Gender differences in quality of life and depression among older people with coronary heart disease

Paola Zaninotto

Thesis submitted for the degree of Doctor of Philosophy

UCL

2012

Declaration of authorship

‘I, Paola Zaninotto confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.'

Abstract

This study explored gender differences in quality of life and depressive symptoms over time among people with coronary heart disease (CHD) and compared them with a healthy population. Using three waves of the English Longitudinal Study of Ageing (2002-03, 2004-05 and 2006-07) methodological problems such as missing data and sources of error and uncertainty which may arise from reliance on a self-reported measure of CHD were addressed.

A simulation study was set up to compare three techniques for dealing with missing data: full information maximum likelihood, multivariate normal imputation and a two-fold fully conditional specification. Results supported the use of the latter technique which outperformed the other two techniques in recovering the targeted parameters, especially with a binary outcome.

Results based on imputed data showed that compared to people from the healthy population, men and women with CHD had on average lower levels of quality of life. Men with CHD were also at higher risk of having depressive symptoms than men from the healthy population. Women with CHD were as likely as women from the healthy population to have depressive symptoms. Trajectories over time of quality of life had a different shape from trajectories of depressive symptoms after the onset of CHD. Men's quality of life declined over time and no changes in depressive symptoms were found. Women's quality of life declined only between baseline and four year follow-up, while in the same period their risk of having depressive symptoms reduced.

A sensitivity analysis based on an external validation study and a deterministic sensitivity analysis helped understand the impact that misclassification of the self-reported CHD measure could have on the results of this thesis. It was found that the reliability of the results presented could be affected by false positive and false negative cases of CHD.

Acknowledgments

I wish to express my warmest gratitude to my supervisors for guiding me throughout this work.

I am thankful to Professor Amanda Sacker, who has given numerous hours of her time to guide my work, who has inspired me and helped me with practical issues. I am grateful to Dr Elizabeth Breeze, for initially giving me the topic of this study. I appreciated her advice, ideas and detailed comments. I am thankful to Dr Anne McMunn for her advice, support and for building up my spirits. Her optimism has been invaluable.

I am sincerely grateful to Professor Mike Kenward for his guidance on the application to this thesis of his method for dealing with missing data, two-fold fully conditional specification. I am most grateful for the time and patience he has generously offered.

I wish to thank Martin Shipley for his help on the validation of the self-reported measure of CHD in the Whitehall II study. I thank Jenny Head for her valuable comments on the chapter on sensitivity analysis and for help with the Whitehall II data.

I would like to thank my PhD colleagues, Dr Lidia Panico, Caroline Coope and Rebecca Landy, and my work colleagues Dr Aparna Shankar, Dr Nicola Shelton and Dr Jenny Mindell, for their help and encouragement especially towards the end of my PhD.

I also would like to acknowledge participants of the English Longitudinal Study of Ageing for their time and co-operation.

Finally, I want to give my special and deepest thanks to my mother, my sisters, my dearest son and my husband.

Table of Contents

List of Tables

List of Figures

List of abbreviations and acronyms

ADL: Activity of daily living

CASP: Control, Autonomy, Self-realisation, Pleasure

CESD: Centre for Epidemiological Study Depression

CHD: coronary heart disease

CSL: Compulsory School Leaving

ELSA: English longitudinal study of ageing

EM: Expectation Maximization

FCS: fully conditional specification

FIML: full information maximum likelihood

HRQOL: Health Related Quality of Life

MAR: Missing at Random

MCAR: Missing Completely at Random

MCMC: Markov Chain Monte Carlo

MI: Multiple Imputation

ML: Maximum Likelihood

MNAR: Missing Not at Random

MVNI: multivariate normal imputation

Post-MI: Post Myocardial Infarction

QoL: Quality of Life

SF-36: Short Form-36 Health Survey Questionnaire

Introduction to the thesis

The present thesis focuses on gender differences in quality of life and depression (depressive symptoms) among older people with coronary heart disease (CHD). Initially, the idea of this topic came from my supervisor Dr Elizabeth Breeze who, at the time I started my PhD (2006), was the UCL manager of the English Longitudinal Study of Ageing (ELSA). She found that gender differences in quality of life and depression while living with angina or a history of myocardial infarction had not been researched systematically. We believed that ELSA could provide a unique opportunity to assess gender differences in quality of life and depression among people with CHD in a large national sample of non-institutionalized older people living in England. Using the ELSA data to explore gender differences in quality of life and depression could contribute to the existing literature, since current findings mainly come from studies that have focussed on patients which tend to have follow-up periods of a year or less.

ELSA was designed to collect longitudinal data on health, disability, economics, social participation and social networks. The advantages of using these data are many, including the large sample size, as opposed to patient studies which usually have small samples; the long follow-up period (four years at the time this thesis started); the wide range of information collected on people's lives allowing inclusion of several important covariates in the analyses; and importantly, the opportunity to use a measure of quality of life specifically developed for older ages, the CASP-19, as opposed to disease-specific or health-related measures of quality of life. The term depressive symptoms will be used throughout the thesis to refer to people who were depressed because experiencing three or more depressive symptoms in the week prior to the interview.

A large population study most certainly offers methodological limitations and challenges. As a statistician, I have an interest in addressing methodological problems of epidemiological studies such as sources of error and uncertainty. Two of the potential sources of error that I am faced with in using ELSA data may arise from missing data and self-reported CHD.

In longitudinal studies, missing data often occur because subjects do not respond to certain questions (item non-response) or do not respond to a particular wave or drop-out

of the study (unit non-response or attrition). Both attrition and unit non-response contribute to lower the sample size and pose serious problems for researchers because missingness can affect properties of estimators and inferences. A large volume of methodological research is now devoted to the application of statistical techniques to handle missing data that arise in several settings and under different assumptions. Amongst the "state of the art" of missing data methods, the most widely used and recommended are maximum likelihood (ML) and multiple imputation methods (MI). I compare the performance of a ML method (full information maximum likelihood) against two MI methods (two-fold fully conditional specification and multivariate normal imputation) to deal with missing data in both continuous and binary outcomes as well as covariates. For doing this, I use a simulation study and the technique that is found to perform best is then applied to the original data. The thesis will then bring together substantive and methodological results.

The other common source of error related to self-reported data is misclassification bias. In ELSA, information on diseases is collected using self-reports. It was not possible to have access to medical records of participants or a linkage with the Hospital Episode Statistics. Therefore the self-reported CHD measure available in the ELSA study could not be validated. For this reason, I decided to undertake a sensitivity analysis in order to quantify the misclassification bias and its impact on the results.

Chapter 1 reviews the literature on gender differences in quality of life and depression amongst people with CHD, on missing data and self-reported CHD, in order to identify current gaps in the knowledge which then form the objectives and hypotheses of this thesis. Chapter 2 describes in detail the English Longitudinal Study of Ageing, the sample design, data collection and baseline sample characteristics. In Chapter 3 a comparison study of three techniques to handle missing data is conducted using a simulation study. First a literature review on missing data is presented, followed by a detailed description of the simulation study, results and discussion. In Chapter 4 the issue of gender differences in quality of life and depressive symptoms among people with CHD are explored longitudinally and missing data are imputed. First gender differences in quality of life and depressive symptoms are explored comparing people with CHD with those from a healthy group; second, men and women are compared with respect to four-year trajectories of quality of life and depressive symptoms. The chapter ends with conclusions drawn from the results. In Chapter 5 the sources of error due to

self-reported CHD are addressed. A sensitivity analysis based on an external validation study and a deterministic sensitivity analysis are conducted in order to quantify the misclassification bias and its impact on the results reported in Chapter 4. Chapter 6 discusses the results, draws the conclusions, highlights the implications of the results found and raises questions for future research.

Chapter 1 Literature Review

This chapter will explore the literature that is relevant to this study in order to identify current gaps in the knowledge which then form the objectives and hypotheses of the thesis. The first two sections of the chapter will review the literature on the existing evidence on gender differences in i) quality of life and ii) depression among people with coronary heart disease. The third section of the chapter reviews the literature on missing data and the robustness of self-reports of disease state in order to address methodological gaps to challenge the validity of quantitative evidence.

1.1 Introduction

Coronary heart disease (CHD) includes myocardial infarction (or heart attack) and angina pectoris. Myocardial infarction is caused by a blood clot that blocks one of the coronary arteries. The coronary arteries bring blood and oxygen to the heart, if the blood flow is blocked the heart muscle is damaged and might die. Angina pectoris is chest pain or discomfort that occurs if an area of the heart muscle does not get enough oxygen-rich blood. Although angina symptoms and those of a myocardial infarction are very similar, the amount of damage caused to the heart differs considerably. Angina is not a disease but a symptom of coronary heart disease which occurs on exertion and is relieved by rest. Therefore an angina episode does not cause any heart muscle damage because the blood flow is only temporarily blocked, while a myocardial infarction can permanently damage the heart muscle (Jevon, 2012).

Traditionally coronary heart disease was considered as a predominantly male disease and women were often excluded from studies. In the past 15 years this belief has been abandoned and some research on heart disease has taken an interest in gender-based differences. Myocardial infarction (also known as acute myocardial infarction or heart attack) is common in middle aged men, while angina is the predominant presentation of CHD among women (Lerner and Kannel, 1986; Wenger, 2002). Later cardiovascular complications of myocardial infarction, which can happen months or years after the first event has occurred, include angina pectoris, aneurysm, congestive heart failure and the risk of another myocardial infarction, while cardiovascular complications of angina are unstable angina and myocardial infarction. The quality of life and mental health of people living with angina or myocardial infarction are also affected; therefore an

improvement in both is considered an important outcome in the life of those people with CHD. While there are established findings of gender differences in the manifestations, incidence and risk factors for CHD (Charney, 1999; Wenger, 2002; Polk and Naqvi 2005; Stramba et al., 2006), the literature on gender difference in quality of life and depression in patients living with angina or myocardial infarction is not consistent.

Both quality of life and depression are important outcomes in the lives of older people. Over the past twenty years, many studies have shown that in later life, worsening in mental health, as measured by both clinical and symptomatic depression, is very common (Baldwin et al., 2003; Beekman et al., 1995; Beekman et al., 1999; Beekman et al., 2002; Copeland et al., 2004; Penninx et al., 1999) with considerable impact on well-being and disability (Beekman et al., 2002) and increased risk of dying (Penninx et al., 1999, Schoevers et al., 2000). However, depression has been found to be more common among women than men. The evidence of higher prevalence of depression (minor and major) in women than in men is one of the most widely documented findings, in both population-based and clinical studies (Dennerstein, 1993; Kessler, 2003). Gender differences in depression have been found to emerge during puberty (Angold et al., 1998) and to persist across the life course and during later life (Beekman et al., 1999; Cole and Dendukuri, 2003).

In recent years, there has been an increasing interest in quality of life assessment among older people in sociological, medical and health research. In the 1970s and 1980s structured dependence theory dominated social gerontology in Europe. Older age was considered a time of potential poverty, dependency, and declining physical and mental health. As a consequence, ageing was perceived to decrease quality of life (William, 1977). More recently, the concept of structured dependency has been abandoned and we have witnessed a shift of emphasis in social gerontology towards a positive view of older age as a period of life in which one is free from structured social roles (e.g. employment, parent of dependent children) and free to explore personal fulfilment (Grundy and Bowling, 1999). Laslett (1996) in ‘A fresh map of life’ recognizes a more positive dimension of ageing. He sees the ‘Third Age’ as the period during which people are freed from work and family constraints and have time to pursue personal interests and a good quality of life. As the population ages and survival from coronary events continues to improve, assessment of quality of life has also become an important and useful outcome measure for evaluating the impact of disease and benefits of

medical interventions. As a consequence, an increasing number of studies have focussed on the quality of life of older people with CHD (Brown et al., 1999; van Jaarsveld et el., 2001; Barbareschi et al., 2009).

1.2 Coronary heart disease, quality of life and depression

1.2.1 Coronary heart diseases and quality of life

Improvements in treatments for coronary heart disease (CHD) have resulted in decreased mortality and morbidity. These treatments serve two primary goals: first, to prevent further progression of the disease or mortality, and, second, to relieve symptoms and improve function and quality of life. It is believed that following non-fatal coronary events patients can go back to a near normal life, which also includes going back to work, suggesting therefore that quality of life after a coronary event should not deteriorate over the long term. However, it has been demonstrated that return to near normal life depends on many factors and may vary by gender.

There is the belief that an improvement in quality of life can be important as a primary outcome and in the determination of therapeutic benefit (Thompson and Yu, 2003). As a consequence, there has been increasing recognition of the importance of measuring quality of life among individuals following a cardiac event. Several classes of instruments have been designed to measure quality of life according to certain dimensions such as health (Turner-Bowker et al., 2002) (for which the term "health-related quality of life" (HRQOL) is used), subjective (Hyde et al., 2003; Power et al., 2005) and objective quality of life defined in terms of objective living conditions and material resources (e.g. standard of living) (Erikson, 1993). Most of the studies that explored quality of life among people with CHD used a health-related quality of life instrument, amongst which the Short Form-36 Health Survey Questionnaire (SF-36) (Turner-Bowker et al., 2002) is the most commonly used. The SF-36 contains 36 items which assess eight dimensions: physical functioning, role limitation due to physical problems, role limitation due to emotional problems, social functioning, depression, energy/vitality, pain, and general health perception. There are also two summary scores, one for mental health and one for physical health. Previous studies have reported negative effects on HRQOL following angina (Lyons et al., 1994) and myocardial infarction (Daly, 2000; Mendes de Leon et al., 1998) and more generally following a coronary event (Norris et al., 2004). However, findings from prospective studies are

mixed. While some found that specific domains of HRQOL improved in the long term (Wiklund et al., 1989; Brown et al., 1999), others showed that there was no improvement in HRQOL (Van Jaarsveld et al., 2001) or that the decline in HRQOL was limited to the immediate post-myocardial infarction period (Mendes de Leon et al., 1998).

Gender differences in quality of life among people with coronary heart disease

The majority of studies on quality of life in patients with CHD suggested that women do not cope as well as men, as concluded by a large review by Brezinka and Kittel (1996). However, the existing literature on the gender differences in quality of life of people with CHD is sparse and somewhat contradictory. Seven studies were found (Table 1.1) which have specifically addressed gender differences in quality of life of people with myocardial infarction (Westin et al., 1999; Bogg et al., 2000; Mendes de Leon et al., 2001; Kristofferzon et al., 2005a and 2005b; Brink et al., 2005; Norris et al., 2007) and only two addressed CHD (angina and/or myocardial infarction) (Norris et al.; 2004, Ford et al., 2008). Some of these studies did not find any significant gender differences in HRQOL among patients with myocardial infarction (Mendes de Leon et al., 2001; Kristofferzon et al., 2005b). Several studies measured HRQOL one month after myocardial infarction in men and women and found that women scored significantly lower than men in the following domains of HRQOL: mental health (Westin et al., 1999; Kristofferzon et al., 2005a), emotional health (Bogg et al., 2000), general health (Westin et al., 1999), physical health and functioning (Kristofferzon et al., 2005a), self-esteem and family interactions (Westin et al., 1999).

One year post-myocardial infarction (post-MI) women scored significantly lower than men on physical function (Brink et al., 2005; Norris et al., 2007), bodily pain, social function (Brink et al., 2005), mental health (Westin et al., 1999; Norris et al., 2007), general health (Westin et al., 1999) and emotional health (Bogg et al., 2000) domains of HRQOL. One year post-MI the gender differences found at one month post-MI in self-esteem and family interaction domains of HRQOL no longer persisted (Westin et al., 1999). Results from one study showed that gender differences in the mental health dimension of HRQOL at one year post-MI did not persist once the model was adjusted for demographic, clinical, co-morbid and psychosocial covariates (Norris et al., 2007).

Results from other studies reported an improvement in HRQOL after one year post-MI as compared to five months in both men and women (Brink et al., 2005) and between four months and one year post-MI (Kristofferzon et al., 2005b).

Results on gender differences in quality of life found in the literature are not consistent, and several limitations of the studies must be acknowledged. First, results cannot be generalised because all studies, except one (Ford et al., 2008) have used samples from selected community hospitals. The second limitation is related to the power of these studies. The study by Mendes de Leon et al., (2001) had a small sample size (41 women and 47 men). Therefore, the non-significant gender differences may be due to low power. Similarly, the studies of Kristofferzon et al., (2005a), Westin et al., (1999) and Brink et al., (2005), although prospective, had small sample sizes (74 women and 97 men in Kristofferzon et al., 2005a; 316 men and 97 women in Westin et al., 1999; 77 men and 37 women in Brink et al., 2005) which further decreased in the follow-up period (60 women and 88 men in Kristofferzon et al., 2005a; 288 men and 88 women in Westin et al., 1999; 65 men and 33 women in Brink et al., 2005). The study by Bogg and colleagues (2000) had a slightly higher sample size (220 participants); however it was limited to participants aged between 37 and 64. The study by Ford et al. (2008) used a large population-based sample; however, the results are limited by the cross-sectional nature of the data. Only the study by Westin et al., (1999) and the study by Ford et al., (2008) compared the results of the CHD sample with those found in a control group (i.e. free from CHD); while the study by Brink et al., (2005) compared the HRQOL scores of the sample with myocardial infarction with Swedish normative scores. Another potential limitation of these studies is related to the measure of quality of life used. All studies have used a specific measure of quality of life which is either disease specific or focuses on the health-related aspects of quality of life. Of the seven studies, four measured health-related quality of life using the SF-36 questionnaire (Mendes de Leon et al., 2001; Kristofferzon et al., 2005a; Brink et al., 2005; Norris et al., 2007).

Ford et al., (2008) used the Centre for Disease Control Prevention (CDC HRQOL-4) measure which had four questions assessing the number of physically and mentally unhealthy days over the month prior to the interview. Westin et al. (1999) constructed a multifaceted questionnaire to measure quality of life, which had six dimensions measuring general physical health, heart-related physical health, depression, experience

of social life and self-esteem. The extent to which this is a well-validated instrument is unknown. As well as the SF-36, Kristofferzon and colleagues (2005a) used Quality of Life Index-Cardiac (QLIC) version, which was a modified version of the Ferrans and Powers Quality of Life Index (Ferrans and Powers, 1985; 1992). The questionnaire was a self-administered disease specific measure and as such it focused on the complaints that were attributable to the specific characteristics of the disease (Smith et al., 2000). The findings of the SF-36 questionnaire in Kristofferzon's study were not supported by the QLIC version. This was most likely due to the low sensitivity of the QLIC to changes over time and to the low Cronbach's alpha coefficient of this instrument (Kristofferzon et al., 2005a). Bogg and colleagues (2000) used a modified version of the Quality-of-Life after Myocardial Infarction (QLMI) questionnaire; a disease-specific measure of health-related quality of life which was developed to evaluate a comprehensive cardiac rehabilitation programme for patients after myocardial infarction (Hillers et al., 1994).

The problem related to disease-specific measures, such as QLMI and QLIC, and the more generic measure of health-related quality of life (like the SF-36), is that when they are applied to a cardiac population, they usually lack sensitivity – the ability of a measure to detect important changes. This limitation of the measure has been confirmed by Smith and colleagues (2000) where they compared the sensitivity of the above mentioned three measures. The SF-36 aims to measure subjective health-status; it is often referred as a health-related quality of life instrument although it does not have any underlying theoretical conceptualisation for quality of life (Carr and Higginson, 2001). This measure is clearly focussed on the impact of poor health on the eight dimensions measuring the various aspect of people's life (Higginson and Carr, 2001). Health-related quality of life measures are based on proxies (such as health) "which draw on a set of normative assumptions about what a particular condition implies for a person's quality of life without necessarily taking close account of a person's current life experience" (Wiggins et al., 2008:4). Health is only one dimension of wider quality of life (Bowling, 2005) and cannot be considered as a proxy for a holistic concept of quality of life. In addition, the domains used to measure health-related quality of life, such as body pain, physical functioning, health perception, depression, are all dimensions that influence the quality of life of individuals and they are often on the pathway between health and quality of life. The result is that the measure of health-related quality of life is not distinct from what influences it.

Table 1.1 Studies on gender differences in quality of life of people with CHD

First author	Design	Number and age range of subjects	Main outcome	CHD variables	Measures	Results	Comparison with disease-free population	Term effect	Length of follow-up
Westin L et al. (1999)	P	413 aged under 70	HRQOL	AMI, PCI, CABG	They have constructed a new measure	One month post AMI women compared to men had decreased HRQOL in the domains of GH, FI, MH, SE. One year post AMI women had poorer HRQOL in the GE and MH domains only.	Yes	Decline in both short and long term	1 month and 1 year
Bogg J et al. (2000)	P	220 aged 37-64	QoL	MI	A	Women had poorer EM of QoL compared with men in the short and long term.	No	Both short and long term decline	3-4 days, 1,3 and 6 months
Mendes de Leon et al. (2001)	CS	88 aged 35-89	HRQOL	AMI	B	No gender differences in HRQOL.	No	-	-
Kristofferzon ML et al. (2005a)	CS	171 aged 30-80	HRQOL, QoL	MI	B, C	Women reported lower HRQOL (MH, PF) than men one month post MI and lower levels of QoL in particular health and function dimensions.	No	-	-

Table 1.1Continued

Kristofferzon ML et al. (2005b)	P	171 aged 30-80	HRQOL, QoL	MI	B, C	Increased HRQOL in PF, ER, V, SF over time in both men and women. The QoL did not increase over time No gender differences in HRQOL and QoL.	No	Improvement in HRQOL at 1 year follow-up and but decline in QLI	1, 4 and 12 months
Norris C et al. (2004)	P	3,392 aged ≥18	HRQOL	CHD	D	Women had poorer HRQOL (PF,RP,BP,SF) compared to men.	No	Decline in the long term	1 year
Brink E et al. (2005)	P	98	HRQOL	AMI	B	Women had poorer HRQOL (MH, PF) compared to men. An improvement in HRQOL was observed for both genders 1 year post-AMI: women had lower scores than men (PF, BP, SF).	No	Decline in HRQOL in the first 5 months and then improvement	1 week, 5 months and 1 year
Norris C et al. (2007)	P	486 Mean age 59 men,66 women	HRQOL	AMI	B	Women had poorer HRQOL (MH, PF) compared with men in unadjusted models and no gender differences after adjustment in MH dimension of HRQOL.	No	Decline in HRQOL overtime	1 year
Ford ES et al. (2008)	CS	50,573 aged ≥18	HRQOL	Self-report CHD	E	Women had poorer HRQOL compared to men.	Yes	-	-

Abbreviations: CS: Cross-sectional P: Prospective R: Review HRQOL: Health-Related Quality of Life. QoL: Quality of Life. MI: myocardial infarction. AMI: acute myocardial infarction. PF: physical functioning SF: social functioning MH: mental health BP: bodily pain V: vitality EM: emotional reaction GH: General health FI: Family interaction SE: self-esteem. A: Quality of life after Myocardial Infarction Questionnaire B: Short-Form Health Survey 36 (SF-36) C: Ferrans & Powers Quality of life Index-Cardiac version D: Seattle Angina Questionnaire (SAQ). E: Centers for Disease and Control Prevention HRQOL-4

1.2.2 Coronary heart disease and depression

Over the past 30 years there has been growing evidence of the relationship between CHD and depression. Depression even at low levels of severity has been found to increase the risk of developing heart diseases (Strike and Steptoe, 2002, Hemingway et al., 2003; Marzari et al., 2005). Reviews of the literature on depression and CHD risk supported the view that psychological problems before and after coronary events not only increase the risk of a cardiac event but also all-cause mortality (Hemingway and Marmot, 1999; Rozanski et al., 1999). Established evidence has demonstrated that depression is common following episodes of myocardial infarction or angina and is associated with increased mortality (Lesperance et al., 1996, 2000; Ziegelstein et al., 2000; Bogg et al., 2000; Ferketich et al., 2000; Bush et al., 2001; Carney et al., 2003, Lane et al., 2005; Parashar et al., 2006), with poor adherence to recommended behaviours and lifestyle changes after the cardiac event (Ziegelstein et al., 2000, 2001) and with an increased risk of readmission because of cardiac complications (Lauzon et al., 2003). Moreover, patients with symptoms of depression after myocardial infarction were less likely to return to work (Stern et al., 1977; Schleifer et al., 1989). Amongst the possible explanations for the mortality risk associated with depression following myocardial infarction, some suggested that depressed patients are less likely to adhere to treatment and lifestyle changes than non-depressed patients (Stansfeld et al., 2002). Another possible explanation was that depression leads to decreased heart rate variability, with a greater risk of fatal arrhythmias (Stansfeld et al., 2002).

Gender differences in depression among people with coronary heart disease

Several studies (Table 1.2) have assessed depression in men and women after myocardial infarction (Forrester et al., 1992; Wiklund et al., 1993; Frasure-Smith et al., 1999; Bjerkeset et al., 2005; Naqvi et al., 2005; Brink et al., 2005; Mallik et al., 2006; Norris et al., 2007). Some studies did not find any gender differences in depression among patients with myocardial infarction (Wiklund et al., 1993; Brink et al., 2005; Norris et al., 2007). Others found that women were more likely to be depressed than men following myocardial infarction (Forrester et al., 1992; Frasure-Smith et al., 1999; Naqvi et al., 2005; Mallik et al., 2006; Bjerkeset et al., 2005).

Longitudinal studies exploring gender differences in depression among people with myocardial infarction, did not report any difference between men and women in their prevalence of depression at five months (Brink et al., 2005) or at one year (Wiklund et al., 1993; Brink et al., 2005) post-MI. However, women had improved in that they reported less depression at one year post-MI than they had at five months post-MI (Brink et al., 2005). Others found that one year post-MI women were more likely than men to be depressed (Norris et al., 2007).

Bjerkeset et al., (2005) specifically addressed gender differences in depression during the five years after myocardial infarction. Women had a high initial risk for depression, with a significant decrease after two years, while in men the risk for depression was only increased in the two to five years post-MI (Bjerkeset et al., 2005).

Most of the studies on depression and myocardial infarction had a focus on depression as a risk factor for the development of myocardial infarction or cardiac mortality (Lesperance et al., 1996, 2000; Frasure-Smith et al., 1999; Ferketich et al., 2000; Bush et al., 2001; Carney et al., 2003, Lane et al., 2005; Marzari et al., 2005; Parashar et al., 2006) or have focussed on the prevalence of depression following heart disease (Wiklund et al., 1993; Brink et al., 2005; Parashar et al., 2006; Mallik et al., 2006; Norris et al., 2007). All of these studies have used samples from selected community hospitals. The search of the literature failed to find any study that has addressed gender differences in depression following angina. Only one study specifically addressed gender differences in depression five years after myocardial infarction using a large population-based study of adults aged from 35 to 79 and used a comparison group of people free from myocardial infarction (Bjerkeset et al., 2005). However, this study had several limitations: first, there was an 11 year gap between the first and the second interview; second, depression post myocardial infarction was only measured once (at follow-up) and, as a consequence, they could not assess the course of depression over time. Moreover, to assess whether depression was present at baseline they used an index of anxiety and depression (available at baseline only), which however was not internationally validated.

Table 1.2 Studies on depression in men and women with CHD

First author	Design	Number and age range of subjects	Main outcome	CHD variables	Measures	Results	Comparison with disease-free population	Term effect	Length of follow-up
Forrester et al. (1992)	CS	129	Major Depression	AMI	-	Female gender was a significant predictor of depression in people with AMI.	No	-	-
Wiklund et al. (1993)	P	595 aged 56-83	Depression	MI	Self-reported depression	No gender differences in the prevalence of depression.	No		1 year
Freasure-Smith et al. (1999)	CS	896 Mean age 58 men, 63 women	Symptoms of depression	MI	F	Women more likely than men to report symptoms of depression in hospital after MI.	No	-	-
Bjerkeset et al. (2005)	P	23,693 aged 35-79	Anxiety and depression	MI	G	Women were more likely than men to have depression in the first 2 years post MI, while men were more likely than women to have depression in the 2 to 5 years post MI.	Yes	-	2 and 5 years
Naqvi et al. (2005)	R	-	Depression	AMI	-	Depression was more prevalent in women post-AMI than in men, and depressive symptoms persisted longer.	-	-	-
Brink E et al. (2005)	P	98 Mean age 71 men, 65 women	Anxiety and depression	AMI	G	No gender differences were found in depression after 5 months and 1 year post-AMI.	No	Women reported less depression at 1 year compared to 5 months	5 months and 1 year

Table 1.2 Continued

Mallik et al. (2006)	CS	2,498	Depression	AMI	H	Women were more likely to have depression than men.	Yes	-	-
Norris C et al. (2007)	P	486 Mean age 59 men,66 women	Depression	AMI	F	Women were more likely than men to be depressed at 1 year post-AMI , no gender differences were found at baseline.	No	Women reported worsening depression at 1 year post-AMI	1 year

Abbreviations: CS: Cross-sectional P: Prospective R: Review HRQOL: Health-Related Quality of Life. QoL: Quality of Life. MI: myocardial infarction. AMI: acute myocardial infarction. PF: physical functioning SF: social functioning MH: mental health BP: bodily pain V: vitality EM: emotional reaction GH: General health FI: Family interaction SE: self-esteem. F: Back Depression Inventory (BDI). G: Hospital Anxiety and Depression Scale (HADS) H: Primary Care Evaluation of Mental Disorders Brief Patient Health Questionnaire

1.2.3 Summary and conclusions

Although the statistics show that CHD is common among women, its recognition in women is still hampered by misconceptions about it being a man's disease. Studies in the 1990s have indicated that morbidity and mortality of women with CHD exceeds that of men (Greenland et al., 1991; Brezinka and Kittel, 1996) and researchers have increasingly focussed on gender differences in CHD. There are established findings of gender differences in the manifestations, incidence and risk factors of myocardial infarction and angina (Sharp, 1994; Charney, 1999; Wenger, 2002; Polk and Naqvi, 2005; Stramba et al., 2006). For example, among people aged over 50 years, myocardial infarction is the most frequent presenting feature of CHD for men (Lerner, 1986) whereas angina is for women (Charney, 1999).

Following a non-fatal myocardial infarction event or angina symptoms most patients can return to a near normal life. For many people this is not the case. Moreover, there may be clear gender differences in adaptation to the disease. Following episodes of myocardial infarction or angina, patients are at risk of reporting symptoms of depression (Forrester et al., 1992; Frasure-Smith et al., 1999; Bjerkeset et al., 2005; Naqvi et al., 2005; Mallik et al., 2006), pain and physical functioning (Mendes de Leon et al., 1998; Brown et al., 1999; Van Jaarsveld et al., 2001), and women are particularly at higher risk than men. Although it is evident that angina or myocardial infarction can have considerable impact on well-being, and that the consequences may vary between men and women in older age, gender differences in depression and quality of life while living with angina or a history of myocardial infarction at older age have not been researched systematically.

The review of the literature on gender differences in quality of life and depression among people living with CHD identified some limitations and gaps in current knowledge. To summarise, there are no studies that have addressed gender differences in depression following angina. This issue is important because angina is the predominant presentation of CHD among women (Lerner and Kannel, 1986; Wenger, 2002), while myocardial infarction is more common in men. To overcome this limitation this thesis will explore angina as well as myocardial infarction.

Among the few studies that have addressed the issue of gender difference in quality of life and depression among people with myocardial infarction, all have used a specific measure of quality of life, the health-related quality of life. For this reason, in this thesis I will adopt a measure of quality of life specifically developed for old age called CASP comprising four domains ('control', 'autonomy', 'pleasure' and 'self-realization'). CASP has been based on theories of need satisfaction (Doyal and Gough, 1991), which assume that quality of life at older ages is conceptualized as the degree to which human needs are satisfied in the above mentioned four domains (Hyde et al., 2003). This measure differs from health-related measures by focusing on positive aspects of quality of life and by being independent of health and other factors that might influence it (Hyde et al., 2003; Wiggins et al., 2008).

Only four studies (Westin et al., 1999; Bjerkeset et al., 2005; Brink et al., 2005; Ford et al., 2008) compared some characteristics of people with CHD with those from a control group or a community sample of people free from disease. However, none of the studies specifically explored whether gender differences in depression and quality of life were found in the control group. Using a healthy population as a comparison group might help to understand whether similar results are found in the two groups and provide highlights into gender differences in quality of life and depressive symptoms among people with CHD. For this reason, in this thesis I will be comparing results with a reference group of healthy individuals free from CHD, chronic diseases and any limiting longstanding illness.

Only one study was conducted in England (Bogg et al., 2005), therefore the extent to which results can be generalised to the population in England is not known. All studies except two (Bjerkeset et al., 2005; Ford et al., 2008), used small samples from selected community hospitals which might affect generalisation of results and none has specifically focussed on an older population. This latter issue is important because most CHD events occur in people aged 50 and over, where declines in quality of life and depression may be perceived as a "normal" consequence of ageing. Moreover, considering that a relatively high proportion of people aged 50 years can now expect to live for further 30 years or more and could potentially enjoy an active and healthy lifestyle, it is important to understand the impact that a non-fatal coronary event can have in later life. To overcome these limitations this thesis will use a large national sample of people aged 50 and over.

Lastly only some of the studies have adjusted their analyses for covariates other than age and gender (Norris et al., 2004, 2007; Bjerkeset et al., 2005; Brink et al., 2005; Ford et al., 2008). The covariates that this thesis will consider for adjustment in the analyses are age, cohabiting status, retirement status, education, wealth, smoking status, physical activity, alcohol consumption, pain, physical functioning, social support and social networks. These variables are fully explained in section 2.2.3 of Chapter 2.

1.3 Common sources of error and uncertainty in epidemiological studies

1.3.1 Missing data

The previous section of this chapter reviewed the literature on gender differences in quality of life and depression among people with CHD. Most of the studies found in the literature used samples from selected community hospitals and only two studies (Ford et al., 2008; Bjerkeset et al., 2005) used population-based samples. Samples from community hospitals are relatively small and non-response and attrition introduce bias and contribute to lower the sample size. Population-based samples on the other hand have higher sample sizes but, they are also affected by non-response and attrition. None of the studies reviewed have dealt with or acknowledged the problem of missing data in their results.

Researchers in epidemiology and other disciplines are often faced with the problem of incomplete data sets, particularly when the study aims at collecting a large number of characteristics for each individual. In longitudinal studies, missing data often occur because subjects do not respond to certain questions. This situation is often referred in the literature as item non-response. Other situations common in longitudinal studies are those in which subjects do not respond to a particular wave or drop-out of the study (because of moving out or death). This situation is known as unit non-response (in contrast to initial unit non-response it is prerequisite that the respondent has participated in at least one wave). The terms attrition, drop-out, loss to follow-up and withdrawal are used interchangeably in the literature to refer to this latter form of missingness. Therefore, in longitudinal studies the problem could be severe since we face several

types of missing data, such as item non-response, unit non-response and drop-out (Clarke and Hardy, 2007).

Missing data can pose serious problems for researchers because missingness can affect properties of estimators and inferences. Ignoring missing data also affects the accuracy and precision of parameter estimation. The seriousness of the problem depends in part on how much data are missing. There is no clear rule regarding how much is too much missing data. This is because potential bias is inherent whenever observations are missing (Kline, 1998). Also the number or proportion of missing observations alone is not sufficient to indicate whether missing data are an issue or not. Rather the impact of missingness is determined by the research question, the information in the observed data, and the reason for the missing data.

The implication of missingness for the analysis depends on the missingness mechanism, which is usually unknown. In handling missing data it is important to differentiate among three missing data mechanisms (Little & Rubin, 2002; Rubin, 1976): missing completely at random (MCAR), missing at random (MAR), missing not at random (MNAR)' (Rubin, 1976, Kline, 1998, Schumacker & Lomax, 1996). MCAR refers to data where the missingness mechanism does not depend on the variable of interest or any other variable (does not depend on observed and unobserved data) (Scheffer, 2002). With MCAR the missing data are a simple random sample of all data values, therefore MCAR reflects the highest degree of randomness of the missing data mechanism and shows no underlying reasons for missing observations that can potentially bias research findings. In practice this means that, under MCAR, the analysis of only those units with complete data gives valid inferences (Musil et al., 2002). For example, MCAR data can occur when respondents accidentally skip a question on a questionnaire or if the participant accidentally discarded the questionnaire. In these situations there is no underlying pattern to the missing observations that would contribute to biased data. With MAR the missingness depends only on the components of a variable that are observed and not on those that are missing (Little and Rubin, 2002). MAR data show some randomness to the pattern of data omission: "For example, in a study of dietary intake, if participants with depression are less likely than those without depression to record their daily intake, then depression is a variable that predicts missing observations" (Musil et al., 2002:816). MAR has a very special and important role in longitudinal studies where, essentially, it implies that future drop-out is conditionally

independent of future values, given all observed past values. Another way of expressing MAR in the longitudinal/drop-out setting is to say that the future statistical behaviour of those who share the same history of measurements is the same whether they drop-out or not. MNAR or non-ignorable missing data occurs when missingness is related to the values that would have been observed. This is the most difficult condition to model. Non-ignorable missing data have systematic non-random factors underlying the occurrence of the missing values that are not apparent or otherwise measured. Non-ignorable missing data affect generalisability of research findings and may bias parameter estimates, also the direction of bias is unpredictable (Musil et al., 2002).

The ELSA study is subject to missing data due to item non-response and attrition. Using a simulation study I will compare three techniques for dealing with missing data in order to find the best method to be applied to the ELSA data.

The reader is referred to Chapter 3 for a literature review on missing data and for a study of three techniques for handling missing data.

1.3.2 Self-reported measure of disease

The findings reviewed in section 1.2 of the chapter mainly come from community hospitals samples, which use a clinical diagnosis of CHD. Only two studies have used population-based samples from which data on CHD diagnosis were self-reported (Ford et al., 2008; Bjerkeset et al., 2005) and was not validated with medical records or a clinical diagnosis. Most epidemiological studies and health surveys assess the presence of chronic diseases from self-report, as opposed to clinical assessments mainly because the collection of self-reported conditions involves lower costs (Kriegsman et al., 1996). However, to use self-reported data to assess CHD with confidence, it is important to know the validity of these measures. Clearly, inaccurate reporting of CHD by surveyed populations may result in people not being identified early for chronic disease-related illnesses or not being offered interventions, such as changes in health behaviours. In terms of findings, results from studies using self-report measures might be subject to misclassification bias.

Misclassification bias is defined as the systematic error due to erroneous classification. When assessing misclassification bias of a test or measure sensitivity and specificity must be considered. The terms sensitivity and specificity are used to measure the

effectiveness of a test procedure in relation to a certain disease. Sensitivity is the proportion of those with the disease that are identified as positive by the test, therefore sensitivity measures how well the test detects a disease. Specificity is the proportion of those without the disease that are identified by the test as not having the disease. It follows that specificity refers to how well the test detects absence of disease (Armitage, Berry and Matthews, 2002). The ideal value of both sensitivity and specificity is 100% indicating no misclassification. However, the relationship between these two measures tend to be inverse that is, the more sensitive a test procedure, the less specific it is likely to be, and vice versa.

Previous studies have reported high values of specificity and sensitivity of self-reported CHD (Haapanen et al., 1997; Baumeister et al., 2010). Although a number of validation studies have suggested that self-reported CHD is reasonably accurate when compared with medical records (Bush et al., 1989; Okura et al., 2004; Merkin et al., 2007; Yamagishi et al., 2009; Barr et al., 2009; Lampe et al., 1999; Baumeister et al., 2010), the extent to which self-reported measures introduced bias in the findings of epidemiological studies is an issue rarely addressed quantitatively (Jurek et al., 2006). Since it is not possible to validate self-reported CHD cases in ELSA using medical records, through a sensitivity analysis I will investigate the extent to which the self-reported measure of CHD used in this thesis may lead to biased estimates and/or different conclusions in the results.

The reader is referred to Chapter 5 for a literature review on validation studies on self-reported CHD and for a sensitivity analysis investigating bias due to misclassification of self-reported CHD.

1.4 Objectives and Hypotheses

In the previous sections of this chapter a review of the literature on gender differences in quality of life and depression among older people living with CHD was presented and common sources of error and uncertainty occurring in epidemiological studies were reviewed. To summarise, the search of the literature identified that previous studies have not examined gender differences in depressive symptoms or quality of life in a national sample of older adults suffering from both myocardial infarction and angina,

while adjusting the analysis for a set of important covariates and making comparisons with a healthy reference population. Additionally, none of the studies reviewed in the literature have dealt with the issues of missing data. Many epidemiological studies involve large numbers of individuals and large numbers of variables (especially with repeated measurements over time), and complete data are rarely available, therefore addressing the problem of missing data has been recommended in order to improve the validity of epidemiological research results and to reduce estimation bias caused by missing data (Sterne et al., 2009).

All studies reviewed in the literature, with the exception of two (Ford et al., 2008; Bjerkeset et al., 2005), used objective measure of disease. However, large epidemiological studies and health surveys assess the presence of chronic diseases from self-report. It is likely that the validity of self-reported diagnosis may vary depending on the severity of the disease and it is possible that misclassification introduces bias in the estimates. Despite these common problems, the studies that have used self-reported CHD have only acknowledged the issue as a limitation (Ford et al., 2008; Bjerkeset et al., 2005).

These gaps identified in current knowledge form the key objectives for investigation.

The specific objectives are:

1. To explore gender differences in quality of life and depressive symptoms among men and women with CHD (angina and/or myocardial infarction) and compare them with healthy people of similar age (who do not have CHD or any limiting long-standing illness or chronic disease), in order to understand whether results are similar to the group of people with CHD.
2. To compare men and women with respect to trajectories over time of quality of life and depressive symptoms over four years once they have experienced CHD.
3. To compare different methods for dealing with missing data in longitudinal studies using the full information maximum likelihood, the multivariate normal imputation and two-fold fully conditional specification techniques as examples, in order to find the best method that yields unbiased results when applied to the data.
4. To assess the sensitivity of the models to different assumptions about the reliability of self-reported CHD measure.

This PhD thesis sets out to test the following hypotheses:

1. People aged over fifty years who have experienced CHD are at higher risk of experiencing depressive symptoms and poor quality of life than those who have not. Among people with CHD there are also significant gender differences in quality of life and depressive symptoms, with women being at higher risk than men of reporting depressive symptoms and lower quality of life.
2. The shapes of trajectories over time of quality of life and depressive symptoms are different in men and women following the CHD event. Women tend towards a time-limited reaction (in terms of depressive symptoms and poor quality of life) to the actual CHD event, while men seem less able to adapt to the long-term consequences of the event.
3. Ignoring missing data will give biased results.
4. Self-reported CHD is a robust and reliable measure.

The hypotheses will be tested on a large national sample of older people living in England, participants of the English Longitudinal Study of Ageing (ELSA). ELSA is the only longitudinal study in England to cover men and women from age 50 onwards with rich data on many different aspects of people's lives. In addition, the Whitehall II study is used for the sensitivity analysis (refer to Chapter 5 for a detailed description of the study).

The next chapter describes in detail the data set used in this thesis which comprises three waves of the English Longitudinal Study of Ageing.

Chapter 2: The English Longitudinal Study of Ageing

The first chapter reviewed the literature on gender differences in quality of life and depressive symptoms among older people living with CHD. From the gaps in the literature hypotheses and objectives were defined. This chapter describes the data sets and measures used in this study to test the hypotheses.

2.1 Data source: The English Longitudinal Study of Ageing

The data sets come from the first three waves of the English Longitudinal Study of Ageing (ELSA), collected between 2002-03 and 2006-07 (Marmot et al., 2003; Banks et al., 2006; Banks et al., 2008). ELSA is a panel study where individuals aged 50 and over are followed and re-interviewed every two years. The aim of ELSA is to explore the unfolding dynamic relationships between health, functioning, social networks and economic position. It is in effect a study of people's quality of life as they age beyond 50 and of the factors associated with it.

Each survey involves a face-to-face interview and every four years there is a subsequent visit by a nurse in which biomedical data are taken such as blood pressure and anthropometric measurements as well as blood and saliva samples.

The ELSA sample was designed to represent people aged 50 and over, living in private households in England and was selected from households that had previously responded to the Health Surveys for England (HSE) in 1998, 1999 or 2001 and had a household member born before March 1952. The HSE is an annual cross-sectional household survey that collects a wide range of health data and biometric measures. Each of the main HSE adopted a multi-stage stratified probability sampling design in which postcode sectors have been the primary sampling units within which addresses were selected with a probability proportional to their size (number of addresses). Within each sector, addresses were then selected systematically. Full details of the sample design and response rates for the Health Surveys for England have been published elsewhere (Erens and Primatesta 1999; Erens, Primatesta, and Prior 2001; Bakejal et al., 2003). Taking the three HSE years used for the ELSA sample together, a total of 31,051 households were sampled. Of these, 23,382 households responded to HSE. The ELSA sample was only selected from households that responded to HSE. Furthermore,

households were only issued to field if they included at least one age–eligible individual who was living in the household at time of the HSE interview, born on or before 29 February 1952 and who, according to administrative records, remained alive and gave permission to be re-contacted in the future.

ELSA wave 1 (2002-03) achieved 11,391 productive interviews with eligible sample members. The survey achieved a household response rate of 70%; approximately 96% of individuals responded within households. This equates to an overall individual response rate of 67%. The survey also completed 636 productive interview with partners aged under 50 and 72 interviews with new partners (whose presence in the household only became known after the sample was issued). For the purpose of this thesis, younger partners and new partners were excluded from the analyses. Eligible sample members who responded at this stage were renamed 'core members' to distinguish them as the core element of the continuing ELSA sample. Core members were eligible for the main interview in wave 2 unless they had since died, had explicitly asked at the end of the first ELSA not to be re-contacted, or had moved out of England. A total of 8,781 core members (response rate 81.5%) participated in wave 2 (2004-05) and at wave 3 (2006-05) there were 7,114 respondents (response rate 73%). The response rate is defined as "total individual respondents to a given wave divided by total individuals eligible for that wave"; participants that moved into institutions or from England to another country at a particular wave were no longer eligible. Inclusion in either numerator or denominator was not conditional upon response in any previous wave. For the purpose of this study, wave 1 will be referred to as "baseline", wave 2 as the two-year follow-up and wave 3 as the four-year follow-up, since the data are collected at two-yearly intervals.

2.2 Measures

2.2.1 Outcomes: depressive symptoms and quality of life

The eight-item version of the Centre for Epidemiologic Study Depression scale (CESD-8) administered during the face-to-face interview was used to measure depressive symptoms (Radloff, 1977). The questions asked the degree to which the respondent had experienced (or not) depressive symptoms, such as restless sleep, being unhappy and so on, over the past week (Table 2.1). The total score ranges from 0 to 8 (items 4 and 6

were reverse coded for scoring), which was recoded as: 0 '0-2 symptoms' of depression and 1 '3+ symptoms' of depression, this cut-off has been used by the Health and Retirement Study to classify respondents as being depressed (Steffick, 2000).

Table 2.1 CESD-8 item wording

	Item statements
Much of the time during the past week,	
1	you felt depressed?
2	you felt that everything you did was an effort?
3	your sleep was restless?
4[a]	you were happy?
5	you felt lonely?
6[a]	you enjoyed life?
7	you felt sad?
8	you could not get going?

[a] Item reverse coded for scoring

Quality of life was measured using the CASP-19 questionnaire available in the self-completion booklet. CASP-19 contains 19 items (Table 2.2) covering four conceptual domains of individual needs that are particularly relevant in later life: Control, Autonomy, Self-realization and Pleasure. The instrument has four items for the control domain and five for each of the others. Each item is assessed on a four-point Likertscale (rated 'this applies to me: often, sometimes, not often, never') numerically coded so that the most positive response was scored as 3 and the most negative response as 0. Items 3, 5, 7, and 10 to 19 were reversed coded for the calculation of the total score so that all item responses were in the same direction. The resulting scale scores are summed to form an index with higher scores indicating better quality of life (Cronbach's alpha=0.67) as recommended by Wiggins et al., (2008).

The psychometric properties of CASP-19 are fully described by Hyde et al., (2003), while Wiggins et al., (2008) evaluated the properties of CASP-19 using data from the English Longitudinal Study of Ageing (ELSA). The total score (theoretically ranging from 0 to 57, but in ELSA ranges from 6 to 57) of CASP-19 (with higher scores indicating better quality of life) was used.

Table 2.2 CASP-19 item wording arranged by domain categories

Item statements	
Control	
1 My age prevents me from doing the things I would like to do	
2 I feel that what happens to me is out of my control	
3 I feel free to plan for the future	*Item reverse coded for scoring*
4 I feel left out of things	
Autonomy	
5 I can do the things I want to do	*Item reverse coded for scoring*
6 Family responsibilities prevent me from doing the things I want to do	
7 I feel that I can please myself what I do	*Item reverse coded for scoring*
8 My health stops me from doing the things I want to do	
9 Shortage of money stops me from doing things I want to do	
Pleasure	
10 I look forward to each day	*Item reverse coded for scoring*
11 I feel that my life has meaning	*Item reverse coded for scoring*
12 I enjoy the things that I do	*Item reverse coded for scoring*
13 I enjoy being in the company of others	*Item reverse coded for scoring*
14 On balance, I look back on my life with a sense of happiness	*Item reverse coded for scoring*
Self-realisation	
15 I feel full of energy these days	*Item reverse coded for scoring*
16 I choose to do things that I have never done before	*Item reverse coded for scoring*
17 I feel satisfied with the way my life has turned out	*Item reverse coded for scoring*
18 I feel that life is full of opportunities	*Item reverse coded for scoring*
19 I feel that the future looks good for me	*Item reverse coded for scoring*

Table adapted from Wiggins et al, 2008

2.2.2 Main exposure: coronary heart disease

During the interview participants were asked whether a doctor had ever told them that they suffered from angina or myocardial infarction/heart attack, and if so, whether they had angina symptoms or myocardial infarction in the past two years. This information was complemented with information on the age at which the respondents first had angina or myocardial infarction.

Having had coronary heart disease (first or recurrent angina and/or myocardial infarction) in the two years preceding the baseline interview (2002-03) is the main exposure variable. In addition to CHD, it would have been ideal to consider angina and myocardial infarction as two separate exposures. However, due to the small number of cases and the lack of power that this would introduce in the analyses, only CHD is used.

People with CHD are compared with a reference population of healthy individuals (Well group) defined as people that at baseline (2002-2003) had never had CHD, stroke, diabetes, pulmonary disease, Alzheimer, Parkinson's, cancer or any limiting longstanding illness.

2.2.3 Covariates

The covariates considered in this study were identified from the literature as potentially influencing each outcome and as potentially being correlated with the exposure. Bivariate analyses of baseline data were used to confirm whether each covariate considered was related to the outcomes and also correlated to the exposure (CHD). The covariates are described in detail below.

Socio-demographic variables

Socio-economic status, employment status and marital status have all been found to relate to well-being among older adults. Low educational qualification, poorest wealth and living without a partner were found to decrease quality of life among older people (Netuveli et al., 2006; Zaninotto et al., 2010; Webb et al., 2009). Several studies have also reported that older women are more likely than men to be exposed to factors associated with depression such as lower education, lower income, less skilled occupations and to widowhood (Arber and Cooper, 1999; Barefoot et al., 2001). Socio-

economic status, employment status and marital status have also been shown to correlate with coronary heart disease (Bunker et al., 2003; Sproston and Primatesta, 2004).

During the face-to-face interview information regarding age, gender, marital status, educational attainment, occupational status as well as total wealth were collected. From the questions on marital status and cohabitation status (contained in the self-completion questionnaire) a cohabiting status variable with three categories was derived as follows: living with a partner (married or not), not living with a partner (previously married), never married and not cohabiting; the variable was then recoded as 0"living with a partner (married or not)" and 1 "not living with a partner (including never married)". The "never married" cases were only 5% of the sample for that reason they were grouped with the "not cohabiting" cases.

The compulsory school-leaving (CSL) age has increased steadily since state-sponsored education was first recognised as a right for all children in the UK. CSL age was originally set at 10, which increased to 13 in 1899 and then to 15 in 1944 (which took effect in 1947). To account for different durations in compulsory education when determining the educational attainment of the sample, I used information on the self-reported age of first leaving full-time education. From responses to these questions I derived a variable of education with three categories: those leaving at or after age 19 (referred to as 'high' education), those leaving school after CSL but before age 19 (referred to as defined as 'mid' education) and those who left at or before CSL (referred to as 'low' education). The variable was then coded as 0 "high and medium education" and 1 "low education". High and medium education categories were grouped because only about 13% of the sample was in the "high" education category.

From the questions on employment status a variable was derived to describe people in paid employment (full-time, part-time, self-employed or semi-retired), completely retired and other (permanently unable to work, not currently in paid employment, looking after home or family).

Total non-pension wealth is defined as the sum of financial worth, physical worth (such as business wealth, land or jewellery), and housing wealth after deducting debts; it

represents a better measure of the permanent economic status of older people than income (Marmot, 2003; Banks et al., 2006; Banks et al., 2008).

Health covariates

There is established evidence of decreased physical functioning following CHD (Mendes de Leon et al., 1998; Brown et al., 1999; van Jaarsveld et al., 2001) and increased bodily pain (Brown et al., 1999). Also, limitations in physical functioning (activity of daily living) has been found to predispose older adults to a decreased quality of life (Netuveli et al., 2006) and increased risk of depression (Stuart-Shor et al., 2003; Stek et al., 2004). Smoking, alcohol consumption and physical inactivity are risk factors for cardiovascular disease (Sproston and Primatesta, 2004). There is evidence of a positive relationship between physical activity and quality of life (Bize et al., 2007) and both inverse and positive relationships between physical activity and depression (van Gool et al., 2003). Among older people alcohol consumption has been found to positively relate to quality of life and negatively relate to depression (Zaninotto et al., 2010). Smoking was found to decrease quality of life (Zaninotto et al., 2010) and increase the risk of depression (Zaninotto et al., 2010; Colman et al., 2011).

A dichotomous variable (no and yes) was used to assess whether respondents were often troubled by pain. Respondents were also asked to report whether because of a physical, mental, emotional or memory problem they have any difficulty with Activities of Daily Living (ADLs), such as dressing (including putting on shoes and socks), walking across a room, bathing or showering, eating (such as cutting up food), getting out of bed, using toilet (including getting up or down). From this question a variable was derived to count the number of difficulties with ADLs (range 0 to 6).

From information collected on smoking status a variable was derived with three categories: never smoked, ex-smoker, current smoker. The mean quality of life and the prevalence of depressive symptoms did not differ significantly between those that never smoked and ex-smokers, therefore for simplicity the variable was then coded as 0 “never smoked and ex-smoker” and 1 “current smoker”.

From the questions on frequency of alcohol consumption a variable with three categories was derived as follows: “not at all or occasionally”, “once or twice a week”, and “three and more times a week and daily”. The prevalence of those not drinking or drinking occasionally was about 27% , for that reason this the variable was then recoded as 0”less than three times a week” and 1”three times a week and more”.

Lastly, from questions on frequency of leisure-time physical activity a variable was derived according to the definition given by McMunn et al., (2003, chapter 6: 212-213) which summarised leisure-time physical activity into three ordinal categories defined as “high activity (vigorous activity)”, “medium (moderate and low moderate activity)”, and “inactive (sedentary)”. Only 16% of people were in the medium category of physical activity, therefore the variable was then recoded as 0 “physically active (high activity and medium)” and 1 “inactive”.

Social support and social networks

There is evidence that lack of social support and poor social networks are associated with CHD risk and recurrence (Brezinka and Kittel, 1996; Wang et al., 2005). On the other hand, evidence has also shown that lack of social support, and social network, independent from CHD, are risk factors for well-being (Victor, 2005; Netuveli et al., 2006).

The self-completion questionnaire included a series of detailed items on the quality of the respondent’s social relationships. Specifically, respondents were asked about the closeness of their marital relationship on a scale from “not at all close”=1 to “very close”=4; about the presence of positive support from their spouse, children, other relatives and friends (how much they understand the way the respondent feels about things, how much they can be relied on if the respondent has a serious problem and how much the respondent can open up to them to talk about worries). Positive support items were scored as 1=”not at all” and 4=”a lot”, such that higher numbers indicate more of each type of support. Responses were summed to create a positive aspects of social relations scale (based on perceptions of empathy, availability of help and being able to confide, Cronbach’s alpha 0.68). The total score ranges from 0 to 36, with higher scores

indicating greater positive support (Stafford et al., 2011). Those respondents without a spouse or children were given the lowest value of positive support from that source.

Respondents were also asked to indicate the number of family members and friends with whom they had a close relationship. From this question a continuous variable was derived that indicates the number of close friends/family in the respondent's social networks, at baseline 1% of people did not have any close friends/family.

2.3 Sample characteristics at baseline (wave 1, 2002-03)

2.3.1 Sample description

This section describes the samples that are being analysed in this thesis, those with CHD and the comparison group of Well people. The group of people reporting CHD in the two years preceding the baseline interview (2002-03) is composed of 518 men and 376 women. The Well group is defined as people that at baseline did not report CHD, stroke, diabetes, pulmonary disease, Alzheimer, Parkinson's, cancer or any limiting longstanding illness (1,701 men and 1,892 women).

In descriptive analyses shown in next section age-standardisation has been used in all tables in which age is not included as a break variable. In comparing categories of the break variable, age-standardisation reweights the sample in each category of the break variable so as to give all categories the same age profile. In this way it is possible to remove the effect of age from comparisons between groups. Direct standardisation was applied for both men and women, with the standards being the age distribution of core members of the whole ELSA sample at baseline (consisting of 11,391 individuals).

All descriptive analyses have also been weighted for non-response and account for the complex survey design. The statistical package used was Stata version 10.

