS)[112]
- The Functional Limitations Profile (FLP) work and ambulatory sub-scales[113]
- The SF12[114]
- The EQ-5D for the economic analysis[115]

These measures were formatted into a single 18 page A4 questionnaire booklet. Return to normal occupation and leisure activities were recorded as single question items on the questionnaire. A short resource use questionnaire was added to the booklet at the 12 week and nine month follow-up to ascertain additional treatment and/or expenses incurred by participants. See appendix 1 and 2 for examples of the CAST background information and outcome measure questionnaires.

The outcome measure questionnaire was completed in the trial clinic on recruitment into the trial. It was then posted to the participants four weeks, 12 weeks and nine months following injury.

4.2.6 Sample size

The sample size estimate was based on a standard sample size calculation for a two-sample t test with equal variances and a significance level of 0.05. The variance was estimated from an ANOVA of the 4 and 12 week data from the pilot phase and initial recruitment. A difference of 10% was taken as the minimal clinically important difference. Estimates of the mean and standard deviations were based on observed values at baseline, 4 and 12 weeks. Sample size estimates were calculated using standard methods.

The total sample size was around 600 participants which included an allowance of 20% loss to follow-up.[1] This was a revised sample size with the approval of the trial Data Monitoring Committee.

4.2.7 Randomisation

Telephone randomisation was used and randomisation was stratified by trial centre. Allocation concealment was ensured by using a remote computer generated randomisation system that was independently administered and quality controlled. This meant that people entering participants into the trial were shielded from discovering future allocations.

Due to the nature of the trial, blinding of the participant and those administering the intervention to treatment allocation was not possible. Personnel responsible for data inputting and outcome assessment were, however, blind to treatment allocation.

4.3 Results

An account of the results of CAST is beyond the contextual explanation required for this thesis. It is however appropriate to include an account of the response issues identified in the pilot phase of the trial. This phase of CAST was used for the purposes of this thesis to assess response rates, clarify follow-up procedures and identify aspects of response to subject to deeper investigation. These aspects of a clinical trial are central to the theme of this thesis. The CAST pilot phase therefore provided a useful initial access to a 'real' clinical trial situation.

4.3.1 CAST pilot study

The pilot phase of CAST took place between April and June 2003 at Coventry and Warwickshire Hospital. Twenty four patients were recruited, 11 females and 13 males. The age range was 16 to 53 with a mean age of 30. Seven patients were randomised to receive tubigrip, six were randomised to plaster, seven to Aircast splint and four to Bledsoe Boot.

4.3.1.1 Response rates

The final response rate at each follow-up time point was established after all efforts to chase reluctant responders. Table 1 shows the response rates at the three follow-up time points for the CAST pilot study.

Table 1 ***Response rates to the CAST pilot postal questionnaire follow-up***

Follow-up point	*Number responding*	*Response rate*
4 weeks	19	79%
12 weeks	16	67%
9 months	15	63%

It became apparent during the CAST pilot that few participants respond without some form of reminder and this is more evident at each time point as shown in Table 2

Table 2 ***Percentage of CAST pilot participants requiring no reminders***

Follow-up point	*Response rate*
4 weeks	37%
12 weeks	29%
9 months	12.5%

Apart from returning unacceptably low response rates there are other dangers of only including first time responders in an analysis. Several authors have identified differences between earlier and later responders therefore only including those who respond to the initial approach could introduce bias.[17] It became apparent, therefore, that a standardised system of follow-up would be required in CAST to promote response. This would give the trial administration team a protocol to follow to chase reluctant responders. It would also give clarification on when to stop chasing and class the participant as a non-responder.

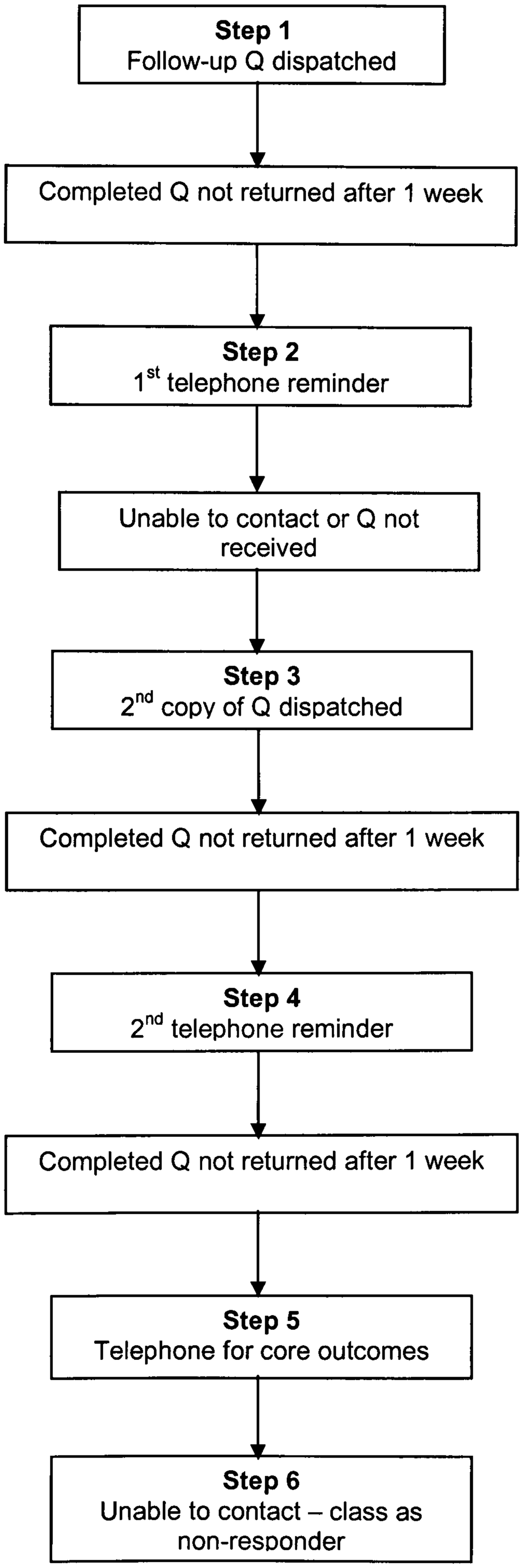
4.3.1.2 Standardised follow-up procedures

To establish the standardised follow-up procedures, the literature surrounding promoting response to postal questionnaires was considered. The overview of this literature (as presented in chapter 2 section 2.3.2) suggested a general consensus on the effectiveness of follow-up contact. An early review in survey research showed that intensive follow-up techniques are particularly effective in promoting response.[55] Subsequent

texts have concurred with this.[24, 116] Dillman[37] has gone so far as to present a 'Total Design Method' which involves three follow-up contacts including a postcard reminder and replacement questionnaires to persistent non-responders. The work of Dillman is well cited in the survey literature as an effective way of improving response. This method has also been tested in the context of a clinical trial and was found to effectively increase the response rate to mailed questionnaire follow-up from 62% to 88% over a 70 day period.[117] In a review of many aspects of enhancing questionnaire response, McColl et al[17] recommend at least one follow-up contact and including a duplicate questionnaire with the reminder. They highlight, however, that every research situation is unique and follow-up efforts should be tailored to suit the research context. They also point out the lack of literature regarding follow-up procedures specific to the health care setting.

Based on the available literature and the time and resource constraints of CAST, Figure 1 shows the standardised protocol that was established and repeated at each follow-up time point:

Figure 1 ***Standardised follow-up protocol used in CAST***



Core outcomes (Step 5) were obtained as a last resort if the participant failed to return their questionnaire after all the previous prompts. This involved asking the participant a brief set of questions over the telephone which mapped onto the primary outcome questions in the main questionnaire. This was only done after every possible attempt to persuade the participant to return their completed questionnaire. Although the information gained from this option was very brief it was seen as better than gaining no information at all from the participant. In subsequent chapters of this thesis which examine response and non-response in CAST, participants who required the collection of core outcomes are classed as non-responders since no questionnaire was ever returned.

4.3.1.3 Analysis of response procedures

Although the standardised follow-up protocol offered an appealingly simplistic flow chart to follow, in reality it was more complicated. The CAST administration team often had to send out more than one repeat mailing of the questionnaire and invariably had to make numerous telephone calls before eventually contacting participants. Participants were either out or calls were diverted to an answer phone. The phone call was therefore only considered to 'count' in the follow-up procedure if the CAST administrator actually spoke to the participant. Some participants were impossible to contact by telephone throughout the whole follow-up procedure. If such participants failed to return their questionnaire by the end of follow-up they were classed as non-responders. It was, therefore, theoretically possible for

a participant to be classed as a non-responder without actually receiving any telephone prompts due to the inability of the trial administrator to make a successful contact. These participants could be different from those who received maximum prompting but still failed to return their questionnaire. This concept has been recognised in the survey literature with a distinction between 'non-contacts' and 'refusals'[71] but has not been investigated in the clinical trial setting. A field was therefore added to the CAST database during the pilot phase enabling the CAST administration team to keep a detailed log of attempts at contact for each participant. This made it possible to distinguish between those who were 'non-contactable' and those who received prompts but still failed to respond. This provided useful data for analysis of the characteristics of responders and non-responders which is detailed in chapter 8.

Table 3 gives a summary of the follow-up efforts made by the CAST administration team. This gives an indication of the amount of input required to follow a simple follow-up protocol. 'Successful' phone calls are those in which the administrator actually spoke to the participant. A total of 172 phone calls (of which 67 which were successful) and 50 repeat mailings were made to the 24 pilot participants by the end of the nine month follow-up to achieve the response rates detailed in Table 1. If the main trial which aimed to recruit 600 patients required the same amount of prompting this would equate to over 4000 telephone calls and over 1000 repeat mailings.

Table 3 Follow-up procedure details for CAST pilot study

Participant ID	4 week follow-up			12 week follow-up			9 month follow-up			Totals per participant	
	Phone calls made	Successful phone calls	Repeat mailings	Phone calls made	Successful phone calls	Repeat mailings	Phone calls made	Successful phone calls	Repeat mailings	Phone calls	Repeat mailings
1002	1	1	0	0	0	0	5	2	2	6	2
1004	0	0	0	5	1	1	10	0	2	15	3
1005	1	1	0	0	0	0	2	1	1	3	1
1006	1	1	1	1	1	0	2	2	1	4	2
1007	3	1	0	0	0	0	4	3	1	7	1
1008	0	0	0	0	0	0	2	2	1	2	1
1009	2	2	2	4	1	2	6	4	1	12	5
1010	1	1	0	1	0	1	4	4	1	6	2
1011	3	0	3	3	0	0	0	0	0	6	3
1012	1	1	1	2	0	2	3	3	0	6	3
1013	1	1	0	2	1	1	0	0	0	3	1
1014	0	0	0	0	0	0	3	1	1	3	1
1015	0	0	0	0	0	0	1	0	0	1	0
1016	0	0	0	1	0	1	1	1	0	2	1
1017	3	1	2	1	1	0	6	4	2	10	4
1018	3	1	1	1	0	1	5	3	1	9	3
1019	0	0	0	1	1	1	6	4	1	7	2
1020	0	0	0	0	0	0	3	3	0	3	0
1021	4	3	1	1	1	0	6	0	1	11	2
1022	0	0	0	3	2	0	3	0	1	6	1
1023	0	0	0	3	2	1	5	1	1	8	2
1024	6	2	2	5	0	1	7	1	1	18	4
1025	5	0	1	6	0	1	4	0	1	15	3
1026	1	0	1	4	0	1	4	1	1	9	3
Total	**36**	**16**	**15**	**44**	**11**	**14**	**92**	**40**	**21**	**172**	**50**
										Mean: 7.16	**Mean: 2.08**

4.3.1.4 Response categories

As CAST was being used as a vehicle to explore response issues, it was felt that a summary of how much prompting was required by each participant would be useful. This would allow a deeper analysis of response behaviour in subsequent chapters. At the end of the follow-up procedures at each time point in CAST, participants were either classed as a 'responder' or 'non-responder'. By having the information available as to how much prompting participants required, it was possible to refine responders into 'keen (or early) responders' or 'reluctant responders'. The concept of these different types of responder is well recognised in survey research and the term 'reluctant responder' was first used by Robins in 1963.[118] Whether keen responders differ from reluctant responders in important ways which could bias the results has also been a debated subject in survey research.[119, 120]

To obtain the information required to class CAST participants as either keen or reluctant responders 'response categories' were formed. These categories mapped directly onto the follow-up procedures protocol shown in Figure 1. Participants were assigned a response category at the end of the follow-up procedures at each time point. Table 4 gives a summary of the response categories:

Table 4 ***Response categories for CAST participants***

Response Category	Questionnaire return	Corresponding step in follow-up procedures (see Figure 1)
1	Returned with no prompting	Step 1
2	Returned after one telephone prompt	Step 2
3	Returned after one telephone prompt and second copy of questionnaire sent	Step 3
4	Returned after two telephone prompts and second copy of questionnaire	Step 4
5	Not returned but core outcome questions answered over telephone	Step 5
6	Not returned and unable to contact for core outcomes	Step 6

There is no detail in the available literature as to how much prompting is required before a participant is classed as a 'reluctant' responder. It was therefore necessary to establish this in a way which would be logical in the context of CAST and the subsequent analysis of response. Table 5 outlines how the response categories were grouped into types of responder:

Table 5 ***Types of responder based on response categories***

Response category	Type of responder
1 and 2	Keen
3 and 4	Reluctant
5 and 6	Non-responder

4.3.1.5 Analysis of response categories

Table 6 gives the response categories assigned to the CAST pilot participants based on the amount of prompting they required at each follow-up time point:

Table 6 ***CAST pilot response categories***

Participant ID	Response category		
	4 weeks	12 weeks	9 months
1002	2	1	3
1004	1	5	6
1005	2	1	3
1006	3	2	4
1007	2	1	4
1008	1	1	3
1009	4	5	5
1010	2	3	5
1011	6	6	1

1012	3	4	2
1013	2	3	1
1014	1	1	3
1015	1	1	1
1016	1	3	2
1017	6	5	5
1018	4	3	4
1019	1	5	5
1020	1	1	2
1021	6	5	6
1022	1	2	3
1023	1	3	6
1024	6	6	6
1025	6	6	6
1026	3	3	4

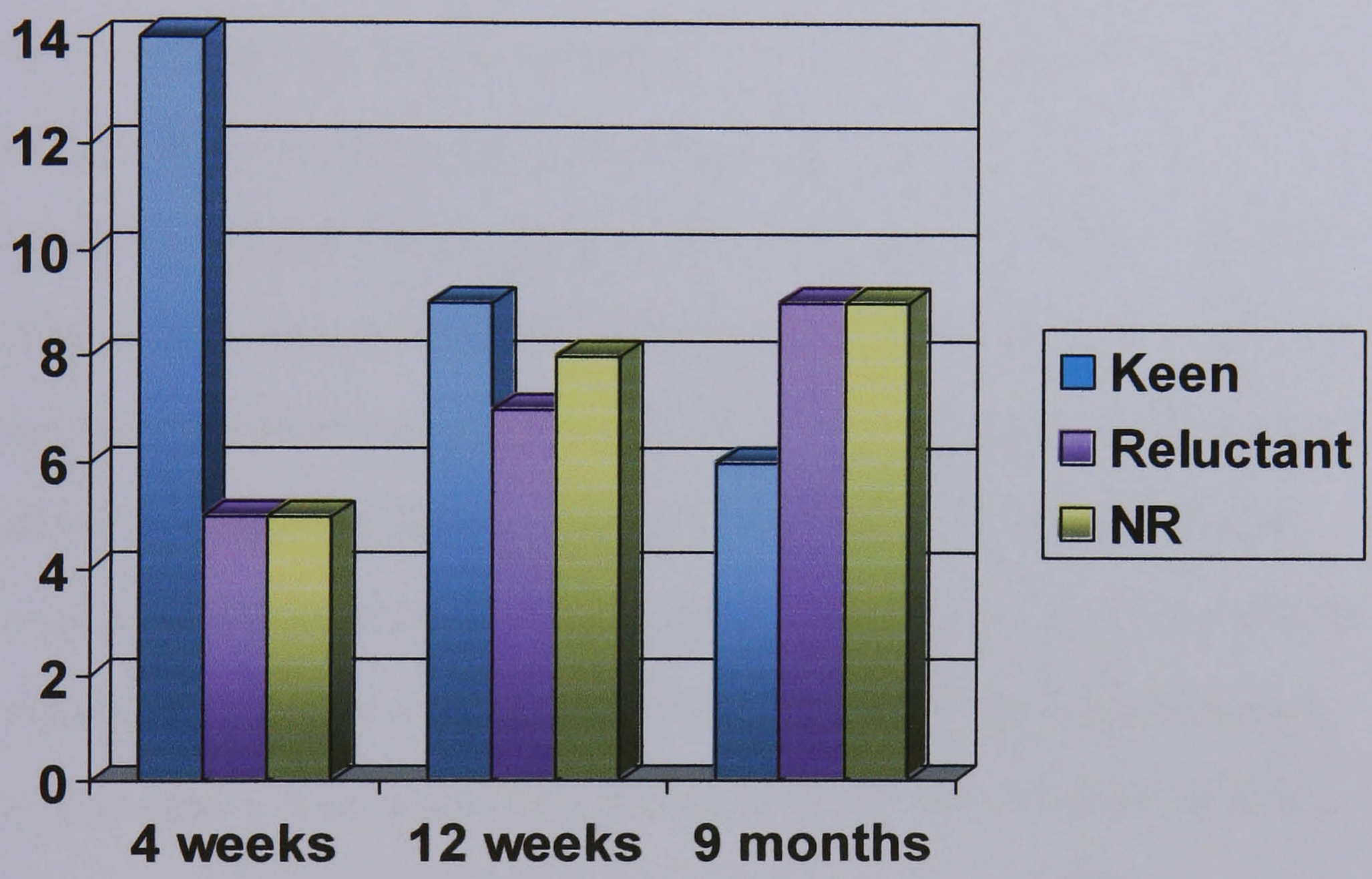
The response categories were also summarised for each time point as shown in Table 7:

Table 7 ***Number of CAST pilot participants in each response category at each time point***

Response category	**Follow-up time point**		
	4 weeks	**12 weeks**	**9 months**
1	9	7	3
2	5	2	3
3	3	6	5
4	2	1	4
5	0	5	4
6	5	3	5

Using the groupings outlined in Table 5, Figure 2 illustrates the pattern of 'keen', 'reluctant' and 'non-responders' (NR) at each time point:

Figure 2 ***Response types of CAST pilot participants at each time point***



4.4 Discussion

An analysis of the response issues in the CAST pilot phase allowed the conceptualisation of several processes to enable subsequent deeper analysis of response to be performed. Firstly, a standardised follow-up protocol was established. This meant that the participants involved in the subsequent investigation of a method of improving response which was nested within CAST (detailed in chapter 6) all received the same amount of

prompting. Standardising the follow-up procedures therefore reduced a potential source of bias in the proposed nested trial. Secondly, by establishing 'response categories' the response behaviour of CAST participants could be summarised at each time point enabling an analysis of how much prompting was required by each participant. Finally, by further refining participants into 'response types', differences in 'keen' and 'reluctant' responders could be studied. This was particularly useful in the analysis of the qualitative study detailed in chapter 7.

A predictable pattern emerged form the CAST pilot data in that the amount of keen responders decreased at each subsequent time point and reluctant and non-responders increased. CAST is a trial of treatments for acute ankle sprain. It is expected that many patients would have recovered from their injury by the 12 week follow up point and that most would have recovered at nine months. Recovery combined with declining enthusiasm to participate as time passes could explain the difficulties in maintaining follow-up at the later time points. This pattern has been seen in previous studies investigating long term outcome following ankle sprain.[121-126]

4.5 Chapter summary

This chapter outlined the methodology of CAST to provide a contextual reference for subsequent chapters. The way in which the CAST pilot phase was used to clarify response procedures was detailed. CAST was an ideal vehicle for investigating various issues surrounding response to postal

questionnaire follow-up in a clinical trial. It provided a 'real-life' trial situation which would make the results of embedded studies looking at response issues pragmatic and relevant.

A logical first step in the investigation of response is to establish what methods have so far been used effectively to improve response rates in clinical trials. This is the subject of the next chapter.

5 Chapter 5

Systematic review and meta-analysis of methods of improving response to postal questionnaire follow-up in health research

This chapter aims to evaluate the current literature regarding methods of improving response to postal questionnaire follow-up in clinical trials. To satisfy this aim a systematic review with a view to meta-analysis was conducted. Systematically reviewing the literature ensures a comprehensive literature search and pre-specified inclusion criteria to avoid bias.

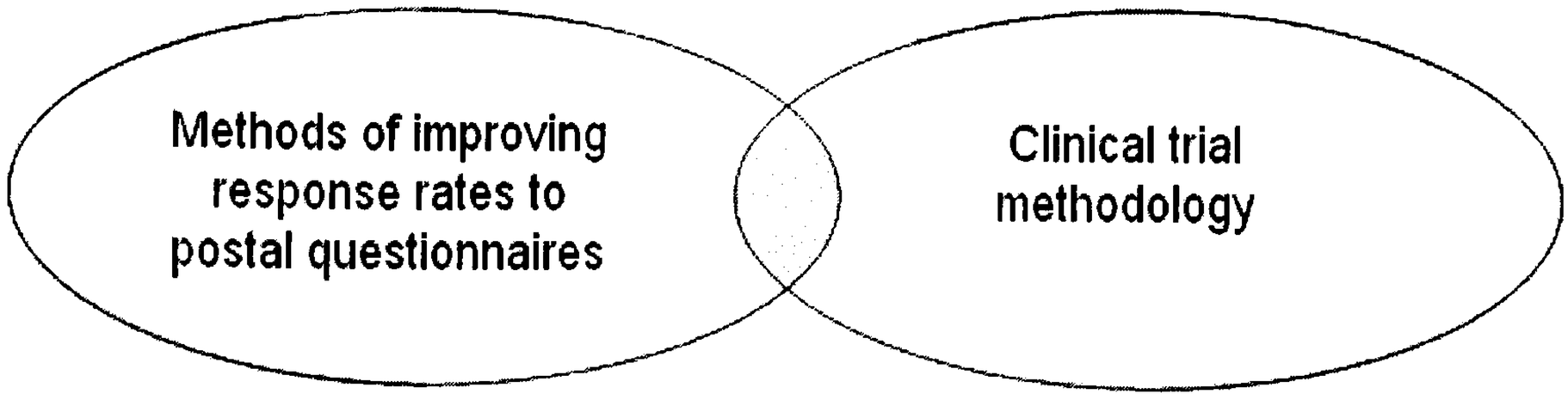
The National Library of Medicine definition of 'meta-analysis' is 'A quantitative method of combining the results of independent studies (usually drawn from the published literature) and synthesizing summaries and conclusions which may be used to evaluate therapeutic effectiveness, plan new studies, etc., with the application chiefly in the areas of research and medicine'. Meta-analysis has become popular in recent years due to the huge increase in available information and number of clinical trials being conducted.[127] Although some authors have highlighted the dangers of bias within meta-analyses,[128-130] the technique is superior to a narrative systematic review by providing an estimate of the overall treatment effect. The decision was therefore made to conduct a systematic review for this

chapter of this thesis together with a meta-analysis if the studies found could be appropriately combined. The appropriate examination of heterogeneity across individual studies can provide in itself useful information with which to guide rational treatment decisions. Furthermore, the results can demonstrate areas with insufficient available evidence indicating that new, adequately sized trials are required.[131]

This systematic review is reported according to the QUOROM statement which aims to improve the quality of reporting of meta-analyses of randomised trials.[132]

5.1 Background

When conducting a search for methods of improving response to postal questionnaires there is a plethora of survey research data which are available for reference. The focus of this thesis is response issues in *clinical trials.* No assumptions can be made that methods of improving response found in survey research can be applied with equal success to the clinical trial setting. This specific area represents the other end of the spectrum with a paucity of literature from which to draw conclusions. This systematic review seeks to refine the reviews previously carried out on methods of improving response rates to postal questionnaires. The following Venn diagram helped to clarify the area of investigation:



The shaded area represents the specific area under investigation for this systematic review. Trials were sought which combined the areas of improving response to postal questionnaires and clinical trial methodology.

Following an initial broad search of the literature it became apparent that virtually no studies exist which deal specifically with postal questionnaire response in clinical trials. A decision was therefore made to widen the search criteria to include any type of health care research methodology. This returned a far greater number of references. Many of these references, however, did not investigate health related questionnaires sent to actual patients. The recipients of the questionnaires were often members of the general public who were not actually receiving medical treatment. A key aim of this thesis is to establish whether clinical trial participants differ in their response behaviour from general survey participants. The search was therefore repeated limiting the studies to health care studies on patient populations. The terms 'health care studies' and 'patient populations' will be defined in the next section.

The findings of this review will, therefore, be of greater relevance to the context of this thesis and to health researchers using postal questionnaires.

5.2 Methods

A systematic review with a meta-analysis.

5.2.1 Search Strategy

Randomised trials of methods of improving response to postal questionnaires in health care research were identified. Six electronic bibliographic health care and medical databases were searched for relevant trials. The search strategy is shown in Table 8.

Table 8 ***Electronic bibliographic databases searched and search strategy used***

Database	*Host*	*Search Strategy*
Medline (1996-2004)	Ovid	**1.** Health care survey* *or* Questionn*
Embase (1980-2004)	Ovid	**2.** Respons* *or* Respons* adj rate *or* follow adj up *or* return
CENTRAL (1980-2004)	Update Software ltd	**3.** Post* *or* mail*
Cochrane database of systematic reviews (1980-2004)	Update Software ltd	**4.** Enhanc* *or* improv* *or* promot* *or* increas* *or* influenc* *or* maximis*
PsycINFO (1990-2004)	Ovid	**5.** Remind* *or* letter* *or* postcard* *or* incentiv* *or* reward *or* money *or* payment *or* lottery *or* prize *or* personalis* *or* sponsor *or* length *or* style *or* format *or* appearance *or* colour *or* color *or* stationary *or* envelope *or* stamp *or* postage *or* certified *or* registered *or* telephone *or* notice *or* dispatch *or* deliver *or* sensitive *or* disseminate
National Research Register (2000-2004)	DoH (Web version)	**6.** Randomi* *or* control* *or* trial*
		7. 1 *and* 2 *and* 3 *and* 4 *and* 5 *and* 6

The reference lists of identified trials and reviews were also searched. Authors of relevant trials and reviews were contacted to identify unpublished trials. The Journal of Epidemiology and Community Health was hand searched because most relevant trials were found in this journal. The BMJ 'Cite Track Alert' service[133] was used to alert for articles citing the most recent relevant review[15] and the 'Biomail' Medline search service[134]

was used with the search terms of ('clinical trial') and ('follow-up' or 'questionnair*'). There were no language restrictions.

5.2.2 Study Selection

All identified randomised trials of any method of improving response to postal questionnaires in a health care context were evaluated for study inclusion. 'Health care research' is defined as the questionnaire being used in a clinical trial, survey or observational study of health state and containing questions relating to aspects of a person's physical, mental or social well-being (based on the WHO definition of health[135]). Only studies that recruited patient populations were included. A 'patient' is defined as a person who is receiving medical or surgical treatment.[136] Studies in which participants were recruited via GP patient lists but were not actively receiving medical treatment were excluded. A list of excluded studies, with reasons for exclusion, is given in appendix 3. The outcome to assess the effect of the interventions was a comparison of the percentage of questionnaires returned after all follow-up efforts. All potentially relevant studies were checked for study quality independently by two reviewers.

5.2.3 Quality assessment

It is recommended that controlled trials selected for inclusion in a systematic review or meta-analysis should be assessed for quality. Such quality assessment limits bias in conducting the review and helps to guide the interpretation of the findings of individual studies.[137] Poor quality empirical

studies have been shown to produce systematically different results (e.g. larger treatment effects).[138] If the results of these studies are incorporated into a meta-analysis the findings of the meta-analysis may be misleading.[139]

Since the early 1960s checklists for assessing trial quality have been available, the first being published in 1961.[140] A relatively recent development in the area of quality assessment is the emergence of quality assessment scales. Scales differ from checklists in that each quality item has a numeric score attached to it enabling the generation of an overall summary score.[141] The first scale was published in 1981[142] and it is estimated that up to 50 scales are now available and the number is likely to keep increasing.[143]

Much of the variation between different quality scales stems from the developers' definition of the term 'quality' and this is a construct that is difficult to define.[144] Jadad et al[145] define the quality of a randomised controlled trial as 'the likelihood of the trial design to generate unbiased results and approach the 'therapeutic truth''. Verhagen et al,[146] however, point out that this definition only covers the dimension of internal validity of a trial and they imply that external validity is also an important component of quality. It is agreed, however, that there are four main sources of systematic bias in clinical trial research. These are systematic differences between comparison groups in terms of (i) the patients' characteristics (selection bias), (ii) the provision of care apart from the intervention (performance

bias), (iii) the assessment of outcomes (detection bias) and (iv) dealing with patient withdrawals (attrition bias).[137]

For researchers embarking on a systematic review or meta-analysis the abundance of available tools to assess trial quality is a source of some dismay. This has been recognised with recent attempts to assess the quality of the quality scales.[137, 139, 141, 147] Moher et al[141] reviewed 25 different quality assessment scales and came to the disappointing conclusion that all but one of the scales have major weaknesses and should therefore be used with caution. The one exception was a scale developed by Jadad et al[145] (who coincidentally was a co-author of the review paper!). A major criticism of the scales reviewed was the lack of rigour in their development. The scale developed by Jadad et al was the only one that used standard scale development techniques although it has been noted that this does not automatically make it better than the other scales.[144] The Jadad scale has also been criticised for placing too much emphasis on the quality of the trial report rather than the methodological quality of the paper. Furthermore, this scale addresses randomisation but not allocation concealment which is also seen as a weakness of the scale.[147]

The problem of using scoring scales for assessing trial quality has been highlighted by several other authors. Juni et al[144] highlight the considerable variation in scales in terms of dimensions covered, size and complexity and how this can lead to contradictory results. Meta-analysis of

the same studies using different quality assessment scales has been shown to produce opposite conclusions of treatment effect.[147]

Scales have also been criticised for placing too much emphasis on the reporting of the trial rather than the methodological quality (as in the Jadad example above). The quality of trial reporting poses a dilemma for anyone reviewing a paper. A well-conducted but badly reported trial will be misclassified as poor. This is because the 'guilty until proven innocent' approach is usually used whereby quality is assumed inadequate unless information is provided to the contrary in the reporting of the trial.[144] Of course, the opposite also applies in that a biased but well reported trial will receive more credit than it deserves. The problems of poor reporting of trials is well recognised and led to the development of the Consolidated Standards of Reporting Trials (CONSORT) statement in 1996.[148] Soon after publication the CONSORT statement was endorsed by several major journals and many more are following this lead.[143] It will take time, though, for the message to spread and research published prior to 1996 had no reporting guidelines.

For the purposes for the systematic review conducted as part of this thesis it was considered necessary to follow the recommendations of the Cochrane Collaboration and perform some form of quality assessment on the selected papers. A decision then had to be made as to which assessment tool to use. For the reasons outlined above, this was not an easy decision to make.

Although much has been written on the quality of quality assessment scales no recommendations have emerged to guide the reviewer to the best scale currently available since no 'gold standard' for the 'true' validity of a trial exists.[137] The Cochrane Handbook for Systematic Reviews of Interventions[137] suggests that although quality scales offer 'appealing simplicity' due to the fact that a summary score is generated, this approach is not supported by empirical evidence.[138] They recommend using 'simple' approaches for assessing trial validity which have not been shown to be any less reliable than more complex scoring scales. Simple approaches have the advantage of taking less time to complete and are less likely to confuse the quality of reporting with the validity of the study.[137] These simple approaches use only a few assessment criteria but the criteria address important threats to the validity of study results (eg allocation concealment). An overall assessment of the validity of the study is obtained rather than a 'score'. Juni et al[144] also advocate this type of approach (which they term the 'component' approach) and point out that the importance of the assessment criteria will vary between the contexts in which the trials are performed. This is something that scoring scales fail to take into account. Juni et al conclude that there is currently no consensus on whether scales or the simple component approach is preferable. The Cochrane Handbook for Systematic Reviews of Interventions, however, concludes that due to the problems with scales outlined above, 'it is preferable to use simple approaches for assessing validity'.[137] This

appears to have been taken on board by Cochrane reviewers. A recent review of the use of quality assessment in 36 reviews published in the Cochrane Database of Systematic Reviews (CDSR) demonstrated that 92% used components and none used scales.[149]

Based on the above arguments the decision was made to use the simple approach for assessing trial quality for the purposes of the systematic review of methods of enhancing response to postal questionnaires. Important aspects of the validity of trials investigating response rates are seen to be;

1. Adequate randomisation (Selection bias)
2. Performance bias (i.e. one group receiving another method of improving response other than the intervention)
3. Blinding of the assessor to intervention allocation (Detection bias)

As attrition or loss to follow-up is the primary outcome of this review it is not relevant to include attrition bias as a method for assessing trial quality.

A fairly recent addition to the debate of what items should be used to assess the quality of a randomised controlled trial has been offered by Verhagen et al.[146] They point out that the external validity of a trial is often overlooked when assessing trial quality with most emphasis being placed on the components of internal validity already mentioned. They developed a criteria list for quality assessment of randomised controlled trials for use in

systematic reviews by Delphi consensus. The resulting 'Delphi List' contains eight items relating to both internal and external validity of a trial and also statistical considerations. These eight items are shown in Table 9:

Table 9 ***Delphi list for quality assessment[146]***

1	Treatment allocation: Randomisation performed Allocation concealed
2	Groups similar at baseline
3	Eligibility criteria specified
4	Outcome assessor blinded
5	Care provider blinded
6	Patient blinded
7	Point estimates and measures of variability presented for primary outcome measures
8	Intention to treat analysis

The authors point out that this list is not intended to replace existing scales or quality criteria lists but to be used alongside existing methods of quality assessment.

The Delphi List was therefore used alongside the simple approach to quality assessment for this review. This created a method of assessment which was based on the available literature and recommendations outlined in the Cochrane Handbook for Systematic Reviews of Interventions.

5.2.4 Quantitative Data Synthesis

The data were analysed using the Cochrane review manager software (RevMan version 4.2; Oxford, UK).

5.2.4.1 Measures of effect

To allow comparison between studies, individual study results have to be expressed in the same format.[150] In this review, the end point in all studies is dichotomous – response or non-response to postal questionnaire follow-up. This means that it is appropriate to calculate either relative risks or odds ratios to evaluate the effects of the interventions. 'Risk' is the number of people incurring an event divided by the total number of people. 'Relative risk' (or risk ratio) is the risk of an event occurring in the intervention group divided by the risk of it occurring in the control group.[150] 'Odds' are defined as the ratio of a number of people incurring an event to the number of people who have non-events.[10] The 'odds ratio' is the odds of an event occurring in the intervention group divided by the odds of it occurring in the control group. The decision on whether to use odds ratios or relative risks to compare trials has been the subject of some debate.[150-153] It has been noted that odds ratios are usually interpreted as being equivalent to the relative risk.[137, 151] This is the case in studies where the risks (or odds) in the two groups being compared are both small (eg less than 20%). However, as the risk in either group rises above this, the gap between the odds ratio and relative risk will widen. If the two terms are seen as interchangeable, this will (and has been shown to) lead to

misinterpretation of study results.[137] It has also been suggested that although odds ratios have the strongest mathematical properties, they are harder to understand and apply to clinical practice than risk ratios. Furthermore, 'risk' is a concept more familiar to patients and health professionals than 'odds'.[137, 150] Some authors, therefore, suggest that the results of trials and systematic reviews should be reported as relative risks 'unless there is a convincing argument otherwise'.[152]

For this review, the results were reported as relative risks. This takes into consideration the arguments outlined above. It is anticipated that the risks and odds will be small in the groups being compared (i.e. changes in response rate with and without the intervention designed to improve it). It is likely, therefore, that the relative risk and odds ratio values will be comparable. To avoid any confusion, however, relative risks were reported together with 95% confidence intervals (CI 95%).

5.2.4.2 Identifying and measuring heterogeneity

The specific nature of the search used in this review reduced a potential source of clinical heterogeneity which could have confounded the results of previous reviews. This was achieved by limiting the setting and population of the included trials to 'health care research' and 'patients'. Furthermore, it was anticipated that the search would identify trials of several different methods of improving response to postal questionnaires. This could be another potential source of clinical heterogeneity. It was, therefore, decided *a priori* that trials of similar methods of improving response would be

grouped for comparison. It was felt important to control such sources of clinical heterogeneity to prevent their contribution to statistical heterogeneity. Statistical heterogeneity is evident when the observed treatment effects are more different from each other than would be expected due to random error alone.[137] A common test for statistical heterogeneity is the chi-square (χ^2) test. This test, however, has been criticised for failing to detect true heterogeneity between studies as significant, especially in meta-analyses which contain small numbers of studies.[154] It has been argued that since meta-analyses always contain clinical and methodological diversity, statistical heterogeneity is inevitable.[137] Because of this, methods have been developed to quantify the effect of heterogeneity rather than just give an indication of its existence.[154, 155] Higgins and Thompson[155] developed such an approach which they termed 'I^2'. This statistic describes the percentage of total variation across studies that is due to heterogeneity rather than chance. The result of an I^2 test lies between 0% and 100% (negative values of I^2 are put equal to zero). A value of 0% indicates no heterogeneity and larger values show increasing heterogeneity.[154] A value greater than 50% may be interpreted a substantial heterogeneity.[137]

For the purposes of this review both the χ^2 and I^2 statistics are presented. A value of $p<0.10$ was used in the χ^2 test to reflect significant heterogeneity.

5.2.4.3 Summarising measures of effect across studies

In a meta-analysis, once the treatment effects for each individual study have been calculated in a standardised format, (as described in section 5.2.4.1) the results are combined. This yields an overall statistic that summarises the effectiveness of the experimental intervention compared with the control intervention, and the uncertainty around it (in the form of a confidence interval).[137] A variety of statistical techniques are available for taking into account the less robust results gained from small studies. Such techniques employ a weighted average of the results. Larger trials generally have more influence than the smaller ones,[150] although this is not always the case. Smaller trials which contain more information (eg higher incidence of outcomes or smaller standard deviation) may be allocated more weight. The way in which trials are weighted differs depending on the method of combining the studies. Two broad categories of statistical techniques for combining studies are the 'fixed effects' model and 'random effects' model. The difference between the two models relates to the way the variability of the results *between* studies is treated.[150] The choice of model to use is determined by certain assumptions about the underlying data. A fixed effect meta-analysis assumes that the 'true' effect of treatment is the same value (ie 'fixed') in every study.[156] This implies that the differences among study results are due solely to the play of chance.[137] A random effects model assumes that the 'true' value varies across studies and that the effects are randomly distributed. This model leads to relatively more weight being given

to smaller studies than the fixed effect model.[150] A random effects model is more conservative resulting in wider confidence intervals.

Taking the above points into account, a random effects model was used for the purposes of this review. It was felt that even though studies investigating similar methods of improving response were analysed separately, there would still be a variation of the 'true' effect of the intervention between studies. This could be due to such things as variations in the settings and populations of the studies. Also, there was some variation in the degree of similarity of the methods of improving response within the comparison groups. For example, one comparison group is 'Follow-up strategies'. This ranges from 'telephone follow-up' to 'mail follow-up' which are similar but clearly not identical interventions. It is recognised, however, that using a random effects model will not simply take into account such heterogeneity so that it can be ignored. Attempts were made to explain any heterogeneity and its effect on the results of the meta-analysis.

The random effects meta-analysis used by RevMan and, therefore, used in this review is the DerSimonian and Laird method.[157]

5.2.4.4 Sensitivity analysis

Conducting a sensitivity analysis tests how robust the results of the review are relative to the assumptions made in the methodological choices and conduct of the review.[137] The meta-analysis was therefore repeated examining the effect of the following factors on the overall results:

- **Replacing the random effects model with a fixed effects model**

A decision was made to analyse the data using a random effects model as outlined above. It is recognised, however, that the assumptions made about the underlying data upon which this decision was based are not necessarily correct.[156] Re-running the analysis using a fixed effects model therefore highlighted any major differences in the estimates of effect using the different models.

- **Methodological quality of included studies**

A major feature of all methods of assessing study quality is the way the study deals with randomisation and allocation concealment. Inadequate concealment of treatment allocation is often associated with larger treatment effects.[138, 158] The effect that this has on overall estimates can be assessed by excluding trials deemed to be of poor quality from the analysis.

- **Study size**

Studies reporting statistically significant results are more likely to get published than non-significant findings and this is a well recognised potential source of bias.[137] A sensitivity analysis excluding small studies can be used to examine such publication bias.[150] Smaller effects can be statistically significant in larger studies. If publication bias is present, it is expected that the larger published studies will report the smaller effects.[150] Re-running the analysis excluding the smaller studies will identify whether their results have any effect on the overall estimate of

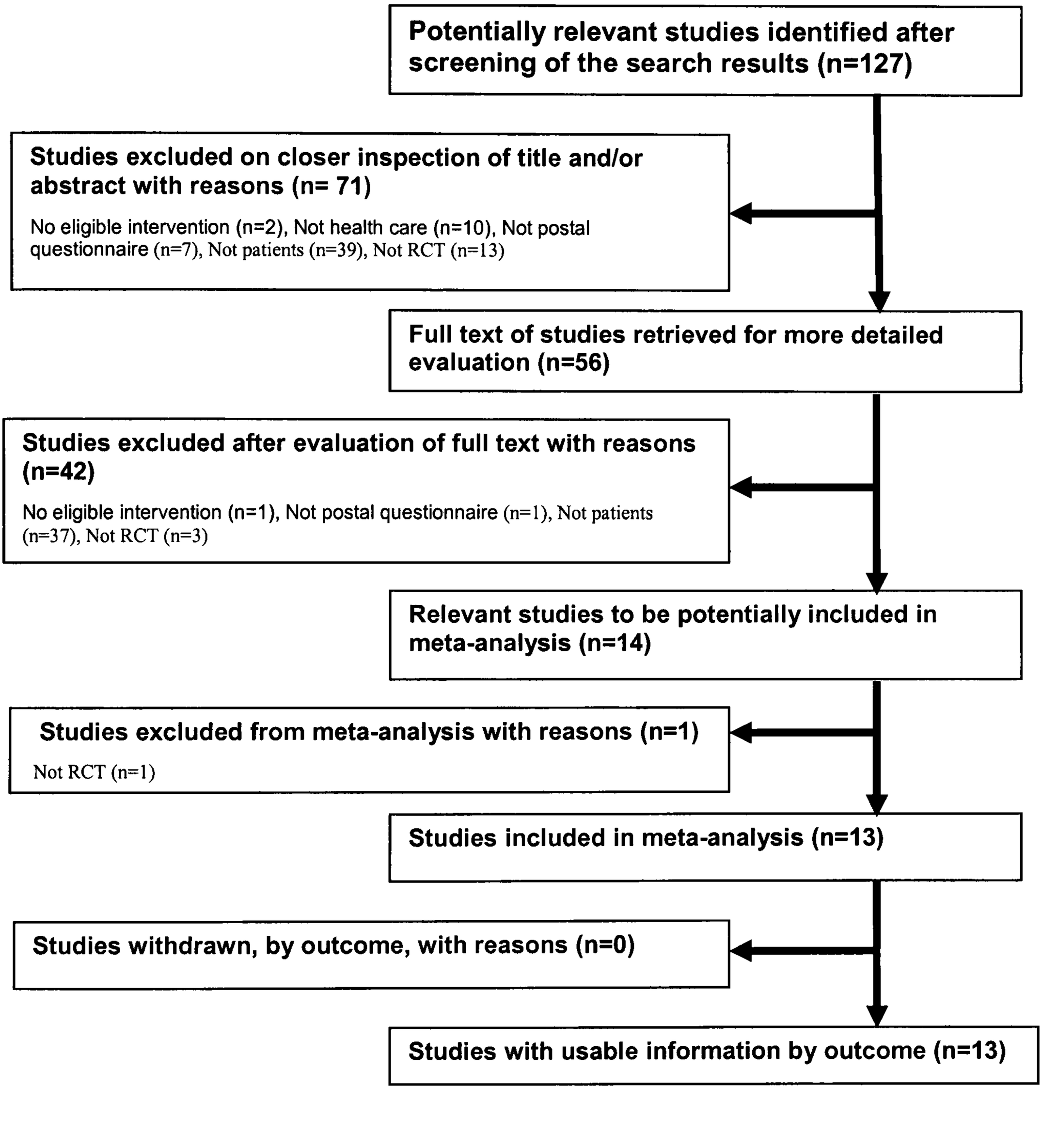
effect. The traditional method for assessing publication bias is the 'funnel plot'. The use of this method in this review is, however, limited due to the small number of included studies.[137, 156]

5.3 Results

5.3.1 Trial Flow

The search identified 13 randomised trials including 25607 participants that fulfilled the inclusion criteria.[22, 117, 159-169] Figure 3 gives a flow chart summarising the study selection process:

Figure 3 ***Flow chart of study selection process***



5.3.2 Study Characteristics

The studies evaluated five different methods of enhancing response to postal questionnaires. These methods were: questionnaire length,

incentives (cash, prize draw, lottery or phone card), question order, reminder strategies and including an information brochure with the questionnaire. One paper reported results in two distinct patient groups (angina and asthma) and these are presented as separate studies.[166] Another paper described two separate interventions (questionnaire length and incentives) and these are also reported as separate studies.[162] Six papers contained information regarding missing data from the returned questionnaires[22, 159-161, 166] but used different interpretations of missing data. All the studies incorporated their randomised trial of methods of improving response into an existing research study. The majority of the studies nested their trial of enhancing response within a patient survey. None of the studies nested their study of methods of improving response into a randomised clinical trial. Table 10 gives details of extracted data.