Prevalence of CHD

Figure 2.1 presents the age-standardised gender distribution in the CHD and Well groups. A higher prevalence of men report having had a CHD event in the two years preceding the baseline interview, compared to women (27%, and 19% respectively,

p<0.01); while more women than men were in the Well group (81% and 73% respectively, p<0.001).

Figure 2.1 Age-standardised gender distribution of the CHD and Well groups (2002-03)

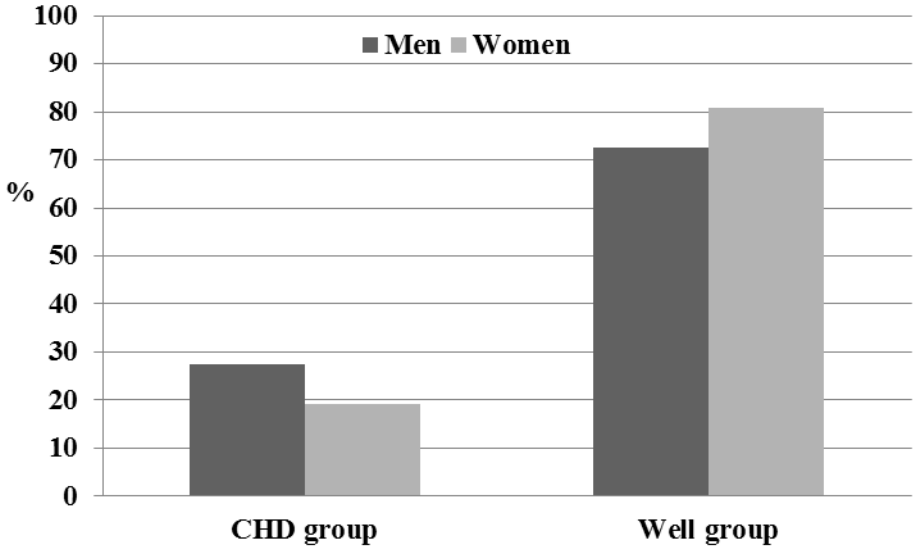


Table 2.3 reports prevalence of CHD by age and gender. The prevalence of CHD increases significantly with age in men and women. For men the prevalence of CHD is highest from the age of seventy, while for women it is highest from the age of eighty.

Table 2.3 Prevalence of CHD, by age and gender (2002-03)

	Men % (95% CI)	Women % (95% CI)
50-54	9.5 (7.2,12.5)	6.0 (4.2,8.4)
55-59	16.6 (13.5,20.3)	8.7 (6.4,11.6)
60-64	19.8 (15.8,24.6)	14.9 (11.4,19.4)
65-69	26.6 (22.1,31.6)	21.7 (17.3,26.8)
70-74	36.8 (31.1,42.9)	27.1 (22.0,32.9)
75-79	36.0 (29.1,43.6)	26.2 (20.3,33.0)
80+	36.1 (29.6,43.2)	29.0 (23.4,35.3)
Total	22.1 (20.4,23.8)	16.6 (15.1,18.3)
Bases unweighted	*518*	*376*

The CHD group is defined as people reporting the first or recurrent angina or myocardial infarction event in the two years preceding the baseline interview.

2.3.2 Quality of life and depressive symptoms

Quality of life

The age-standardised mean quality of life for the ELSA general sample (based on 11,391 respondents) at baseline was 42.3 (S.D. 8.7) in men and 42.8 (S.D. 8.7) in women. Among people with CHD the age-standardised mean quality of life of men was 39.8 (S.D. 11.0) and that of women was 39.9 (S.D. 8.8). Men and women in the Well group had higher mean quality of life than people with CHD ($p<0.001$). There were no gender differences in either group in mean quality of life (Table 2.4).

Table 2.4 Mean quality of life, by disease status and gender (2002-03)

	CHD group	**Well group**
	Mean (S.D)	**Mean (S.D)**
Men	39.8 (11.0)	44.7 (8.7)
Women	39.9 (8.8)	44.5 (10.2)
Bases unweighted		
Men	*456*	*1,516*
Women	*304*	*1,699*

Age-standardised figures.

Linear regression was performed to look at the association between quality of life at baseline and disease status interacted with gender and adjusted for age. Men and women in the CHD group had quality of life that was around four points lower than people in the Well group (β=-4.4 [95% CI:-5.4;-3.4] for men and β=-4.8 [95% CI: -5.9;-3.8] for women). The interaction term between gender and disease group was not significant, suggesting that the association between the disease groups (CHD and Well group) and quality of life did not differ by gender.

Depressive symptoms

At baseline the overall prevalence of depressive symptoms in the general sample was 20.1% in men (95% CI: 18.9; 21.2) and 28.1% in women (95% CI: 26.9; 29.4).

Among women with CHD the age-adjusted prevalence of depressive symptoms was significantly higher than in men (36% and 27% respectively); in the Well group women also reported higher prevalence of depressive symptoms than men.

The prevalence of depressive symptoms of men and women with CHD was over 10 percentage points higher than men and women in the Well group ($p<0.001$) (Table 2.5).

Table 2.5 Prevalence of depressive symptoms, by disease status and gender (2002-03)

	CHD group	**Well group**
	% (95% CI)	**% (95% CI)**
Men	27.3 (23.2,31.8)	17.1 (15.1,19.3)
Women	36.1 (31.2,41.3)	25.6 (23.1,28.2)
Bases unweighted		
Men	*518*	*1,701*
Women	*376*	*1,892*

Age-standardised figures.

Logistic regression was performed to examine the association between depressive symptoms and disease status interacted with gender, adjusted for age. Men and women with CHD had higher odds (OR=1.6 [95% CI:1.3; 2.1] in men and OR=1.8 [95% CI:1.4; 2.3] in women) of reporting depressive symptoms than people in the Well group. However, the interaction term between disease status and gender was not significant.

2.3.3 Socio-demographic characteristics

Women with CHD were on average two years older than men. Also in the Well group women were on average older than men (Table 2.6). Among those with CHD 47% of women were not cohabiting with a partner, a prevalence that was 16 percentage points higher than that of men in the same group ($p<0.001$). These gender differences were also found in the Well group. Men with CHD were more likely than women to be in

paid employment (p<0.001), and less likely to be in the poorest quintile of wealth (p<0.01). In the Well group men were more likely than women to be in paid employment (p<0.001), but there were no gender differences in wealth. Among people with CHD, 62% of men and 65% of women were retired. The proportion of retired people did not differ significantly by gender neither in the CHD nor in the Well group.

Table 2.6 Socio-demographic characteristics of the sample, by disease status and gender (2002-03)

	CHD group		**Well group**	
	Men	**Women**	**Men**	**Women**
Mean Age (S.D)	67.9(11.7)	70.3 (11.6)	62.1 (9.6)	63.5 (10.3)
Bases unweighted	*518*	*376*	*1,701*	*1,892*
	%(95% CI)	**%(95% CI)**	**%(95% CI)**	**%(95% CI)**
Cohabiting status				
Not living with partner	30.7	46.5	27.3	31.9
	(26.5,35.3)	(41.6,51.4)	(25.1,29.6)	(30.2,33.7)
Bases unweighted	*518*	*376*	*1701*	*1892*
Educational level				
Low education	61.3	56.3	51.4	50.2
	(57.0,65.4)	(50.8,61.6)	(48.8,54.0)	(47.7,52.8)
Bases unweighted	*518*	*376*	*1697*	*1889*
Employment status				
In paid employment	25.4	11.9	49.5	42.0
	(21.7,29.4)	(8.8,16.1)	(47.3,51.7)	(39.8,44.4)
Completely retired	61.8	65.3	43.7	42.5
	(58.3,65.2)	(59.9,70.3)	(41.5,45.9)	(40.2,44.9)
Other	12.8	22.7	6.9	15.4
	(9.9,16.3)	(18.6,27.4)	(5.6,8.3)	(13.7,17.3)
Bases unweighted	*518*	*376*	*1701*	*1892*
Total Wealth quintile				
Poorest quintile	22.8	31.7	15.8	18.2
	(19.1,27.0)	(27.4,36.4)	(13.6,18.4)	(16.1,20.6)
2nd	24.4	18.8	17.2	17.6
	(20.5,28.8)	(15.0,23.3)	(15.2,19.5)	(15.6,19.7)
3rd	20	20.8	20.5	19.9
	(16.6,24.0)	(16.8,25.6)	(18.4,22.7)	(17.9,22.0)
4th	17.2	16.2	21.7	21.2
	(14.0,21.0)	(12.7,20.5)	(19.6,23.9)	(18.9,23.6)
Richest quintile	15.6	12.4	24.8	23.2
	(12.4,19.4)	(9.2,16.6)	(22.3,27.4)	(20.9,25.6)
Bases unweighted	*517*	*371*	*1,688*	*1,865*

Age-standardised figures. N figures are unweighted

Men with CHD differed from men in the Well group in that they were more likely to have a low educational qualification, more likely to be completely retired, substantially less likely to be in paid employment and more likely to be in the poorest and second quintile of wealth. Compared to women in the Well group, women with CHD were less likely to be cohabiting with a partner, more likely to have a low educational qualification, less likely to be in paid employment, more likely to be completely retired, and more likely to be in the poorest and second quintile of wealth (Table 2.6).

2.3.4 Health characteristics

Among those with CHD 41% of men and 50% of women reported being often troubled with pain ($p<0.001$). Similarly, women from the Well group were more likely than men to be often troubled with pain. In the group with CHD the gender difference in the prevalence of pain was greater than the Well group (Table 2.7).

Twenty nine percent of men and 34% women with CHD reported having one or more difficulties in performing activities of daily living (ADLs). There was no gender difference in the prevalence of difficulties with ADLs in either the CHD group or in the Well group. There were no gender differences for people with CHD in physical activity, smoking status and alcohol consumption. In the Well group men were significantly less likely than women to be physically inactive and more likely to consume alcohol on three or more days a week.

Men and women with CHD differed from the Well group in that they were more likely to be often troubled with pain, to report one or more difficulty with ADLs, more likely to be physically inactive and to drink alcohol on three or more days a weeks (except for women) (Table 2.7).

Table 2.7 Health characteristics of the sample by disease status and gender (2002-03)

	CHD group		Well group	
	Men	**Women**	**Men**	**Women**
	%(95% CI)	**%(95% CI)**	**%(95% CI)**	**%(95% CI)**
Pain				
Often troubled with pain	40.9	49.9	17.3	21.5
	(36.3,45.6)	(44.4,55.3)	(15.5,19.3)	(19.3,23.9)
N	*506*	*363*	*1,679*	*1,871*
ADL				
No difficulties	71.3	66.4	91.7	91.0
	(67.4,74.9)	(61.4,71.0)	(89.7,93.4)	(88.9,92.7)
1	14.0	16.2	6.4	5.6
	(11.3,17.4)	(12.6,20.6)	(4.9,8.4)	(4.4,7.2)
2	5.6	7.5	0.8	1.9
	(3.8,8.4)	(5.3,10.5)	(0.5,1.3)	(1.1,3.2)
3	3.0	5.5	0.6	0.6
	(1.8,4.9)	(3.7,8.1)	(0.3,1.3)	(0.2,1.4)
4	3.4	2.7	0.3	0.7
	(2.1,5.5)	(1.4,5.3)	(0.1,0.6)	(0.2,2.3)
5	2.2	1.5	0.1	0.1
	(1.7,2.9)	(0.6,3.3)	(0.0,0.4)	(0.0,0.4)
6	0.3	0.3	0.1	0.1
	(0.1,1.3)	(0.1,1.3)	(0.0,0.5)	(0.0,0.3)
N	*506*	*363*	*1,679*	*1,871*
Physical Activity				
Inactive	70.7	74.3	58.7	66.4
	(66.4,74.8)	(69.2,78.9)	(56.2,61.1)	(64.3,68.4)
N	508	363	1,679	1,875
Smoking status				
Current smoker	15.5	17.8	16.5	19.0
	(12.4,19.2)	(13.9,22.5)	(14.8,18.3)	(17.2,20.9)
N	508	363	1,678	1,875
Alcohol consumption				
≥3 days a week	27.8	23.1	33.1	25.8
	(23.7,32.2)	(18.9,28.0)	(30.6,35.6)	(23.6,28.1)
N	*505*	*363*	*1,679*	*1870*

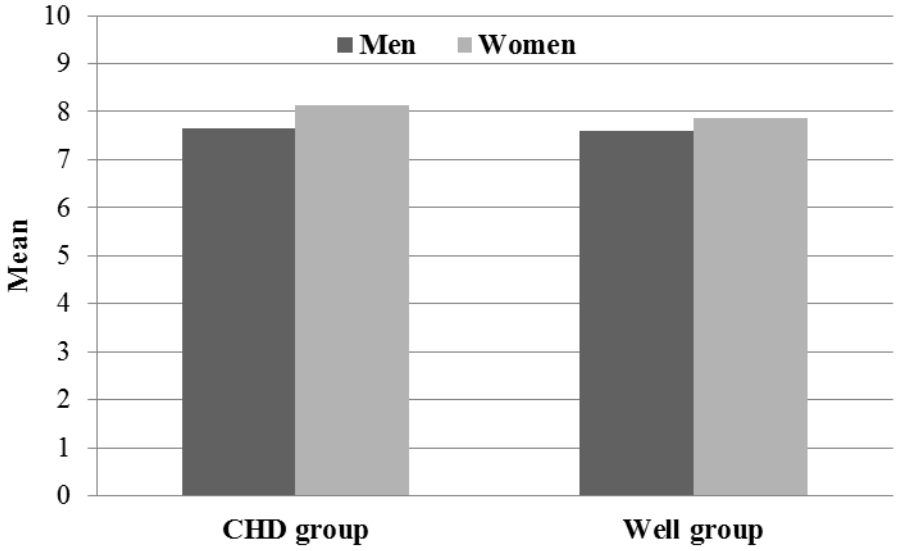
Age-standardised figures. N figures are unweighted

2.3.5 Social networks

Figure 2.2 reports the age-standardised average number of family members and friends with whom respondents have a close relationship. On average, men and women with CHD had 8 close relatives and friends. Similar results were found in the Well group.

There were no gender differences in the number of close relatives and friends, in either group.

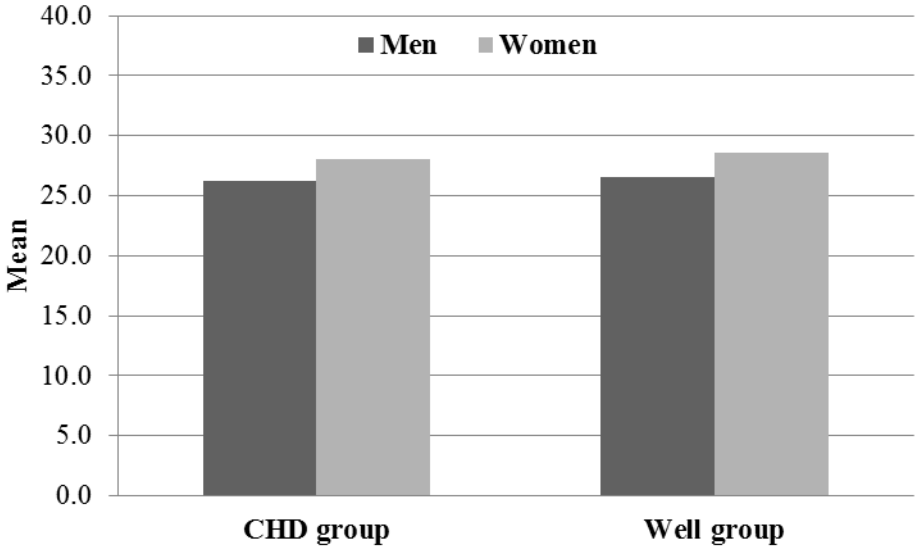
Figure 2.2 Mean number of close family and friends (2002-03)



Age-standardised figures

Women received on average more social support from their spouse, children, immediate relatives or friends, regardless to whether they were in the CHD or Well group. The age-standardised mean social support in people with CHD was similar to that of people in the Well group (Figure 2.3).

Figure 2.3 Mean score of social support*, by gender (2002-03)



Age-standardised figures. *Higher scores indicate more social support

2.4 Summary

This chapter described the English Longitudinal Study of Ageing (ELSA) in terms of data collection, sample, and variables being used. Baseline characteristics of the sample under study were presented. It was found that at baseline (2002-03) men were more likely than women to have CHD. Results showed that woman with CHD (but also women in the Well group) were more likely than men to report depressive symptoms, while there were no gender differences in the mean score of quality of life neither in people with CHD nor in the Well group.

The descriptive analysis has also identified some gender differences in socio-demographic characteristics of men and women with CHD such as cohabiting status, employment status and wealth. With the exception of wealth, the Well group also presented gender differences in cohabiting status and employment status. In terms of health characteristics, the main difference between men and women with CHD was in the prevalence of being often troubled with pain, which was higher in women; a finding that was also true for the Well group. Men and women with CHD did not differ in health behaviours while gender differences in physical activity and alcohol consumption were found in the Well group. Women received more social support than men; this was true in both the CHD and Well groups. Preliminary unadjusted results support the hypothesis that there are gender differences in depressive symptoms of people with CHD but equally to that of men and women in the Well group. However, the hypothesis of gender differences in quality of life of people with CHD was not supported by the cross-sectional analysis of baseline data.

The hypotheses and objectives of this thesis will be tested using three waves of data from ELSA in Chapter 4. Before that the next chapter sets out to compare three recently developed with suitable software techniques for handling missing data in order to find the best technique that can be applied to the ELSA data.

Chapter 3: A simulation study to evaluate three key strategies to handle missing data

The previous chapter described the ELSA data and presented baseline characteristics of the sample under investigation. This chapter aims at finding the best technique to deal with the problem of missing data in the ELSA data set.

3.1 Review of the methods for handling missing data

It has already been highlighted in Chapter 1 that missing data can pose serious problems for researchers because missingness can affect properties of estimators and inferences. The choice of the appropriate method to handle missing data strongly depends on the data available (Durrant, 2005). Several methods for dealing with missing data have been developed and are not mutually exclusive. These methods are grouped into "unprincipled and principled" methods.

Unprincipled methods are:

1. *Procedures based on completely recorded units*: these are called complete-case and available-case methods. Complete-case analysis is limited to observations for which all values are completely recorded (Little and Rubin, 2002).
2. *Last observation carried forward (LOCF)*: this method is specific to longitudinal data problems. For each individual, missing values are replaced by the last observed value of that variable.
3. *Single-imputation procedures*: with these procedures the missing values are filled in and the completed data is analysed using standard methods. The most commonly used single-imputation methods are mean (or median) imputation which replaces missing values with means (or medians) and regression imputation where missing values are replaced by single predicted values from the regression analysis (Little and Rubin, 2002).

A key problem of unprincipled methods is that inferences about parameters do not account for uncertainty therefore the standard errors are systematically underestimated, p-values of tests are too small and confidence intervals are too narrow (Little and Rubin, 2002).

Principled methods are based on a well-defined statistical model for the complete data and on assumptions about the missing data mechanism (MCAR or MAR). The subsequent analysis, inferences and conclusions are valid under these assumptions. This does not mean the assumptions are necessarily true but it does allow the dependence of the conclusions on these assumptions to be investigated (Kenward and Carpenter, 2007).

Principled methods include:

1. *Weighting procedures*: in this approach, a model for the probability of missingness is fit, and the inverse of these probabilities are used as weight for the complete-case analysis.
2. *Simple stochastic imputation:* instead of replacing a value with a mean, this method uses a random draw from a suitable distribution. Provided the distribution is chosen appropriately, consistent estimators can be obtained from methods that would work with the whole data set. In the large survey setting draws are made from units with complete data that are 'similar' to the one with missing values. There are several variations of this approach which use non-parametric estimates of the distribution of the missing data.
3. *Maximum likelihood estimation:* this approach assumes multivariate normality and MCAR or MAR. Maximum likelihood uses all of the available data to identify parameter values that have the higher probability of producing the sample data (Baraldi and Enders, 2010). Inferences are based on the observed data likelihood that links the observed data and the parameters. Formally this is obtained from the complete data likelihood by adding together the likelihood contribution over all possible values of the missing data, also known as integration (averaging) over the missing data from the joint density of the observed values and the missing values (Sinharay et al., 2001). This approach is most commonly known as full information maximum likelihood (FIML), because by using direct ML estimation, parameter estimates are obtained for cases with complete and incomplete data.
4. *Multiple-imputation (MI)*: is a likelihood-based method that incorporates the uncertainty into the imputation process. MI is based on a Bayesian paradigm where the model parameters are independently drawn from the posterior distribution for each imputed data set (Rubin, 1987, 1996). In the frequentist

approach, MI fixes the model parameters at the maximum likelihood estimates for all imputed data sets (Robins and Wang, 2000). MI is comprised of three stages: imputation stage, which generates a specified number of data sets (*m*), each of which contains different estimates of the missing values; analysis stage, in which each of the imputed data set is analysed using the same technique that would have been used had the data been complete; and the pooling stage where the estimates and their standard errors are averaged into a single set of values. The pooling stage is done according to Rubin's formula (Rubin, 1987), in order to yield a final result that combines the uncertainty in the data and the uncertainty due to missing values. Let $\hat{\beta}_i$, $\text{var}(\hat{\beta}_i)$, $i=1, \ldots, M$ be the M complete-data estimators and their associated variances, Rubin's rules for the combined estimate is:

$$\hat{\beta} = \frac{1}{M}\sum_{i=1}^{M}\hat{\beta}_i \qquad (1)$$

which is the average of the $\hat{\beta}_i$, and the estimator of the variance of this is a simple combination of within-and between-imputation variability:

$$\text{var}(\beta) = \frac{1}{M}\sum_{i=1}^{M}\text{var}(\hat{\beta}_i) + \frac{1}{M-1}\sum_{i=1}^{M}(\hat{\beta}_i - \hat{\beta})^2 \qquad (2)$$

To illustrate the pooling stage, I use a data set consisting 20 subjects for which the quality of life score was calculated (first column of Table 3.1). The score of quality of life is missing for 8 subjects. The last five columns of the table show five imputed data sets for quality of life. The multiple imputation mean is the average of the five estimates in Table 3.1, which is

$$\hat{\beta} = \frac{40.6 + 41.8 + 39.9 + 41.7 + 43.6}{5} = 41.5$$

The variance is:

$$\text{var}(\beta) = \frac{(8.7)^2 + (7.0)^2 + (8.2)^2 + (7.9)^2 + (8.1)^2}{5} +$$
$$+ \frac{(40.6-41.5)^2 + (41.8-41.5)^2 + (39.9-41.5)^2 + (41.7-41.5)^2 + (43.6-41.5)^2}{4} = 65.8$$

It is therefore easy to see that the variance is the average variance of the 5 imputed data sets plus a correction factor that quantifies the extent to which the estimates vary across data sets. The pooled standard deviation is simply the square root of the variance (S.D.= 8.1).

Table 3.1 Example of an observed variable with missing data and five imputed data sets

Observed data	Imputed quality of life				
Quality of life	Data set 1	Data set 2	Data set 3	Data set 4	Data set 5
52	52	52	52	52	52
-	46	36	42	39	49
50	50	50	50	50	50
49	49	49	49	49	49
47	47	47	47	47	47
-	18	54	26	51	51
45	45	45	45	45	45
-	34	45	45	33	53
41	41	41	41	41	41
-	51	37	32	51	48
44	44	44	44	44	44
43	43	43	43	43	43
38	38	38	38	38	38
-	42	41	48	42	54
36	36	36	36	36	36
-	40	40	40	42	43
-	42	42	31	46	42
33	33	33	33	33	33
-	36	39	31	28	30
24	24	24	24	24	24
Mean 41.8	40.6	41.8	39.9	41.7	43.6
S.D. 8.0	8.7	7.0	8.2	7.9	8.1

In epidemiologic research missing data are common; however, despite continuing methodological development, most of the studies in the literature do not report the handling of missing data (Chan and Altman, 2005). Traditionally, the analysis of incomplete data has been dominated by unprincipled methods (Enders, 2001; Klebanoff and Cole, 2008). As these approaches lack a principled foundation, it follows that these methods often behave unexpectedly in different settings (Kenward and Carpenter, 2007). For example, procedures based on completely recorded units assume that all

incomplete data arise from a MCAR process. If the missing data mechanism is MCAR, then using these methods in certain settings may result in nonbiased parameter estimates, although there might be loss of precision (Little and Rubin, 2002) and additionally "such settings are typically narrow and often unrealistic and difficult to establish" (Kenward and Carpenter, 2007:202). Therefore if any data are missing due to a MAR or NMAR mechanism, results will be biased and inferences invalid. In single imputation procedures only one value is imputed for each missing item. It follows that the most obvious limitation of single imputation is the underlying assumption that the imputed value is the true value. This limitation leads to underestimation of the variance, which affects confidence intervals and statistical tests.

Principled methods are usually preferred to unprincipled ones because they produce estimates that are superior and also because unprincipled methods have no theoretical rationale. However, these methods are infrequent in published epidemiologic manuscripts (Klebanoff and Cole, 2008). Among the principled methods that have received considerable attention in the methodological literature of the past 20 years are full information maximum likelihood (FIML) and Multiple Imputation (MI). They are considered the "state of the art" missing data techniques (Shafer and Graham, 2002) and are highly recommended in the methodological literature (Shafer and Olsen, 1998; Enders, 2006). These techniques give unbiased estimates with MCAR and MAR data, therefore they are superior to unprincipled methods (such as deletion and single imputation approaches) (Baraldi & Enders, 2010). Amongst the main MI techniques, the multivariate normal imputation (MVNI) and fully conditional specification (FCS) have been widely used and tested in different settings. For this reason they will be reviewed together with FIML.

The FIML technique does not impute, or fill in missing values, but directly estimates model parameters. Cases with incomplete data are included in computations and all available data are employed by the ML algorithm to obtain optimal parameter estimates under the assumption of multivariate normality. FIML estimation is available in structural equation modelling software (e.g., AMOS, MPLUS). In recent years several studies have been reporting the efficiency of FIML estimation relative to unprincipled methods (procedures based on completely recoded units and single imputation) when the missing data is MCAR and MAR (Artbuckle; 1996; Baraldi and Enders, 2010; Enders and Bandalos, 2001; Shin et al., 2009; Wotke, 2000). Others also recommend

the use of FIML when data are MNAR (Enders 2001; Wiggins and Sacker, 2001). It has also been shown that FIML yields superior performance to traditional unprincipled methods in multiple regression analyses. Thus the use of FIML can be extended to statistical analysis other than structural equation modelling (Enders 2001; Baraldi and Enders, 2010). However, to date and to my knowledge, the performance of the FIML has not been explored in the context of clustered data, such as multilevel modelling in the presence of missing data in both outcome and independent variables.

The MVNI procedure is based on the assumption that the data arise from a multivariate normal distribution. Findings arising from analyses of multiply imputed data sets when the joint multivariate normality assumption is violated are robust to this assumption violation, as long as the statistical models used to subsequently analyze the imputed data properly account for the data's non-normality. Schafer (1997) cites simulation studies and provides his own simulation evidence to illustrate the robustness to non-normality of imputation generating models that assume joint multivariate normality when the number of missing data is moderate (eg, < 50%) and the amount of non-normality in variables not severe. Recent development of computer packages performing MVNI (such as SAS and NORM) include variable transformation and categorical variable rounding utilities that may further improve the performance of multiple imputation conducted under the assumption of joint multivariate normality. Shafer (1999) states that MVNI can be easily applied to longitudinal data set with missing values. One study found in the literature supported the MVNI approach in the longitudinal setting with a monotone pattern of missing data (Newman, 2003). A data set is said to have a monotone missing pattern when a variable Y_j is missing for the individual *i* implies that all subsequent variables Y_k, $k>j$, are also missing for the individual *i*. It was shown that MVNI outperformed traditional approaches to missing data (Newman, 2003). However, to date and to my knowledge, no previous studies explored the performance of MVNI in a multilevel modelling setting with non-monotone missing data in continuous and binary outcomes.

The FCS approach does not start from an explicit multivariate density. Instead, it involves a variable-by-variable approach using chained equations (van Buuren et al. 1999; van Buuren et al. 2007). The imputation model is specified separately for each variable according to its type (linear regression for continuous variables, logistic regression for binary variables, ordinal logistic regression for ordinal variables and so

forth), involving the other variables as predictors. This is the main advantage of this approach (Molenberghs and Kenward, 2007). The FCS procedure is useful when the specification of a joint multivariate distribution of all the variables with missing values is difficult. However, from a theoretical standpoint this technique is problematic because the sequence of regression models might not be consistent with a true joint distribution (Shafer and Graham, 2002). This means that the iterative algorithm might never converge because the joint distribution to which they might converge does not exist. Despite the lack of a satisfactory theory, FCS seems to work quite well in many applications. A number of simulation studies provide evidence that FCS generally yields estimates that are unbiased and that possess appropriate coverage, at least in the variety of cases investigated (Brand et al., 2003; Brand, 1999; Raghunathan et al., 2001; van Bureen et al., 2006; Horton and Lipsitz, 2001). Recently, Nevaleinen et al., (2009) proposed an extension of the FCS to repeated measurement settings, the so called two-fold fully conditional specification, which increases the imputation model by conditioning also on variables measured at other times. Using a simulation study on dietary data the authors demonstrate that the two-fold FCS is a suitable approach for imputing time dependent covariates or repeated measurements (Nevaleinen et al., 2009). Despite the fact that could be computationally intensive in the presence of many follow-up years, the efficiency in recovering parameter estimates makes this approach appealing and investigating its performance in the presence of missing data with continuous and categorical variables (both outcomes and covariates) represents an innovative area of research that can contribute to the literature of missing data in the longitudinal setting.

This review is not exhaustive; there are in fact several approaches that can be used to generate MI data sets. For example, it has been suggested that if a data set to be imputed is multilevel, then the imputation model should be multilevel too (Carpenter and Goldstein, 2004). In recent years Carpenter and Goldstein have developed macros that implement multiple imputation in a multilevel data setting in MLwiN for normal and non-normal models of interest under the assumption of missing at random (Carpenter and Goldstein, 2004; Goldstein et al., 2009). The macros set up a multilevel multivariate imputation model with the partially observed variables as responses, and fit this model in a Bayesian framework with uninformative priors using Markov Chain Monte Carlo methods to impute a number of complete data sets. However, one of the main limitations is that the macros cannot handle missing categorical variables. For this and

other reasons that are discussed in details in the final discussion of this chapter, this method has not been considered in this thesis.

3.1.2 Summary of review

Until recently, the analysis of data with missing observations has been dominated by unprincipled methods (Enders, 2001; Klebanoff and Cole, 2008). However, alternative approaches for treating missing data have become increasingly common and are now available in many software packages, including two "state of the art" missing data methods: maximum likelihood and multiple imputation (Shafer and Graham, 2002). These methods have been recommended by the methodological literature as advantageous techniques that yield unbiased estimates with MCAR and MAR data (Baraldi & Enders, 2010). As a consequence, a large volume of methodological research is devoted to the application of these techniques to missing data that arise in several settings and under different assumptions. Amongst the "state of the art" missing data methods, the most widely used and recommended are full information maximum likelihood (FIML), multivariate normal imputation (MVNI) and fully conditional specification (FCS). To summarise, the FIML and MVNI methods assume multivariate normality, while two-fold FCS has the ability to handle different variable types (continuous, binary, unordered and ordered categorical) since each variable is imputed using its own imputation model. FIML does not impute missing values; instead it uses all available data (complete and incomplete) to identify the parameter values that have the highest probability of producing the sample data. FIML is relatively easy for the analyst to use and is widely available in structural equation programs. It is therefore important to understand whether when applying FIML to repeated measures and in particular in the presence of a non-continuous outcome it performs as well as more complex MI techniques.

Recently, the use of the FIML, MVNI and FCS has become increasingly popular in the literature and a number of comparisons between them under different settings and assumptions have been published (Collins et al., 2001; Newman, 2003; Acock, 2005; Ibrahim et al., 2005; van Bureen et al., 2006; van Buuren, 2007; Yu et al., 2007; Buhi et al., 2008; Marshall et al., 2010; Peyre et al., 2011). The FIML has been widely used with longitudinal data; the MVNI can be easily extended to repeated measures and recently the two-fold fully conditional specification has been proposed as an extension

of the FCS to repeated measures (Nevaleinen et al., 2009). Despite this, no recent studies have examined and compared the performance of the FIML, MVNI and FCS techniques in the repeated measures setting. Therefore investigating the performance of these techniques in a repeated measure setting represents an innovative area of research. In particular, the comparisons of performance of these techniques in the presence of missing data in continuous and binary outcomes, as well as in covariates, in a multilevel modelling analysis, can contribute to the literature of missing data in the longitudinal setting.

In order to investigate the usefulness of the ML and MI methods in the context of repeated measures, a simulation study is set up, where full information maximum likelihood (FIML) is compared against two MI techniques: multivariate normal imputation (MVNI) and two-fold fully conditional specification (two-fold FCS). The missing data techniques are explained in detail in the next section.

Based on the results of the simulation study, the technique that performs best will be applied to the original data set for the longitudinal analysis reported in Chapter 4.

3.2 Methods

3.2.1 Description of the data set

For the purpose of making the computation of the simulation study less intensive, a sub-set of covariates where selected from those that will be used in the final analysis of Chapter 4. The data consist of two incomplete dependent variables quality of life and depressive symptoms; two complete independent variables (age and gender) and seven incomplete covariates (CHD, cohabiting status, wealth, depressive symptoms, physical activity, smoking status and alcohol consumption) (Table 3.2). In this chapter cohabiting status, wealth, physical activity, smoking status and alcohol consumption are recoded with three categories (as described in Chapter 2, section 2.2.3) and depressive symptoms is used in the original scale when it is a covariate (see Table 3.2). This choice was made to overcome a problem that arises in Mplus when specifying the variance of dichotomous variables, such as the model may not be identified and the standard errors may not be trustworthy due to a non-positive first order derivative product matrix. The problem is further discussed in section 3.2.4.

In order to make the three techniques as comparable as possible, in the analyses (substantive and imputed) all non-continuous covariates are treated as continuous. As explained in Chapter 2, the sample size is restricted to participants with CHD and to healthy participants (Well group), the total sample size is 4,496 in wave 1 (2002-03); 3,465 in wave 2 (2004-05) and 3,031 in wave 3 (2006-07). Table 3.3 shows the prevalence of missing values due to mortality and unit non-response at each wave of the study. Only 1,998 participants had complete data on all variables at the 3 waves (44.4% of the sample in wave 1). Table 3.4 reports the prevalence of missing values (item non-response) for each of the incomplete variables.

Table 3.2 Prevalence and sample size at each wave of the covariates used in the simulation study

	Wave 1		Wave 2		Wave 3	
	%	**N**	**%**	**N**	**%**	**N**
Cohabiting status						
Cohabiting	71.7	3,221	70.7	2,448	68.8	2,085
Not cohabiting	23.2	1,045	24.7	857	26.3	798
Never married not cohabiting	5.1	229	4.6	160	4.9	148
Total	*100*	*4,495*	*100*	*3,465*	*100*	*3,031*
Wealth						
Richest	16.4	732	16.9	576	15.7	462
Middle	59.8	2,660	60.5	2,057	60.2	1,769
Poorest	23.8	1,057	22.6	768	24.1	709
Total	*100*	*4,449*	*100*	*3,401*	*100*	*3,031*
Smoking status						
Never smoked	36.0	1,596	36.8	1,274	40.2	1,217
Ex-smoker	45.2	2,001	47.2	1,632	45.7	1,385
Current smoker	18.8	832	16.0	553	47.1	426
Total	*100*	*4,429*	*100*	*3,459*	*100*	*3,028*
Physical activity						
Active	16.1	711	24.7	856	23.0	699
Moderate activity	23.0	1,021	16.3	565	16.0	484
Low activity or inactive	60.9	2,699	59.0	2,044	61.0	1,847
Total	*100*	*4,431*	*100*	*3,465*	*100*	*3,030*
Alcohol consumption						
Drinks occasionally	27.7	1,223	23.8	724	24.8	627
Once or twice a week	42.7	1,888	38.5	1,172	37.1	940
≥3 days a week	29.6	1,309	37.7	1,146	38.1	964
Total	*100*	*4,420*	*100*	*3,042*	*100*	*2,531*
Depressive symptoms						
0	46.9	2,047	44.8	1,539	47.8	1,409
1	22.7	989	24.8	854	24.6	725
2	10.3	448	11.1	383	9.7	286
3	6.8	298	6.5	223	6.1	179
4	5.0	216	4.7	160	4.0	117
5	3.5	152	2.9	100	2.9	86
6	2.2	97	2.9	98	1.9	56
7	1.7	76	1.5	51	1.9	55
8	0.9	40	0.9	30	1.1	33
Total	*100*	*4,363*	*100*	*3,438*	*100*	*2,946*

Table 3.3 Unit non-response and deaths

	N	%
Present all waves	2,880	64.1
Missed wave 2	876	19.5
Deaths before wave 2	151	3.4
Missed wave 3	431	9.6
Deaths before wave 3	158	3.5
Total	4,496	100

Table 3.4 Prevalence of missing values for each variable of interest, by wave

	Wave 1		Wave 2		Wave 3	
	N	%	N	%	N	%
CHD						
Complete	4,487	99.8				
Item non-response	9	0.2				
Total	4,496	100				
Quality of life						
Complete	3976	88.4	2,987	86.2	2,581	85.2
Item non-response	520	11.6	478	13.8	450	14.8
Total	4,496	100	3,465	100	3,031	100
Depressive symptoms						
Complete	4,363	97.0	3,438	99.2	2,946	97.2
Item non-response	133	3.0	27	0.8	85	2.8
Total	4,496	100	3,465	100	3,031	100
Cohabiting status						
Complete	4,495	100.0	3,465	100.0	3,031	100.0
Item non-response	1	0.0	0	0.0	0	0.0
Total	4,496	100	3,465	100	3,031	100
Wealth						
Complete	4,449	99.0	3,401	98.2	2,940	97.0
Item non-response	47	1.1	64	1.8	91	3.0
Total	4,496	100	3,465	100	3,031	100
Smoking status						
Complete	4,429	98.5	3,459	99.8	3,028	99.9
Item non-response	67	1.5	6	0.2	3	0.1
Total	4,496	100	3,465	100	3,031	100
Physical activity						
Complete	4,431	98.6	3,465	100.0	3,030	100.0
Item non-response	65	1.5	0	0.0	1	0.0
Total	4,496	100	3,465	100	3,031	100
Alcohol consumption						
Complete	4,420	98.3	3,042	87.8	2,531	83.5
Item non-response	76	1.7	423	12.2	500	16.5
Total	4,496	100	3,465	100	3,031	100

3.2.2 Complete case analysis

The complete data of 1,998 individuals was treated as if it were the underlying population, and the regression coefficients obtained from random intercepts models are the target parameters of interest I wish to recover. Random intercept models were estimated as follows:

$$y_{ij} = \beta_0 + \sum_{p=1}^{3} \beta_p x_{pj} + \sum_{p=1}^{8} \beta_p x_{p_{ij}} + u_j + e_{ij} \qquad (3)$$

Where y_{ij} is the quality of life for individual j at time i, x_{pj} are time-invariant factors such as gender, CHD (at wave 1), and the interaction term between CHD and gender; x_{pij} are time-varying factors such as age (a linear and quadratic term), cohabiting status, depressive symptoms, wealth, smoking status, alcohol consumption and physical activity.

A logit model was estimated for depressive symptoms as follows:

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \sum_{p=1}^{3} \beta_p x_{p\,j} + \sum_{p=1}^{6} \beta_p x_{p_{ij}} + u_j \qquad (4)$$

Where $\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right)$ is the odds that $y_{ij}=1$ (i.e. the probability of having depressive symptoms) at occasion *i* for individual *j*, x_{pj} are time-invariant factors such as gender, CHD (at wave 1), and the interaction term between CHD and gender; x_{pij} are time-varying factors such as age, cohabiting status, wealth, smoking status, alcohol consumption and physical activity. u_j denotes the random error associated with the individual level variation with residual variance equal to σ_u^2. Model (4) does not include a level-one residual because it is an equation for the probability $\frac{\pi_{ij}}{1-\pi_{ij}}$ rather than for the outcome y_{ij} (Goldstein, 2003). The level-one residual variance σ_e^2 cannot change and is conventionally fixed at $\pi^2/3 = 3.29$.

In order to compare trajectories of quality of life and depressive symptoms of men and women, the following random intercept models were estimated:

$$y_{ij} = \beta_0 + \beta_1 x_{1j} + \beta_2 t_{ij} + \beta_3 (t_{ij} * x_{1j}) + \sum_{p=1}^{8} \beta_p x_{p_{ij}} + u_j + e_{ij} \qquad (5)$$

Where y_{ij} is the quality of life for individual j at time i, x_{1j} is gender, t_{ij} denotes time and takes three discrete values denoting the baseline, the second and third waves, t_{ij}* x_{ij} denotes the interaction term between time and gender; x_{pij} are time-varying factors described in equation 3). The model is run separately for people with CHD and people in the Well group because the interaction term between CHD, gender and time was statistically significant, and also to facilitate the interpretation of results.

The same model as in equation 5) was estimated for depressive symptoms as follows:

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 x_{1j} + \beta_2 t_{ij} + \beta_3 (t_{ij} * x_{1j}) + \sum_{p=1}^{6} \beta_p x_{p_{ij}} + u_j \qquad (6)$$

Where x_{1j} is gender, t_{ij} denotes time and takes three discrete values, t_{ij}* x_{1j} denotes the interaction term between time and gender; x_{pij} are time-varying factors described in model 4). The model is run separately for people with CHD and people in the Well group.

All models are estimated in Stata (version 10).

3.2.3 Simulation strategy

To evaluate the performance of the three techniques for handling missing data, an artificial simulation study was set up (Figure 3.1). From the complete data of 1,998 individuals (the reference population for the complete-case analysis), ~55.6% of missing values were generated using random uniform and binomial numbers to reproduce similar probabilities of missingness as occurred in the original data (4,496 individuals). Deletion was repeated 1,000 times to generate 1,000 replicates in which the prevalence of missing values obtained ranged between 54% and 57%. Deletion was performed for each wave as follows:

- For wave 1, the pattern of missing values for the variables of interest was examined, only patterns with over 1% missing values were replicated and these are reported in Table 3.5. In order to create missing values, random uniform

numbers were used (Burton et al., 2006). If the rank of the random number was equal to or less than the proportion specified, a missing value was generated for the variable of interest (Table 3.5). For the remaining proportions of missing values (those patterns with less than 1% coverage reported in italics in Table 3.5) and for the variables not included in the patterns reading across the rows in Table 3.5, cases were deleted using random binomial variables, with the probability of having a missing value being the same as for the original data.

- For wave 2, cases were first deleted to generate missing values due to unit non-response (19.5%) and mortality (3.5%) using random uniform numbers. Then the pattern of missing values for the variables of interest was examined and missing values for item non-response were generated using random uniform number according to the patterns described in Table 3.5. If the rank of the random number was equal to or less than the proportion specified, a missing value was generated for the variable of interest (Table 3.6). For the remaining proportion of missing values (those patterns with less than 1% coverage) and for the remaining variables, missing values were generated using random binomial variables using the same proportion of missing values as for the original data.

- For wave 3, the same procedure as for wave 2 was used to generate missing values (Table 3.7). First missing values were generated to reproduce unit non-response (25.6%) and mortality (7%) using random uniform numbers. Then the pattern of missing values for the variables of interest was examined and missing values for item non-response were generated using random uniform number according to the patterns described in Table 3.7. If the rank of the random number was equal to or less than the proportion specified, a missing value was generated for the variable of interest (Table 3.7). For the remaining proportion of missing values (those patterns with less than 1% coverage) and for the remaining variables, missing values were generated using random binomial variables using the same proportion of missing values as for the original data. The full code used to generate missing values is available in Appendix 3.1.

The deletion procedure was repeated 1,000 times in order to generate 1,000 replicates. Each replication was then analyzed as follows:

1) In Mplus to perform the FIML estimation.
2) In SAS, using the MI procedure to obtain five imputed data sets under MVNI.
3) In Stata, to perform the two-fold FCS to obtain five imputed data sets.

Steps 1 to 3 are described in detail in the following section. Steps 2 and 3 are followed by analysis of the imputed data sets to obtain overall estimate according to Rubin's formula (Rubin, 1987).

The choice of imputing five data sets was made to make the simulation less computationally intensive. Most literature (Rubin, 1987; van Buuren et al.,, 1999) suggests that good inferences can be made with the number of imputed data sets (m) as few as m = 5. Rubin (1987) showed that the efficiency of an estimate based on m imputations, relative to one based on an infinite number, is $(1 + \lambda/m)^{-1}$ where λ is the rate of missing information. In my case, with 56% missing information, m =5 imputations is 90% efficient.

Figure 3.1 Design of simulation study

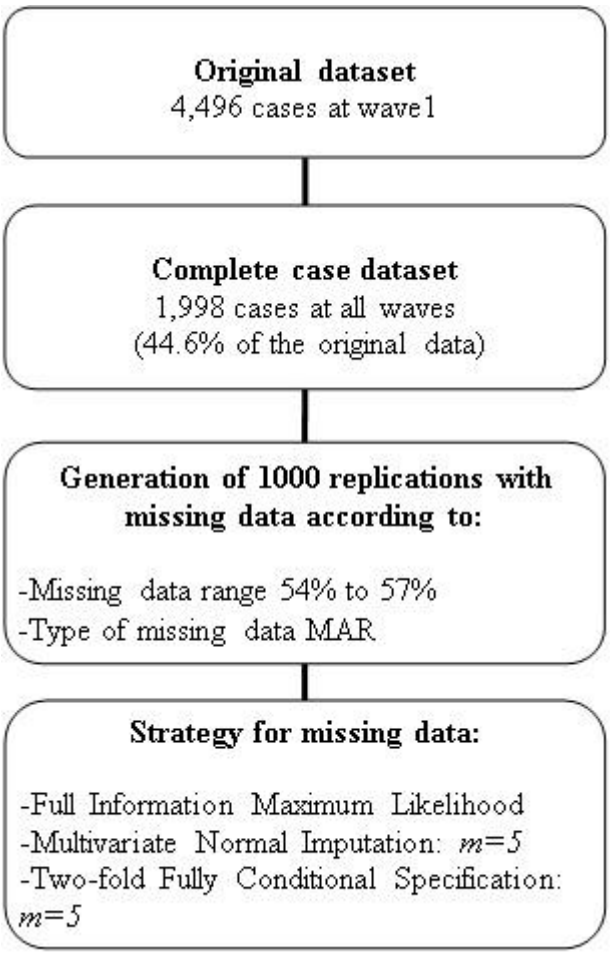


Table 3.5 Missing values patterns, wave 1

Marital status	Wealth	Smoking	Physical Activity	Alcohol consumption	QoL	Depressive symptoms	CHD	Missing values	N	%
+	+	+	+	+	+	+	+	0	3888	86.48
+	+	+	+	+	.	+	+	1	402	8.94
+	+	.	.	.	.	.	+	5	47	1.05
+	.	+	+	+	+	+	+	*1*	*37*	*0.82*
+	+	+	+	+	+	.	+	*1*	*34*	*0.76*
+	+	+	+	+	.	.	+	*2*	*31*	*0.69*
+	+	+	+	.	.	+	+	*2*	*21*	*0.47*
+	+	.	.	+	+	.	+	*3*	*14*	*0.31*
+	.	+	+	+	.	+	+	*2*	*8*	*0.18*
+	+	.	.	.	.	.	.	*6*	*3*	*0.07*
+	+	+	+	.	+	+	+	*1*	*2*	*0.04*
+	+	+	+	.	.	+	.	*3*	*2*	*0.04*
+	+	+	+	+	+	+	.	*1*	*1*	*0.02*
+	+	.	+	+	.	+	+	*2*	*1*	*0.02*
+	+	+	+	+	.	.	.	*3*	*1*	*0.02*
+	.	+	+	+	.	.	+	*3*	*1*	*0.02*
+	+	.	+	+	.	.	.	*4*	*1*	*0.02*
+	+	.	.	+	.	.	+	*4*	*1*	*0.02*
.	.	+	+	.	.	+	.	*5*	*1*	*0.02*

A plus sign indicates that the variable was observed; a dot indicates that the data on the variable were missing.
Missing values patterns in italics were not replicated because missing values <1%.

Table 3.6 Missing values patterns, wave 2

Wealth	Smoking	Alcohol consumption	QoL	Depressive symptoms	Missing values	N	%
+	+	+	+	+	0	2871	82.9
+	+	.	.	+	2	348	10.0
+	+	+	.	+	1	103	3.0
+	+	.	+	+	1	50	1.4
.	+	+	+	+	1	49	1.4
+	+	.	.	.	*3*	*12*	*0.3*
.	+	.	.	+	*3*	*11*	*0.3*
+	+	+	+	.	*1*	*10*	*0.3*
+	.	+	+	+	*1*	*4*	*0.1*
.	+	+	+	.	*2*	*3*	*0.1*
+	+	+	.	.	*2*	*1*	*0.0*
.	+	+	.	+	*2*	*1*	*0.0*
+	.	.	.	+	*3*	*1*	*0.0*
+	.	.	.	.	*4*	*1*	*0.0*

A plus sign indicates that the variable was observed; a dot indicates that the data on the variable were missing. Missing values patterns in italics were not replicated because missing values <1%.

3.2.4 Full information maximum likelihood (FIML)

Table 3.7 Missing values patterns, wave 3

Wealth	Smoking	Physical Activity	Alcohol consumption	QoL	Depressive symptoms	Missing values	N	%
+	+	+	+	+	+	0	2421	79.9
+	+	+	.	.	+	2	311	10.3
+	+	+	.	+	+	1	91	3.0
+	+	+	.	.	.	3	62	2.0
.	+	+	+	+	+	1	58	1.9
+	+	+	+	.	+	1	44	1.5
.	+	+	.	.	+	*3*	*16*	*0.5*
.	+	+	.	.	.	*4*	*15*	*0.5*
+	+	+	+	+	.	*1*	*7*	*0.2*
.	+	+	.	+	+	*2*	*2*	*0.1*
+	.	+	+	+	+	*1*	*1*	*0.0*
+	.	+	.	+	+	*2*	*1*	*0.0*
+	.	+	.	.	+	*3*	*1*	*0.0*
+	+	.	.	.	.	*4*	*1*	*0.0*

A plus sign indicates that the variable was observed; a dot indicates that the data on the variable were missing. Missing values patterns in italics were not replicated because missing values <1%.

The FIML method estimates model parameters and standard errors using all available raw data. It does not involve imputation of missing items but directly estimates the parameters from available items (Enders, 2001). The FIML estimator maximizes a likelihood function that is the sum of *n* casewise likelihood functions (where *n* in the number of respondents). Enders (2001) describes the method as follows: assuming multivariate normality, the casewise likelihood of the observed data is obtained by maximising the function:

$$\log L_i = K_i - \frac{1}{2}\log|\Sigma_i| - \frac{1}{2}(x_i - \mu_i)^T \sum_i^{-1}(x_i - \mu_i) \qquad (7)$$

such that x_i is the vector of complete data for case *i*, μ_i is the vector of mean estimates for those variables that are observed for case *i*, and K_i is a constant that depends on the number of complete points for case *i*. The determinant and inverse of the covariance matrix Σ_i are based only on those variables that are observed for case *i*. Summing over the n casewise functions yields the discrepancy function for the entire sample:

$$\log L(\mu, \Sigma) = \sum_{i=1}^{n} \log L_i \qquad (8)$$

To illustrate of FIML works consider a model with four observed variables: X_1, X_2, X_3 and X_4. The parameters of interest are

$$\mu = [\mu_1, \mu_2, \mu_3, \mu_4] \text{ and } \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{bmatrix}$$

The likelihood value for a subject with missing X_1 would be a function of the values on the observations for the other three variables, X_2, X_3 and X_4, as well as the parameter estimates that involved these three variables. The relevant parameters are shown in the following:

$$\mu = [0, \mu_2, \mu_3, \mu_4] \text{ and } \Sigma = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ 0 & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ 0 & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{bmatrix}$$

By contrast the likelihood value for a subject with missing X_2 and X_4 would be a function of the two other observations (X_1, and X_3) as well as the parameter estimates that involved X_1, and X_3. The relevant parameters are shown in the following:

$$\mu = [\mu_1, 0, \mu_3, 0] \text{ and } \Sigma = \begin{bmatrix} \sigma_{11} & 0 & \sigma_{13} & 0 \\ 0 & 0 & 0 & 0 \\ \sigma_{31} & 0 & \sigma_{33} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Then the value of the overall discrepancy function is obtained by summing the likelihood functions for each individual.

Enders (2001) explains that at a more conceptual level, it is assumed that missing values on a variable X are conditionally dependent on other variables in the data (missing at random (MAR)), and incorporating vectors of partially complete data in the individual level likelihood functions (7) implies probable values for the missing data during the parameter estimation process. Conceptually this is analogous to generating predicted scores for the missing data by regressing X on other variables used in the analysis.

FIML in Mplus

Random intercept models with FIML were performed on each of the 1,000 replicates in Mplus, using a Monte Carlo simulation. Mplus has extensive Monte Carlo simulation facilities for both data generation and data analysis. Since the replicates were generated in Stata, the *external* Monte Carlo simulation study was used, whereby multiple data sets are generated in a first step using another computer program (in my case Stata). The data sets (or replicates) are analysed in a second step in Mplus and the results are summarised (Muthén and Muthén, 2007). In an external Monte Carlo simulation in Mplus it is possible to provide population values (β and SE(β)) for each parameter when specifying the random interaction model. These parameters are used as the population parameter values (i.e. the targeted parameters) for the analysis model (see Mplus coding

in Appendix 3.2). The population parameter values obtained from the complete case analysis are those that I wish to recover. In the logistic random intercept model for depressive symptoms, the variable was specified as being a dummy variable (total score ranged from 0 to 8, cut off 3); while in the linear random intercept model for quality of life the variable measuring depressive symptoms was used as continuous covariate. When using FIML each model is estimated conditioned on the independent variables, therefore cases that are missing on the independent variables will be excluded from the analyses. It is possible to include these cases by treating these variables as dependent variables and distributional assumptions are made about them (i.e. normality). This is achieved by mentioning the variances of the independent variables in the MODEL command. One problem that commonly arises in Mplus when specifying the variance of dichotomous variables is that the model may not be identified and the standard errors may not be trustworthy due to a non-positive first order derivate product matrix. One way to get round this problem is to use independent variables with more than two categories (especially when their variances need to be specified) and treat them as continuous. This is the solution adopted for this simulation study. In some analyses Mplus allows the inclusion of auxiliary variables that are known to predict missingness, for example ethnicity and region of residence. However, when using a multilevel model in Mplus, this option is not available.

In the Monte Carlo simulation setting, random intercept models for continuous outcomes are estimated using maximum likelihood; for binary outcomes random intercept models are estimated using a Monte Carlo algorithm for integration.

3.2.5 Multivariate Normal imputation (MVNI)

Multiple imputations under the normal model assume a joint multivariate normal distribution for all variables. With the multivariate normal model missing data are imputed using simultaneous linear regression models in which each variable potentially depends on all other variables (Schafer and Olsen, 1998). Various methods can be used to fit and make Bayesian draws from the joint distribution. The method of choice depends on the type of missing data pattern, i.e. monotone or arbitrary. A data set is said to have a monotone missing pattern when a variable Y_j is missing for the individual *i* implies that all subsequent variables Y_k, $k>j$, are also missing for the individual *i*. For data sets with arbitrary or non-monotone missing patterns, a Markov Chain Monte Carlo

(MCMC) method (Schafer 1997) can be used. A Markov chain is a sequence of random variables in which the distribution of each element depends only on the value of the previous one. MCMC creates multiple imputations by drawing simulations from a Bayesian predictive distribution for normal data. A regression model is fitted for each variable with missing values, with other variables as covariates. Based on the fitted regression coefficients, a new regression model is simulated from the posterior predictive distribution of the parameters and is used to impute the missing values for each variable (Rubin 1987). The process is repeated sequentially for variables with missing values.

MVNI in SAS

For each of the 1,000 replicates, 5 imputed data sets were generated using PROC MI available in SAS (SAS OnlineDoc™: Version 8; Vargas-Chanes et al., 2003). The missing values were imputed using a Markov Chain Monte Carlo (MCMC) method which is suitable for arbitrary missing data patterns, and which assumes multivariate normality. In MCMC, one constructs a Markov Chain long enough for the distribution of elements to stabilize to a common, stationary distribution. By repeatedly simulating steps of the chain, it simulates draws from the distribution of interest.

In Bayesian inference, information about unknown parameters is expressed in the form of a posterior distribution. MCMC has been applied as a method for exploring posterior distributions in Bayesian inference. That is, through MCMC, one can simulate the entire joint distribution of the unknown quantities and obtain simulation-based estimates of posterior parameters that are of interest. Assuming that the data are from a multivariate normal distribution, data augmentation is applied to Bayesian inference with missing data by repeating a series of imputation and posterior steps. In the Imputation (I) step the missing data are imputed by drawing values from the conditional distribution, given the observed values and the parameters; in the Posterior (P) step new values for the parameters are imputed by drawing them from a Bayesian posterior distribution given the observed data and the most recent estimates (from the I step) for the missing data (Vargas-Chanes et al., 2003). These two steps are iterated long enough for the results to be reliable for a multiply imputed data set (Schafer 1997).

By default, the SAS procedure uses the MCMC method with a single chain to create five imputations. I have specified multiple chains meaning that a separate chain is used for each imputation (data set), because using multiple chains may be computationally more efficient than a single long chain. The posterior mode, the highest observed-data posterior density, with a non-informative prior, is computed from the Expectation Maximization (EM) algorithm and is used as the starting value for the chain. The EM algorithm starts with randomly assigning values to all the parameters to be estimated. It then iterately alternates between two steps, called the expectation step (E-step) where it computes the expected likelihood for the complete data, and the maximization step (M-step) where it re-estimates all the parameters by maximizing the likelihood function for the complete data (Little and Rubin, 2002). The MI procedure takes 200 burn-in iterations before the first imputation and 100 iterations between imputations. In a Markov chain, the information in the current iteration has influence on the state of the next iteration. The burn-in iterations are iterations at the beginning of each chain that are used to eliminate the series of dependence on the starting value of the chain and to achieve a stationary distribution.

In order to monitor the convergence in MCMC to assess whether the number of iterations is enough to achieve convergence, I looked at the time-series and autocorrelation function plots for means of the independent variables. For quality of life and depressive symptoms at each wave, I requested the time-plot of the mean against the iterations, and the autocorrelations (with 95% confidence limits) for the means at various lags in the sequence of iterations. The time-series plots showed that for both variables the series of iterations had converged, as each resembled a horizontal band without long upward or downward trends. Similarly, the autocorrelation plots showed no significant negative or positive correlations.

The imputation model included the same variables as the substantive models (including the interaction term between CHD and gender). Categorical and binary variables were imputed under the normal model and imputed values were rounded to the nearest category. The variable for depressive symptoms was imputed under the normal model, which was then transformed into a dummy variable for use as an outcome variable. Although the regression and MCMC methods assume multivariate normality, inferences based on multiple imputation can be robust to departures from the multivariate normality assumption if the amount of missing information is not large. It often makes

sense to use a normal model to create multiple imputations even when the observed data are somewhat non-normal, as supported by simulation studies described in Schafer (1997) and the original references therein. The imputation of the quality of life measure (CASP-19) and of the depressive symptoms measure (CESD-8) were performed at the level of each summed index and not for the individual items that constitute the two measures.

Imputation of the 1,000 replicates was performed in blocks of 100 (SAS coding for imputation is available in Appendix 3.3). Imputed data sets were then saved and transferred to Stata for the estimation of the random intercept models using Rubin's rules. Because of system limitations, analysis of each random intercept model was also run in blocks of 100 imputed replicates and the estimates stored.

3.2.6 Fully conditional specification (FCS) and two-fold FCS

Van Buuren et al., (2007:1051-1052) describe the fully conditional specification as follows: suppose $Y=(Y_1, Y_2,\ldots,Y_p)$ is a vector of p random variables (explanatory and or dependent) with p-variate distribution $P(Y|\theta)$. We assume that the joint distribution of Y is completely specified by θ, a vector of unknown parameters. For example if Y is multivariate normally distributed, $\theta=(\mu,\Sigma)$, with μ a p-dimensional mean vector and Σ a p x p covariance matrix. Let the matrix $y=(y_1, \ldots y_n)$ with $y_i=(y_{i1}, y_{i2}\ldots, y_{ip})$, i=1,…,$n$ be an independently and identically distributed (i.i.d.) sample of the vector Y. The matrix y is partially observed, in the sense that each column in y has missing data.

The standard procedure for creating multiple imputations y^* of y^{mis} is as follows:

1. Calculate the posterior distribution $p(\theta|y^{obs})$ of θ based in the observed data y^{obs};
2. Draw a value of θ^* from $p(\theta|y^{obs})$;
3. Draw a value of y^* from the conditional posterior distribution of y^{mis} given $\theta=\theta^*$, $p(y^{mis}|y^{obs}, \theta=\theta^*)$.

Steps 2 and 3 are repeated several times for more imputations. FCS proposes to obtain the posterior distribution of θ by sampling iteratively from conditional distribution of the form:

$P(Y_1| Y_{-1}, \theta_1)$,

… (9)

$P(Y_p | Y_{-p}, \theta_p)$.
Y_{-p} *is* defined as $Y_1 \ldots . Y_n$ excluding Y_p.

The parameters $\theta_1, \ldots, \theta_p$ are treated as specific to the respective conditional densities and are not necessarily the product of some factorization of the *true* joint distribution $P(Y|\theta)$. The process is iterative, starting with some simple initial values, and cycles through all variables, with possibly different conditional specifications, a number of times. More precisely, the t^{th} iteration of the method consists of the following successive draws of the Gibbs sampler:

$$
\begin{aligned}
&\theta_1^{*(t)} \sim P(\theta_1 | y_1^{obs}, y_2^{(t-1)}, \ldots, y_p^{(t-1)}) \\
&y_1^{*(t)} \sim P(y_1^{mis} | y_1^{obs}, y_2^{(t-1)}, \ldots, y_p^{(t-1)}, \theta_1^{*(t)}) \\
&\ldots \\
&\theta_p^{*(t)} \sim P(\theta_p | y_p^{obs}, y_1^{(t)}, y_2^{(t)}, \ldots, y_{p-1}^{(t)}) \\
&y_p^{*(t)} \sim P(y_p^{mis} | y_p^{obs}, y_1^{(t)}, y_2^{(t)}, \ldots, y_{p-1}^{(t)}, \theta_1^{*(t)})
\end{aligned}
\qquad (10)
$$

"No information about y_j^{mis} is used to draw $\theta_j^{*(t)}$ which differs from MCMC approaches to joint modelling. The iterations of (10) are executed *m* times in parallel to generate *m* multiple imputations. This procedure simply assumes that the joint distribution is specified by (9), and that the Gibbs sampler in (10) provides draws from it" (van Buuren et al., 2007:1051-1052).

The great advantage of this approach is that each type of variable (continuous, binary, unordered and ordered categorical) is modelled separately (Molenberghs and Kenward, 2007). However, from a theoretical standpoint of view this technique is problematic, because the sequence of regression models might not be consistent with a true joint distribution (Shafer and Graham, 2002), meaning that the iterative algorithm may never converge because the joint distribution to which they may converge does not exist. Nevertheless, simulation work (Brand, 1999) suggests that in some practical applications the method can work well despite the theoretical problems.

The FCS approach differs from the MVNI in that it does not start with the construction of a well-defined joint distribution for the variables to be imputed. FCS starts with a collection of univariate conditional distributions for variables with missing data in terms of all other variables. The main idea is that a univariate conditional model is constructed for each potentially missing variable (dependent and/or explanatory) which is

appropriate to the type. This means that logistic regression can be used for binary variables, linear regression for continuous, ordinal logistic regression for categorical variables and so forth. The other potentially missing variables are used as explanatory variables in each univariate imputation model. The conditional density for the j^{th} missing variable (of p) would be

$$f(Y_j \mid Y_1,...,Y_{j-1},Y_{j+1},...,Y_p) \quad \text{j=1....p} \qquad (11)$$

Univariate posterior draws are made one variable at a time by cycling through all p models given current values of the other variables (Molenberghs and Kenward, 2007). After sufficient cycles (10-20), the imputations are taken from one final cycle through the univariate model.

This approach can be extended to q repeated waves, in which case equation (11) becomes

$$f(Y_{ij} \mid Y_{i1},...,Y_{i(j-1)},Y_{i(k+1)},...,Y_{ip}) \quad \text{j=1....p} \qquad \text{i=1...q} \qquad (12)$$

Nevalainen et al., (2009) proposed using an imputation strategy which is doubly iterative, the so called two-fold fully conditional specification. At time i, Y_i is imputed conditional on the same variable observed at time $i-1$ and $i+1$, and the other variables at time i. One iteration runs over the variables j=1....p, called within-times iterations. The past and future observations (Y_{i-1} *and* Y_{i+1}) are not imputed at this stage, they serve only in the role of predictors in the imputation model. There is also second imputation iteration over waves (i=1...q), called among-times iterations.

Two-fold FCS in Stata

For each of the 1,000 replicates, five imputed data sets were generated using the Stata's user-written program *ice*. The acronym, *ice*, stands for Imputation by Chained Equations (Royston, 2005, 2007, 2009; Carlin et al., 2008; Royston et al., 2009). The two-fold FCS is an extension of the FCS method of *ice*, programming was required to implement the doubly iterative procedure. The imputation model included the same variables as the substantive models (including the interaction term between CHD and gender). For each variable with missing data a univariate conditional model was constructed which was appropriate to the type. The default option for a variable is logistic regression when there are two distinct values, multinomial logistic regression

when there are three to five categories, and linear regression otherwise. It is possible to define the regression command to be used with a specific variable when this is to be imputed. Variables with three to five categories are modelled as linear terms when they are covariates for other variables to be imputed. This choice was made to make the treatment of these variables in the same way as FIML and MVNI. For depressive symptoms, an ordered logistic regression was chosen to be used when it was a dependent variable. For all other variables with three categories, multinomial logistic regression was chosen when these were dependent variables. However, use of multinomial logistic regression may produce unstable estimates or perfect prediction may arise when a predictor variable perfectly predicts success or failure in the outcome variable. In the former situation it is not possible to use multinomial logistic regression, so instead the ordered logistic regression is used, that was the case for cohabiting status, physical activity and smoking status. If perfect prediction arises ice temporarily augments the data with a few extra observations with low weight, in such a way as to remove the perfect prediction.

The two-fold FCS was performed as follows:

1) Variables with missing data at wave 1 were imputed using as predictors all other variable at the same wave, plus the future observation (at wave 2) of the same variable. Although this latter variable is imputed by default in Stata, it is then dropped as it serves only the role of a predictor in the imputation model. For example to impute missing values for quality of life (QoL) at wave 1, the following linear regression model is used:

$$QoL_{wave1} = Sex_{wave1} + CHD_{wave1} + (Sex * CHD)_{wave1} + Age_{wave1} + Age^2_{wave1} + Cohabiting_{wave1} + Wealth_{wave1} + \\ + Smoking_{wave1} + Physicalact_{wave1} + Alcohol_{wave1} + Depression_{wave1} + QoL_{wave2}$$

2) Variables with missing data at wave 2 were imputed using as predictors all other variable at the same wave, plus the past (wave 1 imputed in the previous step) and future observations of the same variable (wave 3 not imputed). These variables from wave 1 and wave 3 are then dropped after the imputation as they serve only the role of predictors in the imputation model.

3) Variables with missing data at wave 3 were imputed using as predictors all other variable at the same wave, plus the past observations (wave 2 only imputed in

the previous step) of each variable to be imputed. Again, the wave 2 variable are then dropped after the imputation as serves only the role of a predictor in the imputation model. One imputed data set is generated.

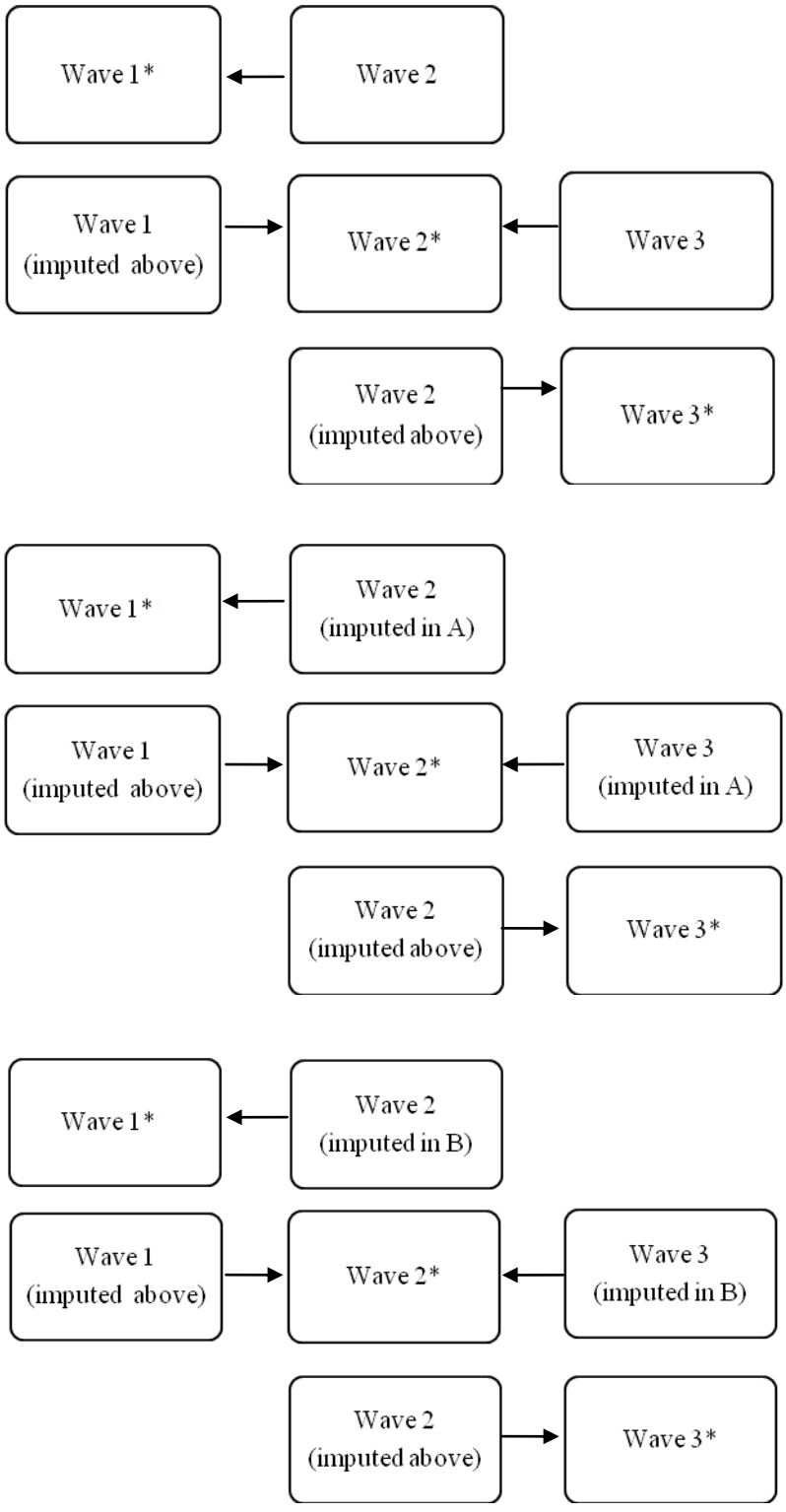
Figure 3.2 gives a graphical explanation of steps 1) 2) and 3).
Within step 1) 2) and 3) it is possible to decide how many within-times iterations are needed to reach convergence. The default option in ice is 10 iterations. Convergence was explored on one of the replicates (simulated data set), by plotting the mean values of the outcomes variables (at each wave) against 100 iterations. From graphical inspection it could be concluded that the pattern of the imputed means of depressive symptoms and quality of life occurred randomly; also the mean estimates obtained from 10 iterations and those obtained from 100 iterations did not differ significantly (Table A3.1, Appendix 3.4). Therefore for all replicates the number of within-times iterations was set to 10.

Steps 1) to 3) form one among-times iteration. It can be decided how many among-times iterations are needed, Nevalainen et al., (2009) showed that increasing the number of iterations from one to five improved the performance of the estimators although the gain due to the increase was relatively small. To establish the number of among-times iterations to be used in this study I used three among-times iterations on one replicate and imputed 5 data sets. The means of the imputed variables at each wave were then compared with the means of the complete case data (Table A3.2 Appendix 5.4). Since the estimates from the five imputed data sets were very close to those of the complete case data (consisting of 1,998 cases), three among-times iterations were used to impute all 1,000 replicates. The imputation of the quality of life measure (CASP-19) and of the depressive symptoms measure (CESD-8) were performed at the level of each summed index and not for the individual items that constitute the two measures.

To summarise, steps 1) to 3) were repeated three times (three among-times iterations) to generate one imputed data set, the procedure was then repeated four more times to obtain five imputed data sets on one replicate. In order to generate five imputed data sets for all 1,000 replicates, loops were used, and replicates were imputed in blocks of 100 at each time. The imputation stage can be computationally intensive especially if repeated on 1,000 replicates. Stata code for two-fold FCS is available in Appendix 3.4.3.

Random intercept model estimates were obtained from the imputed replicates. Because of system limitations, analysis of each random intercept model was also run in blocks of 100 imputed replicates and the estimates stored.

Figure 3.2 Two-fold fully conditional specification



A) First among-times iteration

B) Second among-times iteration

C) Third among-times iteration

The * indicates the variable with missing data at the specific wave that is to be imputed.

3.2.7 Evaluation criteria

After the missing data strategies had been performed on the 1,000 replicates, the estimates from the analysis stage were stored. From these stored estimates, some summary measures were calculated to assess each strategy to handle missing data as follows:

The (average) estimate of interest: $\bar{\hat{\beta}} = \frac{\sum_{i=1}^{n} \hat{\beta}_i}{n}$ where n is the number of replicates (1,000), and $\hat{\beta}_i$ is the estimate of interest within each of the i=1,…,n replicates. When MI is performed, each $\hat{\beta}_i$ is the overall estimate obtained according to Rubin's formula (Rubin, 1987), which is just the average of the $\hat{\beta}_i$, from the 5 combined estimates within each of the i=1,…,n replicates.

The (average) standard error of the estimate of interest: $SE(\hat{\beta}) = \frac{\sum_{i=1}^{n} SE(\hat{\beta}_i)}{n}$ where $SE(\hat{\beta}_i)$ is the standard error of the estimate of interest within each of the i=1,…,n replications. When MI is performed, each $SE(\hat{\beta}_i)$ is the overall standard error of the estimate of interest obtained from the five combined estimates according to Rubin's formula (Rubin, 1987), within each of the i=1,…,n replicates.

In order to evaluate the performance of each procedure employed to deal with missing data and to evaluate to what extent the targeted coefficients (estimates of interest) are recovered, I used assessments of accuracy and precision. Accuracy indicates the degree of closeness of the estimated value to the targeted parameter; precision refers to the repeatability or reproducibility of the measurement.

For the assessment of accuracy the following were used:

Bias: $(\bar{\hat{\beta}} - \beta)$ which is the difference between the average estimate and the population parameter. A bias that varies between $\frac{1}{2} SE(\hat{\beta}_i)$ to $2SE(\hat{\beta}_i)$ is considered troublesome (Shafer and Graham, 2002; Sinharay et al., 2001).

The Mean Square Error (MSE): $(\bar{\hat{\beta}} - \beta)^2 + (SD(\hat{\beta}))^2$ is the average squared difference between the estimate and its target plus its variance, therefore can be seen as a summary of both bias and variability. A value of the MSE close to zero indicates that the average estimator predicts the targeted parameter with good accuracy.

For the assessment of precision the following were used:

Standardised bias percent: $100 \times \frac{(\bar{\hat{\beta}} - \beta)}{SE(\hat{\beta})}$ which is the bias as a percentage of the standard error. A standardised bias is considered to have a large impact on the precision if it exceeds 40 per cent in either direction (Collins et al., 2001).

The (average) standard deviation of the estimate of interest: $SD(\hat{\beta}) = \frac{\sum_{i=1}^{n} SD(\hat{\beta}_i)}{n}$ where $SD(\hat{\beta}_i)$ is the standard error of the estimate of interest within each of the i=1,…,n replicates. When MI is performed, each $SD(\hat{\beta}_i)$ is the overall standard error of the estimate of interest obtained from the five combined estimates within each of the i=1,…,n replicates. The standard deviation of estimates represents the variability across replicates of the parameter estimates.

3.3 Results

3.3.1 Quality of life

The results of the comparisons of missing data techniques for the analysis of the first model of quality of life (linear random intercept model with an interaction between gender and CHD) are shown in Tables 3.8 to 3.10. The bias of the coefficients for CHD

and gender were small and close to zero for all the techniques, although MVNI seemed to recover the parameter for gender slightly better than FIML and two-fold FCS; the standard error of the estimate for gender (target value, 0.27) was fully recovered by two-fold FCS. Largest bias in the coefficient for the interaction term between CHD and gender was obtained under FIML and smallest bias was obtained under MVNI and FCS

Table 3.8 FIML technique for QoL, linear random intercept model with gender and CHD interaction (model 1)

FIML	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand Bias*	*MSE*
CHD	-1.48	0.44	-1.54	0.48	0.18	-0.06	-12.5	0.04
Gender	1.09	0.27	1.14	0.29	0.11	0.05	18.6	0.02
CHD*Gender	-0.38	0.68	-0.54	0.72	0.28	-0.16	-22.0	0.10
Age	-0.06	0.01	-0.03	0.02	0.01	0.03	183.4	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00
Cohabiting status	-0.29	0.20	-0.40	0.22	0.11	-0.10	-45.9	0.02
Wealth	1.94	0.17	1.95	0.19	0.10	0.00	2.6	0.01
Smoking	-0.31	0.17	-0.29	0.19	0.07	0.02	10.3	0.01
Physical Activity	-0.55	0.10	-0.58	0.12	0.06	-0.03	-22.5	0.00
Alcohol consumption	0.61	0.13	0.64	0.15	0.08	0.03	23.1	0.01
Depressive symptoms	-1.30	0.05	-1.38	0.06	0.04	-0.08	-121.0	0.01

Abbreviations: QoL= quality of life. FIML=full information maximum likelihood

although the values of the MSE were similar; this was due to the larger values of the standard deviation under MVNI and FCS. MI techniques recovered the coefficient for the interaction term better than the FIML. This is probably due to the fact that the imputation models of FCS, two-fold FCS and MVNI included the interaction term and in that sense they reflected the substantive models.

Table 3.9 MVNI technique for QoL, linear random intercept model with gender and CHD interaction (model 1)

MVNI	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
CHD	-1.48	0.44	-1.42	0.49	0.21	0.06	12.7	0.05
Gender	1.09	0.27	1.08	0.29	0.13	-0.01	-2.2	0.02
CHD*Gender	-0.38	0.68	-0.40	0.73	0.32	-0.02	-3.3	0.10
Age	-0.06	0.01	-0.06	0.02	0.01	0.01	35.2	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00
Cohabiting status	-0.29	0.20	-0.39	0.22	0.12	-0.10	-43.9	0.02
Wealth	1.94	0.17	1.81	0.19	0.11	-0.13	-68.8	0.03
Smoking	-0.31	0.17	-0.32	0.19	0.08	-0.01	-5.8	0.01
Physical Activity	-0.55	0.10	-0.49	0.12	0.06	0.06	45.9	0.01
Alcohol consumption	0.61	0.13	0.60	0.15	0.09	-0.01	-6.3	0.01
Depressive symptoms	-1.30	0.05	-1.32	0.07	0.05	-0.02	-26.1	0.00

Abbreviations: QoL=quality of life. MVNI=multivariate normal imputation

The FIML and two-fold FCS methods did not recover the coefficient for depressive symptoms as well as MVNI, as indicated by the values of the bias (-0.08 in FIML and -0.05 in two-fold FCS) and mainly by the large values of the standardised bias percent (over 40% in both techniques). All three methods failed to recover the coefficient for cohabiting status and the MVNI and two-fold FCS did not perform as well as the FIML in recovering the coefficient for wealth and physical activity.

Table 3.10 Two-fold FCS technique for QoL, linear random intercept model with gender and CHD interaction (model 1)

Two-fold FCS	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
CHD	-1.48	0.44	-1.47	0.45	0.26	0.02	3.6	0.07
Gender	1.09	0.27	1.13	0.27	0.16	0.04	15.8	0.03
CHD*Gender	-0.38	0.68	-0.42	0.69	0.41	-0.04	-6.0	0.17
Age	-0.06	0.01	-0.05	0.02	0.01	0.01	89.9	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00
Cohabiting status	-0.29	0.20	-0.39	0.21	0.14	-0.09	-46.0	0.03
Wealth	1.94	0.17	1.78	0.17	0.13	-0.16	-91.0	0.04
Smoking	-0.31	0.17	-0.29	0.18	0.10	0.02	9.0	0.01
Physical Activity	-0.55	0.10	-0.60	0.11	0.08	-0.05	-50.0	0.01
Alcohol consumption	0.61	0.13	0.60	0.14	0.11	-0.01	-6.5	0.01
Depressive symptoms	-1.30	0.05	-1.35	0.06	0.06	-0.05	-84.6	0.01

Abbreviations: QoL is quality of life. FCS=fully conditional specification

Results from the second model of quality of life (linear random intercept model with an interaction between gender and time) for the CHD group are shown in Table 3.11 to Table 3.13. The bias of the coefficient for gender was reasonably small for all three methods, and the values of the standardised bias percent were within the acceptable range. The bias and the standardised bias percent of the coefficients for wave 2 and wave 3 were slightly larger in MVNI and two-fold FCS compared to those obtained by the FIML method, but still fairly small. The recovery of the coefficients for the two interaction terms was slightly better under FIML and MVNI compared to two-fold FCS for which the standardised bias percentages were also highest. Also, in all three techniques the values of the MSE of the two interaction terms were not close to zero.

As for the recovery of the parameters for the covariates, FIML and two-fold FCS did not recover the coefficient for depressive symptoms with good accuracy and precision,

as shown by the bias (-0.14 for both techniques) and the large values of the standardised bias percent (-87.0% in FIML and -98.8% in two-fold FCS) MVNI recovered the coefficient for depressive symptoms with good accuracy and precision (bias -0.04, standardised percent bias -27.3%). The MVNI and two-fold FCS methods failed to recover the targeted parameter for wealth, the bias was -0.39 for MVNI and -0.41 for two-fold FCS and the corresponding values of standardised bias were -83.4% and -92.9% respectively, also the values of the MSE (0.20 and 0.25 respectively) suggested together with the bias values lack of accuracy in recovering the targeted parameters. All three techniques showed large bias values for cohabiting status, although the values of the standardised bias were below 40%.

Table 3.11 FIML technique for QoL, linear random intercept model with gender and time interaction, CHD group (model 2)

FIML	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.28	0.80	0.35	0.83	0.29	0.07	8.4	0.09
Wave 2	-0.43	0.53	-0.44	0.63	0.33	-0.01	-1.6	0.11
Wave 3	-2.65	0.55	-2.67	0.67	0.38	-0.02	-2.9	0.15
Wave 2*Gender	0.90	0.82	0.96	0.98	0.57	0.06	5.9	0.33
Wave 3*Gender	1.11	0.82	1.16	1.03	0.60	0.04	4.3	0.37
Age	0.05	0.04	0.04	0.05	0.02	-0.01	-11.7	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00
Cohabiting status	-0.32	0.47	-0.48	0.52	0.26	-0.16	-31.2	0.09
Wealth	2.53	0.42	2.40	0.49	0.25	-0.13	-27.6	0.08
Smoking	-0.02	0.47	-0.02	0.51	0.21	0.00	0.9	0.04
Physical Activity	-0.79	0.29	-0.83	0.33	0.17	-0.04	-12.3	0.03
Alcohol consumption	1.05	0.32	1.18	0.37	0.21	0.13	35.0	0.06
Depressive symptoms	-1.19	0.13	-1.33	0.16	0.10	-0.14	-87.0	0.03

Abbreviations: QoL is quality of life. FIML=full information maximum likelihood

Table 3.12 MVNI technique for QoL, linear random intercept model with gender and time interaction, CHD group (model 2)

MVNI	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.28	0.80	0.23	0.84	0.29	-0.05	-6.2	0.09
Wave 2	-0.43	0.53	-0.52	0.62	0.34	-0.09	-14.5	0.12
Wave 3	-2.65	0.55	-2.70	0.66	0.40	-0.05	-7.0	0.16
Wave 2*Gender	0.90	0.82	0.97	0.96	0.56	0.07	7.4	0.32
Wave 3*Gender	1.11	0.82	1.13	1.01	0.60	0.02	1.7	0.36
Age	0.05	0.04	0.05	0.05	0.02	0.00	7.0	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00
Cohabiting status	-0.32	0.47	-0.45	0.52	0.24	-0.13	-25.0	0.08
Wealth	2.53	0.42	2.14	0.47	0.21	-0.39	-83.4	0.20
Smoking	-0.02	0.47	-0.04	0.51	0.19	-0.02	-3.9	0.04
Physical Activity	-0.79	0.29	-0.72	0.32	0.16	0.07	22.2	0.03
Alcohol consumption	1.05	0.32	1.07	0.36	0.19	0.02	5.6	0.04
Depressive symptoms	-1.19	0.13	-1.23	0.15	0.09	-0.04	-27.3	0.01

Abbreviations: QoL is quality of life. MVNI=multivariate normal imputation

Table 3.13 Two-fold FCS technique for QoL, linear random intercept model with gender and time interaction, CHD group (model 2)

Two-fold FCS	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.28	0.80	0.23	0.82	0.37	-0.06	-6.7	0.14
Wave 2	-0.43	0.53	-0.39	0.57	0.42	0.04	7.7	0.18
Wave 3	-2.65	0.55	-2.59	0.60	0.49	0.06	10.6	0.25
Wave 2*Gender	0.90	0.82	1.06	0.90	0.69	0.16	17.9	0.51
Wave 3*Gender	1.11	0.82	1.24	0.91	0.77	0.13	13.9	0.61
Age	0.05	0.04	0.05	0.05	0.02	0.00	-1.1	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00
Cohabiting status	-0.32	0.47	-0.47	0.48	0.31	-0.15	-30.9	0.12
Wealth	2.53	0.42	2.12	0.44	0.29	-0.41	-92.9	0.25
Smoking	-0.02	0.47	-0.04	0.48	0.25	-0.02	-4.1	0.06
Physical Activity	-0.79	0.29	-0.76	0.30	0.18	0.03	9.4	0.03
Alcohol consumption	1.05	0.32	1.07	0.34	0.24	0.02	6.9	0.06
Depressive symptoms	-1.19	0.13	-1.33	0.14	0.11	-0.14	-98.8	0.03

Abbreviations: QoL is quality of life. FCS=fully conditional specification

Results of the comparisons of missing data techniques of the quality of life model for the Well group are shown in Tables 3.14 to 3.16. The FIML method recovered well the main targeted parameters (gender, wave 2, wave 3 and the two interaction terms) as shown by the small bias values and values of standardised bias well below 40%; also the values of the MSE were fairly close to zero, indicating good accuracy. MVNI did not recover with good accuracy and precision the targeted parameters for wave 2 and wave 3, the values of the bias and standardised bias were large; in contrast the method recovered the coefficients for gender and for the two interaction terms quite well.

Table 3.14 FIML technique for QoL, linear random intercept model with gender and time interaction, Well group (model 2)

FIML	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	1.02	0.33	1.05	0.34	0.11	0.03	9.0	0.01
Wave 2	-0.83	0.25	-0.87	0.29	0.16	-0.03	-11.2	0.03
Wave 3	-2.33	0.26	-2.42	0.31	0.18	-0.09	-27.6	0.04
Wave 2*Gender	0.19	0.33	0.21	0.40	0.22	0.02	4.4	0.05
Wave 3*Gender	0.00	0.33	0.00	0.41	0.25	0.00	NA	0.06
Age	0.03	0.02	0.04	0.02	0.01	0.02	75.7	0.00
Age^2	0.00	0.00	-0.01	0.00	0.00	0.00	NA	0.00
Cohabiting status	-0.34	0.22	-0.41	0.24	0.12	-0.07	-27.1	0.02
Wealth	1.78	0.18	1.78	0.20	0.11	0.00	0.1	0.01
Smoking	-0.39	0.18	-0.37	0.20	0.08	0.02	10.3	0.01
Physical Activity	-0.47	0.11	-0.49	0.13	0.06	-0.02	-15.7	0.00
Alcohol consumption	0.62	0.14	0.64	0.16	0.09	0.02	12.5	0.01
Depressive symptoms	-1.32	0.06	-1.38	0.07	0.05	-0.06	-79.1	0.01

Abbreviations: QoL is quality of life. FIML=full information maximum likelihood

The two-fold FCS method performed reasonably well, but failed to recover the coefficient for wave 3, the bias was -0.14 and the standardised bias was -48.6%.

In terms of the covariates, the values of standardised bias of depressive symptoms suggested that FIML and MVNI did not achieve a good precision, however, the average standard errors were close to the targeted standard errors and the values of the MSE were also close to zero, indicating good accuracy. The two-fold FCS method recovered the coefficient for depressive symptoms well; the value of the standardised bias for physical activity (-42.6%) suggests that this parameter is not recovered with good precision, however, the bias is not to be considered troublesome and the standard error

is fully recovered. Neither two-fold FCS nor MVNI recover well the targeted parameter for wealth.

Table 3.15 MVNI technique for QoL, linear random intercept model with gender and time interaction, Well group (model 2)

MVNI	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	1.02	0.33	1.00	0.34	0.12	-0.02	-5.4	0.01
Wave 2	-0.83	0.25	-1.00	0.29	0.17	-0.17	-58.1	0.06
Wave 3	-2.33	0.26	-2.54	0.31	0.19	-0.21	-66.7	0.08
Wave 2*Gender	0.19	0.33	0.18	0.39	0.23	-0.02	-4.1	0.05
Wave 3*Gender	0.00	0.33	-0.03	0.41	0.25	-0.02	NA	0.06
Age	0.03	0.02	0.04	0.02	0.01	0.02	90.5	0.00
Age^2	0.00	0.00	-0.01	0.00	0.00	0.00	-114.2	0.00
Cohabiting status	-0.34	0.22	-0.38	0.24	0.13	-0.04	-15.8	0.02
Wealth	1.78	0.18	1.65	0.21	0.12	-0.13	-61.0	0.03
Smoking	-0.39	0.18	-0.38	0.20	0.09	0.01	3.1	0.01
Physical Activity	-0.47	0.11	-0.47	0.13	0.08	0.00	0.2	0.01
Alcohol consumption	0.62	0.14	0.61	0.16	0.09	-0.02	-9.7	0.01
Depressive symptoms	-1.32	0.06	-1.28	0.07	0.05	0.04	61.3	0.00

Abbreviations: QoL=quality of life. MVNI=multivariate normal imputation

Table 3.16 Two-fold FCS technique for QoL, linear random intercept model with gender and time interaction, Well group (model 2)

Two-fold FCS	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	1.02	0.33	1.01	0.33	0.14	-0.01	-3.2	0.02
Wave 2	-0.83	0.25	-0.90	0.27	0.20	-0.07	-24.6	0.05
Wave 3	-2.33	0.26	-2.47	0.28	0.24	-0.14	-48.6	0.08
Wave 2*Gender	0.19	0.33	0.24	0.36	0.28	0.05	13.8	0.08
Wave 3*Gender	0.00	0.33	0.05	0.37	0.33	0.06	NA	0.11
Age	0.03	0.02	0.04	0.02	0.01	0.02	96.2	0.00
Age^2	0.00	0.00	-0.01	0.00	0.00	0.00	-119.5	0.00
Cohabiting status	-0.34	0.22	-0.39	0.22	0.15	-0.05	-21.3	0.03
Wealth	1.78	0.18	1.65	0.19	0.14	-0.13	-70.4	0.04
Smoking	-0.39	0.18	-0.37	0.19	0.11	0.02	10.9	0.01
Physical Activity	-0.47	0.11	-0.52	0.11	0.08	-0.05	-42.6	0.01
Alcohol consumption	0.62	0.14	0.62	0.15	0.11	0.00	0.6	0.01
Depressive symptoms	-1.32	0.06	-1.34	0.06	0.06	-0.02	-29.0	0.00

Abbreviations: QoL is quality of life. FCS=fully conditional specification

3.3.2 Depressive symptoms

Before proceeding with the results of the depressive symptoms logistic random intercept models it must be noted that parameter estimates obtained using FIML are compared with the complete case analysis run in Stata. Random intercept models for binary outcomes in Mplus are estimated using a Monte Carlo algorithm for integration while Stata by default uses maximum likelihood estimation with adaptive Gaussian quadrature to approximate the integrals (with 7 integration points). The two technique produced

results that were very close if not the same, therefore for the comparisons with FIML estimates, the targeted parameters reported in the tables are those obtained in Stata. However, to show that the ability of FIML to recover targeted parameters did not depend on the estimation procedure for the random intercept models, results of FIML with the targeted parameters obtained using a Monte Carlo algorithm for integration are presented in Appendix 3.5, but will not be discussed below as the conclusions are exactly the same.

It must also be noted that the treatment of the dependent variable differed in each of the missing data technique as follows: with FIML, depressive symptoms was treated as binary, since there was not an imputation model but only a substantive model; with two-fold FCS depressive symptoms was imputed in its original scale using ordered logistic regression in the imputation model, then recoded into a binary variable for the substantive model when it was used as an outcome; lastly, with MVNI depressive symptoms was treated as continuous (normal) in the imputation model and then recoded into a binary variable when used as an outcome in the substantive model.

Tables 3.17 to 3.19 report the results of comparisons of missing data techniques for the analysis of the first model for depressive symptoms (logistic random intercept model with an interaction between gender and CHD). FIML failed to recover the targeted parameters (coefficients and standard errors) for gender, CHD and the interaction term between gender and CHD, as shown by the large biases and large values of the standardised bias percent. Recovery of the parameter estimates of the covariates under FIML was acceptable, with the exception of cohabiting status (Table 3.17). Two-fold FCS recovered the targeted coefficients and standard errors for gender, CHD and the interaction term between gender and CHD to an impressive extent, with good levels of precision and accuracy as shown by the MSE and standardised bias percent. MVNI did not perform as well as two-fold FCS in recovering the coefficients for gender and CHD, while the recovery of the targeted coefficient and standard error for the interaction term between gender and CHD was acceptable. Both the two-fold FCS and the MVNI techniques did not recover the coefficients for cohabiting status and physical activity with good precision (Tables 3.18 and 3.19).

Table 3.17 FIML technique for depressive symptoms, logistic random intercept model with gender and CHD interaction (model 1)

FIML	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
CHD	0.65	0.25	0.44	0.24	0.14	-0.21	-90.1	0.06
Gender	0.74	0.16	0.50	0.14	0.12	-0.24	-176.3	0.07
CHD*Gender	0.03	0.36	0.20	0.34	0.15	0.18	51.7	0.05
Age	-0.02	0.01	-0.02	0.01	0.00	0.00	29.9	0.00
Cohabiting status	0.53	0.11	0.47	0.11	0.05	-0.05	-49.1	0.01
Wealth	-0.31	0.10	-0.33	0.10	0.05	-0.02	-23.0	0.00
Smoking	0.39	0.10	0.36	0.09	0.05	-0.03	-30.8	0.00
Physical Activity	0.43	0.07	0.40	0.07	0.04	-0.03	-38.7	0.00
Alcohol consumption	-0.14	0.08	-0.15	0.08	0.04	-0.01	-17.2	0.00

Abbreviation: FIML=full information maximum likelihood

Table 3.18 MVNI technique for depressive symptoms, logistic random intercept model with gender and CHD interaction (model 1)

MVNI	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
CHD	0.65	0.25	0.56	0.26	0.10	-0.09	-34.7	0.02
Gender	0.74	0.16	0.66	0.16	0.06	-0.08	-48.3	0.01
CHD*Gender	0.03	0.36	0.00	0.37	0.13	-0.02	-6.7	0.02
Age	-0.02	0.01	-0.01	0.01	0.00	0.01	119.8	0.00
Cohabiting status	0.53	0.11	0.45	0.12	0.05	-0.08	-64.9	0.01
Wealth	-0.31	0.10	-0.29	0.11	0.05	0.02	17.1	0.00
Smoking	0.39	0.10	0.36	0.10	0.04	-0.02	-24.7	0.00
Physical Activity	0.43	0.07	0.33	0.08	0.03	-0.10	-126.3	0.01
Alcohol consumption	-0.14	0.08	-0.13	0.09	0.04	0.01	12.0	0.00

Abbreviation: MVNI=multivariate normal imputation

Table 3.19 Two-fold FCS technique for depressive symptom, logistic random intercept model with gender and CHD interaction (model 1)

Two-fold FCS	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
CHD	0.65	0.25	0.64	0.25	0.12	-0.01	-4.5	0.02
Gender	0.74	0.16	0.74	0.16	0.08	-0.01	-3.6	0.01
CHD*Gender	0.03	0.36	-0.01	0.35	0.17	-0.03	-9.9	0.03
Age	-0.02	0.01	-0.02	0.01	0.00	0.00	47.9	0.00
Cohabiting status	0.53	0.11	0.46	0.11	0.06	-0.06	-57.9	0.01
Wealth	-0.31	0.10	-0.27	0.10	0.07	0.04	34.9	0.01
Smoking	0.39	0.10	0.37	0.10	0.04	-0.02	-17.8	0.00
Physical Activity	0.43	0.07	0.39	0.07	0.04	-0.03	-48.2	0.00
Alcohol consumption	-0.14	0.08	-0.13	0.08	0.05	0.01	13.3	0.00

Abbreviation: FCS=fully conditional specification

Results from the second model of depressive symptoms (logistic random intercept model with an interaction between gender and time) for the CHD group are shown in Table 3.20 to Table 3.22. FIML recovered the coefficients for wave 2, wave 3 and the two interaction terms with acceptable accuracy and precision, although the other two technique produced smaller biases.

Table 3.20 FIML technique for depressive symptoms, logistic random intercept model with gender and time interaction, CHD group (model 2)

FIML	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.80	0.44	0.65	0.42	0.39	-0.15	-35.5	0.17
Wave 2	0.01	0.34	-0.06	0.37	0.19	-0.07	-17.8	0.04
Wave 3	-0.41	0.36	-0.50	0.41	0.24	-0.08	-20.3	0.06
Wave 2*Gender	-0.08	0.49	0.02	0.54	0.33	0.09	17.5	0.12
Wave 3*Gender	-0.14	0.51	-0.04	0.59	0.38	0.09	16.1	0.15
Age	-0.01	0.02	0.00	0.02	0.01	0.01	48.6	0.00
Cohabiting status	0.52	0.24	0.54	0.26	0.16	0.02	7.2	0.02
Wealth	-0.22	0.25	-0.33	0.26	0.14	-0.11	-41.5	0.03
Smoking	0.43	0.25	0.38	0.26	0.14	-0.05	-18.6	0.02
Physical Activity	0.49	0.18	0.50	0.19	0.11	0.01	4.2	0.99
Alcohol consumption	-0.24	0.19	-0.22	0.20	0.11	0.02	10.9	0.01

Abbreviation: FIML=full information maximum likelihood

MVNI failed to recover the coefficients for wave 2 and wave 3 with good accuracy and precision (Table 3.21).

Table 3.21 MVNI technique for depressive symptoms, logistic random intercept model with gender and time interaction, CHD group (model 2)

MVNI	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.80	0.44	0.73	0.43	0.08	-0.07	-15.6	0.01
Wave 2	0.01	0.34	0.22	0.37	0.16	0.21	57.0	0.07
Wave 3	-0.41	0.36	-0.09	0.40	0.19	0.32	79.9	0.14
Wave 2*Gender	-0.08	0.49	-0.14	0.54	0.22	-0.06	-11.8	0.05
Wave 3*Gender	-0.14	0.51	-0.18	0.58	0.27	-0.04	-6.6	0.07
Age	-0.01	0.02	0.00	0.02	0.01	0.01	55.1	0.00
Cohabiting status	0.52	0.24	0.52	0.26	0.10	0.00	-0.2	0.01
Wealth	-0.22	0.25	-0.27	0.26	0.10	-0.06	-21.5	0.01
Smoking	0.43	0.25	0.35	0.26	0.08	-0.08	-28.8	0.01
Physical Activity	0.49	0.18	0.38	0.19	0.07	-0.11	-56.4	0.02
Alcohol consumption	-0.24	0.19	-0.21	0.20	0.08	0.04	18.1	0.01

Abbreviation: MVNI=multivariate normal imputation

Two-fold FCS achieved the smallest biases of the parameters for gender, wave 2, wave 3 and interaction term between wave 2 and gender, compared with the FIML and MVNI, also the standard errors were very close to, if not the same as, the targeted values. MVNI estimated a slightly smaller bias than two-fold FCS for the interaction term between gender and wave 3 (-0.04 in MVNI and 0.06 in two-fold FCS), however, the standard error is larger in MVNI. Two-fold FCS recovered the targeted parameters of the other covariates very well, while FIML failed to recover the coefficient for wealth and MVNI the coefficient for physical activity with good accuracy and precision.

Table 3.22 Two-fold FCS technique for depressive symptoms, logistic random intercept model with gender and time interaction, CHD group (model 2)

Two-fold FCS	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.80	0.44	0.74	0.43	0.10	-0.06	-14.3	0.01
Wave 2	0.01	0.34	-0.01	0.34	0.19	-0.02	-5.2	0.03
Wave 3	-0.41	0.36	-0.41	0.37	0.24	0.01	2.0	0.06
Wave 2*Gender	-0.08	0.49	-0.09	0.50	0.27	-0.02	-3.1	0.07
Wave 3*Gender	-0.14	0.51	-0.08	0.53	0.35	0.06	10.6	0.13
Age	-0.01	0.02	0.00	0.02	0.01	0.01	62.5	0.00
Cohabiting status	0.52	0.24	0.53	0.24	0.12	0.01	3.7	0.02
Wealth	-0.22	0.25	-0.27	0.24	0.14	-0.05	-21.4	0.02
Smoking	0.43	0.25	0.36	0.25	0.11	-0.06	-25.4	0.02
Physical Activity	0.49	0.18	0.45	0.18	0.09	-0.04	-21.0	0.01
Alcohol consumption	-0.24	0.19	-0.20	0.19	0.10	0.05	24.8	0.01

Abbreviation: FCS=fully conditional specification

Results of the comparisons of missing data techniques for the model for the Well group are shown in Tables 3.23 to 3.25. In this model FIML did not recover the targeted parameters for gender, wave 2, wave 3 and the two the interaction terms, while the MVNI failed to recover the coefficients for wave 2 and wave 3. Two-fold FCS performed exceptionally well, compared to FIML and MVNI, it recovered the targeted parameters for gender, wave 2, wave 3 and the two interaction terms, to an impressive extent. Standard errors were almost the same as the targeted standard errors. All three techniques did not perform very well in recovering the coefficients for cohabiting status and physical activity; although for physical activity, the bias produced by two-fold FCS techniques was somewhat less troublesome than the biases produced by FIML and MVNI.

Table 3.23 FIML technique for depressive symptoms, logistic random intercept model with gender and time interaction, Well group (model 2)

FIML	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.97	0.22	0.58	0.17	0.19	-0.38	-229.1	0.19
Wave 2	0.39	0.20	0.20	0.19	0.12	-0.19	-97.5	0.05
Wave 3	0.05	0.21	-0.08	0.21	0.14	-0.13	-61.7	0.03
Wave 2*Gender	-0.37	0.25	-0.12	0.24	0.18	0.25	101.9	0.10
Wave 3*Gender	-0.26	0.26	-0.04	0.26	0.20	0.22	84.2	0.09
Age	-0.02	0.01	-0.02	0.01	0.00	0.00	-7.5	0.00
Cohabiting status	0.54	0.13	0.46	0.12	0.06	-0.07	-60.9	0.01
Wealth	-0.31	0.12	-0.32	0.11	0.05	-0.01	-13.1	0.00
Smoking	0.38	0.11	0.36	0.10	0.05	-0.02	-25.3	0.00
Physical Activity	0.43	0.08	0.39	0.08	0.04	-0.04	-55.2	0.00
Alcohol consumption	-0.12	0.09	-0.13	0.09	0.05	-0.01	-14.0	0.00

Abbreviation: FIML=full information maximum likelihood

Table 3.24 MVNI technique for depressive symptoms, logistic random intercept model with gender and time interaction, Well group (model 2)

MVNI	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.97	0.22	0.90	0.21	0.05	-0.06	-29.3	0.01
Wave 2	0.39	0.20	0.57	0.21	0.09	0.18	85.8	0.04
Wave 3	0.05	0.21	0.40	0.23	0.10	0.34	150.8	0.13
Wave 2*Gender	-0.37	0.25	-0.34	0.27	0.10	0.03	9.6	0.01
Wave 3*Gender	-0.26	0.26	-0.32	0.28	0.12	-0.06	-20.0	0.02
Age	-0.02	0.01	-0.02	0.01	0.00	0.00	27.3	0.00
Cohabiting status	0.54	0.13	0.45	0.14	0.05	-0.09	-68.0	0.01
Wealth	-0.31	0.12	-0.27	0.12	0.05	0.04	33.9	0.00
Smoking	0.38	0.11	0.37	0.11	0.04	-0.01	-10.9	0.00
Physical Activity	0.43	0.08	0.34	0.08	0.03	-0.09	-108.4	0.01
Alcohol consumption	-0.12	0.09	-0.12	0.10	0.04	-0.01	-7.4	0.00

Abbreviation: MVNI=multivariate normal imputation

Table 3.25 Two-fold FCS technique for depressive symptoms, logistic random intercept model with gender and time interaction, Well group (model 2)

Two-fold FCS	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.97	0.22	0.92	0.21	0.05	-0.04	-21.1	0.00
Wave 2	0.39	0.20	0.33	0.20	0.10	-0.06	-29.3	0.01
Wave 3	0.05	0.21	0.07	0.21	0.12	0.01	5.1	0.02
Wave 2*Gender	-0.37	0.25	-0.29	0.25	0.12	0.08	31.9	0.02
Wave 3*Gender	-0.26	0.26	-0.25	0.27	0.16	0.02	6.4	0.03
Age	-0.02	0.01	-0.02	0.01	0.00	0.00	20.9	0.00
Cohabiting status	0.54	0.13	0.46	0.13	0.06	-0.08	-63.4	0.01
Wealth	-0.31	0.12	-0.27	0.12	0.07	0.04	34.0	0.01
Smoking	0.38	0.11	0.38	0.11	0.05	-0.01	-7.5	0.00
Physical Activity	0.43	0.08	0.40	0.08	0.04	-0.03	-40.9	0.00
Alcohol consumption	-0.12	0.09	-0.11	0.09	0.06	0.00	4.3	0.00

Abbreviation: FCS=fully conditional specification

3.4 Discussion

This simulation study used a large, national longitudinal survey to assess the problem of handling an arbitrary pattern of missing data. The data set for this study had incomplete time-dependent outcomes (one continuous and one binary) and incomplete time-dependent and time-independent covariates (of different types). Therefore it was necessary to accommodate missingness for each follow-up survey, as well as unit non-response at each time. In order to investigate which technique could be suitable with this structure of data, the FIML technique was compared with two MI techniques: MVNI and the recently proposed two-fold FCS technique. The performance of each of the technique appeared to vary according to the type of outcome and to the amount of missing data. The results of the comparisons among the missing data techniques for quality of life (continuous outcome) and depressive symptoms (binary outcome) seemed to draw different conclusions on which of the three techniques for dealing with missing data was most suitable.

Table 3.26 summarises the performance of each technique in terms of accuracy and precision (assessed by the bias, MSE and the standardised bias percent) for the main targeted parameters obtained from the three models for quality of life. Although in the models for quality of life, the outcome variable was the variable with the largest proportion of missing data, the three techniques all performed well in recovering the targeted parameters (model 1 with gender and CHD interaction term), even though the two MI techniques outperformed FIML in recovering the interaction term with good precision and smaller bias values. This is an advantage of multiple imputation techniques: the interaction term can and should be accommodated in the imputation model thus reflecting the substantive model.

In model 2 (CHD group), the three techniques produced small biases for gender, wave 2 and wave 3, however, for the interaction terms accuracy was not within acceptable range in all three techniques (Table 3.26). In the model for the Well group the three techniques recovered the targeted parameters for gender and the two interaction terms with good accuracy and precision. MVNI did not perform as well as FIML and two-Fold FCS in recovering with good accuracy and precision the targeted parameters for wave 2 and wave 3 in the model for the Well group (Table 3.26).

Table 3.26 Summary of performance of the three missing data techniques for the models of quality of life

	FIML		MVNI		2-FOLD FCS	
	Accuracy	**Precision**	**Accuracy**	**Precision**	**Accuracy**	**Precision**
Model 1						
CHD	✓	✓	✓	✓	✓	✓
Gender	✓	✓	✓	✓	✓	✓
CHD*Gender	✓	×	✓	✓	✓	✓
Model 2 CHD group						
Gender	✓	✓	✓	✓	✓	✓
Wave 2	✓	✓	✓	✓	✓	✓
Wave 3	✓	✓	✓	✓	✓	✓
Wave 2*Gender	×	✓	×	✓	×	✓
Wave 3*Gender	×	✓	×	✓	×	✓
Model 2 Well group						
Gender	✓	✓	✓	✓	✓	✓
Wave 2	✓	✓	×	×	✓	✓
Wave 3	✓	✓	×	×	✓	×
Wave 2*Gender	✓	✓	✓	✓	✓	✓
Wave 3*Gender	✓	✓	✓	✓	✓	✓

The symbol ✓indicates that the estimates were close enough to infer that the targeted parameter estimates were recovered while the symbol × indicates that the opposite was true.

A different picture was given by the results involving the binary outcome (depressive symptoms). Table 3.27 summarises the performance of each technique in terms of accuracy and precision for the main targeted parameters obtained from the models for depressive symptoms. MVNI generally performed better than the FIML. However, two-fold FCS performed exceptionally well in all models compared to both FIML and MVNI, in terms of accuracy and precision. The relatively lower performance of FIML was observed in particular in the model with the interaction term between gender and time for the Well group, in which the precision and accuracy of the main targeted parameters (gender, wave 2 and 3, and interaction terms) were poor. Although MVNI performed better than FIML, it failed to achieve good precision and accuracy especially in the models with the interaction terms between gender and time (for both the CHD and Well groups). It seemed that both techniques that assumed a multivariate normal

distribution did not perform well with a binary outcome. The flexibility of two-fold FCS became obvious in the presence of a binary outcome for which an appropriate conditional distribution was specified in the imputation stage; also the doubly iterative procedure for imputing each wave of missing data seemed to work better, especially when the amount of missing data increased with time.

Table 3.27 Summary of performance of the three missing data techniques for the models of depressive symptoms

	FIML		MVNI		2-FOLD FCS	
	Accuracy	**Precision**	**Accuracy**	**Precision**	**Accuracy**	**Precision**
Model 1						
CHD	×	×	✓	✓	✓	✓
Gender	×	×	✓	×	✓	✓
CHD*Gender	×	✓	✓	✓	✓	✓
Model 2 CHD group						
Gender	✓	✓	✓	✓	✓	✓
Wave 2	✓	✓	✓	×	✓	✓
Wave 3	✓	✓	×	×	✓	✓
Wave 2*Gender	✓	✓	×	✓	✓	✓
Wave 3*Gender	✓	✓	✓	✓	✓	✓
Model 2 Well group						
Gender	×	×	✓	✓	✓	✓
Wave 2	×	×	✓	×	✓	✓
Wave 3	×	×	✓	×	✓	✓
Wave 2*Gender	×	×	✓	✓	✓	✓
Wave 3*Gender	×	×	✓	✓	✓	✓

The symbol ✓indicates that the estimates were close enough to infer that the targeted parameter estimates were recovered while the symbol × indicates that the opposite was true.

Given that the performance of each missing data technique was perfectly acceptable for the models involving the continuous outcome, while for the models involving the binary outcome two-fold FCS outperformed the other two techniques, the decision regarding which technique should be used for the analysis of Chapter 4 can be made on the basis of several considerations. First, it is recommended to include auxiliary variables predictive of missingness in the imputation model, even if they are not of interest in the substantive model, to reinforce the MAR assumption and to reduce the bias (Sterne et al., 2009). However, in this simulation study it was not possible to add any auxiliary

variables because this option is not available in Mplus when using a hierarchical model. If auxiliary variables were used in the imputation of missing data with the two MI techniques, true comparisons with FIML would not have been possible. Choosing one of the MI techniques over FIML in this particular setting has the advantage of allowing the inclusion of auxiliary variables. Also, when using FIML each model is estimated conditioned on the independent variables, therefore cases that are missing on the independent variables will be excluded from the analyses. In order to include these cases, independent variables were treated as dependent variables. In doing so, dichotomous variables may not always be identified and the standard errors. In the simulation study presented here this problem was solved by using independent variables with more than two categories and by treating them as continuous. However, distributional assumptions were made about them (i.e. normality). When a model includes several independent variables, many of which are binary, the FIML is not recommended.

Second, a further consideration should be made when choosing a technique that assumes multivariate normality. When MVNI was used, categorical and binary variables were imputed under the normal model and imputed values were rounded off to the nearest category. Although multivariate normality was assumed, it has been suggested that inferences based on multiple imputation can be robust to departures from the multivariate normality assumption if the amount of missing information is not large. This is because it often makes sense to use a normal model to create multiple imputations even when the observed data are somewhat non-normal, as supported by simulation studies described in Schafer (1997) and the original references therein. However, in this simulation study, the MVNI technique did not perform at its best in the case of a binary outcome probably due to the fact that the amount of missing data was relatively large (between 54% and 57%).

In the light of these considerations, I will apply the two-fold FCS technique for the treatment of missing data in the analysis of the original data set (presented in Chapter 4). Although the procedure could be computationally time-consuming in the presence of many variables and many waves, and it lacks theoretical underpinnings, its advantages are clear, especially in the presence of a non-continuous outcome, binary and categorical covariates and repeated measures.

Strengths and limitations

A major strength of this study is the use of a real data set to provide a suitable structure for simulating the 1,000 replicates, which simplifies the data generation procedures and avoids arbitrary choices. Also by replicating the patterns of missingness seen in the incomplete data set, a realistic framework was provided for simulating the missing data (Marshall et al., 2010).

Another major strength is that, to date and to my knowledge, this is the first study that compares simultaneously techniques for dealing with missing data in the presence of both continuous and binary outcomes. Furthermore this study addresses missing data (in the outcomes as well as covariates) due not only to item non-response but also to drop-out. This is also the first study that applies the recently proposed two-fold FCS to longitudinal data from a national survey and compares it with the FIML and the MVNI. Most of the studies that have compared FIML with MVNI and FCS used cross-sectional data.