Table 10 - Extracted data of randomised trials of methods of improving response rates to postal questionnaires in health care research

Author	Country	Main Study Type [a]	Patient Characteristics	Method of improving response rate	Questionnaires returned/Questionnaires sent (% response rate) after first mailing	Questionnaires returned/Questionnaires sent (% response rate) after all mailings	Follow up procedures
				Questionnaire length			
Dorman 1997	UK	Survey	Stroke survivors	1. Short Questionnaire (EuroQol)	1. 747/1125 (66%)	1. 905/1125 (80%)	Initial Q sent with pre-paid envelope and personalised letter.
				2. Long Questionnaire (SF36)	2. 679/1128 (60%)	2. 849/1128 (75%)	Reminder letter and Q sent after 2 weeks
					p=0.002	p=0.003	
Ingesias [b] 2000	UK	Survey	Women with osteoporotic risk factors	1. Short Questionnaire (4-5 pg)	1. 270/553 (48.8%)		Information letter and pre-paid envelope sent with questionnaire. No reminders or follow up
				2. Long Questionnaire (7 pg)	2. 119/300 (40.5%) p=0.04, 95%CI=0.3-16.6		
Jenkinson 2003	UK	Survey	Recently discharged in-patients	1. Short questionnaire (4 pgs)		1. 488/721 (67.7%)	Reminder sent to non-responders after 2 weeks. Second questionnaire sent to non-responders after further 2 weeks
				2. Long questionnaire (12 pgs)		2. 461/724 (63.7%)	
Jones (a) [d] 2000	USA	Survey	Psychiatric patients	1. Short questionnaire (SF12, MH5-5 qs of mental health taken from SF36, MH1-1 q of mental health taken from SF36)		1. 122c/219 (55.7%)	Second questionnaire sent 2 weeks after 6 month follow up point. Began telephoning non-responders 2 weeks after 6 month follow up point
				2. Long Q (SF36)		2. 26c/73 (36%)	
				Incentives			
Jones (b) [e] 2000	USA	Survey	Psychiatric patients	1. Cash incentive ($2 or $5) sent with questionnaire		1. 81c/146 (55.4%)	Second questionnaire sent 2 weeks after 6 month follow up point. Began telephoning non-responders 2 weeks after 6 month follow up point
				2. No incentive sent with questionnaire		2. 37c/73 (50.6%)	
Ward 1996	Australia	Survey	Patients on the Age-Sex Disease Register	1. Instant $1 lottery ticket		1. 73/112 (65%)	Prompt letter sent to non-responders 21 days after initial mailing. Second letter and second questionnaire sent to non-responders 30 days after initial mailing

				2. No Lottery ticket		2. 81/119 (68%) p=0.68	
Leigh Brown 1997	UK	Cohort Study	Orthopaedic out-patients	1. Aware of prize draw for £25 gift voucher		1. 461/654 (70.5%	Hand signed cover letter. Second class business reply envelolpe. Postcard reminders sent to non-responders after 10 days. Second reminder after 3 weeks
				2. Unaware of the prize draw		2. 430/653 (65.8%) p=0.07	
Evans 2004	USA	Survey	Prostate cancer patients	1. Immediate incentive (30 min phone card sent with questionnaire)	1. 406/681 (60%)		Pre-paid envelope sent with questionnaire. No reminders or follow up
				2. Delayed incentive (30 min phone card upon receipt of completed questionnaire)	2. 399/669 (60%)		
				Question order			
Dunn 2003	UK	Cohort Study	Back pain sufferers	1. Chronologically ordered questionnaire		1. 129/175 (75%)	Initial Q sent with letter from GP and pre-paid envelope. Reminder postcard sent to non-responders after 2 weeks. Second Q sent to non-responders after 4 weeks
				2. Traditionally ordered questionnaire		2. 55/84 (67%)	
McColl (a) 2003	UK	Survey	Angina Patients	1. Condition specific questions first followed by generic questions		1. 1779/2363 (75%)	Personalised cover letter and info sheet. Second class business reply envelope enclosed. Max 2 reminders to non-responders : 3 weeks after initial mailing and further letter plus second questionnaire 6 weeks after initial mailing
				2. Generic questions first followed by condition specific		2. 1738/2321 (75%)	
McColl (b) 2003	UK	Survey	Asthma patients	1. Condition specific questions first followed by generic questions		1. 1522/2382 (64%)	As above
				2. Generic questions first followed by condition specific		2. 1537/2369 (65%)	
				Reminder strategies			
Tai 1997	UK	Survey	Asthma and Diabetes	1. Telephone reminders		1. 12/94 (16.2%)	NB only non-respnders to previous follow up methods randomised

				2. Recorded delivery reminders	2. 26/98 (35.1%) p=0.008	
Salim Silva 2002	Australia	Survey	Work related neck and upper limb disorders	1. Final reminder letter plus telephone call	1. 11/29 (38%)	Cover letter and pre-paid return envelope. Reminder letter after 2 weeks. Second copy of questionnaire sent out after further 2 weeks.
				2. Final reminder letter only	2. 3/29 (10%) p=0.031	NB The study only randomised non-responders to previous follow up methods
Sutherland 1996	Canada	Pilot Study	Patients in the Canadian Diet and Breast Cancer Prevention Trial	1. Total Design Method (TDM)	1. 100/113*c* (88%)	Customary Group : 1 telephone contat after 1 month. TDM Group : Reminder postcard after 1 week. Second questionnaire to non-responders plus letter after 3 weeks. Third letter after 7 weeks registered mail
				2. Usual follow-up mailing procedure	2. 70/113*c* (62%)	
				Information brochure		
Parkes 2000	Canada	Case Control Study	Cancer Patients	1. Information brochure sent with questionnaire	1. 2829*c*/3732 (75.8%)	Cover letter and stamped return envelope sent with questionnaire. Reminder postcards sent 1-2 weeks after initial mailing.
				2. No information brochure	2. 2816*c*/3755 (75%)	Reminder plus second quetionnaire sent to non-responders 4 weeks after initial mailing. Telephone follow up began 6 weeks after initial mailing

a Type of study in which the randomised trial of response enhancement method is embedded
b Short and medium questionnaire pooled for analysis
c Figures calculated from given information
d SF12, MH5 and MH1 pooled and compared with SF36
e $2 and $5 incentive pooled and compared with standard questionnaire alone

5.3.3 Study Quality

Study quality was assessed independently by two reviewers using the two methods of assessing study quality discussed earlier. There were no major discrepancies between the assessments of study quality and minor disagreements were discussed and agreed upon. There was no need for a third party adjudicator.

5.3.3.1 The Delphi list for quality assessment

The developers of the 'Delphi List' of quality assessment criteria[146] give no instructions on how to obtain a score from their list of quality criteria. This is therefore open to personal interpretation. The Delphi list consists of eight criteria thought to give an indication of study quality. The first criterion of 'Treatment Allocation' has two parts (see Table 9 p76) effectively making the list up to nine criteria. Two of these criteria are not applicable to this review as they are concerned with blinding of the participant and the care provider. With the type of studies under review this type of blinding is not feasible (blind outcome assessment, however, is possible). This leaves seven quality criteria to assess.

5.3.3.2 The Cochrane approach to quality assessment[137]

The method of assessing study quality using the Cochrane approach is more straightforward as there are only four quality criteria (or three as used in this review). If all the criteria were met the study was deemed to be of 'good' quality, two criteria met indicated 'moderate' quality and one or less

criterion met or two or more criteria unmet indicated 'poor' quality. Quality was deemed 'unclear' if two or more criteria could not be assessed due to poor reporting.

5.3.3.3 Combining the two methods of quality assessment

The Cochrane simple approach to quality assessment[137] has some overlap with the Delphi List. Both selection bias and detection bias are covered by the Delphi List. Performance bias, however, is not. This criterion was therefore added to the seven Delphi List criteria to give a total of eight quality criteria in the final quality table used in this review. For the purposes of this review, if five or more quality criteria were met the study quality was deemed to be 'good'. If four or more criteria were unmet, quality was deemed to be 'poor'. If it was not possible to deduce whether the quality criteria had been met due to poor reporting in four or more criteria, study quality was deemed 'unclear'. The quality of studies which did not meet the criteria for 'good' 'poor' or 'unclear' was deemed to be 'moderate'. Using this method, four studies were deemed to be of 'good' quality, six were 'moderate' quality and quality was unclear from the report of three studies. The included studies were given a grade from A to D to represent their performance on the quality assessment tool as shown in Table 11:

Table 11 ***Grading used on quality assessment tool***

Grade	Quality
A	Good
B	Moderate
C	Poor
D	Unclear

This grading was based on recommendations in the Cochrane Handbook for Systematic Reviews of Interventions. It is important to point out that the meanings of A to D as used by the RevMan software are different to this. RevMan only has the facility to include allocation concealment as an indicator of trial quality. 'A' indicates adequate allocation concealment, 'B' indicates unclear allocation concealment, 'C' indicates inadequate allocation concealment and 'D' indicates that allocation concealment was not used. All the forest plots generated by RevMan for the purposes of this meta-analysis have been modified to show the grade of quality as shown in Table 11 above. See Table 12 for details of the quality assessment:

Table12 - Quality assessment scores of included studies

Author	Randomisation performed?	Allocation concealed?	Similar baseline characteristics?	Eligibility criteria specified?	Blind outcome assessment?	Adequate reporting of results?	ITT analysis?	No performance bias?	Quality score
Dorman 1997	√	?	√	√	?	√	?	√	A
Dunn 2003	√	?	√	√	√	X	√	√	A
Evans 2004	√	?	√	√	?	X	?	√	B
Iglesias 2000	√	√	√	X	√	√	√	√	A
Jenkinson 2003	√	?	?	√	?	X	?	√	D
Jones a,b 2000	√	?	?	X	?	X	?	√	D
Leigh Brown 1997	√	?	?	√	X	√	?	√	B
McColl a,b 2003	√	?	√	√	X	√	√	√	A
Parkes 2000	√	?	?	√	X	√	?	√	B
Salim Silva 2002	√	?	?	X	?	√	√	√	B
Sutherland 1996	√	?	√	√	?	X	?	√	B
Tai 1997	√	?	√	X	?	√	?	√	B
Ward 1996	√	?	?	X	?	X	?	√	D

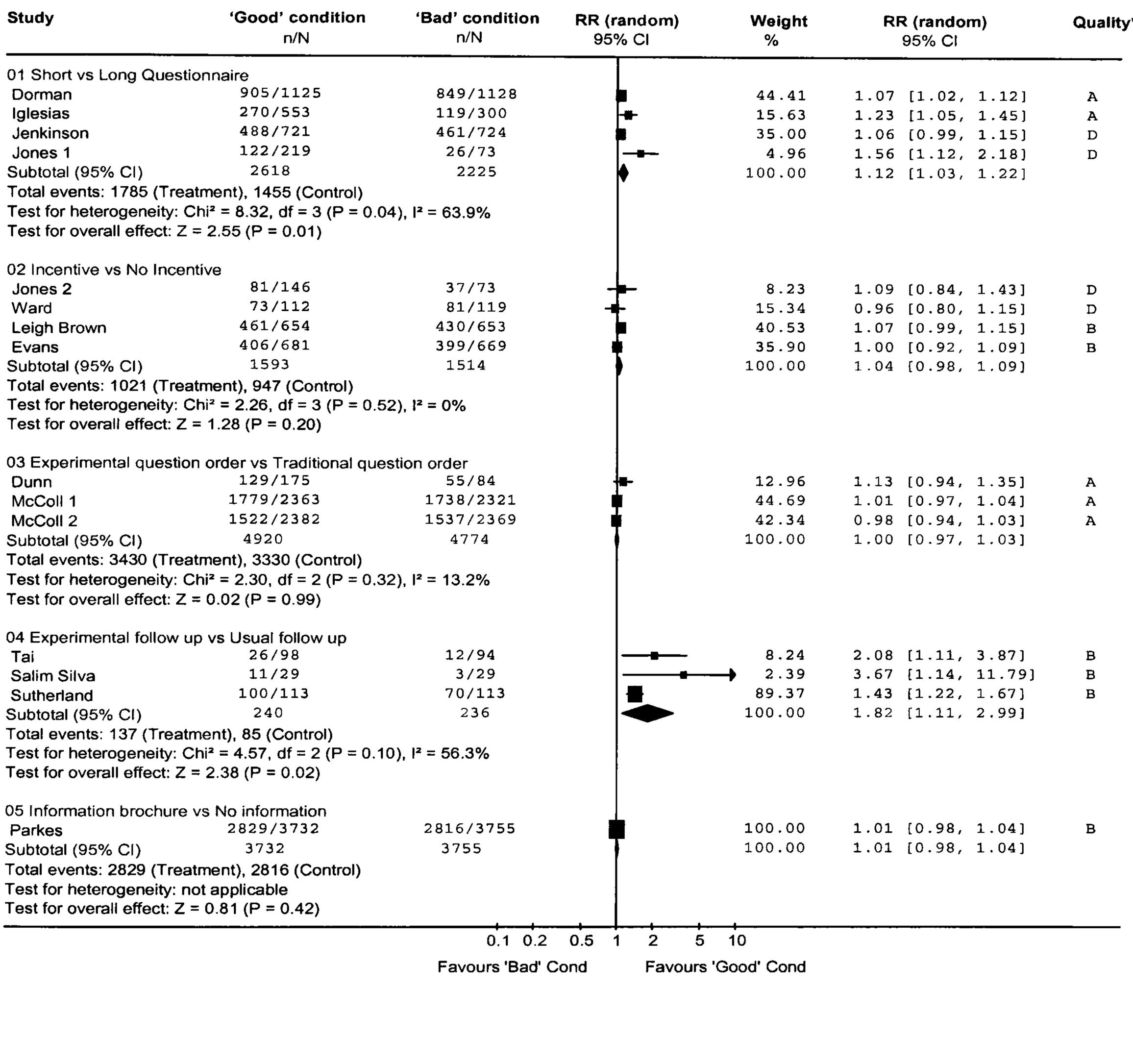
To score: 5 or more √ = 'Good', 4 or more X = 'Poor', 4 or more ? = 'Unclear', others = 'Moderate'

5.3.4 Quantitative data synthesis

Figure 4 shows the pooled relative risks and 95% confidence intervals for the five different strategies investigated for improving response rates.

Follow-up reminder systems had the most significant effect on response rates (RR 1.82, CI 95% 1.11 to 2.99 p=0.02, three trials, 476 participants) There was, however, evidence of heterogeneity between the trials in this group with an I^2 statistic of 56.3%. Shorter questionnaires improved response rates but to a lesser degree (RR 1.12, CI 95% 1.03 to 1.22 p=0.01, four trials, 4843 participants). This group also showed significant heterogeneity with an I^2 of 63.9%. 'Shorter' questionnaires ranged from seven to 47 questions and 'longer' questionnaires ranged from 36 to 123 questions. The studies investigating questionnaire length compared two or more questionnaires. The authors own categorisation of 'shorter' and 'longer' questionnaires was used. The use of incentives (RR 1.04, CI 95% 0.98 to 1.09 p=0.20, four trials, 3107 participants), re-ordering of questions (RR 1.00, CI 95% 0.97 to 1.03 p=0.99, three trials, 9694 participants) and including an information brochure with the questionnaire (RR 1.01, CI 95% 0.98 to 1.04 p=0.42, one trial, 7487 participants) had no significant effect on response rates.

Figure 4 ***Random effects model meta-analyses of methods of improving response rates to postal questionnaires in health care research – response rate after all follow-up efforts***



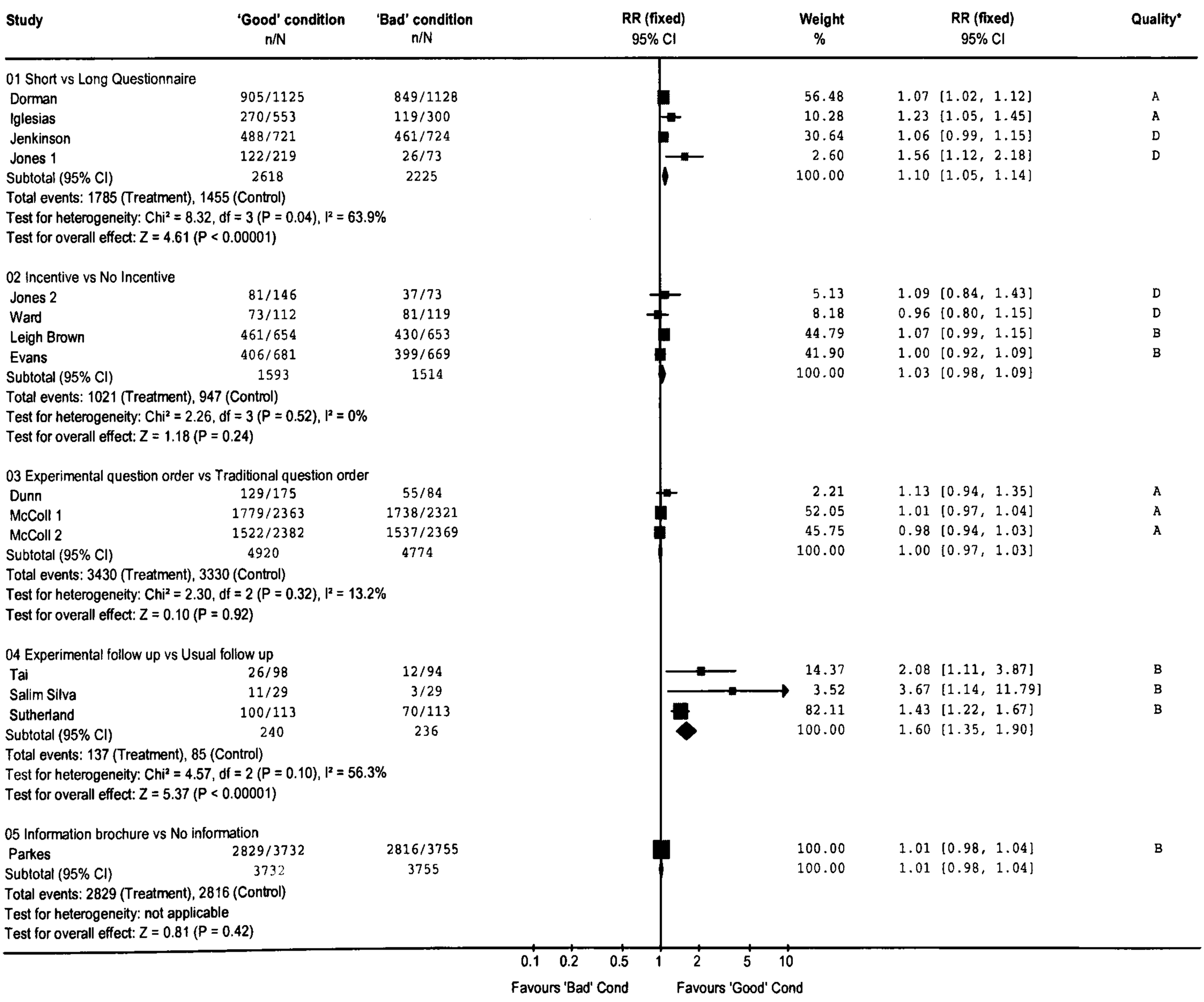
* Quality: A= Good, B= Moderate, C= Poor, D= Unclear from report

5.3.5 Sensitivity analysis

5.3.5.1 Random effects versus fixed effects model

The meta-analysis was re-run using a fixed effect model. The results of this are shown in Figure 5:

Figure 5 ***Fixed effects model meta-analysis of methods of improving response rates to postal questionnaires in health care research***

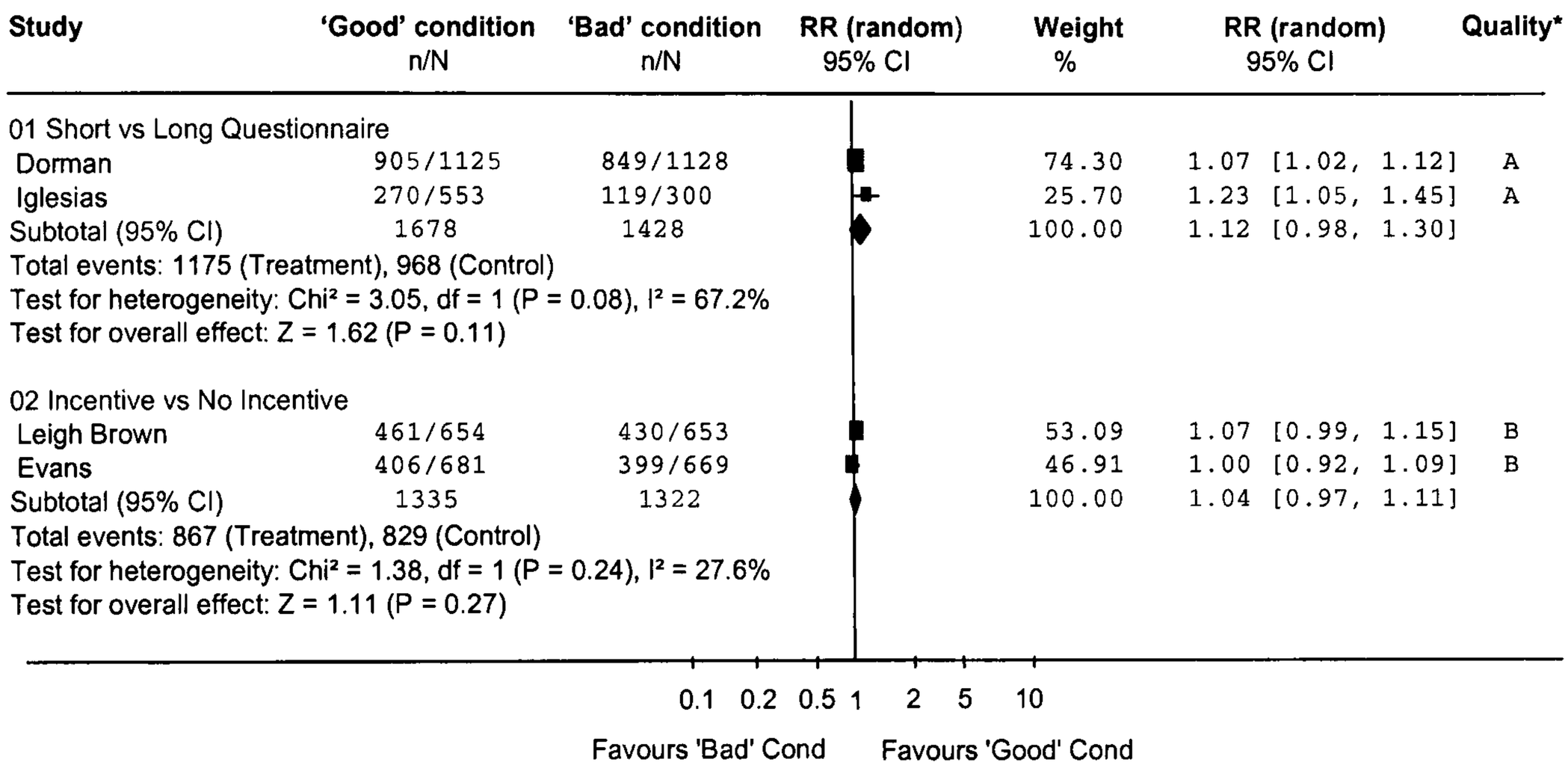


There is very little difference between in the results of the two models. This suggests that incorporating the heterogeneity between the included studies into the analysis makes virtually no difference to the result.

5.3.5.2 Study quality

None of the included studies were classed as 'poor' quality; however the quality of four of the studies was unclear due to poor reporting. Five studies were deemed to be of 'good' quality and six were 'moderate' quality. As can be seen from Figure 4, only two of the comparisons included studies of different quality. These comparison groups are: 'Short versus Long Questionnaire' and 'Incentive versus No Incentive'. It was, therefore, only possible to conduct an analysis based on study quality on these two groups. The meta-analysis was repeated on these groups excluding trials with unclear quality. Figure 6 shows the effect of study quality on overall estimates of effect for these comparisons:

Figure 6 ***Meta-analysis of methods of improving response to postal questionnaires excluding studies of unclear quality***



*Quality: A= Good, B= Moderate

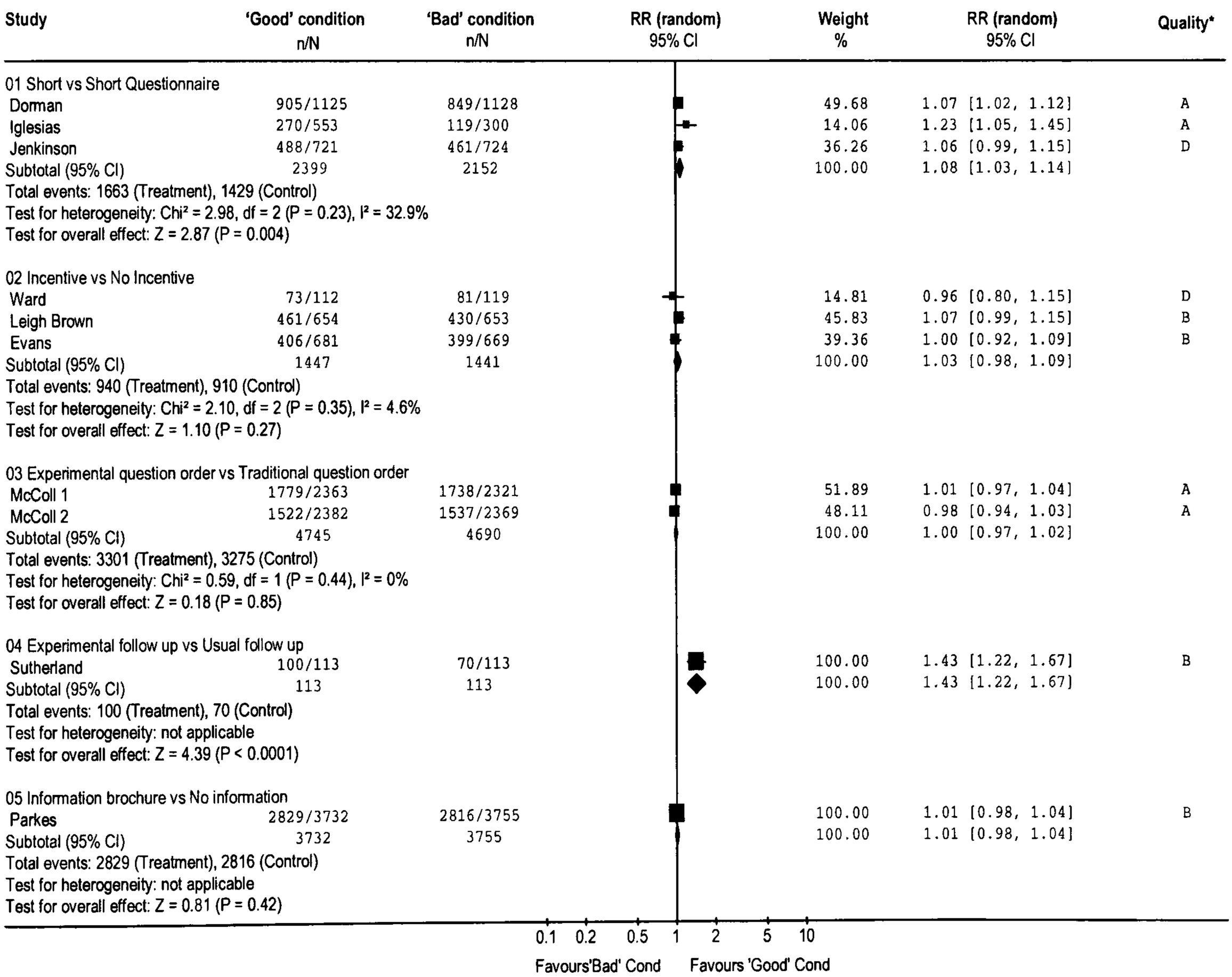
For the two comparison groups for which this analysis was possible, excluding studies of unclear quality had no effect on the overall estimates of effect.

5.3.5.3 Study size

An arbitrary figure of less than one hundred participants in either the experimental or control group was used to categorise 'small' studies. The

meta-analysis was re-run excluding small studies. Figure 7 shows the results of this:

Figure 7 ***Meta-analysis of methods of improving response to postal questionnaires excluding small studies***



*Quality: A= Good, B= Moderate, C= Poor, D= Unclear from report

On examining the overall results of the main meta-analysis as shown in Figure 4 it is evident that the largest effects are generally seen in the small studies. This is evidence of possible publication bias. However, excluding the small studies had very little effect on overall estimates of effect or on the statistical significance of the results.

5.3.5.4 Summary of sensitivity analysis

The sensitivity analysis shows that the results of this meta-analysis are robust to the choice of statistical method and to the exclusion of trials of unknown quality. It also suggests that publication bias is unlikely to have distorted the findings. Because of the small number of studies included in the review, however, it is not possible to come to a firm conclusion about publication bias.

5.4 Discussion

The main findings of this systematic review and meta-analysis are that the implementation of more intense follow-up strategies and shorter questionnaires can improve response rates. The results are more relevant to health care researchers than previous reviews. Since the most recent previous review[15] we included five new relevant studies.

5.4.1 Follow-up strategies

Three studies investigated methods of follow-up to improve response.[117, 167, 168] Although the methods of follow-up differed, all of the trials

compared a more intensive follow-up procedure with a standard method. The three included trials compared telephone, postal or recorded delivery reminders compared to usual follow-up efforts. We therefore carried out an analysis of intensive versus usual follow-up. The results suggest that increased intensity of follow-up effort may improve response rates. This conclusion, however, should be treated with caution as there was significant statistical heterogeneity in this comparison. The most likely explanation for this is the difference between the interventions of the studies in this analysis. Some studies used postal follow-up and others used telephone follow-up. Due to these differences, the possibility of not combining these studies for meta-analysis was considered. It was decided, however, to include a meta-analysis of these studies as summary results for each strategy are presented allowing some interpretation of the usefulness of each intervention. The observed heterogeneity must, however, be considered in this interpretation. One of the studies had a very small sample size[168] but excluding this study and re-running the analysis had little effect on the results. Clinical researchers need to incorporate appropriate follow-up strategies within the budget constraints of their research activities. Due consideration for the patients' privacy is needed, however, to ensure that patients do not feel harassed by the follow-up efforts. Further research is required to determine the acceptability of repeated contact to the patient.

5.4.2 Questionnaire length

A recent review focuses on the effect of questionnaire length on response[7]. Out of twenty seven included trials, fourteen (52%) studied health related topics but only four (15%) studied patients rather than members of the general public. The authors extrapolate that shorter questionnaires should be used in clinical trials to improve response. Since none of the included studies looked specifically at clinical trials, such extrapolation should be viewed with caution. Our findings confirm that shorter questionnaires improve response in the health care setting. Again, however, there was significant observed heterogeneity between the trials with an I^2 value of 63.9%. The reason for this heterogeneity was not as easy to explain as that observed in the follow-up strategies comparison. The analysis was therefore re-run excluding individual trials to assess the effect this had on the I^2 value. It appeared that one trial[162] was contributing most to the observed heterogeneity. Excluding this trial reduced the I^2 value to 32.9% and had very little effect on the overall result of the meta-analysis. The result remained significantly in favour of shorter questionnaires. This trial was much smaller than the three others in this comparison group and was of unclear quality due to poor reporting. Individually excluding each of the other trials in this group had little effect on the I^2 value.

Questionnaires are often used in health care research to answer a research question. There is, however, an inevitable trade off between making the questionnaire comprehensive enough to answer the question adequately,

and not making it so long that it has an adverse effect on response. Careful consideration of the minimum data required when designing the questionnaire is essential. As yet there is insufficient evidence to suggest an optimal questionnaire length in terms of number of questions or pages. Also, optimal length is likely to vary between the populations studied.

5.4.3 Incentives

Previous reviews looking predominantly at market research found incentives to be a useful way of improving response.[12, 15, 36] The largest effect sizes are seen with monetary incentives. The use of incentives in health care research in Europe is uncommon. Trials often have strict budget constraints making the provision of incentives an unacceptable additional cost. Providing incentives in health care research can also raise ethical concerns.[75] A large monetary incentive may be seen as a coercive attempt to encourage cooperation.[170] The health care study participant may view their personal input into the study as the motivator to respond rather than merely responding to an incentive. This review has shown no evidence that incentives are effective in the health care context. This is an area, however, which requires further investigation. The studies included in this review used either small monetary incentives or monetary equivalent incentives (lottery ticket, prize draw or phone card). None of the studies investigated non-financial incentives such as pens. The inclusion of an incentive appropriate for the particular study may have a positive effect on response but this has not been tested. Until this area is investigated more

fully no recommendations can be made on including incentives in health care research as a method of improving response.

5.4.4 Question order

Question order appeared to have little effect on response rate. The three studies looking at question order, however, investigated two different approaches. One study compared a traditionally ordered questionnaire with a chronologically ordered one[160] and the other two studies compared placing condition specific questions either before or after generic questions.[166]

5.4.5 Future research

This review was strict in its definition of a 'patient' and excluded studies which were in the health care setting but involved the general public. It was anticipated that more studies would be found involving patients. The evidence available on which to base conclusions was therefore limited. The review could be repeated including health care research studies of the general public to give a broader perspective of methods of improving response in the health care setting. Previous studies have investigated this area evaluating methods of improving response such as postage stamps[171] and questionnaire length and incentives[171, 172]. The market research literature has investigated many methods of improving questionnaire response. Edwards et al[15] grouped these methods into the following strategies: Incentives, Questionnaire length, Appearance, Delivery,

Contact, Content, Origin and Communication. All these methods need to be tested on patients in the health care setting before extrapolations of their usefulness can be made.

All of the trials included in this review looked at the effect of an intervention in isolation of other interventions. Future studies could use factorial designs to investigate the addition of different methods to improve response.

In any future research it is important that the methods of improving response are well documented and tested in situations that reflect their intended use (i.e. patient populations in health care studies). The effect of the interventions on completeness of the returned questionnaires also requires investigation.

5.5 Conclusions

There is limited evidence of methods to improve response to postal questionnaires in patient populations in health care research. Caution should be taken in utilising the results of previous reviews in clinical study design. Follow-up strategies in the form of repeat mailing or telephone contact offer the most promising method of maximising response to postal questionnaires in health care research. The acceptability of repeated patient contact and ethics relating to this, however, need to be investigated further and guided by research ethics committees. Reducing the length of the questionnaire may also have a positive effect on response.

5.6 Chapter summary

This systematic review has highlighted the fact that there is a paucity of literature surrounding the issues of maximising response to postal questionnaire follow-up in a clinical trial. Although repeated patient contact was shown to improve response rates, such methods threaten to violate basic informed consent principles. It is therefore necessary to identify other features of response which are open to influence without compromising ethical considerations. This finding has resulted in the development of both the quantitative study described in chapter 6 and the qualitative study described in chapter 7.

6 Chapter 6

A randomised controlled trial of a method of improving response to postal questionnaire follow-up in a clinical trial

This chapter presents a randomised controlled trial of a method designed with the intention of improving response rates of participants to clinical trial follow-up. The systematic review presented in chapter 5 highlights the lack of trials which have been conducted to investigate methods of improving response specifically in clinical trials. This chapter will, therefore, significantly add to the body of available evidence in this area.

A novel method of improving response was conceptualised and subsequently developed. The method is a 'Trial Calendar' which is a tool to prompt participants to complete and return their questionnaires. This chapter aims to establish whether the Trial Calendar is effective in improving response to postal questionnaire follow-up in a clinical trial. The effect of the Trial Calendar on missing data will also be considered. The most appropriate methodology to achieve this is to test the Trial Calendar in a controlled experiment. Statistical analysis can then be undertaken to see whether the independent variable (the Trial Calendar) has an effect upon the dependent variable (response rate).

6.1 Background

The importance of maintaining adequate response rates to postal questionnaires in clinical trials has already been stressed. Minimising missing data is also a high priority in clinical trial design. Failure to obtain adequate information from trial participants seriously affects the quality of a trial and reduces the credibility of the results.

This study aims to introduce a method of improving response rates and minimising missing data and test this method within the context of a clinical trial. The proposed method is a customised 'Trial Calendar'. It is hypothesised that this tool will act as a memory aid and reminder for patients to complete and return their questionnaire. The Trial Calendar was designed taking into account relevant issues described in the literature and detailed below. (See appendix 4 for an example of the Trial Calendar).

By 'nesting' the Trial Calendar within an existing clinical trial the main research questions are;

- Does the Trial Calendar improve response rates to postal questionnaires in a clinical trial?
- Does the Trial Calendar reduce the need for prompting participants to return their questionnaires?
- Does the Trial Calendar have a positive influence on minimising missing responses of key recovery events?

Several methods are described in the literature for improving response rates to postal questionnaires. Many of these methods are embedded in the actual questionnaire design. For example, issues such as question wording and sequencing, questionnaire appearance, personalisation and questionnaire length have all been shown to have an effect on response rate.[15] Additional patient contact has also been extensively investigated as a method of improving response rate.[117, 167, 168, 173] The systematic review presented in chapter 5 concluded that follow-up strategies offer the most promising method of improving response in health care research.

CAST offered an ideal setting to add to the body of literature on what methods are effective in improving response in clinical trials. As has been suggested by previous authors, by nesting the method of improving response into an existing clinical trial, a more realistic impression of its effectiveness will be established.[17] The idea for the Trial Calendar was developed in anticipation of difficulties in maintaining adequate response rates in CAST. The idea was developed into a small, user-friendly desk or wall calendar to be given to the participants upon initial recruitment onto the trial. The calendar was designed with the intention of prompting the participant about when to expect a follow-up questionnaire and aiding recall of certain recovery milestones. CAST has three follow-up time points: four weeks, 12 weeks and nine months following injury. As the trial was investigating an acute injury it was expected that many of patients would be

recovered by 12 weeks. Virtually all the patients would be expected to have recovered by nine months. Previous trials of acute ankle sprains either follow the patients up only in the short term (4 to 6 weeks)[174-176] or experienced significant losses to long term follow-up.[121-126] Incorporating the Trial Calendar into CAST was given ethical approval by the Northern and Yorkshire Multi-Centre Research Ethics Committee. Approval was subsequently gained from the relevant Local Research Ethics Committees.

6.2 Methods

6.2.1 Participants

The participants were patients randomised into CAST between November 2003 and July 2005. The Trial Calendar was randomly allocated to consecutive CAST participants at the trial centres local to the Coventry area. Trial centres included were therefore; Coventry and Warwickshire Hospital, Birmingham Heartlands Hospital, Solihull Hospital, Hospital of St Cross Rugby and Alexandra Hospital Redditch. The remote trial centres of Bristol and Oxford were excluded from the calendar trial. The calendar trial required frequent visits to the trial centres involved to instigate the randomisation, replenish the stocks of calendars and generally monitor trial progress. This meant that including the remote trial centres would have involved considerable extra time and cost which was beyond the scope of this project. Warwick Hospital was also not included in the calendar trial.

This centre was a late addition to CAST and the calendar trial was already established and running smoothly. It was, therefore, not necessary to include Warwick in the calendar trial.

6.2.2 Development and design of the Trial Calendar

6.2.2.1 Methods of improving response

The Trial Calendar was designed by drawing on the existing literature (mainly in survey research) and learning from the results of the systematic review. It is in the area of follow-up and patient prompting that the Trial Calendar has relevance. CAST has a standardised system of prompts (both telephone prompts and re-issuing the questionnaire). This system was developed during the pilot phase of CAST with reference to the available literature (see section 4.3.1.2, chapter 4).

In a recent structured review, McColl et al[17] examined (amongst other things) number and timing of patient contacts, pre-notification contacts, and follow-up contacts/reminders. Methods of contact used in the studies reviewed included telephone calls, letters, postcards and replacement questionnaires. None of the studies reviewed used a calendar method of improving response rate and a separate search of the literature has found no studies that have used this method. Also, few of the studies looked specifically at response rates in clinical trials. The main conclusions drawn from the review were that:

- Pre-notification by letter (rather than by phone) is effective in improving response rates.
- Follow-up contact is highly effective in increasing response rates.
- Response rates can be increased through multiple contacts.

As a result of the review the authors recommend instigating these methods into clinical trial design. There are implications on the costs in terms of both time and resources for applying these methods as was found in the CAST pilot. Across all three follow-up points, the 24 CAST pilot patients required an average of seven phone calls in attempts at prompting (range = 1 to18) and two repeat mailings of the questionnaires (range = 0 to 5) (see Table 3 page 57 Chapter 4).

Dillman [37] describes questionnaire response in terms of 'social exchange'. This suggests that response is maximised by optimising the costs and rewards of responding and establishing that this 'contract' would be fully honoured. Dillman incorporated this theory into a framework for inducing response, the 'Total Design Method' (TDM), and this is well documented in the survey literature.

The design of the Trial Calendar aimed to minimise the time and resources required whilst maintaining the important aspects that have been identified as having a positive influence on return rates.

Implications of methods of improving response on Trial Calendar design:

The Trial Calendar was given to the patient on recruitment onto CAST. The calendar was customised on a monthly basis to include a pre-notification caption on the months that the patient was due to receive a questionnaire followed by a reminder caption the following month. This reduced the necessity for multiple contacts which may be considered by ethics committees to be overly intrusive. This design aspect of the calendar accounted for the importance of pre and post questionnaire contact but in a cost effective and researcher time efficient way.

Aspects of Dillman's 'Total Design Method'[37] were incorporated into the design of the Trial Calendar. This included emphasising the rewards to the participant (the social usefulness of the study and the contribution their response would make) and establishing trust by emphasising the credibility of the study (using logos and giving the name and telephone number of the trial manager).

6.2.2.2 Sources and characteristics of non-responders

To understand the reasons why individuals may not respond to questionnaires, investigations have been carried out to evaluate sources of non-response.[24] Other researchers have investigated whether non-responders share certain characteristics.[14] Moser and Kalton [24] have identified the five main sources of non-responders that tend to occur in a survey as:

1. Patients 'unsuitable for inclusion'; for example deaf, blind or illiterate.

2. 'Movers'; It is likely that a number of those included in a study will no longer live at their listed address and will be untraceable to their new address.
3. 'Refusals'; Inevitably, some of the sample will simply decline to participate further in the study and will never return their questionnaire.
4. 'Away from home'; Some individuals may be away from home (eg working away) for periods during the sample time limit.
5. 'Not at home'; This has implications for interview surveys rather than postal ones and relates to individuals who are not at home at the scheduled interview time.

Kanuk and Berenson[14] carried out an extensive literature review of factors influencing response rates to postal surveys although not specifically in clinical trials. They examined the differences between responders and non-responders in terms of a wide range of demographic, socio-economic and personality factors. The only consistent finding was that responders tend to be better educated and have a greater writing ability. Analysis of the CAST pilot data also showed non-responders were most often in the lower educational achievement groups. Cartwright[177] found that non-responders are likely to be in semi-skilled or unskilled manual occupations. Dillman[37] has categorised the methods of maximising response rates into three main areas based on expert opinion:

1. Minimising the costs of responding. (Physical, mental, emotional and economic).
2. Maximise the rewards of responding.
3. Establish trust that the rewards will be delivered.

Implications of sources and characteristics of non-responders on Trial Calendar design:

Taking into account these identified sources and characteristics of non-responders, the Trial Calendar was designed to try and target and address these issues. Starting with Moser and Kalton's[24] five main sources of non-responders; the 'Unsuitable for inclusion' group was not applicable to the trial within which the calendar was used as patients were already recruited as part of the trial design. 'Movers' were targeted in the calendar with a reminder to patients to contact the trial office if they moved house or intended to move house. 'Refusals' were targeted by trying to make the calendar humorous, friendly and approachable to encourage patients to respond. The importance of participant's contribution to the study was also frequently highlighted in the calendar. The 'Away from home' group was difficult to target and the 'Not at home' group was not applicable to a postal questionnaire.

Kanuk and Berenson's[14] finding of non-responders being less well educated and less literate was addressed by designing the calendar with heavy use of pictures and simple text. Michielutte et al[178] investigated the use of illustrations as a means of improving the readability of a health

education brochure and found that patients rated the brochure containing illustrations significantly higher than one with no illustrations.

The Flesch Reading Ease score and The Fog Test were used to assess the readability of the Trial Calendar. The Flesch Reading Ease score rates text on a 100-point scale, the higher the score the easier it is to understand the document. Standard documents should aim for a score between 60 and 70. The Trial Calendar scored 74.1 on this scale. The Fog Test produces a score related to sentence length and the percentage of 'long' words (words of more than three syllables) used in a piece of writing. Scores are grouped into categories ranging from 'very easy reading' to 'almost unreadable'. The score for the calendar was 21 which fell into the category of 'easy reading'. These two scores confirm that the Trial Calendar should be understood by less literate members of the sample.

Finally, the design of the Trial Calendar went some way to targeting the three suggestions of Dillman and colleagues[37] in that:

1. The mental costs to the patient would be reduced by the calendar providing a way of recording significant events (for example; return to work/sport) and also by clarifying when the patient should expect a questionnaire.
2. The calendar would act as a means of maintaining the patient's interest and awareness of the trial over the nine month trial period by confirming the importance of their contribution.

3. The calendar frequently reminded patients how their contribution will potentially benefit others with similar injuries.

The trial calendar idea was developed and nested into CAST early on in the recruitment phase of the trial. It was, therefore, only possible to use data from the 24 pilot patients regarding socio-demographic characteristics of non-responders in the design of the calendar. Much of the data used to guide the design came from the survey literature. There is no available literature specifically detailing the socio-demographic characteristics of non-responders in clinical trials. CAST, however, provided an appropriate situation to study this and the results are presented in chapter 8.

6.2.2.3 *Minimising missing data*

As well as data missing as a result of non-response, some patients may return an incomplete questionnaire. There are two types of missing values found in this situation. Firstly, a question may be left deliberately blank because it did not apply to the individual respondent and secondly a reply may be expected but not given. In clinical trials that investigate the long-term outcomes of a particular treatment intervention, questions are often asked about significant events in the patient's recovery. For example, in CAST patients are asked about when they returned to work and sports activities following their injury. It is easy for a patient to simply skip the question if the event was some time ago and requires some thought to recall.

Patient diaries are commonly used in health care and clinical research to assess patient experiences and help with recall.[179] This method involves the patient keeping a daily record of activities or events (e.g. symptoms) and can be useful when detailed information needs to be collected. Bowling[180] however, suggests that this method is only practical with small, committed samples of people and also that it is unreasonable to expect people to complete diaries for a long length of time. Several authors have also highlighted the problems of non-compliance with the paper diary method.[181, 182] Patient diaries are usually used in clinical trials when detailed qualitative information is required about the patient's health and their perceptions of health. Diary methods do, however, vary and range from symptom diaries in which in-depth details of the patient's symptoms are recorded, to diaries which just collect data about activities and events relating to the patients' health experiences.[180] The Trial Calendar could therefore be considered to be a *type* of diary method as the patients were asked to use it to record significant events in their recovery. An important difference is that the Trial Calendar contains prompts to encourage the patient to record the specific information required for assessing the primary outcomes of CAST. In the context of CAST a detailed patient symptom diary could be considered too in-depth and would generate too much unnecessary information. Also, long-term compliance is likely to be low since the injury under investigation is acute and will usually recover rapidly.