One of the possible limitations of this simulation study is that MAR was assumed. The plausibility of the MAR assumption could have been affected by the fact that auxiliary variables were by design not included in the imputation model. In an earlier version of the simulation study, auxiliary variables were used to impute missing data under MVNI and two-fold FCS. It must be mentioned that the ability to recover the targeted parameters by the MVNI and two-fold FCS techniques did not depend upon the addition of these variables; rather they helped reinforce the MAR assumption. However, for the purpose of choosing the best technique to deal with missing data it was necessary to not use auxiliary variables in order to make the three techniques as comparable as possible as recommended by Mike Kenward (personal communication). Researchers that wish to strengthen the MAR assumptions may decide to opt for one of the MI techniques presented here for this type of analysis rather that FIML which does not allow the inclusion of auxiliary variables in the two-level modelling framework.

Another limitation is that in dealing with attrition, no distinction was made between drop-out due to death and other reasons for loss to follow-up. In longitudinal studies on ageing, an important concern is the potential for bias caused by individuals non-randomly dropping out of the study over time. It is known that selective attrition and mortality selection are intrinsically related to many ageing-related changes introducing

potential bias, and it has been suggested that a key distinction should be made between attrition and mortality selection (Harel et al., 2007). Attrition affects characteristics of the particular sample under study, whereas mortality affects both the definition of the population as well as the sample (Harel et al., 2007). Consequently, there has been an increasing interest in developing techniques for missing data in longitudinal studies that distinguish between attrition and mortality (Dufouil et al., 2004; Harel et al., 2007; Chang et al., 2009). Most of these approaches focus on attrition and death as the only source of missingness in the data. It must be noted that in this study, as in many other longitudinal studies, item-non response is also a considerable source of missingness. This topic is an area of research that is still developing; therefore it was decided that it was beyond the scope of this thesis to deal with the problem of distinguishing between deaths and attrition. Nevertheless, the limitation should be acknowledged and the problem addressed in future research.

In this simulation study by treating death and attrition as the same form of drop-out, I implicitly assumed that trajectories continue beyond death and therefore that the panel is immortal. Another possible approach could have been to exclude deaths from the sample and only deal with missing data due to item non-response and attrition. This approach was not adopted for two reasons: first the number of deaths was a much smaller proportion of missingness compared to drop-out. However, while the prevalence of those dropping out did not differ between the CHD group and the Well group neither at wave 2 (19.4% CHD and 19.5% Well) nor at wave 3 (31.8% CHD and 28.4% Well), the prevalence of death occurring after baseline was significantly higher in the CHD group (13.8%) compared to the Well group (5.3%). Therefore excluding deaths from the analysis could have introduced some bias. Second, the research question of this thesis was to explore gender differences in quality of life and depressive symptoms over time, therefore it seemed appropriate to consider the possible outcome a participant could have had if he or she had not died.

Longitudinal data can be thought of as clustered or two-level data (Goldstein, 2003). It has been suggested that if a data set to be imputed is multilevel, then the imputation model should be multilevel too (Carpenter and Goldstein, 2004). In recent years Carpenter and Goldstein have developed macros that implement multiple imputation in a multilevel data setting in MLwiN for normal and non-normal models of interest under the assumption of missing at random (Carpenter and Goldstein, 2004; Goldstein et al.,

2009; Carpenter et al., 2011). The macros set up a multilevel multivariate imputation model with the partially observed variables as responses, and fit this model in a Bayesian framework with uninformative priors using Markov Chain Monte Carlo techniques to impute a number of complete data sets. The choice not to use a multilevel model structure was made on the basis of the following limitations of the technique: first, although in theory the macros can handle missing data in the all the variables in the model of interest, in practice this may cause convergence difficulties (Carpenter et al., 2011). Therefore using the macros when there are missing data in almost all variables of the model of interest can cause problems. Second, the macros cannot handle missing categorical variables, not ideal especially for the imputation of the outcome depressive symptoms which is categorical. Only recently, the authors have developed REALCOM-IMPUTE software which performs multilevel multiple imputation and handles ordered and unordered categorical data (Carpenter et al., 2011). In the light of these limitations, it was decided not to use this technique to handle missing data in this study.

Only in more recent versions of Stata (11.1 and 12) it is possible to impute missing data under the MVNI technique, which can also be used to impute clustered data. Unfortunately when the analyses of this chapter were undertaken, only version 10 of Stata was available, hence SAS was used.

A last consideration to note is that the theory of multiple imputation for missing data requires that imputations be made conditional on the sampling design. As described in Chapter 2 (Section 1.1) the ELSA sample was drawn using a stratified multi-stage design which was clustered within postal sectors. However, the multiple imputation approaches considered in this study did not account for complex sampling design features, such as stratification and clustering. The choice not to condition on sample design was based on the fact that the outcome variables were not correlated with the design variables. It has been suggested that when this is the case, disregarding the design in multiple imputation models may provide acceptable inference (Reiter et al., 2006).

To conclude, the advantages of two-fold FCS over the FIML and MVNI techniques, especially when dealing with non-continuous variables, justify the required time and effort in implementing this technique for dealing with missing data. Based on the results

of this simulation study the two-fold FCS technique will be used in next chapter to deal with missing data when exploring longitudinally gender differences in quality of life and depressive symptoms in people with CHD.

Appendix 3

Appendix 3.1 Stata code for the generation of missing data

```
forvalues i = 1/1000 {
**********WAVE 1 MISSING ITEMS**********
****replicate the patterns of missing values***
gen x=runiform()
gen patt1mis=0 if x>0.089
replace patt1mis=1 if x<=0.089
tab patt1mis
replace  casp191=. if patt1mis==1
gen patt2mis=0 if x>0.011
replace patt2mis=1 if x<=0.011
tab patt2mis
replace cigst1=. if patt2mis==1
replace physact1=. if patt2mis==1
replace alcoh1=. if patt2mis==1
replace casp191=. if patt2mis==1
replace totcesd1=. if patt2mis==1
***use random binomial numbers  to replicate the remaining patterns of missing values***
gen qolw1=rbinomial(1, 0.016)
tab qolw1
replace  casp191=. if qolw1==1
gen totw1=rbinomial(1, 0.010)
tab totw1
replace  totwq5_b1=. if totw1==1
gen cigw1=rbinomial(1, 0.004)
tab cigw1
```

```
replace  cigst1=. if cigw1==1

gen paw1=rbinomial(1, 0.003)

tab paw1

replace  physact1=. if paw1==1

gen alw1=rbinomial(1, 0.006)

tab alw1

replace  alcoh1=. if alw1==1

gen csdw1=rbinomial(1, 0.016)

tab csdw1

replace  totcesd1=. if csdw1==1

**********WAVE 2 DROP OUT AND MORTALITY******

gen j=runiform()

gen missw2=0 if j>0.195

replace missw2=1 if j<=0.195

tab missw2

gen m=runiform()

gen deadw2=0 if m>0.034

replace deadw2=1 if m<=0.034

tab deadw2

replace  indsex2=. if missw2==1 |  deadw2==1

replace  indager2=.  if missw2==1 |  deadw2==1

replace marital2=.  if missw2==1 |  deadw2==1

replace totwq5_b2=.  if missw2==1 |  deadw2==1

replace cigst2=.  if missw2==1 |  deadw2==1

replace physact2=.  if missw2==1 |  deadw2==1

replace alcoh2=.  if missw2==1 |  deadw2==1

replace casp192=.  if missw2==1 |  deadw2==1

replace totcesd2=.  if missw2==1 |  deadw2==1
```

```
**********WAVE 2 ITEM NON RESPONSE********

****replicate the patterns of missing***

gen y=runiform()

gen patt1mis2=0 if y>0.100

replace patt1mis2=1 if y<=0.100

tab patt1mis2

replace casp192=. if patt1mis2==1

replace alcoh2=. if patt1mis2==1

gen patt2mis2=0 if y>0.030

replace patt2mis2=1 if y<=0.030

tab patt2mis2

replace casp192=. if patt2mis2==1

gen patt3mis2=0 if y>0.014

replace patt3mis2=1 if y<=0.014

tab patt3mis2

replace alcoh2=. if patt3mis2==1

gen patt4mis2=0 if y>0.014

replace patt4mis2=1 if y<=0.014

tab patt4mis2

replace totwq5_b2=. if patt4mis2==1

***use random binomial numbers  to replicate the remaining patterns of missing values***

gen tw2=rbinomial(1, 0.004)

replace  totwq5_b2=. if tw2==1

gen smw2=rbinomial(1, 0.002)

replace   cigst2=. if smw2==1

gen alw2=rbinomial(1, 0.008)

replace   alcoh2=. if alw2==1

gen qolw2=rbinomial(1, 0.008)
```

```
tab qolw2

replace  casp192=. if qolw2==1

gen cesdw2=rbinomial(1, 0.008)

replace  totcesd2=. if cesdw2==1

**********WAVE 3 DROP OUT AND MORTALITY**********

gen missw3=0 if j> 0.256

replace missw3=1 if j<=0.256

tab missw3

gen deadw3=0 if m> 0.070

replace deadw3=1 if m<=0.070

tab deadw3

replace  indsex3=.  if missw3==1 |  deadw3==1

replace  indager3=.  if missw3==1 |  deadw3==1

replace marital3=.  if missw3==1 |  deadw3==1

replace totwq5_b3=.  if missw3==1 |  deadw3==1

replace cigst3=.  if missw3==1 |  deadw3==1

replace physact3=.  if missw3==1 |  deadw3==1

replace alcoh3=.  if missw3==1 |  deadw3==1

replace casp193=.  if missw3==1 |  deadw3==1

replace totcesd3=.  if missw3==1 |  deadw3==1

**********WAVE 3 ITEM NON RESPONSE**********

****replicate the patterns of missing***

gen f=runiform()

gen patt1mis3=0 if f>0.103

replace patt1mis3=1 if f<=0.103

tab patt1mis3

replace casp193=. if patt1mis3==1

replace alcoh3=. if patt1mis3==1
```

```
gen patt2mis3=0 if f>0.030

replace patt2mis3=1 if f<=0.030

tab patt2mis3

replace alcoh3=. if patt2mis3==1

gen patt3mis3=0 if f>0.021

replace patt3mis3=1 if f<=0.021

tab patt3mis3

replace casp193=. if patt3mis3==1

replace alcoh3=. if patt3mis3==1

replace totcesd3=. if patt3mis3==1

gen patt4mis3=0 if f>0.019

replace patt4mis3=1 if f<=0.019

tab patt4mis3

replace totwq5_b3=. if patt3mis3==1

gen patt5mis3=0 if f>0.015

replace patt5mis3=1 if f<=0.015

tab patt5mis3

replace casp193=. if patt5mis3==1

***use random binomial numbers  to replicate the remaining patterns of missing values***

gen cesdw3=rbinomial(1, 0.007)

replace  totcesd3=. if cesdw3==1

gen tw3=rbinomial(1, 0.011)

replace  totwq5_b3=. if tw3==1

gen smw3=rbinomial(1, 0.001)

replace   cigst3=. if smw3==1

gen alw3=rbinomial(1, 0.011)

replace   alcoh3=. if alw3==1

gen qolw3=rbinomial(1, 0.009)
```

```
tab qolw3

replace  casp193=. if qolw3==1

gen _mj=`i'

save simulation`i'.dta

}
```

Appendix 3.2 Mplus coding for the estimations of the multilevel models with FIML

Random intercept model for quality of life with gender and CHD interaction term

```
 DATA: FILE = simulonglist.dat;

TYPE = MONTECARLO;

Variable:

  Names are

     idauniq wave totcesd marital totwq5_b physact al casp19 cigst chd

     cesd indager agec agesq indsex sexchd hseyr alcoh3 wealth3 gor hotenu

     hhtot fqethnr;

  Missing are all (-9999) ;


Usevariables are casp19 agec agesq chd marital

wealth3 cigst physact alcoh3 totcesd indsex

sexchd;


BETWEEN= chd indsex sexchd;

WITHIN = agec agesq marital wealth3 cigst physact alcoh3 totcesd;

CLUSTER = idauniq;

ANALYSIS: TYPE = TWOLEVEL;

estimator=ml;
```

```
MODEL:
  %WITHIN%
   casp19 ON agec*-0.06 agesq*-0.00 marital*-0.29 wealth3*1.94
   cigst*-0.31 physact*-0.55 alcoh3*0.61 totcesd*-1.30;
agec  marital wealth3 cigst physact alcoh3 totcesd;
%BETWEEN%
 casp19 ON chd*-1.48 indsex*1.09 sexchd*-0.38;
output:tech1;
```

Random intercept model for quality of life with gender and time interaction term, CHD group

```
DATA: FILE = simulonglist.dat;
TYPE = MONTECARLO;
Variable:
  Names are
     idauniq wave totcesd marital totwq5_b physact al casp19 cigst chd
     cesd indager agec agesq indsex sexchd hseyr alcoh3 wealth3 gor hotenu
     hhtot fqethnr;
  Missing are all (-9999) ;
Usevariables are casp19 agec agesq marital
wealth3 cigst physact alcoh3 totcesd indsex
wave2  sex  wave3 sex  wave2  wave3;
 USEOBSERVATIONS =chd EQ 1;
BETWEEN= indsex ;
WITHIN = wave2 wave3 wave2sex wave3sex agec agesq marital
wealth3 cigst physact alcoh3 totcesd;
CLUSTER = idauniq;
Define:
if (wave EQ 2) THEN wave2=1;
```

```
if (wave EQ 1 OR wave EQ 3)THEN wave2=0;
if (wave EQ 3)THEN wave3=1;
if (wave EQ 1 OR wave EQ 2)THEN wave3=0;
wave2sex=indsex*wave2;
wave3sex=indsex*wave3;

ANALYSIS: TYPE = TWOLEVEL;
 estimator=ml;
MODEL:
  %WITHIN%
casp19 ON  wave2*-0.43 wave3*-2.65 wave2sex*0.90 wave3sex*1.11 agec*0.05 agesq*0.00
    marital*-0.32 wealth3*2.53 cigst*-0.02 physact*-0.79 alcoh3*1.05 totcesd*-1.19;
agec marital wealth3 cigst physact alcoh3 totcesd;
%BETWEEN%
    casp19 ON indsex*0.28;
output: tech1;
```

Random intercept model for quality of life with gender and time interaction term, Well group

```
  DATA: FILE = simulonglist.dat;
TYPE = MONTECARLO;
Variable:
  Names are
     idauniq wave totcesd marital totwq5_b physact al casp19 cigst chd
     cesd indager agec agesq indsex sexchd hseyr alcoh3 wealth3 gor hotenu
     hhtot fqethnr;
  Missing are all (-9999) ;
Usevariables are casp19 agec agesq marital
```

```
wealth3 cigst physact alcoh3 totcesd indsex

wave2sex wave3sex wave2 wave3;

 USEOBSERVATIONS =chd EQ 0;

BETWEEN= indsex ;

WITHIN = wave2 wave3 wave2sex wave3sex agec agesq marital

wealth3 cigst physact alcoh3 totcesd;

CLUSTER = idauniq;

Define:

if (wave EQ 2) THEN wave2=1;

if (wave EQ 1 OR wave EQ 3)THEN wave2=0;

if (wave EQ 3)THEN wave3=1;

if (wave EQ 1 OR wave EQ 2)THEN wave3=0;

wave2sex=indsex*wave2;

wave3sex=indsex*wave3;

ANALYSIS: TYPE = TWOLEVEL;

 estimator=ml;

MODEL:

  %WITHIN%

    casp19 ON  wave2*-0.83 wave3*-2.33 wave2sex*0.19 wave3sex*-0.00 agec*0.03
agesq*0.00

    marital*-0.34 wealth3*1.78 cigst*-0.39 physact*-0.47 alcoh3*0.62 totcesd*-1.32;

    agec marital wealth3 cigst physact alcoh3 totcesd;

%BETWEEN%

    casp19 ON indsex*1.02;

output: tech1;
```

Random intercept model for depressive symptoms with gender and CHD interaction term

```
DATA: FILE = simulonglist.dat;
TYPE = MONTECARLO;
Variable:
  Names are
     idauniq wave totcesd marital totwq5_b physact al casp19 cigst chd
     cesd indager agec agesq indsex sexchd hseyr alcoh3 wealth3 gor hotenu
     hhtot fqethnr;
  Missing are all (-9999) ;
Usevariables are agec chd marital
wealth3 cigst physact alcoh3 cesd indsex
sexchd;
CATEGORICAL=cesd;
BETWEEN= chd indsex sexchd;
WITHIN = agec marital wealth3 cigst physact alcoh3;
CLUSTER = idauniq;
  ANALYSIS:
    TYPE = TWOLEVEL;
       estimator=ml;
        ALGORITHM=INTEGRATION;
     INTEGRATION=MONTECARLO;
        PROCESSORS = 2;
MODEL:
%WITHIN%
   cesd ON agec*-0.02  marital*0.53 wealth3*-0.31 cigst*0.39 physact*0.43 alcoh3*-0.140;
   agec marital wealth3 cigst physact alcoh3;
%BETWEEN%
```

```
  cesd ON chd*0.65 indsex*0.74 sexchd*0.03;
output: tech1;
```

Random intercept model for depressive symptoms with gender and time interaction term, CHD group

```
DATA: FILE = simulonglist.dat;
TYPE = MONTECARLO;
Variable:
 Names are
    idauniq wave totcesd marital totwq5_b physact al casp19 cigst chd
    cesd indager agec agesq indsex sexchd hseyr alcoh3 wealth3 gor hotenu
    hhtot fqethnr;
 Missing are all (-9999) ;
Usevariables are agec marital
wealth3 cigst physact alcoh3 cesd indsex
 wave2sex wave3sex wave2 wave3;
 USEOBSERVATIONS =chd EQ 1;
Categorical=cesd;
BETWEEN= indsex ;
WITHIN = wave2 wave3 wave2sex wave3sex agec marital
wealth3 cigst physact alcoh3;
CLUSTER = idauniq;
Define:
if (wave EQ 2) THEN wave2=1;
if (wave EQ 1 OR wave EQ 3)THEN wave2=0;
if (wave EQ 3)THEN wave3=1;
if (wave EQ 1 OR wave EQ 2)THEN wave3=0;
wave2sex=indsex*wave2;
```

```
wave3sex=indsex*wave3;
 ANALYSIS:
  TYPE = TWOLEVEL;
    estimator=ml;
     ALGORITHM=INTEGRATION;
   INTEGRATION=MONTECARLO;
     PROCESSORS = 2;
MODEL:
 %WITHIN%
  cesd ON wave2*0.01 wave3*-0.41 wave2sex*-0.08 wave3sex*-0.14 agec*-0.01
marital*0.52  wealth3*-0.22 cigst*0.43 physact*0.49 alcoh3*-0.24;
agec marital wealth3 cigst physact alcoh3;
%BETWEEN%
 cesd ON indsex*0.80 ;
output: tech1;
```

Random intercept model for depressive symptoms with gender and time interaction term, Well group

```
DATA: FILE = simulonglist.dat;
TYPE = MONTECARLO;
Variable:
 Names are
   idauniq wave totcesd marital totwq5_b physact al casp19 cigst chd
   cesd indager agec agesq indsex sexchd hseyr alcoh3 wealth3 gor hotenu
   hhtot fqethnr;
 Missing are all (-9999) ;
Usevariables are agec marital
wealth3 cigst physact alcoh3 cesd indsex
```

```
 wave2sex wave3sex wave2 wave3;
 USEOBSERVATIONS =chd EQ 0;
Categorical=cesd;
BETWEEN= indsex ;
WITHIN = wave2 wave3 wave2sex wave3sex agec marital
wealth3 cigst physact alcoh3;
CLUSTER = idauniq;
Define:
if (wave EQ 2) THEN wave2=1;
if (wave EQ 1 OR wave EQ 3)THEN wave2=0;
if (wave EQ 3)THEN wave3=1;
if (wave EQ 1 OR wave EQ 2)THEN wave3=0;
wave2sex=indsex*wave2;
wave3sex=indsex*wave3;
 ANALYSIS:
   TYPE = TWOLEVEL;
     estimator=ml;
      ALGORITHM=INTEGRATION;
    INTEGRATION=MONTECARLO;
     PROCESSORS = 2;
MODEL:
 %WITHIN%
  cesd ON wave2*0.39 wave3*0.05 wave2sex*-0.37 wave3sex*-0.26 agec*-0.02
  marital*0.54  wealth3*-0.31 cigst*0.38 physact*0.43 alcoh3*-0.12;
agec marital wealth3 cigst physact alcoh3;
%BETWEEN%
 cesd ON indsex*0.97;
output: tech1;
```

Appendix 3.3 SAS coding for the imputation of missing data

The following coding is used for the imputation of the first 100 replicates, it was then repeated 9 more times for the remaining 900

```
options mprint symbolgen mlogic;

%macro loop(count);

%do i=1 %to &count;

 proc mi data=Simulation&i out=midataw&i nimpute=5 seed=1375

  round=   1

  minimum= 12 11 12  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 -13 0    0 0 0

  maximum= 57 57 57  8 8 8 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3  33 1089 1 1 1;

  mcmc chain=multiple initial=em  outiter=outit

    timeplot (mean (casp191 casp192 casp193 totcesd1 totcesd2 totcesd3))

    acfplot (mean (casp191 casp192 casp193 totcesd1 totcesd2 totcesd3));

var casp191 casp192 casp193 totcesd1 totcesd2 totcesd3 marital1 marital2 marital3 cigst1
cigst2 cigst3 physact1 physact2  physact3 alcoh31 alcoh32 alcoh33  wealth31 wealth32
wealth33 agec1 agesq1 indsex1 chd1 sexchd;

run;

%end;

%mend;

%loop(200);
```

Appendix 3.4 Two-fold FCS

Table A3.1 Mean comparisons by number of cycles and wave for the outcome variables

	WAVE 1		WAVE 2		WAVE 3	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Quality of life						
100 cycles	44.9	0.04	44.3	0.1	42.8	0.1
10 cycles	44.9	0.06	44.3	0.1	42.8	0.1
Depressive symptoms						
100 cycles	1.12	0.01	1.20	0.02	1.1	0.0
10 cycles	1.12	0.01	1.18	0.02	1.1	0.0

Table A3.2 Mean comparisons between complete case data and imputed data after 3 among-times iterations, by wave

		WAVE 1		WAVE 2		WAVE 3
	Complete case	**Imputed**	**Complete case**	**Imputed**	**Complete case**	**Imputed**
	Mean (s.e.)	Mean* (s.e.)	Mean* (s.e.)	Mean* (s.e.)	Mean* (s.e.)	Mean* (s.e.)
Quality of life	44.9 (0.17)	44.8 (0.17)	44.3 (0.17)	44.2 (0.18)	42.7 (0.18)	42.6 (0.21)
Depressive symptoms	1.1 (0.04)	1.1 (0.04)	1.2 (0.04)	1.2 (0.04)	1.1 (0.04)	1.2 (0.04)
Cohabiting status	0.3 (0.01)	0.3 (0.01)	0.3 (0.01)	0.3 (0.01)	0.3 (0.01)	0.3 (0.01)
Wealth	2.2 (0.01)	2.2 (0.01)	2.1 (0.01)	2.1 (0.01)	2.2 (0.01)	2.2 (0.02)
Cigarette smoking	0.8 (0.02)	0.8 (0.02)	0.8 (0.02)	0.8 (0.02)	0.7 (0.02)	0.7 (0.02)
Physical activity	1.3 (0.02)	1.3 (0.02)	1.2 (0.02)	1.2 (0.02)	1.3 (0.02)	1.3 (0.02)
Alcohol consumption	1.1 (0.02)	1.1 (0.02)	1.2 (0.02)	1.2 (0.02)	1.2 (0.02)	1.2 (0.02)

* Average mean and standard error over 5 imputed data sets obtained according to Rubin's formula (Rubin, 1987)

Stata coding for the imputation of missing data with two-fold FCS

The following coding generates only 1 imputed data set for each replicate, it was then repeated 4 more times to generate 5 imputed data set for each replicate

```
forvalues i = 1/1000 {

**********IMPUTING WAVE1 VARIABLES**********

use
"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta",
clear

cd "C:\Temp\imput1"

 **********DRYRUN**********

ice totcesd1 totcesd2 chd1 sexchd marital1 cigst1  cigst2 physact1 physact2  alcoh31 alcoh32
casp191 casp192 wealth31 wealth32  agec1 agesq1 indsex1,  ///

cmd(totcesd1 physact1 cigst1 cigst2:ologit) ///

eq(totcesd1: totcesd2 chd1 indsex1 sexchd cigst1 physact1 alcoh31 wealth31 casp191 marital1
agec1, cigst1: cigst2 totcesd1 marital1 chd1 indsex1 sexchd  physact1 alcoh31 wealth31
casp191 agec1 agesq1, physact1: physact2 cigst1 totcesd1 marital1 chd1 indsex1 sexchd
alcoh31 wealth31 casp191 agec1 agesq1, alcoh31: alcoh32 physact1 cigst1 totcesd1 marital1
chd1 indsex1 sexchd  wealth31 casp191 agec1 agesq1, wealth31:  wealth32 alcoh31 physact1
cigst1 totcesd1 marital1 chd1 indsex1 sexchd  casp191 agec1 agesq1, casp191: casp192
wealth31 alcoh31 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  agec1 agesq1)
dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd1 totcesd2 chd1 sexchd marital1 cigst1  cigst2 physact1 physact2  alcoh31 alcoh32
casp191 casp192 wealth31 wealth32  agec1 agesq1 indsex1,  ///

cmd(totcesd1 physact1 cigst1 cigst2:ologit) ///

eq(totcesd1: totcesd2 chd1 indsex1 sexchd cigst1 physact1 alcoh31 wealth31 casp191 marital1
agec1, cigst1: cigst2 totcesd1 marital1 chd1 indsex1 sexchd  physact1 alcoh31 wealth31
casp191 agec1 agesq1, physact1: physact2 cigst1 totcesd1 marital1 chd1 indsex1 sexchd
alcoh31 wealth31 casp191 agec1 agesq1, alcoh31: alcoh32 physact1 cigst1 totcesd1 marital1
chd1 indsex1 sexchd  wealth31 casp191 agec1 agesq1, wealth31:  wealth32 alcoh31 physact1
cigst1 totcesd1 marital1 chd1 indsex1 sexchd  casp191 agec1 agesq1, casp191: casp192
wealth31 alcoh31 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  agec1 agesq1)
match(casp191) seed(1285964) cycles(10)  m(1) saving(impw1, replace)



use "C:\Temp\imput1\impw1.dta", clear

tab totcesd1 if _mj==1
```

```
tab casp191 if _mj==1

tab alcoh31 if _mj==1

keep idauniq totcesd1 casp191 physact1 cigst1 wealth31 alcoh31 _mj

keep if _mj==1

sum totcesd1 casp191 physact1 cigst1 wealth31 alcoh31

replace totcesd1= round(totcesd1)

replace alcoh31= round(alcoh31)

replace cigst1= round(cigst1)

replace physact1=round(physact1)

replace wealth31=round(wealth31)


save "C:\Temp\imput1\impw1.dta", replace

use
"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta",
clear

drop totcesd1 casp191 physact1 cigst1 wealth31 alcoh31

merge 1:1 idauniq using "C:\Temp\imput1\impw1.dta"

drop  _mj _merge

save "C:\Temp\imput1\simulimpw1.dta", replace

use "C:\Temp\imput1\simulimpw1.dta", clear

**********IMPUTING WAVE2 VARIABLES**********

 **********DRYRUN**********

ice totcesd1 totcesd2 totcesd3 chd1 sexchd marital1 marital2 marital3 cigst1  cigst2 cigst3
physact1 physact2  physact3 alcoh31 alcoh32 alcoh33 casp191 casp192 casp193 wealth31
wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd2 physact2 marital2 marital3 cigst2 cigst3:ologit) ///

eq(totcesd2: totcesd3 totcesd1 chd1 indsex1 sexchd  cigst2 physact2 alcoh32 wealth32
casp192 agec1, marital2:marital1 marital3 totcesd2 chd1 indsex1 sexchd cigst2 physact2
alcoh32 wealth32 casp192 agec1 agesq1,cigst2: cigst1 cigst3 totcesd2 chd1 indsex1 sexchd
physact2 alcoh32 wealth32 casp192 agec1 agesq1, physact2:  physact3 physact1 cigst2
totcesd2 chd1 indsex1 sexchd  alcoh32 wealth32 casp192 agec1 agesq1,alcoh32: alcoh31
alcoh33 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  wealth32 casp192 agec1 agesq1,
```

```
wealth32:  wealth31 wealth33 alcoh32 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  casp192 agec1 agesq1, casp192: casp191 casp193 wealth32 alcoh32 physact2 cigst2 totcesd2 chd1 indsex1 sexchd agec1 agesq1)dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd1 totcesd2 totcesd3 chd1 sexchd marital1 marital2 marital3 cigst1  cigst2 cigst3 physact1 physact2  physact3 alcoh31 alcoh32 alcoh33 casp191 casp192 casp193 wealth31 wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd2 physact2 marital2 marital3 cigst2 cigst3:ologit) ///

eq(totcesd2: totcesd3 totcesd1 chd1 indsex1 sexchd  cigst2 physact2 alcoh32 wealth32 casp192 agec1, marital2:marital1 marital3 totcesd2 chd1 indsex1 sexchd cigst2 physact2 alcoh32 wealth32 casp192 agec1 agesq1,cigst2: cigst1 cigst3 totcesd2 chd1 indsex1 sexchd physact2 alcoh32 wealth32 casp192 agec1 agesq1, physact2:  physact3 physact1 cigst2 totcesd2 chd1 indsex1 sexchd  alcoh32 wealth32 casp192 agec1 agesq1,alcoh32: alcoh31 alcoh33 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  wealth32 casp192 agec1 agesq1, wealth32:  wealth31 wealth33 alcoh32 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  casp192 agec1 agesq1, casp192: casp191 casp193 wealth32 alcoh32 physact2 cigst2 totcesd2 chd1 indsex1 sexchd agec1 agesq1) match(casp192) seed(1285964) cycles(10)  m(1) saving(impw2, replace)

use "C:\Temp\imput1\impw2.dta", clear

tab totcesd2 if _mj==1

tab casp192 if _mj==1

tab alcoh32 if _mj==1

keep idauniq totcesd2 casp192 physact2 cigst2 wealth32 alcoh32 marital2  _mj

keep if _mj==1

replace totcesd2= round(totcesd2)

replace alcoh32= round(alcoh32)

replace cigst2= round(cigst2)

replace physact2=round(physact2)

replace wealth32=round(wealth32)

replace marital2=round(marital2)

sum totcesd2 casp192 physact2 cigst2 wealth32 alcoh32 marital2

save "C:\Temp\imput1\impw2.dta",replace
```

```
use
"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta",
clear

drop totcesd2 casp192 physact2 cigst2 wealth32 alcoh32 marital2

merge 1:1 idauniq using "C:\Temp\imput1\impw2.dta"

drop  _mj _merge

save "C:\Temp\imput1\simulimpw2.dta", replace

use "C:\Temp\imput1\simulimpw2.dta", clear

**********IMPUTING WAVE3 VARIABLES**********

 **********DRYRUN**********

ice totcesd2 totcesd3 chd1 sexchd marital2 marital3 cigst2 cigst3 physact2  physact3 alcoh32
alcoh33 casp192 casp193 wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd3 physact3 marital3 cigst3:ologit) ///

eq(totcesd3: totcesd2 marital3 chd1 indsex1 sexchd  cigst3 physact3 alcoh33 wealth33
casp193 agec1, marital3:marital2 totcesd3 chd1  indsex1 sexchd  cigst3 physact3 alcoh33
wealth33 casp193 agec1 agesq1, cigst3: cigst2 totcesd3 marital3 chd1 indsex1 sexchd
physact3  alcoh33 wealth33 casp193 agec1 agesq1, physact3: physact2 cigst3 totcesd3
marital3 chd1 indsex1 sexchd alcoh33 wealth33 casp193 agec1 agesq1, alcoh33: alcoh32
physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  wealth33 casp193 agec1 agesq1,
wealth33: wealth32 alcoh33 physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  casp193
agec1 agesq1, casp193:casp192 wealth33 alcoh33 physact3 cigst3 totcesd3 marital3 chd1
indsex1 sexchd  agec1 agesq1) dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd2 totcesd3 chd1 sexchd marital2 marital3 cigst2 cigst3 physact2  physact3 alcoh32
alcoh33 casp192 casp193 wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd3 physact3 marital3 cigst3:ologit) ///

eq(totcesd3: totcesd2 marital3 chd1 indsex1 sexchd  cigst3 physact3 alcoh33 wealth33
casp193 agec1, marital3:marital2 totcesd3 chd1  indsex1 sexchd  cigst3 physact3 alcoh33
wealth33 casp193 agec1 agesq1, cigst3: cigst2 totcesd3 marital3 chd1 indsex1 sexchd
physact3  alcoh33 wealth33 casp193 agec1 agesq1, physact3: physact2 cigst3 totcesd3
marital3 chd1 indsex1 sexchd alcoh33 wealth33 casp193 agec1 agesq1, alcoh33: alcoh32
physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  wealth33 casp193 agec1 agesq1,
wealth33: wealth32 alcoh33 physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  casp193
agec1 agesq1, casp193:casp192 wealth33 alcoh33 physact3 cigst3 totcesd3 marital3 chd1
indsex1 sexchd  agec1 agesq1) match(casp193) seed(1285964) cycles(10)  m(1) saving(impw3,
replace)
```

```
use "C:\Temp\imput1\impw3.dta", clear

tab totcesd3 if _mj==1

tab casp193 if _mj==1

sum casp193

tab alcoh33 if _mj==1

keep idauniq totcesd3 casp193 physact3 cigst3 wealth33 alcoh33 marital3 _mj

keep if _mj==1

replace totcesd3= round(totcesd3)

replace alcoh33= round(alcoh33)

replace cigst3= round(cigst3)

replace physact3=round(physact3)

replace wealth33=round(wealth33)

replace marital3=round(marital3)

sum totcesd3 casp193 physact3 cigst3 wealth33 alcoh33 marital3

save "C:\Temp\imput1\impw3.dta",replace


***START SECOND ROUND of imputation using the new imputed values****

****impute wave1 using new imputed wave2 variables******

use
"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta",
clear

drop totcesd2 casp192 physact2 cigst2 wealth32 alcoh32

merge 1:1 idauniq using "C:\Temp\imput1\impw2.dta"

drop  _mj _merge

save "C:\Temp\imput1\w2imp.dta" , replace

use "C:\Temp\imput1\w2imp.dta" , clear

cd "C:\Temp\imput1\round2"

 **********DRYRUN**********

ice totcesd1 totcesd2 chd1 sexchd marital1 cigst1  cigst2 physact1 physact2  alcoh31 alcoh32
casp191 casp192 wealth31 wealth32  agec1 agesq1 indsex1,  ///
```

```
cmd(totcesd1 physact1 cigst1:ologit) ///

eq(totcesd1: totcesd2 chd1 indsex1 sexchd cigst1 physact1 alcoh31 wealth31 casp191 marital1 agec1,  cigst1: cigst2 totcesd1 marital1 chd1 indsex1 sexchd  physact1 alcoh31 wealth31 casp191 agec1 agesq1, physact1: physact2 cigst1 totcesd1 marital1 chd1 indsex1 sexchd alcoh31 wealth31 casp191 agec1 agesq1, alcoh31: alcoh32 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  wealth31 casp191 agec1 agesq1, wealth31:  wealth32 alcoh31 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  casp191 agec1 agesq1, casp191: casp192 wealth31 alcoh31 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  agec1 agesq1) dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd1 totcesd2 chd1 sexchd marital1 cigst1  cigst2 physact1 physact2  alcoh31 alcoh32 casp191 casp192 wealth31 wealth32  agec1 agesq1 indsex1,  ///

cmd(totcesd1 physact1 cigst1:ologit) ///

eq(totcesd1: totcesd2 chd1 indsex1 sexchd cigst1 physact1 alcoh31 wealth31 casp191 marital1 agec1,  cigst1: cigst2 totcesd1 marital1 chd1 indsex1 sexchd  physact1 alcoh31 wealth31 casp191 agec1 agesq1, physact1: physact2 cigst1 totcesd1 marital1 chd1 indsex1 sexchd alcoh31 wealth31 casp191 agec1 agesq1, alcoh31: alcoh32 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  wealth31 casp191 agec1 agesq1, wealth31:  wealth32 alcoh31 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  casp191 agec1 agesq1, casp191: casp192 wealth31 alcoh31 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  agec1 agesq1) match(casp191) seed(1285964) cycles(10)  m(1) saving(impw1, replace)


replace totcesd1= round(totcesd1)

replace alcoh31= round(alcoh31)

replace cigst1= round(cigst1)

replace physact1=round(physact1)

replace wealth31=round(wealth31)


use "C:\Temp\imput1\round2\impw1.dta", clear

keep idauniq totcesd1 casp191 physact1 cigst1 wealth31 alcoh31 _mj

keep if _mj==1

sum totcesd1 casp191 physact1 cigst1 wealth31 alcoh31

save "C:\Temp\imput1\round2\impw1.dta", replace

use"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta", clear
```

```
drop totcesd1 casp191 physact1 cigst1 wealth31 alcoh31 totcesd3 casp193 physact3 cigst3
wealth33 alcoh33 marital3

merge 1:1 idauniq using "C:\Temp\imput1\round2\impw1.dta"

drop  _merge

merge 1:1 idauniq using "C:\Temp\imput1\impw3.dta"

drop  _merge

save "C:\Temp\imput1\round2\w1(r2)w3(r1).dta", replace

use "C:\Temp\imput1\round2\w1(r2)w3(r1).dta", clear

**********IMPUTING WAVE2 VARIABLES**********

 **********DRYRUN**********

ice totcesd1 totcesd2 totcesd3 chd1 sexchd marital1 marital2 marital3 cigst1  cigst2 cigst3
physact1 physact2  physact3 alcoh31 alcoh32 alcoh33 casp191 casp192 casp193 wealth31
wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd2 physact2 marital2 cigst2:ologit) ///

eq(totcesd2: totcesd3 totcesd1 chd1 indsex1 sexchd  cigst2 physact2 alcoh32 wealth32
casp192 agec1, marital2:marital1 marital3 totcesd2 chd1 indsex1 sexchd cigst2 physact2
alcoh32 wealth32 casp192 agec1 agesq1,cigst2: cigst1 cigst3 totcesd2 chd1 indsex1 sexchd
physact2 alcoh32 wealth32 casp192 agec1 agesq1, physact2:  physact3 physact1 cigst2
totcesd2 chd1 indsex1 sexchd  alcoh32 wealth32 casp192 agec1 agesq1,alcoh32: alcoh31
alcoh33 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  wealth32 casp192 agec1 agesq1,
wealth32:  wealth31 wealth33 alcoh32 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  casp192
agec1 agesq1, casp192: casp191 casp193 wealth32 alcoh32 physact2 cigst2 totcesd2 chd1
indsex1 sexchd agec1 agesq1)dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd1 totcesd2 totcesd3 chd1 sexchd marital1 marital2 marital3 cigst1  cigst2 cigst3
physact1 physact2  physact3 alcoh31 alcoh32 alcoh33 casp191 casp192 casp193 wealth31
wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd2 physact2 marital2 cigst2 :ologit) ///

eq(totcesd2: totcesd3 totcesd1 chd1 indsex1 sexchd  cigst2 physact2 alcoh32 wealth32
casp192 agec1, marital2:marital1 marital3 totcesd2 chd1 indsex1 sexchd cigst2 physact2
alcoh32 wealth32 casp192 agec1 agesq1,cigst2: cigst1 cigst3 totcesd2 chd1 indsex1 sexchd
physact2 alcoh32 wealth32 casp192 agec1 agesq1, physact2:  physact3 physact1 cigst2
totcesd2 chd1 indsex1 sexchd  alcoh32 wealth32 casp192 agec1 agesq1,alcoh32: alcoh31
alcoh33 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  wealth32 casp192 agec1 agesq1,
wealth32:  wealth31 wealth33 alcoh32 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  casp192
agec1 agesq1, casp192: casp191 casp193 wealth32 alcoh32 physact2 cigst2 totcesd2 chd1
```

```
indsex1 sexchd agec1 agesq1) match(casp192) seed(1285964) cycles(10)  m(1) saving(impw2, replace)

use "C:\Temp\imput1\round2\impw2.dta", clear

tab totcesd2 if _mj==1

tab casp192 if _mj==1

tab alcoh32 if _mj==1

keep idauniq totcesd2 casp192 physact2 cigst2 wealth32 alcoh32 marital2  _mj

keep if _mj==1

replace totcesd2= round(totcesd2)

replace alcoh32= round(alcoh32)

replace cigst2= round(cigst2)

replace physact2=round(physact2)

replace wealth32=round(wealth32)

replace marital2=round(marital2)

sum totcesd2 casp192 physact2 cigst2 wealth32 alcoh32 marital2  _mj


save "C:\Temp\imput1\round2\impw2.dta",replace

use"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta", clear

drop totcesd1 casp191 physact1 cigst1 wealth31 alcoh31  totcesd2 casp192 physact2 cigst2 wealth32 marital2 alcoh32

merge 1:1 idauniq using "C:\Temp\imput1\round2\impw1.dta"

 drop  _mj _merge

 merge 1:1 idauniq using "C:\Temp\imput1\round2\impw2.dta"

 drop  _mj _merge

save "C:\Temp\imput1\round2\w1w2imp.dta", replace

use "C:\Temp\imput1\round2\w1w2imp.dta", clear

**********IMPUTING WAVE3 VARIABLES**********

 **********DRYRUN**********
```

```
ice totcesd2 totcesd3 chd1 sexchd marital2 marital3 cigst2 cigst3 physact2  physact3 alcoh32
alcoh33 casp192 casp193 wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd3 physact3 marital3 cigst3:ologit) ///

eq(totcesd3: totcesd2 marital3 chd1 indsex1 sexchd  cigst3 physact3 alcoh33 wealth33
casp193 agec1, marital3:marital2 totcesd3 chd1  indsex1 sexchd  cigst3 physact3 alcoh33
wealth33 casp193 agec1 agesq1, cigst3: cigst2 totcesd3 marital3 chd1 indsex1 sexchd
physact3  alcoh33 wealth33 casp193 agec1 agesq1, physact3: physact2 cigst3 totcesd3
marital3 chd1 indsex1 sexchd alcoh33 wealth33 casp193 agec1 agesq1, alcoh33: alcoh32
physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  wealth33 casp193 agec1 agesq1,
wealth33: wealth32 alcoh33 physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  casp193
agec1 agesq1, casp193:casp192 wealth33 alcoh33 physact3 cigst3 totcesd3 marital3 chd1
indsex1 sexchd  agec1 agesq1) dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd2 totcesd3 chd1 sexchd marital2 marital3 cigst2 cigst3 physact2  physact3 alcoh32
alcoh33 casp192 casp193 wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd3 physact3 marital3 cigst3:ologit) ///

eq(totcesd3: totcesd2 marital3 chd1 indsex1 sexchd  cigst3 physact3 alcoh33 wealth33
casp193 agec1, marital3:marital2 totcesd3 chd1  indsex1 sexchd  cigst3 physact3 alcoh33
wealth33 casp193 agec1 agesq1, cigst3: cigst2 totcesd3 marital3 chd1 indsex1 sexchd
physact3  alcoh33 wealth33 casp193 agec1 agesq1, physact3: physact2 cigst3 totcesd3
marital3 chd1 indsex1 sexchd alcoh33 wealth33 casp193 agec1 agesq1, alcoh33: alcoh32
physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  wealth33 casp193 agec1 agesq1,
wealth33: wealth32 alcoh33 physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  casp193
agec1 agesq1, casp193:casp192 wealth33 alcoh33 physact3 cigst3 totcesd3 marital3 chd1
indsex1 sexchd  agec1 agesq1) match(casp193) seed(1285964) cycles(10)  m(1) saving(impw3,
replace)

use "C:\Temp\imput1\round2\impw3.dta", clear

tab totcesd3 if _mj==1

tab casp193 if _mj==1

sum casp193

tab alcoh33 if _mj==1

keep idauniq totcesd3 casp193 physact3 cigst3 wealth33 alcoh33 marital3   _mj

keep if _mj==1

replace totcesd3= round(totcesd3)

replace alcoh33= round(alcoh33)

replace cigst3= round(cigst3)
```

```
replace physact3=round(physact3)

replace wealth33=round(wealth33)

replace marital3=round(marital3)

sum totcesd3 casp193 physact3 cigst3 wealth33 alcoh33 marital3

save "C:\Temp\imput1\round2\impw3.dta",replace

**********ROUND 3**********

******impute wave1 variables using wave2 variables imputed in round2***

use"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta",
clear

drop totcesd2 casp192 physact2 cigst2 wealth32 alcoh32

merge 1:1 idauniq using "C:\Temp\imput1\round2\impw2.dta"

drop  _mj _merge

save "C:\Temp\imput1\round3\w2(r2).dta" , replace

use "C:\Temp\imput1\round3\w2(r2).dta" , clear

cd "C:\Temp\imput1\round3"

 **********DRYRUN**********

ice totcesd1 totcesd2 chd1 sexchd marital1 cigst1  cigst2 physact1 physact2  alcoh31 alcoh32
casp191 casp192 wealth31 wealth32  agec1 agesq1 indsex1,  ///

cmd(totcesd1 physact1 cigst1:ologit) ///

eq(totcesd1: totcesd2 chd1 indsex1 sexchd  cigst1 physact1 alcoh31 wealth31 casp191
marital1 agec1,  cigst1: cigst2 totcesd1 marital1 chd1 indsex1 sexchd  physact1 alcoh31
wealth31 casp191 agec1 agesq1, physact1: physact2 cigst1 totcesd1 marital1 chd1 indsex1
sexchd  alcoh31 wealth31 casp191 agec1 agesq1, alcoh31: alcoh32 physact1 cigst1 totcesd1
marital1 chd1 indsex1 sexchd  wealth31 casp191 agec1 agesq1, wealth31:  wealth32 alcoh31
physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  casp191 agec1 agesq1, casp191:
casp192 wealth31 alcoh31 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  agec1
agesq1) dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd1 totcesd2 chd1 sexchd marital1 cigst1  cigst2 physact1 physact2  alcoh31 alcoh32
casp191 casp192 wealth31 wealth32  agec1 agesq1 indsex1,  ///

cmd(totcesd1 physact1 cigst1:ologit) ///

eq(totcesd1: totcesd2 chd1 indsex1 sexchd  cigst1 physact1 alcoh31 wealth31 casp191
marital1 agec1,  cigst1: cigst2 totcesd1 marital1 chd1 indsex1 sexchd  physact1 alcoh31
```

```
wealth31 casp191 agec1 agesq1, physact1: physact2 cigst1 totcesd1 marital1 chd1 indsex1
sexchd  alcoh31 wealth31 casp191 agec1 agesq1, alcoh31: alcoh32 physact1 cigst1 totcesd1
marital1 chd1 indsex1 sexchd  wealth31 casp191 agec1 agesq1, wealth31:  wealth32 alcoh31
physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  casp191 agec1 agesq1, casp191:
casp192 wealth31 alcoh31 physact1 cigst1 totcesd1 marital1 chd1 indsex1 sexchd  agec1
agesq1) match(casp191) seed(1285964) cycles(10)  m(1) saving(impw1, replace)

 use "C:\Temp\imput1\round3\impw1.dta", clear

tab totcesd1 if _mj==1

tab casp191 if _mj==1

tab alcoh31 if _mj==1

keep idauniq totcesd1 casp191 physact1 cigst1 wealth31 alcoh31 _mj

keep if _mj==1

sum totcesd1 casp191 physact1 cigst1 wealth31 alcoh31

replace totcesd1= round(totcesd1)

replace alcoh31= round(alcoh31)

replace cigst1= round(cigst1)

replace physact1=round(physact1)

replace wealth31=round(wealth31)

save "C:\Temp\imput1\round3\impw1.dta", replace


**********IMPUTING WAVE2 VARIABLES**********

**use wave1 variables imputed at round3 and wave3 variables imputed at round2*****

use"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta",
clear

drop totcesd1 casp191 physact1 cigst1 wealth31 alcoh31  totcesd3 casp193 physact3 cigst3
wealth33 alcoh33 marital3

merge 1:1 idauniq using "C:\Temp\imput1\round3\impw1.dta"

drop  _merge

 merge 1:1 idauniq using "C:\Temp\imput1\round2\impw3.dta"

drop  _merge

save "C:\Temp\imput1\round3\w1(r3)w3(r2).dta", replace
```

```
use "C:\Temp\imput1\round3\w1(r3)w3(r2).dta", clear

**********IMPUTING WAVE2 VARIABLES**********

 **********DRYRUN**********

ice totcesd1 totcesd2 totcesd3 chd1 sexchd marital1 marital2 marital3 cigst1  cigst2 cigst3
physact1 physact2  physact3 alcoh31 alcoh32 alcoh33 casp191 casp192 casp193 wealth31
wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd2 physact2 marital2 cigst2:ologit) ///

eq(totcesd2: totcesd3 totcesd1 chd1 indsex1 sexchd  cigst2 physact2 alcoh32 wealth32
casp192 agec1, marital2:marital1 marital3 totcesd2 chd1 indsex1 sexchd cigst2 physact2
alcoh32 wealth32 casp192 agec1 agesq1,cigst2: cigst1 cigst3 totcesd2 chd1 indsex1 sexchd
physact2 alcoh32 wealth32 casp192 agec1 agesq1, physact2:  physact3 physact1 cigst2
totcesd2 chd1 indsex1 sexchd  alcoh32 wealth32 casp192 agec1 agesq1,alcoh32: alcoh31
alcoh33 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  wealth32 casp192 agec1 agesq1,
wealth32:  wealth31 wealth33 alcoh32 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  casp192
agec1 agesq1, casp192: casp191 casp193 wealth32 alcoh32 physact2 cigst2 totcesd2 chd1
indsex1 sexchd agec1 agesq1)dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd1 totcesd2 totcesd3 chd1 sexchd marital1 marital2 marital3 cigst1  cigst2 cigst3
physact1 physact2  physact3 alcoh31 alcoh32 alcoh33 casp191 casp192 casp193 wealth31
wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd2 physact2 marital2 cigst2:ologit) ///

eq(totcesd2: totcesd3 totcesd1 chd1 indsex1 sexchd  cigst2 physact2 alcoh32 wealth32
casp192 agec1, marital2:marital1 marital3 totcesd2 chd1 indsex1 sexchd cigst2 physact2
alcoh32 wealth32 casp192 agec1 agesq1,cigst2: cigst1 cigst3 totcesd2 chd1 indsex1 sexchd
physact2 alcoh32 wealth32 casp192 agec1 agesq1, physact2:  physact3 physact1 cigst2
totcesd2 chd1 indsex1 sexchd  alcoh32 wealth32 casp192 agec1 agesq1,alcoh32: alcoh31
alcoh33 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  wealth32 casp192 agec1 agesq1,
wealth32:  wealth31 wealth33 alcoh32 physact2 cigst2 totcesd2 chd1 indsex1 sexchd  casp192
agec1 agesq1, casp192: casp191 casp193 wealth32 alcoh32 physact2 cigst2 totcesd2 chd1
indsex1 sexchd agec1 agesq1) match(casp192) seed(1285964) cycles(10)  m(1) saving(impw2,
replace)

use "C:\Temp\imput1\round3\impw2.dta", clear

keep idauniq totcesd2 casp192 physact2 cigst2 wealth32 alcoh32 marital2  _mj

keep if _mj==1

replace totcesd2= round(totcesd2)

replace alcoh32= round(alcoh32)

replace cigst2= round(cigst2)
```

```
replace physact2=round(physact2)

replace wealth32=round(wealth32)

replace marital2=round(marital2)

sum totcesd2 casp192 physact2 cigst2 wealth32 alcoh32 marital2  _mj

save "C:\Temp\imput1\round3\impw2.dta", replace

**********IMPUTE WAVE3 VARIABLES**********

********use wave2 variables imputed at round3**********

use"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta",
clear

drop   totcesd2 casp192 physact2 cigst2 wealth32 marital2 alcoh32

 merge 1:1 idauniq using "C:\Temp\imput1\round3\impw2.dta"

 drop _merge

  save "C:\Temp\imput1\round3\w2(r3).dta", replace

  use "C:\Temp\imput1\round3\w2(r3).dta", clear

**********IMPUTING WAVE3 VARIABLES**********

 **********DRYRUN**********

ice totcesd2 totcesd3 chd1 sexchd marital2 marital3 cigst2 cigst3 physact2  physact3 alcoh32
alcoh33 casp192 casp193 wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd3 physact3 marital3 cigst3:ologit) ///

eq(totcesd3: totcesd2 marital3 chd1 indsex1 sexchd  cigst3 physact3 alcoh33 wealth33
casp193 agec1, marital3:marital2 totcesd3 chd1  indsex1 sexchd  cigst3 physact3 alcoh33
wealth33 casp193 agec1 agesq1, cigst3: cigst2 totcesd3 marital3 chd1 indsex1 sexchd
physact3  alcoh33 wealth33 casp193 agec1 agesq1, physact3: physact2 cigst3 totcesd3
marital3 chd1 indsex1 sexchd alcoh33 wealth33 casp193 agec1 agesq1, alcoh33: alcoh32
physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  wealth33 casp193 agec1 agesq1,
wealth33: wealth32 alcoh33 physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  casp193
agec1 agesq1, casp193:casp192 wealth33 alcoh33 physact3 cigst3 totcesd3 marital3 chd1
indsex1 sexchd  agec1 agesq1) dryrun

**********REAL RUN 10 CYCLES**********

ice totcesd2 totcesd3 chd1 sexchd marital2 marital3 cigst2 cigst3 physact2  physact3 alcoh32
alcoh33 casp192 casp193 wealth32 wealth33  agec1 agesq1 indsex1, ///

cmd(totcesd3 physact3 marital3 cigst3:ologit) ///
```

```
eq(totcesd3: totcesd2 marital3 chd1 indsex1 sexchd  cigst3 physact3 alcoh33 wealth33
casp193 agec1, marital3:marital2 totcesd3 chd1  indsex1 sexchd  cigst3 physact3 alcoh33
wealth33 casp193 agec1 agesq1, cigst3: cigst2 totcesd3 marital3 chd1 indsex1 sexchd
physact3  alcoh33 wealth33 casp193 agec1 agesq1, physact3: physact2 cigst3 totcesd3
marital3 chd1 indsex1 sexchd alcoh33 wealth33 casp193 agec1 agesq1, alcoh33: alcoh32
physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  wealth33 casp193 agec1 agesq1,
wealth33: wealth32 alcoh33 physact3 cigst3 totcesd3 marital3 chd1 indsex1 sexchd  casp193
agec1 agesq1, casp193:casp192 wealth33 alcoh33 physact3 cigst3 totcesd3 marital3 chd1
indsex1 sexchd  agec1 agesq1) match(casp193) seed(1285964) cycles(10)  m(1) saving(impw3,
replace)

use "C:\Temp\imput1\round3\impw3.dta", clear

tab totcesd3 if _mj==1

tab casp193 if _mj==1

tab alcoh33 if _mj==1

keep idauniq totcesd3 casp193 physact3 cigst3 wealth33 alcoh33 marital3   _mj

keep if _mj==1

replace totcesd3= round(totcesd3)

replace alcoh33= round(alcoh33)

replace cigst3= round(cigst3)

replace physact3=round(physact3)

replace wealth33=round(wealth33)

replace marital3=round(marital3)

sum totcesd3 casp193 physact3 cigst3 wealth33 alcoh33 marital3

save "C:\Temp\imput1\round3\impw3.dta", replace

use
"C:\Users\paoletta\Documents\PHD\ANALYSES\longitudinal\simulations\simulation`i'.dta",
clear

drop  agec1 agesq1 agec2 agesq2 agec3 agesq3 al1 al2 al3 totwq5_b1 totwq5_b2 totwq5_b3
sexchd indsex2 indsex3 totcesd1 totcesd2 totcesd3 alcoh31 alcoh32  alcoh33 casp191 casp192
casp193 physact1 physact2 physact3 cigst1 cigst2 cigst3 wealth31 wealth32 wealth33 alcoh31
alcoh32 alcoh33 marital2 marital3

merge 1:1 idauniq using "C:\Temp\imput1\round3\impw1.dta"

drop _mj _merge

merge 1:1 idauniq using "C:\Temp\imput1\round3\impw2.dta"
```

```
drop _mj _merge

merge 1:1 idauniq using "C:\Temp\imput1\round3\impw3.dta"

drop _mj _merge

gen indsex2=indsex1

gen indsex3=indsex1

***reshape data***

gen cesd1=0 if totcesd1>=0 & totcesd1<=2

replace cesd1=1 if totcesd1>=3

replace cesd1=. if totcesd1==.

tab cesd1

gen cesd2=0 if totcesd2>=0 & totcesd2<=2

replace cesd2=1 if totcesd2>=3

replace cesd2=. if totcesd2==.

tab cesd2

gen cesd3=0 if totcesd3>=0 & totcesd3<=2

replace cesd3=1 if totcesd3>=3

replace cesd3=. if totcesd3==.

tab cesd3

tab alcoh31

tab alcoh32

tab alcoh33

drop fqethnr1 hhtot1 hotenu1 gor1 hseyr1 hseyr2 hseyr3 gor2 gor3 hotenu2 hotenu3 hhtot2
hhtot3

reshape long chd cesd totcesd  marital  indsex  indager cigst wealth3 physact casp19 alcoh3, i(
idauniq ) j( wave )

sum indager if wave==3

sum indager

sum indager

gen agec=indager-63
```

```
gen agesq=agec*agec

tab indsex

sum indsex

lab val indsex sexlab

tab indsex

gen sexchd= chd*indsex

save "C:\Temp\r`i'imputation1.dta", replace

}
```

Appendix 3.5 Depressive symptoms results of FIML with the targeted parameters obtained using a Monte Carlo algorithm for integration

Table A3.3 Table A3.3 FIML technique for depressive symptoms, random intercept model with gender and CHD interaction (model 1)

	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
CHD	0.66	0.25	0.44	0.24	0.14	-0.22	-92.5	0.06
Gender	0.74	0.16	0.50	0.14	0.12	-0.24	-173.4	0.07
Gender*CHD	-0.02	0.36	0.20	0.34	0.15	0.22	64.8	0.07
Age	-0.02	0.01	-0.02	0.01	0.00	0.00	21.5	0.00
Cohabiting status	0.52	0.11	0.47	0.11	0.05	-0.04	-37.6	0.00
Wealth	-0.31	0.10	-0.33	0.10	0.05	-0.02	-21.7	0.00
Smoking	0.38	0.10	0.36	0.09	0.05	-0.02	-20.3	0.00
Physical Activity	0.43	0.07	0.40	0.07	0.04	-0.03	-46.0	0.00
Alcohol consumption	-0.14	0.08	-0.15	0.08	0.04	-0.01	-15.4	0.00

Estimates obtained by using Monte Carlo algorithm for integration

Table A3.4 FIML technique for depressive symptoms, random intercept model with gender and time interaction, CHD group (model 2)

	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.82	0.361	0.65	0.42	0.39	-0.17	-40.0	0.18
Wave 2	0.02	0.333	-0.06	0.37	0.19	-0.08	-21.5	0.04
Wave 3	-0.40	0.354	-0.50	0.41	0.24	-0.10	-25.0	0.07
Wave 2*Gender	-0.11	0.463	0.01	0.54	0.33	0.12	22.6	0.12
Wave 3*Gender	-0.17	0.478	-0.05	0.59	0.37	0.12	20.9	0.16
Age	-0.01	0.018	0.00	0.02	0.01	0.01	43.5	0.00
Cohabiting status	0.54	0.236	0.54	0.26	0.16	0.00	-1.7	0.02
Wealth	-0.24	0.24	-0.33	0.26	0.14	-0.08	-31.6	0.03
Smoking	0.44	0.242	0.38	0.26	0.14	-0.06	-24.6	0.02
Physical Activity	0.50	0.18	0.50	0.19	0.11	0.00	-1.5	1.01
Alcohol consumption	-0.26	0.181	-0.22	0.20	0.11	0.04	17.6	0.01

Estimates obtained by using Monte Carlo algorithm for integration

Table A3.5 FIML technique for depressive symptoms, random intercept model with gender and time interaction, Well group (model 2)

	β	$SE(\beta)$	$\bar{\hat{\beta}}$	$SE(\hat{\beta})$	$SD(\hat{\beta})$	*Bias*	*Stand. Bias*	*MSE*
Gender	0.85	0.203	0.58	0.17	0.19	-0.26	-156.7	0.11
Wave 2	0.36	0.194	0.20	0.19	0.12	-0.16	-81.7	0.04
Wave 3	0.04	0.204	-0.08	0.21	0.14	-0.12	-54.8	0.03
Wave 2*Gender	-0.33	0.242	-0.12	0.24	0.18	0.20	84.7	0.08
Wave 3*Gender	-0.23	0.249	-0.04	0.26	0.20	0.19	71.8	0.08
Age	-0.02	0.009	-0.02	0.01	0.00	0.00	-14.0	0.00
Cohabiting status	0.54	0.122	0.46	0.12	0.06	-0.07	-58.5	0.01
Wealth	-0.31	0.111	-0.32	0.11	0.05	-0.02	-13.9	0.00
Smoking	0.37	0.101	0.36	0.10	0.05	-0.01	-8.8	0.00
Physical Activity	0.42	0.074	0.39	0.08	0.04	-0.03	-34.2	0.00
Alcohol consumption	-0.13	0.088	-0.13	0.09	0.05	0.00	1.5	0.00

Estimates obtained by using Monte Carlo algorithm for integration

Chapter 4: A study of gender differences in quality of life and depressive symptoms among older people with CHD using the English Longitudinal Study of Ageing

In the previous chapter three techniques for handling missing data were compared in order to find the best technique to be applied to the ELSA data for the analysis of this chapter. Results from this comparison supported the two-fold FCS technique. In this chapter gender differences in quality of life and depressive symptoms among people with CHD are explored using three waves of ELSA, from 2002-03 to 2006-07. First gender differences in quality of life and depressive symptoms are explored comparing people with CHD to the Well group; second men and women are compared with respect to trajectories over four years in quality of life and depressive symptoms once they have experienced CHD. The same trajectory model is run in the Well group to understand whether similar results are found in a disease-free group.