In a search of the literature of the use of calendars to aid patient recall, reference could only be found to 'Event History Calendars' (EHCs). Event History Calendars have been used for retrospective data collection of occurrence, timing and sequencing of a variety of life events.[183] The design of an EHC is determined by the researcher and the aims of the research. Data for the EHC is, however, usually recorded in an interview session making this method of gaining information beyond the methodological scope of CAST. The trial was not designed to collect and evaluate detailed qualitative information and the primary outcomes were the time it took for the patient to return to pre-injury mobility and work. This information was gathered as single questionnaire items. A secondary outcome of CAST was cost implications of the injury to both the patient and the NHS. Some form of tool/method for helping patients recall such information would consequently be very valuable. The Trial Calendar therefore contained prompts in the form of questions attached to the months when these events were likely to occur and mirrored questions that were included in subsequent questionnaires. These prompts were:

- Have you gone back to work yet?
- Are you doing your usual sports/activities?
- Have you spent any money on medicines etc because of your ankle?
- Have you needed to return to the hospital or see your doctor because of your ankle injury?

Patients were encouraged to circle the date on the calendar when these events happened and then use the calendar to help them complete the subsequent questionnaires. The prompts were placed in the calendar at the time points they were considered to be most relevant. For example, the 'return to work' prompt was placed in the calendar one month and two months after recruitment. It was felt unlikely that patients with an ankle sprain would be off work for longer than two months.

The general format of the Trial Calendar was the same as a normal decorative calendar. The Trial Calendar was an easier and less time consuming way to record information about return to pre-injury function than a symptom diary. It was thought, therefore, that compliance would be higher. The aim was that the Trial Calendar would help patients recall information and complete the follow-up questionnaires more completely thereby minimising missing responses.

6.2.2.4 Summary and justification for Trial Calendar

The available literature surrounding methods of improving response to postal questionnaires has demonstrated that patient contact both before and after sending the questionnaire can influence response rates.[17] The Trial Calendar is an idea which incorporates this into a single tool. If the Trial Calendar is effective, this will reduce the time and resources required by the research team to chase reluctant responders. The cost of producing a Trial Calendar is less than the cost of several phone calls, repeat posting and researcher time. By incorporating prompts for answers to primary outcome

questions the Trial Calendar, if effective, will also reduce missing data and therefore have a dual role. This is an original idea and no reference can be found in the literature to a similar tool.

6.2.3 Objectives

The objectives of this randomised controlled trial are:

- To assess the effect the Trial Calendar has upon amount of prompting required and response rates to postal questionnaire follow-up in CAST.
- To assess the effect the Trial Calendar has on the amount of missing data in primary and secondary outcome questions in CAST questionnaires.

6.2.4 Outcome measures

The effect of the Trial Calendar in the areas of response rate and missing data was measured in the following ways:

1. *Percentage final response rate to questionnaires of patients supplied with calendar compared to patients not supplied with calendar measured at each time point*

Final response rate is that which occurs after the standardised sequence of prompts used in CAST (see Figure 1, p54). If patients failed to return their questionnaire after the standardised sequence of prompts, attempts were

made to gather core information over the telephone. Patients who required core information gathering by telephone were *not* classed as responders for this analysis. This type of data gathering is seen as a 'last resort' and is desirable to avoid if possible.

2. *Amount of prompting required by participants to return their questionnaires at each time point.*

CAST categorised each participant according to how much prompting was necessary for them to return their questionnaire (see Table 4 p59). These data were used to assess the effect of the Trial Calendar on the amount of prompting required.

3. *Percentage of missing data of the pre-defined core outcome questions for which the calendar 'reminder captions' are used (see page 118).*

The analysis was repeated at each follow-up time point in CAST which were four weeks, 12 weeks and nine months following injury.

6.2.5 Sample size

The sample size calculation was governed by the limitations of the size of the CAST sample but also took into account what was considered to be a desirable increase in questionnaire response rate. This was based on

figures obtained from the pilot phase of CAST. The sample size calculations used in the trials included in the systematic review of methods of improving response in health care research (reported in chapter 5) were also used as reference. Of the 13 trials included in the systematic review, six failed to report how they had calculated their sample size.[117, 163, 166-169] The remaining trials based their sample size on expected increases in response rates of between five and 20% following the introduction of the intervention. Expected changes in actual response rates varied from 45% to 65%[162] to 75% to 80%[22].

It has been suggested that the minimum acceptable response rate in a clinical trial is 80% and efforts should be made to get this figure as close to 100% as possible.[9] Responses to the four and 12 week and 9 month follow-up points of the pilot phase of CAST were 79%, 67% and 63% respectively.

Previous research and the CAST pilot data were therefore taken into consideration and the following decisions were made:

- The percentage response rate expected without the Trial Calendar was set at 65%. This reflects response at the later follow-up points in the CAST pilot and is in keeping with other similar studies. This figure is termed **P1**

- The percentage response rate expected with the Trial Calendar was set at 80%. This reflects an acceptable level of response to follow-up in a clinical trial. This figure is termed **P2**
- To detect a difference between the two conditions a 5% level of significance will be used. This is termed the **alpha** and is the probability of detecting an effect when one does not exist (type I error).
- The degree of certainty that the difference between P1 and P2, if present, would be detected is termed the power of the study. This is denoted as (1 – **beta**). This was set at 80% therefore beta is equal to 0.2

These figures were then inserted into the following standard sample size calculation equation:

$$n = \frac{P1 \times (100 - P1) + P2 \times (100 - p2) \times f(alpha, beta)}{(P2 - P1)^2}$$

Where 'f(alpha, beta)' is a function of the ordinates of a normal distribution and is given as 7.9 for the alpha and beta used here.[184]

Inserting the values chosen for this study into the equation gives:

$$n = \frac{65 \times (100 - 65) + 80 \times (100 - 80) \times 7.9}{(80 - 65)^2}$$

The result of this calculation is n = 136. This means that 136 participants are needed in each of the two arms of the trial giving a total sample size of 272 participants.

6.2.6 Randomisation

A computer generated random sequence was used to allocate CAST participants to either the 'Calendar' or 'No Calendar' group. Once the participant was randomised onto CAST they were allocated a numbered 'Baseline Pack' which contained all the necessary forms for CAST. This included the consent form and baseline questionnaires. The baseline packs were compiled and numbered in advance and stored at the trial centres. When a patient was randomised onto CAST (using computer generated randomisation obtained by telephone) the next consecutively numbered pack from the pile was taken. The number on the pack became the patient's unique 'Centre ID Number' for the remainder of the trial period. The randomly generated sequence of 'calendar' or 'no calendar' was mapped onto the consecutive sequence of centre ID numbers. It was then possible to insert a sheet into the baseline packs indicating whether the participant was allocated to receive a calendar or not. A colour coded sticker was also attached to the CAST baseline questionnaire to indicate the allocation which could then be blindly entered onto a database (see next section). Allocation was, therefore, concealed until the participant was recruited onto CAST and the baseline pack was opened.

6.2.7 Implementation of Trial Calendar Procedures

Recruitment of CAST participants at the local trial centres was carried out by members of the CAST research team who were all physiotherapists. All members of the team were trained in the calendar trial procedures. The Trial Calendars were customised each month to reflect when follow-up questionnaires were due. Calendars were therefore stored at the trial centres in envelopes according to the month in which the patient was recruited. If the recruiting physiotherapist found a sheet in the baseline pack indicating that the participant was allocated to receive a calendar, a calendar was taken from the current month's envelope. The calendar was then given to the patient at the end of the CAST recruitment process. A standardised explanation as to the purpose and use of the Trial Calendar was printed on the front page of the calendar. This explanation was also given verbally by the recruiting physiotherapist (see appendix 5).

As part of CAST, the completed baseline packs were then returned to the university and the data was inputted into a Microsoft Access database. A field was added to the database on commencement of the calendar trial entitled 'Calendar Colour'. This corresponded to the coloured sticker added to the baseline questionnaire which indicated whether the participant had received a calendar or not. An orange sticker indicated that the participant had been allocated to the 'calendar' group and a purple sticker indicated allocation to the 'no calendar' group. This colour code was not revealed to the person inputting the data. This ensured the 'blind' inputting of data which

reduced a possible source of bias. All participants in the calendar trial received the same CAST systematic follow-up procedure. A detailed record of the amount of prompting required at each follow-up point was kept.

6.2.8 Data cleaning

A screening procedure was carried out to check for errors in data entry prior to analysis of the study data. This involved checking categorical data to make sure that no values fell outside the stated parameters. Continuous data was checked to ensure that it fell within the anticipated range. Any data which appeared unusual was investigated by checking against the original study documents and questionnaires.

6.2.9 Statistical methods

The response category for the CAST follow-up questionnaires detailed in chapter 4 section 4.3.1.4 was logged for each participant at each time point. The data were entered into a Microsoft Access database by independent personnel blind to calendar allocation. The subsequent data analysis was carried out using SPSS for Windows 14.0.[185] Calendar allocation was re-coded by the inputting personnel ensuring that the data analysis was also blind to calendar allocation.

Univariate analysis

Overall response rates for participants who had received a Trial Calendar were compared to those who had not received a Trial Calendar. The response categories were analysed to see if patients receiving a Trial

Calendar required less prompting than patients not receiving a calendar.

The chi-squared test was used to test for significance.

Multivariate analysis

Certain demographic characteristics have been shown to influence postal questionnaire response rates in survey research.[71] These characteristics are discussed in detail in chapter 8. For the purposes of this study a logistic regression analysis was performed on the response data from each follow-up time point. The aim was to develop a model to assess and explain the effect of the Trial Calendar taking into account certain variables which were felt to be important in the context of CAST. These variables were: age, sex, employment, trial centre, CAST treatment received, education level and recurrent sprain. Logistic regression is appropriate as it allows for the testing of models in which the dependent variable is categorical rather than continuous. In this study the dependent variable is categorical: questionnaire response or non-response.

Missing data

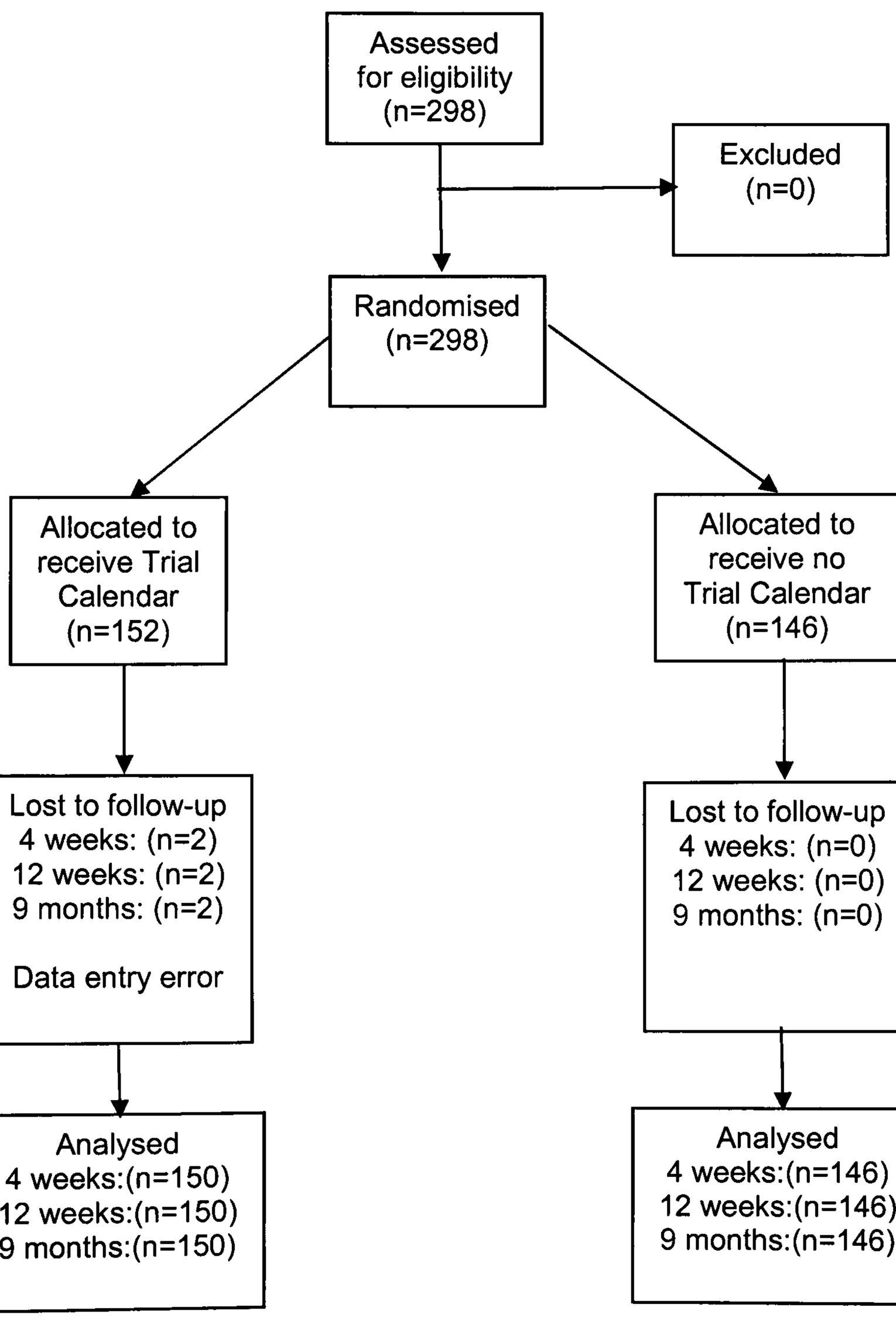
The amount of missing data in key recovery event questions was compared in participants receiving a Trial Calendar and those not. Crosstabulations were used to analyse the effect of the Trial Calendar and the chi-squared test was used to test for significance.

6.3 Results

6.3.1 Trial Flow

In accordance with the CONSORT guidelines[186] Figure 8 is a flow diagram showing the progress of the trial through the phases of recruitment and analysis.

Figure 8 ***Flow diagram of trial progress***



6.3.2 Descriptive statistics

Descriptive data of the baseline characteristics and study baseline variables are presented prior to inferential statistical analysis. A total of 298 CAST participants aged 16 to 57 were recruited onto the calendar trial.

Table 13 shows the frequency distribution of the baseline characteristics of participants randomised to receive a Trial Calendar or no Trial Calendar.

Table 13 ***Baseline characteristics of participants randomised to receive a Trial Calendar or no Trial Calendar. Values are numbers (%) (except for mean age).***

	Calendar (n=152)	**No Calendar** (n=146)
Mean (SD) age (years)	28.9 (10.14)	30.1 (10.52)
Male/Female	91/61	85/61
CAST treatment received		
Tubigrip	39 (26)	38 (26)
Aircast splint	33 (22)	44 (30)
Bledsoe Boot	44 (29)	35 (25)
Plaster cast	36 (24)	29 (20)
Trial centre		
Coventry	68 (45)	69 (47)
Heartlands	24 (16)	23 (16)
Redditch	27 (18)	18 (12)
Solihull	18 (12)	19 (13)
Rugby	15 (10)	17 (12)
Employment type		
Unskilled	25 (17)	18 (12)
Skilled	53 (35)	56 (39)
Professional	36 (24)	33 (23)
Other	7 (4)	6 (3)
Not at work	31 (21)	33 (23)
Education level		
Nil	16 (11)	16 (11)

CSE/O Level/GCSE	67 (44)	51 (35)
A Level/Degree/Other	69 (45)	79 (54)
Recurrent sprain		
Yes	17 (11)	16 (11)
No	135 (89)	130 (89)

It is now generally agreed that the use of hypothesis testing to compare baseline characteristics of experimental groups is not appropriate.[187, 188] This is, however, still common practice in the literature. Altman [187] argues that if the randomisation is adequately and fairly performed, any differences between the treatment groups must be due to chance.

Table 13 shows that the baseline characteristics of participants receiving a Trial Calendar and those not are generally evenly distributed.

Although baseline information was available for all 298 participants a data entry error meant that no further information was available regarding subsequent response for two participants. The data from a total of 296 participants were therefore subsequently analysed.

6.3.3 Inferential statistics

6.3.3.1 Outcome One

- Effect of the Trial Calendar on percentage final response rate to questionnaires of patients supplied with calendar compared to patients not supplied with calendar measured at each time point.

This was the main outcome measure. Table 14 shows the effect of the Trial Calendar on final response rates at each of the follow-up time points in CAST. The chi-squared test (χ^2) was used to test for significance.

Table 14 ***Final response rates at each follow-up point for participants supplied with a Trial Calendar compared to those not supplied with a Trial Calendar. Values are numbers responding and non-responding (percentage).***

	Calendar	No Calendar	χ^2
Four week follow-up point	**n=150**	**n=146**	
Responder	117 (78)	114 (78)	
Non-responder	33 (22)	32 (22)	0.986
12 week follow-up point	**n=150**	**n=146**	
Responder	96 (64)	96 (66)	
Non-responder	54 (36)	50 (34)	0.752
Nine month follow-up point	**n=150**	**n=146**	
Responder	89 (59)	74 (51)	
Non-responder	61 (41)	72 (49)	0.135

The results of the analysis show that the Trial Calendar had no significant effect on response rates at any follow-up time point. The difference in response at nine months appeared to be bigger but still not significantly different from chance.

Logistic Regression

Although the simple univariate analysis described above showed that the Trial Calendar had no effect on response, it does not take into account the effect on response of the various covariates such as age and sex as described in Table 13. Simple analysis also does not analyse the interaction of the covariates and possession of a Trial Calendar. To describe the relationship (if any) between the dependent variable (response or non-response) and the independent variables (the covariates) a logistic regression model was used. Logistic regression is used rather than linear regression since the outcome variable is dichotomous.[189] Because there is more than one independent variable in which we are interested, the logistic regression will be 'multiple'. The independent variables of age, sex, employment and education have been shown to have an effect on response to questionnaires in the survey literature. The variables of trial centre, CAST treatment received and recurrent sprain are considered to be possibly relevant to response rates and are therefore included in the model. A recurrent sprain meant that a person had sprained their ankle three times or more in the past with the most recent sprain being within the last year.

The first step in the construction of a logistic regression model is to select the variables to include in the model. It has been suggested that the selection process should begin with a univariate analysis of each variable.[189] Any variable whose univariate test had a p value of < 0.25

was then included in the multivariate analysis and the 'Enter' method in SPSS was used to enter the variable into the logistic regression . This screening criterion for selection of variables is based on the work of Mickey and Greenland[190] who showed that traditional levels of significance (e.g. 0.05) often failed to identify variables known to be important. The interaction of the various independent variables with the effect of whether the participant received a Trial Calendar is fundamental to this analysis. For this reason the interaction effects of the Trial Calendar on variables included in the multivariate analysis was assessed.

Each variable was coded for analysis using SPSS 14.0. The code sheet for each variable is shown in Table 15.

Table 15 ***Code sheet for variables in Trial Calendar data set***

Variable (coding)	**Abbreviation**
Response (0=No, 1=Yes)	RESP
Received Trial Calendar (0=No, 1=Yes)	CAL
Age (1= 16-24, 2=25-34, 3= 35-44, 4=45-57)	AGE
Sex (0=female, 1=male)	SEX
CAST treatment received (1=Tubigrip, 2=Plaster, 3=Aircast, 4=Bledsoe)	TMT
Trial centre (1=Redditch, 2=Heartlands, 3=Coventry, 4=Solihull, 5= Rugby)	TC
Employment (0=No, 1=Yes)	EMP
Education level (0=nil, 1=CSE/O Level/GCSE 2= A level/degree/other	EDU
Recurrent Sprain (0=No, 1=Yes)	REC

SPSS 14.0 uses the 'maximum likelihood' method of logistic regression estimation. This method yields values for the unknown parameters which maximise the probability of obtaining the observed set of data.[189] For mathematical ease the log of the maximum likelihood equation is calculated resulting in the **log likelihood**. The **likelihood ratio** is a function of the log likelihood (-2x log likelihood) and is calculated so that a quantity is obtained which can be used for hypothesis testing purposes. The difference between the likelihood ratio obtained for the model without the variable, compared to that obtained with the variable, is termed 'G' in SPSS. Under the null hypothesis (i.e. the covariate has no effect on response) this quantity will follow the chi-squared distribution with the same number of degrees of freedom as the covariate. If the G value is greater than the chi-squared value the result is statistically significant. In this analysis the 0.05 level of significance was used.

To test how well the resulting regression models fitted the data, the 'Hosmer and Lemeshow's goodness if fit test' was used in SPSS. A non-significant result of this chi-squared test indicated that the model had adequate fit.

Results of four week follow-up logistic regression

Table 16 ***Four week follow-up: Summary of univariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)***

Variable	-2 Log Likelihood (Likelihood Ratio)	G	Degrees of freedom	χ^2 value*	Sig
Constant	311.625				
CAL	311.625	0.000	1	3.841	0.986
AGE	305.001	6.624	3	7.815	0.085
SEX	311.363	0.262	1	3.841	0.609
TMT	308.678	2.947	3	7.815	0.400
TC	308.525	3.100	4	9.488	0.541
EMP	309.177	2.448	1	3.841	0.118
EDU	311.217	0.408	2	5.991	0.815
REC	311.515	0.110	1	3.841	0.739

* 0.05 significance level

Two variables; AGE and EMP (see Table 15 for definitions of codes) reach the p <0.25 criterion for inclusion in the multivariate analysis. Table 17 shows the results of the multivariate logistic regression including these two variables.

Table 17 ***Four week follow-up: Multivariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)***

Variable	Regression Co-eff	Standard Error	Odds Ratio	95% CI
AGE(1)	-1.201	0.642	0.301	0.085,1.059
AGE(2)	-0.659	0.676	0.517	0.137,1.946
AGE(3)	-1.151	0.678	0.316	0.084, 1.196
EMP	-0.446	0.324	0.640	0.339,1.207

Younger people and those who are employed are somewhat less likely to respond. The likelihood ratio (-2 log likelihood) for the multivariate model is 303.155. Comparing the difference of this value to the likelihood ratio of the constant only model of 311.625 gives a G value of 8.471 with a significance value of 0.076. This indicates that the model is not significantly different from the constant only model.

Fundamental to this RCT was the effect of the Trial Calendar on response. Although univariate analysis did not indicate that the Trial Calendar had an effect on response it is possible that there may have been an interaction between the Trial Calendar and the other variables. The effects of adding interactions between (CAL and AGE) and (CAL and EMP) to the previous main effects model including CAL are shown in Table 18.

Table 18 ***Four week follow-up: Effect of addition of interaction effects to previous multivariate model***

Interaction	-2 Log Likelihood (Likelihood Ratio)	G	Degrees of freedom	χ^2 value*	Sig
Constant	311.625				
Main effects model	303.155	8.471	4	9.488	0.076
Main effects model + CAL	303.154	8.472	5	11.070	0.132
CAL X AGE	298.623	4.530	3	7.815	0.210
CAL X EMP	303.142	0.012	1	3.841	0.912

*0.05 significance level

Logistic regression of the four week data has not resulted in the development of a predictive model of response. This model adequately fits the data as indicated by a non-significant result of the Hosmer and Lemeshow goodness if fit test. Age is the only variable which approaches significance and appears to have an association with possession of a Trial Calendar in its effect on response. The data were therefore explored and tabulated and presented in Table 19.

Table 19 ***Crosstabulation of four week response rates showing age group and Trial Calendar possession. Figures are observed and expected numbers responding and non-responding***

Age		No Calendar		Calendar	
		Observed	***Expected***	***Observed***	***Expected***
16-24	Responder	39	42	51	48
n=123	Non-responder	18	15	15	18
	Total	**57**	**57**	**66**	**66**
25-34	Responder	37	34	32	35
n=83	Non-responder	4	7	10	7
	Total	**41**	**41**	**42**	**42**
35-44	Responder	23	23	21	21
n=59	Non-responder	8	8	7	7
	Total	**31**	**31**	**28**	**28**
45-57	Responder	15	15	13	13
n=31	Non-responder	2	2	1	1
	Total	**17**	**17**	**14**	**14**

This table shows that in every age group except the 25-34 year olds response rates were slightly (but not significantly) higher in the group possessing a Trial Calendar. Amongst the 25-34 year olds the possession of the Trial Calendar appears to have a negative effect on response. The numbers in each age group, however, are small so the power to detect differences is low. The trial was not powered for sub-group analysis.

Results of twelve week follow-up logistic regression

Table 20 ***Twelve week follow-up – Summary of the univariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)***

Variable	-2 Log Likelihood (Likelihood ratio)	G	Degrees of freedom	χ^2 value*	Sig
Constant	383.781				
CAL	383.681	0.100	1	3.841	0.752
AGE	376.834	6.947	3	7.815	0.074
SEX	376.425	7.356	1	3.841	0.007
TMT	383.573	0.208	3	7.815	0.976
TC	378.844	4.937	4	9.488	0.294
EMP	382.308	1.473	1	3.841	0.225
EDU	382.567	1.214	2	5.991	0.545
REC	383.490	0.291	1	3.841	0.589

* 0.05 significance level

The result of the univariate logistic regression of the twelve week follow-up data shows that the variables of SEX, AGE, and EMP fall within the $p<0.25$ criterion for inclusion in the multivariate analysis. Results of the multivariate logistic regression are shown in Table 21.

Table 21 ***Twelve week follow-up: Multivariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)***

Variable	Regression Co-eff	Standard Error	Odds Ratio	95% CI
AGE(1)	-0.887	0.499	0.412	0.155,1.096
AGE(2)	-0.699	0.518	0.497	0.180,1.371
AGE(3)	-0.523	0.542	0.593	0.205,1.716
SEX	0.650	0.273	1.915	1.122, 3.267
EMP	-0.448	0.303	0.353	0.639,1.156

Women are rather more likely to respond. The likelihood ratio (-2 log likelihood) for the multivariate model is 369.783. Comparing the difference of this value to the likelihood ratio of the constant only model of 383.781 gives a G value of 13.998 with a significance value of 0.016. This indicates that the model is significantly different from the constant only model and is useful for predicting response. The model was refined by taking the most significant predictive variable (SEX) and adding the other variables (AGE and EMP) in turn. The results are shown in Table 22.

Table 22 ***Twelve week follow-up: Refined multivariate model***

Variables	-2 Log Likelihood (Likelihood Ratio)	G	Degrees of freedom	χ^2 value*	Sig
Constant	383.781				
SEX	376.425	7.356	1	3.841	0.007
SEX + AGE	371.957	4.468	3	7.815	0.215
SEX + EMP	373.615	2.810	1	3.841	0.094

* 0.05 significance level

SEX is the only variable which has a significant effect on response at 12 weeks. Adding AGE and EMP does not contribute significantly to the predictive ability of the model. The effect of employment might be significant in a larger sample and is significant at the 10% level.

To put the model into the context of this trial, the effect of the interaction between the main effects (SEX, AGE and EMP) and Trial Calendar possession (CAL) was added to the model. The CAL variable was first added to the main effects model before assessing the interaction effect. The results are shown in Table 23.

Table 23 ***Twelve week follow-up: Effect of addition of interaction effects to previous main effects model***

Interaction	-2 Log Likelihood (Likelihood Ratio)	G	Degrees of freedom	χ^2 value*	Sig
Constant	383.781				
Main effects	369.783	13.998	5	11.070	0.016
Main effects + CAL	369.745	14.037	6	12.592	0.029
SEX X CAL	368.379	1.366	1	3.841	0.243
AGE X CAL	369.225	0.520	3	7.815	0.914
EMP X CAL	367.283	2.462	1	3.841	0.117

* 0.05 significance level

Adding interaction effects to the main effects plus CAL model did not significantly change the predictive power of the model. This indicates that there was no significant interaction effect between CAL and any of the main effects variables at the 12 week follow-up point.

The regression model therefore shows that SEX is the only variable to significantly affect response at 12 weeks. The odds ratio of 1.915 (see Table 21) indicates that females have almost double the odds of responding as men to respond to postal questionnaire follow-up at 12 weeks. This model adequately fits the data as indicated by a non-significant result of the Hosmer and Lemeshow goodness if fit test.

Results of nine month follow-up logistic regression

Table 24 ***Nine month follow-up: Univariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)***

Variable	-2 Log Likelihood (Likelihood ratio)	G	Degrees of freedom	χ^2 value*	Sig
Constant	407.297				
CAL	405.058	2.239	1	3.841	0.135
AGE	391.064	16.233	3	7.815	0.001
SEX	405.384	1.913	1	3.841	0.167
TMT	406.622	0.675	3	7.815	0.879
TC	402.757	4.540	4	9.488	0.338
EMP	404.635	2.662	1	3.841	0.103
EDU	406.454	0.843	2	5.991	0.656
REC	407.109	0.188	1	3.841	0.664

* 0.05 significance level

The results of the univariate logistic regression of the nine month follow-up data shows that the variables of CAL, AGE, SEX, and EMP fall within the $p<0.25$ inclusion criteria for multivariate analysis. A multivariate logistic regression model was therefore built containing these variables. The results are shown in Table 25.

Table 25 ***Nine month follow-up: Multivariate logistic regression model for Trial Calendar data (dependent variable: questionnaire response/non-response)***

Variable	Regression Co-eff	Standard Error	Odds Ratio	95% CI
CAL	-0.424	0.244	0.654	0.406, 1.055
AGE(1)	-1.637	0.498	0.195	0.073, 0.517
AGE(2)	-1.096	0.514	0.334	0.122, 0.915
AGE(3)	-0.989	0.531	0.372	0.131, 1.053
SEX	0.240	0.259	1.271	0.764, 2.113
EMP	-0.442	0.299	0.643	0.358, 1.155

The likelihood ratio (-2 log likelihood) for the multivariate model is 385.244 Comparing the difference of this value to the likelihood ratio of the constant only model of 407.297 gives a G value of 22.053 with a significance value of 0.001. This indicates that the model is significantly different from the constant only model and is useful for predicting response. The model was refined by taking the most significant predictive variable (AGE) and adding the other variables in the multivariate model in turn (EMP, CAL and SEX). The results of this are shown in Table 26.

Table 26 ***Nine month follow-up: Refined multivariate model***

Variables	-2 Log Likelihood (Likelihood Ratio)	G	Degrees of freedom	χ^2 value*	Sig
Constant	407.297				
AGE	391.064	16.233	3	7.815	0.001
AGE + EMP	389.134	1.930	1	3.841	0.165
AGE + CAL	387.919	3.144	1	3.841	0.076
AGE + SEX	390.605	0.459	1	3.841	0.498

* 0.05 significance level

At nine months AGE is the only variable which has a significant effect on response. Adding EMP, CAL and SEX does not contribute significantly to the predictive ability of the model. The refined model including AGE and CAL, however, is approaching significance. The relationship between AGE, CAL and response is shown in Table 27

Table 27 ***Crosstabulation of nine month response rates showing relationship between age and Trial Calendar possession. Figures are observed and expected numbers responding and non-responding***

Age		No Calendar		Calendar	
		Observed	***Expected***	***Observed***	***Expected***
16-24	Responder	26	25	28	29
n=123	Non-responder	31	32	38	37
	Total	**57**	**57**	**66**	**66**
25-34	Responder	19	24	29	24
n=83	Non-responder	22	17	13	18
	Total	**41**	**41**	**42**	**42**
35-44	Responder	15	19	21	17
n=59	Non-responder	16	12	7	11
	Total	**31**	**31**	**28**	**28**
45-57	Responder	14	14	11	11
n=31	Non-responder	3	3	3	3
	Total	**17**	**17**	**14**	

Possession of a Trial Calendar appears to have no effect on response in the older and younger age groups but significantly improves response in the middle age groups.

Finally, the interaction between the main effects of the logistic regression and CAL was again assessed and the results are shown in Table 28.

Table 28 ***Nine month follow-up: Effect of addition of interaction effects to previous main effects model***

Interaction	-2 Log Likelihood (Likelihood Ratio)	G	Degrees of freedom	χ^2 value*	Sig
Constant	407.297				
Main effects	385.244	22.053	6	12.592	0.001
AGE X CAL	379.533	5.711	3	7.815	0.127
EMP X CAL	383.085	2.159	1	3.841	0.142
SEX X CAL	377.610	7.634	1	3.841	0.006

* 0.05 significance level

Adding interaction effects of CAL to the main effects model shows that there is a significant relationship between SEX and CAL at the nine month follow-up point. Again, this model adequately fits the data as indicated by a non-significant result of the Hosmer and Lemeshow goodness if fit test.

The data were explored in more detail and are presented in Table 29.

Table 29 ***Crosstabulation of nine month response rates showing the relationship between sex and Trial Calendar possession. Figures are observed and expected numbers responding and non-responding***

Sex		No Calendar		Calendar	
		Observed	***Expected***	***Observed***	***Expected***
Male	Responder	46	44	44	46
n=174	Non-responder	39	41	45	43
	Total	**85**	**85**	**89**	**89**
Female	Responder	28	37	45	37
n=122	Non-responder	33	24	16	24
	Total	**61**	**61**	**61**	**61**

Exploring the data this way shows a significant difference in response rates between females possessing a Trial Calendar and those not. Of the females in possession of a Trial Calendar, 45/61 (74%) responded to the nine month questionnaire. The Trial Calendar had no significant effect on response rate amongst males.

6.3.3.2 Outcome Two

- Effect of the Trial Calendar on the amount of prompting required by participants to return their questionnaires at each time point.

The effect of the Trial Calendar on the amount of prompting required to encourage questionnaire return was analysed using the response categories assigned to CAST participants (see Chapter 4 section 4.3.1.4). The results of this analysis are shown in Table 30.

Table 30 ***Effect of the Trial Calendar on amount of prompting required to return questionnaire at each follow-up point. Values are numbers responding (percentage).***

		No Calendar	Calendar	χ^2
Four week follow-up point		**n=146**	**n=150**	
Response Category :	**1**	55 (38)	50 (33)	
	2	38 (26)	38 (25)	
	3	9 (6)	19 (13)	
	4	12 (8)	10 (7)	
	5	8 (6)	2 (1)	
	6	24 (16)	30 (20)	
Withdrawal from trial		0 (0)	1 (1)	
				0.162
12 week follow-up point		**n=146**	**n=150**	
Response Category :	**1**	46 (32)	38 (25)	
	2	35 (24)	34 (23)	
	3	7 (5)	15 (10)	
	4	8 (6)	9 (6)	
	5	25 (17)	21 (14)	
	6	25 (17)	31 (21)	
Withdrawal from trial		0 (0)	2 (1)	
				0.374
9 month follow-up point		**n=146**	**n=150**	
Response Category :	**1**	26 (18)	32 (21)	
	2	23 (16)	30 (20)	
	3	10 (7)	14 (9)	
	4	15 (10)	13 (9)	
	5	28 (19)	24 (16)	
	6	44 (30)	35 (23)	
Withdrawal from trial		0 (0)	2 (1)	
				0.465

The results of the analysis shows that the Trial Calendar had no significant effect on the amount of prompting required to return the questionnaire at any of the follow-up time points.

CAST also categorised responders as either 'keen' or 'reluctant' by collapsing the response categories into groups (see Table 5, chapter 4). Response categories 1 and 2 were grouped as 'keen' responders and response categories 3 and 4 were grouped as 'reluctant' responders. Table 31 shows the effect of the Trial Calendar in terms of these two categories.

Table 31 ***Effect of the Trial Calendar on categorisation of 'keen' or 'reluctant' responder at each follow-up point. Values are numbers responding (percentage).***

		No Calendar	Calendar	χ^2
Four week follow-up point		**n=114**	**n=117**	
Response Category :	**Keen**	93 (82)	88 (75)	
	Reluctant	21 (18)	29 (25)	0.240
12 week follow-up point		**n=96**	**n=96**	
Response Category :	**Keen**	81 (84)	72 (75)	
	Reluctant	15 (16)	24 (25)	0.106
9 month follow-up point		**n=74**	**n=89**	
Response Category :	**Keen**	49 (66)	62 (70)	
	Reluctant	25 (34)	27 (30)	0.638

The Trial calendar had no significant effect on classification of participants as 'keen' or 'reluctant' responders at any follow-up point.

6.3.3.3 Outcome Three

- Effect of the Trial Calendar on amount of missing data of the pre-defined core outcome questions for which the calendar 'reminder captions' are used (see section 6.2.2.3).

There were some issues regarding the design of the CAST questionnaires that made an analysis of missing data more difficult than expected. A full explanation of this is given in section 6.4. The data available for analysis is presented below:

- Return to work

The Trial Calendar contained a reminder caption the month after recruitment asking participants if they had returned to work. If they had they were prompted to use the calendar to remind them how many days they were absent from work due to their ankle injury. Data are available at the four week follow-up point for 158 responders who were employed either full or part time and had returned to work. The 'return to work' prompt mapped to question 57 on the follow-up questionnaire (see appendix 2). This question asks participants to note how many days they were off work. Table 32 shows the effect of the Trial Calendar on the amount of missing data for this question.

Table 32 ***Effect of Trial Calendar on amount of missing data for 'number of days off work' question at four week follow-up. Figures are number of questionnaires with answers missing or filled in (percentage).***

	No Calendar n=81	Calendar n=77	χ^2
Missing	17 (21)	22 (30)	
Filled in	64 (79)	55 (70)	0.269

The Trial Calendar had no significant effect on the amount of missing data for the return to work core outcome question.

- Return to sport

The Trial Calendar contained a reminder caption about return to sport four months after recruitment. This prompt mapped onto question 87 on the follow-up questionnaires (see appendix 2). This question is in two parts and asks if the participant had fully returned to sport and if so how long this took. If a participant had returned to sport it was necessary for them to answer the 'time taken to return to sport' part of the question. The effect of the Trial Calendar on how well this question was completed is shown in Table 33.

Table 33 ***Effect of Trial Calendar on amount of missing data for 'time taken to return to sport' question. Figures are number of questionnaires with answers missing or filled in (percentage).***

		No Calendar	Calendar	χ^2
Four week follow-up point		**n=12**	**n=11**	
	Missing	1 (8)	0 (0)	
	Filled in	11 (92)	11 (100)	0.328
12 week follow-up point		**n=19**	**n=24**	
	Missing	5 (26)	4 (17)	
	Filled in	14 (74)	20 (83)	0.440
9 month follow-up point		**n=31**	**n=33**	
	Missing	4 (13)	4 (12)	
	Filled in	27 (87)	29 (88)	0.925

At the four week follow-up only 23 responders to the questionnaire had returned to sport. This number increased at each time point as would be expected although at nine months only 64 responders had returned to sport. The Trial Calendar had no significant effect on the amount of missing data for the 'return to sport' core outcome question at any follow-up point.

- Resource use

Key secondary outcomes in CAST were whether the participant had to spend any resources on medicines or required further medical input for their ankle sprain. These data were collected at the 12 week and 9 month follow-up points. The reminder caption for resource use mapped onto question 89 in the twelve week and 9 month follow-up questionnaires (see appendix 2) This question asked if the participant had received any further treatment for their ankle injury apart from the treatment received as part of CAST. If the answer was yes to this question, participants were asked to elaborate on this by providing details of who they received treatment from and how many times. The effect of the Trial Calendar on how well this question was completed is shown in Table 34

Table 34 ***Effect of Trial Calendar on amount of missing data for 'further treatment' question. Figures are number of questionnaires with answers missing or filled in (percentage).***

		No Calendar	Calendar	χ^2
12 week follow-up point		**n=28**	**n=39**	
	Missing	3 (11)	2 (5)	
	Filled in	25 (89)	37 (95)	0.391
9 month follow-up point		**n=12**	**n=20**	
	Missing	1 (8)	5 (25)	
	Filled in	11 (92)	15 (75)	0.242

At the 12 week follow-up 67 participants indicated that they required further treatment. At the nine month follow-up an additional 32 participants required further treatment. The Trial Calendar had no significant effect on the amount of missing data for the 'further treatment required' secondary outcome question at either the 12 week or nine month follow-up.

6.3.4 Summary of results

The response behaviour of a total of 296 CAST participants was analysed at three follow-up points over a nine month period. Participants were aged between 16 and 57 years (mean 29.5, SD 10.3) including 174 males and 122 females.

At the four week follow-up point the Trial Calendar had no significant effect on response. Logistic regression including interaction effects of independent variables approaching significance and possession of a Trial Calendar also failed to demonstrate any significant effect of the Trial Calendar.

At the 12 week follow-up point females are significantly better at responding than males. Logistic regression, however, demonstrated no interaction effect between sex and the possession of a Trial Calendar.

Logistic regression of the nine month follow-up data showed an interaction effect between sex and Trial Calendar possession at this follow-up point. Females with a Trial Calendar have approximately double the odds of responding that those without a Trial Calendar.

The Trial Calendar had no effect on the amount of prompting required to encourage questionnaire return at any follow-up point. There was also no effect of the Trial Calendar on the amount of missing data of core primary and secondary outcome questions.

6.4 Discussion

This randomised controlled trial was undertaken to assess the effect of the Trial Calendar on response rates in a clinical trial. The principal finding was that at the four week and twelve week follow-up points the Trial Calendar had no significant effect on response. At the nine month follow-up point females with the Trial Calendar were significantly better at responding. This finding may be a chance event due to the large number of tests performed, and caution is therefore necessary in drawing conclusions. Such results may be most appropriately used for generating hypotheses to be tested in future studies.

The utility of postal questionnaires has stimulated many studies on how to improve response to this method of data collection. Very few studies, however, have concentrated on assessing response to postal questionnaires in health care research and fewer still in the context of follow-up in a clinical trial. This study used a 'real' clinical trial setting (CAST) and nested the Trial Calendar within the trial to assess its effect on response. The Trial Calendar was designed as a pre-notification, prompting and post-notification tool to remind participants to return their

questionnaires. Such participant contact has been shown to be effective in health care research.[191] The aim of the Trial Calendar was to combine the principles of pre and post questionnaire contact within a single tool. This would be cost effective in terms of both time and resources. At the same time the calendar made appeals to the 'social exchange' involved in responding.[37] It was expected that the Trial Calendar would significantly increase response rates in CAST. Previous health care research into the effect of participant contact on response rates has used either telephone or postal contact to prompt questionnaire return. Sutherland et al (1996)[117] used Dillman's 'Total Design Method'[37] in the postal follow-up of women in a cancer prevention trial. Response was increased by almost one third using this method. Salim Silva et al (2002)[168] used telephone reminders in a survey of women with 'work related neck and upper body disorders'. They found that telephone reminders significantly improved response rates. The sample size in this study, however, was small and the results should therefore be viewed with caution. Both these studies were confined to female populations. This study showed that at long term follow-up (nine months) the Trial Calendar had a significant effect on response in females. This raises the possibility that females are more susceptible to follow-up strategies than males.

CAST, overall, had excellent response rates at each follow-up point: 83%, 82% and 76% at four week, 12 week and nine months respectively. The CAST sample of 296 participants used for the RCT of the Trial Calendar

demonstrated comparable response rates when including those participants who gave core outcome responses by telephone. These return rates are, in part, reflective of the dedication of the trial team in following the agreed protocol of follow-up. The Trial Calendar was used in addition to the comprehensive follow-up protocol which could be an explanation for its disappointing effect on increasing response further. A 'saturation' point will come in a trial when further improvement in response is not possible due to participants being 'immune' to further follow-up strategies. Reasons for this immunity could be simply a refusal to cooperate, lost contact, moved house or died etc. This possible explanation for the ineffectiveness of the Trial Calendar is strengthened by the response rate at the nine month follow-up point. At this time point overall response rate is at its lowest and the effect of the Trial Calendar on improving response is approaching significance (and is significant in the female population).

The relevance of the questionnaire to the participants is another factor which is likely to have contributed to the overall high response rates in CAST. In an extensive literature review, Heberlein and Baumgartner (1978)[20] found that surveys with non-salient questions averaged a 42% response rate while questionnaires judged to be salient to the respondent averaged a 77% return rate. It could be argued that a questionnaire used in any clinical trial as a method of follow-up will be highly relevant to the participant. This results in a 'positive disposition' to respond.[57] It has been noted in the survey literature that as the number of positively disposed

participants increases, the effect of efforts to persuade non-responders to cooperate is diminished.[57]

The Trial Calendar contained pre and post notification prompts to warn participants of an impending questionnaire delivery and then to prompt them to return their completed questionnaire. A further hypothesis was that the Trial Calendar would therefore reduce the amount of additional prompting required by participants to return their questionnaire. This was, however, not the case and there was no significant difference between the prompting required by participants in possession of a Trial Calendar and those not. For the calendar to act as a prompting tool it necessitates participants keeping the calendar visible and utilising it as a desk or wall calendar. Data on how well the Trial Calendar was used were not collected in this study. However, qualitative data about the participant's views of the Trial Calendar were collected as part of a qualitative study (described in chapter 7). A purposive sample of 22 CAST participants was interviewed. The sample contained participants who had received a Trial Calendar and those who had not. Those participants receiving a Trial Calendar were asked whether they found it useful and actually used it for its intended purposes. Those participants who had not received a Trial Calendar were shown an example of the calendar, given a description of its use and then asked for their thoughts on the idea. More than half the participants who received a Trial Calendar reported that they used it for its intended purpose and found it useful. The rest either lost the calendar or put it in a drawer and forgot about

it. All the participants who did not receive a Trial Calendar thought the idea was good. Many thought that they would have used the calendar, had they received one, to help them answer the questionnaires.

These results show that for prompts to be effective they need to be high profile and timely – for example a telephone call the day after a participant receives a questionnaire. Prompts embedded within a tool such as the Trial Calendar are ineffective. It is evidently much easier to ignore a written prompt than a human voice appealing for a questionnaire to be returned.