4.1 Introduction

Chapter 1 reviewed the literature on gender differences in quality of life and depressive symptoms among people with CHD. In this chapter, I summarise briefly the findings from studies that have looked at gender differences in quality of life and depressive symptoms over time.

One year post-myocardial infarction women scored significantly lower than men on the following domains of HRQOL: physical function (Brink et al., 2005; Norris et al., 2007), bodily pain, social function (Brink et al., 2005), mental health (Westin et al., 1999; Norris et al., 2007), general health (Westin et al., 1999) and emotional health (Bogg et al., 2000). One year post-myocardial infarction the gender differences found at one month post-myocardial infarction in self-esteem and family interaction domains of HRQOL no longer persisted (Westin et al., 1999). Results from one study shown that gender differences in the mental health dimension of HRQOL at one year post-myocardial infarction did not persist once the model was adjusted for demographic, clinical, co-morbid and psychosocial covariates (Norris et al., 2007). Results from other studies reported an improvement in HRQOL after one year post-myocardial infarction as compared to five months in both men and women (Brink et al., 2005) and between four months and one year post-MI (Kristofferzon et al., 2005b).

Longitudinal studies exploring gender differences in depressive symptoms among people with CHD, did not report any difference between men and women in their prevalence of depression at five months (Brink et al., 2005) and at one year (Wiklund et al., 1993; Brink et al., 2005) post-myocardial infarction. However, women had improved in that they reported less depression than they had at five months post-myocardial infarction (Brink et al., 2005). Others found that one year post-myocardial infarction women were more likely than men to be depressed (Norris et al., 2007).

Bjerkeset et al., (2005) specifically examined gender differences in depression during the five years after myocardial infarction. Women had a high initial risk for depression, with a significant decrease after two years, while in men the risk for depression was only increased in the two to five years post-MI (Bjerkeset et al., 2005).

The review of the literature identified some limitations and gaps in current knowledge. To summarise, there are no studies that have addressed gender differences in depression following angina. All studies except two (Bjerkeset et al., 2005; Ford et al., 2008), used small samples from selected community hospitals which might affect generalisation of results and none has specifically focussed on an older population. Only one study had a follow-up greater than a year. Lastly, only some of the studies have adjusted their analyses for covariates other than age and sex (Norris et al., 2004, 2007; Bjerkeset et al., 2005; Brink et al., 2005; Ford et al., 2008) and none have accounted for missing data.

The first objective of this chapter is to explore gender differences in quality of life and depressive symptoms among men and women with CHD compared to healthy people. It is hypothesised that people aged over fifty years who had had a CHD event would be at higher risk of experiencing depressive symptoms and poor quality of life than those who have not.

The second objective is to compare men and women with respect to trajectories over four years of quality of life and depressive symptoms once they have experienced CHD, while adjusting for several important covariates such as age, gender, cohabiting status, retirement status, education, wealth, smoking status, physical activity, alcohol consumption, pain, physical functioning, social support and social networks. These

variables were selected from the literature and were found to correlate with CHD and quality of life and depressive symptoms.

It is hypothesised that the shapes of trajectories over time of quality of life and depressive symptoms are different in men and women following the CHD event. Women with CHD are at higher risk than men of reporting depressive symptoms and lower quality of life. However, women tend towards a time-limited reaction (in terms of depressive symptoms and poor quality of life) to the actual CHD event, while men seem less able to adapt to the long-term consequences of the event. The analyses are repeated for people in the Well group to understand whether similar gender differences are found in their trajectories of quality of life and depressive symptoms.

4.2 Methods

4.2.1 Data

As explained in Chapter 2, the sample size is restricted to participants with a CHD event occurred in the two years preceding the baseline interview (2002-03) and to healthy participants (Well group). The total sample size is 4,496 in wave 1 (2002-03); 3,465 in wave 2 (2004-05) and 3,031 in wave 3 (2006-07). A total of 894 people were in the CHD group and 3,601 in the Well group. Missing data were imputed as described in section 4.2.3.

The variables used for the analyses in this chapter have been described in details in Chapter 2, in the measures section (2.2). Briefly, the main outcome measures are quality of life (measured by the CASP-19) and depressive symptoms (measured by the CESD-8). For quality of life a total score is used (ranging from 6 to 57 with higher scores indicating better quality of life); while depressive symptoms is a dummy variable coded as: 0 '0-2 symptoms' of depression and 1 '3+ symptoms' of depression. The use of a cut-off point of three or more depressive symptoms to indicate symptomatic depression is in line with previous studies that have used this abridged version of the scale (Steffick, 2000). Covariates used to adjust the models are: gender, age, cohabiting status (cohabiting with a partner vs not cohabiting), employment status (in paid employment, completely retired, other such as permanently unable to work, not currently in paid employment, looking after home or family), educational attainment (high and medium vs low), quintile of non-pension wealth, smoking status (never smoked and ex-smoker

vs current smoker), alcohol consumption (less than three times a week vs three times a week and more), physical activity (high and medium activity vs sedentary), pain (not troubled with pain vs often troubled with pain), ADLs (0 to 6 limitations with activities of daily living), positive support (score) and number of close friends.

4.2.2 Models

In longitudinal studies an individual's responses over time are correlated with each other. Ignoring the dependence between observations results in an underestimation of the standard error of the parameter estimates and provides inefficient estimates of the parameters of interest (Goldstein, 2003). Multilevel models can be used to explicitly model the clustered structure of the longitudinal data, which is considered as two-level data with *i= 1,2,...n* denoting occasions and *j=1,2,...m* denoting individuals (or units); the occasions are therefore clustered within individuals that represent the level-two units with measurement occasions as the level-one units. To model individual trajectories of the outcome variable over time taking into account the hierarchical structure of the data, it is possible to fit a two-level multilevel model with a random intercept as follows:

$$
\begin{aligned}
&y_{ij} = \beta_0 + \beta_1 x_{ij} + (u_j + e_{ij}) \\
&e_{ij} \sim N(0, \sigma_e^2) \\
&u_i \sim N(0, \sigma_u^2)
\end{aligned}
\qquad (1)
$$

where u_j denotes the random error associated with the individual level variation, and e_{ij} denotes the random error specific to each occasion *i* for an individual *j*. Each of the error term is assumed to be identically and independently normally distributed with mean 0 and variances equal to σ_u^2 and σ_e^2 at the individual and occasion levels. It follows that the variance of y_{ij} is $\sigma_u^2 + \sigma_e^2$. In model (1) $\beta_0+\beta_1 x_{ij}$ is the fixed part and is the equation for the overall average line, the slope is not allowed to be random, while the intercept $\beta_0+ u_j$ is allowed to vary from individual to individual because it includes the parameter u_j from the random part (Reise and Duan, 2003). By having a random intercept, both the variance between individuals and the variance between repeated observations are estimated. The random intercept model (1) only expresses the dependent variable y_{ij} as a function of a single predictor x_{ij}. It is possible to extend the model by adding time-constant and time-varying covariates (continuous or categorical), interaction terms and also a measure of time (continuous or categorical). The multilevel

model (1) can also be fitted when the dependent variable has a distribution other than normal. So a model for a binary dependent variable is expressed as follows:

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 x_{ij} + u_j \qquad (2)$$

$u_i \sim N(0, \sigma_u^2)$ where $\frac{\pi_{ij}}{1-\pi_{ij}}$ is the odds that $y_{ij}=1$ at occasion *i* for individual *j*, u_j denotes the random error associated with the individual level variation with residual variance equal to σ_u^2. Model (2) does not include a level-one residual e_{ij} because it is an equation for the probability $\frac{\pi_{ij}}{1-\pi_{ij}}$ rather than for the outcome y_{ij} (Goldstein, 2003). The level-one residual variance σ_e^2 cannot change and is conventionally fixed at $\pi^2/3 = 3.29$.

Random intercept models were estimated as follows:

$$y_{ij} = \beta_0 + \sum_{p=1}^{3} \beta_p x_{pj} + \sum_{p=1}^{14} \beta_p x_{p_{ij}} + u_j + e_{ij} \qquad (3)$$

Where y_{ij} is quality of life for individual j at time i, x_{pj} are time-invariant factors gender, CHD (at wave 1), and the interaction term between CHD and gender; x_{pij} are time-varying factors age (a linear and quadratic term), cohabiting status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends. Age, number of close friends and positive support were all centred to the mean.

A logit model was estimated for depressive symptoms as follows:

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \sum_{p=1}^{3} \beta_p x_{p_j} + \sum_{p=1}^{12} \beta_p x_{p_{ij}} + u_j \qquad (4)$$

Where $\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right)$ is the log odds of having depressive symptoms, for individual j at time i, x_{pj} are time-invariant factors such as gender, CHD (at wave 1), and the interaction term between CHD and gender; x_{pij} are time-varying factors such as age, cohabiting status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of

close friends). Age, number of close friends and positive support were all centred to the mean.

In order to compare trajectories of quality of life and depressive symptoms of men and women, the following random intercept models were estimated:

$$y_{ij} = \beta_0 + \beta_1 x_{1j} + \beta_2 t_{ij} + \beta_3 (t_{ij} * x_{1j}) + \sum_{p=1}^{14} \beta_p x_{p_{ij}} + u_j + e_{ij} \qquad (5)$$

Where y_{ij} is quality of life for individual j at time i, x_{1j} is gender, t_{ij} denotes time and takes three discrete values denoting the first, second and third wave, t_{ij}* x_{1j} denotes the interaction term between time and gender; x_{pij} are time-varying factors described in model 3). The model was run separately for people with CHD and people in the Well group.

The same model as 3) was estimated for depressive symptoms as follows:

$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 x_{1j} + \beta_2 t_{ij} + \beta_3 (t_{ij} * x_{1j}) + \sum_{p=1}^{12} \beta_p x_{p_{ij}} + u_j \qquad (6)$$

Where $\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right)$ is the log odds of having depressive symptoms, for individual j at time i, x_{1j} is gender, t_{ij} denotes time and takes three discrete values denoting the first, second and third wave, t_{ij}* x_{1j} denotes the interaction term between time and gender; x_{pij} are time-varying factors described in model 4). The model was run separately for people with CHD and people in the Well group.

4.2.3 Treatment of missing data

The two-fold FCS technique to impute missing data described in Chapter 3 is used here. Five imputed data sets were created and results estimated were combined according to Rubin's rule (1987) which averages the estimates and their standard errors into a single set of values. When using a MI method it is recommended to include auxiliary variables predictive of missingness in the imputation model even if they are not of interest in the substantive model. This strategy helps to reinforce the MAR assumption (Clarke and Hardy, 2007) and to reduce the bias (Sterne et al., 2009). Therefore five variables

(measured at wave 1) predictive of non-response (Taylor et al., 2007; Scholes et al., 2009) were added to the imputation model, these variables are: year of HSE interview, Government Office Region (GOR), housing tenure, number of people in the household and ethnicity (Table 4.1). These variables were used to create cross-sectional non-response weights in ELSA (Scholes et al., 2009), non-responders were more likely than responders to have the following characteristics: sampled from HSE 1999 (rather than 1998 or 2001), living in London, renting or other 'non-owning' category compared with owner-occupiers (recorded in wave 1, or HSE if missing in wave 1), increasing household size; non-white ethnicity.

For comparison, results from the augmented samples are presented together with those obtained from the observed sample (with missing data). The term "observed" implies that the sample has missing data and therefore the panel is unbalanced. Analyses were carried out in Stata 10 using *xtmixed* and *xtmelogit* commands for multilevel modelling which accommodate unbalanced panels using maximum likelihood. All analyses for the observed sample were restricted to participants with valid answers on both quality of life and depressive symptoms.

Table 4.1 Auxiliary variables used in the imputation model

Year of HSE interview	**N**	**%**
1998 and 2001	3,615	80.4
1999	881	19.6
Total	4,496	100
GOR	**N**	**%**
Other regions	4,064	90.4
London	431	9.6
Missing	1	0.0
Total	4,496	100
Housing tenure	**N**	**%**
Own it	3,712	82.6
Renting and other non-owing	784	17.4
Total	4,496	100
Number of people in the household	**N**	**%**
1	1,038	23.1
2	2,490	55.4
3	600	13.4
4	269	6.0
5	71	1.6
6	23	0.5
7 and more	5	0.1
Total	4,496	100
Ethnicity	**N**	**%**
White	4,360	97.0
Non-white	136	3.0
Total	4,496	100

4.3 Results

4.3.1 Comparisons of characteristics between the sample with missing data and the augmented sample

Table 4.2 compares the characteristics at each wave of the study for the sample with missing data and the augmented sample (imputed). The mean quality of life and standard deviation of the sample with missing data and augmented samples were very similar, if not the same, at all waves. There was also no difference in the means of positive support and close friends between the augmented sample and the sample with missing data at each wave. These three variables were collected in the self-completion booklet, and participants were requested to post it back. Therefore there are usually more missing values on these variables than in the variables collected through the face-

to-face interview. It was expected to see more differences in the sample means obtained from the augmented sample and those obtained from the sample with missing data. For example, another variable collected in the self-completion booklet was alcohol consumption (at wave 2 and wave 3). The prevalence of those drinking on 3 or more days a week was the same at wave 1 and wave 2, while at wave 3 was lower in the augmented sample (36.6%) than in the sample with missing data (38.1%) but the difference was not statistically significant ($p>0.05$). This is also surprising, given that people tend to under-report their alcohol consumption, it was expected that people who drink more were less likely to answer this question, and that the prevalence was then underestimated.

For depressive symptoms we see no differences in the prevalence at wave 1, but at wave 2 and wave 3 the prevalence of depressive symptoms is over 1.6 percentage points higher for the augmented sample (wave 2: 19.3% sample with missing data and 20.9% augmented sample; wave 3: 17.9% sample with missing data and 19.6% augmented sample), although the difference is not statistically significant ($p>0.05$).

Similarly, the prevalence of those not cohabiting, completely retired, and those with 1 or more ADLs, are on average over 2 percentage points higher in the augmented sample at wave 2 and wave 3 compared to the sample with missing data. The differences were statistically significant only for not cohabiting at wave 3 ($p<0.05$). At wave 3 the prevalence of those physically inactive is 2.9 percentage points higher in the augmented sample than the sample with missing data, although not statistically significant. These results are line with what expected.
At all waves, the prevalence of those in the poorest wealth quintile are significantly higher in the augmented sample compared to the sample with missing data ($p<0.05$), while the prevalence of those in the 4^{th} wealth and 3^{rd} wealth quintiles are slightly lower, but not significant, for the augmented sample at wave 1 and wave 3.

Surprisingly, the prevalence of people with low education was over 11 percentage points lower in the augmented sample than in the sample with missing data at wave 2 ($p<0.05$) and wave 3 ($p<0.001$). This difference is consistent with the higher rates of attrition among people in the lowest education category.

For all the other variables, figures based on the sample with missing data are very similar if not the same of those in the augmented sample.

Table 4.2 Comparisons of sample characteristics for the sample with missing data and augmented samples

	Wave 1		Wave 2		Wave 3	
	Observed	**Imputed**	**Observed**	**Imputed**	**Observed**	**Imputed**
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Quality of life	43.8 (8.2)	43.8 (8.0)	42.9 (8.6)	42.7 (8.7)	41.3 (8.8)	41.2 (8.7)
Positive support	27.1 (5.4)	27.1 (5.4)	27.0 (5.5)	27.1 (5.4)	27.6(5.3)	27.5 (5.4)
Close friends	8.0 (5.7)	7.9 (5.4)	8.5 (5.8)	8.5 (6.0)	9.4 (7.7)	9.1 (7.4)
	%(95%CI)	**%(95%CI)**	**%(95%CI)**	**%(95%CI)**	**%(95%CI)**	**%(95%CI)**
Depressive symptoms	20.3 (19.1; 21.5)	20.3 (19.1; 21.5)	19.3 (17.9; 20.6)	20.9 (19.7; 22.1)	17.9 (16.5; 19.2)	19.6 (18.4; 20.7)
Not cohabiting	28.4 (27.0; 29.7)	28.4 (27.0; 29.7)	29.4 (27.8; 30.9)	31.6 (30.3; 33.0)	31.2 (29.6; 32.9)	34.4 (33.0; 35.8)
Completely retired	42.4 (41.0; 43.9)	42.4 (41.0; 43.9)	47.7 (46.0; 49.4)	49.4 (47.9; 50.8)	52.7 (50.9; 54.5)	55.4 (53.9; 56.8)
Other*	12.5 (11.6; 13.5)	12.5 (11.6; 13.5)	12.8 (11.6; 13.9)	13.3 (12.3; 14.3)	11.0 (9.9; 12.1)	11.7 (10.8; 12.7)
Low education	50.8 (49.3; 52.2)	50.8 (49.3; 52.2)	48.4 (46.7; 50.1)	37.3 (35.9; 38.7)	47.3 (45.5; 49.1)	34.4 (33.0; 35.8)
Wealth 4th	18.7 (17.5; 19.8)	18.7 (17.5; 19.8)	20.7 (19.4; 22.1)	19.6 (18.4; 20.7)	21.6 (20.1; 23.1)	19.7 (18.5; 20.9)
Wealth 3rd	20.4 (19.2; 21.6)	20.4 (19.2; 21.6)	20.6 (19.3; 22.0)	20.4 (19.2; 21.6)	20.7 (19.2; 22.1)	19.9 (18.7; 21.0)
Wealth 2nd	20.7 (19.5; 21.9)	20.7 (19.5; 21.9)	19.1 (17.8; 20.4)	19.9 (18.8; 21.1)	17.9 (16.5; 19.3)	20.6 (19.4; 21.7)
Wealth Poorest	23.8 (22.6; 25.0)	23.8 (22.6; 25.0)	16.9 (15.7; 18.2)	20.5 (19.3; 21.7)	15.7 (14.4; 17.0)	21.8 (20.5; 23.0)
Current smoker	18.7 (17.6; 19.9)	18.7 (17.6; 19.9)	16.0 14.8; 17.2	16.3 (15.2; 17.3)	14.1 (12.8; 15.3)	14.4 (13.4; 15.4)
Physically inactive	61.0 (59.6; 62.5)	61.0 (59.6; 62.5)	59.0 (57.4; 60.6)	60.3 (58.8; 61.7)	61.0 (59.2; 62.7)	62.9 (61.5; 64.3)
Alcohol ≥3 days a week	29.3 (28.0; 30.7)	29.3 (28.0; 30.7)	37.7 (36.0; 39.4)	37.4 (36.0; 38.8)	38.1 (36.2; 40.0)	36.6 (35.2; 38.0)
Often troubled with pain	25.4 (24.1; 26.7)	25.4 (24.1; 26.7)	15.4 (14.2; 16.6)	15.7 (14.6; 16.8)	13.2 (12.0; 14.4)	14.0 (13.0; 15.0)
1 or more ADLs	11.8 (10.8; 12.7)	11.8 (10.8; 12.7)	13.5 (12.4; 14.7)	15.6 (14.6; 16.7)	14.8 (13.5; 16.1)	17.3 (16.2; 18.4)

Reference categories: Well, Male, Cohabiting, In paid employment High/medium education, Richest wealth, Non-smoker, Physically active, Drinks <3 days a week, Not troubled with pain, No ADLs, No depressive symptoms. The total sample size of the observed sample is 4,496 in wave 1; 3,465 in wave 2 and 3,031 in wave 3 (which varies for each variable according to item non-response).
*Permanently unable to work, not currently in paid employment, looking after home or family

4.3.2 Quality of life

Table 4.3 reports the combined results from the five imputed data sets as well as observed data, for the linear random intercept model for the quality of life outcome, with an interaction term between CHD and gender. The adjusted average quality of life (i.e. for someone aged 63, cohabiting, in paid employment, with medium and high educational attainment, in the richest quintile of wealth, non-smoker, physically active, drinks on less than 3 days a week, not troubled with pain, without ADLs and without depressive symptoms) for men in the CHD group, was 45.7, while for women in the CHD group the mean quality of life was 47.1. Men and women in the CHD group had an adjusted mean quality of life that was on average over one point lower than men and women in the Well group ($p<0.001$) (47.7 and 48.1 respectively). The interaction term between CHD and gender was statistically significant ($p<0.05$).

Covariates in the model negatively related to quality of life were: increasing age, not cohabiting, retired and other employment status, poorest wealth, smoking, physical inactivity, increasing number of difficulties with ADLs and increasing number of depressive symptoms. Low education, increased positive support and number of close friends and family were positively related to quality of life. The between-individual variance was 19.4 (41% of the total variance), which expresses the variation in quality of life due to unobserved differences between individuals after controlling for covariates. The within individual variance was 27.5 (59% of the total variance), and expresses the variation in quality of life due to differences within individuals over time after controlling for covariates. We can conclude that a greater proportion of the unexplained variability in quality of life was due to differences within individuals over time.

The last two columns of Table 4.3 report the results of the same model based on observed data (with missing data), results were consistent with those based on the imputed data and of similar magnitude, with the exception of the coefficient for gender and the interaction term that were not significant. The standard errors of estimates based on observed data were larger than those obtained from the imputed data. Also, in the results obtained from the sample with missing data, gender, the quadratic effect of age, education, smoking status, alcohol consumption and limitations with ADLs were not significantly related to quality of life.

Table 4.3 Linear random intercept model for quality of life

	Imputed			Observed		
	Coef.	Std. Err.	P-value	Coef.	Std. Err.	P-value
CHD	-2.00	0.27	0.000	-2.51	0.40	0.000
Gender	0.41	0.19	0.027	0.38	0.25	0.123
CHD*Gender	0.98	0.40	0.016	0.47	0.60	0.431
Age (centred 63)	-0.04	0.01	0.003	-0.03	0.02	0.044
Age squared	0.00	0.00	0.414	0.00	0.00	0.018
Not cohabiting	-0.49	0.16	0.002	-0.24	0.25	0.343
Completely retired	-0.87	0.16	0.000	-0.64	0.24	0.009
Other*	-2.09	0.21	0.000	-2.43	0.32	0.000
Low education	0.61	0.14	0.000	0.14	0.11	0.210
Wealth 4th	-1.03	0.18	0.000	-0.92	0.26	0.000
Wealth 3rd	-1.88	0.19	0.000	-1.61	0.28	0.000
Wealth 2nd	-2.52	0.20	0.000	-2.60	0.31	0.000
Wealth Poorest	-3.95	0.23	0.000	-4.04	0.37	0.000
Current smoker	-0.52	0.20	0.008	-0.32	0.28	0.258
Physically inactive	-1.04	0.13	0.000	-1.05	0.19	0.000
Drinks alcohol ≥3 days a week	0.43	0.14	0.001	0.30	0.21	0.138
Often troubled with pain	-1.46	0.15	0.000	-1.88	0.24	0.000
ADLs	-1.75	0.07	0.000	-0.00	0.11	0.974
Positive support (centred at 27)	0.27	0.01	0.000	0.30	0.02	0.000
Close friends (centred at 8)	0.05	0.01	0.000	0.03	0.01	0.032
Depressive symptoms	-4.25	0.15	0.000	-4.84	0.25	0.000
Constant	47.69	0.22	0.000	47.98	0.32	0.000
Model						
Between variance	19.4	0.03		20.92	1.08	
Within variance	27.5	0.02		23.75	0.72	

Reference categories: Well, Male, Cohabiting, In paid employment High/medium education, Richest wealth, Non-smoker, Physically active, Drinks <3 days a week, Not troubled with pain, No depressive symptoms.

* Permanently unable to work, not currently in paid employment, looking after home or family

Table 4.4 reports the results of the linear random intercept model with gender and time interaction, among people with CHD. Results of trajectories over time of quality of life by gender among people with CHD and the Well group are also summarised graphically in Figure 4.1. Results based on the imputed data showed that among men with CHD the adjusted mean quality of life at baseline was 47.4, which decreased significantly to 46.5 at two year follow-up ($p<0.05$) and to 44.6 at four year follow-up (wave 3) ($p<0.001$). The adjusted mean quality of life at baseline for women with CHD was 48.2, which remained stable at two year follow-up (48.4), and then decreased significantly to 46.5 at four year follow-up ($p<0.001$). The interaction terms between wave and gender were not significant, suggesting no gender difference in the rate of change over time in quality of life. The p-value for the coefficient of gender was non-significant, indicating that at baseline the quality of life of men and women did not differ. However, at two year follow-up (wave 2) and at four year follow-up (wave 3) women had significantly higher quality of life than men (wave 2: 46.5 [95%CI:45.3; 47.8] men and 48.4 [95%CI:47.4; 49.3] women Wald Test $p<0.001$; wave 3: 44.6 [95%CI:43.4; 45.9] and 46.5 [95%CI:45.5; 47.5] women Wald Test $p<0.001$).

Increasing age, not cohabiting (vs cohabiting), low educational attainment (vs medium and high), 4th quintile of wealth (vs richest wealth quintile) and smoking (vs ex-smoker and never smoked) were not significantly related to quality of life. About 60% of the unexplained variability in quality of life was due to differences within individuals over time (within variance 30.6).

Results from the analysis based on observed data were not all consistent with those based on multiple imputations. In general the coefficients obtained from the observed data were larger than those obtained from imputed data sets, showing stronger relationships. Standard errors were smaller for parameters obtained from imputed data sets. In the model of people with CHD obtained from the observed data, wave 2 and the interaction term between wave 3 and gender was statistically significant; the quadratic effect of age, retirement status, limitations with ADLs and number of close friends were not significant while education was significant. The coefficient from smoking was positive in the results from the observed data and negative in the imputed results, although non-significant in both.

Table 4.4 Linear random intercept model for quality of life with time for the CHD group

	Imputed			Observed		
	Coef.	Std. Err.	P-value	Coef.	Std. Err.	P-value
Gender	0.83	0.50	0.094	0.54	0.77	0.489
Wave 2	-0.84	0.36	0.020	-0.85	0.65	0.192
Wave 3	-2.73	0.37	0.000	-3.58	0.73	0.000
Wave 2*Gender	0.98	0.53	0.065	1.18	1.01	0.243
Wave 3*Gender	0.98	0.53	0.066	2.56	1.10	0.020
Age (centred 63)	0.04	0.03	0.248	0.00	0.05	0.941
Age^2	-0.01	0.00	0.000	0.00	0.00	0.315
Not cohabiting	-0.60	0.37	0.102	-0.83	0.64	0.194
Completely retired	-1.28	0.44	0.004	-0.52	0.76	0.491
Other*	-2.30	0.53	0.000	-2.88	0.95	0.002
Low education	-0.47	0.32	0.146	-0.62	0.31	0.047
Wealth 4^{th}	-0.36	0.47	0.440	-0.12	0.80	0.886
Wealth 3^{rd}	-2.28	0.49	0.000	-2.30	0.83	0.006
Wealth 2^{nd}	-2.21	0.51	0.000	-3.45	0.89	0.000
Wealth Poorest	-3.61	0.53	0.000	-3.70	0.95	0.000
Current smoker	-0.27	0.49	0.587	1.41	0.80	0.078
Physically inactive	-1.37	0.34	0.000	-2.02	0.56	0.000
Drinks alcohol ≥3 days a week	0.76	0.33	0.023	1.58	0.60	0.008
Often troubled with pain	-1.84	0.31	0.000	-2.14	0.57	0.000
ADLs	-1.46	0.12	0.000	-0.07	0.30	0.819
Positive support (centred at 27)	0.29	0.03	0.000	0.29	0.05	0.000
Close friends (centred at 8)	0.10	0.02	0.000	0.06	0.04	0.150
Depressive symptoms	-4.32	0.33	0.000	-5.33	0.63	0.000
Constant	47.38	0.64	0.000	47.39	0.99	0.000
Model						
Between variance	20.65	0.04		19.98	3.23	
Within variance	30.62	0.02		29.92	2.53	

Reference categories: Well, Male, Cohabiting, In paid employment High/medium education, Richest wealth, Non-smoker, Physically active, Drinks <3 days a week, Not troubled with pain, No depressive symptoms.

* Permanently unable to work, not currently in paid employment, looking after home or family

Table 4.5 shows the results of the random intercept model for quality of life with an interaction term between time and gender for the Well group. Results are also summarised graphically in Figure 4.1. Among men and women in the Well group, quality of life decreased significantly between baseline and years two and four of follow-up. The interaction terms between wave and gender were not significant suggesting no difference by gender in the rate of change over time in quality of life. However, at two year follow-up women had higher quality of life compared to men (47.9 [95%CI: 47.4; 48.4] men and 48.4 [95%CI: 48.0; 48.9] women $p<0.05$]. The quadratic effect of age and low educational attainment were not significantly related to quality of life. With the exception of alcohol consumption, positive support and close friends and family, all the other covariates were negatively related to quality of life. About 57% of the total unexplained variability in quality of life was due to differences within individuals over time.

The main differences between the results based on augmented data and those based on observed data were that the interaction term between wave 2 and gender was negative in the results based on observed data and positive in the results based on the imputed data (although not significant in both models). Also age, cohabiting status, alcohol consumption and limitations with ADLs were not significant (Table 4.5), while the quadratic effect of age was significant.

Table 4.5 Linear random intercept model for quality of life with time for the Well group

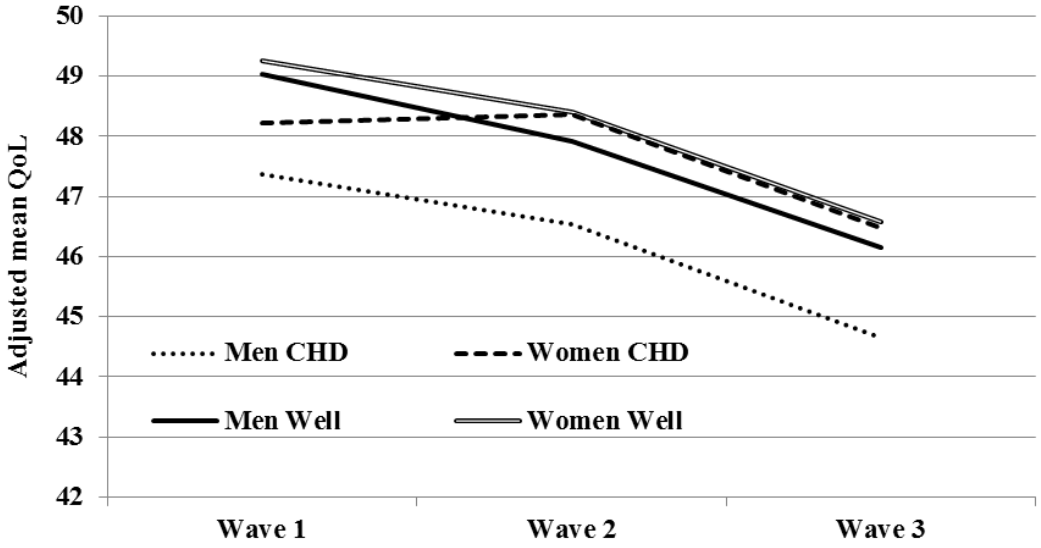
	Imputed			Observed		
	Coef.	**Std. Err.**	**P-value**	**Coef.**	**Std. Err.**	**P-value**
Gender	0.24	0.23	0.289	0.37	0.30	0.213
Wave 2	-1.11	0.18	0.000	-1.20	0.27	0.000
Wave 3	-2.88	0.18	0.000	-3.02	0.29	0.000
Wave 2*Gender	0.25	0.24	0.299	-0.03	0.36	0.929
Wave 3*Gender	0.18	0.24	0.443	0.05	0.38	0.896
Age (centred 63)	0.04	0.01	0.009	0.03	0.02	0.096
Age^2	0.00	0.00	0.212	-0.01	0.00	0.000
Not cohabiting	-0.35	0.18	0.045	-0.15	0.27	0.577
Completely retired	-0.73	0.17	0.000	-0.58	0.25	0.019
Other*	-2.05	0.22	0.000	-2.33	0.34	0.000
Low education	0.15	0.16	0.335	0.18	0.12	0.129
Wealth 4^{th}	-1.03	0.19	0.000	-0.93	0.27	0.000
Wealth 3^{rd}	-1.58	0.20	0.000	-1.33	0.29	0.000
Wealth 2^{nd}	-2.34	0.21	0.000	-2.31	0.32	0.000
Wealth Poorest	-3.82	0.25	0.000	-3.85	0.39	0.000
Current smoker	-0.73	0.21	0.001	-0.79	0.30	0.008
Physically inactive	-0.96	0.13	0.000	-0.81	0.19	0.000
Drinks alcohol ≥3 days a week	0.60	0.15	0.000	0.35	0.21	0.099
Often troubled with pain	-1.87	0.17	0.000	-2.18	0.25	0.000
ADLs	-1.68	0.09	0.000	0.03	0.11	0.790
Depressive symptoms	-4.19	0.17	0.000	-4.65	0.27	0.000
Positive support (centred at 27)	0.27	0.01	0.000	0.31	0.02	0.000
Close friends (centred at 8)	0.06	0.01	0.000	0.04	0.01	0.004
Constant	49.03	0.25	0.000	49.11	0.36	0.000
Model						
Between variance	19.17	0.02		20.90	1.07	
Within variance	24.95	0.01		20.68	0.67	

Reference categories: Well, Male, Cohabiting, In paid employment High/medium education, Richest wealth, Non-smoker, Physically active, Drinks <3 days a week, Not troubled with pain, No depressive symptoms.

* Permanently unable to work, not currently in paid employment, looking after home or family

In general the gender specific trajectories found in people with CHD were not the same as those found in the Well group. Figure 4.1 shows graphically the adjusted mean quality of life over time among people in the CHD and Well group (obtained from the imputed data), by gender. Women with CHD had stable quality of life between baseline and two year follow-up (wave 2), which then decreased at four year follow-up (wave 3) to reach a similar quality of life to women in the Well group and a higher quality of life than men with CHD. The quality of life of men with CHD decreased significantly over time. Men and women in the Well group had similar trajectories of quality of life and only at two year follow-up (wave 2) the quality of life of women was significantly higher than the quality of life of men, although the difference was small compared to the difference in quality of life between men and women with CHD.

Figure 4.1 Trajectories over time of quality of life among people with CHD and the Well group, by gender (imputed data)



Results adjusted for gender, age, quadratic effect of age, cohabiting status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends and family.

In order to understand the effect of covariates on the estimates of trajectories of quality of life, five linear random intercept models with an interaction term between gender and time were run by sequentially expanding the set of covariates as follows: model 1 is a model with interaction term between gender and time plus age and its quadratic effect (model 1); model 2 is model 1 plus socio-demographic factors (cohabiting status, education and wealth); model 3 is model 2 plus health behaviour factors (physical

activity, smoking, alcohol consumption); model 4 is model 3 plus health factors (pain and ADLs); model 5 is model 4 plus psychosocial factors (depressive symptoms, social support and social networks). Results are presented in Tables 4.6 and 4.7.

In models adjusted for age (model 1), socio-demographic characteristics (model 2), and health behaviours (model 3) quality of life for men with CHD was not significantly lower at two year follow-up (wave 2) compared to baseline (wave 1). However, after adjusting for psychosocial (model 4) and health factors (model 5), the difference in quality of life between baseline and two year follow-up (wave 2) increased in magnitude and became significant (Table 4.6). The factor that contributed to the increase in the difference in quality of life between baseline and two year follow-up for men was pain. After controlling for the effect of this factor, the quality of life of men decreased between baseline and two year follow-up. If this factor was not accounted for in the model, then we would have concluded that the quality of life of men was not decreasing significantly.

At four year follow-up (wave 3), quality of life of men was significantly lower than baseline, independent of other covariates. However, the magnitude of the difference in quality of life between baseline and four year follow-up (wave 3) increased from -1.55 (model 1) in the model adjusted for age only to -2.73 in the fully adjusted model (model 5).

The interaction terms between wave 2 and gender and wave 3 and gender were statistically significant in models 1 to 4 and models 2 to 4 respectively. When psychosocial factors were accounted for (model 5), these interaction terms were no longer significant. The psychosocial factor that contributed to the attenuation of the differences in the rate of change of quality of life between men and women at wave 2 and wave 3 was depressive symptoms.

Among people from the Well group, covariate adjustment did not change the conclusions about their trajectories of quality of life (Table 4.7). The only exception was that adjustment for psychosocial factors (model 5) attenuated the gender difference in baseline quality of life found in models 1 to 4. In particular, after adjusting for positive social support, women no longer had higher quality of life than men at baseline (models 1 to 4).

Table 4.6 Sequentially adjusted linear random intercept model for quality of life with time for the CHD group, imputed data

	Model 1		Model2		Model3		Model4		Model5	
	b	**se**	**b**	**se**	**b**	**se**	**b**	**se**	**b**	**se**
Gender	-0.63	0.62	0.19	0.59	0.40	0.59	0.78	0.55	0.83	0.50
Wave 2	-0.21	0.36	-0.62	0.36	-0.73	0.36	-1.12**	0.36	-0.84*	0.36
Wave 3	-1.55***	0.37	-2.08***	0.38	-2.18***	0.38	-2.54***	0.37	-2.73***	0.37
Wave 2*Gender	1.30*	0.54	1.50*	0.54	1.59**	0.54	1.29*	0.53	0.98	0.53
Wave 3*Gender	1.04	0.54	1.31*	0.54	1.44*	0.54	1.12*	0.53	0.99	0.53
Age (centred 63)	0.03	0.04	0.06	0.04	0.07	0.04	0.07*	0.04	0.04	0.03
Age^2	-0.01***	0.00	-0.01***	0.00	-0.01***	0.00	-0.01***	0.00	-0.01***	0.00
Not cohabiting			-1.43**	0.42	-1.38**	0.42	-1.20**	0.40	-0.61	0.37
Completely retired			-1.51**	0.48	-1.56**	0.48	-1.30**	0.47	-1.28**	0.44
Other			-2.94***	0.57	-2.98***	0.57	-2.35***	0.56	-2.30***	0.53
Low education			-0.79*	0.37	-0.62	0.37	-0.74*	0.35	-0.47	0.32
Wealth 4^{th}			-0.44	0.50	-0.33	0.50	-0.08	0.49	-0.36	0.47
Wealth 3^{rd}			-2.74***	0.53	-2.54***	0.53	-2.20***	0.51	-2.28***	0.49
Wealth 2^{nd}			-2.96***	0.55	-2.66***	0.55	-2.22***	0.54	-2.21***	0.51
Wealth Poorest			-4.79***	0.58	-4.49***	0.59	-3.83***	0.56	-3.61***	0.53
Current smoker					-0.06	0.58	-0.32	0.55	-0.27	0.49
Physically inactive					-1.99***	0.36	-1.55***	0.35	-1.38***	0.34
Drinks alcohol ≥3 days a week					0.84*	0.37	0.73*	0.36	0.76*	0.33
Often troubled with pain							-1.98***	0.33	-1.84***	0.31
ADLs							-1.63***	0.13	-1.46***	0.12
Positive support (centred at 27)									0.29***	0.03
Close friends (centred at 8)									0.10***	0.02
Depressive symptoms									-4.32***	0.33
Constant	41.22***	0.43	45.44***	0.65	46.19***	0.72	46.92***	0.69	47.38***	0.64
Model										
Between variance	51.62	1.06	43.12	1.06	40.77	1.06	32.66	1.07	20.66	1.08
Within variance	32.07	1.03	31.56	1.03	31.63	1.03	30.69	1.03	30.63	1.03

* for p<.05, ** for p<.01, and ***for p<.001

Table 4.7 Sequentially adjusted linear random intercept model for quality of life with time for the Well group, imputed data

	Model 1		Model2		Model3		Model4		Model5	
	b	**se**	**b**	**se**	**b**	**se**	**b**	**se**	**b**	**se**
Gender	0.15	0.27	0.58*	0.26	0.74**	0.25	0.64**	0.24	0.24	0.23
Wave 2	-1.21***	0.18	-1.02***	0.18	-1.15***	0.18	-1.16***	0.18	-1.11***	0.18
Wave 3	-2.61***	0.19	-2.60***	0.19	-2.71***	0.19	-2.62***	0.18	-2.88***	0.18
Wave 2*Gender	-0.03	0.25	0.14	0.24	0.22	0.24	0.30	0.24	0.25	0.24
Wave 3*Gender	-0.26	0.25	-0.02	0.24	0.06	0.24	0.09	0.24	0.18	0.24
Age (centred 63)	0.02	0.01	0.06***	0.02	0.07***	0.02	0.06***	0.01	0.04**	0.01
Age^2	-0.01*	0.00	-0.01	0.00	-0.01	0.00	-0.01	0.00	-0.01	0.00
Not cohabiting			-1.02***	0.20	-1.01***	0.19	-0.91***	0.19	-0.36*	0.18
Completely retired			-1.02***	0.18	-1.05***	0.18	-0.74***	0.18	-0.73***	0.17
Other			-2.90***	0.24	-2.88***	0.24	-2.10***	0.23	-2.05***	0.22
Low education			-0.06	0.18	0.12	0.18	0.06	0.17	0.15	0.16
Wealth 4^{th}			-1.04***	0.20	-0.92***	0.20	-0.91***	0.20	-1.03***	0.19
Wealth 3^{rd}			-1.82***	0.22	-1.63***	0.22	-1.54***	0.21	-1.58***	0.20
Wealth 2^{nd}			-2.83***	0.23	-2.52***	0.23	-2.31***	0.23	-2.34***	0.21
Wealth Poorest			-4.61***	0.27	-4.20***	0.27	-3.92***	0.26	-3.82***	0.25
Current smoker					-0.95***	0.24	-0.96***	0.23	-0.73**	0.21
Physically inactive					-1.42***	0.14	-1.19***	0.14	-0.96***	0.13
Drinks alcohol ≥3 days a week					0.56***	0.16	0.50**	0.15	0.60***	0.15
Often troubled with pain							-2.07***	0.17	-1.87***	0.17
ADLs							-1.87***	0.10	-1.68***	0.09
Positive support (centred at 27)									0.28***	0.01
Close friends (centred at 8)									0.06***	0.01
Depressive symptoms									-4.19***	0.17
Constant	45.73	0.21	48.11	0.26	48.55	0.28	48.64	0.27	49.03	0.25
Model										
Between variance	36.74	1.03	32.07	1.03	30.69	1.03	27.61	1.03	19.18	1.04
Within variance	26.95	1.02	26.26	1.02	26.21	1.02	25.28	1.02	24.93	1.02

* for p<.05, ** for p<.01, and ***for p<.001

4.3.3 Depressive symptoms

The results of the logistic random intercept model for depressive symptoms are reported in Table 4.8. Overall, women in the CHD group had higher odds for depressive symptoms than men with CHD (adjusted OR: 1.65, $p<0.01$) independent of other covariates. Similarly, in the Well group women were more likely than men to have depressive symptoms (adjusted OR: 2.2, $p<0.001$). The adjusted odds of having depressive symptoms was 1.40 times higher in men with CHD than men in Well group ($p<0.01$), while women with CHD were as likely as women in the Well group to have depressive symptoms. The interaction term between gender and disease status was not significant, mainly due to the fact that women in the CHD group are as likely to have depressive symptoms as women in the Well group.

Being retired as compared to being in paid employment, being in the 4^{th}, 3^{rd} and 2^{nd} wealth quintiles compared to the richest wealth quintile, drinking on three days a or more a week compared to drinking on less than three days a week, and increasing number of close friends and family were not significantly related to depressive symptoms, while all the other covariates were. The between-individual variance is 5.38 (62% of the total variance), which expresses the unexplained variation in depressive symptoms due to differences between individuals after controlling for covariates. The within-individual variance is fixed at 3.29 (38% of the total variance) and expresses the variation in depressive symptoms within individuals over time due to unmeasured differences after controlling for covariates. We can conclude that a greater proportion of the residual variability in depressive symptoms is due to differences between individuals.

Results based on observed data show a similar pattern to the results based on values for missing data. There are however few exceptions: increasing age, being in the “other” category of retirement status (vs in paid employment), poorest wealth quintile (vs richest), current smoker (vs ex-smoker and never smoked) and limitations with ADLs were not significantly related to depressive symptoms (Table 4.8).

Table 4.8 Logistic random intercept model for depressive symptoms

	Imputed			Observed		
	OR	**Std. Err.**	**P-value**	**OR**	**Std. Err.**	**P-value**
CHD	1.40	0.19	0.015	1.79	0.42	0.013
Gender	2.18	0.21	0.000	1.96	0.30	0.000
CHD*Gender	0.76	0.15	0.146	0.82	0.27	0.549
Age (centred 63)	1.01	0.00	0.006	1.00	0.01	0.975
Not cohabiting	2.04	0.17	0.000	2.56	0.37	0.000
Completely retired	0.92	0.08	0.388	0.86	0.13	0.332
Other*	1.26	0.14	0.040	1.37	0.26	0.099
Low education	1.31	0.10	0.000	1.16	0.08	0.027
Wealth 4th	0.86	0.09	0.157	0.77	0.14	0.146
Wealth 3rd	1.12	0.12	0.302	1.25	0.22	0.218
Wealth 2nd	1.23	0.14	0.064	1.25	0.24	0.252
Wealth Poorest	1.54	0.19	0.000	1.43	0.31	0.093
Current smoker	1.53	0.15	0.000	1.25	0.20	0.174
Physically inactive	1.87	0.14	0.000	1.90	0.24	0.000
Drinks alcohol ≥3 days a week	1.00	0.08	0.977	0.91	0.12	0.488
Often troubled with pain	2.04	0.16	0.000	2.81	0.39	0.000
ADLs	1.57	0.06	0.000	1.04	0.07	0.620
Positive support (centred at 27)	0.93	0.01	0.000	0.94	0.01	0.000
Close friends (centred at 8)	0.99	0.01	0.155	0.99	0.01	0.249
Model						
Between variance	5.38	0.06		3.57	0.55	
Within variance	3.29			3.29		

Reference categories: Well, Male, Cohabiting, In paid employment High/medium education, Richest wealth, Non-smoker, Physically active, Drinks <3 days a week, Not troubled with pain, No limitations with ADLs.
* Permanently unable to work, not currently in paid employment, looking after home or family

Table 4.9 reports results from the logistic random intercept model with a time and gender interaction among people with CHD. Among men with CHD the odds of having depressive symptoms did not differ significantly at years two and four of follow-up (OR: 1.20 [95%CI: 0.71; 1.8] and OR: 1.06 [95%CI: 0.70; 1.58] respectively) compared to baseline (wave 1). At two year follow-up (wave 2) the risk of having depressive symptoms did not differ significantly from baseline (wave 1) (OR: 1.06 [95%CI: 0.70; 1.61]), while at four year follow-up (wave 3) they were significantly less likely to have depressive symptoms (OR: 0.63[95%CI: 0.40; 0.97], $p<0.05$) than at baseline. Women with CHD were more likely than men to have depressive symptoms at baseline

(p<0.001), and at two year follow-up (OR 1.8 [95%CI: 1.1; 2.9] p<0.05) but there was no gender difference at four year follow-up (OR 1.2 [95%CI: 0.7; 1.9] p=0.517).

Only not cohabiting, physical inactivity, pain, limitation with ADLs and positive support were significantly related to depressive symptoms.

The results obtained from the sample with missing data (observed) were very similar to those obtained from analyses based on augmented samples. However, standard errors obtained from analyses based on the augmented samples were smaller. The only difference found between observed data results and augmented samples results was that limitation with ADLs was not significantly related to depressive symptoms.

Table 4.9 Logistic random intercept model for depressive symptoms with time for the CHD group

	Imputed			Observed		
	OR	Std. Err.	P-value	OR	Std. Err.	P-value
Gender	1.99	0.48	0.004	2.40	0.99	0.035
Wave 2	1.20	0.24	0.359	1.48	0.57	0.312
Wave 3	1.06	0.22	0.779	0.67	0.32	0.397
Wave 2*Gender	0.88	0.25	0.665	0.55	0.31	0.290
Wave 3*Gender	0.59	0.17	0.071	0.82	0.53	0.766
Age (centred 63)	1.01	0.01	0.168	0.99	0.02	0.555
Not cohabiting	2.17	0.38	0.000	2.62	0.89	0.004
Completely retired	0.88	0.21	0.579	1.03	0.43	0.935
Other*	1.22	0.33	0.474	1.03	0.53	0.948
Low education	1.27	0.20	0.132	1.24	0.21	0.192
Wealth 4th	0.79	0.21	0.374	0.52	0.27	0.212
Wealth 3rd	1.23	0.32	0.423	1.67	0.78	0.274
Wealth 2nd	1.33	0.35	0.282	2.30	1.14	0.093
Wealth Poorest	1.50	0.41	0.132	1.57	0.80	0.377
Current smoker	1.17	0.27	0.498	0.60	0.26	0.235
Physically inactive	2.16	0.41	0.000	2.67	0.92	0.004
Drinks alcohol ≥3 days a week	0.80	0.14	0.198	0.93	0.32	0.839
Often troubled with pain	2.11	0.33	0.000	2.49	0.76	0.003
ADLs	1.53	0.09	0.000	0.74	0.14	0.105
Positive support (centred at 27)	0.92	0.01	0.000	0.90	0.03	0.000
Close friends (centred at 8)	0.99	0.01	0.361	1.00	0.02	0.887
Model						
Between variance	4.38	0.12		3.47	1.41	
Within variance	3.29			3.29		

Reference categories: Well, Male, Cohabiting, In paid employment High/medium education, Richest wealth, Non-smoker, Physically active, Drinks <3 days a week, Not troubled with pain, No limitations with ADLs.
*Permanently unable to work, not currently in paid employment, looking after home or family.

Similar results to those of men in the CHD group were found among men in the Well group (Table 4.10): there was not a significant difference over time in the odds of having depressive symptoms. Women from the Well group were more likely to have depressive symptoms than men at each time point (wave 1 OR: 2.31 p<0.001; wave 2 OR: 1.8 p<0.001; wave 3 OR: 2.6 p<0.001).

Observed sample results and augmented samples results were slightly different, in the former results increasing age, low education (vs medium and high) and poorest wealth quintile (vs richest) were not significantly related to depressive symptoms.

Table 4.10 Logistic random intercept model for depressive symptoms with time for the Well group

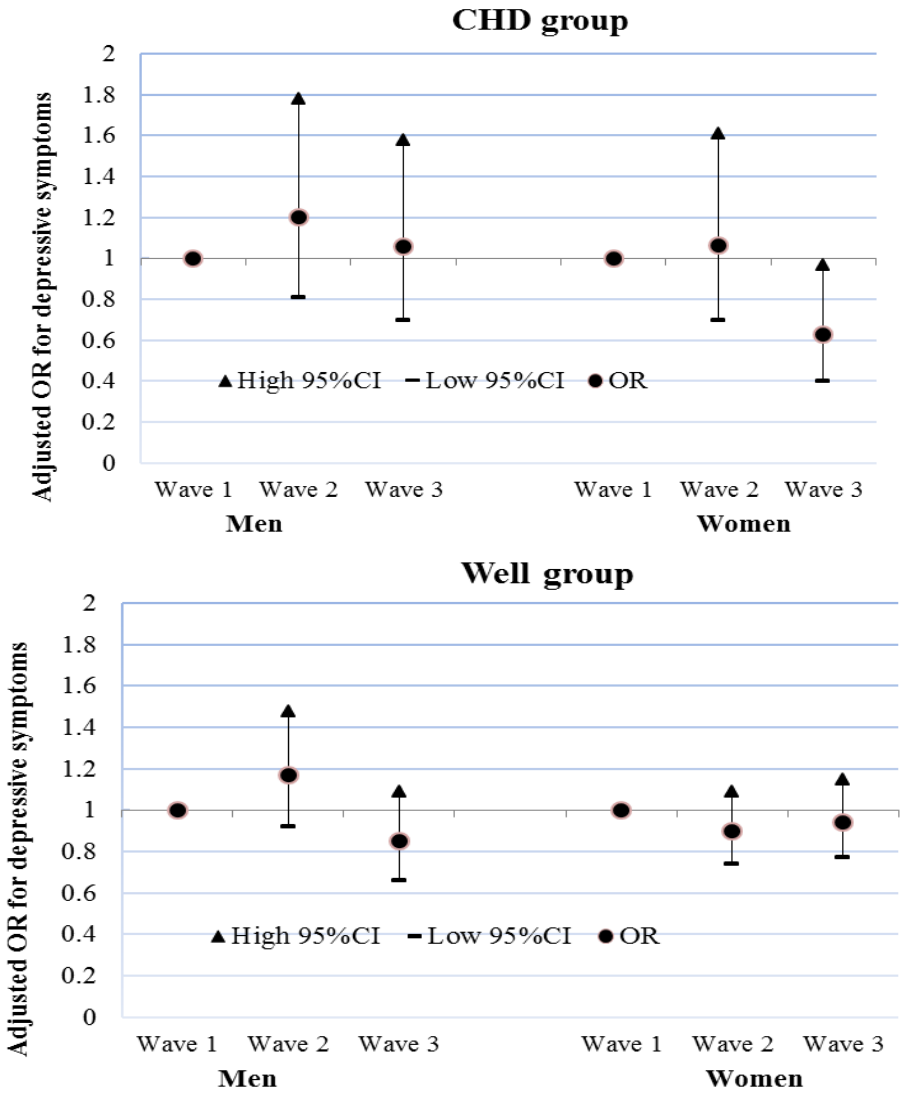
	Imputed			Observed		
	OR	Std. Err.	P-value	OR	Std. Err.	P-value
Gender	2.31	0.30	0.000	1.78	0.35	0.004
Wave 2	1.17	0.14	0.212	0.97	0.21	0.891
Wave 3	0.85	0.11	0.203	0.80	0.19	0.343
Wave 2*Gender	0.77	0.12	0.091	1.08	0.29	0.786
Wave 3*Gender	1.11	0.18	0.519	1.29	0.38	0.394
Age (centred 63)	1.01	0.01	0.010	1.00	0.01	0.725
Not cohabiting	2.04	0.19	0.000	2.60	0.42	0.000
Completely retired	0.93	0.09	0.472	0.80	0.13	0.185
Other*	1.27	0.16	0.058	1.41	0.30	0.105
Low education	1.30	0.11	0.002	1.14	0.08	0.085
Wealth 4th	0.88	0.11	0.296	0.81	0.16	0.273
Wealth 3rd	1.09	0.13	0.475	1.17	0.23	0.430
Wealth 2nd	1.21	0.15	0.138	1.07	0.23	0.733
Wealth Poorest	1.54	0.21	0.002	1.45	0.35	0.125
Current smoker	1.62	0.18	0.000	1.42	0.25	0.049
Physically inactive	1.83	0.15	0.000	1.83	0.25	0.000
Drinks alcohol ≥3 days a week	1.06	0.09	0.531	0.93	0.13	0.622
Often troubled with pain	1.98	0.18	0.000	2.88	0.46	0.000
ADLs	1.60	0.08	0.000	1.10	0.08	0.226
Positive support (centred at 27)	0.93	0.01	0.000	0.95	0.01	0.000
Close friends (centred at 8)	0.99	0.01	0.322	0.99	0.01	0.288
Model						
Between variance	5.38	0.07		3.62	0.61	
Within variance	3.29			3.29		

Reference categories: Well, Male, Cohabiting, In paid employment High/medium education, Richest wealth, Non-smoker, Physically active, Drinks <3 days a week, Not troubled with pain, No limitations with ADLs. *Permanently unable to work, not currently in paid employment, looking after home or family

Trajectories of depressive symptoms reported in Table 4.10 are presented graphically in Figure 4.2 (imputed data). Men with CHD and men from the Well group did not report

significant changes over time in the odds of having depressive symptoms. Women with CHD were less likely to have depressive symptoms at four year follow-up (wave 3) than baseline, while women from the Well group did not report significant changes over time in their odds of having depressive symptoms.

Figure 4.2 Trajectories over time of depressive symptoms among people with CHD and the Well group, by gender (imputed data)



Results adjusted for gender, age, cohabiting status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends and family.

In order to understand the effect of covariates on the estimates of trajectories of depressive symptoms, five logistic random intercept models with an interaction term between gender and time were run by sequentially expanding the set of covariates as follows: model 1 is a model with interaction term between gender and time plus age and

its quadratic effect (model 1); model 2 is model 1 plus socio-demographic factors (cohabiting status, education and wealth); model 3 is model 2 plus health behaviour factors (physical activity, smoking, alcohol consumption); model 4 is model 3 plus health factors (pain and limitation with ADLs); model 5 is model 4 plus social factors (social support and social networks). Results are presented in Tables 4.11 and 4.12. Results are presented as log odds.

Table 4.11 reports the sequentially adjusted random intercept models for depressive symptoms among people with CHD. At baseline, the odds ratio for having depressive symptoms in women compared to men was 2.8 (log odds 1.04) in the model adjusted for age only (model 1). After full adjustment the odds ratio decreased to 1.99 (log odd 0.68), but was still significant. Adjustment for covariates did not change the conclusions about gender differences in depressive symptoms at baseline and two year follow-up, but did change the conclusions at four year follow-up. It was found that women had higher odds of having depressive symptoms than men at four year follow-up (OR: 1.9 [95%CI: 1.1; 3.2]; $p<0.05$) only in a model adjusted for age (Model 1), when further adjustment was made to the model (Models 2 to 5) the gender differences was no longer significant.

For people in the Well group, adjustment did not change the conclusions of trajectories about depressive symptoms.

Table 4.11 Sequentially adjusted random intercept model for depressive symptoms with time for the CHD group, imputed data

	Model 1		Model2		Model3		Model4		Model5	
	b	**se**	**b**	**se**	**b**	**se**	**b**	**se**	**b**	**se**
Gender	1.04***	0.27	0.76**	0.26	0.69**	0.25	0.57*	0.24	0.69**	0.24
Wave 2	-0.05	0.19	0.06	0.20	0.09	0.20	0.24	0.20	0.18	0.20
Wave 3	-0.32	0.20	-0.15	0.20	-0.14	0.20	0.03	0.21	0.06	0.21
Wave 2*Gender	-0.17	0.28	-0.24	0.28	-0.27	0.28	-0.19	0.28	-0.12	0.28
Wave 3*Gender	-0.40	0.29	-0.54	0.29	-0.58*	0.29	-0.51	0.29	-0.52	0.29
Age (centred 63)	0.04***	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Not cohabiting			0.88***	0.19	0.85***	0.18	0.82***	0.18	0.78***	0.18
Completely retired			0.07	0.25	0.07	0.25	-0.08	0.24	-0.13	0.24
Other			0.56*	0.28	0.51	0.28	0.23	0.27	0.20	0.27
Low education			0.29	0.17	0.19	0.17	0.23	0.16	0.24	0.16
Wealth 4th			-0.12	0.28	-0.15	0.28	-0.24	0.27	-0.24	0.27
Wealth 3rd			0.46	0.27	0.36	0.27	0.21	0.26	0.21	0.26
Wealth 2nd			0.66*	0.27	0.51	0.28	0.32	0.27	0.29	0.27
Wealth Poorest			0.93***	0.28	0.76**	0.28	0.49	0.27	0.41	0.27
Current smoker					0.17	0.25	0.17	0.24	0.16	0.23
Physically inactive					1.02***	0.20	0.79***	0.19	0.77***	0.19
Drinks alcohol ≥3 days a week					-0.24	0.18	-0.19	0.17	-0.22	0.17
Often troubled with pain							0.71***	0.16	0.75***	0.16
ADLs							0.43***	0.06	0.43***	0.06
Positive support (centred at 27)									-0.09***	0.02
Close friends (centred at 8)									-0.01	0.01
Constant	-2.34***	0.21	-3.16***	0.33	-3.61***	0.37	-3.64***	0.36	-3.62***	0.36
Model										
Between variance	2.33	1.07	2.10	1.07	2.02	1.07	1.79	1.08	1.72	1.08
Within variance	3.29		3.29		3.29		3.29		3.29	

* for p<.05, ** for p<.01, and ***for p<.001. Estimates are expressed in log odds

Table 4.12 Sequentially adjusted random intercept model for depressive symptoms with time for the Well group, imputed data

	Model 1		Model2		Model3		Model4		Model5	
	b	**se**	**b**	**se**	**b**	**se**	**b**	**se**	**b**	**se**
Gender	0.87***	0.14	0.71***	0.13	0.66***	0.13	0.67***	0.13	0.84***	0.13
Wave 2	0.12	0.12	0.13	0.12	0.17	0.12	0.17	0.12	0.15	0.12
Wave 3	-0.26*	0.13	-0.20	0.13	-0.17	0.13	-0.20	0.13	-0.16	0.13
Wave 2*Gender	-0.16	0.16	-0.24	0.16	-0.25	0.16	-0.26	0.16	-0.26	0.16
Wave 3*Gender	0.20	0.16	0.11	0.16	0.09	0.16	0.11	0.16	0.10	0.16
Age (centred 63)	0.03***	0.00	0.02**	0.01	0.01*	0.01	0.009	0.01	0.01**	0.01
Not cohabiting			0.83***	0.10	0.82***	0.10	0.78***	0.10	0.71***	0.09
Completely retired			0.01	0.10	0.03	0.10	-0.06	0.10	-0.07	0.10
Other			0.53***	0.13	0.50***	0.12	0.26*	0.13	0.24	0.12
Low education			0.29**	0.09	0.23*	0.09	0.25**	0.09	0.26**	0.09
Wealth 4th			-0.09	0.12	-0.14	0.12	-0.14	0.12	-0.13	0.12
Wealth 3rd			0.16	0.13	0.09	0.13	0.07	0.12	0.09	0.12
Wealth 2nd			0.37**	0.13	0.24	0.13	0.16	0.13	0.19	0.13
Wealth Poorest			0.71***	0.14	0.51***	0.14	0.42**	0.14	0.43**	0.14
Current smoker					0.51***	0.12	0.51***	0.11	0.48***	0.11
Physically inactive					0.76***	0.08	0.64***	0.08	0.60***	0.08
Drinks alcohol ≥3 days a week					0.02	0.09	0.06	0.09	0.05	0.09
Often troubled with pain							0.69***	0.09	0.68***	0.09
ADLs							0.47***	0.05	0.47***	0.05
Positive support (centred at 27)									-0.07***	0.01
Close friends (centred at 8)									-0.01	0.01
Constant	-2.97***	0.12	-3.44***	0.15	-3.85***	0.17	-3.84***	0.17	-3.86***	0.16
Model										
Between variance	2.09	0.04	1.94	0.04	1.88	0.04	1.76	0.04	1.68	0.04
Within variance	3.29		3.29		3.29		3.29		3.29	

* for p<.05, ** for p<.01, and ***for p<.001. Estimates are expressed in log odds

4.4 Discussion

This chapter set out to explore longitudinally gender-specific trajectories of quality of life and depressive symptoms among people with CHD, and compare the results with those from a Well group.

It was hypothesised that people aged over fifty years who had had experienced CHD would be at higher risk of experiencing depressive symptoms and poor quality of life than those who have not. Results suggested that men and women with CHD reported lower levels of quality of life compared to people in the Well group. Men with CHD were more likely to have depressive symptoms than men in the Well group. The hypothesis was not supported for women: it was found that women with CHD were equally likely to have depressive symptoms women from the Well group. A possible explanation might be that women are in general more prone to depression irrespective of the disease status (Forrester et al., 1992; Freasure-Smith et al., 1999; Mallik et al., 2006).

Gender differences in quality of life and depressive symptoms among people with CHD

It was hypothesised that women with CHD would report lower levels of quality of life and be at higher risk of having depressive symptoms compared to men. It was found that women with CHD reported the same quality of life as men at baseline (wave 1, 2002-03), and higher quality of life at years two (wave 2, 2004-05) and four (wave 3, 2005-07) of follow-up independent of other covariates. No previous studies have reported these findings. A possible explanation might lie in the measure of quality of life used in this study, the CASP-19, which is neither disease-specific nor health-related. Previous studies showing lower quality of life among women used disease-specific measures, such as Quality of Life After Myocardial Infarction (Hillers et al, 1994) and Quality of Life Index-Cardiac (Ferrans and Powers, 1985; 1992), and a more generic measure of health-related quality of life, the SF-36 (Turner-Bowker et al., 2002). It has been shown that the problem related to these measures is that when they are applied to a cardiac population, they usually lack sensitivity – the ability of a measure to detect important changes (Smith et al., 2000). Also health-related measures of quality of life aim to measure quality of life through the use of health domains such as mental health,

physical functioning, body pain and health perception. These are all dimensions that might influence the subjective experience of quality of life of individuals. It has been shown that women with CHD are at higher risk of having depression (Naqvi et al., 2005) and have lower physical functioning than men (Richardson, 2003). Therefore the lower levels of health-related quality of life among women reported in the literature might be attributed to higher depression and lower physical functioning.

As for depressive symptoms, women with CHD reported higher odds of having depressive symptoms than men at baseline and two year follow-up but not at four year follow-up. It was when the model was adjusted for covariates other than age that the gender difference disappeared. This result is consistent with one previous study that showed gender differences in the mental health dimension of HRQOL at one year post-myocardial infarction did not persist once the model was adjusted for demographic, clinical, comorbid and psychosocial covariates (Norris et al., 2007).

One of the possible explanations for not finding any gender difference in depressive symptoms at four year follow-up might lie in the measure of CHD used in this thesis which included myocardial infarction as well as angina symptoms. It is possible that the long term effects of angina lead to fewer impairments than those of myocardial infarction. Angina is not a disease but a symptom of coronary heart disease which occurs on exertion and is relieved by rest. On the other hand, myocardial infarction is an event caused by a blood clot that blocks one of the coronary arteries, whereby the part of the heart muscle is damaged and might die. Women in this sample experienced more angina symptoms than myocardial infarction (57% of women with CHD experienced angina, 28% myocardial infarction and 15% both); it is possible that the gender difference in depressive symptoms attenuated after four years as a result of adaptation to the coronary event. Mallik et al. (2006) suggested that gender differences in depressive symptoms among people with CHD no longer persist at older ages. They found significant gender differences in depression up to the age of 60 among men and women who had experienced myocardial infarction, but did not find any gender difference in depression among older women (aged 60 and over) compared to older men.

It was also hypothesised that shapes of trajectories over time of quality of life and depressive symptoms would be different in men and women with CHD. This study supported this hypothesis. It was found that, while the quality of life of men with CHD

decreased significantly over time, the quality of life of women with CHD was stable at two years after baseline, followed by a decline between baseline and four year follow-up.

Trajectories in depressive symptoms also differed by gender. The odds of having depressive symptoms did not change significantly over time for men. Among women with CHD the odds of having depressive symptoms at two year follow-up were the same as at baseline but at four year follow-up women were significantly less likely to report depressive symptoms than baseline.

The results of stable levels of quality of life at two year follow-up and of improvement in depressive symptoms among women with CHD at four year follow-up compared to baseline suggest that the adaptation to the CHD event might differ by gender. A possible explanation is that women might be able to recover quickly to the event in terms of quality of life and mental health, while men might not be coping well with the long-term consequences of CHD. It has been shown that women have, in general, more coping strategies for stressful life events than men (Hobfoll et al., 1994). After experiencing a myocardial infarction women are more likely to adopt both problem-focused and emotional-focused coping strategies to try to change their situation such as seeking help, learning new skills and other cognitive and behavioural efforts (Bogg et al., 2000). Women are also more likely than men to share the experiences of the cardiac event and to simply accept what has happened. Household activities are reported as helping the recovery among women (Kristofferzon et al., 2003). Men with CHD are usually more likely than women to deny anger, depression and anxiety (Ketterer et al., 2004). The coping strategy often adopted by men is focusing on work and keeping physically active (Kristofferzon et al., 2003). The use of such strategies by men might also explain the findings showing a long term decline in quality of life. In fact a high proportion of men with CHD in the ELSA sample is retired and physically inactive. Therefore it might be that not being able to work or to do physical activity affects the capacity of men to cope with the illness.

A final remark should be made on why the quality of life of women with CHD decreased between baseline and four year follow-up, while the risk of having depressive symptoms reduced. It is possible that subgroups of women with certain characteristics (and changes over time therein) contributed to these results. Another possible

explanation might be that women experienced depressive symptoms as a consequence of the disease, therefore at four years of follow-up they might have adjusted to the CHD event and consequently their mental health improved. The decrease in quality of life occurred at four year follow-up could have been a consequence of ageing. This can be supported by the finding of this thesis which showed a decrease in quality of life over time among the healthy population (Well group) used as a reference group. Previous studies exploring longitudinal changes in quality of life in a general population of older individuals have shown that there is a trend of worsening quality of life over time at older ages (Zaninotto et al., 2009; Webb et al. 2010).

Gender differences in trajectories of quality of life and depressive symptoms: comparing the CHD and the Well groups

In the Well group men and women reported similar trajectories of quality of life and depressive symptoms. The quality of life of people from the Well group decreased significantly between baseline and years two and four of follow-up, while the odds of having depressive symptoms did not change significantly over time. People from this group had levels of quality of life that were on average higher than the general ELSA population, and the prevalence of depressive symptoms was lower. It is possible that the decline in quality of life levels is a consequence of ageing or development or progression of other disease. About 3% of people in this group experienced a CHD event after the baseline interview and 7% experienced other diseases (such as diabetes, stroke, pulmonary disease, Alzheimer, Parkinson's and cancer). Stable levels of depressive symptoms over time might reflect the relatively good health reported in this group of people at baseline; this group of people could also be more resilient to any event that might have happened to them.

Gender differences in quality of life were found at two year follow-up only, when women had higher quality of life than men. Women from the Well group were more likely than men to have depressive symptoms at each time point. It is difficult to understand why gender differences in quality of life and depressive symptoms found in the CHD group were different from those found in the healthy group at four year follow-up, even after controlling for the same covariates. It is possible that some of the covariates and their change over time had a different impact on the outcomes considered according to their disease status.

Gender differences in quality of life and depressive symptoms found among people with CHD were different from those found in the reference population of healthy individuals (Well group) at four year follow-up. It cannot be concluded that the results are specific to people with CHD because individuals with a disease other than CHD were not studied. However, for the purpose of this study it was important to see that gender differences in quality of life and depressive symptoms found among people with CHD were not the same as those in the healthy group, even after adjusting for the same covariates.

Covariate adjustment

Covariate adjustment was shown to be important in the analysis of quality of life for both the CHD and Well groups and for the CHD group only in the analysis of depressive symptoms.

In the models for quality of life, adjustment for depressive symptoms changed the conclusions about trajectories of quality of life among people with CHD. If the model was not adjusted for depressive symptoms it would have been concluded that there was a significant gender difference in the rate of change of quality of life at years two and four of follow-up. This is mainly attributable to the fact that for men the risk of depressive symptoms did not change over time, therefore not adjusting for depressive symptoms would have resulted in a steeper rate of change in quality of life.

Among men with CHD, adjustment for pain (reference category is "not being troubled with pain") resulted in a significant decrease in quality of life between baseline and two year follow-up. It is difficult to understand the underlying processes of adjustment. One possible reason might be that men with CHD who were often troubled with pain already had low quality of life therefore they experienced a small change in their quality of life between baseline and two year follow-up (baseline mean quality of life 37.3 S.D. 9.7; two year follow-up mean quality of life 35.7 S.D. 9.3).

Also, in the Well group adjustment for social support (simultaneously with the other covariates) attenuated the difference in baseline levels of quality of life between men

and women, the baseline quality of life would have otherwise been higher among women.

Thus depressive symptoms, pain and possibly social support are mediators of the relationships between disease status and quality of life.

For the CHD group, it was found that after adjusting for covariates women were no longer more likely than men to have depressive symptoms at four year follow-up. That was true after adjustment for socio-demographic factors, health behaviour factors, health factors and social factors. Suggesting that adjustment for covariates in addition to age is important, otherwise the results could have led to different conclusions.

Comparisons with other studies

It is difficult to compare the results from this study with those from the literature as most of the latter studies had a shorter length of follow-up than ELSA and also the samples came mainly from selected community hospitals, with the exception of two studies (Bjerkeset et al., 2005; Ford et al., 2008). Nevertheless, the finding of no gender differences in quality of life at baseline is in agreement with two other studies reporting no gender differences post-myocardial infarction at baseline (Mendes de Leon et al., 2001; Kristofferzon et al., 2005b). This study is the first to show that women's quality of life was higher than men's quality of life at years two and four of follow-up. This study is also the first to show that women's quality of life did not deteriorate following the baseline report of a CHD event, while men's quality of life deteriorated over time.

The results showing that, following a CHD event, women are more likely than men to have depressive symptoms is well-known (Forrester et al., 1992; Freasure-Smith et al., 1999; Mallik et al., 2006), however, this study was the first to show that these gender differences were found at baseline and at two year follow-up but not at four year follow-up. It was explained earlier that the gender difference disappeared after adjustment for covariates. Only one previous study reported that gender differences in the mental health dimension of HRQOL found at one year post-myocardial infarction did not persist once the model was adjusted for demographic, clinical, comorbid and psychosocial covariates (Norris et al., 2007).

This is also the first study to show that gender differences in depressive symptoms and quality of life are not the same in the CHD group and the Well group, at four year follow-up. This is an important finding, because previous studies have not shown whether gender differences in depressive symptoms were also found in a population free from any disease. A possible explanation could be that no previous studies comparing the results of the CHD group with a reference population had a length of follow-up longer than a year. The result of an improvement in depressive symptoms in women is somewhat in line with the study of Bjerkeset et al., (2005) which showed that women had a significant decrease in the risk of depression two year post-myocardial infarction.

Strengths and limitations

One of the strengths of this analysis is the use of a large sample of older people living in private households in England. The study has been designed to collect information on topics necessary to understand the economic, social, psychological and health elements of the ageing process. Some of the advantages of using this data set include: information on angina symptoms as well as myocardial infarction; the ability to measure well-being with two distinct measures such as depressive symptoms (experienced well-being) and quality of life (evaluative well-being); and the ability to compare the results for the CHD population with those for a healthy population.

The treatment of missing data constitutes a further strength of this analysis. Missing data often occur in epidemiological studies where non-response is a major problem. In addition to non-response longitudinal studies also face attrition due to death or drop-out from the study. The development of sophisticated missing data techniques allow researchers to improve the validity of epidemiological research results and to reduce estimation bias caused by missing data (Sterne et al., 2009; Jelicic et al., 2009). The technique used to impute missing data in this analysis is particularly suitable for repeated measures. The inclusion of several auxiliary variables helped reinforce the MAR assumption. Results for depressive symptoms based on augmented samples were similar to those obtained from the sample with missing data (observed). More differences were found in the results of quality of life. This finding most probably reflects the fact that the missing data mechanism was not missing completely at random. Missingness in quality of life depended on observed characteristics such as increasing age, CHD, low education, poor wealth, not cohabiting with a partner, permanently

unable to work, not currently in paid employment and looking after home or family and being physically inactive. Those reporting these characteristics are more likely to drop-out and to report a lower quality of life than those who were more advantaged (i.e. those in the healthy group, with high education, cohabiting with a partner and so forth).
Another strength is the use of multilevel models, which enabled modelling individual trajectories of quality of life and depressive symptoms over time taking into account the hierarchical structure of the data and has the advantage of providing efficient estimates of the parameters of interest.

One possible limitation of this analysis is the use of a self-reported measure of CHD. Unfortunately ELSA does not collect objective measures of CHD nor does it link the respondents' information with medical or hospital records. The impact that misclassification bias might have on the results presented here is addressed in the next chapter. It should also be mentioned that the trajectories of quality of life and depressive symptoms were studied for people with CHD and those in the Well group according to their disease status at baseline. About 12% of people with CHD experienced a repeat event at subsequent waves which might have affected their quality of life and depressive symptoms. Further, no distinction was made between the first and recurrent CHD event. It was implicitly assumed that a recent recurrence of a CHD event was as important as the first onset.

Conclusions

To summarise, it was found that, compared to people from the Well group, men and women with CHD had, on average, lower levels of quality of life. Men with CHD were also at higher risk of having depressive symptoms than men from the Well group. The findings from this thesis did not support the hypothesis that women with CHD were more likely to have depressive symptoms than women from the Well group.

The results of this analysis supported the hypothesis of differently shaped trajectories over time of quality of life and depressive symptoms in men and women following the onset of CHD. However, this analysis did not support the hypothesis of women with CHD reporting lower quality of life than men at any time. Significant gender differences in depressive symptoms were found at baseline and at two year follow-up only. The issue of the self-reported measure of CHD is addressed in the next chapter which

investigates the impact that misclassification bias might have on the results presented here.

Chapter 5: A Sensitivity analysis investigating bias due to misclassification of self-reported CHD

In the previous chapter gender differences in quality of life and depressive symptoms among people with CHD were explored longitudinally. One of the acknowledged limitations of the analysis was the use of the self-reported measure of CHD. The impact that self-report bias might have on the results presented in Chapter 4 is addressed in this chapter.

5.1 Introduction

Most epidemiological studies and health surveys assess the presence of chronic disease from self-report, as opposed to clinical assessments mainly because the collection of self-reported conditions involves lower costs (Kriesgsman et al., 1996). Several studies have assessed the value of a self-reported measure of myocardial infarction (Bush et al., 1989; Okura et al., 2004; Merkin et al., 2007; Yamagishi et al., 2009; Lampe et al., 2009), angina (Bush et al., 1989) and CHD (angina and/or myocardial infarction) (Kehoe et al., 1994; Haapanen et al., 1997; Lampe et al., 2009; Baumeister et al., 2010) by comparing self-reports with medical records. Table 5.1 gives a summary of findings from studies that have validated self-reported measures of CHD (angina, myocardial infarction or both). These studies have focussed on sensitivity and/or specificity of self-reported angina and/or myocardial infarction and/or on the agreement (percent and kappa) between self-reports and medical records or disease registries. Sensitivity is the proportion of true positives (defined as the presence of the condition according to clinical assessment of medical records) that are correctly identified by the self-report measure; specificity is defined as the proportion of true negatives (absence of the condition) that are correctly identified by the self-report measure (Altman and Bland 1994). Total agreement percent is defined as correctly reported positive and negative self-assessments over total reports or records. Cohen's kappa is a measure of inter-rater agreement between self-reports and clinical assessment calculated as the amount by which the observed agreement (p) exceeds that expected by chance alone (p_e), divided by the maximum which this difference could be (1- p_e). As suggested by Landis and Koch (1977) a kappa value of <0.40 is considered poor-to-fair agreement, a kappa value

of 0.41 to 0.60 is considered moderate agreement, a kappa value of 0.61 to 0.80 is considered substantial agreement, and a kappa value of 0.81 to 1.00 is considered excellent agreement.

Substantial agreement between self-reported myocardial infarction and/or angina and medical records was found in several studies with kappa over 0.70 and/or percent agreement greater than or equal to 80 (Bush et al., 1989; Okura et al., 2004; Lampe et al., 2009; Barr et al., 2009). Some of the studies reported moderate sensitivity (60%) (Merkin et al., 2007) and high sensitivity (over 89%) of self-reported myocardial infarction (Okura et al., 2004; Yamagishi et al., 2009) and angina (Barr et al., 2009). The studies reporting specificity of self-reported myocardial infarction found that it was considerably higher (over 93%) than the sensitivity (Okura et al., 2004; Merkin et al., 2007; Yamagishi et al., 2009). Results from studies comparing self-reported medical history of CHD with medical records (or physician's records) found a kappa greater than or equal to 0.80 (Haapanen et al., 1997; Lampe et al., 2009; Baumeister et al., 2010). The specificity of self-reported CHD was greater than or equal to 96%, and the sensitivity was greater than or equal to 88% (Haapanen et al., 1997; Baumeister et al., 2010) with the exception of two studies that found sensitivity values equal to 60% and 64% (Kehoe et al., 1994; Merkin et al., 2007).

Overall, from results reported in the literature there is evidence that the assessment of CHD by self-reports is a valid alternative when clinical assessment is not feasible. Given that self-reported measures of chronic diseases are widely used in epidemiology, the next question is to what extent a misclassification in a self-reported exposure introduces bias in model estimates in spite of relatively high agreement, high sensitivity and specificity. Self-reported CHD is used in the ELSA study. This information is not validated by a clinical screening or verified against medical records. Therefore in ELSA is not possible to ascertain the potential accuracy of self-reported CHD as compared to medical records. Since the definition of CHD used in this thesis has a two-year recall, it is possible that only a low rate of positive diagnosis was incorrect. However, it would be interesting to investigate the extent to which the self-reported measure of CHD used in this thesis may lead to biased estimates and/or different conclusions in the results presented in the previous chapter.

Table 5.1 Review of studies validating self-reported measures of angina and/or myocardial infarction

First author	Number and age of subjects	Setting	Condition being validated	Validation	Agreement % and kappa	Sensitivity	Specificity
Bush T.L. 1989	107 aged 65+	Florida Screening-based	Angina and MI	Medical records	85% angina k=0.57 94% MI k=0.70	-	-
Kehoe R. 1994	1,389 mean age 65 (S.D.8.2)	Boston Cataract case-control study	CHD	Medical records	-	64%	96%
Haapanen N. 1997	596 aged 45 to 73	Finland	CHD	Medical records	k=0.80	88%	96%
Okura Y. 2004	1,950 aged 45+	Minnesota Population-based	MI	Medical records	97.8% k=0.80	89.5%	98.2%
Merkin S.S 2007	1,041 aged 18+	USA Patient-based	MI	Medical records and physician reports	k=0.55 Medical records k=0.33 Physician reports	60%	93%
Yamagishi K. 2009	90,102 aged 50+	Japan Screening-based	MI	Medical records	-	82%	-
Lampe F.C. 2009	5,701 men aged 52 to 75	Britain population-based	CHD	Medical records	80% k=0.82	-	-
Baumister H. 2010	7,124 aged 18 to 79	Germany population-based	CHD	Physician reports	k=0.81	91.9%	98.0%

Abbreviations: MI: myocardial infarction; CHD: coronary heart disease; PPV: positive predictive value; NPV: negative predictive value

When validation of a self-reported measure is not available, several methods have been proposed to adjust findings and to investigate the potential bias of a misclassified exposure (Lash and Fink, 2003; Fox et al., 2005; Chu et al., 2006; Orsini et al., 2008; Lyles and Lin, 2010). The focus of these studies varies from standard tabular data (case-control study setting), with a binary risk factor subject to misclassification, to estimated log odds ratios in logistic regression adjusted for misclassification based on assumed sensitivity and specificity parameters (Fox et al., 2005; Lyles and Lin, 2010). These methods allowed for both deterministic and probabilistic sensitivity analysis. In a deterministic sensitivity analysis the idea is to adjust risk ratios or odds ratios by assuming several pairs of sensitivity and specificity. Probabilistic sensitivity analysis allows uncertainty about the bias parameters, i.e. sensitivity and specificity (Fox et al., 2005; Chu et al., 2006; Orsini et al., 2008; Lyles and Lin, 2010), and requires the definition of a distribution for the sensitivity and specificity. Therefore, the accuracy of a correction for misclassification depends on how accurate the definition is of such a distribution (Fox et al., 2005).

However, the proposed methodologies for performing a sensitivity analysis are not suitable in the setting of this thesis. First, I have a continuous outcome (quality of life) as well as a binary one (depressive symptoms); to date and to my knowledge, the methods proposed in the literature were developed for binary indicator misclassified variables (either exposure or outcome); second, my binary exposure (CHD) is interacted with gender. The methods to perform sensitivity analysis proposed in the literature mentioned above can accommodate a confounder but not an interaction term (Lash and Fink, 2003; Fox et al., 2005; Chu et al., 2006; Orsini et al., 2008; Lyles and Lin, 2010).

One possible way of conducting a sensitivity analysis is to investigate the impact of misclassification of self-reported CHD in ELSA using an external validation study. For the purpose of this chapter, the Whitehall II study is used, which validates the self-reported measure of CHD with clinical assessment. From the validation of the self-reported CHD measure performed in the Whitehall II study two scenarios, under which misclassification might have occurred, are hypothesised and applied to the ELSA data. Results obtained under these two scenarios are compared to the results presented in the previous chapter in order to quantify the potential impact of CHD misclassification bias and assess potential accuracy. The Whitehall II study has many similarities with the ELSA study, not only in terms of setting and age range of the sample but also in terms of data collection and measures. However, in the Whitehall II study only positive self-

reported cases of CHD could be verified, since follow-up of clinical records and validation using clinically verified events was only carried out for the subset of Whitehall II participants in whom there was a suggestion of a CHD event. Therefore sensitivity and specificity could not be calculated.

In order to assess the impact that also false negatives could have on the results, a deterministic sensitivity analysis (Greenland, 1996) is performed. In a deterministic sensitivity analysis the idea is to back-calculate the data that would have been observed without bias, assuming several pairs of sensitivity and specificity. The sensitivity and specificity values reported in Table 5.1 come from studies that are not directly comparable with the ELSA sub-sample used in this thesis. However, results from these studies can suggest educated values of specificity and sensitivity. For example, all studies found higher specificity than sensitivity; the lowest sensitivity value among those reported in Table 5.1 was 60%. The population with age range similar to ELSA was the one used by Yamagishi et al. (2009) and they found a sensitivity of self-reported MI equal to 82%. For the deterministic sensitivity analysis two values of sensitivity (60% and 80%) and two values of specificity (90% and 95%) are used. All possible combinations of these values are calculated to create four scenarios (scenario 1: sensitivity 60% and specificity 90%; scenario 2: sensitivity 60% specificity 95%; scenario 3: sensitivity 80% and specificity 90%; scenario 4: sensitivity 80% and specificity 95%) under which the prevalence of CHD in ELSA is modified.

Even a very simple sensitivity analysis can shed light on the robustness of the results presented in Chapter 4. It is hypothesised that the self-report measure of CHD is a reliable alternative when clinical assessment is not available and therefore the results will not draw different conclusions.

5.2 Sensitivity analysis based on Whitehall II validation study

5.2.1 Methods

The Whitehall II study

Whitehall II is a longitudinal study of 10,308 women and men, all of whom were employed in the London offices of the British Civil Service at the time they were recruited to the study in 1985. The baseline survey (phase 1) included a clinical

examination. Since then, ten phases of data collection have been completed, of which every odd-numbered phase has included a medical examination in addition to a questionnaire. The Whitehall II study was set up with the explicit purpose of testing hypotheses as to the causes of the social gradient in cardiovascular and other diseases (Marmot and Brunner, 2005; Marmot et al., 1991). For the purpose of the sensitivity analysis and matching the sample with ELSA, phase 7 data are used. This phase includes the questionnaire and a clinical screening carried out between 1st of October 2002 and 30th of September 2004, when participants were aged 50 to 74. The sample size at phase 7 was 6,761 (66% of Phase 1 responders).

Measures

In the phase 7 (2003-2004) questionnaire, participants were asked the following questions: "Since 2001 has a doctor told you that you have had angina?" and "Since 2001 has a doctor told you that you have had a heart attack?". From these questions I derived self-reported doctor diagnosed CHD (which occurred on average two years preceding the phase 7 interview). A variable on validated CHD events was already available in the data set and it was derived by the Whitehall II team. Briefly, CHD diagnosis was based on clinically verified events. Non-fatal myocardial infarction was defined following MONICA (MONItoring of trends and determinants in CArdiovascular disease Project) criteria based on questionnaires, study electrocardiograms, hospital acute electrocardiograms (ECGs), cardiac enzymes and physician records (Britton and Shipley, 2010). Angina was assessed on the basis of participants' reports of symptoms and diagnoses with corroboration in medical records or abnormalities on a resting ECG, exercise ECG or coronary angiogram. Classification was carried out independently by two trained coders, with adjudication in the event of disagreement (Britton and Shipley, 2010).

Only positive self-reported cases of CHD could be verified, since follow-up of clinical records and validation using clinically verified events was only carried out for the subset of Whitehall II participants in whom there was a suggestion of a CHD event.

Additional measures used for this sensitivity analysis were age and educational attainment based on years of full time education.

Data analysis

First, the age and education adjusted prevalence of self-reported CHD in the two years preceding the interview was calculated for Whitehall II and for ELSA (Table 5.2). Using a direct standardisation method, with the standards being the age and education distribution of ELSA at wave 1, the analysis was restricted to participants aged up to 74 in order to make the standardisation sample comparable to that of Whitehall II. Since it was not possible to create a similar Well group in the Whitehall II data, for comparison purposes the prevalence of CHD in ELSA reported in this Table is the proportion of those reporting CHD in the two years preceding the interview over the total of the population (aged up to 74 n=9,347). Among participants of Whitehall II study, the adjusted prevalence of those reporting having had a CHD event in the two years preceding the interview was about half that of ELSA (Table 5.2).

Table 5.2 Unadjusted and adjusted prevalence of self-reported CHD in ELSA and Whitehall II (age range 50 to 74)

	No CHD	**CHD**
	% (95% CI)	**% (95% CI)**
Unadjusted		
ELSA[a]	93.1 (92.5, 93.6)	6.9 (6.4, 7.5)
N	*8,697*	*650*
Whitehall II	97.3 (96.9, 97.6)	2.7 (2.4, 3.1)
N	*6,717*	*189*
Adjusted[b]		
ELSA[a]	93.0 (92.5, 93.5)	7.0 (6.5, 7.5)
N	*8,680*	*653*
Whitehall II	96.6 (95.7, 97.3)	3.4 (23.1, 28.2)
N	*4, 938*	*174*

a ELSA whole sample at wave 1 (not restricted to CHD and Well group). b Age and education-standardised figures using ELSA wave 1 as standard population.

Second, the prevalence of both true positives (correctly self-reported CHD) and false positives (incorrectly self-reported CHD) was calculated based on the validation of CHD in Whitehall II. A logistic regression was performed in order to explore the factors associated with the false positives in Whitehall II (Table A5.1 in appendix). In Whitehall II, gender was not associated with higher odds of reporting a false CHD diagnosis; older age ($p<0.001$) and lowest educational level (compared to highest $p<0.05$) were both associated with higher odds of reporting a false CHD diagnosis.

Participants of Whitehall II are white-collar civil servants, therefore results are usually affected by the healthy worker effect at baseline (Ferrie et al., 2009). To partly adjust for the age and education influence on self-reported CHD true positive and false positive were standardised for age and education.

To quantify the impact of self-reported CHD misclassification bias on the parameter estimates obtained from the random intercept models reported in Chapter 4, the age and education adjusted prevalence of false positive in Whitehall II (see Table 5.3 in results) was applied to each of the ELSA imputed data sets (5 data sets). These data sets were re-analysed under two possible scenarios. The first scenario assumes that the number of falsely reporting CHD is a random sample of people of the whole population; therefore the incorrect diagnosis occurred randomly and did not depend on CHD status. Using this definition gives 17% of people in the CHD group misreporting their status (see Appendix 5.1 for a detailed description of how the prevalence is derived).

The second scenario is more realistic and assumes that the number of those incorrectly reporting self-reported CHD is a random sample of people in the self-reported CHD population. Using this definition gives 35% of people in the CHD group misreporting their status (see Appendix 5.1 for a detailed description of how the prevalence is derived).

To introduce some randomness in the alteration of the CHD prevalence in ELSA, random uniform numbers were generated (Appendix 5.1). The prevalence of people with CHD was altered in each imputed data set according to scenario 1 and to scenario 2 as follows: people with a self-reported CHD diagnosis were recoded as not having CHD if the random number was less than or equal to 0.17 for scenario 1 and less than or equal to 0.35 for scenario 2 (see Appendix 5.2 for a Stata code of how the numbers were generated and recoded). The newly derived prevalence of individual reports of CHD was then used to estimate random intercept models (based on 5 imputed data sets) described in chapter 4. For comparison, the results from imputed data sets reported in chapter 4 were re-analysed limiting the upper age to 74. For simplicity, results are shown only for the main parameters of interest and omitted for covariates.

Evaluation criterion

To quantify to what extent the results obtained from the ELSA original sample are similar to those which are obtained under the two scenarios, I used bias for assessments of the potential accuracy.

Bias: $(\hat{\beta} - \beta)$ which is the difference between the estimate obtained from the original ELSA data and the estimate obtained under one of the two scenarios.

5.2.2 Results

Self-reported CHD and validation

Table 5.3 reports the age and education standardised prevalence of true positive and false positive cases of self-reported CHD in Whitehall II participants. About 2% of people in the sample correctly reported a CHD diagnosis (70% of those who self-reported CHD). About 30% of those who self-reported a CHD event did not have their diagnosis confirmed by a clinical screening or medical record (1% of the total sample).

Table 5.3 Adjusted[a] prevalence of validated CHD in Whitehall II

	% (95% CI)
No CHD	96.6 (95.7, 97.3)
N	*4,938*
True positive	2.2 (1.7, 2.9)
N	*113*
False positive	1.2 (0.8, 1.8)
N	61

a Age and education-standardised figures using ELSA as standard population. Age range 50 to 74

Table 5.4 shows the unadjusted prevalence of CHD among ELSA people aged up to 74, comparing the original data with the data obtained under scenario 1 and scenario 2. The results are combined from the 5 imputed data sets. The original prevalence of self-reported CHD was 17.2%. Under scenario 1 this prevalence decreased to 14.3%. Under scenario 2 there is a lower prevalence of people with CHD (11.3%) compared to

scenario 1 and the original data and consequently a higher prevalence of people in the Well group.

Table 5.4 Prevalence of self-reported CHD in ELSA (imputed data)

	Well group	CHD group
	%	%
ELSA original data	82.8	17.2
(95% CI)	(82.1,83.5)	(16.5,17.9)
Scenario 1[a]	85.7	14.3
(95% CI)	(84.6,86.9)	(13.1,15.4)
Scenario 2[b]	88.9	11.1
(95% CI)	(87.7, 89.9)	(9.8, 12.3)

Figures based on 5 imputed data sets. Age range 50 to 74.
a False positives is a random sample of people of the whole population. b False positives is a random sample of people of the self-reported CHD population

Sensitivity analysis of quality of life models

Table 5.5 shows the results of the linear random intercept model of quality of life of the original ELSA data compared to the results obtained under scenario 1 and scenario 2. The results in the first column of Table 5.5 differ slightly from those reported in chapter 4, Table 4.3, due to the sample being restricted to those aged of 74 or younger at baseline. The main difference compared with the results based on the whole sample (aged 50 and over) is that the interaction term between CHD and gender in Table 5.5 was no longer statistically significant. Results from scenario 1, compared to the original results of ELSA were quite similar for the coefficient for gender and the interaction term, while the coefficient for CHD has decreased in magnitude (-1.57 for the original data and -1.49 for scenario 1). Overall, the potential accuracy of the original estimates under the first scenario was good. Assuming that false positives is a random sample of the whole population led in general to smaller parameter estimates. Under scenario 2, the coefficient for CHD decreased in magnitude, compared to the coefficient from the original ELSA data (-1.37 and -1.57 respectively), but the standard error was larger. The coefficient for gender was slightly bigger compared to the coefficient for the original

ELSA data, but the standard errors were the same. The coefficient for the interaction term and its standard error decreased in magnitude (0.47 in the original data and 0.36 in scenario 2). The baseline values of quality of life (constant) obtained under the two scenarios were almost identical to the original data. Conclusions about the gender specific relationships between CHD and quality of life did not change according to results from scenario 1 and scenario 2.

Table 5.5 Sensitivity analysis of the linear random intercept model for quality of life, sample aged up to 74 (imputed data)

	ELSA original		**Scenario 1[a]**		**Scenario 2[b]**	
	Coef.	***Std. Err.***	**Coef.**	***Std. Err.***	**Coef.**	***Std. Err.***
CHD	-1.57	*0.31*	-1.49	*0.37*	-1.37	*0.40*
Bias			-0.08		-0.19	
Gender	0.56	*0.20*	0.60	*0.20*	0.65	*0.20*
Bias			-0.04		-0.08	
CHD*Gender	0.47	*0.47*	0.43	*0.57*	0.36	*0.78*
Bias			0.04		0.11	
Constant	47.36	*0.24*	47.32	*0.24*	47.27	*0.24*
Bias			0.04		0.09	

a False positives is a random sample of people of the whole population. b False positives is a random sample of people of the self-reported CHD population. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends.

Table 5.6 reports the results of the linear random intercept model for quality of life with time among people with CHD, comparing the results from the original ELSA data with those obtained under scenario 1 and scenario 2. Results in the first column of Table 5.6 differ from those reported in Chapter 4, Table 4.4, in that the interaction term between wave 3 and gender was significant in the analysis based on the sample restricted to the age of 74 or less, meaning that there was a significant gender difference in the rate of change in quality of life. The potential accuracy of the original estimates under the first scenario was good, as shown by the small values of bias. Standard errors were a bit larger than those presented in the first column. However, under this scenario it was not

found that women had significantly higher quality of life than men at four year follow-up (wave 3). Result that contradicted the finding based on the original ELSA data.

Under scenario 2, the parameter estimates for gender, wave 3 and the interaction term between gender and wave 2 were smaller than those obtained from the original data. However, the constant was slightly higher than that obtained from the original data. Therefore, if the false positives cases are assumed to be a random sample of the CHD population the levels of quality of life at each wave found in the original results would be slightly lower for both men and women. Also, under scenario 2 the coefficient for wave 2 (two year follow-up) was not statistically significant; suggesting that compared to baseline (wave 1) the quality of life of men was not significantly lower. This means that misclassification might introduce type I error (which occurs when the null hypothesis is wrongly rejected). For women under scenario 2, quality of life also did not change significantly at wave 3 (four year follow-up). Lastly, under this scenario the quality of life of women with CHD at wave 2 (two year follow-up) was not significantly higher than that of men.

Table 5.6 Sensitivity analysis of the linear random intercept model for quality of life with time, people with CHD aged up to 74 (imputed data)

	ELSA original		**Scenario 1[a]**		**Scenario 2[b]**	
	Coef.	***Std. Err.***	**Coef.**	***Std. Err.***	**Coef.**	***Std. Err.***
Gender	0.30	*0.59*	0.31	*0.73*	0.15	*0.87*
Bias			0.00		0.16	
Wave 2	-1.05	*0.42*	-1.04	*0.49*	-0.97	*0.66*
Bias			-0.01		-0.08	
Wave 3	-3.01	*0.44*	-3.01	*0.53*	-2.89	*0.60*
Bias			0.00		-0.12	
Wave 2*Gender	1.16	*0.63*	1.20	*0.72*	1.02	*1.01*
Bias			-0.04		0.14	
Wave 3*Gender	1.77	*0.63*	1.72	*0.79*	1.85	*0.89*
Bias			0.05		-0.08	
Constant	47.76	*0.71*	47.74	*0.79*	47.88	*1.27*
Bias			0.02		-0.12	

a False positives is a random sample of people of the whole population. b False positives is a random sample of people of the self-reported CHD population. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends.

Table 5.7 shows the results of the linear random intercept model for quality of life with time among people from the Well group, comparing the results from the original ELSA data with those obtained under scenario 1 and scenario 2. The results in the first column of Table 5.7 differ slightly from those reported in Chapter 4, Table 4.5, due to the age of the sample being restricted to less than 74, but conclusions are the same. For the Well group, the results obtained under scenario 1 were very similar to those of the ELSA original data, and standard errors were the same. Bias values obtained under scenario 1 were small. This implies that changing the Well population by adding those people with a false positive CHD diagnosis did not change quality of life trajectories in this group, under scenario 1. Under scenario 2 values of bias were small and standard errors were the same as the ELSA original data. However, results obtained under this scenario suggest that the coefficient for gender was underestimated in the original data, but the constant was slightly higher.

Table 5.7 Sensitivity analysis of the linear random intercept model for quality of life with time, people from the Well group aged up to 74 (imputed data)

	ELSA original		**Scenario 1[a]**		**Scenario 2[b]**	
	Coef.	***Std. Err.***	**Coef.**	***Std. Err.***	**Coef.**	***Std. Err.***
Gender	0.24	*0.24*	0.27	*0.24*	0.35	*0.25*
Bias			-0.03		-0.10	
Wave 2	-1.15	*0.19*	-1.14	*0.19*	-1.13	*0.19*
Bias			-0.01		-0.02	
Wave 3	-2.81	*0.19*	-2.79	*0.19*	-2.80	*0.19*
Bias			-0.01		-0.01	
Wave 2*Gender	0.47	*0.25*	0.48	*0.25*	0.50	*0.25*
Bias			-0.01		-0.03	
Wave 3*Gender	0.40	*0.25*	0.45	*0.25*	0.45	*0.25*
Bias			-0.05		-0.05	
Constant	48.82	*0.27*	48.79	*0.27*	48.72	*0.28*
Bias			0.03		0.09	

a False positives is a random sample of people of the whole population. b False positives is a random sample of people of the self-reported CHD population. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends.

Sensitivity analysis of depressive symptoms models

Table 5.8 shows the results of the logistic random intercept model for depressive symptoms of the original ELSA data compared to the results obtained under scenario 1 and scenario 2. Although results restricted to age 74 or less showed in the first column of Table 5.8 are similar to those reported in chapter 4, Table 4.6, conclusions were the same. Results of the sensitivity analysis showed little change in the log odds under either scenario 1 or scenario 2, compared to ELSA original data. Also the values of the bias were all small, suggesting good potential accuracy. The coefficients for CHD were not statistically significant under scenarios 1 and 2, contradicting that obtained in the ELSA original data.

Table 5.8 Sensitivity analysis of the logistic random intercept model for depressive symptoms, sample aged up to 74 (imputed data)

	ELSA original		**Scenario 1[a]**		**Scenario 2[b]**	
	Coef.	***Std. Err.***	**Coef.**	***Std. Err.***	**Coef.**	***Std. Err.***
CHD	0.35	*0.16*	0.30	*0.19*	0.31	*0.25*
Bias			0.05		-0.01	
Gender	0.73	*0.11*	0.72	*0.11*	0.70	*0.11*
Bias			0.01		0.01	
CHD*Gender	-0.14	*0.23*	-0.12	*0.27*	-0.10	*0.41*
Bias			-0.02		-0.02	

a False positives is a random sample of people of the whole population. b False positives is a random sample of people of the self-reported CHD population. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends. Results presented as log odds.

Results of the logistic random intercept model for depressive symptoms with time among people with CHD are reported in Table 5.9, where comparisons of results from the original ELSA data with those obtained under scenario 1 and scenario 2 are presented. In this restricted age sample it was not found that women with CHD were significantly less likely to have depressive symptoms at four year follow-up compared to baseline (see Table 4.8). The rest of the results led to the same conclusions. The parameter estimates for gender, wave 2 and wave 3 obtained from the ELSA original

data would be overestimated if we assume that 17% of people with CHD falsely reported their disease (scenario 1), as suggested by the values of the bias. The parameter estimates of the interaction terms were not very different. Instead, if we assume that 35% of people with CHD misreported their disease status, then the results obtained from the ELSA original data for gender and wave 2, and the interaction term between gender and wave 2 would be underestimated and that of the interaction term between wave 3 and gender would be overestimated. Standard errors obtained under scenario 1 and scenario 2 were larger than those of the original data.

In terms of the conclusions, the most important difference was that under scenario 2, the log odds for gender were no longer statistically significant, implying that women with CHD were not at higher risk of having depressive symptoms than men at baseline. Also at wave 2 (two year follow-up) women did not have significantly higher odds of reporting depressive symptoms than men, result that was not found in the original data.

Table 5.9 Sensitivity analysis of the logistic random intercept model for depressive symptoms with time, people with CHD aged up to 74 (imputed data)

	ELSA original		**Scenario 1[a]**		**Scenario 2[b]**	
	Coef.	***Std. Err.***	**Coef.**	***Std. Err.***	**Coef.**	***Std. Err.***
Gender	0.71	*0.28*	0.65	*0.33*	0.76	*0.47*
Bias			0.06		-0.11	
Wave 2	0.09	*0.24*	0.04	*0.26*	0.22	*0.35*
Bias			0.05		-0.18	
Wave 3	0.06	*0.25*	-0.05	*0.28*	-0.01	*0.44*
Bias			0.11		-0.04	
Wave 2*Gender	-0.02	*0.34*	0.01	*0.39*	-0.17	*0.50*
Bias			-0.04		0.18	
Wave 3*Gender	-0.44	*0.35*	-0.29	*0.40*	-0.39	*0.62*
Bias			-0.15		0.11	

a False positives is a random sample of people of the whole population. b False positives is a random sample of people of the self-reported CHD population. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends. Results presented as log odds.

Table 5.10 shows the results of the logistic random intercept model for depressive symptoms with time among people in the Well group, comparing the results from the original ELSA data with those obtained under scenario 1 and scenario 2. The log odds and standard errors of the parameter estimates under scenario 1 were very close if not

the same as those of the original data. Under scenario 2, all the coefficients were smaller than those of the original data (except that of the interaction term between wave 3 and gender), also the standard errors were very similar. Conclusions remained unchanged under both scenarios.

Table 5.10 Sensitivity analysis of the logistic random intercept model for depressive symptoms with time, people from the Well group aged up to 74 (imputed data)

	Elsa original		Scenario 1[a]		Scenario 2[b]	
	Coef.	***Std. Err.***	**Coef.**	***Std. Err.***	**Coef.**	***Std. Err.***
Gender	0.84	*0.15*	0.84	*0.14*	0.80	*0.14*
Bias			0.00		0.04	
Wave 2	0.22	*0.14*	0.22	*0.13*	0.17	*0.13*
Bias			0.00		0.05	
Wave 3	-0.03	*0.14*	-0.01	*0.14*	-0.02	*0.14*
Bias			-0.03		0.02	
Wave 2*Gender	-0.29	*0.17*	-0.29	*0.17*	-0.24	*0.17*
Bias			0.00		-0.05	
Wave 3*Gender	-0.01	*0.18*	-0.06	*0.17*	-0.05	*0.18*
Bias			0.05		-0.01	

a False positives is a random sample of people of the whole population. b False positives is a random sample of people of the self-reported CHD population. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends. Results presented as log odds.

5.2.3 Summary of results

The results of the sensitivity analysis of trajectories of quality of life presented in Tables 5.6 and 5.7 are summarised graphically in Figure 5.1 for the CHD group and 5.2 for the Well group. The graphs provided a better insight on how biased the trajectories of quality of life might have been under different assumptions about misclassification. Assuming that 17% of people with CHD (scenario 1 where the false positives was assumed to be a random sample of the total population) misreported their status (and therefore they should have been classified as not having CHD) did not change the trajectories of quality of life of men with CHD (top part of Figure 5.1). Trajectories of

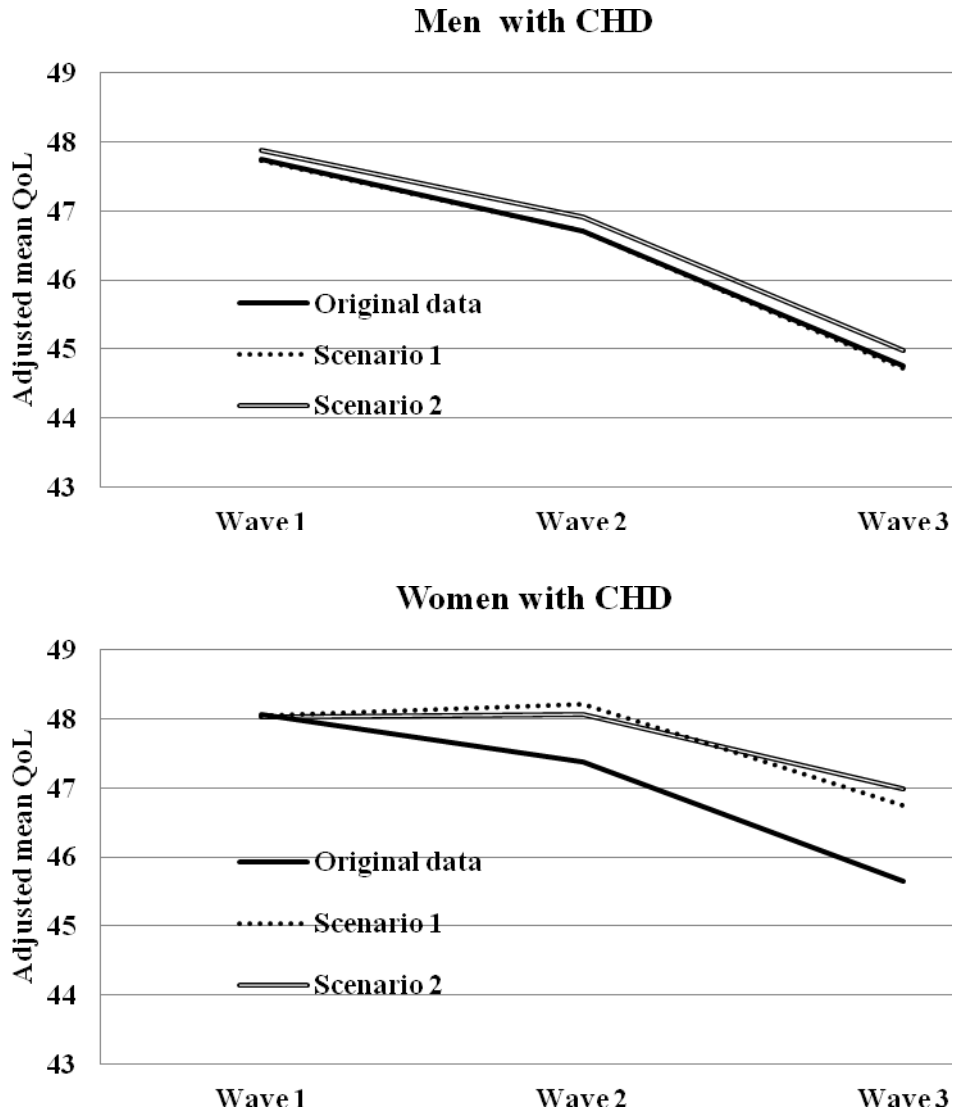
quality of life under scenario 1 overlapped with those found from the original data. If it was assumed that 35% of people in the CHD group falsely self-reported their disease status (scenario 2 where the false positives is assumed to be a random sample of the CHD population), the trajectories of quality of life of men obtained from the original data would be slightly underestimated. In fact, under scenario 2, the shape of trajectories of quality of life of men with CHD was the same as in the original data, but levels of quality of life were slightly higher at each wave (0.1 at wave 1, 0.2 at wave 2 and 0.3 at wave 3).

The shape of trajectories of quality of life of women with CHD did change according to the first scenario, and the levels of quality of life were higher at years two (0.8) and four of follow-up (0.7), than those obtained in the original data, but conclusions were the same as those obtained from the original data: women's quality of life was stable between baseline and two year follow-up (non-significant small decrease) and then decreased significantly at four year follow-up.