A final hypothesis of this trial was that the Trial Calendar would reduce the amount of missing data of core primary and secondary outcome questions. This was to be achieved by the calendar containing prompts to record key recovery events such as return to work and sport. Attempts, however, to analyse the effect of the Trial Calendar on missing data of core outcome questions proved to be problematic for a number of reasons. The root cause of the problems with this analysis stemmed from the timing of the development and design of the Trial Calendar. The idea for the Trial Calendar developed during the initial recruitment phase of CAST. At this point in CAST the outcome questionnaires had already been developed and formatted. This process was overseen by the Trial Steering Committee and the questionnaires had been approved by the relevant research ethics committees. It was therefore not possible to alter the questionnaires without causing significant disruption to the CAST administration. The primary outcomes for CAST were return to pre-injury mobility and return to pre-injury

occupation/activities. These data were collected using a range of outcome measures but also using some single questions. The Trial Calendar contained prompts to remind participants to record when key events such as return to sport and return to work occurred. It was anticipated that the effect of the Trial Calendar on completion of these single questions could be analysed. This would indicate whether the calendar was useful in helping participants recall and record key recovery events thus reducing the amount of missing data. The way that some of these single questions were worded, however, enabled participants to skip the question if it did not apply to them. This caused difficulties in the analysis as it was impossible to know whether a participant had skipped a question because it was not applicable or whether an answer was appropriate but missing. Had the Trial Calendar and the questionnaires been developed simultaneously the implications of having 'skippable' questions on the subsequent analysis would have been apparent. It would then have been possible to word the key questions so as to make the analysis of missing data more straightforward. It was, however, possible to salvage some information on the amount of missing data from the core outcome questions as is presented in the results. From the available data there is no evidence that the Trial Calendar is effective in reducing the amount of missing data. An explanation for this ineffectiveness could be similar to the possible reasons for the ineffectiveness of the Trial Calendar in improving response rates as a whole. If the prompts are not

visible enough and the Trial Calendar is not used then the possible effectiveness of the calendar on reducing missing data will be lost.

The body of evidence concerning methods of preventing missing data in clinical trials is small compared to studies which investigate how to handle missing data once it occurs.[192] Several authors have, however, recently considered the potential for preventing missing data in trial design.[193, 194] Wisniewski et al in 2005[192] suggests several steps to minimising the amount of missing data including clarity of the study documentation and participant contact. These studies, however, have not specifically investigated missing data from postal questionnaires. To address this issue a Cochrane review has recently been proposed which aims to quantify the effects of methods to influence the completeness of response to self-administered questionnaires.[195] Once completed this review will offer useful information for clinical researchers using questionnaires as a method of follow-up.

6.5 Conclusions

Despite its anticipated usefulness, this study failed to demonstrate any effect of the Trial Calendar in terms of improving response and reducing missing data in postal questionnaire follow-up in a clinical trial. Although the overall result is disappointing some useful conclusions can still be drawn. A logistic regression analysis did suggest a significant effect of the Trial Calendar in improving response in females at the nine month follow-up. This

suggests that females may be more susceptible to this type of follow-up technique in the long term. This result, however, was arrived at after multiple testing of the data and the possibility remains that the difference seen is due to chance.

The results also suggest that clinical trial participants require follow-up contact and prompting to be high profile and timely. Although participants felt that the Trial Calendar was a good idea it is possible that its use as a prompting tool was too subtle.

Additionally, this study has been useful in highlighting the need for self-administrated questionnaires to be as unambiguous as possible. This includes careful wording of questions to avoid participants being able to skip questions if they are not applicable. This will help to reduce item non response.

Finally, the design, production and implementation of the Trial Calendar into CAST involved an additional cost to the trial. This study has demonstrated in a well conducted, adequately powered randomised controlled trial that this approach to improving response and minimising missing data is ineffective. Allocating funds to this type of follow-up approach is therefore not recommended in future acute injury clinical trials.

6.6 Chapter summary

This chapter described, in detail, the methodology and results of a randomised controlled trial of a method of improving response to postal questionnaire follow-up in a clinical trial.

Although the results failed to support the hypothesis, valuable information regarding the response behaviour of clinical trial participants has been gained. This will be integrated with the findings from previous and subsequent chapters and will contribute to the overall conclusions of this thesis.

In contrast to this quantitative methodological approach to investigating factors affecting questionnaire response, the next chapter details a qualitative study of clinical trial participant's response behaviour.

7 Chapter 7

Issues surrounding response and non-response to postal questionnaires in a clinical trial – the patients perspective

This chapter addresses the lack of literature concerning the patient's perspective and motivations to respond to postal follow-up in a clinical trial. No published studies can be found which address the issues of response and non-response in clinical trials by asking the participants themselves. The way that theories of human behaviour have been applied to survey research to explain response decisions was discussed in chapter 2. The systematic review reported in chapter 5 revealed a wealth of literature concerning response issues in survey research. It cannot be assumed, however, that the response behaviour of survey participants is transferable to clinical trial participants. Previous chapters have also highlighted and explained the importance of maintaining adequate response rates in clinical trials. This chapter aims to identify the motivations and thought processes people utilise in their decision to respond or not to questionnaire follow-up in a clinical trial. To achieve this, a qualitative methodological approach was used to allow clinical trial participants to verbalise their response decisions in their own words.

The essential questions are therefore:

I. Why do people respond to postal follow-up in a clinical trial?

and, more importantly,

II. Why do people choose not to respond?

If these questions can be answered it may then be possible to suggest ways to improve response rates by understanding the deeper thought processes used by individuals in their response decisions.

7.1 Background

Fundamental to survey research is the long-standing practice of finding things out by asking people questions.[196] Scientific social surveys were pioneered by Booth in the late 19th century in the investigation of the living conditions of Londoners. Surveys are now used in a wide variety of investigations ranging from Gallup Polls, to town planning surveys and market research as well as investigations sponsored by research institutes, universities and the government.[24]

Survey researchers have often collaborated with social psychologists in attempts to understand the response decisions of survey participants.[74, 86] Some of the most well cited behavioural theories which have been applied to survey research have been elaborated upon in chapter 2. Groves and Couper,[86] for example, have used the theory of 'compliance' to help explain response behaviour. The personal relevance of the questionnaire to the respondent is seen as a major factor in whether the respondent

complies and returns the questionnaire. Groves and Couper believe that a potential survey respondent does not usually have a large personal interest in the subject of the survey. Questionnaires used in a clinical trial, however, usually have a strong personal relevance for the participant. A higher response rate should therefore be expected from clinical trial participants compared to survey participants to whom the issues in the questionnaire are of minimal concern.[197] Relevance of the questionnaire is a major difference between survey research and clinical trials. Clinical trial participants may consider the reasons and benefits of responding more deeply due to a more personal involvement in the research process. Theories of 'Reactance',[72] 'Dissonance'[76] and the 'Functional' theory of attitudes[81] have also been used in explanations of survey response behaviour. More recently, the 'Leverage-Saliency' theory has been developed specifically to explain survey response.[57] No reference can be found of any application of these theories to explain the response behaviour of clinical trial participants.

The methodology of a clinical trial differs to that of a survey. There may also be differences in the motivations of subjects recruited onto a clinical trial to respond to follow-up. A fundamental similarity, however, is that researchers conducting clinical trials often want to find out about the effectiveness of a treatment intervention by asking the patient questions. These questions could be in the format of interviews or questionnaires. Although no previous work has focussed specifically on clinical trial participants, the similarities of

the methods of data collection with survey research should not be overlooked. It may be possible to apply concepts surrounding response issues identified in survey research to the clinical trial setting. Before this assumption can be made, however, clinical trial participants need to be evaluated independently. Only then can response theories identified in survey research be accepted or rejected. This is the main focus of this chapter.

7.2 Methods

To gain an insight into clinical trial response issues a qualitative study was conducted. This involved interviewing participants who were currently involved in a clinical trial.

7.2.1 Sample selection

The sample was selected from participants in a randomised clinical trial of different mechanical supports for severe ankle sprains (CAST). A purposive sample of participants was sought to represent the diversity of CAST participants in various dimensions which are outlined below. This type of sampling ensures a wide range of experience and views.[198]

The CAST database was searched to locate participants with a range of characteristics including age, sex, level of education, occupation and type of ankle support received as part of the trial. The effects of these characteristics on response are discussed in chapter 8.

CAST used postal questionnaires to collect information on a range of outcome measures (see chapter 4 section 4.2.5). Participants were selected if they had been sent at least one follow-up questionnaire. The sample contained both responders and non-responders. CAST has a standardised protocol for chasing late questionnaires using postal and/or telephone prompts. This is explained in detail in chapter 4. In summary, participants were classed as a responder if they had returned all the questionnaires sent to them (even if this required some prompting). Non-responders were those participants who had failed to send back at least one of their questionnaires despite prompting. To provide detail on the amount of prompting required, participants were coded into a 'response category' according to their response to each follow-up questionnaire. The coding method is detailed in chapter 4 section 4.3.1.4. To summarise, the codes are:

1 Questionnaire returned with no prompting

2 Questionnaire returned after one telephone prompt

3 Second copy of questionnaire sent and returned with no further prompts

4 Second copy of questionnaire returned following further telephone prompt

5 Questionnaire not returned but core data obtained over the telephone

6 Questionnaire not returned. No data obtained by telephone

Categories one to four are classified as 'responders' and categories five and six are classified as 'non-responders'. Categories one and two were further grouped as 'keen responders' and categories three and four were grouped as 'reluctant responders'.

7.2.2 Data collection

To gain an insight into the views of clinical trial participants, semi-structured interviews were conducted. Ethical approval to interview CAST participants was gained from the Northern and Yorkshire Multi-Centre Research Ethics Committee. The participants were approached either by letter or telephone to ascertain their willingness to take part in the qualitative study. Once this was confirmed the date and venue of the interview was arranged. All the interviews were audio-tape recorded with the participant's consent and later transcribed verbatim.

The interviews were conducted either in the participant's home, work place or in a neutral venue. In all, 22 participants were interviewed and eight of these were interviewed by telephone. This mixture of interview locations and formats was partly to make the process as convenient as possible for the participants and partly due to the geographical spread of the CAST trial centres. In order not to exclude participants from outlying trial centres, telephone interviews were conducted. It was beyond the time and resource constraints of this qualitative study to travel to the more distant locations to conduct all the interviews face to face. The limitations of telephone interviews compared to face to face interviews were, however, considered.

Establishing and maintaining rapport and noticing non-verbal communication in a telephone interview is more difficult.[199] However, confining the interviews to just participants in the Coventry area, in order to conduct them all face to face, would be more limiting than mixing the interview methods. The nature of the research question was to establish reasons for response decisions rather than to collect in-depth patient experiences of, for example, a chronic or sensitive medical condition. Conducting some of the interviews by telephone was therefore considered to be acceptable.

7.2.2.1 Development of the interview technique and questions

Semi-structured interviews were appropriate as clearly defined areas of interest had been identified (i.e. reasons for response/non-response). The open ended nature of semi-structured interview questions defines the topic but allows discussion of the topics in more detail than the structured approach.[200] To establish a rapport with the interviewees and put them at ease, a relaxed, polite manner was adopted.[198] The aim of the study was explained to the participants with an emphasis on the confidentiality of their responses to the questions. The preparation of the interview questions was guided by the research questions. Care was taken, however, not to overly structure the questions so as to constrain the respondent.[201] Open-ended, clear questions were used. Questions that required only a dichotomous response were avoided as they would not encourage the participant to expand on the topic. If needed, probes were used after a

participant's response such as 'can you explain that a bit more'. This encouraged participants to elaborate their responses providing more depth to the interview. The interviews commenced with an 'ice-breaking' question to relax the participant.[101] This question pertained to the participant's ankle injury: 'So, how is your ankle now?' and enabled the participant to talk freely about their recovery. This provided a natural lead to bring the questioning around to their participation in CAST.

The basic structure of the interview questions is shown in Table 35. See appendix 6 for full details of the interview schedule.

Table 35 ***Basic structure of interview questions***

Question 1	How is your ankle now?
Question 2	What was your understanding of the trial when you were asked to take part at the ankle trial clinic?
Question 3	Suppose you have a questionnaire sent to you 'out of the blue' like a household survey. Do you usually respond to it? (Explain)
Question 4	**Responders:** So what encouraged you fill in our questionnaire and send it back? **Non-responders:** Can you tell me any specific reasons why you didn't send your questionnaire back.
Question 5	What do you think we could have done to make it easier for you to complete and return our questionnaire?
Question 6	What are your thoughts about the whole questionnaire? For example; layout, appearance, length and time to complete.

An interpretative theoretical standpoint was taken during the interviews. The intention was to uncover reasons and meanings behind the motivations of participants to either respond or not to postal questionnaires in a clinical trial. This draws on the social interactionist perspective, a branch of interpretivism which focuses on the meanings individuals attribute to their activities and their environments.[202] It was recognised, however, that participants, could be influenced by the interaction with the interviewer. Such a perspective has been termed 'hermeneutics' which argues that there is an essential interaction between the interpreter and the object of interpretation.[197] This could lead interviewees to say what they thought the interviewer wanted to hear rather than their true feelings. This was a special concern for the non-responders who could have felt awkward when probed for reasons for their non-compliance. These potential issues were addressed by the participants being assured that there were no right or wrong answers. It was also stressed that the reason for the interview was a genuine attempt to understand response behaviour and was in no way a means of reprimanding the participant for not responding. Participants were assured that whatever they said was confidential and would not affect their health care or continued participation in the clinical trial.

7.2.3 Analytic framework

The approach taken to data analysis is 'Framework Analysis'. This is a recent approach to qualitative analysis which is gaining popularity in health related research.[203] In contrast to an analysis approach such as grounded

theory, framework analysis was developed to meet specific research information needs and provide outcomes or recommendations. A rationale for this choice of analytic framework is given in section 7.6.

7.2.3.1 The development of 'framework analysis'

Framework analysis was developed in the context of conducting applied qualitative research by a team from the National Centre for Social Research, London. The approach was initially developed in the context of applied policy research. However, it has been refined and developed over the years and is now used across a wide range of studies including health care.[204] There are two key features of the context in which the use of qualitative methods within social policy has developed. Firstly, research within this area is commissioned and funded by public bodies which have certain requirements of the research they commission. Secondly, organisations conducting applied policy research have strong quantitative research traditions. These features mean that emphasis is placed on producing qualitative evidence that is collected and analysed with rigour. A key element in the development of framework analysis is the 'visibility' of the qualitative method. This leads to data which are valid, unbiased and clear in how interpretations have been reached.[205] Qualitative methodology receives criticism from advocates of quantitative methods who question how the findings of the research have been obtained.[206] The systematic and explicit nature of framework analysis responds to this. The developers of this approach suggest that this brings a greater confidence in the

methodology and a deeper understanding of what qualitative research can achieve.[204] By adhering to these developmental principles, framework analysis does not fit into any of the traditional 'schools' of qualitative research. The approach is an eclectic mix of different traditions which is seen by the developers as a significant strength.[205]

Another key feature in the development of framework analysis is that it is grounded in, and driven by, the original accounts and observations of the people it is about.[204] This is one of the major differences between framework analysis and grounded theory which deliberately attempts to 'fracture' the data in order to open up new avenues for analysis.[207]

7.2.3.2 The epistemological stance of framework analysis

The epistemological stance of framework analysis very much reflects the quantitative context in which it has developed. Aspects of the scientific method have been adapted to suit the goals of qualitative research. Framework analysis therefore draws on 'positivism' - a conceptualisation of research which assumes that methods used to measure the physical world can be modified and used to study the social world. This adoption of positivist ideals is most apparent in the dedication to be as objective as possible in the collection, interpretation and presentation of the data. This would initially seem a rather dogmatic approach but the developers are keen to stress an ontological stance of 'subtle realism' as described by Hammersley.[208] They accept that it is impossible for the researcher to

escape the social world in order to study it. According to a subtle realist view, true objectivity in qualitative research can never be fully achieved and the importance of the personal interpretations of both the study participant and the researcher is acknowledged.

Interpretivism and pragmatism are also embraced epistemologically by framework analysis. The acceptance of interpretivism is reflected in the understanding that meaning emerges through interaction and is not standardised across social and cultural groups. A pragmatic view is taken by emphasising the value of choosing the most appropriate methods of research to address specific questions rather than being limited to a consistent philosophical approach.

The purest qualitative researcher may criticise this approach for 'borrowing' concepts from different traditions within social research. The degree of consistency between the researchers' beliefs and research practices often guides the evaluation of the quality of qualitative research. The inability to assess this consistency in such an eclectic approach may be seen by some as a weakness.[205] The developers of framework analysis, however, describe in detail the key parameters within which they carry out qualitative research. (see Snape and Spencer in Ritchie and Lewis 2003[205])

7.2.3.3 Stages of 'framework analysis'

(See Ritchie and Spencer in Bryman and Burgess 1994 Chapter 9 [204])

Framework analysis has five key stages which can be carried out in a linear fashion allowing all the data to be collected before analysis begins:

1. Familiarisation

Familiarisation involves immersion in the data. The tapes were listened to and the transcriptions were read several times to gain a feel for the material as a whole. Notes were made of key ideas and recurrent themes.

2. Identifying a thematic framework

This was an initial coding framework which was developed from issues emerging from the familiarisation stage. Notes were made recording the range of responses to questions posed and recurrent themes. To enhance reliability, this process was repeated by a second researcher on a sub-set of five randomly chosen interview transcripts. Any differences in the interpretation of emergent themes were discussed. The framework was then applied to a few transcripts. Categories were then refined and became more responsive to emergent and analytical themes.

3. Indexing

This is the process of applying the thematic framework to the data. This is commonly called 'coding' in other qualitative approaches. Numerical indexing references were recorded in the margins of the transcripts which linked back to the thematic framework. Indexing involves judgements as to

the meaning and significance of the data. This is a subjective exercise but by annotating the transcripts the process was made visible and accessible to others.

4. Charting

Headings from the thematic framework were used to create charts of the data so that the whole dataset could be easily read. This built up a picture of the data as a whole. Data were lifted from their original context and rearranged according to the appropriate thematic reference. This charting facilitated comparisons both within and between cases.

5. Mapping and interpretation

The final stage of framework analysis involved searching for patterns, associations, concepts and explanations in the data and, if appropriate, developing typologies. Typologies are specific forms of classification that help to explain the way that phenomena can be characterised or differentiated.[209]

Although not specifically a process of framework analysis, frequency counts of emergent themes have been used in the analysis. There is some debate within qualitative research as to the appropriateness of counting the frequency with which themes are represented in the data. The size of the sample used in qualitative research and the way the sample is selected mean that any statements about the prevalence or distribution of a theme

are only applicable to the study sample itself.[210] There is a danger when presenting numbers that erroneous statistical inferences are drawn to the wider population. Some authors stress that 'qualitative research should be explaining patterns of recurrence, not simply stating that they exist'.[210] Green and Thorogood[207], however, suggest that simple frequency counts can increase the reader's faith in the validity of the interpretations of the researcher. They argue that counts are useful in giving some perspective on how common various views or experiences were and this can defend against anecdotalism.

A full critique of framework analysis and its application in the context of this study is given at the end of this chapter.

7.3 Results

The sample consisted of 22 CAST participants, 11 males and 11 females. The age range was from 16 to 62 with a mean age of 34. Table 36 gives details of the participant characteristics.

Table 36 Interview Participant Characteristics

Date	Place	Int by	ID	Location	Gender	Age	Job	Profession	Driver	Sports	Time since	Mech	Treatment	Resp 4wks	Resp 12wks
28/01/2004	C&W	RN	1063 MH	Cov	M	19	student	student	yes	cycling walking skateboarding	4mths	fell down stairs	Tubigrip	1	1
04/02/2004	C&W	RN	1078 MH	Cov	M	38	professional	IT analyst	yes	cycling jogging football raquets	4mths	football	Plaster	2	2
09/02/2004	Home	RN	1072 MF	Cov	F	50	Housewife	Housewife	no	walking	4mths	fell pot hole	Bledsoe	2	4
14/02/2004	C&W	RN	1075 GK	Cov	M	44	skilled man	skilled man	yes	swimming walking DIY	4mths	uneven ground	Aircast	2	2
14/02/2004	Home	RN	1062 JB	Cov	F	62	retired	retired	yes	nil	5mths	tripped and fell	Aircast	3	5
14/02/2004	Home	RN	1061 AO	Cov	F	52	skilled man	soft furnisher	yes	swimming walking	5mths	fell kerb	Bledsoe	3	2
17/02/2004	C&W	RN	1076 IH	Cov	M	34	skilled non man	Retail	yes	swimming jogging football	4mths	football	Tubigrip	2	2
22/04/2004	C&W	RN	1027 JO	Cov	F	37	unskilled manual	unskilled manual	no	cycling swimming	9mths	?	Bledsoe	6	5
28/05/2004	Tel	RN	3017 A	JRH	M	25	professional	racing driver	yes	jogging weights aerobics	6mths	tripped up stairs	Aircast	6	5
10/06/2004	Tel	RN	3020 C	JRH	F	23	skilled man	skilled man	yes	weights aerobics cycling walking sailing	6mths	In the pub	Plaster	6	5
10/06/2004	Tel	RN	4009 J	BHH	F	38	unskilled non man	unskill non man	yes	walking	9mths	walking in heels	Aircast	6	5
07/10/2004	Work	RN	1165 SC	Cov	F	28	unskilled non man	unskilled non man	yes	netball walking	2mths	netball	Plaster	1	
07/10/2004	C&W	RN	1155 MF	Cov	F	52	Unemp	Unemp	yes	walking	4mths	fell bottom step	Tubigrip	1	2
07/10/2004	Home	RN	1157 CS	Cov	F	30	Housewife	Housewife	No	walking	4mths	fell off a chair	Aircast	3	2
07/10/2004	Tel	RN	1152 PL	Cov	M	20	skilled man	Fork lift driver	No	walking	4mths	fell kerb	Aircast	2	1
07/10/2004	C&W	RN	1154 LS	Cov	M	25	unskilled manual	warehouse operative	yes	swimming walking	4mths	fell kerb	Bledsoe	2	1
11/10/2004	Tel	RN	6003 CW	Rugby	F	32	Housewife	Housewife	yes	walking	10mths	went over	Bledsoe	6	5
11/10/2004	Tel	RN	1120 TB	Cov	F	36	Housewife	Housewife	no	swimming cycling walking	6mths	at home	Tubigrip	6	5
11/10/2004	Tel	RN	1103 SG	Cov	M	24	professional	insurance	no	swimming cycling football walking	8mths	football	Plaster	6	6
11/10/2004	Tel	RN	5012 AS	Redditch	M	16	student	student	no	weights cycling football walking	8mths	football	Aircast	6	5
12/10/2004	Work	RN	1161 JH	Cov	M	28	professional	hotel manager	no	swimming aerobics	3mths	fell down steps	Plaster	2	4
15/10/2004	C&W	RN	1163 NM	Cov	M	34	professional	engineer	yes	DIY	3mths	trampolining	Bledsoe	1	

The majority of participants had reached two follow-up time points (four weeks and 12 weeks) and had therefore received two questionnaires. Fourteen participants had responded to all the questionnaires sent to them leaving eight participants who had not responded to at least one of their follow-up questionnaires.

7.3.1 Thematic framework

In accordance with the stages of framework analysis outlined above, a thematic framework was constructed following familiarisation with the interview transcriptions. Five interview transcripts were given to a second researcher to code independently. Many different reasons for response and non-response were identified in the transcripts. These were also identified by the second researcher. Following discussion the response issues were grouped into agreed similar themes. There was good corroboration between the coding of the two researchers and there were no major discrepancies.

The final thematic framework is shown in Table 37

Table 37 ***Thematic framework of response issues of clinical trial participants***

1. Personal details (current)
- 1.1 Age
- 1.2 Sex
- 1.3 Occupation
- 1.4 Type of ankle support
- 1.5 Responder/non-responder

2. General survey response
- 2.1 Respond
- 2.2 Don't respond

3. Reasons for responding
- 3.1 Personal relevance
- 3.2 Agreed to take part
- 3.3 Obliged to respond
- 3.4 Altruism
- 3.5 Important project

4. Reasons for non-response
- 4.1 Internal aspects of trial
- 4.2 External factors beyond control
- 4.3 Personal 'blame'
- 4.4 Life events

5. Treatment preference
- 5.1 Expectations/beliefs
- 5.2 Tubigrip yes/no
- 5.3 Plaster yes/no
- 5.4 Bledsoe yes/no
- 5.5 Aircast yes/no
- 5.6 Satisfaction

6. Recovery
- 6.1 Full recovery
- 6.2 Residual symptoms
- 6.3 Initial symptoms
- 6.4 Work
- 6.5 Referred on

7. Suggestions to improve response
- 7.1 Questionnaire design
- 7.2 Delivery and follow-up
- 7.3 Incentives

7.3.2 Mapping and Interpretation

7.3.2.1 General survey response

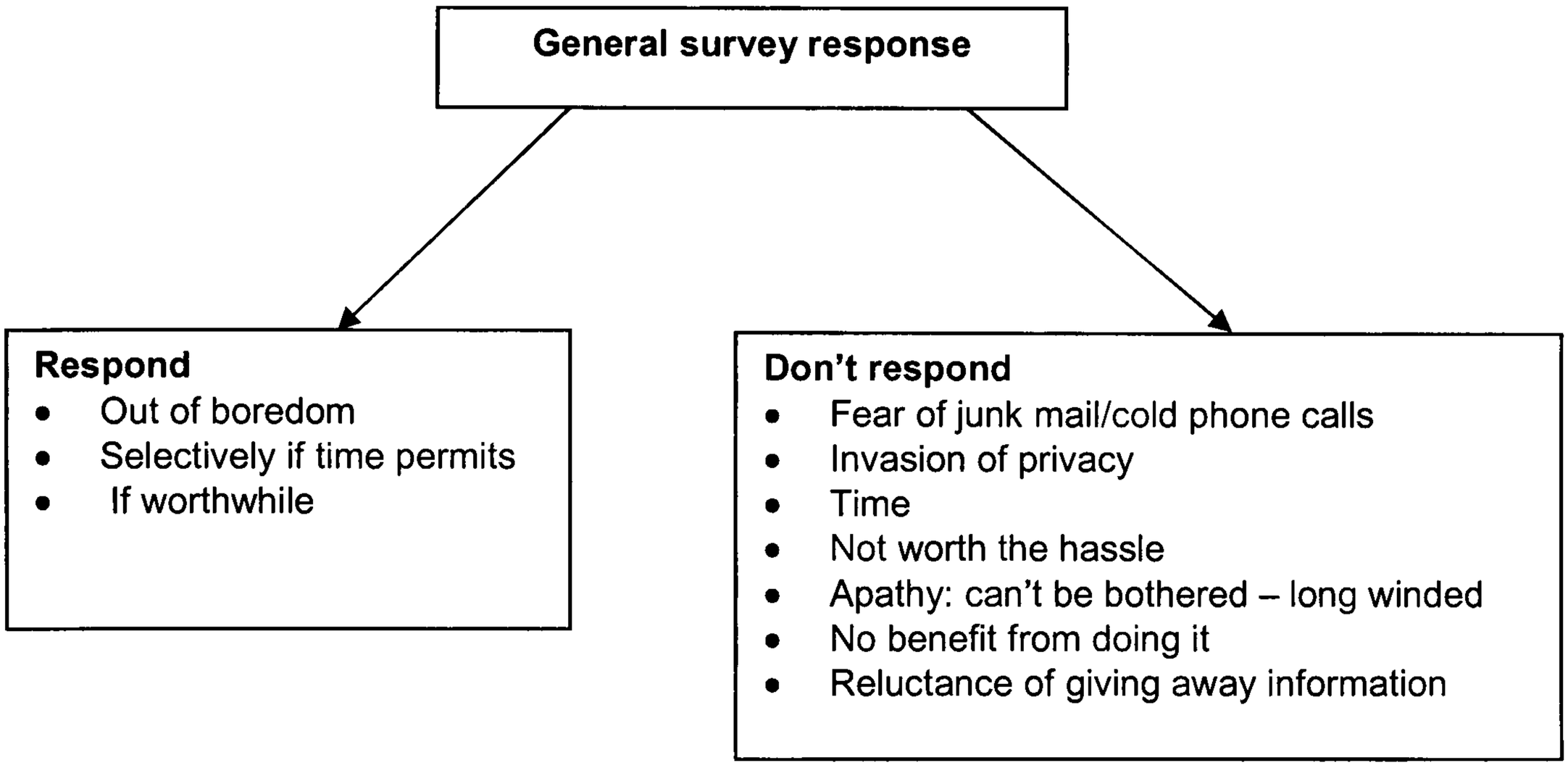
In order to establish possible differences in the motivations to respond to a survey compared to follow-up in a clinical trial, participants were asked to imagine receiving a cold questionnaire survey through the post. The majority of participants admitted that they don't usually complete such surveys. The over-riding reason for this reluctance was a perceived fear of receiving 'junk mail' following survey completion. For example:

…if you just get a normal survey, not thinking of the trial now, but if you get one of these surveys that some through the post, do you normally fill those in?

no… 'cause you end up getting tons of junk mail come in as a consequence (1076 IH p3)

A cold survey questionnaire was viewed with the suspicion of an ulterior motive. Participants felt that such instruments were a way for companies to get a 'foot in the door' and subsequently bombard them with unwanted attention. Cold surveys were also seen as time consuming and with no personal relevance which ignited an attitude of 'can't be bothered'. Of the few participants who admitted that they did usually complete household surveys, the most common reason was to alleviate boredom. Figure 9 gives a diagrammatic representation of response to general surveys.

Figure 9 ***Diagrammatic representation of response to general surveys***



These identified factors in general survey response led to the next question. Participants who were responders were asked what factors motivated them to respond to postal questionnaire follow-up in CAST.

7.3.2.2 Reasons for responding

Several themes emerged from participants explaining their reasons and motivations for responding to their questionnaire. Most participants cited more than one reason and response decisions appeared to be based on multiple factors.

- ***Personal relevance***

Personal relevance of the questionnaire was one of the most prevalent features in reasons for response. Participants felt that by responding to the questionnaires they would benefit directly from the research by receiving better treatment for any future ankle injuries. For example:

'Well, yeh and I think its really important because like I don't know if I'm gonna sprain my ankle again in the future obviously erm there will be a good chance that I will……But em I would like to think that there would be better treatment for me so yeh I think its really important that em I should send it back' (1155 MF p9)

Most participants found the questionnaire easy to complete. Some of the questions, however, were seen as irrelevant or repetitive. For example:

'Some of the "how do you feel at this present time" with your state of mind… Things I didn't think they were relevant to a foot injury erm so some of them were a bit mmm I don't know a bit difficult' (1165 SC p6)

and

'Er I mean there were some questions I mean I can't remember which which almost felt as if they were asking the same question again' (1078 MH p7)

There was no association, however, between participants who felt this way and the decision to respond or not.

- ***Agreed to take part***

Another prevalent motivator for participants to respond was the fact that they had agreed to take part in the trial and felt that they should see it through. This is reflected in comments such as:

I said I would (fill in questionnaire) so I stuck to my word' (1076 IH p2)

and

Otherwise why take part in the first place……. If you've got no intention of doing it (1155 MF p9)

This implies an understanding of the procedures of the trial and the necessity for follow-up at various time points. Participants were questioned about their understanding of the trial early on in the interviews. The majority of the participants who responded to the questionnaires demonstrated a 'good understanding' of the trial and the procedures involved. A good understanding was considered if the participant was able to verbalise the main aims of the trial. For example:

…what was your understanding of what we were trying to do with our research?

to see which support or er brace or whatever worked best (1076 IH p1)

as opposed to participants who seemed to be unaware of the trial aims:

…what was your understanding of our research when you agreed to take part in the trial?

It was just to give you information back on how I was doing with the boot and all that (1027 JO p2)

Trial understanding was an emergent sub-theme which had strong links with response. If participants were aware of the aims of the research and hence the necessity in responding to the questionnaires they were more likely to comply. Two deviant cases to this theory emerged from the data. Two responders demonstrated a poor understanding of the trial.

- ***Obliged to respond***

Some participants simply felt obliged to respond to the questionnaire partly because we had specifically requested this. For example:

…that's my next question what do you think encouraged you to fill ours in?

well 'cause you asked me to do it and er it's well worth thing to do isn't it I think (1075 GK)

Three participants felt that their response to the questionnaire was a repayment for good service or perceived 'better' treatment. For example:

What do you think motivated you to send ours (questionnaire) back so promptly?

I was actually it was actually to do with like I thought the service I got with my ankle etc was good so at the end of the day (1161 JH p3)

and

Can you explain that a bit more, what motivated you to send ours back straight away

Well a) cus I was in excruciating pain when I first injured my ankle and erm accepted the responsibilities to get involved in the trial…..And b) you were good enough to give me something that helped me to recover (1163 NM p8)

There was no consistency, however, between which particular treatments were perceived as 'good'. Furthermore, the treatment received did not emerge as a theme when participants considered their reasons for response.

- ***Altruism***

Many participants included altruistic reasons for questionnaire response. Altruistic reasons were either directed at the research institution or at others with similar injuries. For example:

…and also it would help you guys to ascertain…future needs you know and best practices… for you

good

so that's why I've done it yeah (1078 MH p6)

and

(Questionnaire was)…relevant to me but not only to me cos I am helping the cause of other people with ….. that's how I thought of it (1154 LS p6)

Altruism is also connected to understanding of the trial. If participants are aware of the ultimate purpose of the trial, the decision to help is more likely to be evoked. Most participants felt happy about being part of the trial. Some participants were probed as to whether their likelihood of response

would have changed if they had been unhappy with the trial or treatment received. These participants felt that this would have actually increased their motivation to respond so that they could air their grievances.

- ***Important project***

It emerged that some participants responded to the questionnaires because they felt that CAST was an important and worthwhile project. One participant verbalised how he thought that his motivation for responding in the trial was different to responding to a household survey. One participant explained how she responded to the questionnaire because:

…I know its part of an important project for research (1155 MF p10)

Almost all the responders cited more than one reason for their decision to respond. An indication of the frequency of the emergent response themes and the relationship between participants' multiple responses is shown in Table 38.

Table 38 ***Frequency of reported response themes***

	Emergent themes for response decisions of responders				
Participant ID	**Personal relevance**	**Agreed to take part**	**Obliged to respond**	**Altruism**	**Important project**
1063 MH		√			
1078 MH	√		√	√	
1072 MF				√	
1075 GK			√		√
1062 JB			√		
1061 AO	√				√
1076 IH		√			
1027 JO[1]	√			√	
1165 SC		√		√	
1155 MF	√	√	√		√
1157 CS	√	√		√	
1152 PL	√	√			
1154 LS	√			√	
1161 JH	√	√	√		
1163 NM		√	√		√
TOTAL	8	8	6	6	4

[1] Although participant 1027 JO was a non-responder she also gave reasons why she thought it important to respond therefore the data is included in this table.

Whilst analysing the reasons for responding, it emerged that there were patterns in actual response behaviour. If a participant was a 'keen responder' at the four week follow up point (returned the questionnaire with minimal or no prompting) they also tended to be a keen responder at the 12 week follow up point. The same pattern was also true of 'reluctant responders' (those needing a fair amount of prompting). These two groups

of participants could, therefore, be looked at separately to establish if there were any common response themes within the groups. Figure 10 and Figure 11 show the results of this analysis.

Figure 10 ***Frequency of emergent response themes for 'keen responders' (n=9)***

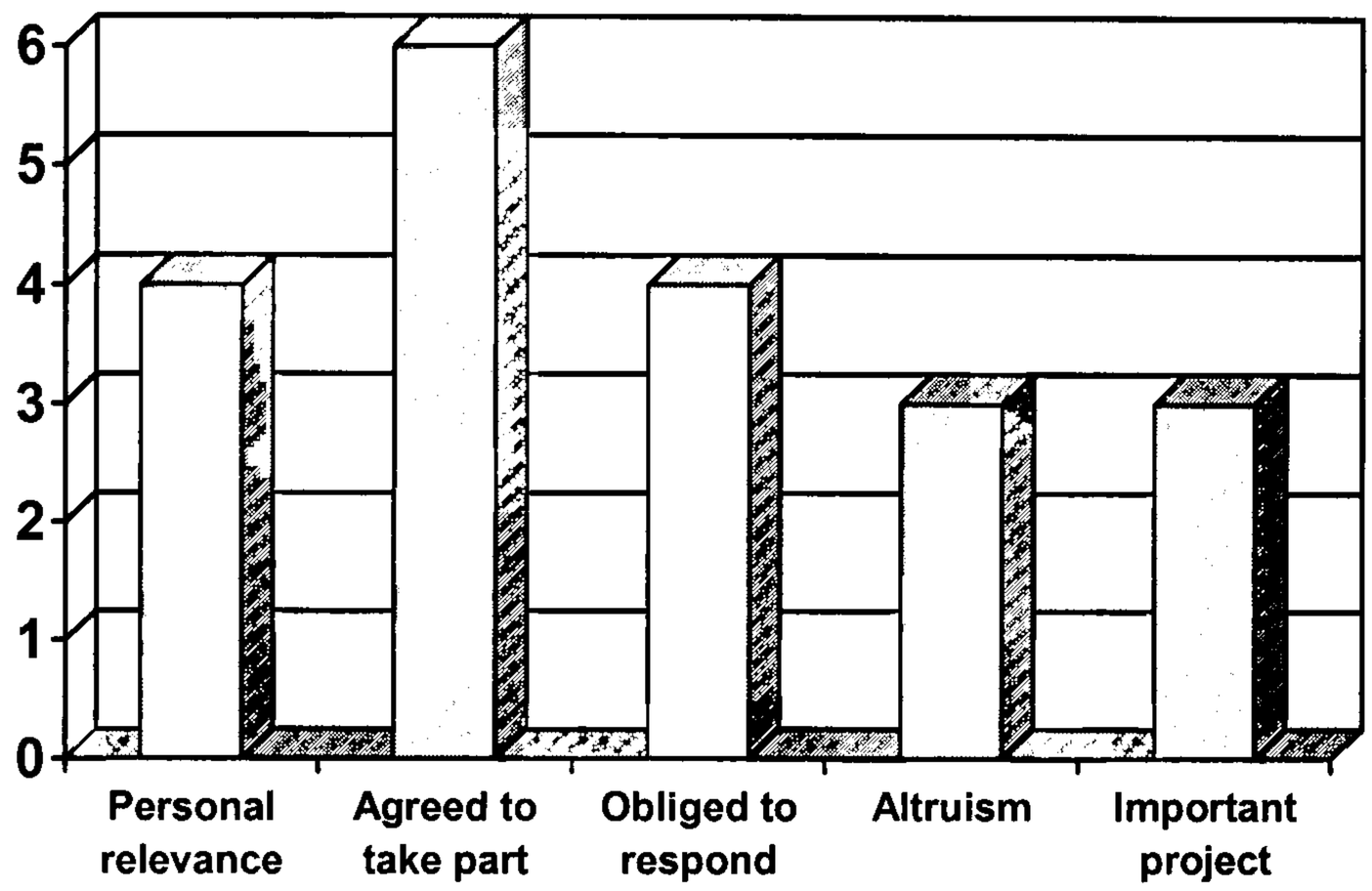
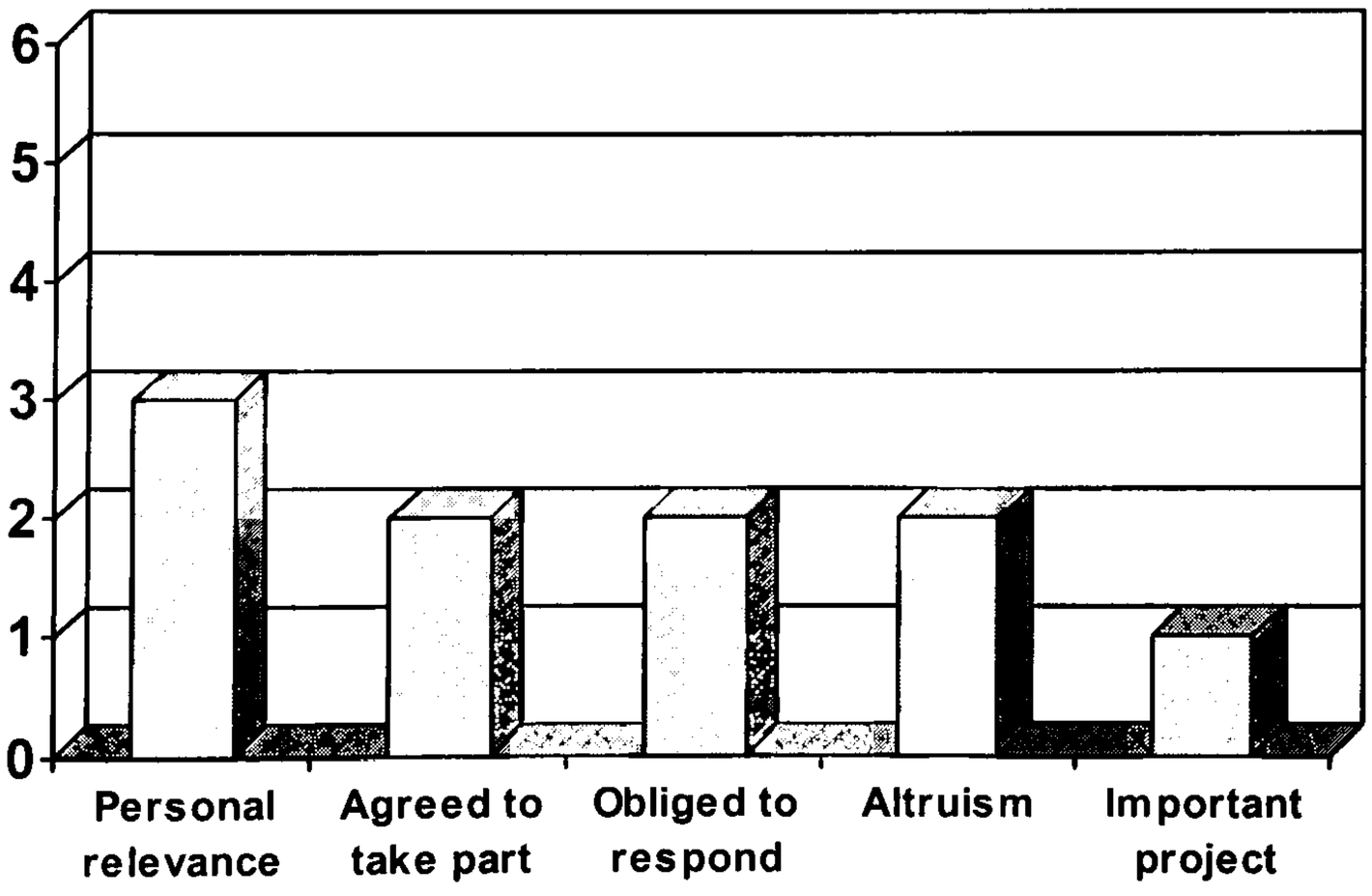


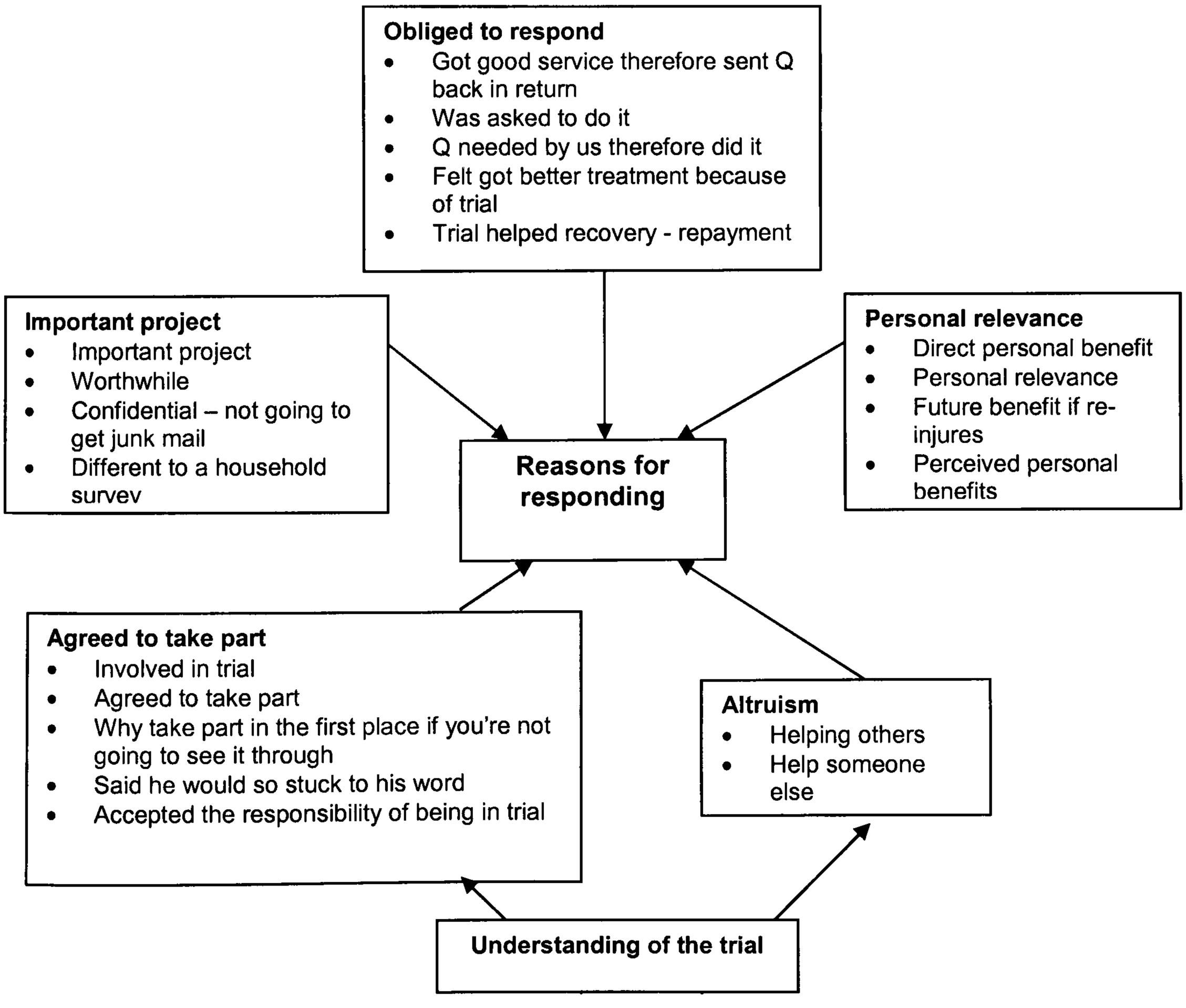
Figure 11 ***Frequency of emergent response themes for 'reluctant responders' (n=5)***



The most frequently reported response theme for 'keen responders' was that they had agreed to take part in the trial. Two thirds of all the 'keen responders' cited this as one of their reasons for responding.

A diagrammatic summary of the emergent themes for reasons for response is shown in Figure 12.

Figure 12 ***Diagrammatic summary of reasons for response to clinical trial follow-up***



7.3.2.3 Reasons for non-response

Emergent themes for reasons for non-response given by participants were grouped into four categories: 'internal aspects of the trial', 'external factors beyond control', 'personal blame' and due to 'life events'.

- ***Internal aspects of the trial***

These reasons for non-response were those given which related to internal aspects of the trial design or procedures. Of the eight non-responders, two gave internal reasons. One participant was unhappy with the treatment received as part of the trial and consequently felt reticent to comply with the rest of the trial. A second participant felt that the questionnaire was too long and time consuming and this discouraged her from responding.

- ***External factors beyond control***

External reasons were beyond the influence of the trial and included such things postal strikes. One participant received a separate questionnaire regarding another study which led to confusion and ultimately non-response.

- ***Personal 'blame'***

These reasons were those which the participant attributed to themselves. Descriptions such as 'forgetful', 'disorganised' and 'lazy' were used by

participants to describe themselves. These personal attributes were given as contributing factors to failure to respond to the questionnaires.

- ***Life events***

These were reasons such as pregnancy, exams or work commitments which participants felt were the cause of their non-response.

All the participants who had not responded to their questionnaires had all reached two follow-up time points and had therefore received two questionnaires. All had failed to respond to both questionnaires. The majority of participants who failed to respond were quite happy with the trial. They were keen to stress that their lack of response was through no fault of the trial itself and they were happy to continue to be involved. This suggests a lack of understanding of the importance of complying with follow-up to the outcome of the trial. A link between good understanding of the trial and responders has already been noted as a sub-theme. It would seem appropriate to extrapolate, therefore, that participants with a poor understanding of the trial would be non-responders. This was investigated and the results are presented in Table 39.

Table 39 ***Understanding of the trial and reason for non-response***

Participant ID	Understanding	Reason for non-response
1027 JO	Poor	Confusion with another questionnaire
3017 A	Good	Postal strike
4009 J	Questionable	Forgetful and pregnant
3020 C	Good	Unhappy with treatment
1103 SG	Good	Lazy
1120 TB	Poor	Difficulty with literacy
6003 CW	Good	Questionnaire too long
5012 AS	Questionable	Exams

Of the eight participants who had failed to respond to the questionnaires, four had either a poor or questionable understanding of the trial. Two of the non-responders, who demonstrated a good understanding, had very specific internal reasons for non-response (unhappy with treatment and questionnaire too long).

Sub-themes emerged form the data which were linked to non-response but were not identified by participants as their reasons for non-response:

- ***Treatment preference***

Treatment preferences were explored to establish whether participants who did not receive their treatment of choice were less likely to comply with follow-up. It emerged that many participants had strong views on the two treatments that were familiar to them (tubigrip and plaster). Participants had a strong preference either for or against these two treatments. These beliefs

about the perceived benefits of these two treatments were often embedded in personal experiences of the treatments. The association between treatment preference and response was investigated according to whether the participant was a 'keen responder', 'reluctant responder' or a non-responder. The results are shown in Table 40.

Table 40 ***Treatment preference, treatment received and response***

Keen responders

Participant ID	Treatment preference	Treatment received
1063 MH	Plaster or tubigrip	Tubigrip
1078 MH	**Not** Tubigrip	Plaster
1075 GK	**Not** Plaster	Aircast
1076 IH	Tubigrip/**Not** plaster	Tubigrip
1165 SC	Plaster/ **Not** tubigrip	Plaster
1155 MF	**Not** Tubigrip	Tubigrip
1152 PL	**Not** plaster or tubigrip	Aircast
1163 NM	**Not** tubigrip	Bledsoe
1154 LS	Plaster/ Not tubigrip	Bledsoe

Reluctant responders

Participant ID	Treatment preference	Treatment received
1072 MF	Plaster	Bledsoe
1062 JB	Plaster	Aircast
1061 AO	**Not** plaster	Bledsoe
1157 CS	**Not** plaster	Aircast
1161 JH	Nil	Plaster

Non-responders

Participant ID	Treatment preference	Treatment received
1027 JO	**Not** plaster	Bledsoe
3017 A	**Not** plaster	Aircast
3020 C	Aircast/**Not** plaster	Plaster
4009 J	Nil	Aircast
6003 CW	Nil	Bledsoe
1120 TB	Plaster/ **Not** tubigrip	Tubigrip
1103 SG	**Not** plaster	Plaster
5012 AS	Nil	Aircast

There was some association between treatment preference and response. All but one of the 'keen' responders either received the treatment they had a preference for, or avoided the treatment they had a preference against.

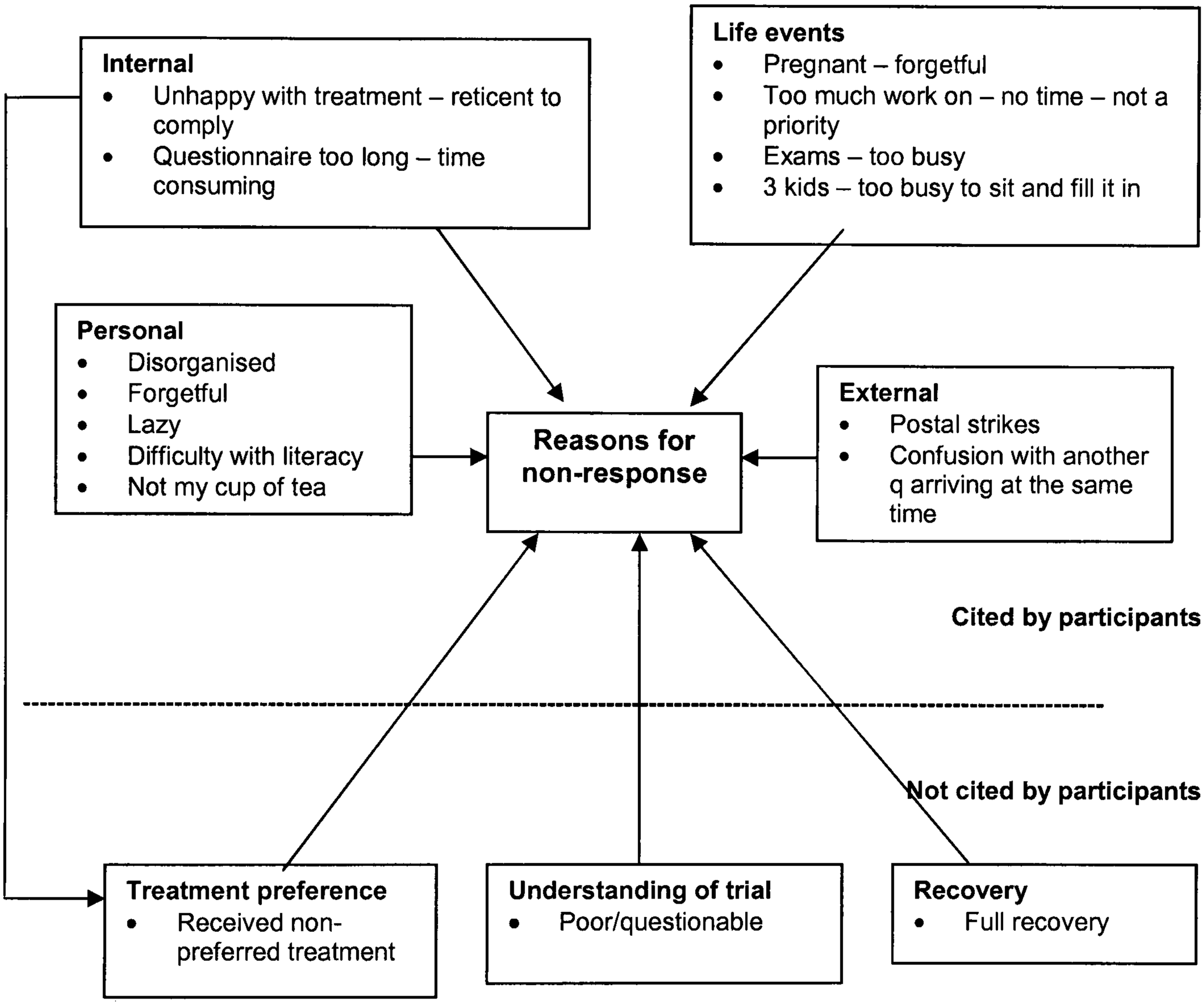
Three of the non-responders were randomised to receive the treatment which they explicitly expressed a preference against and had hence been unsatisfied with. For most this was not cited as the reason for their non-response. Only one of the non-responders cited dissatisfaction with treatment as her reason for not responding. This participant commented that she expected to receive the 'best' treatment for her ankle but didn't get it. Her belief about what was the 'best' treatment was grounded in her own prior experience of ankle injury. Many of the participants who were responders, and demonstrated a good understanding of the trial, commented that they would have complied with the trial even if the treatment received was not their preferred one.

- ***Recovery***

Almost half of the participants who did not respond to follow-up considered themselves to have made a full recovery by the second time point (ie 12 weeks post injury). The term 'full recovery' is used to describe those participants who, on questioning, used terminology such as 'back to normal' or 'perfect' to describe their ankle. The effect of recovery on response is an important consideration in acute injury trials. Full recovery is likely to occur in the majority of participants before the end of the follow-up period. Participants may then feel that their further input is unnecessary and hence fail to respond to follow-up attempts. This appears to be the case with CAST participants.

Figure 13 gives a diagrammatical summary of the emergent themes for non-response.

Figure 13 ***Diagrammatic representation of reasons for non-response in clinical trial follow-up***



Comparing both within case and between cases of the non-responders and searching for associations and links it emerged that there were fundamentally two types of non-responders. There were those non-responders who were quite happy with the trial and those who had some issue with the trial which they identified as their reason for non-response.

These emergent typologies could be described as 'happy non-responders' and 'unhappy non-responders':

Happy non-responders

These participants were keen to express that the reason for their non-response was through no fault of the trial. This group most commonly cited personal or external reasons for their non-response. For example:

Yeh, is there a particular reason why you found it difficult to do them for us or….

Do you know what…laziness I'm just gonna put it down to that

Ok and em it wasn't because you were disgruntled about part of the project

Definitely not no

Ok

No definitely not (1103 SG p5)

Unhappy non-responders

These participants all cited internal reasons as the cause of their non-response. This could be interpreted as this group using non-response as a 'punishment' to the research institute or collaborating hospital for a perceived injustice in some aspect of the conduct of the trial. For example the following quote comes from a participant who was unhappy with the treatment she was randomised to. When asked for her reasons for non-response her reply was:

And it's not been a priority and partly as well because yes I didn't I was annoyed that I felt that I was not……..getting didn't have the level of treatment….that I would have liked…… that kind of made me a bit reticent to make time although I have intended to look (3020 C p4)

7.3.2.4 Suggestions to improve response

Participants were finally asked for their suggestions to improve response to postal questionnaire follow-up in clinical trials. Ideas ranged from sending the questionnaires electronically to dropping off and collecting the questionnaire from the participant's home. Several participants suggested more phone calls and reminders. Participants were asked for their views on whether they felt an incentive would encourage them to respond. There was a mix of opinions. Some felt that an incentive wasn't expected as they had agreed to be in the trial and felt that this was unnecessary. For example:

would that (an incentive) have made a difference to you filling this in or would an incentive like that not have made any difference at all

Well at the beginning its quite clear what you was doing isn't it it ain't worthwhile paying people to give their opinions so obviously it's a voluntary thing isn't it so you don't expect anything back for it. (1154 LS p6)

Others felt that incentives would invoke some suspicion as to an ulterior motive of the research:

If they come with an incentive if they some with a pen or if they say you are gonna get into a prize draw…..If you do it does that make you more likely to return them

Them ones I don't usually do

Don't you

No

That's interesting why why is that

I dunno I always think there's suddenly gonna get junk mail all the time (1027 JO p4)

An exception to this view was amongst the non-responders. All non-responders who expressed an opinion about incentives felt that an incentive would probably have encouraged them to respond.

7.4 Discussion

This study was undertaken to identify factors which have an influence on the likelihood of a participant responding to questionnaire follow-up in a clinical trial. Theories of behaviour have been adopted by survey researchers to explain response decisions of survey participants. These theories are outlined in Chapter 2. It cannot be assumed, however, that these theories can be applied to clinical trial participants. The following section discusses the findings of this study and revisits the survey literature to evaluate any similarities or differences.

7.4.1 Response

Personal relevance of the subject and agreeing to take part in the trial featured highly in the decision making processes of participants in CAST.

Feelings of obligation and altruism were also response features as was the feeling of being involved in an important project.

- **Personal relevance**

The personal relevance of the questionnaires to the participants in this study was identified as an important feature in the response decision. The 'adjustive' function of the 'Functional' theory of response behaviour[81] helps to explain this. This function describes how people seek to maximise the rewards in their external environment. Because of the direct relevance of the questionnaire, participants felt they would ultimately benefit from complying with the request for completion. The 'Leverage-Saliency' theory of survey participation[57] is also applicable in this context. The relevance of the follow-up questionnaires would have been apparent to the participant at recruitment onto CAST (the 'saliency'). On questioning, participants appear to assign a high level of importance to this attribute in their decision to respond (the 'leverage'). Clinical trials have an advantage over general surveys in respect to the personal relevance of the questionnaire. The content of questionnaires used in clinical trials will always have relevance to the participant. Care must be taken in questionnaire design, however, to remain focussed on the necessary data required for evaluation. The apparent redundancy of some of the questions used in the CAST questionnaires was commented upon by participants. This was mainly in relation to some of the SF12 questions which asked about the participant's emotional state. This issue did not appear to affect response in this study.

There may be a point, however, at which the number of questions perceived as 'irrelevant' begins to have a negative effect on response. The SF instruments are well validated measures of quality of life and provide useful information. Rather than omit such a measure due to the perceived irrelevance by participants, it may be better to offer a better explanation of the reason for including the measure.

- **Agreed to take part**

Agreeing to take part has similarities with the 'Compliance' principle of consistency principle identified by Cialdini[85] and used by Groves and Couper in their theory of survey cooperation.[86] Once the patient had agreed to take part in the trial they felt they had a commitment to stay involved. A key feature which invoked such consistency appeared to be an understanding of the nature of the clinical trial and its procedures. This encouraged participants to see the trial through to its completion which meant responding to all questionnaires. Participants were considered to have good comprehension of the trial if they understood the randomisation procedure and long term goals of the trial. Establishing the extent of participants' understanding of the trial was not a principal consideration when developing the interview questions. This was a sub-theme which emerged during coding of the transcripts. The data surrounding trial understanding, therefore, lacks depth. Enough data were gathered, however, to suggest an association between trial understanding and consistency of involvement. This association requires further exploration

and should only be seen as an explanatory hypothesis at this stage. However, ensuring that participants fully understand the trial procedures is an important consideration and is, in part, the responsibility of the recruiting researcher. There is some evidence that patients' understanding of clinical trials is generally poor.[211, 212] A thorough explanation of the trial should be given to the patient at the recruitment stage. This could include an explanation of why complying with follow-up is essential to the success of the trial. It is common in clinical trials for staff responsible for recruitment to change frequently. Ideally, formal quality assessment of the recruitment approach should be undertaken. In CAST, new members of the recruitment team were individually trained and observed, if possible, during recruitment of their first patients. Augmenting written trial information with a verbal accompaniment has been suggested by other authors as an important consideration in the investigation of ways to maximise participants' understanding of clinical trials.[213, 214]

- **Obliged to respond**

An obligation to respond has parallels with the 'Compliance' principle of reciprocation. This principle suggests that 'one should be more willing to comply with a request to the extent that compliance constitutes the repayment of a perceived gift, favour or concession'.[86] The theories of 'Dissonance'[76] and the 'ego-defensive' function of the 'Functional Theory'[81] could also explain the urge for participants to respond to reduce cognitive conflict. Patients felt that as they had received a good service or

treatment as part of the trial they would reciprocate by returning their questionnaires. Others felt obliged to respond simply because we had asked for their help. This is a difficult principle to incorporate into a clinical trial to improve response. Inevitably, there will be people who are not happy with the service or treatment received as part of the trial. The effect that this has on response is discussed later in this chapter.

- **Altruism**

Altruism has similarities with the 'value-expressive' function of the 'Functional Theory'[81] which has been identified in the survey literature as a possible feature in participants' response behaviour.[50] This function explains how individuals give positive expression to their central values and to the type of person they conceive themselves to be. A link to trial understanding was also evident here. The altruistic notion of helping both the research institution and others with similar injuries was only possible if the participant understood the reasons for the trial. Appealing to the altruistic tendencies of participants can have significant effects on response. By adding the words 'it would really help us out' to the end of their request for survey participation, Mowen and Cialdini[215] obtained a 19% increase in uptake. This study, however, was in the field of market research. In health care research there are strict ethical codes of conduct to adhere to. It is important that the participants in a clinical trial are not subjected to coercive tactics to encourage them to comply. The Central Office for Research Ethics Committees (www.corec.org.uk) provides 'Guidelines for Researchers'

which includes suggested wording of patient information sheets. COREC stipulate that 'the tone should be invitational and not coercive or overly persuasive'.[92] Appeals for help need to be worded carefully in order not to be considered inappropriate in this context.

- **Important project**

Several participants mentioned that they felt that CAST was an important project and this encouraged them to respond. This has similarities with the 'Compliance' principle of authority. Appeals can be made to the participant's perceived sense of importance of the project at the recruitment stage of a clinical trial. The fact that the trial is being run by a reputable and legitimate institution should be emphasised in the trial literature and explanation. In CAST all written material was presented to the participants on either hospital or university headed paper. In addition, recruiting researchers (where possible) wore a badge identifying them as members of the institution conducting the research.

In a typical clinical trial using postal questionnaires as a method of follow-up, the only time the participant has face-to-face contact with the researcher is at the initial recruitment stage. The results of this study show that this stage is crucial to invoking the compliance and altruistic principles which appear to affect response. Recruitment clinics are often busy and under strict time constraints. The priority is to recruit and randomise the participant and, if appropriate, administer the treatment. Adequate time, however,

should be allocated to giving the participant a full and detailed explanation of the aims of the trial and the importance of follow-up. A systematic review of methods of improving response to postal questionnaires in health care research was conducted prior to this qualitative study and is presented in chapter 5. The results showed that reminder strategies either by post or telephone had the most significant affect on improving response. This is also an opportunity to appeal to the compliance principles of participants. Reminders could include reference to the principles of consistency and authority as well as an appeal for the participants continued 'help' with the research.

7.4.2 Non-response

Similar to previous research in the survey literature,[71] when questioned about their reasons for non-response participants had a range of explanations. A common theme, however, was that all the non-responders had reached the second follow-up time point and all had failed to respond to both questionnaires. In this sample, therefore, this suggests that 'once a non-responder, always a non-responder'. This is especially true for those participants who gave reasons for non-response such as impending exams and postal strikes. The two questionnaires would have been separated by at least 8 weeks. Presumably the impending exam would have taken place and it is unlikely that a postal strike would have lasted so long. The postal strike excuse also loses credibility when the participant comments that

'…bills, they seem to come through all right'. This suggests a selective postal strike which is implausible.

Only a minority of non-responders were unhappy with the trial in some respect. The majority were quite happy with the trial and their continued participation. The framework analysis approach allows for the description of such typologies as 'happy non-responders' and unhappy non-responders'. Happy non-responders may be receptive to strategies to convert them into a responder. More than half the happy non-responders demonstrated a lack of understanding of the trial. These participants could be targeted and contacted after their first episode of non-response. The purpose and procedures of the trial could be re-emphasised. The compliance principle of consistency, which is apparent in responders, could then be invoked in this group. Requests for 'help' could also be included in follow-up contact of non-responders. According to the 'Functional Theory' of the study of attitudes,[81] this will appeal to the 'value-expressive' attitude function or altruism of participants.

Studies of why people 'help' have shown that feelings of anger can have a negative effect on helping behaviour.[94] This is reflected in this study with the 'unhappy non-responders'. These participants expressed some form of resentment of the trial or the treatment received as the reason for their non-response. Unhappy non-responders are also an example of the negative effect of the compliance principle of authority.[87] This group may feel they have been poorly treated by the hospital or research institution and hence

refuse to cooperate. It is inevitable in a large clinical trial that some participants are going to be dissatisfied with some aspect of the trial. This study shows that this leads to a reluctance to continue to participate. Some of these unhappy non-responders may be persuaded to respond if they feel that this is a means for them to 'air their grievances'. Unhappy non-responders should be assured that their comments are welcomed and will be taken seriously. Such feedback can be used to improve future trial design.

It is important to emphasise to clinical trial participants that their response to follow-up questionnaires is vital at every time point. This is true even if the patient considers them self to be fully recovered. Participants may feel that answering questions about symptoms that are no longer present is unnecessary. This study suggests that full recovery has a negative effect on response. This is therefore an important point to consider in the initial trial explanation and subsequent participant contact.

The effect of participants in clinical trials having a preference for a particular intervention has been widely studied.[216-218] In this study, those participants who received their preferred treatment or avoided their non-preferred treatment were more likely to be 'keen responders'. One participant stated her reason for non-response as due to not receiving her preferred treatment but as this was an isolated opinion it cannot be seen as representative. King et al[218] found that although participant preferences

may affect perceptions of the interventions and satisfaction they appear not to affect further participation in a clinical trial.

The 'Reactance' theory first introduced by Brehm in 1965[72] has been applied to survey research as an explanatory model for non-response.[73, 74] This theory suggests that if an individual feels that his freedom of behavioural choice is eliminated, he will experience 'reactance' to re-establish his freedom to choose how to behave. This is applied to survey response by hypothesising that reactance would be manifested by an individual refusing to respond to a questionnaire if they feel that they have no choice in the matter.[73]

Reactance did not appear as a theme in clinical trial participants reasons for non-response to postal follow-up. The reason for this may be that, in a clinical trial, participants have already 'chosen' to be involved by agreeing to initial recruitment. This may, therefore, negate any reactance being invoked in requests for follow-up questionnaire completion. The informed consent process ensures that potential participants are aware of their freedom of choice as to whether to participate in the trial. If reactance is manifest in a clinical trial it is likely that it would be at the recruitment stage rather than at follow-up stages.

7.4.3 Suggestions to improve response

Incentives are used frequently in survey research and appear to be effective in improving response.[12, 15, 36] Dissonance theory has often been given as an explanation for this.[77, 78] The systematic review conducted prior to

this study (chapter 5) found no evidence, however, that incentives are effective in health care research. The opinions of most participants in this study appear to verify this. Some participants felt incentives to be unnecessary in a clinical trial. Others viewed incentives with the suspicion of an ulterior motive of the research. This implies that incentives would effectively 'demote' the questionnaire into the realms of a survey with the associated suspicion as to the motives of the research. The 'Leverage-Saliency' theory of survey non-response may offer a more theoretical insight into the apparent ineffectiveness of incentives in health care research. This theory argues that in the absence of intrinsic motives for survey participation (such as self-interest in the topic), incentives are extrinsic substitutes.[57] In clinical trial follow-up it could be argued that participants are motivated to respond to follow-up due to such things as the personal relevance of the questionnaire and feelings of altruism as described in section 7.4.1. Lower positive 'leverage' is therefore assigned to incentives rendering them ineffective in this context.

The views on incentives outlined above were given by responders to the CAST follow-up questionnaires. Non-responders, however, had different views on the use of incentives. They admitted that some form of incentive may have encouraged them to respond. There may be a case, therefore, for selective incentives targeted at first time non-responders or those displaying characteristics which put them at risk of non-responding. This was investigated in a recent health care study which found this method

increased response rate amongst initial non-responders by 11.7%.[219] The selective use of incentives in groups at risk of non-responding has also been suggested in the survey literature.[27, 170] An important addendum is that those groups at risk of non-response first need to be identified before differential effort can be assigned to encourage response.[27] This is the subject of chapter 8 which presents an analysis of the characteristics of non-responders in CAST.

7.5 Conclusions

This study found some common themes which emerged in the reasons participants in a clinical trial gave for responding or not to postal questionnaire follow-up. These response reasons were compared to reasons identified in the literature on survey response. This study found that aspects of the behavioural theory of 'Compliance', with its associated principles and Katz's 'Functional' theory are, in part, applicable to the clinical trial setting. The compliance principles of scarcity, liking and social validation are absent form this study of clinical trial participants. The principles of consistency, authority and reciprocation, however, do feature in participants' reasons for responding to postal questionnaire follow-up. Consistency appears to be closely related to understanding although this is a concept which requires further exploration. It is therefore important to ensure that adequate time is taken to fully explain the trial to participants at recruitment. This includes explanations of the follow-up procedure and why

response is so important. Trial information, both in written and verbal form, should be worded to invoke compliance principles. The ethics surrounding 'coercive' language, however, need to be considered and accounted for. Dissonance theory[76] was also evident in participants who felt they had received good service or treatment as part of the trial. This encouraged them to return their questionnaire as a form of repayment for this.

Personal relevance of the questionnaire topic to the participant is also an important issue. This study shows that this is one of the most prevalent influences on response to follow-up in a clinical trial. One way of increasing the relevance of the questionnaire to participants could be to reduce the number of seemingly 'redundant' questions. However, rather than reducing the amount of information collected it may be better practice to offer a more thorough explanation of why the data is required. The 'Leverage-Saliency' theory of survey response[57] is a relatively new concept in survey research which offers, perhaps, the most relevant theoretical basis for encouraging response to postal questionnaires.

Steps can be taken to try to convert non-responders into responders. 'Happy non-responders' may be receptive to compliance principles if their understanding of the trial is improved. 'Unhappy non-responders' may welcome the opportunity to air their grievances by responding. The importance of responding even if full recovery has been reached should be emphasised to all participants. The selective use of incentives on non-

responders may be an option but the ethical considerations of this need to be addressed.

This qualitative study is the first to specifically investigate the response decisions of clinical trial participants from their own perspective. It has revealed that there are many similarities in the way survey participants and clinical trial participants make their response decisions. It has also revealed, however, that there are also many differences. Clinical trial participants comprise a very specific group of individuals which is unique to the subject matter of each trial. These specific circumstances need to be taken into consideration when attempting to maximise response to follow-up.

7.6 Rationale for choice of analytic framework

The methodological options available for analysing the data generated in this study were carefully considered. Three approaches were seen as potentially appropriate. These were: thematic content analysis, grounded theory and framework analysis. These approaches, and the reasons why framework analysis was eventually chosen as the method of data analysis, are outlined below:

- **Thematic content analysis**

Thematic content analysis is considered to be the most basic type of qualitative analysis.[207] This approach is favoured by quantitative researchers who find themselves in a position of having to analyse written material.[206] Precise categories are established and compared and often

the number of instances falling into each category is counted. This method is often criticised for losing the 'richness' of the qualitative data. The process of categorisation can 'disembody' the person who produced the data from the interactive nature of the interview.[220]

For this study, more depth of analysis was required than a simple thematic content analysis could deliver. Rather than just establish the reasons for response and non-response, this study aimed to explore relationships and links between themes. This would then give a richer understanding of the response behaviour of participants in a clinical trial.

- **Framework analysis**

Framework analysis has many similarities with methods of thematic content analysis. Although framework analysis uses a similar method of categorisation, the 'mapping and interpretation' of the data is an absorbing and detailed process. It is this aspect of framework analysis which moves it beyond even a sophisticated thematic content analysis.[207] Through immersion in the data, concepts and connections are formed by means of intuition and imagination. The explicit way in which the methodology of framework analysis is presented is also seen as an advantage over thematic content analysis.

A key feature of framework analysis is that it allows for the inclusion of *a priori* as well as emergent concepts. This aspect of framework analysis is open to criticism. Grounded qualitative researchers argue that theories and concepts should emerge from the setting under study rather than being

imposed from the outside.[221] There is a historical wariness amongst qualitative researchers of imposing prior and possibly inaccurate frames of reference on the people they study.[222] More recently, however, the concept and theory generation of earlier work is being used as a catalyst for further research. It is important, though, that these concepts and theories should be held lightly and be subject to change and possible rejection as the study progresses.[223] In this study, theories of response were identified in the survey literature but the data generated form the interviews was analysed independently before comparisons were made with these existing response theories.

- **Grounded Theory**

The nature of the data and the research questions made the grounded theory approach, as described by Glaser and Strauss,[221] a possible option as the analytical framework. Grounded theory allows social theory to be generated from the data through a process of rigorous and structured analysis. A key element of grounded theory is that the selection of subjects, data collection and data analysis are concurrent, ongoing and interrelated.[203] With the present study, theories of response behaviour were identified in the literature concerning survey response. It was therefore felt unnecessary to generate a new theory using a methodology such as the grounded theory approach. The aim of this study was to explore response behaviour in participants in a clinical trial. It was proposed that comparisons could then be made to existing theories of survey response behaviour. Any

new concepts specific to the clinical trial environment generated from the data would also be explored.

During the search for the most appropriate analytical method to use in this study, published articles which utilised 'grounded theory' were referred to. It became apparent that often researchers claiming to have used grounded theory had actually done no more that a superficial content analysis. This is a view shared by Green and Thorogood who note that 'grounded theory' is perhaps one of the most abused phrases in the qualitative health literature'.[207] A study using a true grounded theory approach is a detailed and time consuming process requiring a thorough understanding of the procedures involved. Rather than contributing to the misuse of grounded theory it was felt that framework analysis was more appropriate for the timescale of this study. For this study, grounded theory had no major advantage over framework analysis. Had a grounded theory approach been used, however, it is anticipated that similar conclusions would have been drawn from the data.

Framework analysis was, therefore, chosen as the analytical method for this study after careful consideration of the various options available. There were several reasons for this choice which are outlined below:

- Framework analysis allowed a deeper and more detailed analysis of the data than simple thematic content analysis.

- The subject of the interviews and nature of the data collected lacked the depth required to rationalise the use of a grounded theory approach.
- The explicit nature of the methodology of framework analysis was seen as advantageous in making the results of the study more robust.
- Framework analysis is suited to research asking specific questions with limited timescales.[203] These factors were relevant to this study.
- The pragmatic theoretical standpoint of framework analysis, drawing on both positivism and interpretivism, is consistent with the theoretical underpinnings of this thesis.

A common criticism of qualitative research, regardless of the analytical approach taken, is that the results are anecdotal.[203] Reliability and validity are, however, important issues in qualitative research. These terms are interpreted differently in qualitative research compared to quantitative research. Steps can be taken, however, to ensure that qualitative data analysis is rigorous. The logic of generalisability of the findings of qualitative research is also somewhat different form quantitative analysis. The two issues of ensuring rigour and generalisability of the results of this study are discussed in the following sections. The way in which the choice of framework analysis as the analytic approach relates to these concepts is also considered.

7.7 Ensuring rigour

The credibility of qualitative analysis is improved if a number of general principles of 'good practice' are adhered to.[207] Table 41 summarises the criteria that characterises rigorous analysis. These criteria are then each discussed in more depth.

Table 41 ***Features of rigorous qualitative analysis[207]***

Criteria	*Possible methods*
Transparent	Provide a clear account of procedures used An 'audit trail' that others could follow
Maximises validity	Analysis of deviant cases Simple frequency counts of key themes 'Member validation' Including enough context for the reader to judge interpretation
Maximises reliability	Analysis of whole data set Using more than one analyst/coder
Comparative	Compares data between and within the data set Compares findings to other studies
Reflexivity	Accounts for the role of the researcher in the research

7.7.1 Transparency

This concept relates to the clarity and openness of the methods used to analyse qualitative data. This is one of the key elements of framework

analysis and has been faithfully adhered to throughout this study. The method of sample selection is clearly described. The original transcripts were indexed following development of the thematic framework. An extensive chart was developed to allow the whole dataset to be easily read. The ensuing results and thought processes behind their interpretation are described in detail in the text and by diagrammatic representation. Every attempt has been made to provide an honest and clear account of the analytic process.

7.7.2 Validity

The traditional understanding of 'validity' refers to the 'truth' or 'correctness' of the interpretation of a piece of research.[224] In qualitative research 'truths' are immersed in the social world and the positivist idea of one fixed truth is rejected.[207] The problems of applying the concept of validity to qualitative research are well recognised. This is not an excuse, however, to neglect the issue. Qualitative researchers should endeavour to justify why their analysis should be considered a legitimate and credible one.[207] A number of different ways have been proposed to increase faith in the validity of qualitative data:

7.7.2.1 Deviant case analysis

Disconfirming evidence or 'outliers' should be actively sought and accounted for. Such cases can be used as an important resource in aiding understanding or theory development.[224] In this study this concept was

used throughout the analysis. Where deviant cases were identified this was highlighted in the text and possible explanations suggested.

7.7.2.2 Frequency counts

There is some debate within qualitative research as to the appropriateness of counting the frequency with which themes are represented in the data. The size of the sample used in qualitative research and the way the sample is selected mean that any statements about the prevalence or distribution of a theme are only applicable to the study sample itself.[210] There is a danger when presenting numbers that erroneous statistical inferences are drawn to the wider population. Some authors stress that 'qualitative research should be explaining patterns of recurrence, not simply stating that they exist'.[210] Green and Thorogood[207], however, suggest that simple frequency counts can increase the reader's faith in the validity of the interpretations of the researcher. They argue that counts are useful in giving some perspective on how common various views or experiences were and this can defend against anecdotalism.

Frequency counts have been used sparingly in this study. This was to indicate how reasons for response related to *a priori* concepts found in the literature of survey response.

7.7.2.3 Member validation

This involves taking the findings of the research back to the research participants to see if they agree with the interpretation. There are, again,

differing views as to the usefulness of this method as a check of validity. Member checking assumes that the research participant will analyse their own accounts in the same way as the researcher. Unless the analysis has taken the form of merely *reporting* the participant's account of the world there is questionable value in this exercise.[207]

Member validation was not used in this study. The decision not to carry out member validation was made after assessing both the usefulness of the technique and the time constraints of the study.

7.7.2.4 Including enough context

Validity can be enhanced by providing enough context for the reader to judge interpretations of the data. The nature of the framework analysis approach means that the context of the research is clearly communicated. This has been adhered to in this study. Efforts were made to provide a clear description of both the research methods and the findings. This included details of the research setting and background to the study. This concept is related to the concept of 'transparency' outlined previously.

7.7.3 Reliability

Reliability concerns the 'repeatability' of research findings and whether or not they would be repeated if another study was undertaken using the same methods.[224] Like validity, there are a number of issues surrounding the application of reliability in the context of qualitative research. An extreme view is held by advocates of the 'constructivist' school who argue that

replication is an artificial goal to pursue since there is no single reality to be captured in the first place.[225] The more generally held mainstream view, however, is that reliability should not be seen as an alien concept in qualitative research.[224] Attention to reliability ensures that whatever interpretation is followed through is credible and the codes and themes are identifiable.[207] The nature of framework analysis allows visibility at all stages of the research process. This was adhered to in this study. This allows readers to have access to the whole data set ranging from how the sample was selected to how the interpretations were derived. This study also used a second person to repeat the coding process which enhances reliability further.

7.7.4 Comparative

Comparison is what drives qualitative analysis.[207] Comparing data both between and within cases allows key themes to emerge, deviant cases to be identified and the contextual meanings of accounts to be explored. Theoretical analysis also develops as a result of comparison. The 'charting' stage of framework analysis used in this study involved rearranging the data according to thematic content either case by case or by theme. This facilitated comparison both between and within cases and allows the reader to see the whole dataset. The comparison of findings with other findings from the field is also good practice in enhancing the rigour of a study. This study frequently makes reference and comparisons to previous work in survey research.

7.7.5 Reflexivity

Reflexivity is the recognition that the researcher is part of the process of producing the data and their meanings. A reflexive awareness of the research process can be demonstrated by some 'good practice' approaches.[207] This increases the rigour of the analysis:

7.7.5.1 Methodological openness

This entails being explicit about the steps taken in the data production and analysis and why alternative approaches were not pursued. In this study the reasons for choosing framework analysis over other approaches has been discussed at some length.

7.7.5.2 Theoretical openness

The theoretical starting points and the way in which they shaped the study should be accounted for. The theoretical stand point of framework analysis and how this relates to the philosophical underpinnings of this thesis has been described.

7.7.5.3 Awareness of the social setting itself

In interviews, such as those used in this study, the 'data' are largely the results of interactions between the researcher and the participant. There needs to be a constant awareness of the ways in which the data result from these interactions. This was recognised and elaborated upon in the 'Data Collection' section.

7.7.5.4 Awareness of the wider social context

This is recognised throughout this study by relating the emergent themes to prior issues in survey research. The importance of maximising response rates in clinical trials is recognised.

7.8 Generalisability

7.8.1 Definitions of generalisability in qualitative research

Generalisability refers to the applicability the findings of a study to the wider population or to different contexts. In quantitative research, random sampling generates a sample which is likely to be statistically representative of the larger population of interest.[207] There is much debate among authors as to whether the findings of qualitative research are capable of supporting wider inference. This is due to the non-random sampling usually associated with qualitative methods and also whether the results have any 'reality' beyond the context in which they were derived.[224] Furthermore, the concept of reality is strongly influenced by the epistemological and ontological views of the researchers.[226] Some authors, therefore, argue that generalisability is an irrelevant concept in qualitative research. Qualitative work is seen as providing rich description rather than typical accounts. A more recent view, however, is that the issue of generalisability of qualitative work in health research does need to be addressed.[207]

Without tackling this issue researchers will be unable to attach any usefulness to their findings.

Applying the concept of generalisability to qualitative work is not simply a matter of adopting the methods used in quantitative research. The kind of relationship that the study findings have with other populations and settings needs to be considered. Lewis and Ritchie[224] clarify the way in which generalisation is applied to qualitative research by suggesting three concepts. These concepts are outlined below and the generalisability of this study is considered.

7.8.1.1 Representational generalisation

This concept refers to whether the results of a study can be generalised to the parent population from which the study is drawn. Qualitative research usually involves study samples which are small and purposefully selected. This is often used as an argument to undermine the scope for representational generalisation in qualitative research.[227] Lewis and Ritchie, however, recognise that qualitative research cannot be generalised on a statistical basis. Wider inference is not drawn from the *prevalence* of views or experiences in qualitative research. It is the *range* of views or experiences and the factors that influence them that can be inferred to the researched population. Generalisation takes place at the level of categories, concepts and explanation. Assessing representational generalisation in qualitative research relies on the accuracy of data collection and interpretation. It is also important that the sample contains the diversity of

dimensions that are central to the area of study (rather than a statistical match).[224]

This qualitative study was conducted within the context of a randomised controlled trial of severe ankle sprains (CAST). The aim of the study was to gain an insight into the response issues relevant to clinical trial participants. The key question is whether the same themes and categories which emerged from the data are believed to exist in the rest of the CAST participants. The purposive sample was selected to represent the diversity of the CAST participants in several dimensions which were seen as important factors in responding to postal questionnaires. The original interviews were transcribed verbatim and the ensuing analysis was conducted with rigour. As the analysis developed, conceptual themes emerged and the relevance and validity of these could be questioned. The analytic routes used, however, are described both in the text and by diagrammatical representation. This allows readers to asses the evidence. It is likely that the broad categories of reasons for response and non-response would be the same or similar if another sample of the CAST participants were interviewed. It is felt, therefore, that this study allows representational generalisation to the rest of the CAST participants.

7.8.1.2 Inferential generalisation

This concept relates to whether the findings of a study can be generalised to other settings beyond the sampled one. Other authors have used the terms 'transferability' or 'empirical generalisation' to describe this concept. This is

an important and pragmatic issue for readers of qualitative research who will want to know if the findings can be applied to their own practice.[207] To allow readers to assess the meanings attached to the original observations, and the environments in which they occurred, sufficient detail needs to be provided by the researcher. This involves the researcher providing thick description of the research context and findings. This allows others to assess the transferability to another setting.[224]

The context of this study was a clinical trial. Inferential representation would suggest that the findings of this study may be applicable to other clinical trials. Clinical trials vary greatly in their size, methodology, target sample and methods of assessing outcomes. The reasons for non-response given by participants in an acute injury trial (such as CAST) may be very different to those given, for example, by young patients in a cancer trial. Thorough descriptions of the research context and findings of this study are, however, provided. This allows the reader to evaluate the inferential generalisability for themselves.

7.8.1.3 Theoretical generalisation

This concept draws theoretical principles from the findings of a study for more general application. This concept is open to different interpretations related to the ontological base from which the qualitative study originates.[224] Lewis and Ritchie believe that *'qualitative research studies can contribute to social theories where they have something to tell us about*

the underlying social processes and structures that form part of the context of, and the explanation for, individual behaviours or beliefs'.[224]

The extent to which the data from a study supports existing theories can be established by comparing how well different cases 'fit' within an established theory.

Theoretical generalisability is, perhaps, the most relevant form of generalisability for this study. The issues surrounding response to postal questionnaire follow-up in a clinical trial were compared with existing theories in survey research. New emergent themes were analysed in depth and suggestions for improving response were conceptualised. The results of the analysis offer some interesting concepts in response to postal questionnaires. Although it is recognised that every clinical trial is different, theoretical generalisation of the findings of this study is seen as a legitimate hypothesis. It is also recognised that this claim of theoretical generalisability is equally open to challenge by other researchers.

7.9 Chapter Summary

This chapter rationalised the need for, and presented the results of, a qualitative study of clinical trial participants. Response issues of clinical trial participants, as verbalised in their own words, have been identified. These issues have been compared to existing theories of response presented in the literature on survey research. The results of this chapter have been interpreted and discussed. Recommendations as to how the findings could

be used to improve response to postal questionnaires have been offered. This chapter goes some way to answering the question '*why* do people choose to either respond or not respond to postal questionnaire follow-up in a clinical trial. The next chapter examines a different perspective which survey researchers have widely investigated: '*who* responds to postal questionnaires'.

8 Chapter 8

Characteristics of responders and non-responders

Many investigations have been conducted in the field of survey research which attempt to establish common personal characteristics which can be used to predict whether or not an individual will respond to a survey. Such information can be used to suggest appeals and procedures which may reduce the number of non-responders. Additionally, studies of non-responders are important to determine what biases exist in the survey.[99] A methodological problem that survey researchers face when investigating socio-demographic determinants of response is that little is known about the group who have been sent the questionnaire. This is very different to the clinical trial setting. Participants are usually recruited onto a trial in a face to face clinical setting and at this point detailed background information is collected. Trials should always have such information collected to assess comparability of the groups during subsequent data analyses. Much information is therefore available to analyse response behaviour from a socio-demographic perspective. This gives clinical trial researchers wishing to investigate this aspect of response behaviour a big advantage over survey researchers. Despite this, little work has been carried out looking specifically at the socio-demographic characteristics of clinical trial participants. There is a limited body of research looking at response

characteristics in health related surveys which may offer a more relevant discussion than the general survey research.

This chapter begins with a note on the importance of the distinction between 'refusals' and 'non-contacts' in response to postal questionnaires. An outline is then given of the general survey literature regarding the socio-demographic correlates of response behaviour together with a summary of the available literature in this area specific to health related surveys. This chapter will then conclude with an analysis of the socio-demographic characteristics of the CAST participants.

8.1 The difference between 'refusals' and 'non-contacts'

The terms 'refusals' and 'non-contacts' are used widely in survey research to distinguish between different types of non-responders. This is largely in relation to non-response in interview surveys. Refusals are those people who decline to cooperate with the interview process whilst non-contacts are those in the sample who cannot be contacted for interview. Both groups add to the pool of non-responders. That non-contacts have a different socio-demographic profile to refusals has received much emphasis in the survey literature. For example, the young middle class may be well disposed towards surveys but they are also the hardest to reach either by telephone or at the door.[71] Although there is an increased awareness of the difference between non-contacts and refusals in survey research there is an inherent neglect in portraying this distinction.[228, 229]

The interpretation of refusals and non-contacts in response to postal questionnaires has received much less attention in the survey literature. Non-response to a mail survey is evidenced only by non-return of the questionnaire. It is difficult to establish whether this non-response is due to a refusal to cooperate or because the participant for some reason did not receive the questionnaire.[230] No literature can be found which specifically categorises the types of non-responders found in clinical trials. One of the administrative tasks in CAST was to keep a detailed log of the amount of follow-up contact required by participants to return their questionnaires. How the CAST participants responded to this contact was also recorded. An analysis of these data is presented in section 8.3.

8.2 Socio-demographic characteristics of responders and non-responders to general and health related surveys

This section is a discussion (rather than a systematic review) of issues around questionnaire response in relation to socio-demographic characteristics of survey participants.

Socio-demographic variables which have received attention in the survey literature regarding response to postal questionnaires include socio-economic status, age, sex, marital status, work status and urban versus rural dwelling. These factors are not thought to be directly *causal* to the

decision to cooperate with the survey. Instead some believe that they are associated with a set of psychological predispositions that affect the response decision.[87] This section gives a summary of the effect of these variables in both the general survey literature and the health related survey literature.

8.2.1 Socio-economic status

Socio-economic status (SES) is usually measured by occupation, income or education.[71] Historically there has been a concern about non-response bias in postal surveys because it was considered that literacy and its social correlates would introduce important biases.[231] One of the earliest reviews in this area found that non-responders to postal surveys had lower mean educational levels than responders.[232] Two, more recent, reviews reaffirm this.[14, 233] The latter of these reviews found that the highest probability of response was among the highly educated in all 26 studies included in the review. Although SES correlates positively for response decisions after contact, the effect does not, however, extend to accessibility. Accessibility refers to how easy it is to initially contact an individual to request survey participation. Several studies have shown that the higher SES groups are inaccessible to initial contact.[71, 234]

When looking at the literature specific to health related surveys, the same proxies for SES as those identified in survey research have been studied (i.e. occupation, income or education).The health related survey literature suggests a similar effect of SES on response to that found in general

surveys. Hoffman et al (1998) found response rates to a health related questionnaire survey to be higher amongst persons with more than 13 years of education.[235] Two separate Scandinavian studies [101, 236] found that people with lower educational levels required more intensive follow-up efforts to encourage response. Using occupation as an indicator of SES, Cartwright and Windsor (in a study of new mothers) found lower response rates for those whose partners were in manual occupations compared to non-manual.[177] Kotaniemi et al (2002) found that non-response was significantly higher amongst self-employed men.[237] Their explanation for this is that 'young businessmen' or 'entrepreneurs' may be too busy or not happy answering letters from the health care system. The authors appear to extrapolate that self employed men are likely to have a higher SES by describing them as 'businessmen' and 'entrepreneurs'. If this is the case, then SES had a negative effect on response in this study.

In summary, the evidence about the effects of SES on response is variable; different effects are seen in different studies with plausible explanations.

8.2.2 Age

The probability of response to face-to-face interviews and telephone interviews has often been shown to correlate negatively with age.[119, 120] There is much more ambiguity in the literature regarding the effect of age on response to postal questionnaires. Some studies demonstrate a negative effect of increasing age on response[238] whilst others show a positive effect of increasing age on response.[239] One possible explanation for this

inconsistency is that the effect of age may be confounded by the effect of SES. Older people tend to have had less formal schooling than the young due to historical upgrading in educational systems.[71] This may lead to lower literacy levels and therefore less confidence in completing a questionnaire. Studies which demonstrate better response rates in older people may be due to this cohort having more free time to sit and complete a questionnaire if they have retired from work. It is also plausible for age to have a more complex shaped relationship with response; response might be highest in middle aged people and lower in the young and old.