Under scenario 2, women's quality of life was stable between baseline and two year follow-up (wave 2) and then decreased slightly at four year follow-up (wave 3) but the change over time was not significant. Also at years two and four of follow-up the levels of quality of life were higher than those obtained from the original data (1.1 and 1.3 respectively). This means that if instead of assuming that people with CHD correctly self-report their disease status, we assume that about 35% of them self-reported their disease status incorrectly, women's trajectories of quality of life would be different. Women with CHD would not have decreased quality of life over time.

As for gender differences in quality of life, the finding of higher quality of life in women with CHD compared to men at four year follow-up could not be replicated under scenario 1 and that of higher quality of life at two year follow-up could not be replicated under scenario 2.

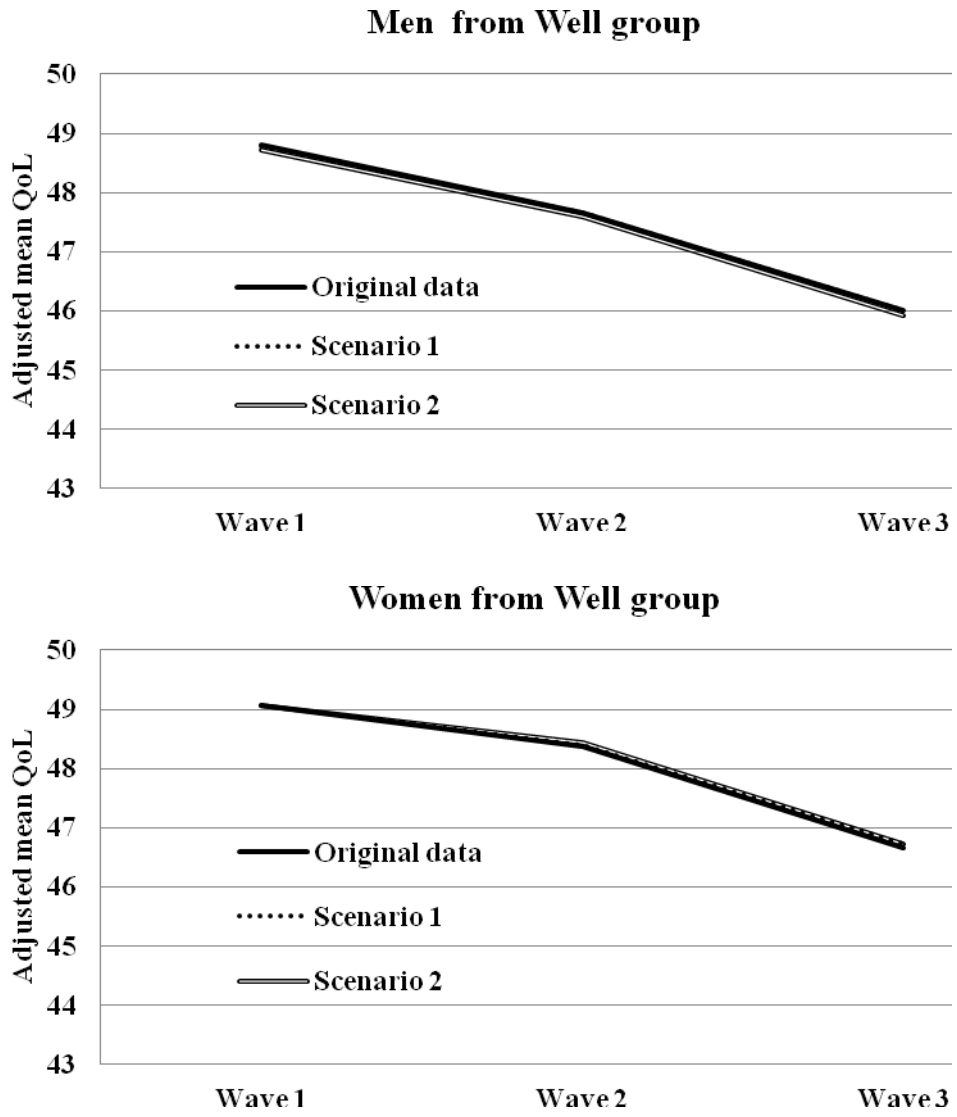
Figure 5.1 Sensitivity analysis of trajectories over time of quality of life for men and women with CHD, aged up to 74 (imputed data)



Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends. Age range 50 to 74.

From figure 5.2 we can see that the trajectories of quality of life of people from the Well obtained from the original data are quite robust, therefore the results for the Well group are not biased by false positives.

Figure 5.2 Sensitivity analysis of trajectories over time of quality of life for men and women from the Well group, aged up to 74 (imputed data)

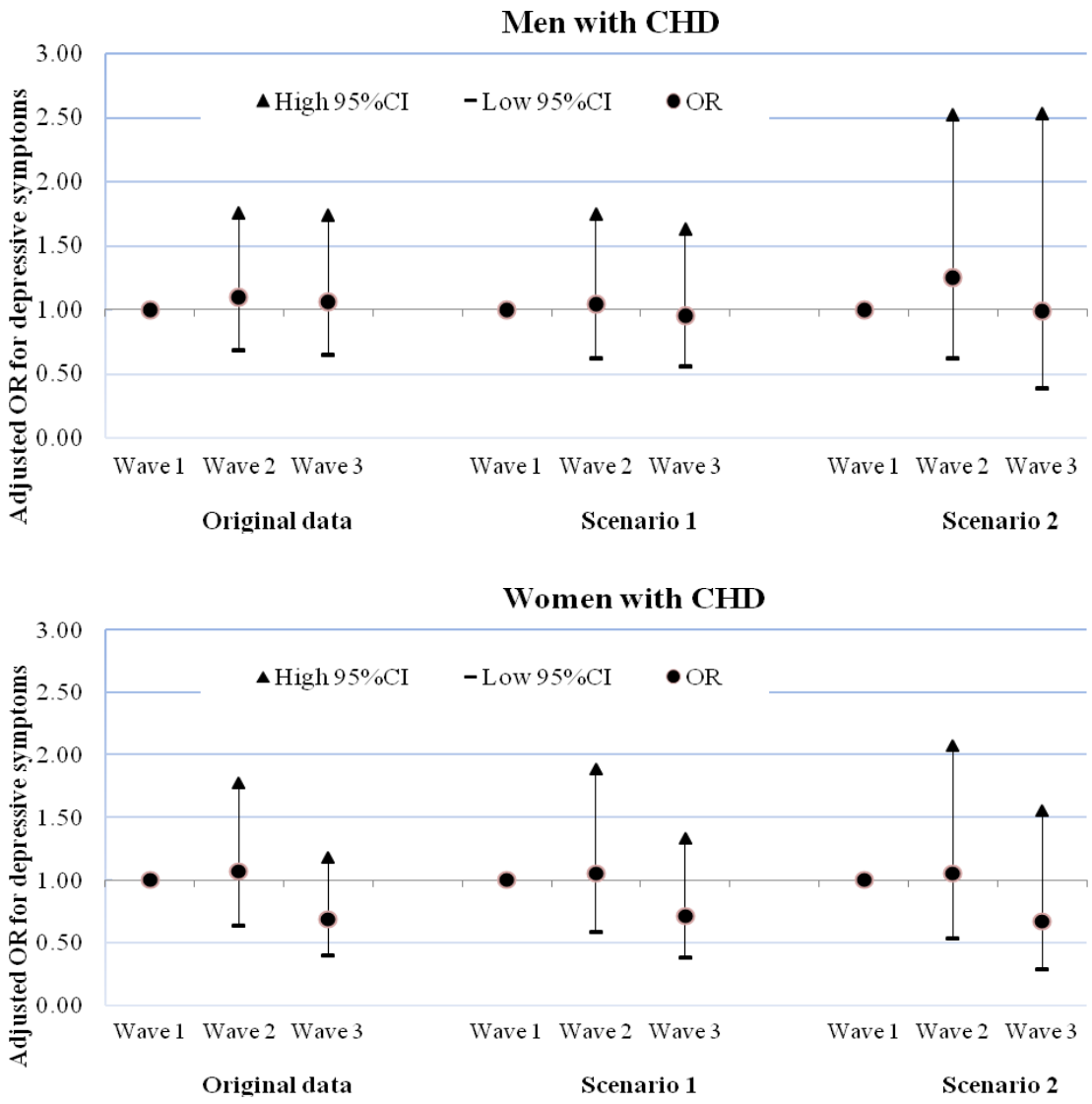


Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends. Age range 50 to 74.

Results of the trajectories of depressive symptoms presented in Tables 5.9 and 5.10 are presented graphically as odds ratios in Figure 5.3 for the CHD group and figure 5.4 for the Well group. The graphs showed that for men and women with CHD trajectories of depressive symptoms obtained from the original data and those obtained under scenarios 1 and 2 led to the same conclusions (Figure 5.3). However, under scenario 2 women with CHD were not more likely than men to have depressive symptoms at any point in time. Also under this scenario the confidence intervals of the odds ratios for the CHD group are larger than those of the original data.

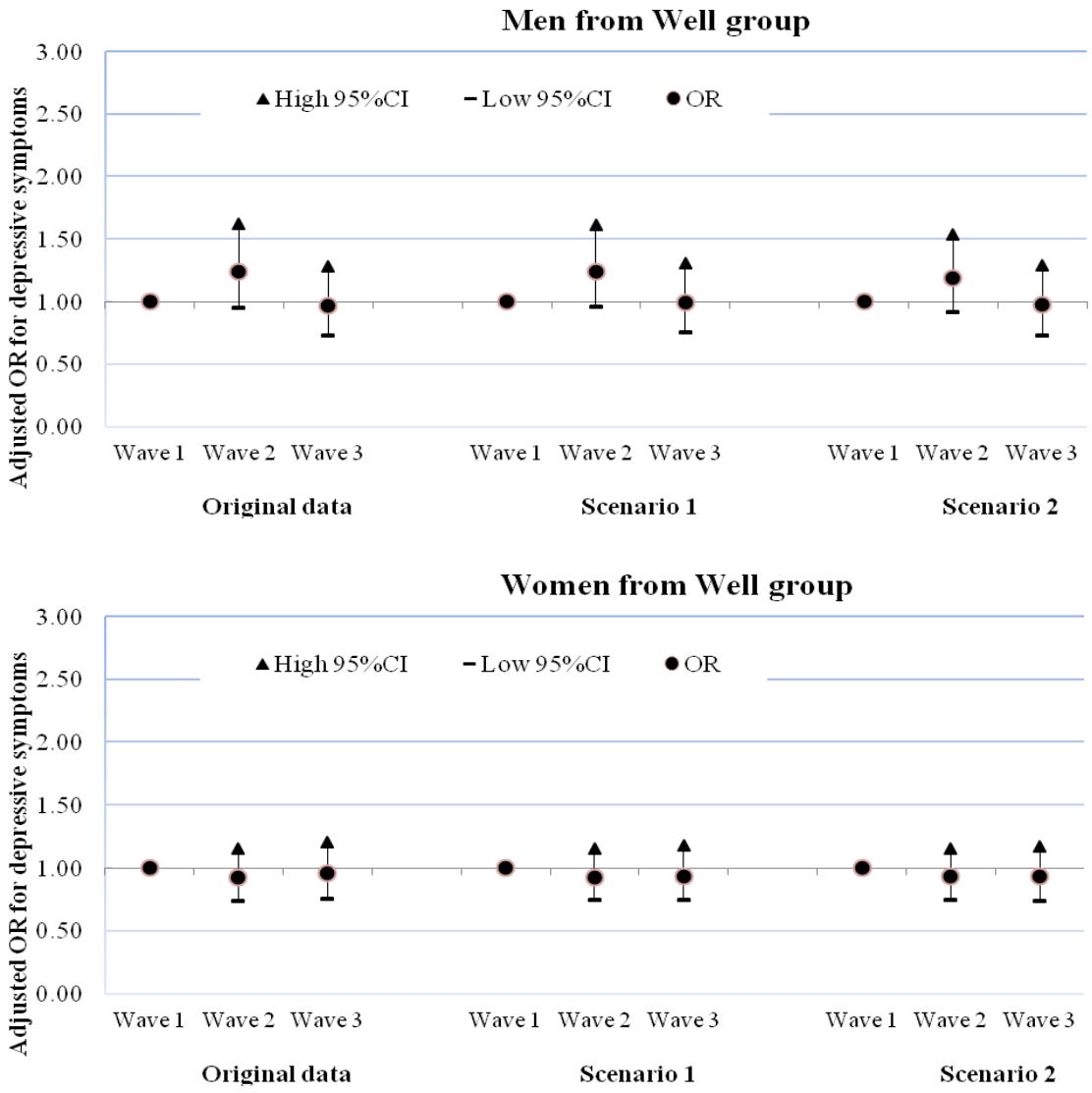
The ELSA original results obtained for the Well group were quite robust, as shown in figure 5.4.

Figure 5.3 Sensitivity analysis of trajectories over time of depressive symptoms, for men and women with CHD, aged up to 74 (imputed data)



Scenario 1: False positives is a random sample of people of the whole population. Scenario 2: False positives is a random sample of people of the self-reported CHD population. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends. Results presented as odds ratios.

Figure 5.4 Sensitivity analysis of trajectories over time of depressive symptoms, for men and women from the Well group, aged up to 74 (imputed data)



Scenario 1: False positives is a random sample of people of the whole population. Scenario 2: False positives is a random sample of people of the self-reported CHD population. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends. Results presented as odds ratios.

5.3 Deterministic sensitivity analysis

In the previous sensitivity analysis an external validation study was used to estimate misclassification probabilities and to assess the impact of bias due to false positives. However, from the validation study it was not possible to obtain values of sensitivity and specificity. In order to address the impact that misclassification error has, not only on the false positive cases of CHD but, also on the false negative cases of CHD a deterministic sensitivity analysis is conducted. In a deterministic sensitivity analysis it is estimated what the true parameter estimate(s) would be in light of the observed data and some level of hypothetical bias. Deterministic sensitivity analysis can be seen as a series of educated guesses about the bias parameters (Greenland 1996). The idea is to back-calculate the data that would have been observed without bias, assuming several pairs of rates of sensitivity and specificity.

5.3.1 Methods

Two values of sensitivity equal to 60% and 80% and two values of specificity equal to 90% and 95% are used. All the possible combinations of these values are calculated to create the followings four scenarios:

- scenario A: sensitivity 60% and specificity 90%;
- scenario B: sensitivity 60% specificity 95%;
- scenario C: sensitivity 80% and specificity 90%;
- scenario D: sensitivity 80% and specificity 95%.

The variable representing CHD status was manipulated according to the four scenarios described above. To introduce some randomness in the alteration of the variable, random uniform numbers were generated (appendix 5.3). Four new variables for CHD status were obtained in each imputed data set according to the each scenarios as follows: for scenario A cases from the CHD category were moved into the Well group if the random uniform number was less or equal than 0.40 and cases from the Well group were moved into the CHD category if the random uniform number was less or equal than 0.10; as for scenario A cases were moved if the probabilities were ≤ 0.40 and ≤ 0.05 for scenario B; ≤ 0.20 and ≤ 0.10 for scenario C; and ≤ 0.20 and ≤ 0.05 for scenario D.

Each newly derived variable for CHD was then used to estimate random intercept models (based on 5 imputed data sets) described in chapter 4.

Evaluation criterion

To quantify to what extent the results obtained from the ELSA original sample and those obtained under the two scenarios are similar I used the bias for the assessments of potential accuracy defined in section 5.2.1.

5.3.2 Results

Table 5.11 reports the pooled prevalence (based on 5 imputed data sets) of self-reported CHD obtained under each scenario. Scenario A assumed that the sensitivity (someone correctly classified as having CHD) was 60% and the specificity (someone correctly classified as not having CHD) was 90% which gave a prevalence of CHD equal to 19.2% very similar to the original prevalence (19.9%). This could be explained by the similar number of people misclassified in the CHD and Well group (40% of 895= 358 CHD and 10% of 3601=360 Well group). Holding the sensitivity constant (to 60%) and changing the specificity to a higher value (95%) led to a lower prevalence of CHD (15.9%) compared to the original data (scenario B). The prevalence was lower under scenario B because only 5% of people in the Well group were wrongly classified as being healthy while a greater number of people with CHD were misclassified as having the disease. Assuming a sensitivity of 80% and a specificity of 90% (scenario 3) led to a prevalence of CHD that was almost 4 percentage points higher than that of the original data. A sensitivity of 80% assumed that a higher number of people with CHD were correctly classified as having the disease compared to a sensitivity of 60%. Therefore under this scenario the number of healthy people wrongly classified was higher than the number of people wrongly classified as having CHD. The last scenario was probably the most realistic one, where the sensitivity and specificity were highest reflecting results found from most of the studies in the literature. Sensitivity set at 80% and specificity set at 95%, led to a prevalence of CHD equal to 21.3%, very close to the ELSA original prevalence.

Table 5.11 Corrected prevalence of the self-reported CHD under various assumptions about the CHD sensitivity (Se) and specificity (Sp) (imputed data)

	Well group	CHD group
	% (95% CI)	% (95% CI)
ELSA	80.1 (79.4, 80.8)	19.9 (19.2, 20.6)
ScenarioA Se: 60% Sp: 90%	80.8 (80.1, 81.4)	19.2 (18.6, 19.9)
Scenario B Se: 60% Sp: 95%	84.1 (83.5, 84.8)	15.9 (15.2, 16.5)
Scenario C Se: 80% Sp: 90%	76.2 (75.5, 76.9)	23.8 (23.1, 24.5)
Scenario D Se: 80% Sp: 95%	79.8 (79.1, 80.5)	20.2 (19.5, 20.9)

Figures based on the sample aged 50 and over.

Deterministic sensitivity analysis of quality of life models

Table 5.12 shows the results of the deterministic sensitivity analysis of the linear random intercept model for quality of life with the interaction term between CHD and gender. The four scenarios were compared with the original results reported in chapter 4. The coefficient for CHD changed from -2.00 to -1.12 under scenario A, to -1.38 under scenario B, to -1.35 under scenario C and to -1.65 under scenario D. This implied that the difference between men from the Well group and men with CHD in quality of life (adjusted) was larger in the ELSA original data where it was assumed that no misclassification of the self-reported CHD occurred.

The coefficient for gender increased in magnitude under scenarios A, B and C compared to the original data (0.41). These results mean that the difference in quality of life between men with CHD and men from the Well group obtained in the original results would be overestimated.

Under scenario A the coefficient for the interaction term was about a third of that of the original data: if there had been misclassification the difference in quality of life between men and women from the Well group changed from 0.98 in the original results to 0.32 under scenario A. Under scenario B and C the coefficients for the interaction term were

also lower than that obtained from the original data. Therefore the difference in quality of life between men and women from the Well group would be slightly overestimated in the original data.

Assuming highest values of sensitivity and specificity (scenario D) gave similar results to those obtained from the original data (with the exception of value of the interaction term which was larger under scenario D). This scenario showed that the results obtained from the original data could possibly be underestimated.

Table 5.12 Deterministic sensitivity analysis of the linear random intercept model for quality of life with time (imputed data)

	ELSA original		**Scenario A**		**Scenario B**		**Scenario C**		**Scenario D**	
	Coeff.	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***
CHD	-2.00	*0.27*	-1.12	*0.28*	-1.38	*0.30*	-1.35	*0.26*	-1.65	*0.27*
Bias			-0.88		-0.62		-0.65		-0.35	
Gender	0.41	*0.19*	0.62	*0.18*	0.56	*0.18*	0.50	*0.19*	0.39	*0.19*
Bias			-0.21		-0.15		-0.09		0.02	
CHD*gender	0.98	*0.40*	0.32	*0.41*	0.78	*0.44*	0.76	*0.38*	1.30	*0.40*
Bias			0.65		0.20		0.21		-0.32	
Constant	47.69	*0.22*	47.60	*0.22*	47.59	*0.22*	47.69	*0.22*	47.71	*0.22*
Bias			0.09		0.10		0.00		-0.02	

Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%.
Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends and family. Figures based on the sample aged 50 and over.

The results of the deterministic sensitivity analysis of the linear random intercept model of quality of life with and interaction term between gender and time for the CHD group are presented in Table 5.13. Under scenarios A and B the coefficients for gender changed from 0.86 to 0.38 and 0.64 respectively, meaning that the difference in quality of life at baseline between men and women with CHD was overestimated in the original

results. However, under scenarios C and D the coefficients for gender increased in magnitude and became significant. Therefore if sensitivity of 80% was assumed together with specificity values of 90% and 95% it would be concluded that women with CHD had significantly higher quality of life than men at baseline, a result that was not found in the original data. Under scenarios A, B and C the coefficients for wave 2 increased in magnitude compared to the original data, however conclusions remained the same (men having lower quality of life at wave 2 compared to wave 1). Under scenario D, the coefficient for wave 2 became smaller and was no longer significant. Under all four scenarios the coefficients for wave 3 were similar to the coefficient for the original data and conclusions were the same (men had lower quality of life at wave 3 compared to wave 1). The values of the interaction terms would be overestimated in the original data, compared to those obtained under the four scenarios. One exception was for the interaction term between wave 2 and gender obtained under scenario B which was larger than that of the original data and also statistically significant. This implied that assuming the lowest value of sensitivity and highest value of specificity led to a significant difference between men and women with CHD in the rate of change of quality of life at wave 2, a result that was not found in the original data.

The baseline quality of life of men with CHD (constant) was higher under scenarios A, B, and C compared to the original data and slightly lower under scenario D. For women the baseline quality of life was lower in the original data compared to scenarios A, C and D, but slightly higher compared to scenario B.

Table 5.13 Deterministic sensitivity analysis of the linear random intercept model for quality of life with time, people with CHD (imputed data)

	ELSA original		Scenario A		Scenario B		Scenario C		Scenario D	
	Coeff.	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***
Gender	0.83	*0.50*	0.38	*0.50*	0.64	*0.54*	1.11	*0.44*	1.39	*0.48*
Bias			0.46		0.19		-0.28		-0.56	
Wave 2	-0.84	*0.36*	-1.46	*0.37*	-1.16	*0.41*	-1.04	*0.33*	-0.67	*0.36*
Bias			0.62		0.32		0.20		-0.17	
Wave 3	-2.73	*0.37*	-2.91	*0.37*	-2.61	*0.42*	-2.93	*0.34*	-2.48	*0.36*
Bias			0.18		-0.13		0.20		-0.25	
Wave 2*Gender	0.98	*0.53*	0.88	*0.52*	1.32	*0.59*	0.26	*0.47*	0.72	*0.52*
Bias			0.10		-0.33		0.72		0.26	
Wave 3*Gender	0.98	*0.53*	0.74	*0.52*	0.69	*0.59*	0.31	*0.47*	0.64	*0.52*
Bias			0.25		0.29		0.67		0.35	
Constant	47.38	*0.64*	48.75	*0.58*	47.21	*0.67*	48.18	*0.53*	47.42	*0.59*
Bias			-1.37		0.17		-0.80		-0.04	

Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%.
Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends and family. Figures based on the sample aged 50 and over.

The results of the deterministic sensitivity analysis of trajectories of quality of life for the Well group are presented in Table 5.14. Results obtained under scenarios A and B showed that the difference in baseline quality of life of men and women from the Well group would be higher than that found from the original data and also statistically significant. Under scenario C the coefficient for gender was very similar to that of the original data (0.27 and 0.24 respectively); under scenario D the coefficient was slightly lower (0.16). Like in the original data, scenario C and D reported non-significant gender difference in baseline quality of life. Coefficients of wave 2 and wave 3 were slightly

higher under scenarios A, B, and C and slightly lower under scenario D compared to the original data. However, conclusions remain unchanged. Smallest values of bias were found under scenarios C and D, suggesting that when the sensitivity is highest, the results of the original data tend to be more robust. Interaction terms of the original data were slightly overestimated according to scenarios A and B and slightly underestimated according to scenarios C and D, however conclusions were the same. Standard errors were all very similar to those obtained from the original data.

The baseline quality of life of men from the Well group slightly decreased compared to the original data. The baseline quality of life of women from the Well group was almost the same under the four scenarios as that obtained from the original data.

Table 5.14 Deterministic sensitivity analysis of the linear random intercept model for quality of life with time, people from the Well group (imputed data)

	ELSA original		**Scenario A**		**Scenario B**		**Scenario C**		**Scenario D**	
	Coeff.	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***
Gender	0.24	*0.23*	0.52	*0.23*	0.45	*0.22*	0.27	*0.23*	0.16	*0.23*
Bias			-0.28		-0.21		-0.03		0.08	
Wave 2	-1.11	*0.18*	-0.90	*0.18*	-0.97	*0.17*	-1.05	*0.18*	-1.16	*0.18*
Bias			-0.21		-0.13		-0.06		0.05	
Wave 3	-2.88	*0.18*	-2.76	*0.18*	-2.82	*0.18*	-2.78	*0.18*	-2.93	*0.18*
Bias			-0.12		-0.06		-0.10		0.05	
Wave 2*Gender	0.25	*0.24*	0.18	*0.24*	0.15	*0.23*	0.39	*0.24*	0.32	*0.24*
Bias			0.07		0.10		-0.15		-0.07	
Wave 3*Gender	0.18	*0.24*	0.16	*0.24*	0.22	*0.23*	0.29	*0.24*	0.26	*0.24*
Bias			0.02		-0.04		-0.11		-0.08	
Constant	49.03	*0.25*	48.69	*0.25*	48.91	*0.25*	48.89	*0.26*	49.06	*0.25*
Bias			0.34		0.12		0.14		-0.03	

Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends and family. Figures based on the sample aged 50 and over.

Deterministic sensitivity analysis of depressive symptoms models

Table 5.15 shows the results of the deterministic sensitivity analysis of depressive symptoms. The coefficient for CHD decreased in magnitude and was no longer significant in any of the four scenarios compared to the original results. Therefore, results of the deterministic sensitivity analysis suggest that men with CHD were not significantly more likely to have depressive symptoms than men from the Well group. The coefficient for gender slightly decreased in magnitude in all four scenarios compared to the original results, but conclusions remained unchanged. The coefficient for the interaction term between CHD and gender decreased in magnitude under scenarios A, C and D, and was very similar to that of the original data under scenario B; as in the original data the interaction term was not statistically significant under any of the four scenarios. Values of bias were large for the coefficients of gender and interaction term under all scenarios, suggesting lack of potential accuracy of the original data.

Table 5.15 Deterministic sensitivity analysis of the logistic random intercept model for depressive symptoms (imputed data)

	ELSA original		**Scenario A**		**Scenario B**		**Scenario C**		**Scenario D**	
	Coeff.	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***
CHD	0.35	*0.16*	0.25	*0.14*	0.21	*0.16*	0.11	*0.13*	0.24	*0.14*
Bias			0.10		0.14		0.24		0.11	
Gender	0.78	*0.11*	0.73	*0.09*	0.75	*0.09*	0.75	*0.10*	0.74	*0.09*
Bias			0.05		0.03		0.03		0.04	
CHD*gender	-0.27	*0.23*	-0.10	*0.20*	-0.28	*0.21*	-0.18	*0.18*	-0.11	*0.19*
Bias			-0.17		0.01		-0.09		-0.16	

Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%.
Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends and family. Figures based on the sample aged 50 and over.

Results of the deterministic sensitivity analysis of the trajectories of depressive symptoms for the CHD group are shown in Table 5.16. Under scenarios A and D the

coefficient for gender increased in magnitude compared to the original data, but standard errors were very similar. This means that at baseline the gender difference in depressive symptoms obtained in the original data would be slightly underestimated if sensitivity and specificity were at the lowest and highest values. Under scenario B and C the coefficient decreased in magnitude (from 0.69 original data to 0.44 scenario B and to 0.52 scenario C), but under scenario B was no longer significant. Therefore if it was assumed that the self-reported measure of CHD had low sensitivity and high specificity, women would no longer be more likely than men to have depressive symptoms at baseline. The coefficient for wave 2 increased in magnitude under scenarios A and D compared to the original data. The coefficient decreased in magnitude under scenario C and under scenario B became negative (meaning that men were less likely to have depressive symptoms than at baseline). However, none of the coefficients were statistically significant, leading to the same conclusions as the original data (men at wave 2 were not significantly more likely than baseline to have depressive symptoms).

Standard errors were close to those obtained from the original data, with the exception of the standard error obtained under scenario B, which was larger. Under scenarios A to D, the coefficient for wave 3 decreased in magnitude; however none of the coefficients was statistically significant, implying that the odds of having depressive symptoms did not change significantly in men from baseline to four year follow-up, a result that was also found in the original data. Large bias values were found for the interaction terms according to all scenarios, suggesting lack of potential accuracy of the results based on the original data. None of the interaction terms obtained under the four scenarios were statistically significant, in line with results of the original data. In general, the results obtained from the original data would be most biased if lowest sensitivity and highest specificity are assumed (scenario B).

In terms of trajectories of depressive symptoms of women with CHD, the deterministic sensitivity analysis showed that the finding obtained from the original data of women being significantly less likely to have depressive symptoms at four year follow-up (wave 3) compared to baseline was not replicated when misclassification was assumed. Also, under scenarios A and C the finding that women were not at higher risk of having depressive symptoms than men was not replicated.

Table 5.16 Deterministic sensitivity analysis of the logistic random intercept model for depressive symptoms with time, people with CHD (imputed data)

	ELSA original		Scenario A		Scenario B		Scenario C		Scenario D	
	Coeff.	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***
Gender	0.69	*0.24*	0.78	*0.25*	0.44	*0.27*	0.52	*0.22*	0.76	*0.25*
Bias			-0.09		0.25		0.17		-0.07	
Wave 2	0.18	*0.20*	0.27	*0.21*	-0.18	*0.24*	0.13	*0.20*	0.24	*0.20*
Bias			-0.09		0.36		0.05		-0.06	
Wave 3	0.06	*0.21*	-0.03	*0.22*	-0.06	*0.24*	-0.22	*0.20*	0.01	*0.21*
Bias			0.09		0.12		0.28		0.05	
Wave 2*Gender	-0.12	*0.28*	-0.29	*0.29*	0.23	*0.33*	-0.01	*0.26*	-0.18	*0.29*
Bias			0.17		-0.35		-0.11		0.06	
Wave 3*Gender	-0.52	*0.29*	-0.15	*0.30*	-0.33	*0.33*	0.02	*0.27*	-0.41	*0.29*
Bias			-0.37		-0.19		-0.54		-0.11	

Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%.
Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends and family. Figures based on the sample aged 50 and over.

The results of the deterministic sensitivity analysis for depressive symptoms in the Well group are shown in Table 5.17. Under the four scenarios the conclusions about depressive symptoms trajectories of men and women from the Well group remained unchanged compared to those obtained from the original data. Some of the values of the bias suggested good potential accuracy.

Table 5.17 Deterministic sensitivity analysis of the logistic random intercept model for depressive symptoms with time, people from the Well group (imputed data)

	ELSA original		Scenario A		Scenario B		Scenario C		Scenario D	
	Coeff.	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***	**Coeff.**	***s.e.***
Gender	0.84	*0.13*	0.86	*0.13*	0.86	*0.13*	0.89	*0.13*	0.81	*0.13*
Bias			-0.02		-0.02		-0.05		0.03	
Wave 2	0.15	*0.12*	0.16	*0.12*	0.24	*0.12*	0.18	*0.12*	0.13	*0.12*
Bias			-0.01		-0.09		-0.03		0.02	
Wave 3	-0.16	*0.13*	-0.09	*0.12*	-0.13	*0.12*	-0.06	*0.13*	-0.14	*0.13*
Bias			-0.07		-0.03		-0.10		-0.02	
Wave2*Gender	-0.26	*0.16*	-0.26	*0.15*	-0.33	*0.15*	-0.33	*0.16*	-0.25	*0.15*
Bias			0.00		0.07		0.07		-0.01	
Wave3*Gender	0.10	*0.16*	-0.06	*0.16*	0.04	*0.15*	-0.06	*0.16*	0.06	*0.16*
Bias			0.16		0.06		0.16		0.04	

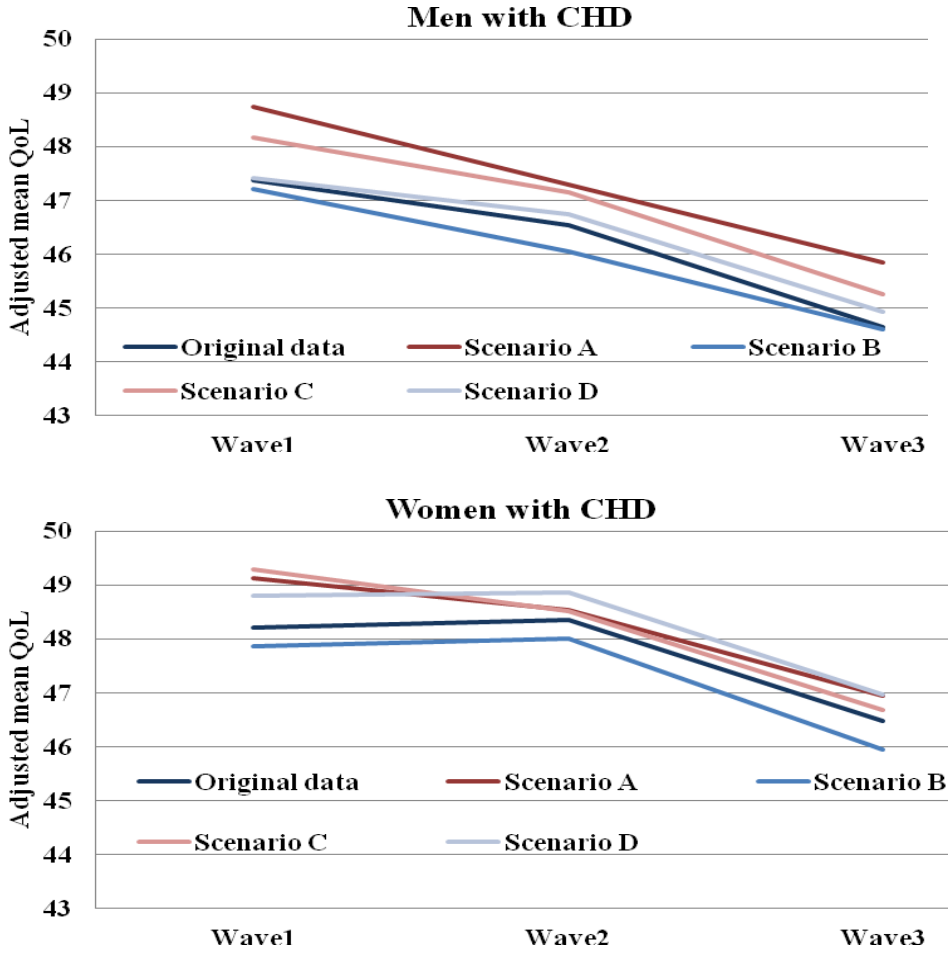
Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends and family. Figures based on the sample aged 50 and over.

5.3.3 Summary of results

The results of the sensitivity analysis of trajectories of quality of life presented in Tables 5.13 and 5.14 are summarised graphically in Figure 5.5 for the CHD group and 5.6 for the Well group. Regardless of the sensitivity, lower specificity (90%) of the CHD measure would lead to underestimated trajectories of quality of life of men and women with CHD. If the measure of CHD was assumed to have the lowest sensitivity (60%) and the highest specificity (95%), then the levels of quality of life of men with CHD would be overestimated at wave 2 and for women would be overestimated at each time. If the self-reported measure of CHD was assumed to have a sensitivity of 80% and a specificity of 95%, the trajectories of quality of life of men with CHD obtained from the original data would be almost the same. For women, the results of the original data would be underestimated but the shapes of trajectories would be the same.

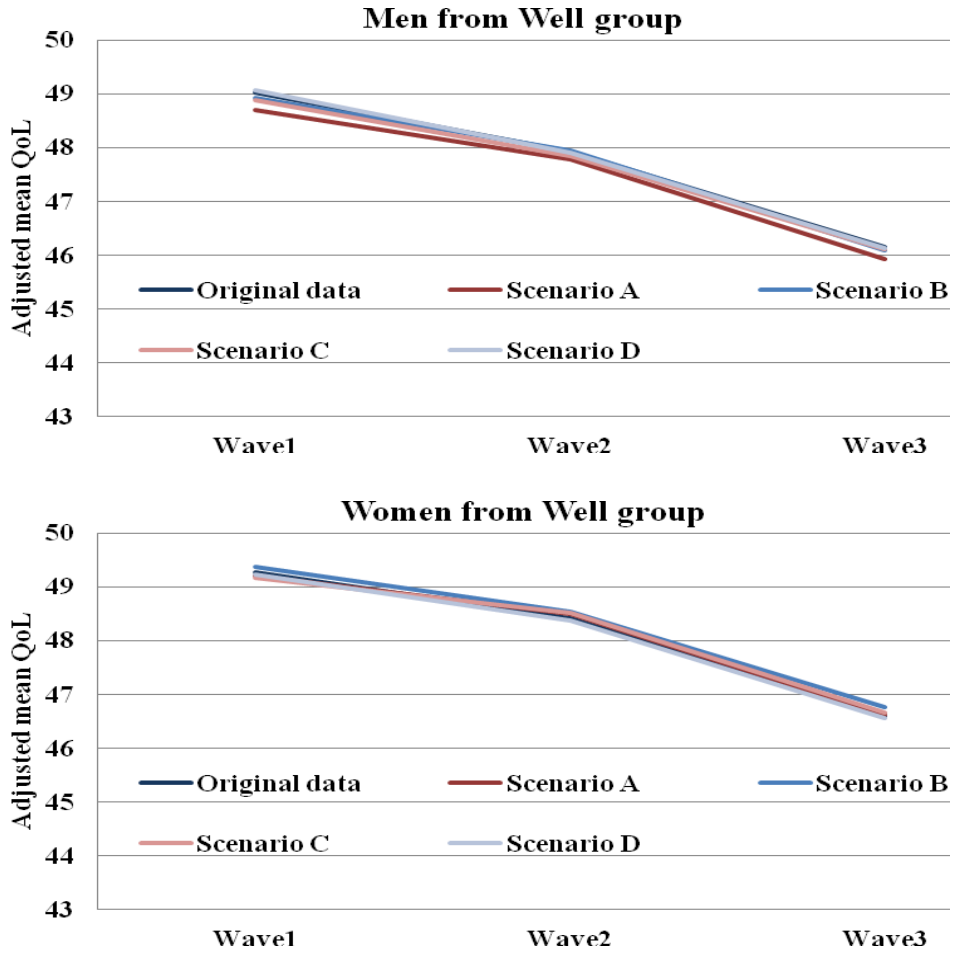
Assuming different values of the specificity and sensitivity led to negligible bias in the original results of trajectories of quality of life of men and women in the Well group (Figure 5.6). The only exception was when the lowest levels of sensitivity (scenario A and B) were assumed, which lead to a significant gender difference in baseline levels of quality of life of people from the Well group not found in the original data. If a high proportion of people with CHD were misreporting their disease status, women from the Well group would have had significantly higher quality of life than men at wave 1.

Figure 5.5 Trajectories over time of quality of life for men and women with CHD, comparing the four scenarios obtained from the deterministic sensitivity analysis (imputed data)



Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%.
Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends and family. Figures based on the sample aged 50 and over.

Figure 5.6 Trajectories over time of quality of life for men and women from the Well group, comparing the four scenarios obtained from the deterministic sensitivity analysis (imputed data)



Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%.
Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, depressive symptoms, positive support and number of close friends and family. Figures based on the sample aged 50 and over.

Results obtained in Tables 5.16 and 5.17 are presented graphically as odds ratios in figures 5.7 for the CHD group and figure 5.8 for the Well group. From the graphs it is easy to see that the trajectories of depressive symptoms obtained from the original data and those obtained under four scenarios led to the same conclusions for men with CHD, and men and women from the Well group. For women with CHD, under the four scenarios the results would be biased towards the null, as suggested by the non-significant change over time in the odds of having depressive symptoms.

Figure 5.7 Trajectories over time of depressive symptoms for men and women with CHD, comparing the four scenarios obtained from the deterministic sensitivity analysis, odds ratios (imputed data)

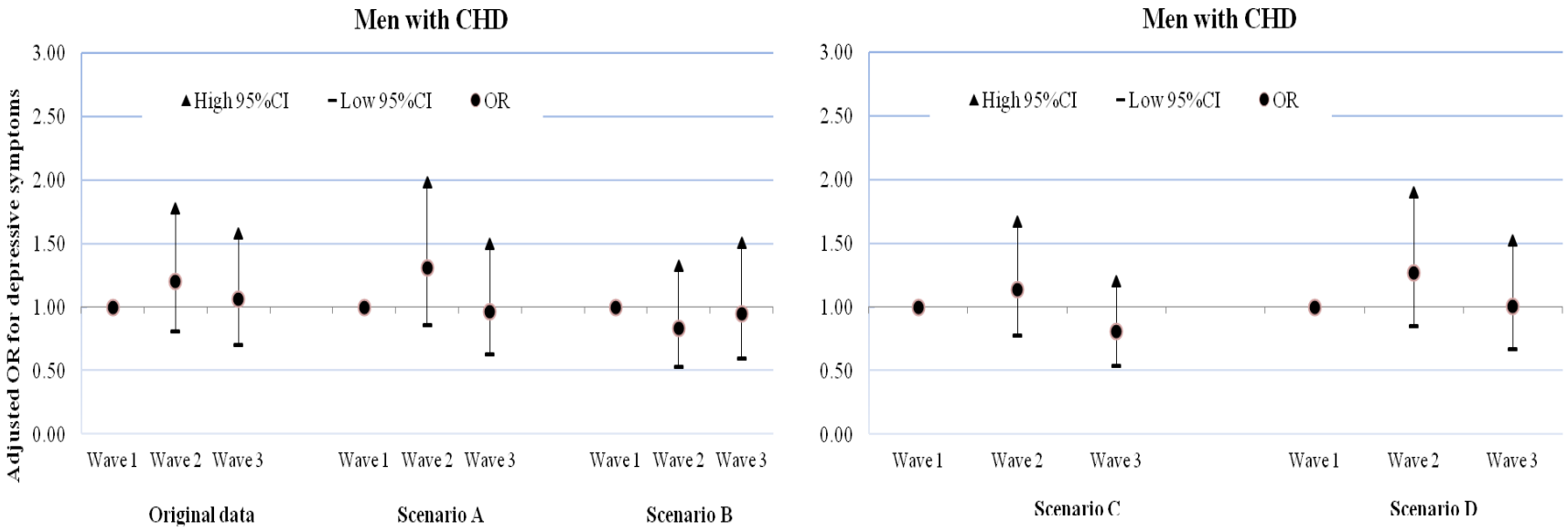
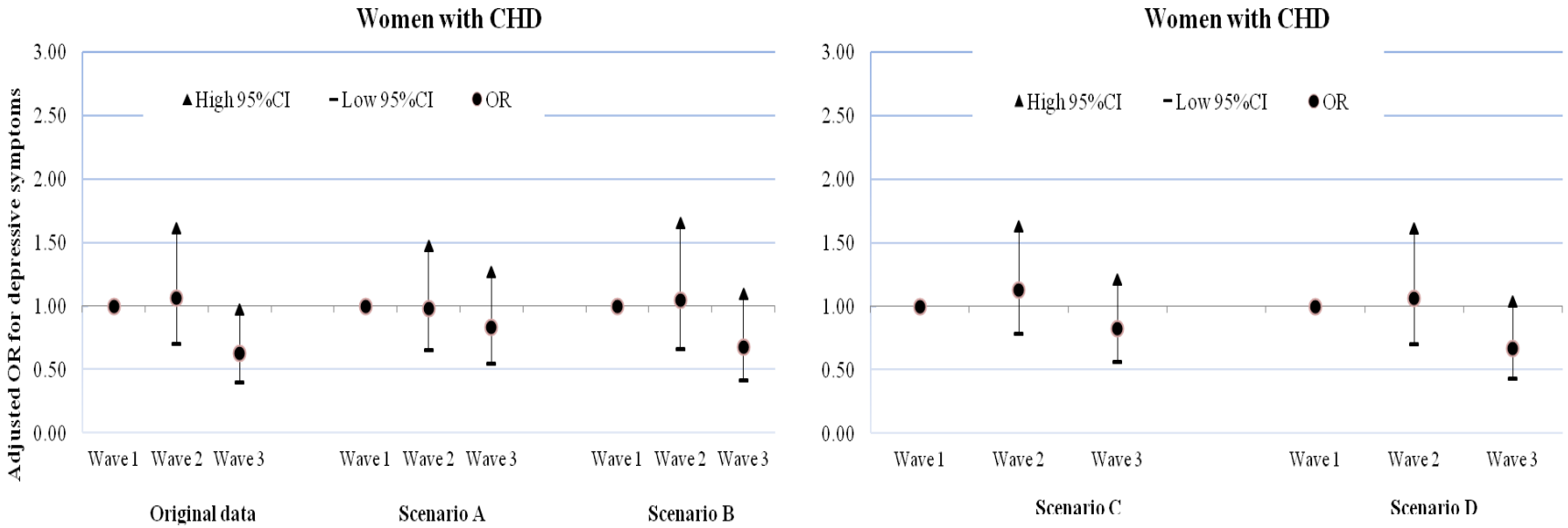


Figure 5.7 Continued



Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%.Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends and family. Figures based on the sample aged 50 and over. Results presented as odds ratios

Figure 5.8 Trajectories over time of depressive symptoms for men and women from the Well group, comparing the four scenarios obtained from the deterministic sensitivity analysis, odds ratios (imputed data)

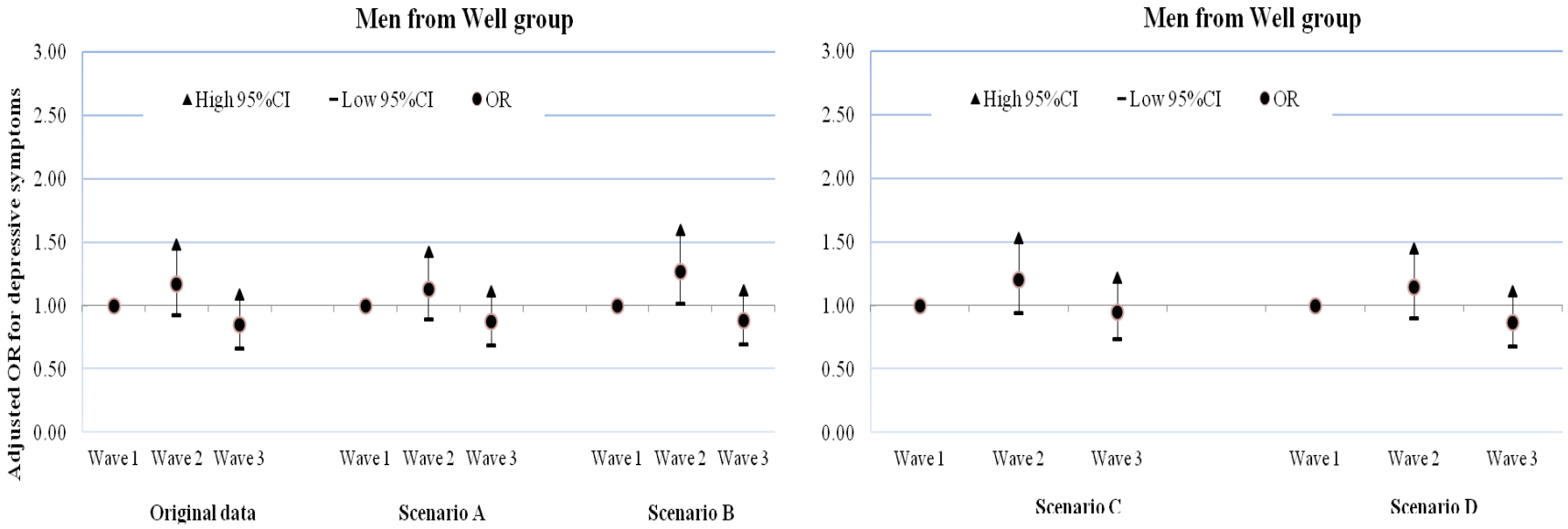
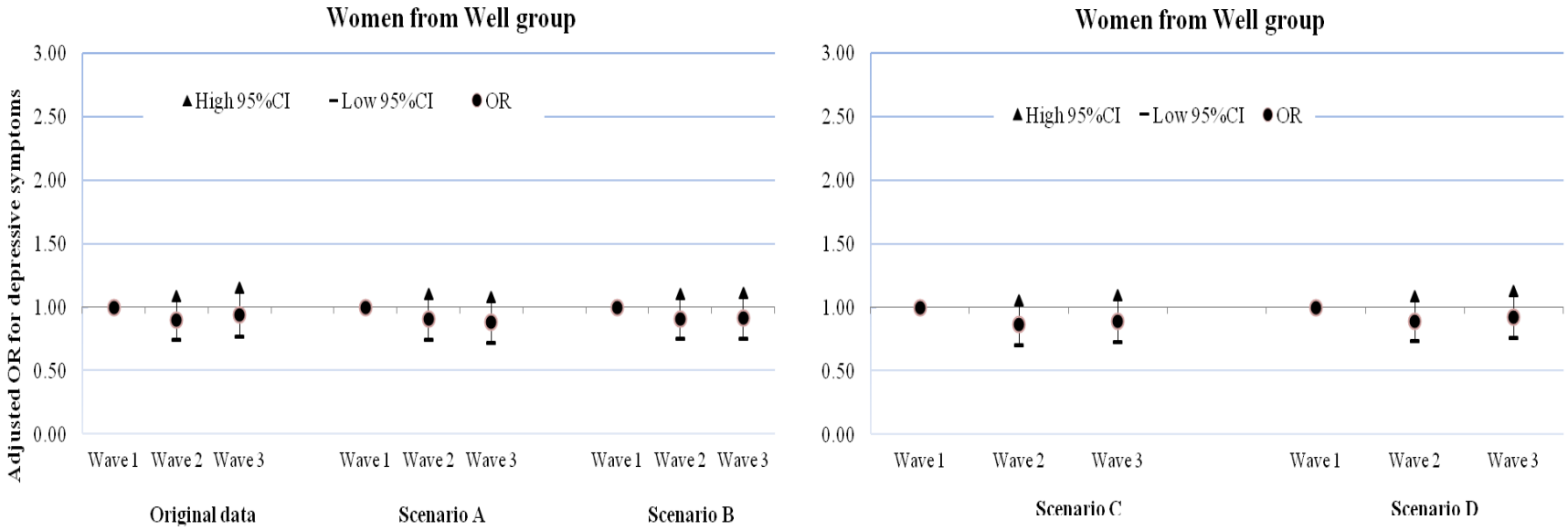


Figure 5.8-Continued



Scenario A: sensitivity 60% and specificity 90%. Scenario B: sensitivity 60% and specificity 95%. Scenario C: sensitivity 80% and specificity 90%. Scenario D: sensitivity 80% and specificity 95%. Results adjusted for gender, age, marital status, employment status, educational attainment, wealth, smoking status, alcohol consumption, physical activity, pain, ADLs, positive support and number of close friends and family. Figures based on the sample aged 50 and over. Results presented as odds ratios.

5.4 Discussion

Misclassification bias of self-reported measures can occur in epidemiological studies. Although previous studies have shown that the assessment of CHD by self-reports is a valid alternative when clinical assessment is not feasible, addressing quantitatively the effect of bias has been recommended (Jurek et al, 2006). The objective of this chapter was to follow this recommendation and quantify the bias that potential misclassification of self-reported CHD has on the results and learn how the observed findings may change by varying the assumed values of the misclassification parameters. It was hypothesised that the self-report measure of CHD is a reliable alternative when clinical assessment is not available and therefore the results would not draw different conclusions.

The main problem with the ELSA data is that it does not validate the self-reported measure of CHD using clinical screening or medical records. To date it has not been possible to link the survey data with hospital records statistics. In order to be able to quantify the misclassification bias of the self-reported CHD measure first a sensitivity analysis was conducted using an external validation study, the Whitehall II study. The study was chosen because it is similar to ELSA in terms of mode and year of data collection, and the age range of the sample. The Whitehall II study validated the self-reported measure of CHD with a clinical screening only for those with positive self-reported CHD. From this study the prevalence of false positive self-reports of CHD were estimated and these were applied randomly to the ELSA data in order to simulate two possible scenarios under which misclassification could have occurred. The first scenario assumed that the prevalence of false positives was a random sample of the whole population. The second scenario was more realistic and assumed that the prevalence of false positives was a random sample of the self-reported CHD population.

The prevalence of people with CHD was 17.2% in the original data. Assuming that falsely reporting CHD occurred randomly in the total population resulted in a lower prevalence of CHD (14.3% under first scenario) compared to the original. Assuming that the false positive group is a random sample of the CHD population led to a decrease of nearly 6 percentage points (11.1%) in the prevalence of CHD (second scenario). Results from the sensitivity analysis are summarised in Table 5.18. In general, results obtained under this first scenario and those from the original data led to the same conclusions about trajectories of quality of life of people with CHD and those

from the Well group: quality of life decreased over time in men from the CHD group and men and women from the Well group, while for women with CHD quality of life was stable between baseline and two year follow-up (wave 2) and then decreased. In terms of gender differences in quality of life, scenario 1 led to a type I error: women with CHD did not have higher quality of life than men at four year follow-up (wave 3).

Under scenario 2, type I errors (occurring when the null hypothesis is wrongly rejected) were introduced at two year follow-up (wave 2) in both quality of life and depressive symptoms: women did not have higher quality of life than men and also were not more likely than men to have depressive symptoms. The results from the second scenario suggest that levels of quality of life for men with CHD would be slightly underestimated at each wave (0.1 at baseline, 0.2 at wave 2 and 0.2 at wave 3) and a type I error would be introduced at two year follow-up, since the quality of life did not decrease significantly. Under scenario 2 a type I error could be introduced by misclassification of CHD for women. In fact women's quality of life would not change significantly at four year follow-up if 35% of people with CHD were misreporting their disease status. Results of quality of life of people from the Well group were unbiased, suggesting that misclassification does not affect the original results for this group. For depressive symptoms, results of men and women with CHD obtained under the first scenario were slightly biased. Assuming that 17% of people with CHD were wrongly classified as having CHD led to a type I error: men with CHD would not be at higher risk of having depressive symptoms than men from the Well group. The original estimates of the trajectories of depressive symptoms of people with CHD were slightly biased under scenario 2 and standard errors were larger, however conclusions remain unchanged. Results of trajectories of depressive symptoms for the Well group were robust as shown by the small values of bias obtained under the two scenarios and by the unchanged conclusions.

Table 5.18 Summary of findings of the sensitivity analysis based on the validation study

	Scenario 1	Scenario 2
Men with CHD	QoL: unbiased results. Depression: Type I error for the differences with "Well" men. Slightly biased results but same conclusions.	QoL: slightly underestimated levels at each wave. Type I error wave 2. Depression: Type I error for the differences with "Well" men. Biased estimates and large standard errors but same conclusions.
Women with CHD	QoL: underestimated levels at each wave, but same conclusions. Depression: slightly biased results but same conclusions.	QoL: underestimated levels at each wave. Type I error wave 4. Depression: biased estimates and large standard errors.
Gender differences in QoL CHD group	QoL: Type I error, women did not have higher QoL than men at wave 3.	QoL: Type I error, women did not have higher QoL than men at wave 2. Depression: Type I error, women did not have higher risk of depressive symptoms at baseline and wave2
Men-Well group	QoL: unbiased results. Depression: unbiased results.	QoL: unbiased results. Depression: unbiased results.
Women -Well group	QoL: unbiased results. Depression: unbiased results	QoL: unbiased results. Depression: unbiased results

Abbreviations: QoL quality of life
Scenario 1: False positives is a random sample of people of the whole population. Scenario 2: False positives is a random sample of people of the self-reported CHD population.

To summarise, if 17% of people with CHD were wrongly classified as having the disease, this would not have any impact in the results obtained from the original data for men with CHD, the only exception being that men with CHD would not be more likely than men from the Well group to have depressive symptoms. For women with CHD the levels of quality of life at wave 2 and wave 3 would be slightly higher than the original data where it was assumed that the self-reported measure of CHD was reliable.

On the contrary, if 35% of people with CHD were wrongly classified as having the disease, this would change the conclusions of the original results in that the quality of life levels of men and women with CHD at each wave were slightly underestimated. Also, men would not have decreased quality of life at wave 2 and women would have stable quality of life over time, contrary to what was found in the original results. Results for depressive symptoms would be slightly biased and variability around the

estimates would be greater than the original results. The main impact of the misclassification on the depressive symptoms results would be on the baseline difference in depressive symptoms between men and women with CHD and on the difference between men with CHD and men from the Well group.

The results of the Well group were not affected by the false positives.

The sensitivity analysis conducted using an external validation study is a novel approach, especially in longitudinal data analysis. The main advantage of using the Whitehall II study was that the measures used in the analysis were collected in the same way as in ELSA. Also the year of collection was the same and the age range of the samples were almost the same. Nevertheless, the use of the Whitehall II study for the validation study had some limitations. Participants of the Whitehall II study were white-collar civil servants and although at study entry they covered a wide range of grades, the "healthy worker effect" at baseline might have influenced the generalisability of any findings (Ferrie et al., 2009). A recent study on non-response and mortality in Whitehall II participants (Ferrie et al., 2009) showed that non-responders or partial responders had increased hazard for mortality compared to responders. Even after adjustment for age and education the overall prevalence of CHD in the Whitehall II study at phase 7 was about half of that of ELSA. This reflects the fact that the ELSA sample is a population-based sample which therefore includes blue-collar as well as white-collar workers. The extent to which the prevalence of false positive found in Whitehall II could be applied at the same level to ELSA is not known. It is also possible that the prevalence of false positives is positively related to the true prevalence of CHD and therefore might be higher in the ELSA study than the Whitehall II study. To some extent the second scenario could be seen as a more realistic scenario. Given the scarce availability of clinical screening and linkage to medical records in health surveys, this is the best that could have been done in order to explore the bias introduced by self-reported measure of CHD. A better way to have done the correction would have been to allow the probability of a false positive report of CHD for an individual to depend on age and education as predicted by the logistic regression for Whitehall II. However, due to the small number of false positives it was not possible to predict the probabilities using this approach.