The health care literature is similar to the general survey literature in that the effect of age on postal questionnaire response is equivocal. Some studies have found no effect of age on response.[89, 231] None of these studies, however, performed a multivariate analysis on the correlates of response. Studies that have included a multivariate analysis tend to show that the likelihood of response increases with age.[32, 236, 240] Two studies were found which showed that the likelihood of response decreases with age. One of these studies, however, was a study of elderly people with the youngest age in the study being 65 years.[241] The other study analysed responses to a questionnaire regarding sexual behaviour.[101] A possible explanation of the effect of age on response rates in this study could be that older people feel less comfortable answering questions of such a sensitive nature. The ambiguity in the health care literature regarding the effect of age on response highlights the complex interactions between response

variables and the subject of the questionnaire. It also demonstrates the necessity for multivariate analysis to be performed.

8.2.3 Sex

Sex is one of the most ambiguous of the socio-demographic correlates of survey response.[71] Being one of the easiest variables to establish, much research has been conducted on the male/female bias in surveys. The majority of studies in this area have found either no gender effect on survey cooperation or the tendency for males to have lower cooperation rates.[87, 242] Groves and Couper(1998)[87] have offered several theoretical explanations for this gender effect relating to role differentiation of males and females. There is no available evidence suggesting a gender effect in postal questionnaire surveys.

The health related survey literature suggests a similar pattern. The effect of gender on response to postal health related questionnaires is either negligible[101, 231, 235] or indicative of better response in women.[32, 236, 237] Closer inspection of this literature, however, again reveals a lack of multivariate analyses. Only one[101] of the studies found which suggested no effect of gender on response conducted a multivariate analysis to control for other variables.

8.2.4 Marital status

Some evidence suggests that single people are harder to contact but once reached are no less cooperative with surveys than married people.[243,

244] Little information is available for the effect of marital status specifically on response to postal questionnaires. One of the few studies found indicates little difference between married and single people.[245] Only one study was found relating to the effect of marital status on response to health related questionnaires.[101] This study also indicated no difference in response between married and single people. Reasons for the effect, if any, of marital status on response are not easy to explain. It is likely that, in this context, marital status is a proxy for other things such as lack of time to respond due to family commitments. Conversely, married people may be more likely to respond due to being more settled and less likely to change address.

8.2.5 Work status

Much of the survey research investigating the effect of work status on survey cooperation uses this as a proxy for the contactability of individuals to participate in face to face surveys. Individuals who are gainfully employed are elusive to this type of contact. Non-working people may have more free time for answering surveys but may also feel little personal involvement in many survey research topics.[71] Non-workers are also a heterogeneous group including a range of reasons for this status such as retired, long term unemployed, ill-health and prime family carers. Any effect of work status has been shown to diminish when subjected to multivariate modelling.[71] No evidence can be found to suggest an effect of work status on cooperation with a general survey postal questionnaire.

One study was found which included work status in an analysis of the demographic characteristics of respondents in a health survey.[246] This was a telephone survey, however, and univariate analysis showed responders to be significantly more likely to be currently employed.

8.2.6 Urban versus rural dwelling

The problem of securing face to face interviews with inner city dwellers is well documented in the survey literature. Residents in rural areas have the highest survey response rates, followed by small town dwellers then city dwellers.[242, 247] Definitions of 'rural', 'small town' and 'city' dwellers vary in their complexity from simple definitions based on populations[247] to more complex sub-categories of central city, large suburbs, smaller suburbs, other urban and rural.[242] Urbanicity is commonly observed as a response correlate because this variable is readily available from the sampling frame of a survey.[87] Although there is much evidence describing this effect there is little evidence to explain it. Some have suggested that the effect is possibly due to the inherent features of urban life - the faster pace and looser ties of community in such areas.[87] An early study found some evidence of urban dwellers resistance to postal questionnaire surveys but this was extrapolated in the discussion to occupational differences between farmers and non-farmers.[248] Two studies were found which specifically investigated response to health related questionnaires. One found better response rates from rural dwellers compared to city dwellers (OR 1.31

CI95% 1.02 to 1.68),[32] and the other found no difference between these two groups.[101]

8.2.7 Symptoms

A factor in the likelihood of response to a health care questionnaire is whether responders and non-responders differ in the very characteristic that the questionnaire is seeking to investigate. In general surveys, questionnaires are often used to collect information from a broad section of the population often regarding subjects which are of little personal concern to them. In health related research, and even more so in clinical trial follow-up, the questionnaire is likely to be of far more interest to the participant. There is some suggestion, however, that in symptom orientated questionnaires, those with health problems may be more likely to respond than those without.[249, 250] Other studies, however, have found that persons feeling well are more likely to respond.[251, 252] Different populations, therefore, appear to have different reasons for non-response.[237] In clinical trial follow-up this effect may manifest itself in participants not responding because they have recovered from the health state under investigation. The qualitative study presented in chapter 7 suggests that this may be happening in CAST. If this is the case, the final sample will be biased towards participants who still have symptoms, possibly leading to a misleading interpretation of the effect of the intervention. Norquist et al, [253] however, found the opposite effect. In a study of patients lost to follow-up in a clinical trial of shoulder muscle

injuries, they found that at the last known time of questionnaire response, non-responders reported significantly worse shoulder function. The only other study found relating specifically to clinical trial follow-up suggested that participants receiving the experimental treatment were more likely to respond without prompting.[254] An inherent problem in assessing symptom related effects on non-response is the difficulty in obtaining this information. The fact that the participant is a non-responder means that they have been lost to follow-up. Their symptoms and recovery status are therefore unknown. Unlike other potential sources of bias (such as the socio-demographic factors already discussed) symptoms and recovery are not fixed entities. Although initial symptoms are documented and recorded on recruitment onto a trial, these are likely to change through the course of the trial. No studies can be found which relate severity of initial symptoms to non-response at later attempts at follow-up. This highlights the necessity for achieving as little loss to follow-up as possible in clinical trials of health interventions to minimise this potential source of bias.

8.2.8 Summary

In summary, there exists a massive body of literature concerning the socio-demographic correlates of survey response behaviour, a limited proportion of which relates specifically to health related surveys. A very brief overview only is provided in the previous section. Unfortunately, this literature is somewhat unclear and poorly integrated. Suggested reasons for this lack of clarity are an over-emphasis on case studies and insufficient multivariate

analysis.[71] The contradictory findings of many of the studies of the effect of socio-demographic correlates on response could be explained by the effect of the involvement of the participant with the topic of the survey.[71] Very little research has looked specifically at clinical trial participants and the effect that the subject of the investigation has on follow-up response.

8.3 Socio-demographic characteristics of CAST participants

8.3.1 Definitions of CAST non-responders

The Collaborative Ankle Support Trial (CAST) provided an ideal opportunity to study the socio-demographic characteristics of participants in a clinical trial. The methodology of CAST is described in chapter 4. The CAST database contains detailed information on the socio-demographic characteristics of participants which was collected on recruitment onto the trial. A log of participant contact was also kept detailing the amount of prompting (if any) required by each participant at each follow-up point to encourage questionnaire return. The way in which participants who failed to return their questionnaires responded to this prompting, before being classed a non-responder, was also documented. A detailed perusal of this information reveals a slightly more complex situation than the simple 'refusal' or 'non-contact' conditions which have been reported in the survey literature. The CAST database revealed a further category of non-

responder; 'non-persuaders'. Table 42 gives a summary of the definitions of these categories.

Table 42 ***Definitions of non-responders in CAST***

Category	Definition
Non-contact	Questionnaire not returned but attempts at making follow-up contact unsuccessful
Non-persuader	Follow-up contact successful and participant agrees to return questionnaire but then still fails to respond (includes 'core outcome' data collected over the telephone).
Refuser	Participant actively refuses to return questionnaire and/or expresses the wish to withdraw from the trial

Section 8.2.7 highlighted the special case of bias in clinical trial follow-up relating to whether those with different outcomes of the trial are more or less likely to respond to follow-up. Since no information is available from non-responders regarding the outcome of the treatment intervention alternative methods have to be conceptualised to ascertain this information. Several authors have advocated using reluctant responders to surveys as 'proxies' for non-responders. The theory behind this is that those people who require extensive follow-up efforts are more similar to non-responders than they are to people who require minimal encouragement to respond.[71, 87] It may, therefore, be possible to infer something about the characteristics of non-responders from data gathered from reluctant responders.[17] This is a controversial subject in the survey literature with as many opponents to this

view as there are advocates. Some authors feel that the technique is flawed due to fundamental differences between people who are eventually persuaded to cooperate and those who continue to refuse.[255]

In CAST, those participants who were chased several times for their questionnaire but still failed to return it were asked 'core outcome' questions over the telephone. These participants were counted as responders in the final CAST response rate calculation because they supplied information for the trial's main outcomes. The information obtained in this manner was sufficient for the trial statistician to incorporate or impute enough information for useful analysis. For the purposes of the detailed analysis of response and non-response which constitutes this thesis, these same participants were classed as non-responders as no questionnaire was ever obtained form them (categorised as a 'non-persuader' in Table 42). This provided, therefore, an ideal situation to use 'core outcome' responders as a proxy for non-responders. Information was available for 'core outcome' responders at the 12 week and nine month follow-up points on how much benefit they felt they had obtained from treatment as part of CAST. It was therefore possible to compare the benefit obtained by 'core outcome' responders (in their capacity as non-responders) to that obtained by responders. Any differences may suggest non-response bias as a result of outcome of treatment.

8.3.2 Methods

8.3.2.1 Sample

CAST recruited 555 participants (excluding those recruited during the pilot phase of the trial). Data were available at each follow-up time point regarding whether the participant was a responder or a non-responder. This information was then combined with the CAST database of background information enabling comparisons to be made between responders and non-responders in terms of certain socio-demographic characteristics.

8.3.2.2 Comparisons

1. Non-responders versus Responders

The main analysis compared non-responders and responders at each follow-up point in terms of the following characteristics: Age, Sex, CAST treatment received, Employment type and Education level achieved. Apart from treatment received, these were the most commonly investigated characteristics in the survey literature.

2. *Categories of non-responder*

Table 42 gives the definition of three different types of non-responder identified in CAST; 'non-contact', 'non-persuader' and 'refuser'. The identification of these categories, however, was a post-hoc observation made at the end of data collection. For this reason these categories were not used in the main analysis but were the subject of a secondary analysis. The same correlates were used as those used in the main analysis described above and the three categories of non-responder were compared.

3. Benefit obtained from treatment

As described in section 8.3.1, participants giving core outcome information over the telephone could be considered 'proxies' for non-responders. In the 12 week and nine month follow-up questionnaires, question 88 asked participants to rate their benefit obtained form the treatment given as part of CAST from 1 to 10 (see appendix 2). A score of 0 indicated 'no benefit' and a score of 10 indicated 'maximum benefit'. It was therefore possible to compare the mean scores for this question given by responders and the 'core outcome' proxies for non-responders.

8.3.2.3 Statistical methods

Univariate analysis of the comparisons was carried out using chi-squared tests where appropriate. The data from comparison one were subjected to a multivariate analysis using logistic regression. The same procedure was followed as that used, and described in detail, in chapter 6. Variables were selected for inclusion in the multivariate analysis if they had a p value of <0.25 on univariate analysis. The 'maximum likelihood' method of logistic regression was then used to assess multiple variables.

8.3.3 Results

In the CAST target population of 555 participants the mean age was 29.9 years (range 16-72, SD 10.8) and 58% were male. The number of participants in each of the four treatment arms of the trial was fairly similar although there was a slight under representation in those receiving a plaster cast. The majority of participants were employed and the most common

type of employment was skilled. Just over half the participants had achieved an educational level of A levels or higher.

1. Non-responders versus Responders

The results of these analyses at each follow-up point in CAST are shown in Table 43,Table 44 and Table 45

Univariate analysis

There were significantly more non-responders in the youngest age group at the four week (OR 0.23, CI 95% 0.12 to 0.73) and nine month (OR 0.19, CI 95% 0.09 to 0.40) follow-up points. Although the same trend was seen at the 12 week follow-up point the difference was not significant. There was a tendency at every follow-up point for females to respond better than males but this was only significant at 12 weeks (OR 1.5, CI 95% 1.1 to 2.2). There were no significant differences at any follow-up point between non-responders and responders in terms of the treatment received as part of CAST or education level. Type of employment was not associated with non-response at the four week follow-up point but at 12 weeks non-responders were more likely to be unemployed (OR 0.36, CI 95% 0.12 to 1.1) or in skilled employment (OR 0.34, CI 95% 0.11 to 1.02). At nine months the same pattern emerged in the unemployed (OR 0.29, CI 95% 0.10 to 0.82) and skilled workers (OR 0.30 CI 95% 0.11 to 0.84). Professionals were less likely to be non-responders at these time points.

Multivariate analysis

At the four week follow-up point the variables of Age and Employment type reached the p<0.25 criterion for inclusion in the multivariate model. When controlling for employment type, there were significantly more non-responders in both the youngest age group (OR 0.31, CI 95% 0.12 to 0.75) and the 35 to 44 year olds (OR 0.38, CI 95% 0.14 to 0.98). Employment type continued to have no significant effect on response even when age was controlled for.

At the 12 week follow-up point the variables of Age, Sex, Employment and Education level reached the entry criterion for multivariate analysis. Age and Education level added little to the multivariate model. Although both Employment type and Sex significantly affected response, the results of the multivariate model combining these two variables were very similar to the univariate analysis.

At the nine month follow-up point the variables of Age, Sex and Employment type were entered into the multivariate analysis. The model containing Age and Employment proved to be useful for predicting response but controlling for Sex contributed little to the predictive power of the model. The interaction between Age and Sex was tested at each time point. The interaction between these two variables was not significant at any follow-up point.

Table 43 ***Distribution of socio-demographic characteristics in non-responders compared with responders at the CAST 4 week follow-up point (%)***

	Non-responders	Responders	Target Population	χ^2
N	122	433	555	
Mean age (yrs)	27.3	30.6	29.9	
Age				
16-24	62 (27)	166 (73)	228	
25-34	28 (19)	121 (81)	149 118	
35-44	26 (22)	92 (78)	60	
44+	6 (10)	54 (90)		0.023
Sex				
Male	73 (23)	246 (77)	319	
Female	49 (21)	187 (79)	236	0.551
Treatment				
Tubigrip	37 (24)	117 (76)	154 114	
Plaster	20 (18)	94 (82)	144	
Aircast	37 (26)	107 (74)	143	
Bledsoe	28 (20)	115 (80)		0.345
Employment				
Nil	31 (25)	92 (75)	123	
Unskilled	20 (25)	61 (75)	81	
Skilled	46 (23)	151 (77)	197	
Professional	19 (15)	110 (85)	129	
Other	6 (24)	19 (76)	25	0.256
Education				
Nil	14 (24)	44 (76)	58	
CSE/GCSE	46 (21)	169 (79)	215	
Alevel/degree/	62 (22)	220 (78)	282	
Other				0.905

Table 44 ***Distribution of socio-demographic characteristics in non-responders compared with responders at the CAST 12 week follow-up point (%)***

	Non-responders	Responders	Target Population	χ^2
N	171	384	555	
Mean age (yrs)	28.4	30.6	29.9	
Age				
16-24	81 (36)	147 (64)	228 149	
25-34	47 (32)	101 (68)	118	
35-44	30 (25)	89 (75)	60	
44+	13 (22)	47 (78)		0.092
Sex				
Male	110 (35)	208 (65)	319 (58)	
Female	61 (26)	175 (74)	236 (42)	0.024
Treatment				
Tubigrip	45 (29)	108 (71)	154 114)	
Plaster	35 (31)	79 (69)	144	
Aircast	53 (37)	91 (63)	143	
Bledsoe	38 (27)	105 (73)		0.299
Employment				
Nil	43 (35)	80 (65)	123 81	
Unskilled	30 (37)	51 (63)	197 129	
Skilled	70 (36)	126 (64)	25	
Professional	24 (19)	105 (81)		
Other	4 (16)	21 (84)		0.002
Education				
Nil	24 (41)	34 (59)	58 215	
CSE/GCSE	70 (33)	145 (77)	282	
Alevel/degree/	77 (27)	204 (73)		
Other				0.098

Table 45 ***Distribution of socio-demographic characteristics in non-responders compared with responders at the CAST 9 month follow-up point (%)***

	Non-responders	Responders	Target Population	χ^2
N	212	343	555	
Mean age (yrs)	27.2	31.5	29.9	
Age				
16-24	110 (48)	118 (52)	228 149	
25-34	55 (37)	94 (63)	118	
35-44	38 (32)	80 (68)	60	
44+	9 (15)	51 (85)		<0.001
Sex				
Male	132 (41)	187 (59)	319	
Female	80 (34)	156 (66)	236	0.073
Treatment				
Tubigrip	57 (37)	97 (63)	154	
Plaster	44 (39)	70 (61)	114	
Aircast	59 (41)	85 (59)	144	
Bledsoe	52 (36)	91 (64)	143	0.856
Employment				
Nil	57 (46)	66 (54)	123 81	
Unskilled	28 (35)	53 (65)	197	
Skilled	89 (45)	108 (55)	129	
Professional	33 (26)	96 (74)	25	
Other	5 (20)	20 (80)		<0.001
Education				
Nil	26 (45)	32 (55)	58	
CSE/GCSE	82 (62)	133 (38)	215	
Alevel/degree/ other	104 (37)	178 (63)	282	0.525

2. Categories of non-responder

The results of comparisons between the different categories of non-responder at each follow-up point are shown in Table 46, Table 47 and Table 48

The analyses revealed that the refusal rate in the CAST follow-up was very low. By the end of the nine month follow-up only 12 out of the sample of 555 actively refused to cooperate. At each successive time point the number of non-persuaders was fairly similar. The number of non-contacts, however, almost doubled at each successive time point. The most prevalent non-response category in all age groups at four weeks and 12 weeks was non-persuader. At the nine month follow-up, however, this pattern changed and the younger age groups were more likely to be non-contacts. There were no major gender differences in any of the non-response categories at any follow-up point. Participants who received a plaster were more likely to be non-contacts at the four week follow-up point. There were no other major differences in response categories by treatment received at any follow-up point. Although the total number of refusers was low, the majority of this category consisted of unemployed people. All the refusers at four weeks, and the majority at the nine month follow-up, were from the highest education level.

Table 46 ***Distribution of socio-demographic characteristics in distinct groups of non-responders compared with responders at the CAST 4 week follow-up point (%)***

	Type of non-response				Total	Responders	Target Population
	Non-Persuader (A)	*Non-Contact (B)*	*Refuser (C)*	*A+B+C*			
N	82	37	3	122	433	555	
Mean age (yrs)	27.5	26.5	32.7	27.3	30.6	29.9	
Age							
16-24	40 (48)	22 (60)	0 (0)	62 (51)	166 (38)	228 (41)	
25-34	18 (22)	8 (22)	2 (67)	28 (23)	121 (28)	149 (27)	
35-44	20 (24)	5 (14)	1 (33)	26 (21)	92 (21)	118 (21)	
44+	4 (5)	2 (5)	0 (0)	6 (5)	54 (13)	60 (11)	
Sex							
Male	48 (59)	24 (65)	1 (33)	73 (60)	246 (57)	319 (58)	
Female	39 (41)	13 (35)	2 (67)	49 (40)	187 (43)	236 (42)	
Treatment							
Tubigrip	29 (35)	6 (16)	2 (67)	37 (30)	117 (27)	154 (28)	
Plaster	9 (11)	10 (27)	1 (33)	20 (16)	94 (22)	114 (21)	
Aircast	25 (31)	12 (32)	0 (0)	37 (30)	107 (25)	144 (26)	
Bledsoe	19 (23)	9 (24)	0 (0)	28 (23)	115 (27)	143 (26)	
Employment							
Nil	21 (26)	8 (22)	2 (67)	31 (25)	92 (21)	123 (22)	
Unskilled	15 (18)	5 (13)	0 (0)	20 (16)	61 (14)	81 (15)	
Skilled	30 (37)	15 (41)	1 (0)	46 (38)	151 (35)	197 (35)	
Professional	12 (15)	7 (19)	0 (33)	19 (16)	110 (25)	129 (23)	
Other	4 (5)	2 (5)	0 (0)	6 (5)	19 (4)	25 (4)	
Education							
Nil	7 (9)	7 (19)	0 (0)	14 (11)	44 (10)	58 (10)	
CSE/GCSE	33 (40)	13 (35)	0 (0)	46 (38)	169 (39)	215 (39)	
Alevel/degree/ other	42 (51)	17 (46)	3 (100)	62 (51)	220 (51)	282 (51)	

Table 47 ***Distribution of socio-demographic characteristics in distinct groups of non-responders compared with responders at the CAST 12 week follow-up point (%)***

	Type of non-response			**Total**	**Responders**	**Target Population**
	Non-Persuader (A)	*Non-Contact (B)*	*Refuser (C)*	*A+B+C*		
N	100	65	6	171	384	555
Mean age (yrs)	28.5	27.1	39.7	28.4	30.6	29.9
Age						
16-24	44 (44)	36 (55)	1 (17)	81 (47)	147 (38)	228 (41)
25-34	30 (30)	16 (25)	1 (17)	47 (28)	101 (26)	149 (27)
35-44	17 (17)	10 (15)	3 (50)	30 (18)	89 (23)	118 (21)
44+	9 (9)	3 (5)	1 (17)	13 (8)	47 (12)	60 (11)
Sex						
Male	64 (64)	42 (65)	4 (67)	110 (64)	208 (54)	319 (58)
Female	36 (36)	23 (35)	2 (33)	61 (36)	175 (46)	236 (42)
Treatment						
Tubigrip	28 (28)	15 (23)	2 (33)	45 (26)	108 (28)	154 (28)
Plaster	20 (20)	13 (20)	2 (33)	35 (21)	79 (21)	114 (21)
Aircast	34 (34)	18 (28)	1 (17)	53 (31)	91 (24)	144 (26)
Bledsoe	18 (18)	19 (30)	1 (17)	38 (22)	105 (27)	143 (26)
Employment						
Nil	24 (24)	14 (22)	5 (83)	43 (25)	80 (21)	123 (22)
Unskilled	19 (19)	11 (17)	0 (0)	30 (18)	51 (13)	81 (15)
Skilled	39 (39)	31 (48)	0 (0)	70 (41)	126 (33)	197 (35)
Professional	15 (15)	8 (12)	1 (17)	24 (14)	105 (27)	129 (23)
Other	3 (3)	1 (1)	0 (0)	4 (2)	21 (6)	25 (4)
Education						
Nil	13 (13)	9 (14)	2 (33)	24 (14)	34 (9)	58 (10)
CSE/GCSE	43 (43)	25 (38)	2 (33)	70 (41)	145 (38)	215 (39)
Alevel/degree/ Other	44 (44)	31 (48)	2 (33)	77 (45)	204 (53)	282 (51)

Table 48 ***Distribution of socio-demographic characteristics in distinct groups of non-responders compared with responders at the CAST 9 month follow-up point (%)***

	Type of non-response			Total	Responders	Target Population
	Non-Persuader (A)	*Non-Contact (B)*	*Refuser (C)*	*A+B+C*		
N	91	109	12	212	343	555
Mean age (yrs)	27.4	26.7	31.0	27.2	31.5	29.9
Age						
16-24	45 (50)	61 (57)	4 (33)	110 (52)	118 (34)	228 (41)
25-34	22 (24)	29 (27)	4 (33)	55 (26)	94 (27)	149 (27)
35-44	21 (23)	14 (13)	3 (25)	38 (18)	80 (23)	118 (21)
44+	3 (3)	5 (5)	1 (8)	9 (4)	51 (15)	60 (11)
Sex						
Male	58 (64)	68 (62)	6 (50)	132 (62)	187 (55)	319 (58)
Female	33 (36)	41 (38)	6 (50)	80 (38)	156 (45)	236 (42)
Treatment						
Tubigrip	23 (25)	29 (27)	5 (42)	57 (27)	97 (28)	154 (28)
Plaster	20 (22)	23 (21)	1 (8)	44 (21)	70 (20)	114 (21)
Aircast	26 (29)	28 (26)	5 (42)	59 (28)	85 (25)	144 (26)
Bledsoe	22 (24)	29 (27)	1 (80	52 (24)	91 (27)	143 (26)
Employment						
Nil	23 (25)	28 (26)	6 (50)	57 (27)	66 (19)	123 (22)
Unskilled	13 (14)	12 (11)	3 (25)	28 (13)	53 (15)	81 (15)
Skilled	37 (41)	49 (45)	3 (25)	89 (42)	108 (32)	197 (35)
Professional	17 (19)	16 (15)	0 (0)	33 (16)	96 (28)	129 (23)
Other	1 (1)	4 (4)	0 (0)	5 (2)	20 (6)	25 (4)
Education						
Nil	9 (10)	13 (12)	4 (33)	26 (12)	32 (9)	58 (10)
CSE/GCSE	36 (40)	45 (41)	1 (8)	82 (39)	133 (39)	215 (39)
Alevel/degree/ other	46 (50)	51 (47)	7 (58)	104 (49)	178 (52)	282 (51)

3. Benefit obtained from treatment

The results of this analysis are shown in Table 49

Table 49 ***Comparison of mean scores (standard deviations) for 'benefit of treatment' question on CAST follow-up questionnaires at 12 weeks and nine months***

	Mean score for benefit gained from treatment (SD)		Difference in means	95% CI*
	Core outcomes	**Responders**		
12 week follow-up	(n=73) 6.7 (2.3)	(n=384) 6.1 (3.0)	0.60	-0.13 to 1.33
9 month follow-up	(n=79) 7.2 (2.3)	(n=343) 6.3 (3.0)	0.90	0.19 to 1.61

* Confidence Interval

There was no significant difference in the mean score for benefit gained from treatment between 'core outcome' responders and responders at the 12 week follow-up. At the nine month follow-up point, however, the difference in the mean scores was significant with core outcome responders having a higher perceived benefit of treatment.

8.3.4 Discussion

In the general health care literature, the effect of age on response to postal questionnaires is unclear. The analysis of the effect of age on response in CAST demonstrates a positive effect of age on response with the best

responders in the 44+ age bracket. This may reflect that the older participants have more spare time to sit and complete the questionnaires and may be more motivated and interested in health issues than younger age groups. Having more spare time may not be the most plausible explanation, however, as it could be argued that age 44+ is not that old and many participants in this age bracket will probably still be working or be busy with family commitments.

The nature of CAST was to investigate treatments for severe ankle sprains. In the youngest age group over 50% of the injuries were sports related (compared to 12% in the oldest age group). A possible theory for the higher rate of non–response in this age group could be that these participants had less time to complete the questionnaires due to work but also sport/leisure commitments. Females were generally better at responding than males, a pattern which has been noted in the health related survey literature.[236, 237] In CAST participants there may also be a similar association with the amount of sporting activity as that seen with age. Sports related injuries accounted for 51% of all the injuries in the male population compared to 17% in the females.

The fact that there were no differences between responders and non-responders in terms of the treatment received as part of CAST is a considerable positive factor of the trial. Other authors have noted that the relationship between readiness to respond and treatment group has

particular implications for unblinded trials in which a greater enthusiasm may be generated by the 'new' treatment.[254]

The similarity of mean scores of benefit of treatment of 'core outcome' responders and responders was a positive finding at the CAST 12 week follow-up. At the nine month follow-up, however, the 'core outcome' proxies for non-responders rated their benefit of treatment significantly higher than responders. This suggests that a better recovery leads to a lower likelihood of response. This finding supports the emergent theme in the qualitative study presented in chapter 7 that suggested an association between recovery and non-response.

The fact that no difference was found in the education level of responders and non-responders may reflect the way that a clinical trial follow-up questionnaire is administered. In the survey literature, education level is one of the few variables which appear to be consistently related to response. Studies have shown that people with less formal education are less likely to respond.[231, 232] In CAST, the questionnaire is completed by the participant at the recruitment stage in the presence of the researcher. A similar questionnaire is then posted to the participant at each follow-up point. At this initial completion, the participant is able to clarify any points with the researcher who is available to answer any questions. This may make those less well educated or less literate participants more confident in completing their subsequent questionnaires. In a survey, with no prior contact, a lengthy questionnaire may appear very daunting to this group of

people. This is reflected in higher levels of non-response in this group. The other SES indicator which it was possible to ascertain from the CAST database was employment type. At the four week follow-up point employment type made no difference to response. At 12 weeks and 9 months, however, type of employment did have some affect on response.

A notable difference between the CAST follow-up data and that seen in survey research is the number of participants actively refusing to cooperate with the follow-up questionnaires. In CAST very few participants fell into this group. A recent trend in survey research is that refusals now contribute more to non-response than non-contacts.[26] This is likely to reflect a fundamental difference between clinical trial participants and survey participants. In a clinical trial, prior to receiving a postal questionnaire, participants have already consented to be involved in the research activity. Consent is gained after the potential participant is informed of the procedures involved in the trial. This includes an explanation of the follow-up process. Clinical trial participants are, therefore, aware that questionnaires will be sent to them at future follow-up points. Participants who are unwilling to be part of a clinical trial usually refuse at the recruitment stage. Potential survey participants, however, usually have no prior knowledge of the arrival of a questionnaire. Refusal rates to postal questionnaires once they arrive are therefore much higher.

Other types of non-response categories identified in CAST were 'non-contacts' and 'non-persuaders'. The number of non-contacts increased at

each time point until at the nine month follow-up point more than half the non-responders fell into this group. The log kept by the CAST administration team enabled a rudimentary analysis of reasons for the loss of contact. This revealed that contact was lost in the majority of these participants because the telephone number given was incorrect or had changed or the participants had moved house with no forwarding address. Non-contacts offer, perhaps, the most potential for boosting response rates in clinical trial follow-up. It is, of course, imprudent to assume that all non-contacts can be converted into responders. There may well be other factors apart from non-contact which feature in the reasons for a participant's non-response. However, at the nine month follow-up in CAST, if even half the non-contacts were converted into responders this would have increased the final response rate from 76% to 86%. Careful attention should be paid at the recruitment stage to ensure that contact details are correct and to obtain as many different contact numbers as possible (i.e. home, work and mobile numbers). This may reduce the amount of non-response as a result of non-contact. Many participants were classed as non-contacts because they failed to answer follow-up phone calls. To make contact with such 'hard to reach' participants it may be useful to obtain an email address and use email prompts or to use text messaging prompts. These modern methods of clinical trial follow-up have not been tested experimentally but offer interesting avenues for future research. Much time and effort was expended

in CAST in attempts to contact 'non-contacts'. If this group can be kept to a minimum the time and expense saved would be an added benefit.

8.3.5 Conclusions

A univariate and multivariate analysis of the CAST database has shown that, apart from age, there were few significant differences in the characteristics of responders compared to non-responders. This indicates that the results of CAST are unlikely to be affected by non-response bias. By categorising non-responders into different types it was possible to establish that non-response due to non-contact was an ever increasing factor throughout the progression of the trial. This offers, perhaps, the most promising area to target in efforts to maximise response rates. Refusal rates to postal questionnaire follow-up in CAST were consistently low at each follow-up point. This reflects the fundamental differences in the way that clinical trial participants are approached to comply with a postal questionnaire compared to survey participants.

8.4 Chapter summary

This chapter has focussed on the characteristics and categories of *non-responders* to postal questionnaires. It became evident during the progress of this chapter that this is an area of investigation which could make up a substantial bulk of a thesis such as this in its own right. This was not anticipated at the outset. There are many interesting issues which have

come to light. For example, as well as non-responder categories, different categories of *responders* have also been identified in the survey literature relating to the amount of prompting required to encourage questionnaire return.[118] Data relating to categories of responder (e.g. 'keen' or 'reluctant') were collected as part of CAST. The characteristics of the CAST responders, however, have not been subjected to detailed analysis as part of this thesis. This is a tantalising prospect for the development of the research area established during the evolution of this thesis.

9 Chapter 9

Conclusions and recommendations

9.1 Summary

This study was undertaken to identify methods of improving response to follow-up by postal questionnaire in a clinical trial. A further aim was to understand the deeper issues surrounding the response decisions made by clinical trial participants. This is the first such in depth investigation into an area which can have important implications on the quality and interpretation of the outcomes of a clinical trial. A cohort of active clinical trial participants was studied from both a quantitative and qualitative perspective. This allowed for a deep analysis of the response behaviour of clinical trial participants. The aim of this research was to develop recommendations to help clinical trialists deliver robust and unbiased research findings. Because there is such little previous work in this area specific to clinical trials, the literature surrounding response to postal surveys was used as a theoretical reference. The survey literature is vast and it was necessary to be very selective with the use of this literature in order to remain focussed on the central issues of this thesis. The challenge was to make use of the survey literature without making assumptions as to its direct relevance to the context of clinical trial follow-up. An important revelation to emerge from an evaluation of the survey literature surrounding response to postal

questionnaires is the fact that it constitutes a massive but poorly integrated body of evidence. Criticisms have been made relating to the literature regarding the socio-demographic determinants of response[71] and also the literature suggesting methods of improving response to postal questionnaires.[55, 256] In a comprehensive review of methods of improving response to mailed surveys, Linsky (1975)[55] concluded that 'Given the substantial body of research on the mail questionnaire, much of it experimental, it seems surprising that more has not been learned' (p100) There appears to have been little progress over subsequent years. Childers and Skinner (1996)[256] postulated that the knowledge on survey response behaviour remains unchanged from the view expressed by Linsky. The atheoretical nature of research into postal questionnaires has also been identified by Herberlein and Baumgartner.[20] In a substantial review of factors affecting response to mailed questionnaires they concluded that 'what is not needed is another study reporting the effects of contacts, postage stamps or colour of paper on a single instrument on a single population' (p460). The results of the randomised controlled trial of the Trial Calendar presented in chapter 6 reflect this view. It was disappointing not to have shown a positive effect of the Trial Calendar on response. However, it could be argued that the results confirm that such experiments add little to the 'big picture' of clinical trial response. A recent study has shown that including a pen with the questionnaire significantly improved follow-up in a trial of the management of women with abnormal cervical smears.[240] This

is the only other study found which randomised a method of improving response into an existing clinical trial (this study was conducted after completion of the systematic review presented in chapter 5 and is therefore not included in the review). If subsequent research follows the same pattern as the survey research there are likely to be as many studies demonstrating positive effects of such methods as negative. A hypothetical situation could arise in years to come akin to the current climate in the survey research of methods of improving response to postal questionnaires. That is, a disjointed and equivocal body of research which offers little practical recommendations for clinical trialists striving to maximise their follow-up response rates.

The lessons learned from nearly 100 years of survey research deliver, perhaps, the most useful contribution of the survey literature to the clinical trial setting. Research in this specific aspect of clinical trials is in its infancy. Steps can be taken, therefore, to ensure that a body of evidence regarding issues of maximising response to questionnaire follow-up is developed which considers the theoretical issues of response.

Contemporary research in survey non-response has reached an era which reflects a reluctant recognition that survey response rates are declining and declining at an increasing rate.[257] Survey researchers are now turning their focus on evaluating the effect low response rates have on inferences to the target population. Some authors proclaim that there is not necessarily a connection between non-response rate and non-response bias after all.[27]

This is another area in which it may be prudent to not to rely too heavily on survey research to inform conduct in a clinical trial. An important 'characteristic' of clinical trial participants which is not present in survey participants is their outcome as a result of the intervention being tested in the trial. Since this is an unknown factor in non-responders, every effort should be made to ensure as complete a follow-up as possible. The issues of non-response and bias, however, are not confined to differences in responders and non-responders in a clinical trial. Good quality clinical trials calculate *a priori* the sample size required to enable meaningful statistical analysis of the outcome data. Loss to follow-up reduces the effective sample size and can therefore have serious implications on the power of the trial.

9.2 Practical conclusions

The systematic review of the health care literature presented in chapter 5 concluded that follow-up contact offers the most promising method of improving response to postal questionnaires. The randomised controlled trial of the Trial Calendar presented in chapter 6 highlighted that such prompts are ineffective if they are too subtle and may need to be 'high profile' to work. High profile prompts are such things as making telephone contact with participants and verbally appealing for questionnaire return. This method appears to have been effective in achieving overall high response rates in CAST (although it was not tested experimentally). Due to

the direct relevance of the postal questionnaires to participants in clinical trials, response rates will be naturally higher than those of a survey.[20] It has also been suggested that as response rates increase, the effect of efforts to persuade non-responders to cooperate diminishes.[57] Relying too heavily on the survey literature as a reference for methods of improving response may therefore be misguided. Such literature often reports methods of improving response from levels which would be totally unacceptable in a clinical trial. Response rates as low as 40% are not uncommon in surveys.[20] For the results of a clinical trial to be considered valid it has been suggested that the lowest acceptable response rate is 80%.[10] Studies investigating maximising response in clinical trials should therefore have a sample size based on increasing response to this acceptable level. For example, the calendar trial presented in chapter 6 based its sample size on improving the response rate from 65% to 80%. The only other study found which investigates a method of improving response in a clinical trial based its sample size on improving the response rate from 62% to 70%.[240] This study enclosed a pen with the questionnaire as an incentive to respond which resulted in a significant improvement in response to 68.5%. Additional follow-up efforts as part of the main trial consisted of postal reminders but no telephone contact. Although any improvement in response rate is worthwhile in reducing the risk of bias, this study is of little value in helping clinical trialists achieve the recommended 80% follow-up. It is, however, a good example of the type of

research which will add to the hypothetical (and undesirable) scenario outlined at the end of the previous section.

Chapter 8 presented an analysis of the different types of non-responder identified in CAST. This chapter concluded that the 'non-contacts' constitute a large number of participants with the potential to convert into responders. If only a relatively small proportion of these participants were contacted and persuaded to respond this could have a significant impact on the final response rates.

9.3 Theoretical conclusions

During the development of the research theme of this thesis it became evident that the theoretical aspects of clinical trial response offer more scope for understanding how to maximise it. This is also the conclusion which was eventually arrived at by survey researchers and in recent years attempts have been made at conceptualising mail survey response behaviour.[256]

The qualitative study presented in chapter 7 highlighted many ways in which the response decisions of clinical trial participants could be explained by theories of human behaviour. Response to postal questionnaire follow-up in a clinical trial may therefore be improved by making subtle appeals to psychological processes used by participants. These appeals can be incorporated into the written material supplied to participants but also in the verbal instructions given by the researcher when the participant is recruited

onto the trial. Interviews with CAST participants suggest that emphasising the relevance of the questionnaire and making appeals to feelings of 'consistency' (i.e. a commitment to seeing the trial through to completion) will encourage response. Nurturing altruistic feelings associated with response and emphasising the legitimacy of the research institution may also positively influence response rates. The importance that clinical trial participants assign to these aspects of their response decisions is likely to vary depending on the characteristics of the population being studied and the nature of the trial. The 'Leverage-Saliency' theory of survey response[57] offers, therefore, the most relevant theoretical insight into clinical trial follow-up response. This theory allows for the interaction of many circumstances which may affect response. Unlike the other theories of response behaviour, the 'Leverage-Saliency' theory was developed specifically to help explain the response decisions of survey participants. Rather than just explaining the possible psychological processes involved in decision making, this theory goes a step further. It conceptualises a method of appealing to these decision making processes in the context of the set of circumstances under which they occur. In the context of a clinical trial, the way appeals for response are incorporated into the design of the trial should be developed with the target population and subject of the trial in mind. This theory helps to explain the apparent ineffectiveness of incentives to encourage response in the health care setting. By attempting to understand the target population *a priori,* appeals can be made to their unique attributes

at follow-up. All that remains then is to identify how different clinical trial populations make their response decisions. It is proposed that the best way to establish this is to ask the clinical trial participants themselves as was done as part of this thesis. It is theorised that this area represents the most appropriate direction for future research into the area of response to postal questionnaire follow-up in clinical trials. It is recognised that no two populations or clinical trial situations will be the same. However, it may be possible to develop models which can be generalised across trials with similar populations (e.g. age or sex) or trials investigating specific disease conditions (e.g. cancer or diabetes).

A link between response and the participants understanding of the trial procedures also emerged from the qualitative study. There was a suggestion that some participants did not fully understand that the trial had three follow-up points and it was essential to respond to each one. These participants were more likely to be non-responders. This link requires further investigation but offers another potential avenue for maximising response.

Ultimately, it is concluded that rather than anticipating low response rates and striving to devise methods of converting non-responders into responders, efforts should be directed at preventing participants becoming non-responders in the first place. This thesis argues for the area of follow-up to postal questionnaires in clinical trials to become a theoretical research issue in it own right. The lessons learned form the survey literature support

this view. If further research in this fledgling area adheres to this principle the hypothetical future scenario will be considerably more pleasing. Clinical trialists will be able to base their trial design in this area on an informative, well integrated and theoretical body of literature.

9.4 Recommendations

The aim of this research endeavour was to identify methods of maximising response to postal questionnaires used as a method of follow-up in clinical trials. Based on the findings of this research the following recommendations are made:

9.4.1 Recommendations for clinical trialists

- Clinical trialists should consider carefully and *a priori* the follow-up protocols to be used to chase reluctant responders. This should include appropriately timed telephone prompts although, at present, there is no literature to suggest the optimum number of telephone prompts.
- Non-response as a result of 'non-contact' should be kept to an absolute minimum. As many contact numbers as possible should be obtained from participants at the recruitment stage. Email addresses could be included as another avenue for prompting. It may be worth stressing to the participant the length of the follow-up period and ascertaining whether any of the given contact details are likely to change during that period.

- The use of incentives to encourage response is not recommended.

  There is no practical or theoretical evidence that this is effective in the clinical trial setting.

- The wording of both written and verbal trial information should be considered carefully to appeal to deeper theoretical issues of response. This could include emphasising the relevancy of the follow-up questionnaires to the participant. It is also necessary to ensure that participants fully understand what their involvement in the trial will entail in terms of follow-up. If patients are unwilling, incapable or have some other reason for not being able to complete every follow-up questionnaire, it may be more appropriate to exclude them from the trial prior to randomisation.[8] The legitimacy of the research institution should be conveyed to the participant and trial information should also emphasise ways in which their involvement will help the advancement of medical knowledge in the area under investigation. The language used in such appeals, however, needs to be carefully worded so as not to appear coercive.
- Clinical trialists should strive to understand their target populations. The previous literature concerning trials of similar populations or interventions may offer useful insights into issues which appear to affect response. These issues can then be made more or less salient in appeals to encourage response in new studies.

9.4.2 Recommendations for further research

Several areas of this thesis generated ideas for the development of the analyses already carried out. Work, however, has to stop somewhere and in the avoidance of making this thesis several volumes long the following areas have been saved for future enjoyment!:

- Developing the theme of using 'core outcome' responders as proxies for non-responders which was established in chapter 8, it would be useful to assess whether outcome affected response. Scores gained by responders on measures of recovery could be compared to those gained by 'core outcome' responders. This would help to establish whether outcome in CAST biased the results.
- Chapter 8 could also be developed to include an analysis of the different types of 'responders'. CAST identified 'keen' and 'reluctant' responders. These two groups could be studied to evaluate any differences in the characteristics of these two types of responders. This would generate useful information for the design of future trials. For example, if it appeared that reluctant responders displayed certain characteristics it may be possible to identify these at recruitment and tailor subsequent appeals for response accordingly.
- The emergent link between trial understanding and response which was identified in chapter 7 needs deeper investigation. A further qualitative study of clinical trial participants could be conducted to focus specifically on just what participants understand about the procedures and purpose

of the trial they are involved in. This would give a more robust indication of any links between trial understanding and response or non-response.

As a result of this research, several general areas have been identified which require further investigation. The following suggestions are recommended as ways to establish response to postal questionnaire follow-up in clinical trials as a well-grounded theoretical subject.

- The systematic review presented in chapter 5 concluded that follow-up efforts are the most effective way of boosting response rates to postal questionnaire in health care research. Experience from CAST suggests that telephone prompts may be the most appropriate method of follow-up. To make this recommendation more relevant to the clinical trial setting it would be useful to systematically review the literature of clinical trials which used telephone prompts in their follow-up procedures. This would generate information on the number and timing of telephone prompts enabling recommendations on this method of follow-up to be more robust.
- Much work has been carried out to identify common socio-demographic characteristics of survey participants which appear to affect response. No such work has investigated this area specific to clinical trial participants. Responders and non-responders to follow-up in a clinical trial may demonstrate quite different socio-demographic characteristics to survey participants. A systematic review of the literature in this area

will therefore help clinical trialists understand if any particular group of individuals are 'at risk' of not responding.

- The 'Leverage-Saliency' theory of survey response[57] appears to be readily applicable to the clinical trial setting. This requires further application in this setting, however, to be adopted fully into the theoretical evidence base of clinical trial response.
- Rather than future studies investigating isolated methods of improving questionnaire response (e.g. manipulating aspects of the questionnaire design and administration), it is recommended that the theoretical nature of clinical trial response be the focus of attention. More studies are needed which gain the participants perspective of response in many different clinical trial situations. This is seen as the most promising direction for future research in this important area.

References

[1] Lamb SE, Nakash RA, Withers EJ, Clark M, Marsh JL, Wilson S, et al. Clinical and cost effectiveness of mechanical support for severe ankle sprains: design of a randomised controlled trial in the emergency department. BMC Musculoskeletal Disorders. 2005;6(1).
[2] Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess. 1998;2(14):i-iv, 1-74.
[3] Slevin ML, Plant H, Lynch D, Drinkwater J, Gregory WM. Who should measure quality of life, the doctor or the patient? Br J Cancer. 1988 Jan;57(1):109-12.
[4] Kaasa S. Measurement of quality of life in clinical trials. Oncology. 1992;49(4):289-94.
[5] Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? Stat Med. 1984 Oct-Dec;3(4):409-22.
[6] Maxim P. Quantitative Research Methods in the Social Sciences. New York, Oxford: Oxford University Press.
[7] Edwards P, Roberts I, Sandercock P, Frost C. Follow-up by mail in clinical trials: does questionnaire length matter? Control Clin Trials. 2004 Feb;25(1):31-52.
[8] Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet. 2002 Mar 2;359(9308):781-5.
[9] Sackett D, Richardson W, Rosenberg W, Haynes R. Evidence-based medicine: how to practice and teach EBM. New York: Churchill Livingstone 1997.
[10] Sackett D, Straus S, Richardson W, Rosenberg W, Haynes R. Evidence-based medicine: how to practice and teach EBM. Edinburgh: Churchill Livingstone 2000.
[11] Fox R, Crask M, Kim J. Mail Survey Response Rate: A Meta-Analysis of Selected Techniques for Inducing Response. Puble Opinion Quarterly. 1988;52(4):467-91.
[12] Yammarino F, Skinner S, Childers T. Understanding Mail Survey Response Behaviour - A Meta-Analysis. Public Opinion Quarterly. 1991;55(4):613-39.
[13] Harvey L. Factors Affecting Response Rates to Mailed Questionnaires: A Comprehensive Literature review. Journal of the Market Research Society. 1987;29:341-53.
[14] Kanuk L, Berenson C. Mail surveys and response rates: a literature review. Journal of Marketing Research. 1975;12:440-53.
[15] Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, et al. Methods to increase response rates to postal questionnaires. The Cochrane Database of Methodological Reviews. 2003(Issue 4):Art. No.: MR000008. DOI: 10.1002/14651858.MR000008.pub2.
[16] Smeeth L, Fletcher AE. Improving the response rates to questionnaires. Bmj. 2002 May 18;324(7347):1168-9.

[17] McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Steen N, et al. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. Health Technol Assess. 2001;5(31):1-256.
[18] O'Cathain A. Further Analyses Would Make the Review More Helpful - Rapid Response to Edwards et al Increasing Response Rates to Postal Questionnaires: Systematic Review. British Medical Journal. 2002;324.
[19] Iglesias C. Changing the layout of questionnaires increases response rates. Rapid responses for Edwards et al., 324 (7347) 1183. *BMJ* *http://bmjbmjjournalscom/cgi/eletters/324/7347/1183* 2002.
[20] Heberlein T, Baumgartner R. Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature. American Sociological Review. 1978;43(4):447-62.
[21] Ludemann R, Watson DI, Jamieson GG. Influence of follow-up methodology and completeness on apparent clinical outcome of fundoplication. Am J Surg. 2003 Aug;186(2):143-7.
[22] Dorman PJ, Slattery J, Farrell B, Dennis MS, Sandercock PA. A randomised comparison of the EuroQol and Short Form-36 after stroke. United Kingdom collaborators in the International Stroke Trial. BMJ. 1997 Aug 23;315(7106):461.
[23] Booth C, ed. Labour and life of the people of London. 17 volumes. London: Macmillan 1889-1902.
[24] Moser C, Kalton G. Survey methods in social investigation. 2nd ed. Aldershot: Gower 1971.
[25] Toops H. Validating the questionnaire method. Journal of Personnel Research. 1924;2:153-69.
[26] Singer E. Nonresponse bias in household surveys. Public Opinion Quarterly. 2006;70(5):637-45.
[27] Groves R. Nonresponse rates and nonresponse bias in household surveys. Public Opinion Quarterly. 2006;70(5):646-75.
[28] Nonresponse bias in household surveys. Public Opinion Quarterly.70(5).
[29] Bourque LB, Fielder EP. How to conduct self-administered and mail surveys. London: Sage 1995.
[30] Goyder JC. Further evidence on factors affecting response rates to mailed questionnaires. American Sociological Review. 1982;47:550-3.
[31] Layne BH, Thompson DN. Questionnaire page length and return rate. Journal of Social Psychology. 1981;113:291-2.
[32] Eaker S, Bergstrom R, Bergstrom A, Adami HO, Nyren O. Response rate to mailed epidemiologic questionnaires: a population-based randomized trial of variations in design and mailing routines. Am J Epidemiol. 1998 Jan 1;147(1):74-82.
[33] Etter JF, Cucherat M, Perneger TV. Questionnaire color and response rates to mailed surveys. A randomized trial and a meta-analysis. Eval Health Prof. 2002 Jun;25(2):185-99.
[34] Jobber D, Sanderson S. The effects of a prior letter and coloured questionnaire paper on mail survey response rates. J Market Res Soc. 1983;25:339-49.

[35] Matteson M. Type of transmittal letter and questionnaire colour as two variables influencing response rates in a mail survey. Journal of Applied Psychology 1974;59(4):535-6.
[36] Fox R, Crask M, Kim J. Mail Survey Response Rate: A Meta-Analysis of Selected Techniques for Inducing Response. Public Opinion Quarterly. 1988;52(4):467-91.
[37] Dillman D. Mail and telephone surveys: the total design method. New York: Wiley 1978.
[38] LaGarce R, Kuhn L. The effect of visual stimuli on mail survey response rates. Industrial Marketing Management. 1995;24:11-8.
[39] Sudman S, Bradburn N. Asking questions: A practical guide to questionnaire design. San Francisco: Jossey-Bass 1982.
[40] Rasinski K. The efect of question wording on public support for government spending. Public Opinion Quarterly. 1989;53:388-94.
[41] Schuman H, Presser S. Questions and Answers in Attitude Surveys. New York: Academic Press 1981.
[42] Mullner R, Levy P, Byre C, Matthews D. Effects of characteristics of the survey instrument on response rates to a mail survey of community hospitals. . Public Health Reports. 1982;97(5):465-9.
[43] Giles W, Feild H. Effects of amount, format, and location of demographic information on questionnaire return rate and response bias of sensitive and non sensitive items. Personnel Psychology 1978;31:549-59.
[44] Green R, Murphy K, Snyder S. Should demographics be placed at the end or at the beginning of mailed questionnaires? An empirical answer to a persistent methodological question.
. Social Work Research 2000;24(4):237-40.
[45] Willits F, Ke B. Part-whole question order effects. Public Opinion Quarterly. 1995;59:392-403.
[46] Martin J, McConnell J. Mail questionnaire response induction: The effect of four variables on the response of a random sample to a difficult questionnaire. Social Science Quarterly. 1970;51:409-14.
[47] Pourjalali H, Kimbrell J. Effects of four instrumental variables on survey response. . Psychological Reports. 1994;75:895-8.
[48] Roberson M, Sundstrom E. Questionnaire design, return rates and response favourableness in an emplyee atitude questionnaire. Journal of Applied Psychology. 1990;75(35):4-7.
[49] Wagner WG, O'Toole WM. The effects of covering letter format on faculty response rate in mail survey research. Educational and Psychological Research. 1985;5:29-37.
[50] McKillip J, Lockhart D. The effectiveness of cover letter appeals. Journal of Social Psychology. 1984;122(1):85-91.
[51] Childers TL, Pride WM, Ferrell OC. A reassesment of the effects of appeals on response to mail surveys. Journal of Marketing Research. 1980;17:365-70.
[52] Nederhof AJ. Effects of preliminary contacts on volunteering in mail surveys. Perceptual and Motor Skills. 1982;54:1333-4.

[53] Faria AJ, Dickinson JR, Filipic TV. The effect of telephone versus letter pre-notification on mail survey response rate, speed, quality and cost. Journal of the Market Research Society. 1990;32:551-68.
[54] Kahle LR, Sales BD. Personalisation of the outside envelope in mail surveys. Public Opinion Quarterly. 1978;42:547-50.
[55] Linsky AS. Stimulating responses to mailed questionnaires: a review. Public Opinion Quarterly. 1975;39:82-101.
[56] Jones WH, Linda G. Multiple criteria effects in a mail survey experiment. Journal of Marketing Research. 1978;15:280-4.
[57] Groves R, Singer E, Corning A. Leverage-Saliency theory of survey participation. Public Opinion Quarterly. 2000;64:299-308.
[58] Weltzien RT, Jr., McIntyre TJ, Ernst JA, Walsh JA, Parker JK. Cross validation of some psychometric properties of the CSQ and its differential return rate as a function of token financial incentive. Community Ment Health J. 1986 Spring;22(1):49-55.
[59] Trice AD. Maximising participation in surveys: hotel ratings VII. Journal of Social Behaviour and Personality. 1986;1:137-41.
[60] Brennan M. The effect of a monetary incentive on mail survey response rates: new data. Journal of the Market Research Society. 1992;34:173-7.
[61] Little RE, Davis AK. Effectiveness of various methods of contact and reimbursement on response rates of pregnant women to a mail questionnaire. Am J Epidemiol. 1984 Jul;120(1):161-3.
[62] Shaw MJ, Beebe TJ, Jensen HL, Adlis SA. The use of monetary incentives in a community survey: impact on response rates, data quality, and cost. Health Serv Res. 2001 Feb;35(6):1339-46.
[63] Hansen RA. A self-perception interpretation of the effect of monetary and non-monetary incentives on mail survey respondent behaviour. Journal of Marketing Research. 1980;17:77-83.
[64] Halpern SD, Ubel PA, Berlin JA, Asch DA. Randomized trial of 5 dollars versus 10 dollars monetary incentives, envelope size, and candy to increase physician response rates to mailed questionnaires. Med Care. 2002 Sep;40(9):834-9.
[65] Hubbard R, Little EL. Promised contributions to charity and mail survey responses: replication with extension. Public Opinion Quarterly. 1988;52:223-30.
[66] VanGeest JB, Wynia MK, Cummins DS, Wilson IB. Effects of different monetary incentives on the return rate of a national mail survey of physicians. Med Care. 2001 Feb;39(2):197-201.
[67] Church AH. Estimating the effect of incentives on mail survey response rates: A meta-analysis. Public Opinion Quarterly. 1993;57(1):62-79.
[68] Armstrong JS. Monetary incentives in mail surveys. Public Opinion Quarterly. 1975;39:111-16.
[69] Hopkins KD, Gullickson AR. Response rates in survey research: a meta-analysis of the effects of monetary gratuities. Journal of Experimental Education. 1992;61:52-62.

[70] Jobber D, Saunders J. Modeling the effects of pre-paid monetary incentives on mail survey response. Journal of the Operational Research Society. 1988;39(3):365-72.
[71] Goyder J. The Silent Minority. Cambridge: Polity Press 1987.
[72] Brehm JW. A Theory of Psychological Reactance. New York: Academic Press 1965.
[73] Biner PM. Effects of cover letter appeal and monetary incentives on survey response: A reactance theory application. Basic and Applied Social Psychology. 1988;9(2):99-106.
[74] Biner PM, Barton DL. Justifying the enclosure of monetary incentives in mail survey cover letters. Psychology and Marketing. 1990;7(3):153-62.
[75] Bowling A. Research Methods in Health. Buckingham, Philadelphia: Open University Press 1997:234-235.
[76] Festinger L. A theory of cognitive dissonance. Stanford: Stanford University Press 1957.
[77] Furse DH, Stewart DW. Manipulating dissonance to improve mail survey response. Psychology and Marketing. 1984;1(2):79-94.
[78] Hackler JC, Bourgette P. Dollars, dissonance and survey returns. Public Opinion Quarterly. 1973;37(2):276-81.
[79] Furse DH, Stewart DW, Rados DL. Effects of foot-in-the-door, cash incentives and followups on survey response. Journal of Marketing Research. 1981;18:473-8.
[80] Adams JS. Towards an understanding of inequity. Journal of Abnormal Social Psychology. 1963;67:422-36.
[81] Katz D. The functional approach to the study of attitudes. Public Opinion Quarterly. 1960;24(2):163-204.
[82] Jones WH, Lang JR. Sample composition bias and response bias in a mail survey: a comparison of inducement methods. Journal of Marketing Research. 1980;17:69-76.
[83] Cialdini RB, Ascani K. Test of concession procedure for inducing verbal, behavioural and further compliance with a request to give blood. Journal of Applied Psychology. 1976;61:295-300.
[84] Cialdini RB, Vincent JE, Lewis SK, Catalan J, Wheeler D, Darby BL. Low-ball procedure for producing compliance: Commitment then cost. Journal of Personality and Social Psychology. 1978;31:206-15.
[85] Cialdini R. Influence: The New Psychology of Modern Persuasion. New York: Quill 1984.
[86] Groves R, Cialdini R, Couper M. Understanding the decision to participate in a survey. Public Opinion Quarterly. 1992;56:475-95.
[87] Groves R, Couper M. Nonresponse in Household Interview Surveys. New York: Wiley-Interscience 1998.
[88] Henzlova MJ, Blackburn GH, Bradley EJ, Rogers WJ. Patient perception of a long-term clinical trial: Experience using a close-out questionnaire in the studies of left ventricular dysfunction (SOLVD) trial. Control Clinical Trials. 1994;15:284-93.

[89] Schweitzer M, Asch DA. Timing payments to subjects of mail surveys: cost-effectiveness and bias. J Clin Epidemiol. 1995 Nov;48(11):1325-9.
[90] Brehm JW, Cole A. Effect of a favour which reduces freedom. Journal of Personality and Social Psychology. 1966;3:420-26.
[91] Baumeister RF. A self-presentational view of social phenomena. Psychological Bulletin. 1982;91:3-26.
[92] COREC. Guidelines for Researchers. Information Sheets and Consent Forms. Nov 2005 [cited; version 2:[Available from: http://www.corec.org.uk/applicants/help/guidance.htm#consent
[93] Dovidio JF. Helping and altruism: An empirical and conceptual overview. In: Berkowitz L, ed. *Advances in Experimental Social Psychology*. New York: Academic Press 1984:361-427.
[94] Clark MS, Waddell BA. Effects of moods on thoughts about helping, attraction and information acquisition. Social Psychology Quarterly. 1983;46:31-5.
[95] Pope JE, Tingey DP, Arnold JM, Hong P, Ouimet JM, Krizova A. Are subjects satisfied with the informed consent process? A survey of research participants. J Rheumatol. 2003;30(4):815-24.
[96] Groves R, Presser S, Dipko S. The role of topic of interest in survey participation decisions. Public Opinion Quarterly. 2004;68(1):2-31.
[97] Singer E. Public reactions to some ethical issues of social research: attitudes and behaviour. Journal of Consumer Research. 1984;11:501-9.
[98] Sharp LM, Frankel J. Respondent burden: a test of some common assumptions. Public Opinion Quarterly. 1983;47:36-53.
[99] DeMaio TJ. Refusals: who, where and why? Public Opinion Quarterly. 1980;44:223-33.
[100] Goyder J. Attitudes and actions in response behaviour. *The Silent Minority*. Cambridge: Polity Press 1987.
[101] Biggar RJ, Melbye M. Responses to anonymous questionnaires concerning sexual behavior: a method to examine potential biases. Am J Public Health. 1992 Nov;82(11):1506-12.
[102] Pocock SJ. Clinical Trials - A Practical Approach. Chichester: John Wiley and Sons 1983.
[103] Kanuk L. Mail Surveys and Response Rates: A Literature Review. Journal of Marketing Research. 1975;12:440-53.
[104] Trundle H. Physiotherapy: The contribution. In: Skinner D, Swain A, Peyton R et al eds Cambridge textbook of accident and emergency medicine. Cambridge: Cambridge University Press 1997.
[105] Pijnenburg AC, Van Dijk CN, Bossuyt PM, Marti RK. Treatment of ruptures of the lateral ankle ligaments: a meta-analysis.[comment]. Journal of Bone & Joint Surgery - American Volume. 2000;82(6):761-73.
[106] Struijs P, Kerkhoffs G. Ankle sprain. Clin Evid. 2002 Jun(7):945-53.
[107] Keeman JN. Commentaar Enkelspecial. Reuma en Trauma. 1990;1:34-5.
[108] Anandacoomarasamy A, Barnsley L. Long term outcomes of inversion ankle injuries. Br J Sports Med. 2005;39.
[109] Konradsen L, Bech L, Ehrenbjerg M, Nickelsen T. Seven years follow-up after ankle inversion trauma. Scand J Med Sci Sports. 2002 Jun;12(3):129-35.

[110] Zoch C, Fialka-Moser V, Quittan M. Rehabilitation pf ligamentous ankle injuries: a review of recent studies. Br J Sports Med. 2003;37:291-5.
[111] Cooke MW, Lamb SE, Marsh J, Dale J. A survey of current consultant practice of treatment of severe ankle sprains in emergency departments in the United Kingdom. Emergency medicine journal : EMJ. 2003 Nov;20(6):505-7.
[112] Roos EM, Brandsson S, Karlsson J. Validation of the Foot and Ankle Outcome Score for Ankle Ligament Reconstruction. Foot and Ankle International. 2001 Oct;22(10):788-94.
[113] Bergner M, Bobbitt R, Carter W, Gilson BS. The sickness impact profile: development and final revision of a health status measure. Medical Care. 1981;19:787-805.
[114] Ware JE, Kosinski M, Keller SD. A 12-item Short-Form Health Survey. Construction of scales and preliminary tests of reliability and validity. Medical Care. 1996;34:220-33.
[115] EuroQolGroup.
. EuroQol-A new facility for the measurement of health-related quality of life. Health Policy. 1990;16:199-208.
[116] Czajar R, Blair J. Designing surveys: A guide to decisions and procedures. Thousand Oaks, CA: Pine Forge Press 1995.
[117] Sutherland HJ, Beaton M, Mazer R, Kriukov V, Boyd NF. A randomized trial of the total design method for the postal follow-up of women in a cancer prevention trial. Eur J Cancer Prev. 1996 Jun;5(3):165-8.
[118] Robins L. The reluctant respondent. Public Opinion Quarterly. 1963;27:276-86.
[119] O'Neil MJ. Estimating the non-response bias due to refusals in telephone surveys. Public Opinion Quarterly. 1979;43:218-32.
[120] Smith T. The hidden 25 per cent: an analysis of non-response on the 1980 General Social Survey. Public Opinion Quarterly. 1985;47:386-404.
[121] Brakenbury PH, Kotowski J. A comparative study of the management of ankle sprains. British Journal of Clinical Practice. 1983;37(5):181-5.
[122] Brooks SC, Potter BT, Rainey JB. Treatment for partial tears of the lateral ligament of the ankle: a prospective trial. Br Med J (Clin Res Ed). 1981 Feb 21;282(6264):606-7.
[123] Eiff MP, Smith AT, Smith GE. Early mobilization versus immobilization in the treatment of lateral ankle sprains. Am J Sports Med. 1994 Jan-Feb;22(1):83-8.
[124] Korkala O, Rusanen M, Jokipii P, Kytomaa J, Avikainen V. A prospective study of the treatment of severe tears of the lateral ligament of the ankle. International Orthopaedics. 1987;11(1):13-7.
[125] Munk B, Holm-Christensen K, Lind T. Long-term outcome after ruptured lateral ankle ligaments. A prospective study of three different treatments in 79 patients with 11-year follow-up. Acta Orthop Scand. 1995 Oct;66(5):452-4.
[126] Roycroft S. Treatment of inversion injuries of the ankle by early active management. Physiotherapy. 1983;69(10):355-6.
[127] Olkin I. Statistical and theoretical considerations in meta-analysis. J Clin Epidemiol. 1995;48(1):133-46.

[128] Egger ME, Davey Smith G, Phillips AN. Meta-analysis: Principles and procedures. BMJ. 1997;315:1533-7.
[129] Feinstein AR. Meta-analysis: Statistical alchemy for the 21st century. J Clin Epidemiol. 1995;48(1):71-9.
[130] Thompson S, Pocock SJ. Can meta-analysis be trusted? Lancet. 1991;338:1127-30.
[131] Egger M, Davey Smith G, O'Rourke K. Rationale, potentials and promise of systematic reviews. In: Egger M, Davey Smith G, Altman D, eds. *Systematic Reviews in Health Care: Meta-analysis in context*. 2nd ed. London: BMJ 2001.
[132] Moher D, Cook DJ, Eastwood S, Olkin I, Drummond R, Stroup DF, et al. Improving the quality of reports of maeta-analyses of randomised controlled trials: the QUOROM statement. Lancet. 1999;354:1896-900.
[133] BMJ Cite Track Alert. [cited; Available from: http://bmj.com/cgi/alerts/ctmain
[134] BioMail Medline Search. [website] 2000-2003 [cited; version 0.72pre:[Available from: http://biomail.sourceforge.net/biomail
[135] McCaul LA, Cooper PG. Techniques to increase the response rate in follow-up studies: results of a pilot test. Int J Rehabil Res. 1979 Dec;2(4):465-70.
[136] The Chambers Dictionary: Chambers Harrap Publishers Ltd 1998.
[137] Higgins J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [Updated May 2005]. *The Cochrane Library, Issue 3, 2005*. Chichester, UK: John Wiley and Sons Ltd.
[138] Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. Jama. 1995 Feb 1;273(5):408-12.
[139] Berlin JA, Rennie D. Measuring the quality of trials: the quality of quality scales. Jama. 1999 Sep 15;282(11):1083-5.
[140] Badgley RF. An assessment of research methods reported in 103 scientific articles from two Canadian medical journals. Can Med Assoc J. 1961 Jul 29;85:246-50.
[141] Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. Control Clin Trials. 1995 Feb;16(1):62-73.
[142] Chalmers TC, Smith H, Jr., Blackburn B, Silverman B, Schroeder B, Reitman D, et al. A method for assessing the quality of a randomized control trial. Control Clin Trials. 1981 May;2(1):31-49.
[143] Jadad AR. Assessing the quality of RCTs: why, what, how, and by whom? In: Jadad AR, ed. *Randomised Controlled Trials - A user's guide*. London: BMJ Books 1998.
[144] Juni P, Altman D, Egger M. Chapter 5 Assessing the quality of randomised controlled trials. In: Egger M, Smith G, Altmen D, eds. *Systematic Reviews in Health Care - Meta-analysis in context*. Second ed. London: BMJ publishing group 2001:87-108.
[145] Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 1996 Feb;17(1):1-12.

[146] Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomised clinical trials for conducting systematic reviews developed by Delphi consensus. J Clin Epidemiol. 1998 Dec;51(12):1235-41.
[147] Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. Jama. 1999 Sep 15;282(11):1054-60.
[148] Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. BMJ. 1996 Sep 7;313(7057):570-1.
[149] Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, et al. Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses. Health Technol Assess. 1999;3(12):i-iv, 1-98.
[150] Egger M, Davey Smith G. Principles of and procedures for systematic reviews. In: Egger M, Davey Smith G, Altman D, eds. *Systematic Reviews in Health Care - Meta-analysis in context*. 2nd ed. London: BMJ 2001.
[151] Davies H, Crombie K, Tavakoli M. When can odds ratios mislead? BMJ. 1998;316:989-91.
[152] Deeks J. When can odds ratios mislead? BMJ. 1998;317:1155.
[153] Altman D, Deeks J. Odds ratios should be avoided when events are common. BMJ. 1998;317:1318.
[154] Higgins J, Thompson S, Deeks J, Altman D. Measuring inconsistency in meta-analysis. BMJ. 2003;327:557-60.
[155] Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002 Jun 15;21(11):1539-58.
[156] Glasziou P, Irwig L, Bain C, Colditz G. Systematic reviews in health care - A practical guide. Cambridge: Cambridge University Press 2001.
[157] DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials. 1986;7:177-88.
[158] Chalmers TC, Celano P, Sacks HS, Smith H. Bias in treatment assignment in controlled clinical trials. N Engl J Med. 1983;309:1358-61.
[159] Iglesias C, Torgerson D. Does length of questionnaire matter? A randomised trial of response rates to a mailed questionnaire. J Health Serv Res Policy. 2000 Oct;5(4):219-21.
[160] Dunn KM, Jordan K, Croft PR. Does questionnaire structure influence response in postal surveys? J Clin Epidemiol. 2003 Jan;56(1):10-6.
[161] Jenkinson C, Coulter A, Reeves R, Bruster S, Richards N. Properties of the Picker Patient Experience questionnaire in a randomised controlled trial of long versus short form survey instruments. J Public Health Med. 2003 Sep;25(3):197-201.
[162] Jones R, Zhou M, Yates WR. Improving return rates for health-care outcome. Psychol Rep. 2000 Oct;87(2):639-42.
[163] Ward J, Boyle C, Long D, Ovadia C. Patient surveys in general practice: a randomised trial of an instant lottery ticket to increase return rate. Aust Fam Physician. 1996 Jan;Suppl 1:S19-20.
[164] Leigh Brown AP, Lawrie HE, Kennedy AD, Webb JA, Torgerson DJ, Grant AM. Cost effectiveness of a prize draw on response to a postal questionnaire:

results of a randomised trial among orthopaedic outpatients in Edinburgh. J Epidemiol Community Health. 1997 Aug;51(4):463-4.
[165] Evans BR, Peterson BL, Demark-Wahnefried W. No Difference in Response Rate to a Mailed Survey among Prostate Cancer Survivors Using Conditional versus Unconditional Incentives. Cancer Epidemiol Biomarkers Prev. 2004 Feb;13(2):277-8.
[166] McColl E, Eccles MP, Rousseau NS, Steen IN, Parkin DW, Grimshaw JM. From the generic to the condition-specific?: Instrument order effects in Quality of Life Assessment. Med Care. 2003 Jul;41(7):777-90.
[167] Tai SS, Nazareth I, Haines A, Jowett C. A randomized trial of the impact of telephone and recorded delivery reminders on the response rate to research questionnaires. J Public Health Med. 1997 Jun;19(2):219-21.
[168] Salim Silva M, Smith WT, Bammer G. Telephone reminders are a cost effective way to improve responses in postal health surveys. J Epidemiol Community Health. 2002 Feb;56(2):115-8.
[169] Parkes R, Kreiger N, James B, Johnson KC. Effects on subject response of information brochures and small cash incentives in a mail-based case-control study. Ann Epidemiol. 2000 Feb;10(2):117-24.
[170] Erwin W, Wheelright L. Improving mail survey response rates through the use of monetary incentive. Journal of Mental Health Counseling. 2002;24(3):247-9.
[171] Harrison RA, Holt D, Elton PJ. Do postage-stamps increase response rates to postal surveys? A randomized controlled trial. Int J Epidemiol. 2002 Aug;31(4):872-4.
[172] Kalantar JS, Talley NJ. The effects of lottery incentive and length of questionnaire on health survey response rates: a randomized study. J Clin Epidemiol. 1999 Nov;52(11):1117-22.
[173] Roberts H, Pearson JC, Dengler R. Impact of a postcard versus a questionnaire as a first reminder in a postal lifestyle survey. J Epidemiol Community Health. 1993 Aug;47(4):334-5.
[174] Boyce S, Quigley M, Campbell S. Management of ankle sprains: a randomised controlled trial of the treatment of inversion injuries using an elastic support bandage or an Aircast ankle brace. Br J Sports Med. 2005;39:91-6.
[175] Leanderson J, Bergqvist M, Rolf C, Westblad P, Wigelius-Roovers S, Wredmark T. Early influence of an ankle sprain on objective measures of ankle joint function. A prospective randomised study of ankle brace treatment. Knee Surg Sports Traumatol Arthrosc. 1999;7(1):51-8.
[176] Avci S, Sayli U. Comparison of the results of short-term rigid and semi-rigid cast immobilization for the treatment of grade 3 inversion injuries of the ankle. Injury. 1998 Oct;29(8):581-4.
[177] Cartwright A. Who responds to postal questionnaires? J Epidemiol Community Health. 1986 Sep;40(3):267-73.
[178] Michielutte R, Bahnson J, Dignan MB, Schroeder EM. The use of illustrations and narrative text style to improve readability of a health education brochure. J Cancer Educ. 1992;7(3):251-60.

[179] Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. Control Clin Trials. 2003 Apr;24(2):182-99.
[180] Bowling A. Research Methods in Health; Investigating Health and Health Services. Buckingham: Open University Press 1997.
[181] Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. Bmj. 2002 May 18;324(7347):1193-4.
[182] Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. Pain. 2001 Apr;91(3):277-85.
[183] Martyn KK, Belli RF. Retrospective data collection using event history calendars. Nurs Res. 2002 Jul-Aug;51(4):270-4.
[184] Campbell M, Machin D. Medical Statistics - A Common Sense Approach. Third ed. Chichester: Wiley and Sons 1990.
[185] SPSS for Windows. 13.0.1 ed. Chicago: SPSS Inc 2004.
[186] Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. The Lancet. 2001;357:1191-4.
[187] Altman D. Practical Statistics for Medical Research. London: Chapman and Hall 1991.
[188] Peat J, Elliot E, Baur L, Keena V. Scientific Writing - Easy when you know how. London: BMJ Books 2002.
[189] Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: Wiley 1989.
[190] Mickey J, Greenland S. A study of the impact of confounder-selection criteria on effect estimation. American Journal of Epidemiology. 1989;129:125-37.
[191] Nakash RA, Hutton JL, Jorstad-Stein EC, Gates S, Lamb SE. Maximising response to postal questionnaires - A systematic review of randomised trials in health research. BMC Medical Research Methodology. 2006;6(5).
[192] Wisniewski SR, Leon AC, Otto MW, Trivedi MH. Prevention of missing data in clinical research studies. Biological Psychiatry. 2005;59(11):997-1000.
[193] Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomised controlled trials in major medical journals. Clinical Trials. 2004;1:368-76.
[194] Sprague S, Leece P, Bhandari III M, Tornetta P, Schemitsch E, Swiontkowski MF, et al. Limiting loss to follow-up in a multicentre randomised trial in orthopaedic surgery. Controlled Clinical Trials. 2003(24):719-25.
[195] Edwards P, Cooper R, Wentz R, Fernandes J. Methods to influence the completeness of response to self-administered questionnaires (Protocol). Cochrane Database of Methodology Reviews. 2002(4):Art. No.: MR000015. DOI: 10.1002/14651858.MR000015.
[196] Tourangeau R, Rips L, Rasinski K. The Psychology of Survey Response. Cambridge: Cambridge University Press 2000.
[197] Sim J, Wright C. Research in Health Care: Concepts, Designs and Methods. Cheltenham: Stanley Thornes Ltd 2000.

[198] Patton MQ. Qualitative evaluation and research methods. London: Sage 1990.
[199] Turner A, Williams B, Barlow J. The imapct of an arthritis self-management programme on psychsocial wellbeing. Health Education. 2002;102(3).
[200] Hancock B. Trent Focus for Research and Development in Primary Health Care: An Introduction to Qualitative Research
Trent Focus 1998.
[201] Erlandson DA, Harris EL, Skipper BL, Allen SD. Doing Naturalistic Inquiry: A Guide to Methods. Newbury Park, California: Sage 1993.
[202] Flick U. An Introduction to qualitative Research. London: Sage 2002.
[203] Lacey A, Luff D. Trent Focus for Research and Development in Primary Health Care: Qualitative Data Analysis: Trent Focus 2001.
[204] Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman, Burgess, eds. *Analysing Qualitative Data*. London: Routledge 1994:173-94.
[205] Snape D, Spencer L. The Foundations of Qualitative Research. In: Ritchie J, Lewis J, eds. *Qualitative Research Practice*. London: Sage 2003.
[206] Silverman D. Doing Qualitative Research. London: Sage 2000.
[207] Green J, Thorogood N. Qualitative Methods for Health Research. London: Sage 2004.
[208] Hammersley M. What's Wrong with Ethnograpy? London: Routledge 1992.
[209] Spencer L, Ritchie J, O'Connor W. Analysis: Practices, Principles and Processes. In: Ritchie J, Lewis J, eds. *Qualitative Research Practice*. London: Sage 2003.
[210] White C, Woodfield K, Ritchie J. Reporting and presenting qualitative data. In: Ritchie J, Lewis J, eds. *Qualitative Research Practice*. London: Sage 2003:287-320.
[211] Elbourne D, Snowdon C, Garcia J. Informed consent. Patients may not understand concept of clinical trials. BMJ. 1997;315(7102):248-9.
[212] Snowdon C, Garcia J, Elbourne D. Making sense of randomisation; responses of parents of critically ill babies to random allocation of treatment in a clinical trial. Soc Sci Med. 1997;45(9):1337-55.
[213] Robinson EJ, Kerr CEP, Stevens AJ, Lilford RJ, Braunholtz DA, Edwards SJ, et al. Lay public's understanding of equipoise and randomisation in randomised controlled trials. Health Technol Assess. 2005;9(8).
[214] Flory J, Emanuel E. Interventions to improve research participants' understanding in informed consent for research. JAMA. 2004;292(13):1593-601.
[215] Mowen J, Cialdini RB. On implementing the door-in-the face compliance technique in a business context. Journal of Marketing Research. 1980;17:253-58.
[216] Torgerson D, Klaber-Moffett J, Russell I. Patient preferences in randomised trials: threat or opportunity? J Health Serv Res Policy. 1996;1:194-7.
[217] Brewin CR, Bradley C. Patient preferences and randomised clinical trials. BMJ. 1989;299:313-15.
[218] King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, et al. Conceptual framework and systematic review of the effects of participants' and

professionals' preferences in randomised controlled trials. Health Technol Assess. 2005;9(35).
[219] Kenyon S, Pike K, Jones D, Taylor D, Salt A, Marlow N, et al. The effect of a monetary incentive on return of a postal health and development questionnaire: a randomised trial. BMC Health Services Research. 2005 August 2005;5(55).
[220] Bowling A. Research Methods in Health. Buckingham: Open University Press 1997.
[221] Glaser B, Strauss A. The Discovery of Grounded Theory. London: Weidenfield and Nicolson 1967.
[222] Bryman A. Quantity and Quality in Social Research. London: Unwin Hyman 1988.
[223] Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P. Qualititive research methods in health technology assessmemt. Health Technol Assessment. 1998;2(16).
[224] Lewis J, Ritchie J. Generalising from Qualitative Research. In: Ritchie J, Lewis J, eds. *Qualitative Research Practice*. London: Sage 2003:263-86.
[225] Marshall C, Rossman GB. Designing Qualitative Research. 3rd ed. Thousand Oaks: Sage 1999.
[226] Seale C. The Quality of Qualitative Research. Oxford: Blackwell 1999.
[227] Miles MB, Huberman AM. Qualitative Data Analysis: An Expanded Sourcebook. London: Sage 1994.
[228] Rizzo L, Kalton G, Brick MJ. A comparison of some weighting adjustment methods for panel nonresponse. Survey Methodology. 1996;22:143-53.
[229] Zabel JE. An analysis of attrition in the panel study of income dynamics and the survey of income and program participation with an application to a model of labor market behavior. Journal of Human Resources. 1998;33:2479–506.
[230] Moore DL, Tarnai J. Evaluating non-response error in mail surveys. In: Groves RM, Dillman DA, Eltinge JL, Little RA, eds. *Survey Nonresponse*. New York: Wiley and Sons 2002.
[231] Siemiatycki J, Campbell S. Nonresponse bias and early versus all responders in mail and telephone surveys. American Journal of Epidemiology. 1984;120(2):291-301.
[232] Scott C. Research on mail surveys. Journal of the Royal Statistical Society. 1961;124:143-205.
[233] Rosenthal R, Rosnow R. The Volunteer Subject. New York: Wiley 1975.
[234] Dunkleberg WC, Day GS. Nonresponse bias and callbacks in sample surveys. Journal of Marketing Research. 1973;10:160-8.
[235] Hoffman SC, Burke AE, Helzlsouer KJ, Comstock GW. Controlled trial of the effect of length, incentives, and follow-up techniques on response to a mailed questionnaire. Am J Epidemiol. 1998 Nov 15;148(10):1007-11.
[236] Brogger J, Bakke P, Eide GE, Gulsvik A. Contribution of follow-up of nonresponders to prevalence and risk estimates: a Norwegian respiratory health survey. Am J Epidemiol. 2003 Mar 15;157(6):558-66.
[237] Kotaniemi JT, Hassi J, Kataja M, Jonsson E, Laitinen LA, Sovijarvi AR, et al. Does non-responder bias have a significant effect on the results in a postal questionnaire study? Eur J Epidemiol. 2001;17(9):809-17.

[238] Gannon M, Northern J, Carroll S. Characteristics of nonrespondents among workers. Journal of Applied Psychology. 1971;55:586-8.
[239] Filion F. Estimating bias due to nonresponse in mail surveys. Public Opinion Quarterly. 1975;39:482-92.
[240] Sharp L, Cochran C, Cotton SC, Gray NM, Gallagher ME. Enclosing a pen with a postal questionnaire can significantly increase response rate. Journal of Clinical Epidemiology. 2006;59:747-54.
[241] Victor CR. Some methodological aspects of using postal questionnaires with the elderly. Arch Gerontol Geriatr. 1988 Jun;7(2):163-72.
[242] Smith T. The hidden 25 per cent: an analysis of non-response on the 1980 General Social Survey. Public Opinion Quarterly. 1983;47:386-404.
[243] Hawkins DF. Estimation of non-response bias. Sociological Methods and Research. 1975;3:461-88.
[244] Barnes R, Birch F. The census as an aid in estimating the characteristics of non-response in the GHS. New Methodology Series. 1975;NM1, London: Office of Population Censuses and Surveys
[245] Goudy WJ. Nonresponse effects on relationships between variables. Public Opinion Quarterly. 1976;40:360-9.
[246] Purdie DM, Dunne MP, Boyle FM, Cook MD, Najman JM. Health and demographic characteristics of respondents in an Australian national sexuality survey: comparison with population norms. Journal of Epidemiology and Community Health. 2002;56:748-53.
[247] Steeh CG. Trends in non-response rates, 1952-1979. Public Opinion Quarterly. 1981;45:40-57.
[248] Reuss CF. Differences between persons responding and not responding to a mailed questionnaire. American Sociological Review. 1943;8(4):433-8.
[249] Bakke P, Gulsvik A, Lilleng P, Overa O, Hanoa R, Eide GE. Postal survey on airborne occupational exposure and respiratory disorders in Norway: Causes and consequences of non-response. Journal of Epidemiology and Community Health. 1990;44:316-20.
[250] Locker D, Wiggins D, Sittampalam Y, Patrick DL. Estimating the prevalence of disability in the community: The influence of sample design and response bias. Journal of Epidemiology and Community Health. 1981;35:208-12.
[251] Criqui MH, Barrett-Connor E, Austin M. Differences between respondents and non-respondents in a population-based cardiovascular disease study. American Journal of Epidemiology. 1978;108(5):367-72.
[252] Ronmark E, Lundqvist A, Lundback B, Nystrom L. Non-responders to a postal questionnaire on respiratory symptoms and diseases. European Journal of Epidemiology. 1999;15:293-9.
[253] Norquist BM, Goldberg BA, Matsen FA. Challenges in evaluating patients lost to follow-up in a clinical trial of rotator cuff tears. Journal of Bone and Joint Surgery America. 2000;82(6):838-42.
[254] Parker C, Dewey M. Assessing research outcomes by postal questionnaire with telephone follow-up. TOTAL Study Group. Trial of Occupational Therapy and Leisure. Int J Epidemiol. 2000 Dec;29(6):1065-9.

[255] Lin I-F, Schaeffer NC. Using survey participants to estimate the impact of nonparticipation. Public Opinion Quarterly. 1995;59(2):236-58.
[256] Childers TL, Skinner SJ. Toward a conceptualisation of mail survey response behaviour. Psychology and Marketing. 1996;13(2):185-209.
[257] Curtin R, Presser S, Singer E. The effects of response rate changes on the index of consumer sentiment. Public Opinion Quarterly. 2005;64:13-28.

Appendices

Appendix 1 – CAST Background information questionnaire

Background information

These questions ask for some background details about yourself

Q1. Age: ☐☐ Years

Q2. Sex: Male ☐ Female ☐

Q3. Ethnic Group: **(Please tick one box)**

1. White ☐
2. Black-Caribbean ☐
3. Black-African ☐
4. Black-Other ☐
5. Indian ☐
6. Pakistani ☐
7. Bangladeshi ☐
8. Chinese ☐
9. Other (Please specify) ☐

Q4. What is your first language? (Please tick one box)

1. English ☐
2. Other European ☐
3. Gujarati ☐
4. Hindi ☐
5. Punjabi ☐
6. Urdu ☐
7. Bengali ☐
8. Other (please specify) ☐

Will you be able to fill in questionnaires in English? YES/NO

Q5. Employment Status:

An important part of the study is to determine how much your ankle injury has affected you in terms of days off work. This is why the next question asks about your employment.

5.1 Are you currently employed?
(If you are a full-time student but also work, complete this section and also tick question **5.6** on page 3)

Yes - part time ☐

Yes - full time ☐

No (**go to Q5.2 on page 3**) ☐

a) Is this employment Paid ☐

Unpaid ☐

b) How many hours a week do you work? Less then 10 ☐

10-25 ☐

25-40 ☐

More than 40 ☐

c) Which of the following categories do you think best describes your employment?

Unskilled manual ☐

Skilled manual ☐

Unskilled non-manual ☐

Skilled non-manual ☐

Professional ☐ Please describe: ----------

Other ☐ Please describe: ----------

Decline to answer ☐

If you are **not** currently employed which of the following applies to you:

5.2 Retired ☐

5.3 At home and not looking for paid employment (eg looking after home, family or others) ☐

5.4 Unable to work due to illness or disability ☐

5.5 Unemployed and looking for work ☐

5.6 In full time education ☐

5.7 Other (please specify)---------- ☐

Q6. What is the highest qualification you have achieved?

CSE (or equivalent) ☐

O-Level/GCSE (or equivalent) ☐

A-Level (or equivalent) ☐

Degree (or equivalent) ☐

Higher Degree (or equivalent) ☐

Other (Please specify)..... ☐

Q7.During your usual daily routine (eg work, caring for others, daily activities), approximately how much time do you spend:

a) On your feet? Most of the day ☐

More than 4 hours a day ☐

Less than 4 hours a day ☐

Not much time – mostly sitting ☐

b) Driving? Most of the day ☐

More than 4 hours a day ☐

Less than 4 hours a day ☐

Usually just to/from work ☐

Don't drive ☐

Q8. Are you currently taking any medication for pain or inflammation?

Only since ankle injury ☐

Prior to injury for a separate condition ☐

No ☐

Did not answer ☐

Q9. Which of the following activities do you participate in: (before injuring your ankle)

	More than once weekly	**Less than** once weekly	Never
1. Swimming	☐	☐	☐
2. Weight Training	☐	☐	☐
3. Aerobics/keep-fit	☐	☐	☐
4. Cycling	☐	☐	☐
5. Jogging/running	☐	☐	☐
6. Team sport (eg football, rugby, hockey, netball)	☐	☐	☐
7. Racquet sport (eg tennis, squash, badminton)	☐	☐	☐
8. Yoga	☐	☐	☐
9. Athletics	☐	☐	☐
10. Walking (2 miles or more)	☐	☐	☐
11. Heavy DIY, housework, gardening	☐	☐	☐

Other sports or exercise (please specify) ----------

Q10. How tall are you? ___ feet ___ inches *or* ____cms

Q11. How much do you weigh? ___ stone ___pounds *or*____kgs

Q12. Pain:

a) **Before your injury**, did you **usually** have any pain in your ankle?

YES ☐

NO ☐ (go to Q13)

b) If 'YES' when did you get this pain?

During exercise ☐

Walking on even ground ☐

Walking on uneven ground ☐

Constantly ☐

Other (please specify) .....

c) How often did you experience this pain?

Never ☐

Monthly ☐

Weekly ☐

Daily ☐

Always ☐

Q13. Instability

a) Before your injury did you **usually** have any feelings of instability in your ankle as if you might 'go over on it' or it would let you down?

YES ☐

NO ☐ (go to Q14)

b) If 'YES' would you describe these feelings as:

☐ Mild ☐ Moderate ☐ Severe

c) How frequently did you experience these feelings?

☐ Rarely ☐ Sometimes ☐ Frequently ☐ Always

Q14. Previous injuries:

a) Apart from your current injury, have you sprained or twisted your ankle before?

YES ☐ NO ☐ (go to question 15)

b) If 'YES' how many times?

☐ ☐

Once or twice 3 times or more

c) When was the last time you injured this ankle?

☐ ☐

Less than 1 year ago More than 1 year ago

d) Have you needed to attend A&E for any previous injury to this ankle?

YES ☐ NO ☐

Q15. Weight bearing:

Using the weighing scales, whilst sitting in a chair, how much weight are you able to put through your ankle at the moment?

_____ kgs

Appendix 2 – CAST Outcome measure questionnaire

Important…….

Before you start answering the questionnaire it is important for us to find out if you have unfortunately injured your ankle again since you completed the last questionnaire 8 weeks ago

Please answer the following question:

Have you had another injury (within the last 8 weeks) to the same ankle?

YES ☐ ……Please see below*

NO ☐ ……Please go on and complete the questionnaire

* If 'Yes', approximately how long after your injury 12 weeks ago did this new injury occur?

……

Did you need to return to A&E for your new injury?

Yes ☐ No ☐

Please go on and complete the questionnaire taking into account your new injury.

Foot and Ankle Outcome Score

These questions ask about how your ankle has felt and how well you are able to do your usual activities **in the last week**. Answer every question by ticking the appropriate box, only one box for each question
If you are unsure about how to answer a question, please give the best answer you can.

Please tick **one** box for each question

Symptoms
These questions should be answered thinking of your foot/ankle symptoms **during the last week**.

Q1. Do you have swelling in your foot/ankle?

Never	**Rarely**	**Sometimes**	**Often**	**Always**
☐	☐	☐	☐	☐

Q2. Do you feel grinding, hear clicking or any other type of noise when your foot/ankle moves?

Never	**Rarely**	**Sometimes**	**Often**	**Always**
☐	☐	☐	☐	☐

Q3. Does your foot/ankle catch or lock when moving?

Never	**Rarely**	**Sometimes**	**Often**	**Always**
☐	☐	☐	☐	☐

Q4. Can you straighten your foot/ankle fully? (point toes away from you)

Always	**Often**	**Sometimes**	**Rarely**	**Never**
☐	☐	☐	☐	☐

Q5. Can you bend your foot/ankle fully? (pull toes up towards you)

Always	**Often**	**Sometimes**	**Rarely**	**Never**
☐	☐	☐	☐	☐

Stiffness

The following questions concern the amount of joint stiffness you have experienced **in the last week** in your foot/ankle. Stiffness is a sensation of restriction or slowness in the ease with which you move your joints.

Q6. How severe is your foot/ankle stiffness after first wakening in the morning?

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q7. How severe is your foot/ankle stiffness after sitting, lying or resting later in the day?

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Pain

Q8. How often do you usually experience foot/ankle pain?

Never	Monthly	Weekly	Daily	Always
☐	☐	☐	☐	☐

What amount of foot/ankle pain have you experienced **in the last week** during the following activities?

Q9. Twisting/pivoting on your foot/ankle

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q10. Straightening foot/ankle fully (pointing toes away from you)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q11. Bending foot/ankle fully (pulling toes up towards you)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q12. Walking on flat surface

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q13. Going up or down stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q14. At night while in bed

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q15. Sitting or lying

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q16. Standing upright

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Function, daily living

The following questions concern your physical function. By this we mean your ability to move around and to look after yourself. For each of the following activities please indicate the degree of difficulty you have experienced **in the last week** due to your ankle injury.

Q17. Descending stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q18. Ascending stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q19. Rising from sitting

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q20. Standing

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q21. Bending to floor/pick up an object

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q22. Walking on flat surface

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q23. Getting in/out of car

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q24. Going shopping

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q25. Putting on socks/stockings

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q26. Rising from bed

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q27. Taking off socks/stockings

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q28. Lying in bed (turning over, maintaining knee position)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q29. Getting in/out of bath

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q30. Sitting

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q31. Getting on/off toilet

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q32. Heavy domestic duties (moving heavy boxes, scrubbing floors, etc)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q33. Light domestic duties (cooking, dusting, etc)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Function, sports and recreational activities

The following questions concern your physical function when being active on a higher level. The questions should be answered thinking of what degree of difficulty you have experienced **in the last week** due to your ankle injury.

Q34. Squatting

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q35. Running

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q36. Jumping

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Q37. Twisting/pivoting on your injured foot/ankle

None ☐ **Mild** ☐ **Moderate** ☐ **Severe** ☐ **Extreme** ☐

Q38. Kneeling

None ☐ **Mild** ☐ **Moderate** ☐ **Severe** ☐ **Extreme** ☐

Quality of Life

Q39. How often are you aware of your foot/ankle problem?

Never ☐ **Monthly** ☐ **Weekly** ☐ **Daily** ☐ **Constantly** ☐

Q40. Have you modified your life style to avoid potentially damaging activities to your foot/ankle?

Not at all ☐ **Mildly** ☐ **Moderately** ☐ **Severely** ☐ **Totally** ☐

Q41. How much are you troubled with lack of confidence in your foot/ankle?

Not at all ☐ **Mildly** ☐ **Moderately** ☐ **Severely** ☐ **Extremely** ☐

Q42. In general, how much difficulty do you have with your foot/ankle?

Not at all ☐ **Mildly** ☐ **Moderately** ☐ **Severely** ☐ **Extremely** ☐

Q43. Pain Scale

On a scale of 0 to 100, where 0 is no pain and 100 is the worst pain you can imagine, mark on each line how painful your ankle is now at rest and with weight bearing.

No pain *Worst pain*

a) At rest: 0 ⌊——————————————————————⌋ 100

b) Weight bearing: 0 ⌊——————————————————————⌋ 100

Functional Limitations Profile

These questions refer to your function **today**, and ask about limitations in getting about and working arising due to your ankle injury.

Please tick all the boxes that apply to you **at the moment**

Getting about:

Q44. I have no problems getting about ☐

Q45. I walk shorter distances or often stop for a rest. ☐

Q46. I do not walk up or down hills. ☐

Q47. I only use stairs with a physical aid; for example a handrail, stick or crutches. ☐

Q48. I only go up and down stairs with assistance from someone else. ☐

Q49. I get about in a wheelchair. ☐

Q50. I do not walk at all. ☐

Q51. I walk by myself but with some difficulty; for example I limp, wobble, stumble or have a stiff leg. ☐

Q52. I only walk with help from someone else. ☐

Q53. I go up and down stairs more slowly; for example, one step at a time or I have to stop. ☐

Q54. I do not use stairs at all. ☐

Q55. I get about only by using a walking frame, crutches, stick, walls or hold onto furniture. ☐

Q56. I walk more slowly. ☐

Work:

Q57. If you are off work at the moment due to your ankle injury but intend to return to work, tick this box and leave the rest of this section. ☐

If not working at all due to health (eg, retired, unemployed for health reasons), tick question 58 and leave the remaining questions.

If not working for other reasons (eg, retired, unemployed, looking after home, in full time education), leave this section.

If you are working please tick all the boxes that apply to you at the moment

Q57A. I have no problems with my usual work activities ☐

Q58. I do not work at all (includes retired due to health). ☐

Q59. I do part of my job at home. ☐

Q60. I am not getting as much work done as usual. ☐

Q61. I often get irritable with my workmates; for example, I snap at them or criticise them easily. ☐

Q62. I work shorter hours. ☐

Q63. I only do light work. ☐

Q64. I only work for short periods of time or often stop to rest. ☐

Q65. I work at my usual job but with some changes; for example I use different tools or special aids, or I swap jobs with someone else. ☐

Q66. I do not do my job as carefully and accurately as usual. ☐

Quality of Life (1)

This section asks for your views about your health and how well you are able to carry out your usual activities

Q67. In general, would you say your health is:

Excellent	Very good	Good	Fair	Poor
☐	☐	☐	☐	☐

The following questions are about activities you might do during a typical day. Does your health **now** limit you in these activities? If so, how much?

Q68. **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling or playing golf

Yes, limited a lot	Yes, limited a little	No, not limited at all
☐	☐	☐

Q69. Climbing several flights of stairs

Yes, limited a lot	Yes, limited a little	No, not limited at all
☐	☐	☐

<u>During the past 4 weeks</u>, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

	Yes	No
Q70. Accomplished less than you would like	☐	☐
Q71. Were limited in the kind of work or other activities	☐	☐

<u>During the past 4 weeks</u>, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

	Yes	No
Q72. Accomplished less than you would like	☐	☐
Q73. Didn't do work or other activities as carefully as usual	☐	☐

Q74. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

Not at all	A little bit	Moderately	Quite a bit	Extremely
☐	☐	☐	☐	☐

These questions are about how you feel and how things have been with you during the past 4 weeks.

For each question, please give the one answer that comes closest to the way you have been feeling.

How much of the time during the past 4 weeks:

	All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time
Q75. Have you felt calm and peaceful?	☐	☐	☐	☐	☐	☐
Q76. Did you have a lot of energy?	☐	☐	☐	☐	☐	☐
Q77. Have you felt downhearted and low?	☐	☐	☐	☐	☐	☐

Q78. During the past 4 weeks, how much of the time has your physical health OR emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

All of the time	Most of the time	Some of the time	A little of the time	None of the time
☐	☐	☐	☐	☐

Quality of Life (2)

The following questions are to ask about your general health state at the moment. By placing a tick in one box in each group below, please indicate which statement best describes your own health state today.

Please tick **one** box for each question

Q79. Mobility:
I have no problems in walking about ☐
I have some problems in walking about ☐
I am confined to bed ☐

Q80. Self-Care:
I have no problems with self-care ☐
I have some problems washing or dressing myself ☐
I am unable to wash or dress myself ☐

Q81. Usual Activities (e.g. work, study, housework, family or leisure activities):
I have no problems with performing my usual activities ☐
I have some problems with performing my usual activities ☐
I am unable to perform my usual activities ☐

Q82. Pain / Discomfort:
I have no pain or discomfort ☐
I have moderate pain or discomfort ☐
I have extreme pain or discomfort ☐

Q83. Anxiety / Depression:
I am not anxious or depressed ☐
I am moderately anxious or depressed ☐
I am extremely anxious or depressed ☐

Q84. Compared with my general level of health over the past 6 months, my health state today is (please tick one box):

Better ☐ Much the same ☐ Worse ☐

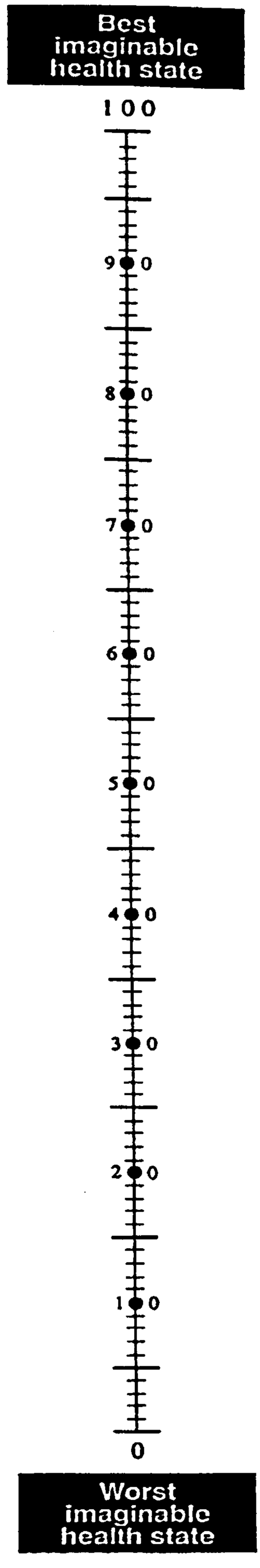
To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked by 100 and the worst state you can imagine is marked by 0.

We would like you to indicate on this scale **how good or bad is <u>your own health today</u>, in your opinion.**

Please do this by <u>drawing a line from the box below</u>, to whichever <u>point on the scale indicates how good or bad your current health state is **today**</u>.

Your own health state TODAY

Your own health state today



Q85. Are you still wearing the support applied to your ankle at the Ankle Trial clinic 12 weeks ago?

Not all – **please answer Q86** ☐

Now and again for sport or other activities ☐

Frequently for sport/activities but not during normal daily activities ☐

Frequently during normal daily activities but not all day ☐

Most of the time during the day ☐

Q86. If you are no longer wearing your support at all, please indicate approximately how long it was before you discarded it.

Less than one week ☐

1-2 weeks ☐

2-3 weeks ☐

3-4 weeks ☐

4-5weeks ☐

5-6 weeks ☐

Longer than 6 weeks (please specify time).....

Q87. Have you returned to your usual sports/activities?

No, not at all ☐

Yes but only gently or modified ☐

Yes, fully* ☐

*Approximately how long did it take for you to feel that your ankle had recovered enough to allow you to participate fully in your usual sports/activities?

.....

Resource Use

These questions refer to the past 12 weeks since you injured your ankle.

Q88. Is your ankle better, just the same or worse after the treatment you received 12 weeks ago?

Better ☐ same ☐ worse ☐

On a scale of 0-10, how much benefit do you think you have gained from the treatment? Circle your answer (0 = no benefit, 10 = maximum benefit).

0 1 2 3 4 5 6 7 8 9 10

Q89. During the past 12 weeks, have you consulted a doctor or therapist or received any further treatment for your ankle (apart from the treatment you received as part of the trial)?

Yes ☐ No ☐

If 'yes', please specify which treatment by placing a tick in the appropriate box:

A&E staff eg plaster technician ☐ how many times? ___

NHS consultant ☐ how many times? ____

Private Consultant ☐ how many times? ____

GP ☐ how many times? ____

Osteopathy ☐ how many times? ____

Chiropractic ☐ how many times? ____

NHS Physiotherapy ☐ how many times? ____

Private Physiotherapy ☐ how many times? ____

Other (please specify) _____

Did you pay for this treatment? Yes ☐ No ☐

If 'Yes' was payment made by yourself or a private insurance company?

Self ☐ Insurance company ☐

How much did it cost? _____

Q90. Over the past 12 weeks have you had any scans or xrays because of your ankle (**apart from xrays you may have had in A&E when you first injured your ankle**)?

Yes ☐ No ☐

If 'Yes' what type of xray or scan? (tick more than one box if needed)

Normal xray ☐

MRI scan ☐

Ultrasound scan ☐

Did you pay for this/these scan(s)? Yes ☐ No ☐

If 'Yes' was payment made by yourself or a private insurance company?

Self ☐ Insurance company ☐

How much did it cost? _____

Q91. Over the past 12 weeks, have you been admitted to hospital because of your ankle?

Yes ☐ No ☐

If 'Yes' how many days did you spend in hospital? _____

Q92. Has your ***doctor prescribed*** any medicines, creams or other treatments (eg brace/strapping) for your ankle over the past 12 weeks? (**Do not** include the brace/support you wore as part of the trial)

Prescribed medicines/creams:

Item description	*Name of item (eg Ibuprofen)*	*Cost to you (eg prescription charge or other cost)*
Painkillers		£
Anti-inflammatories		£
Creams/gels		£
Aids/braces/strapping		£
Injection		£
Other		£

Q93. Over the past 12 weeks, have you ***bought*** any medicines, creams or other treatment (eg brace) for your ankle?

Medicines/creams bought ***without prescription***:

Item description	*Name of item*	*Cost to you*
Painkillers		£
Anti-inflammatories		£
Creams/gels		£
Aids/braces/strapping		£
Herbal remedies		£
Other		£

Q94. Over the past 12 weeks, have you had to take any sick leave from work because of your ankle?

Yes ☐ No ☐ Not applicable ☐

If 'Yes' how many sick days did you take? _____

Q95. Have you been involved in any exercise/sport over the last 12 weeks?

Yes ☐ No ☐

If 'yes' which ones? (please tick all boxes that apply).

Swimming ☐

Weight training ☐

Aerobics/keep-fit ☐

Cycling ☐

Jogging/running ☐

Team sport (eg football ,rugby,hockey,netball) ☐

Raquet sport (eg tennis, squash, badminton) ☐

Yoga ☐

Athletics ☐

Walks of 2 miles or more ☐

Heavy housework/DIY/gardening ☐

Other sports or exercise (please specify):

_____

Approximately how many times in the last 12 weeks have you done any of these activities?

1 ☐

2-3 ☐

3-4 ☐

5 or more ☐

Please check that you have completed all relevant sections.

Please return your completed questionnaire to us in the envelope provided.

Thank you very much for your time.

Appendix 3 – List of studies excluded from systematic review and reasons for exclusion

1. Addington-Hall J, Walker L, Jones C, Karlsen S, McCarthy M. A randomised controlled trial of postal versus interviewer administration of a questionnaire measuring satisfaction with, and use of, services received in the year before death. J Epidemiol Community Health 1998;52(12):802-7.

2. Asch DA. Use of a coded postcard to maintain anonymity in a highly sensitive mail survey: cost, response rates, and bias. Epidemiology 1996;7(5):550-1.

3. Asch DA, Christakis NA. Different response rates in a trial of two envelop styles in mail survey research. Epidemiology 1994;5(3):364-5.

4. Asch DA, Christakis NA, Ubel PA. Conducting physician mail surveys on a limited budget. A randomized trial comparing $2 bill versus $5 bill incentives. Med Care 1998;36(1):95-9.

5. Baker CM. Maximizing mailed questionnaire responses. Image J Nurs Sch 1985;17(4):118-21.

6. Barker PJ, Cooper RF. Do sexual health questions alter the public's response to lifestyle questionnaires? J Epidemiol Community Health 1996;50(6):688.

7. Barriball KL, While AE. Exploring variables underlying non-response in a survey of nurses and nurses' aides in practice. J Adv Nurs 1999;29(4):894-904.

8. Becker H, Cookston J, Kulberg V. Mailed survey follow-ups--are postcard reminders more cost-effective than second questionnaires? West J Nurs Res 2000;22(5):642-7.

9. Bhandari M, Devereaux PJ, Swiontkowski MF, Schemitsch EH, Shankardass K, Sprague S, et al. A randomized trial of opinion

leader endorsement in a survey of orthopaedic surgeons: effect on primary response rates. Int J Epidemiol 2003;32(4):634-6.

10. Biggar RJ, Melbye M. Responses to anonymous questionnaires concerning sexual behavior: a method to examine potential biases. Am J Public Health 1992;82(11):1506-12.

11. Blomberg J, Sandell R. Does a material incentive affect response on a psychotherapy follow-up questionnaire? Psychotherapy Research 1996;6(3):155-163.

12. Brogger J, Bakke P, Eide GE, Gulsvik A. Contribution of follow-up of nonresponders to prevalence and risk estimates: a Norwegian respiratory health survey. Am J Epidemiol 2003;157(6):558-66.

13. Bruce T, Salkeld G, Short L, Solomon M, Ward J. A randomised trial of telephone versus postcard prompts to enhance response rate in a phased population-based study about community preferences. Aust N Z J Public Health 2000;24(4):456-7.

14. Campbell MJ, Waters WE. Does anonymity increase response rate in postal questionnaire surveys about sensitive subjects? A randomised trial. J Epidemiol Community Health 1990;44(1):75-6.

15. Camunas C, Alward RR, Vecchione E. Survey response rates to a professional association mail questionnaire. J N Y State Nurses Assoc 1990;21(3):7-9.

16. Cartwright A. Some experiments with factors that might affect the response of mothers to a postal questionnaire. Stat Med 1986;5(6):607-17.

17. Cartwright A. Interviews or postal questionnaires? Comparisons of data about women's experiences with maternity services. Milbank Q 1988;66(1):172-89.

18. Chapman S, Wong WL. Incentives for questionnaire respondents. Aust J Public Health 1991;15(1):66-7.

19. Choi BC, Pak AW, Purdham JT. Effects of mailing strategies on response rate, response time, and cost in a questionnaire study among nurses. Epidemiology 1990;1(1):72-4.

20. Clark TJ, Khan KS, Gupta JK. Effect of paper quality on the response rate to a postal survey: a randomised controlled trial. ISRCTN 32032031. BMC Med Res Methodol 2001;1(1):12.

21. Clark TJ, Khan KS, Gupta JK. Provision of pen along with questionnaire does not increase the response rate to a postal survey: a randomised controlled trial. J Epidemiol Community Health 2001;55(8):595-6.

22. Clarke R, Breeze E, Sherliker P, Shipley M, Youngman L, Fletcher A, et al. Design, objectives, and lessons from a pilot 25 year follow up re-survey of survivors in the Whitehall study of London Civil Servants. J Epidemiol Community Health 1998;52(6):364-9.

23. Clayton AE, McNutt LA, Homestead HL, Hartman TW, Senecal S. Public health in managed care: a randomized controlled trial of the effectiveness of postcard reminders. Am J Public Health 1999;89(8):1235-7.

24. Collins RL, Ellickson PL, Hays RD, McCaffrey DF. Effects of incentive size and timing on response rates to a follow-up wave of a longitudinal mailed survey. Eval Rev 2000;24(4):347-63.

25. Deehan A, Templeton L, Taylor C, Drummond C, Strang J. The effect of cash and other financial inducements on the response rate of general practitioners in a national postal study. Br J Gen Pract 1997;47(415):87-90.

26. Del Valle ML, Morgenstern H, Rogstad TL, Albright C, Vickrey BG. A randomized trial of the impact of certified mail on response rate to a physician survey, and a cost-effectiveness analysis. Eval Health Prof 1997;20(4):389-406.

27. Donaldson GW, Moinpour CM, Bush NE, Chapko M, Jocom J, Siadak M, et al. Physician participation in research surveys. A randomized study of inducements to return mailed research questionnaires. Eval Health Prof 1999;22(4):427-41.

28. Doody MM, Sigurdson AS, Kampa D, Chimes K, Alexander BH, Ron E, et al. Randomized trial of financial incentives and delivery methods for improving response to a mailed questionnaire. Am J Epidemiol 2003;157(7):643-51.

29. Duffy DL, Martin NG. Increasing the response rate to a mailed questionnaire by including more stamps on the return envelope: a cotwin control study. Twin Res 2001;4(2):71-2.

30. Eaker S, Bergstrom R, Bergstrom A, Adami HO, Nyren O. Response rate to mailed epidemiologic questionnaires: a population-based randomized trial of variations in design and mailing routines. Am J Epidemiol 1998;147(1):74-82.

31. Elliott AM, Hannaford PC. Third mailings in epidemiological studies: are they really necessary? Fam Pract 2003;20(5):592-4.

32. Erwin W, Wheelright L. Improving mail survey response rates through the use of monetary incentive. Journal of Mental Health Counseling 2002;24(3):247-249.

33. Etter JF, Cucherat M, Perneger TV. Questionnaire color and response rates to mailed surveys. A randomized trial and a meta-analysis. Eval Health Prof 2002;25(2):185-99.

34. Etter JF, Perneger TV, Laporte JD. Unexpected effects of a prior feedback letter and a professional layout on the response rate to a mail survey in Geneva. J Epidemiol Community Health 1998;52(2):128-9.

35. Etter JF, Perneger TV, Ronchi A. Collecting saliva samples by mail. Am J Epidemiol 1998;147(2):141-6.

36. Faria M, Mateus C, Coelho F, Martins R, Barros H. Postal questionnaires: a useful strategy for the follow up of stroke cases. Acta Medica Portugesa 1997;10:61-65.

37. Feild H. Effects of sex of investigator on mail survey response rates and response bias. Journal of Applied Psychology 1975;60(6):772-773.

38. Field TS, Cadoret CA, Brown ML, Ford M, Greene SM, Hill D, et al. Surveying physicians: do components of the "Total Design Approach" to optimizing survey response rates apply to physicians? Med Care 2002;40(7):596-605.

39. Futrell C, Lamb C. Effect on mail survey return rates of including questionnaires with follow up letters. Perceptual and Motor Skills 1981;52:11-15.

40. Gasquet I, Falissard B, Ravaud P. Impact of reminders and method of questionnaire distribution on patient response to mail-back satisfaction survey. J Clin Epidemiol 2001;54(11):1174-80.

41. Gattellari M, Ward JE. Will donations to their learned college increase surgeons' participation in surveys? A randomized trial. J Clin Epidemiol 2001;54(6):645-9.

42. Gerace TA, George VA, Arango IG. Response rates to six recruitment mailing formats and two messages about a nutrition

program for women 50-79 years old. Control Clin Trials 1995;16(6):422-31.

43. Gibson PJ, Koepsell TD, Diehr P, Hale C. Increasing response rates for mailed surveys of Medicaid clients and other low-income populations. Am J Epidemiol 1999;149(11):1057-62.

44. Gilbart E, Kreiger N. Improvement in cumulative response rates following implementation of a financial incentive. Am J Epidemiol 1998;148(1):97-9.

45. Gilbert GH, Longmate J, Branch LG. Factors influencing the effectiveness of mailed health surveys. Public Health Rep 1992;107(5):576-84.

46. Gordon SE, Stokes SA. Improving response rate to mailed questionnaires. Nurs Res 1989;38(6):375-6.

47. Gore-Felton C, Koopman C, Bridges E, Thoresen C, Spiegel D. An example of maximizing survey return rates. Methodological issues for health professionals. Eval Health Prof 2002;25(2):152-68.

48. Gupta L, Ward J, D'Este C. Differential effectiveness of telephone prompts by medical and nonmedical staff in increasing survey response rates: a randomised trial. Aust N Z J Public Health 1997;21(1):98-9.

49. Halpern SD, Ubel PA, Berlin JA, Asch DA. Randomized trial of 5 dollars versus 10 dollars monetary incentives, envelope size, and candy to increase physician response rates to mailed questionnaires. Med Care 2002;40(9):834-9.

50. Harrison RA, Holt D, Elton PJ. Do postage-stamps increase response rates to postal surveys? A randomized controlled trial. Int J Epidemiol 2002;31(4):872-4.

51. Hoffman SC, Burke AE, Helzlsouer KJ, Comstock GW. Controlled trial of the effect of length, incentives, and follow-up techniques on response to a mailed questionnaire. Am J Epidemiol 1998;148(10):1007-11.

52. Iglesias CP, Birks YF, Torgerson DJ. Improving the measurement of quality of life in older people: the York SF-12. Qjm 2001;94(12):695-8.

53. Jacoby A. Possible factors affecting response to postal questionnaires: findings from a study of general practitioner services. J Public Health Med 1990;12(2):131-5.

54. John EM, Savitz DA. Effect of a monetary incentive on response to a mail survey. Ann Epidemiol 1994;4(3):231-5.

55. Kalantar JS, Talley NJ. The effects of lottery incentive and length of questionnaire on health survey response rates: a randomized study. J Clin Epidemiol 1999;52(11):1117-22.

56. Kaplan S, Cole P. Factors affecting response to postal questionnaires. Br J Prev Soc Med 1970;24(4):245-7.

57. Kissinger P, Kopicko JJ, Myers L, Wustrack S, Elkins W, Farley TA, et al. The effect of modest monetary incentives on follow-up rates in sexually transmitted disease studies. Int J STD AIDS 2000;11(1):27-30.

58. Koo MM, Rohan TE. Printed signatures and response rates. Epidemiology 1995;6(5):568.

59. Koo MM, Rohan TE. Types of advance notification in reminder letters and response rates. Epidemiology 1996;7(2):215-6.

60. Little RE, Davis AK. Effectiveness of various methods of contact and reimbursement on response rates of pregnant women to a mail questionnaire. Am J Epidemiol 1984;120(1):161-3.

61. Lund E, Gram IT. Response rate according to title and length of questionnaire. Scand J Soc Med 1998;26(2):154-60.

62. Maheux B, Legault C, Lambert J. Increasing response rates in physicians' mail surveys: an experimental study. Am J Public Health 1989;79(5):638-9.

63. Mailey SK. Increasing your response rate for mail survey data collection. SCI Nurs 2002;19(2):78-9.

64. Markush RE. Relation of form length to response in mailed epidemiologic inquiries. Public Health Rep 1966;81(2):191-5.

65. Marrett LD, Kreiger N, Dodds L, Hilditch S. The effect on response rates of offering a small incentive with a mailed questionnaire. Ann Epidemiol 1992;2(5):745-53.

66. McCaul LA, Cooper PG. Techniques to increase the response rate in follow-up studies: results of a pilot test. Int J Rehabil Res 1979;2(4):465-70.

67. McCloskey MI. Maximising response to postal surveys of adult dental health. Community Dent Health 1995;12(4):244.

68. McKillip J, Lockhart D. The effectiveness of cover letter appeals. Journal of Social Psychology 1984;122(1):85-91.

69. Morrison DS, Thomson H, Petticrew M. Effects of disseminating research findings on response rates in a community survey: a randomised controlled trial. J Epidemiol Community Health 2003;57(7):536-7.

70. Morrison TC, Wahlgren DR, Hovell MF, Zakarian J, Burkham-Kreitner S, Hofstetter CR, et al. Tracking and follow-up of 16,915 adolescents: minimizing attrition bias. Control Clin Trials 1997;18(5):383-96.

71. Mortagy AK, Howell JB, Waters WE. A useless raffle. J Epidemiol Community Health 1985;39(2):183-4.

72. Murawski M, Carroll N. Direct mail performance of selected health related quality of life scales. Journal of Pharmacoepidemiology 1996;5(1):17-38.

73. Nagata C, Hara S, Shimizu H. Factors affecting response to mail questionnaire: research topics, questionaire length and non-response bias. Journal of Epidemiology 1995;5(5):81-85.

74. Oden L, Price JH. Effects of a small monetary incentive and follow-up mailings on return rates of a survey to nurse practitioners. Psychol Rep 1999;85(3 Pt 2):1154-6.

75. Osborn M, Ward J, Boyle C. Effectiveness of telephone prompts when surveying general practitioners: a randomised trial. Aust Fam Physician 1996;Suppl 1:S41-3.

76. Parker C, Dewey M. Assessing research outcomes by postal questionnaire with telephone follow-up. TOTAL Study Group. Trial of Occupational Therapy and Leisure. Int J Epidemiol 2000;29(6):1065-9.

77. Patten SB, Li FX, Cook T, Hilsden RJ, Sutherland LR. Irritable bowel syndrome: are incentives useful for improving survey response rates? J Clin Epidemiol 2003;56(3):256-61.

78. Perneger TV, Etter JF, Rougemont A. Randomized trial of use of a monetary incentive and a reminder card to increase the response rate to a mailed health survey. Am J Epidemiol 1993;138(9):714-22.

79. Peters TJ, Harvey IM, Bachmann MO, Eachus JI. Does requesting sensitive information on postal questionnaires have an impact on response rates? A randomised controlled trial in the south west of England. J Epidemiol Community Health 1998;52(2):130.

80. Picavet HS. National health surveys by mail or home interview: effects on response. J Epidemiol Community Health 2001;55(6):408-13.

81. Pirotta M, Gunn J, Farish S, Karabatsos G. Primer postcard improves postal survey response rates. Aust N Z J Public Health 1999;23(2):196-7.

82. Price JH, Dake JA, Akpanudo S, Kleinfelder J. Effect on survey return rates of having a signed or unsigned postcard as the third-wave mailing. Psychol Rep 2003;92(3 Pt 2):1099-102.

83. Rikard-Bell G, Ward J. Maximizing response rates to a survey of dentists: a randomized trial. Aust Dent J 2000;45(1):46-8.

84. Rimm EB, Stampfer MJ, Colditz GA, Giovannucci E, Willett WC. Effectiveness of various mailing strategies among nonrespondents in a prospective cohort study. Am J Epidemiol 1990;131(6):1068-71.

85. Rissel C. Improving response rates with incentives. Med J Aust 1989;150(7):411.

86. Roberts H, Pearson JC, Dengler R. Impact of a postcard versus a questionnaire as a first reminder in a postal lifestyle survey. J Epidemiol Community Health 1993;47(4):334-5.

87. Roberts PJ, Roberts C, Sibbald B, Torgerson DJ. The effect of a direct payment or a lottery on questionnaire response rates: a randomised controlled trial. J Epidemiol Community Health 2000;54(1):71-2.

88. Roberts R, McCrory O, Forthofer R. Further evidence of using a deadline to stimulate responses to a mail survey. The Public Opinion Quarterly 1978;42(3):407-410.

89. Rolnick SJ, Gross CR, Garrard J, Gibson RW. A comparison of response rate, data quality, and cost in the collection of data on sexual history and personal behaviors. Mail survey approaches and in-person interview. Am J Epidemiol 1989;129(5):1052-61.

90. Rosenfeld AA, Smith CR, Brewster W, Wenegrat AO, Haavik DK. Increasing questionnaire responses. Am J Psychiatry 1979;136(11):1480-1.

91. Rudy EB, Estok PJ, Kerr ME, Menzel L. Research incentives: money versus gifts. Nurs Res 1994;43(4):253-5.

92. Russell ML, Mutasingwa DR, Verhoef MJ, Injeyan HS. Effect of a monetary incentive on chiropractors' response rate and time to respond to a mail survey. J Clin Epidemiol 2003;56(10):1027-8.

93. Schweitzer M, Asch DA. Timing payments to subjects of mail surveys: cost-effectiveness and bias. J Clin Epidemiol 1995;48(11):1325-9.

94. Shackleton VJ, Wild JM, Wolffe M. Screening optometric patients by questionnaire: methods of improving response. Am J Optom Physiol Opt 1980;57(6):404-6.

95. Shahar E, Bisgard KM, Folsom AR. Response to mail surveys: effect of a request to explain refusal to participate. The ARIC Study Investigators. Epidemiology 1993;4(5):480-2.

96. Shaw MJ, Beebe TJ, Jensen HL, Adlis SA. The use of monetary incentives in a community survey: impact on response rates, data quality, and cost. Health Serv Res 2001;35(6):1339-46.

97. Siemiatycki J. A comparison of mail, telephone, and home interview strategies for household health surveys. Am J Public Health 1979;69(3):238-45.

98. Smeeth L, Fletcher AE, Stirling S, Nunes M, Breeze E, Ng E, et al. Randomised comparison of three methods of administering a screening questionnaire to elderly people: findings from the MRC trial of the assessment and management of older people in the community. Bmj 2001;323(7326):1403-7.

99. Smith WC, Crombie IK, Campion PD, Knox JD. Comparison of response rates to a postal questionnaire from a general practice and a research unit. Br Med J (Clin Res Ed) 1985;291(6507):1483-5.

100. Spry VM, Hovell MF, Sallis JG, Hofsteter CR, Elder JP, Molgaard CA. Recruiting survey respondents to mailed surveys: controlled trials of incentives and prompts. Am J Epidemiol 1989;130(1):166-72.

101. Streiff MB, Dundes L, Spivak JL. A mail survey of United States hematologists and oncologists: a comparison of business reply versus stamped return envelopes. J Clin Epidemiol 2001;54(4):430-2.

102. Subar AF, Ziegler RG, Thompson FE, Johnson CC, Weissfeld JL, Reding D, et al. Is shorter always better? Relative importance of questionnaire length and cognitive ease on response rates and data quality for two dietary questionnaires. Am J Epidemiol 2001;153(4):404-9.

103. Szirony TA, Price JH, Telljohann SK, Wolfe E. Survey return rates using a covering letter signed by a graduate student or faculty member. Psychol Rep 2002;91(3 Pt 2):1174-6.

104. VanGeest JB, Wynia MK, Cummins DS, Wilson IB. Effects of different monetary incentives on the return rate of a national mail survey of physicians. Med Care 2001;39(2):197-201.

105. Vetter NJ, Ford D. The collection of health data from the elderly in primary care using a postal questionnaire. Arch Gerontol Geriatr 1989;9(1):53-8.

106. Victor CR. Some methodological aspects of using postal questionnaires with the elderly. Arch Gerontol Geriatr 1988;7(2):163-72.

107. Ward J, Bruce T, Holt P, D'Este K, Sladden M. Labour-saving strategies to maintain survey response rates: a randomised trial. Aust N Z J Public Health 1998;22(3 Suppl):394-6.

108. Weltzien RT, Jr., McIntyre TJ, Ernst JA, Walsh JA, Parker JK. Cross validation of some psychometric properties of the CSQ and its differential return rate as a function of token financial incentive. Community Ment Health J 1986;22(1):49-55.

109. Wensing M, Mainz J, Kramme O, Jung HP, Ribacke M. Effect of mailed reminders on the response rate in surveys among patients in general practice. J Clin Epidemiol 1999;52(6):585-7.

110. Whiteman MK, Langenberg P, Kjerulff K, McCarter R, Flaws JA. A randomized trial of incentives to improve response rates to a mailed women's health questionnaire. J Womens Health (Larchmt) 2003;12(8):821-8.

111. Windsor J. What can you ask about? The effect on response to a postal screen of asking about two potentially sensitive questions. J Epidemiol Community Health 1992;46(1):83-5.

112. Woodward A, Douglas B, Miles H. Chance of free dinner increases response to mail questionnaire. Int J Epidemiol 1985;14(4):641-2.

113. Wunder G, Wynn G. The effects of address personalisation on mailed questionnaires response rate, time and quality. Journal of the Market Research Society 1988;30(1):95-101.

114. Yacovone JA. Evaluation of response rates to two differing types of questionnaire surveys: hypothesis that those responding to one questionnaire are likely to respond to others is apparently true. R I Med J 1979;62(7):270-5.

Reasons for exclusion

Not postal q = Data collection technique was not via postal questionnaire.

Not pts = Population studied was not defined as 'patients' (ie not receiving medical or surgical treatment).

Not HC = The questionnaire was not being used as a tool in a health care research study.

Not RCT = The method of improving response was not being tested using a randomised trial.

No eligible intervention = On closer inspection the study does not describe a method of improving questionnaire response.

First author	Reasons for exclusion	Excluded on full text (FT) or abstract (A)
Addington-Hall	Not postal q	A
Asch	Not pts	A
Asch	Not pts	A
Asch	Not pts	A
Baker	Not pts	FT
Barker	Not pts	A
Barriball	Not pts	A
Becker	Not pts	FT
Bhandari	Not pts	A
Biggar	Not pts	FT
Blomberg	Not pts	A
Brogger	Not pts	FT
Bruce	Not HC	A
Campbell	Not pts	FT
Camunus	Not pts	A

Cartwright	Not pts	FT
Cartwright	Not postal q	A
Chapman	Not pts	FT
Choi	Not pts	A
Clark	Not pts	FT
Clark	Not pts	A
Clarke	Not pts	A
Clayton	Not postal q	A
Collins	Not HC	A
Deehan	Not pts	A
Del Valle	Not pts	A
Donaldson	Not pts	A
Doody	Not pts	A
Duffy	Not HC	A
Eaker	Not pts	FT
Elliot	Not RCT	A
Erwin	Not RCT	FT
Etter	Not postal q	FT
Etter	Not pts	FT
Etter	Not pts	A
Faria	Not RCT	A
Feild	Not RCT	A
Field	Not pts	A
Futrell	Not HC	A
Gasquet	Not RCT	A
Gatellari	Not pts	A
Gerace	Not pts	A
Gibson	Not pts	FT
Gilbart	Not pts	A
Gilbert	Not RCT	A
Gordon	Not pts	FT
Gore-Felton	Not pts	A
Gupta	Not pts	A
Halpern	Not pts	A
Harrison	Not pts	FT
Hoffman	Not pts	FT
Iglesias	Not RCT	FT
Jacoby	Not pts	FT
John	Not pts	A
Kalantar	Not pts	FT
Kaplan	Not RCT	A
Kissinger	Not postal q	A
Koo	Not pts	FT
Koo	Not pts	FT
Little	Not pts	FT
Lund	Not pts	FT
Maheux	Not pts	A
Mailey	Not RCT	A
Markush	Not pts	A
Marrett	Not pts	A

McCaul	Not postal q	A
McCloskey	Not pts	A
McKillip	Not Hc	A
Morrison	Not pts	A
Morrison	Not pts	FT
Mortagy	Not pts	FT
Murawski	No eligible intervention	A
Nagata	Not RCT	A
Oden	Not pts	A
Osborn	Not pts	A
Parker	Not RCT	FT
Patten	Not pts	FT
Perneger	Not pts	FT
Peters	Not pts	FT
Picavet	Not postal q	A
Pirotta	Not pts	A
Price	Not HC	A
Rikard-Bell	Not pts	A
Rimm	Not pts	A
Rissel	Not RCT	A
Roberts	Not pts	FT
Roberts	Not HC	A
Roberts	Not pts	FT
Rolnick	No eligible intervention	A
Rosenfeld	Not RCT	A
Rudy	Not RCT	A
Russell	Not pts	A
Schweitzer	Not pts	FT
Shackleton	Not pts	A
Shahar	Not pts	FT
Shaw	Not pts	FT
Siemiatycki	Not postal q	A
Smeeth	No eligible intervention	FT
Smith	Not pts	FT
Spry	Not pts	FT
Streiff	Not pts	A
Subar	Not pts	FT
Szirony	Not HC	A
VanGeest	Not pts	A
Vetter	Not RCT	A
Victor	Not RCT	FT
Ward	Not pts	A
Weltzein	Not HC	A
Wensing	Not pts	FT
Whiteman	Not pts	FT
Windsor	Not pts	FT
Woodward	Not pts	FT
Wunder	Not HC	A
Yacovone	Not RCT	A

Appendix 4 – The Trial Calendar

COLLABORATIVE ANKLE SUPPORT TRIAL

CALENDAR

This calendar is for you to use as a reminder about things that happen over the next few weeks/months due to your ankle injury. In the questionnaires that will follow over the next 9 months, we will be asking you:

- When you stopped wearing you ankle support
- When you went back to work (if you work)
- When you were able to play sports/activities again
- If you had to return to the hospital or see your doctor or have any other treatment because of your ankle injury
- If you have had to buy any medicines etc because of your ankle injury

Please circle the date on the calendar when these things occur and this will make your questionnaires much easier to answer

THANK YOU FOR YOUR HELP

By accurately filling in and returning the questionnaires you will help us to understand how best to treat people with severe ankle sprains.

Rachel Nakash 024 7657 4650

THE UNIVERSITY OF WARWICK

THE UNIVERSITY OF BIRMINGHAM

MARCH

SUN	MON	TUE	WED	THU	FRI	SAT
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

THANK YOU FOR BEING IN OUR TRIAL

We hope your ankle is feeling much better!

<u>Have you gone back to work yet?</u>

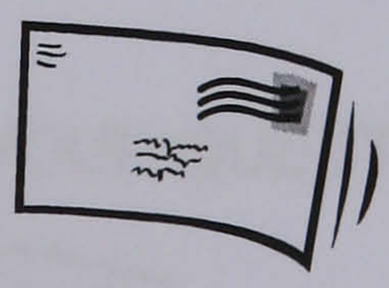
Please make a note on the calendar how many days you were off work due to your ankle

APRIL

SUN	MON	TUE	WED	THU	FRI	SAT
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

Have you gone back to work yet?

Please mark on the calendar how many days you were off work due to your ankle injury



Please remember to send your questionnaire back in the envelope

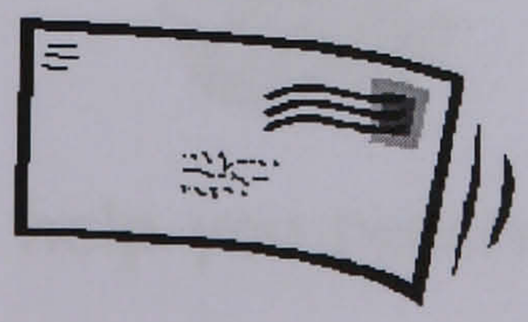
MAY

SUN	MON	TUE	WED	THU	FRI	SAT
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

Are you still wearing your support?

Please try and remember roughly when you stopped wearing it and make a note on the calendar

DID YOU REMEMBER TO SEND YOUR QUESTIONNAIRE BACK?



JUNE

SUN	MON	TUE	WED	THU	FRI	SAT
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30		

Questionnaire time again!

There are some extra questions this time, asking about time/money you have spent on your ankle injury.

Use your calendar to help you record time/money you have

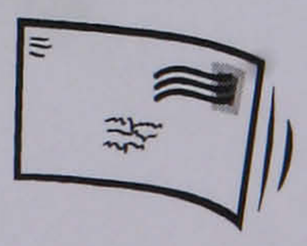
JULY

SUN	MON	TUE	WED	THU	FRI	SAT
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

Are you doing your usual sports/activities?

Use your calendar to record when you felt able to return to sport/activites

Did you remember to send your questionnaire back last month?



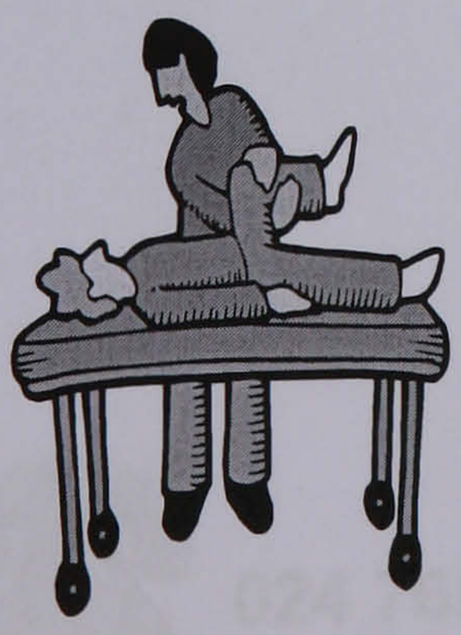
Did you remember to send your questionnaire back last month?

AUGUST

SUN	MON	TUE	WED	THU	FRI	SAT
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

Have you needed to return to the hospital or see your doctor because of your ankle injury?

Use the calendar to record any appointments relating to your ankle injury.



SEPTEMBER

SUN	MON	TUE	WED	THU	FRI	SAT
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

Have you moved house recently?

Please don't forget to Inform the trial office If you haved moved or are about to move

024 7657 4656

Thank you

OCTOBER

SUN	MON	TUE	WED	THU	FRI	SAT
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

Hope you have not had any more problems with your ankle

NOVEMBER

SUN	MON	TUE	WED	THU	FRI	SAT
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

Final questionnaire coming this month!

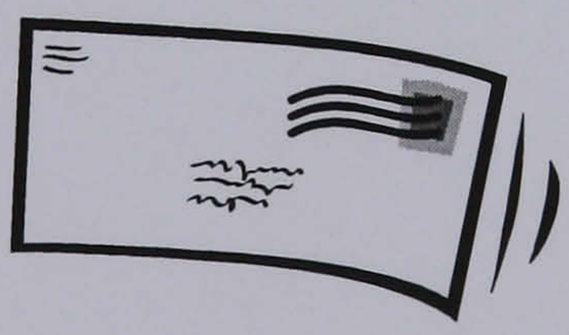
Please answer this final questionnaire to help us complete our records. The long-term outcome of our study is very important to help us learn how best to treat other patients who have severe ankle sprains

DECEMBER

SUN	MON	TUE	WED	THU	FRI	SAT
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

Merry Christmas

Did you remember to return your final questionnaire?



We hope that you are now fully recovered and thank you for being part of the CAST project

Appendix 5 – Trial Calendar instructions for use

Standardised instructions on issue of Trial Calendar

As part of the CAST trial we are giving some people a calendar like this.

This calendar is for you to use as a reminder about things that happen over the next few weeks/months due to your ankle injury. In the questionnaires that will follow over the next 9 months, we will be asking you:

- When you stopped wearing your ankle support
- When you went back to work (if you work)
- When you were able to play sports/activities again
- If you had to return to the hospital or see your doctor or have any other treatment because of your ankle injury
- If you have had to buy any medicines etc because of your ankle injury

Please circle the date on the calendar when these things occur and this will make your questionnaires much easier to answer

Thank you for your help

By accurately filling in and returning the questionnaires you will help us to understand how best to treat people with severe ankle sprains.

Appendix 6 – Interview schedule for qualitative study

Interview Schedule

Four main areas of interest;

- Issues surrounding response to postal questionnaires from patient's perspective.
- Patient's thoughts on the usefulness of the Trial Calendar.
- The relevance and acceptability of the follow-up questionnaire.
- Thoughts on being involved in the trial

Response Issues

Responders:

Suppose you have a questionnaire sent to you 'out of the blue' like a household survey. Do you usually respond to it?

Probe: If 'yes' or 'no', why do you think that is?

So what encouraged you fill in our questionnaire and send it back?

What are your general thoughts about our questionnaire? For example; layout, appearance, length and time to complete.

Can you tell me your thoughts on how easy or difficult you found the questions to answer.

Probe: Understanding of the questions, ease of retrieval of relevant info.

What about the questions that asked about specific events such as when you returned to work and how many days you had off work. How did you remember these things?

Probe: Accurate?

Are there any ways that the process of completing and returning your questionnaire could have been made easier for you?

Non-responders:

Suppose you have a questionnaire sent to you 'out of the blue' like a household survey. Do you usually respond to it?

Probe: If 'yes' or 'no', why do you think that is?

What do you think we could have done to make it easier for you to complete and return our questionnaire?

Probe: more contact, more reminders, incentive.

What are your thoughts about the whole questionnaire? For example; layout, appearance, length and time to complete.

Can you tell me any specific reasons why you didn't send your questionnaire back.

Probe: Dissatisfaction with treatment received, understanding, ease of retrieval of info.

Trial Calendar

What are your thoughts about the design of the calendar?

Is a calendar like this something you would normally use in everyday life?

Can you tell me what you understand about the purpose of the Trial Calendar.

Has the Trial Calendar helped you with any aspect of being in the ankle trial? If so please explain.

Questionnaire acceptability

Looking at the first section of the questionnaire, what are your thoughts about the questions that ask you about your ankle injury?

Are there any areas of your life and daily activities that were affected by your injury that are not covered in the first section of the questionnaire that you feel should have been?

Thoughts on trial

What was your understanding of the trial when you were asked to take part?

What did you think about the way the treatments were given out and the fact that you couldn't choose which treatment you received?

Which ankle support did you receive and how did you cope with it in your day to day life?

Would you have preferred one of the other treatments? If so, which one and why?

Probe: If didn't receive what they wanted did this affect their willingness to complete and return the questionnaire