A major limitation was that the Whitehall II study only validated positive cases of self-reported CHD, and it was not possible to ascertain the validity of negative cases of

CHD and calculate the specificity. It follows that the sensitivity analysis conducted in the first section of this chapter is a somewhat crude analysis. The ideal validation study to use for this sensitivity analysis would have been one that was performed on the ELSA sample, where the information provided by participants (or a subsample) was validated using hospital or physicians records. However, these data were not available and I am unaware of published papers or studies other than Whitehall II that reported appropriate data.

To address the impact that misclassification error had, not only on the false positive cases of CHD but also on the negative cases of CHD a deterministic sensitivity analysis was then conducted. The prevalence of people with CHD was 19.9% in the original data. Assuming that the sensitivity of the self-reported measure of CHD was 60% and the specificity was 90% (scenario A) led to a similar prevalence of CHD (19.2%). A lower prevalence of CHD (15.9%) was obtained under scenario B where the specificity was increased to 95% (and sensitivity was kept at 60%). When the sensitivity was assumed to be 80% and the specificity 90% (scenario C), the prevalence of CHD increased to 23.8% compared to that of the original data, while the prevalence of CHD was almost the same (20.2%) as the prevalence of the original data when the specificity was increased to 95% (and the sensitivity was kept at 80%, scenario D).

The results of the deterministic sensitivity analysis are summarised in Table 5.19. Regardless of the sensitivity, lower specificity (90%) of the CHD measure has been shown to lead to underestimated levels of quality of life at each time for men and at baseline for women with CHD. High specificity (95%) and low sensitivity (60%) of the CHD measure led to overestimated levels of quality of life at wave 2 for men and at each time for women with CHD. With both high sensitivity and specificity values, the trajectories of quality of life of men with CHD obtained from the original data could be slightly underestimated (quality of life levels would be lower at each wave) and misclassification would introduce a type I error (quality of life at wave 2 would not be significantly lower than baseline). For women, the results of the original data could be underestimated but the shapes of trajectories would be the same. Keeping the sensitivity constant to a high value, and varying the specificity values led to a type II error (which occurs when the null hypothesis is not rejected when it is in fact false) in the gender difference in quality of life at baseline among people with CHD: women had significantly higher quality of life than men.

Table 5.19 Summary of findings of the deterministic sensitivity analysis

Men with CHD		**Specificity**	
		Low	High
Sensitivity	Low	QoL: overestimated difference with men from Well group. Underestimated levels of QoL at each wave Depression: Type I error for the differences with men from Well group.	QoL: overestimated difference with men from Well group. Overestimated levels of QoL at wave 2 Depression: Type I error for the differences with men from Well group
	High	QoL: overestimated difference with men from Well group. Underestimated levels of QoL at each time Depression: Type I error for the differences with men from Well group Biased estimates but same conclusions	QoL: overestimated difference with men from Well group. Trajectories almost the same. Type I error QoL at wave 2. Depression: Type I error for the differences with men from Well group Biased estimates but same conclusions
Women with CHD		**Specificity**	
		Low	High
Sensitivity	Low	QoL: baseline level underestimated trajectories less steep. Depression: Type I error for decreased depression at wave 3	QoL: Overestimated levels of at each time, but same shape Depression: Type I error for decreased depression at wave 3
	High	QoL: baseline level underestimated trajectories less steep. Depression: Type I error for decreased depression at wave 3	QoL: Underestimated levels of at each time, but same shape Depression: Type I error for decreased depression at wave 3
Men and Women from the Well group		**Specificity**	
		Low	High
Sensitivity	Low	QoL and depression estimates slightly biased, but same conclusions.	QoL and depression estimates slightly biased, but same conclusions. Women's QoL at baseline: type II error
	High	QoL and depression estimates slightly biased, but same conclusions.	QoL and depression estimates slightly biased, but same conclusions. Women's QoL at baseline: type II error.

Abbreviations: QoL quality of life.
Sensitivity low 60%, high 80%. Specificity low 90%, high 95%.

Assuming different values of the specificity and sensitivity did not identify any bias in the original results of trajectories of quality of life of men and women in the Well group. The only exception was when the lowest level of sensitivity (regardless of the specificity) was assumed, which suggested that a type II error was introduced by misclassification (a significant gender difference in baseline levels of quality of life of people from the Well group that was not found in the original data).

The original results of trajectories of depressive symptoms were quite robust, as shown by the small bias obtained under the four scenarios; results led to same conclusions for men with CHD, and men and women from the Well group. For women with CHD type I error could be introduced by misclassification (under the four scenarios) as suggested by the non-significant change over time in the odds of having depressive symptoms.

The results of the deterministic sensitivity analysis showed that no specific values of sensitivity and specificity but rather a combination of both could introduce some bias in the results obtained from the original data (mainly those of the people with CHD) where it was assumed that the self-reported measure of CHD was correctly identifying people with and without the disease.

A major strength of the deterministic sensitivity analysis was that it gave an idea of how the results might change according to different assumptions about the sensitivity and specificity of the self-reported CHD measure. Researchers might use this approach when a validation study is not available. Educated values of sensitivity and specificity can be applied to the data and robustness of results can be assessed under different assumptions. A more sophisticated way of conducting this deterministic sensitivity analysis would have allowed the sensitivity and specificity to depend on age, gender and perhaps education. Some studies showed that agreement and/or sensitivity and specificity varied according to gender (Okura et al., 2004; Yamagishi et al., 2009), age (Kehoe et al., 1994; Okura et al., 2004; Yamagishi et al., 2009) and education (Okura et al., 2004). However, it was decided not to take this approach for two reasons: first, it was not possible to find from the literature a study that was similar to ELSA from which to base the sensitivity and specificity assumptions by age, gender and education; second, even in the simplest case of applying different levels of sensitivity and specificity to a sample divided in two age groups (younger vs older) would have

resulted in 8 different scenarios. In order to make the deterministic sensitivity analysis simpler and the interpretation of results straightforward it was assumed that misclassification occurred randomly in the whole population.

Researchers should address quantitatively the potential bias that a self-reported measure of disease could introduce in their results and conclusions, despite high levels of sensitivity and specificity. It was shown that even a simple sensitivity analysis could shed some light about the robustness of results based on a self-reported measure of CHD.

In conclusion, the sensitivity analyses reported in this chapter helped understand the impact that misclassification of the self-reported CHD measure could have on the conclusions presented in Chapter 4. Contrary to what was hypothesised, the reliability of the results presented in Chapter 4 could be affected by the misclassification of self-reported measure of CHD. If misclassification of any kind (false positives and a combination of both false positive and false negatives) occurred, then the result that men with CHD were at higher risk of having depressive symptoms compared to men from the Well group could not be replicated and therefore the robustness of the findings is compromised. High prevalence of false positives (35%) could compromise the robustness of the finding of decreased quality of life and lower risk of depressive symptoms at four year follow-up in women with CHD. Men with CHD would have stable quality of life at two year follow-up if high specificity (regardless of the sensitivity) and high prevalence of false positives (35%) occurred. Among people with CHD false positives (in both sensitivity analyses) led to non-significant gender differences in quality of life at four year follow-up and also to a non-significant gender difference in depressive symptoms at two year follow-up.

In the next chapter the findings from the thesis will be discussed as a whole.

Appendix 5

Table A5.1 Logistic regression of factors related to number of false positives in Whitehall II

	Odds Ratio	Std. Err.	P-value
Age	1.11	0.03	0.000
Female	0.67	0.25	0.274
Medium education	1.32	0.51	0.465
Low education	2.50	1.14	0.046

Age range 50 to 74

Appendix 5.1 Calculation of the prevalence of CHD according to scenario 1 and scenario 2

The number of people in ELSA population aged 50 to 74 is 9,347.

The first scenario assumes that the prevalence of false positives is a random sample of the total population. Therefore 1.2%, the adjusted prevalence of false positives obtained from the Whitehall II study, is applied to the total ELSA population as follows:

1.2% of 9,347=112.

In order to find the proportion of people with CHD that are misreporting and should be recoded into the healthy group, 112 is divided by 650 (which is the number of people with CHD) giving 0.17 (17%).

The second scenario assumes that the prevalence of false positives is a random sample of the CHD population. Therefore the proportion of people with CHD that are misreporting according to this definition is found as follows:

61/174=0.35 (35%)

where 61 is the adjusted number of false positives in Whitehall II and 174 is the adjusted number of people with CHD. 35% of people with CHD in the ELSA are then recoded into the healthy group.

Appendix 5.2 Stata syntax for generating the two scenarios of the sensitivity analysis based on the validation study

****The following was applied to each of the 5 imputed data sets separately *****

```
set seed 123456789
gen x=runiform()

gen scen1=0 if x>0.176
replace scen1=1 if x<=0.176
tab scen1

tab scen1 chd1

gen y=runiform()

gen scen2=0 if y>0.35
replace scen2=1 if y<=0.35
tab scen2

tab scen2 chd1



gen sc1chd1=chd1
replace sc1chd1=0 if chd1==1 & scen1==1
tab sc1chd1

gen sc1chd2=sc1chd1
gen sc1chd3=sc1chd1

gen sc2chd1=chd1
replace sc2chd1=0 if chd1==1 & scen2==1
tab sc2chd1
```

Appendix 5.3 Stata syntax for generating the four scenarios of the deterministic sensitivity analysis

****The following was applied to each of the 5 imputed data sets separately *****

```
set seed 123456789

gen x=runiform() if chd1==0

sum x

gen y=runiform() if chd1==1

sum y

gen sc1chd1=chd1
replace sc1chd1=1 if x<=0.1
replace sc1chd1=0 if y<=0.4
tab sc1chd1



gen j=runiform() if chd1==0

sum j

gen k=runiform() if chd1==1

sum k


gen sc2chd1=chd1
replace sc2chd1=1 if j<=0.05
replace sc2chd1=0 if k<=0.4
tab sc2chd1



gen w=runiform() if chd1==0

sum w

gen z=runiform() if chd1==1

sum z

gen sc3chd1=chd1
```

```
replace sc3chd1=1 if w<=0.1
replace sc3chd1=0 if z<=0.2
tab sc3chd1


gen u=runiform() if chd1==0

sum u

gen v=runiform() if chd1==1

sum v

gen sc4chd1=chd1
replace sc4chd1=1 if u<=0.05
replace sc4chd1=0 if v<=0.2
tab sc4chd1
```

Chapter 6: Discussion

This thesis focused on gender differences in quality of life and depressive symptoms among older people with coronary heart disease (CHD) and addressed common methodological problems of epidemiological studies, such as sources of error and uncertainty which may arise from missing data and self-reported CHD.

Gender differences in quality of life and depressive symptoms while living with angina or a history of myocardial infarction had not previously been researched systematically. Findings from previous studies were mixed, and were mainly based on samples from hospitals, with a short period of follow-up. One of the contributions to the literature made by this thesis is the use of a large national sample of non-institutionalized older people living in England, followed over four years.

Substantive results are discussed first, followed by methodological findings.

6.1 Summary of substantive findings

A cardiac event such as myocardial infarction or angina symptoms is a critical experience for an individual, with considerable impact upon mental health and quality of life. There has been a growing recognition of the importance of exploring the impact of coronary heart disease on patients' well-being, with an emphasis on possible gender differences. To date, findings from research on gender differences in quality of life and/or depression among people with coronary heart disease have mainly been based on small samples with short follow-up periods. Additionally, the majority of research has had a selection bias with regards to the focus on myocardial infarction and not angina. Therefore this study presents a unique opportunity to explore gender differences in quality of life and depressive symptoms among people with coronary heart disease (CHD), using a large longitudinal sample of older people living in England.

Based on findings from the literature, this study hypothesised that among people with CHD there would be significant gender differences in quality of life and depressive symptoms, with women being at higher risk than men of reporting depressive symptoms

and having lower quality of life. This study did not support this hypothesis for quality of life. Results showed no gender differences in quality of life at baseline (wave 1, 2002-03), while at two year follow-up (wave 2, 2004-05) and four year follow-up (wave 3, 2005-07) women had significantly higher quality of life than men. This was also true after adjusting for other covariates. The thesis did support the hypothesis of gender differences in depressive symptoms at baseline and at two year follow-up: it was found that women with CHD were at higher risk of having depressive symptoms than men with CHD, independent of other factors. Nevertheless, at four year follow-up there was no gender difference in depressive symptoms among men and women with CHD.

Findings from this work supported the hypothesis of differently shaped trajectories over time for quality of life and depressive symptoms among men and women with CHD. For men with CHD, quality of life decreased significantly over time. Quality of life of women with CHD was stable at two years after baseline, followed by a decline at four year follow-up. Trajectories in depressive symptoms were also different according to gender. The odds of having depressive symptoms did not change significantly over time for men. Among women with CHD, the odds of having depressive symptoms at two year follow-up were the same as at baseline, while at four year follow-up women were significantly less likely to report depressive symptoms than at baseline.

Current debates about well-being suggest that it is a multifaceted concept from which three aspects can generally be distinguished: evaluative well-being, experienced well-being and eudemonic well-being (Dolan, Layard, & Metcalfe 2011; Kahneman & Krueger, 2010). The measure of quality of life used in this study reflects evaluative well-being, since it involves global assessments of how people evaluate their lives. Depressive symptoms reflect experienced well-being as they reflect the experience of emotions, such as sadness and happiness. Results from this thesis highlight the difference between these two aspects of people's well-being and contribute to the current debate on the importance of measuring them separately to develop a broader appreciation of people's lives. By exploring both quality of life and depressive symptoms after the onset of CHD it was possible to untangle aspects of people's well-being never formally identified before.

This study found that the risk of depression was constant over time among men and an improvement in depressive symptoms was found in women, while quality of life decreased in both men and women in the long term. It is possible that depression

reflected the immediate psychological reaction to the cardiac event, while the long term decline in quality of life was the consequence of the burdens that the disease placed on the health and socioeconomic status of individuals. A CHD event often involves changes to an individual's lifestyles, therefore recovery from poor quality of life might require a long time, especially in those who as a result of the disease have experienced loss of control and autonomy. In this sample of older people, the long term decline in quality of life could also be a consequence of ageing, and not only the result of experiencing the disease. This can be supported by the results of this thesis of decline over time in quality of life among healthy individuals and it is consistent with previous studies reporting a trend of worsening quality of life over time especially at older ages (Zaninotto et al., 2009; Webb et al. 2010). On the other hand, the improvement in depression seen in women might reflect a process of adaptation to the disease. It has been shown that women have in general more coping strategies for stressful life events than men (Hobfoll et al., 1994). In particular, after experiencing myocardial infarction women are more likely than men to adopt problem-focused and emotion-focused coping strategies (Bogg et al., 2000). It is possible that in order to cope with CHD, women in this sample have adopted both problem-focused and emotion-focused strategies which helped to improve their mental health. It is also possible that women with depression were more likely to ask for help and be offered interventions by the general practitioner. It has been suggested that women are more likely than men to acknowledge depression and to seek for help from their primary care provider (Young et al., 1990). Either of these explanations show that the results of this thesis are consistent with previous findings; and also suggest that we might need to learn more about how women cope with CHD in order to help men adopt similar coping strategies for a long-term recovery in their quality of life.

Previous studies focussing on gender differences in quality of life and depression among people with CHD did not systematically compare them with a control group of healthy persons. This study addressed this limitation by comparing the results of the CHD group with those from a healthy control group. It was hypothesised that people aged fifty and over who had had a CHD event (first or recurrent) would be at higher risk of experiencing depressive symptoms and poor quality of life than those from a healthy group. It was found that, compared with "healthy" people, men and women with CHD had, on average, lower levels of quality of life. Men with CHD were also at higher risk of having depressive symptoms than men from the healthy group. The findings from

this thesis did not support the hypothesis that, compared to women from the healthy group, women with CHD were more likely to have depressive symptoms. Previous studies have shown that women are in general more prone to depression than men irrespective of the disease status (Forrester et al., 1992; Freasure-Smith et al., 1999; Mallik et al., 2006), which might explain the finding of this study that women with CHD were equally likely to have depressive symptoms as "healthy" women.

Among people from the healthy group, gender differences in depressive symptoms were found at each survey year; gender differences in quality of life were found at two year follow-up only, when women had higher quality of life than men. It is difficult to understand why gender differences in quality of life and depressive symptoms found in the CHD group were different from those found in the healthy group at four year follow-up, even after controlling for the same covariates. It is possible that some of the covariates and their change over time had a different impact on the outcomes considered according to disease status.

Men and women from the healthy group reported similar trajectories of quality of life and depressive symptoms: quality of life decreased over time, while the odds of having depressive symptoms did not change significantly over time for both men and women. People from this healthy group had levels of quality of life that were on average higher than the general ELSA population; therefore, the decline in quality of life levels may have been a consequence of ageing or of development or progression of other disease. Stable levels of depressive symptoms over time could reflect the relatively good health of these individuals. It is also possible that individuals that are healthy are in general more resilient to stressors of life and therefore their mental health is maintained as they age.

As seen for people in the CHD group, in this healthy population the results based on quality of life and those based on depressive symptoms go in different directions, strengthening the case for treating these two outcomes as distinct aspects of people's well-being.

6.2 Summary of methodological findings

6.2.1 Missing data

Analyses of longitudinal data offer powerful and insightful approaches to understanding changes over time in a certain outcome, and what might be driving that change. However, researchers using longitudinal data are faced with some methodological problems. The most common problem of epidemiological studies is non-response. In addition to item non-response, longitudinal studies also face attrition due to death or drop-out from the study.

The longitudinal data used in this thesis had over 50% of information missing that had accumulated over three waves due to item non-response and attrition. The data set therefore had incomplete time-dependent outcomes (one continuous and one binary) and incomplete time-dependent and time-independent covariates (of different types). One of the objectives of this thesis was to compare different techniques for dealing with missing data in longitudinal studies using full information maximum likelihood, multivariate normal imputation and two-fold fully conditional specification, in order to find the best model that yields unbiased results when applied to the data.

A simulation study was set up to compare the performance of the three techniques for dealing with missing data, in order to investigate which technique was most suitable with this data structure. A maximum likelihood based method, full information maximum likelihood, was compared with two multiple imputation techniques: multivariate normal imputation and the two-fold fully conditional specification. Comparisons among the performance of each missing data technique in recovering parameter estimates of the continuous outcome (quality of life) and binary outcome (depressive symptoms) appeared to draw different conclusions on which of the three methodsmissing for dealing with missing data was most suitable for the data structure used in this thesis. Results showed that for the continuous outcome, the three missing data techniques performed equally well in recovering the targeted parameters. However, the two multiple imputation techniques performed better than the maximum likelihood based method in recovering the targeted parameter of the interaction term. This finding most probably reflected one of the advantages of using a multiple imputation method,

over full information maximum likelihood, where the interaction term can be accommodated during the imputation stage.

For the binary outcome the two-fold fully conditional specification technique outperformed the full information maximum likelihood and multivariate normal imputation techniques. The two-fold fully conditional specification method performed exceptionally well in recovering targeted parameters with good accuracy and precision. The multivariate normal imputation technique lacked precision in recovering some of the targeted parameters, especially in the models where there was an interaction term with time. However, this method performed better than full information maximum likelihood which did not recover the targeted parameters in the model for the healthy population with good accuracy and precision. These results suggested that both methods that assume a joint multivariate normal distribution do not perform well with a binary outcome.

Based on the results of the simulation study the two-fold fully conditional specification technique was then used to impute missing data for the substantive analyses, and proved to be particularly suitable for repeated measures. Results for depressive symptoms based on augmented samples were similar to those obtained from observed data (with missing data).

More differences were found in the results for quality of life. For example, results obtained from observed data did not show differences in quality of life among people with CHD compared with people from the healthy population. Also, among men with CHD quality of life did not decrease significantly at two year follow-up in the observed data. This difference between analyses based on the augmented samples and those based on the sample with missing data in the quality of life results most probably reflects the fact that the missing data mechanism was not missing completely at random. Missingness in the quality of life measure depended on observed characteristics such as increasing age, CHD, low education, poor wealth, not cohabiting with a partner, permanently unable to work, not currently in paid employment and looking after home or family, and being physically inactive. Those reporting these characteristics are more likely to drop-out and to report a lower quality of life than those who were more advantaged (i.e. those in the healthy group, with high education, cohabiting with a partner and so forth).

These results supported the hypothesis that ignoring missing data would give biased results, especially when the missing data mechanism is not missing completely at random.

6.2.2 Self-reported measure of CHD

This thesis used a self-reported measure of CHD which was not validated against medical records or clinical assessment. Although previous studies have shown that the assessment of myocardial infarction by self-reports is a valid alternative when clinical assessment is not feasible, it has been recommended to address quantitatively the effect of bias (Jurek et al, 2006). One of the objectives of this study was to follow this recommendation and quantify the bias that potential misclassification of self-reported CHD has on the substantive results and learn how these may change by varying the assumed values of the misclassification parameters. It was hypothesised that the self-report measure of CHD is a reliable alternative when clinical assessment is not available and therefore the results would not draw different conclusions.

A quantitative sensitivity analysis was done using an external study, the Whitehall II study, which collected a self-reported measure of CHD and validated it with a clinical screening. From this study the prevalence of false positives was estimated and then applied to the ELSA data in order to assess the impact of the bias on the estimates obtained from analysis on gender differences in quality of life and depressive symptoms. Two possible scenarios under which misclassification could have occurred were considered. The first scenario assumed that the prevalence of false positives is a random sample of the whole population. The second scenario was more realistic and assumed that the prevalence of false positives is a random sample of the self-reported CHD population.

In general, results obtained under the first scenario and those from the original data led to the same conclusions about trajectories of quality of life of people with CHD: quality of life decreased over time in men from the CHD group, while for women with CHD quality of life was stable between baseline and two year follow-up and then decreased. In terms of gender differences in quality of life, this scenario led to a type I error:

women with CHD did not have higher quality of life than men at four year follow-up. A type I error could also be introduced under this scenario for the results on depressive symptoms: men with CHD would not be at higher risk of having depressive symptoms than men from the healthy population.

The results from the second scenario suggested that levels of quality of life for men with CHD would be slightly underestimated at each survey year and a type I error would be introduced at two year follow-up, since the quality of life of men with CHD did not decrease significantly. Under this scenario a type I error could also be introduced by misclassification of CHD for women. In fact women's quality of life would not change significantly at four year follow-up. The original estimates of the trajectories of depressive symptoms of people with CHD were slightly biased under the second scenario and standard errors were larger. Under this scenario women with CHD were not at higher risk of having depressive symptoms compared to men at any survey year, a result that contradicted the original findings.

The sensitivity analysis showed that the results of the original data were quite robust for men and women from the healthy group. It could be concluded that the results obtained for this group were not affected by false positive cases of CHD.

It was only when the analysis was completed that I realised that not being able to assess the impact of misclassification due to false negatives (as well as false positives) was a main limitation. In the Whitehall II data only positive self-reported cases of CHD could be verified. Follow-up of clinical records and validation using clinically verified events was only carried out for the subset of Whitehall II participants in whom there was a suggestion of a CHD event. Therefore sensitivity and specificity could not be calculated. For that reason in I decided to conduct also a deterministic sensitivity analysis.

From the deterministic sensitivity analysis, it can be concluded that regardless of the sensitivity, lower specificity of the CHD measure has been shown to lead to underestimated levels of quality of life at each time-point for men and at baseline for women with CHD. High specificity and low sensitivity of the CHD measure led to overestimated levels of quality of life at two year follow-up for men and at each time for women with CHD. With both high sensitivity and specificity values, the trajectories of

quality of life of men with CHD obtained from the original data could be slightly under estimated (quality of life levels would be lower at each time) and misclassification would introduce a type I error (quality of life at wave 2 would not be significantly lower than baseline). For women, the results of the original data could be underestimated but the shapes of trajectories remain unchanged.

The original results of trajectories of depressive symptoms were quite robust, as shown by the small bias obtained under the four scenarios; results led to same conclusions for men with CHD. For women with CHD, a type I error could be introduced by misclassification (under the four scenarios) as suggested by the non-significant change over time in the odds of having depressive symptoms.

Assuming different values of the specificity and sensitivity did not identify any bias in the original results of trajectories of quality of life and depressive symptoms of men and women in the healthy population.

Contrary to what was hypothesised, the reliability of the results of this thesis could be affected by the misclassification of the self-reported measure of CHD. Unfortunately, the sensitivity analyses conducted did not give a clear cut answer on which values of false positives or false negatives could bias the results. Therefore specific recommendations about the use of the self-reported measure of CHD cannot be made based on the sensitivity analyses reported here. Rather the results of the sensitivity analyses presented in this thesis highlighted the fact that using a self-reported measure of disease is something that requires more attention and it is therefore important to address quantitatively the effect of any possible bias introduced by its use.

6.3 Implications

The findings from this thesis can be used to inform caregivers that after CHD, deterioration in quality of life among women may not occur in the immediate time following the CHD event. On the other hand, men seem to be less able to cope with the disease in the long term, in terms of their quality of life. Quality of life in men should be monitored in the years following the event in order to reduce the risk of long term deterioration.

Results from this thesis suggest that the mental health of men who have experienced a CHD event does not necessarily deteriorate over time, and the mental health of women could possibly improve in the long term. Caregivers could advise patients and their immediate relatives on effective strategies for coping with the cardiac event to help maintain good mental health. Especially for men they could advise adopting problem-focused coping strategies which may be more useful for long-term coping with lifestyle changes caused by the disease.

This research also has implications for the academic community. Researchers that wish to explore well-being are encouraged to measure separately evaluative and experience components of well-being. Researchers wishing to measure quality of life and its relationships with health, among older people, should consider the use of a measure that is independent of health and other factors that might influence it, such as CASP. This measure has been specifically developed for old age and it is based on the idea that quality of life should stand alone from the factors or influences that shape it. It is probably this characteristic of the measure that, in this thesis, allowed discovering trajectories of quality of life among people with CHD never reported before.

The results from the simulation study for comparing techniques to deal with missing data can inform researchers that multiple imputation methods perform better than maximum likelihood based methods in the presence of a binary outcome and interaction terms. The two-fold fully conditional specification method is particularly recommended with repeated measures and in the presence of variables of different nature. Although this method is less feasible than the other two techniques in the presence of many waves, it proved to be invaluable when used to impute missing data for the substantive analyses. Researchers wishing to address the problem of missing data in longitudinal studies can be reassured that using this method will improve the validity of their results and reduce estimation bias caused by missing data. Recent advances in the use of this technique for missing data include the implementation of an algorithm in Stata, which will help researchers without technical expertise to apply this method easily. The Stata code will be available in the near future (Irene Petersen, personal communication).

In terms of the self-reported measure of CHD, the results from the sensitivity analysis showed that results could lead to different conclusions when different values of

sensitivity and specificity are assumed. Therefore researchers using a self-reported measure of disease should address quantitatively the potential bias that such measures could introduce to their results and conclusions. It was shown that even a simple sensitivity analysis could shed some light about the robustness of results based on a self-reported measure of CHD.

Results from the sensitivity analysis do not suggest that the use of self-reported measures of disease inevitably introduce misclassification bias and should not discourage researchers from using survey data for medical and epidemiological research. Rather results suggested that perceived disease status could be of importance in itself.

6.4 Strengths and limitations

One of the strengths of this thesis included the use of a large sample of older people living in private households in England. The study has been designed to collect information on topics necessary to understand the economic, social, psychological and health elements of the ageing process. Some of the advantages of using this data set include: the study collects information on angina as well as myocardial infarction; a measure of depressive symptoms as well as a measure of quality of life specifically designed for older people (CASP); and the possibility to compare the results for the CHD population with those for a healthy population.

A further strength of this study was the possibility to adjust analyses for a wide range of covariates relevant to the outcomes and exposure.

A major strength of the simulation study was that, to date and to my knowledge, for the first time techniques for dealing with missing data in the presence of both continuous and binary outcomes were simultaneously compared. Furthermore, missing data (in the outcomes as well as covariates) due not only to item non-response but also to drop-out was addressed. This was also the first study to apply the recently proposed two-fold fully conditional specification to longitudinal data from a national survey and compared it with full information maximum likelihood and multivariate normal imputation.

The sensitivity analysis based on the external validation study to test misclassification due to false positive cases of CHD had the strength of being a novel approach, especially in the context of longitudinal data analysis. A strength of deterministic sensitivity analysis was that it helped understand how the results might change according to different assumptions about the sensitivity and specificity of the self-reported CHD measure.

However, some limitations of this thesis should also be acknowledged. First, in the missing data analysis missing at random was assumed. This assumption is often difficult to verify. To reinforce the missing at random assumption and to reduce the bias caused by missing data, auxiliary variables predictive of missingness were included in the imputation stage.

Another limitation of the missing data analysis was that in dealing with attrition, no distinction was made between drop-out due to death and other reasons for loss to follow-up. By treating death and attrition as the same form of drop-out, it was assumed that trajectories continue beyond death and therefore that the panel was immortal. In the context of this research it seemed appropriate to consider the possible outcome a participant could have had if he or she had not died.

A further limitation of this study is that no information about depression and quality of life was available before the baseline interview. It is not known whether poor quality of life and increased risk of depressive symptoms began before the coronary heart event. This is an important issue since recent reviews of the literature (Frasure-Smith 2005, 2006) showed that there is evidence of both an etiologic role for depression in CHD (that is, depression preceding development of CHD) and a prognostic role for depression in CHD (that is, depression predicting prognosis in established CHD). It could be possible that those who were depressed before the baseline interview were more likely to suffer a CHD event, and it is also possible that those who were depressed were less likely to take part in the study. If people with CHD had low pre-CHD quality of life levels and were depressed then attributing low levels of quality of life and high risk of depressive symptoms to a recent CHD event would be erroneous.

Lastly, due to lack of power it has not been possible to consider angina and myocardial infarction as two separate exposures. Angina is the predominant presentation of CHD among women while myocardial infarction is predominant among males. Therefore considering these two exposures together might have resulted in underestimated gender differences in quality of life and depressive symptoms.

6.5 Future research

During the work on this doctoral thesis some new questions on the methods used to adjust for missing data were raised. Attrition and mortality are intrinsically related to many ageing-related changes in longitudinal studies. Therefore the analysis and interpretation of results should consider the bias introduced by these two different forms of missingness. This is an area of research that is still under development, although several methods have been proposed.

In future research I would like to expand on the problem of missing data to distinguish between attrition and mortality, and possibly incorporate missingness due to item non-response. Another question that was raised during the simulation study for missing data was about the plausibility of the missing at random assumption. It would be interesting in future research to investigate the robustness of key inferences to possible departures from the missing at random assumption, by doing a sensitivity analysis assuming a missing not at random mechanism.

Adjustment for socio-demographic factors, health behaviour factors, health factors and social factors were also shown to be important in the analysis presented in this thesis. The next step for this work is to investigate the effect of independent variables using mixture modelling. Under this approach, individuals are classified into subpopulations based on heterogeneity in the data. Individuals who are similar along a dimension of interest (e.g., health behaviour) are grouped into the same class, and those who are different are grouped into different classes.

The sensitivity analysis of self-reported CHD could be extended when the link between the ELSA data and the Hospital Episode Statistics has been achieved. A small validation study could be performed on the ELSA sample using hospital records.

6.6 Conclusions

The main conclusions that can be drawn from this study are:

Men with CHD were at higher risk of having poor quality of life and depressive symptoms than men from the healthy population. Women with CHD had lower levels of quality of life than women from the healthy population, but they were not more likely to have depressive symptoms.

Different shapes of trajectories of quality of life and depressive symptoms among men and women aged fifty and over who had experienced CHD were found. At baseline, following the experience of CHD, men and women had same levels of quality of life. However, while for men quality of life decreased significantly over time, the quality of life of women with CHD was stable at two years after baseline, followed by a decline at four year follow-up. In terms of depressive symptoms, men did not report any change in their mental health over time. Women at two year follow-up were as likely as at baseline to have depressive symptoms, but at four year follow-up were significantly less likely to report depressive symptoms compared with baseline.

Women were more likely to have depressive symptoms than men at baseline at two year follow-up only. No significant gender differences in quality of life were found at

baseline, while at two year follow-up and at four year follow-up women had higher quality of life than men.

Women with CHD did not report similar trajectories of quality of life and depressive symptoms as men with CHD and men and women from the healthy group. Gender differences in quality of life and depressive symptoms found in the CHD group at year four of follow-up were not found in the healthy population, even after adjusting for the same covariates.

Ignoring missing data would have given biased results, especially for quality of life where a large amount of data was missing.

Reliability of the results could be affected by misclassification bias in the self-reported measure of CHD.

At this culmination of my PhD thesis, the following quote seems to capture the view I now hold about the research process:

"The more we learn about the world, and the deeper our learning, the more conscious, specific, and articulate will be our knowledge of what we do not know; our knowledge of our ignorance. For this indeed, is the main source of our ignorance - the fact that our knowledge can be only finite, while our ignorance must necessarily be infinite." (Popper, 2002:38)

This work is just the beginning of my knowledge on this topic - it can never be the end.

References

Acock, A.C. 2005. Working with missing values. *Journal of Marriage and Family*, 67, (4) 1012-1028

Altman, D.G., & Bland, J.M. 1994. Diagnostic tests 1: sensitivity and specificity. *BMJ* 308, 1552.

Angold, A., Costello, E.J., & Worthman, C.M. 1998. Puberty and depression: the roles of age, pubertal status and pubertal timing. *Psychol.Med.*, 28, (1) 51-61

Arber, S. & Cooper, H. 1999. Gender differences in health in later life: the new paradox? *Social Science & Medicine*, 48, (1) 61-76

Armitage P, Berry G, Matthews JNS. 2002. Statistical methods in medical research. 4th edition. Malden, MA: Blackwell Science.

Artbuckle J.L. 1996, "Full information estimation in the presence of incomplete data ," G. A. Marcoulides & R. E. Schumacker, eds., Mahwah, N.J: L. Erlbaum Associates, pp. 243-277.

Bajekal, M., Primatesta, P., & Prior, G. 2001. *Health Survey for England 2001* London, The Stationery Office.

Baldwin, R.C., Anderson, D., Black, S., Evans, S., Jones, R., Wilson, K., & Iliffe, S. 2003. Guideline for the management of late-life depression in primary care. *Int.J.Geriatr.Psychiatry*, 18, (9) 829-838

Banks, J., Breeze, E., Lessof, C., & Nazroo, J. 2006. *Retirement, health and relationships of the older population in England: ELSA 2004 (Wave 2)* London, The Institute for Fiscal Study.

Banks, J., Breeze, E., Lessof, C., & Nazroo, J. 2008. *Living in the 21st century: older people in England ELSA 2006 (Wave 3)* London, The Institute for Fiscal Study.

Baraldi, A.N. & Enders, C.K. 2010. An introduction to modern missing data analyses. *J.Sch Psychol.*, 48, (1) 5-37

Barbareschi, G., Sanderman, R., Kempen, GI, & Ranchor, A.V. 2009. Socioeconomic status and the course of quality of life in older patients with coronary heart disease. *Int.J.Behav.Med.*, 16, (3) 197-204

Barefoot, J.C., Mortensen, E.L., Helms, M.J., Avlund, K., & Schroll, M. 2001. A longitudinal study of gender differences in depressive symptoms from age 50 to 80. *Psychology and Aging*, 16, (2) 342-345

Barr, E.L., Tonkin, A.M., Welborn, T.A., & Shaw, J.E. 2009. Validity of self-reported cardiovascular disease events in comparison to medical record adjudication and a statewide hospital morbidity database: the AusDiab study. *Intern.Med.J.*, 39, (1) 49-53

Baumeister, H., Kriston, L., Bengel, J., & Harter, M. 2010. High agreement of self-report and physician-diagnosed somatic conditions yields limited bias in examining mental-physical comorbidity. *J.Clin.Epidemiol.*, 63, (5) 558-565

Baumeister, H., Kriston, L., Bengel, J., & Harter, M. 2010. High agreement of self-report and physician-diagnosed somatic conditions yields limited bias in examining mental-physical comorbidity. *J.Clin.Epidemiol.*, 63, (5) 558-565

Beekman, A.T., Deeg, D.J., van Tilburg, T., Smit, J.H., Hooijer, C., & van Tilburg, W. 1995. Major and minor depression in later life: a study of prevalence and risk factors. *J.Affect.Disord.*, 36, (1-2) 65-75

Beekman, A.T., Deeg, D.J., Braam, A.W., Smit, J.H., & van Tilburg, W. 1997. Consequences of major and minor depression in later life: a study of disability, well-being and service utilization. *Psychol.Med.*, 27, (6) 1397-1409

Beekman, A.T., Copeland, J.R., & Prince, M.J. 1999. Review of community prevalence of depression in later life. *Br.J.Psychiatry*, 174, 307-311

Beekman, A.T., Geerlings, S.W., Deeg, D.J., Smit, J.H., Schoevers, R.S., de Beurs, E., Braam, A.W., Penninx, B.W., & van Tilburg, W. 2002. The natural history of late-life depression: a 6-year prospective study in the community. *Arch.Gen.Psychiatry*, 59, (7) 605-611

Bize, R., Johnson, J.A., & Plotnikoff, R.C. 2007. Physical activity level and health-related quality of life in the general adult population: a systematic review. *Prev.Med.*, 45, (6) 401-415

Bjerkeset, O., Nordahl, H.M., Mykletun, A., Holmen, J., & Dahl, A.A. 2005. Anxiety and depression following myocardial infarction: gender differences in a 5-year prospective study. *J.Psychosom.Res.*, 58, (2) 153-161

Bogg, J., Thornton E.., & Bundred, P. 2000. Gender variability in mood, quality of life and coping following primary myocardial infarction. *coronary health care*, 4, 163-168

Bowling Ann 2005. *Ageing well: quality of life in old age*, first ed. Maidenhead, Open University Press.

Brand, J.P.L. 1999. Development, implementation and evaluation of multiple imputation strategies for the statistical analysis of incomplete data sets. Thesis University of Rotterdam/TNO Prevention and Health. http://repub.eur.nl/res/pub/19790/990408_BRAND,%2520Jacob%2520Pieter%2520Laurens.pdf accessed November 2008.

Brezinka, V. & Kittel, F. 1996. Psychosocial factors of coronary heart disease in women: a review. *Soc.Sci.Med.*, 42, (10) 1351-1365

Brink, E., Grankvist, G., Karlson, B.W., & Hallberg, L.R. 2005. Health-related quality of life in women and men one year after acute myocardial infarction. *Qual.Life Res.*, 14, (3) 749-757

Britton, A. & Shipley, M.J. 2010. Bored to death? *Int.J.Epidemiol.*, 39, (2) 370-371

Brown, N., Melville, M., Gray, D., Young, T., Munro, J., Skene, A.M., & Hampton, J.R. 1999. Quality of life four years after acute myocardial infarction: short form 36 scores compared with a normal population. *Heart*, 81, (4) 352-358

Brown, W.J., Ford, J.H., Burton, N.W., Marshall, A.L., & Dobson, A.J. 2005. Prospective study of physical activity and depressive symptoms in middle-aged women. *Am.J.Prev.Med.*, 29, (4) 265-272

Buhi, E.R., Goodson, P., & Neilands, T.B. 2008. Out of sight, not out of mind: Strategies for handling missing data. *American Journal of Health Behavior*, 32, (1) 83-92

Bunker, S.J., Colquhoun, D.M., Esler, M.D., Hickie, I.B., Hunt, D., Jelinek, V.M., Oldenburg, B.F., Peach, H.G., Ruth, D., Tennant, C.C., & Tonkin, A.M. 2003. "Stress" and coronary heart disease: psychosocial risk factors National Heart Foundation of Australia position statement update. *Medical Journal of Australia*, 178, (6) 272

Burton, A., Altman, D.G., Royston, P., Holder, R.L. 2006.The design of simulation studies in medical statistics. *Statistics in Medicine* Dec 30;25(24):4279-92.

Bush, D.E., Ziegelstein, R.C., Tayback, M., Richter, D., Stevens, S., Zahalsky, H., & Fauerbach, J.A. 2001. Even minimal symptoms of depression increase mortality risk after acute myocardial infarction. *Am.J.Cardiol.*, 88, (4) 337-341

Bush, T.L., Miller, S.R., Golden, A.L., & Hale, W.E. 1989. Self-report and medical record report agreement of selected medical conditions in the elderly. *Am.J.Public Health*, 79, (11) 1554-1556

Carlin, J.B., Galati, J.C., & Royston, P. 2008. A new framework for managing and analyzing multiply imputed data in Stata. *Stata Journal*, 8, (1) 49-67

Carney, R.M., Blumenthal, J.A., Catellier, D., Freedland, K.E., Berkman, L.F., Watkins, L.L., Czajkowski, S.M., Hayano, J., & Jaffe, A.S. 2003. Depression as a risk factor for mortality after acute myocardial infarction. *Am.J.Cardiol.*, 92, (11) 1277-1281

Carpenter, J.R., & Goldstein, H., 2004. Multiple Imputation using MLwiN. *Multilevel Modelling Newsletter*. 4:9-18.

Carpenter, J.R., Goldstein, H., & Kenward, M.G. 2011. REALCOM-IMPUTE Software for Multilevel Multiple Imputation with Mixed Response Types. *Journal of Statistical Software*, 45, (5) 1-14

Carr, A.J. & Higginson, I.J. 2001. Are quality of life measures patient centred? *BMJ*, 322, (7298) 1357-1360

Chan, A.W. & Altman, D.G. 2005. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet*, 365, (9465) 1159-1162

Chang, C.C., Yang, H.C., Tang, G., & Ganguli, M. 2009. Minimizing attrition bias: a longitudinal study of depressive symptoms in an elderly cohort. *Int.Psychogeriatr.*, 21, (5) 869-878

Charney, P., Walsh, J., & Nattinger, A.B. 1999. Update in women's health. *Ann.Intern.Med.*, 131, (12) 952-958

Chu, H.T., Wang, Z.J., Cole, S.R., & Greenland, S. 2006. Sensitivity analysis of misclassification: A graphical and a Bayesian approach. *Annals of Epidemiology*, 16, (11) 834-841

Clarke, P., Hardy, R. 2007. Methods for Handling Missing Data. In: Pickles A, Maughan B, Wadsworth M, editors. Epidemiological Methods in Life Course Research. Oxford University Press; pp. 157-79

Cole, M.G. & Dendukuri, N. 2003. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. *Am.J.Psychiatry*, 160, (6) 1147-1156

Collins, L.M., Schafer, J.L., & Kam, C.M. 2001. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6, (4) 330-351

Colman, I., Naicker, K., Zeng, Y., Ataullahjan, A., Senthilselvan, A., & Patten, S.B. 2011. Predictors of long-term prognosis of depression. *CMAJ.*, 183, (17) 1969-1976

Copeland, J.R., Beekman, A.T., Braam, A.W., Dewey, M.E., Delespaul, P., Fuhrer, R., Hooijer, C., Lawlor, B.A., Kivela, S.L., Lobo, A., Magnusson, H., Mann, A.H., Meller, I., Prince, M.J., Reischies, F., Roelands, M., Skoog, I., Turrina, C., deVries, M.W., & Wilson, K.C. 2004. Depression among older people in Europe: the EURODEP studies. *World Psychiatry*, 3, (1) 45-49

Daly, J., Elliott, D., Cameron-Traub, E., Salamonson, Y., Davidson, P., Jackson, D., Chin, C., & Wade, V. 2000. Health status, perceptions of coping, and social support immediately after discharge of survivors of acute myocardial infarction. *Am.J.Crit Care*, 9, (1) 62-69

Dennerstein, L. 1993. Psychosocial and mental health aspects of women's health. *World Health Stat.Q.*, 46, (4) 234-236

Dolan, P., Layard, R., & Metcalfe, R. (2011). Measuring Subjective Well-Being for Public Policy. London: Office for National Statistics. http://cep.lse.ac.uk/pubs/download/special/cepsp23.pdf accessed on June 2012

Doyal, L., & Gough, I. 1991. A theory of human need. Basingstoke, Hamphshire: Macmillan

Dufouil, C., Brayne, C., & Clayton, D. 2004. Analysis of longitudinal studies with death and drop-out: a case study. *Stat.Med.*, 23, (14) 2215-2226

Durrant, GB. 2005 *Imputation methods in the social sciences: a methodological review.* NCRM (NCRM Working Paper Series, (NCRM-002)

Enders, C.K. 2001. The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychol.Methods*, 6, (4) 352-370

Enders, C.K. & Bandalos, D.L. 2001. The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models. *Structural Equation Modeling-A Multidisciplinary Journal*, 8, (3) 430-457

Enders, C.K. 2006. A primer on the use of modern missing-data methods in psychosomatic medicine research. *Psychosom.Med.*, 68, (3) 427-436

Erens, B. & Primatesta, P. 1999. *Health Survey for England 1998* London, The Stationery Office.

Erens, B., Primatesta, P., & Prior, G. 2001. *Health Survey for England 1999: The Health of Minority Ethnic Groups* London, The Stationery Office.

Erikson R. 1993, "Description of inequality: the Swedish approach to welfare research," *In The Quality of life*, M. C. Nussbaum, A. Sen, & World Institute for Development Economics Research, eds., Oxford England: Clarendon Press, pp. 67-83.

Ferketich, A.K., Schwartzbaum, J.A., Frid, D.J., & Moeschberger, M.L. 2000. Depression as an antecedent to heart disease among women and men in the NHANES I

study. National Health and Nutrition Examination Survey. *Arch.Intern.Med.*, 160, (9) 1261-1268

Ferrans, C.E. & Powers, M.J. 1985. Quality of life index: development and psychometric properties. *ANS Adv.Nurs.Sci.*, 8, (1) 15-24

Ferrans, C.E. & Powers, M.J. 1992. Psychometric assessment of the Quality of Life Index. *Res.Nurs.Health*, 15, (1) 29-38

Ferrie, J.E., Kivimaki, M., Singh-Manoux, A., Shortt, A., Martikainen, P., Head, J., Marmot, M., Gimeno, D., De Vogli, R., Elovainio, M., & Shipley, M.J. 2009. Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. *Int.J.Epidemiol.*, 38, (3) 831-837

Ford, E.S., Mokdad, A.H., Li, C., McGuire, L.C., Strine, T.W., Okoro, C.A., Brown, D.W., & Zack, M.M. 2008. Gender differences in coronary heart disease and health-related quality of life: findings from 10 states from the 2004 behavioral risk factor surveillance system. *J.Womens Health (Larchmt.)*, 17, (5) 757-768

Forrester, A.W., Lipsey, J.R., Teitelbaum, M.L., DePaulo, J.R., & Andrzejewski, P.L. 1992. Depression following myocardial infarction. *Int.J.Psychiatry Med.*, 22, (1) 33-46

Fox, M.P., Lash, T.L., & Greenland, S. 2005. A method to automate probabilistic sensitivity analyses of misclassified binary variables. *Int.J.Epidemiol.*, 34, (6) 1370-1376

Frasure-Smith, N., Lesperance, F., Juneau, M., Talajic, M., & Bourassa, M.G. 1999. Gender, depression, and one-year prognosis after myocardial infarction. *Psychosom.Med.*, 61, (1) 26-37

Frasure-Smith N, Lesperance F. 2005. Reflections on depression as a cardiac risk factor. *Psychosom Med* May;67 Suppl 1:S19-S25.

Frasure-Smith N, Lesperance F. 2006. Recent evidence linking coronary heart disease and depression. *Can J Psychiatry* Oct;51(12):730-7.

Goldstein, H., Carpenter, J., Kenward, M.G., & Levin, K.A. 2009. Multilevel models with multivariate mixed response types. *Statistical Modelling*, 9, (3) 173-197

Goldstein, H. 2003. Multilevel statistical models., 3rd ed. London: Edward Arnold

Greenland, P., Reicher-Reiss, H., Goldbourt, U., & Behar, S. 1991. In-hospital and 1-year mortality in 1,524 women after myocardial infarction. Comparison with 4,315 men. *Circulation*, 83, (2) 484-491

Greenland, S. 1996. Basic methods for sensitivity analysis of biases. *International Journal of Epidemiology*, 25, (6) 1107-1116

Grundy Emily & Bowling Ann 1999. Enhancing the quality of extended life years. Identification of the oldest old with a very good and very poor quality of life. *Aging Ment.Health*, 3, (3) 199-212

Haapanen, N., Miilunpalo, S., Pasanen, M., Oja, P., & Vuori, I. 1997. Agreement between questionnaire data and medical records of chronic diseases in middle-aged and elderly Finnish men and women. *Am.J.Epidemiol.*, 145, (8) 762-769

Harel, O., Hofer, S M., Hoffman, L., and Pedersen, N.L., (2007) "Population Inference with Mortality and Attrition in Longitudinal Studies on Aging: A Two-Stage Multiple Imputation Method". Faculty Publications, Department of Psychology. Paper 434. http://digitalcommons.unl.edu/psychfacpub/434 Accessed October 2008.

Hemingway, H. & Marmot, M. 1999. Clinical Evidence: Psychosocial factors in the etiology and prognosis of coronary heart disease: systematic review of prospective cohort studies. *West J.Med.*, 171, (5-6) 342-350

Hemingway, H., Shipley, M., Britton, A., Page, M., Macfarlane, P., & Marmot, M. 2003. Prognosis of angina with and without a diagnosis: 11 year follow-up in the Whitehall II prospective cohort study. *BMJ*, 327, (7420) 895

Higginson, I.J. & Carr, A.J. 2001. Measuring quality of life: Using quality of life measures in the clinical setting. *BMJ* , 322, (7297) 1297-1300

Hillers, T.K., Guyatt, G.H., Oldridge, N., Crowe, J., Willan, A., Griffith, L., & Feeny, D. 1994. Quality of life after myocardial infarction. *J.Clin.Epidemiol.*, 47, (11) 1287-1296

Hobfoll, S.E., Dunahoo, C.L., Benporath, Y., Monnier, J. Gender and Coping - the Dual-Axis Model of Coping. *American Journal of Community Psychology* 1994 Feb;22(1):49-82.

Hyde, M., Wiggins, R.D., Higgs, P., & Blane, D.B. 2003. A measure of quality of life in early old age: the theory, development and properties of a needs satisfaction model (CASP-19). *Aging Ment. Health*, 7, (3) 186-194

Ibrahim, J.G., Chen, M.H., Lipsitz, S.R., & Herring, A.H. 2005. Missing-data methods for generalized linear models: A comparative review. *Journal of the American Statistical Association*, 100, (469) 332-346

Jelicic, H., Phelps, E., & Lerner, R.M. 2009. Use of missing data methods in longitudinal studies: the persistence of bad practices in developmental psychology. *Dev.Psychol.*, 45, (4) 1195-1199

Jevon, P. *Angina and Heart Attack* (2012). Oxford University Press

Jurek, A.M., Maldonado, G., Greenland, S., & Church, T.R. 2006. Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. *European Journal of Epidemiology*, 21, (12) 871-876

Kahneman, D., & Krueger, A.B. 2006. Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1), 3-24.

Kalilani, L. & Atashili, J. 2006. Measuring additive interaction using odds ratios. *Epidemiol.Perspect.Innov.*, 3, 5

Kehoe, R., Wu, S.Y., Leske, M.C., & Chylack, L.T., Jr. 1994. Comparing self-reported and physician-reported medical history. *Am.J.Epidemiol.*, 139, (8) 813-818

Kenward, M.G. & Carpenter, J. 2007. Multiple imputation: current perspectives. *Statistical Methods in Medical Research*, 16, (3) 199-218

Kessler, R.C. 2003. Epidemiology of women and depression. *J.Affect.Disord.*, 74, (1) 5-13

Ketterer, M.W., Denollet, J., Chapp, J., Thayer, B., Keteyian, S., Clark, V., John, S., Farha, A.J., & Deveshwar, S. 2004. Men deny and women cry, but who dies? Do the

wages of "denial" include early ischemic coronary heart disease? *Journal of Psychosomatic Research*, 56, (1) 119-123

Klebanoff, M.A. & Cole, S.R. 2008. Use of multiple imputation in the epidemiologic literature. *Am.J.Epidemiol.*, 168, (4) 355-357

Kline, R.B. 1998. *Principles and practice of structural equation modeling* New York, Guilford Press.

Kriegsman, D.M., Penninx, B.W., van Eijk, J.T., Boeke, A.J., & Deeg, D.J. 1996. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly. A study on the accuracy of patients' self-reports and on determinants of inaccuracy. *J.Clin.Epidemiol.*, 49, (12) 1407-1417

Kristofferzon, M.L., Lofmark, R., & Carlsson, M. 2003. Myocardial infarction: gender differences in coping and social support. *Journal of Advanced Nursing*, 44, (4) 360-374

Kristofferzon, M.L., Lofmark, R., & Carlsson, M. 2005a. Perceived coping, social support, and quality of life 1 month after myocardial infarction: a comparison between Swedish women and men. *Heart Lung*, 34, (1) 39-50

Kristofferzon, M.L., Lofmark, R., & Carlsson, M. 2005b. Coping, social support and quality of life over time after myocardial infarction. *J.Adv.Nurs.*, 52, (2) 113-124

Lampe, F.C., Walker, M., Lennon, L.T., Whincup, P.H., & Ebrahim, S. 1999. Validity of a self-reported history of doctor-diagnosed angina. *J.Clin.Epidemiol.*, 52, (1) 73-81

Lampe, F.C., Walker, M., Lennon, L.T., Whincup, P.H., & Ebrahim, S. 1999. Validity of a self-reported history of doctor-diagnosed angina. *J.Clin.Epidemiol.*, 52, (1) 73-81

Lane, D., Ring, C., Lip, G.Y., & Carroll, D. 2005. Depression, indirect clinical markers of cardiac disease severity, and mortality following myocardial infarction. *Heart*, 91, (4) 531-532

Landis, J.R. & Koch, G.G. 1977 The measurement of observer agreement for categorical data. *Biometrics, 33, 159-174.*

Lash, T.L. & Fink, A.K. 2003. Semi-automated sensitivity analysis to assess systematic errors in observational data. *Epidemiology*, 14, (4) 451-458

Laslett, P. 1996. *A fresh map of life the emergence of the Third Age* Basingstoke: Macmillan.

Lauzon, C., Beck, C.A., Huynh, T., Dion, D., Racine, N., Carignan, S., Diodati, J.G., Charbonneau, F., Dupuis, R., & Pilote, L. 2003. Depression and prognosis following hospital admission because of acute myocardial infarction. *CMAJ.*, 168, (5) 547-552

Lerner, D.J. & Kannel, W.B. 1986. Patterns of coronary heart disease morbidity and mortality in the sexes: a 26-year follow-up of the Framingham population. *Am.Heart J.*, 111, (2) 383-390

Lesperance, F., Frasure-Smith, N., & Talajic, M. 1996. Major depression before and after myocardial infarction: its nature and consequences. *Psychosom.Med.*, 58, (2) 99-110

Lesperance, F. & Frasure-Smith, N. 2000. Depression in patients with cardiac disease: a practical review. *J.Psychosom.Res.*, 48, (4-5) 379-391 available from: PM:10880660

Little, R.J.A. & Rubin, D.B. 2002. *Statistical analysis with missing data*, 2nd ed. Hoboken, N.J, Wiley.

Lyles, R.H. & Lin, J. 2010. Sensitivity analysis for misclassification in logistic regression via likelihood methods and predictive value weighting. *Statistics in Medicine*, 29, (22) 2297-2309

Lyons, R.A., Lo, S.V., & Littlepage, B.N. 1994. Comparative health status of patients with 11 common illnesses in Wales. *J.Epidemiol.Community Health*, 48, (4) 388-390

Maldonado, M. & Charney, P. 1995. The prevention of coronary heart disease in women. *N.Engl.J.Med.*, 333, (23) 1571-1572

Mallik, S., Spertus, J.A., Reid, K.J., Krumholz, H.M., Rumsfeld, J.S., Weintraub, W.S., Agarwal, P., Santra, M., Bidyasar, S., Lichtman, J.H., Wenger, N.K., & Vaccarino, V. 2006. Depressive symptoms after acute myocardial infarction: evidence for highest rates in younger women. *Arch.Intern.Med.*, 166, (8) 876-883

Marmot, M., Banks, J., Blundell, R., Lessof, C., & Nazroo, J. 2003. *Health, wealth and lifestyles of the older population in England: ELSA 2002* London, The Institute for Fiscal Study.

Marmot, M. & Brunner, E. 2005. Cohort Profile: the Whitehall II study. *Int.J.Epidemiol.*, 34, (2) 251-256

Marmot, M.G., Smith, G.D., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., & Feeney, A. 1991. Health inequalities among British civil servants: the Whitehall II study. *Lancet*, 337, (8754) 1387-1393

Marshall, A., Altman, D.G., Royston, P., & Holder, R.L. 2010. Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. *BMC.Med.Res.Methodol.* , 10, 7

Marzari, C., Maggi, S., Manzato, E., Destro, C., Noale, M., Bianchi, D., Minicuci, N., Farchi, G., Baldereschi, M., Di Carlo, A., & Crepaldi, G. 2005. Depressive symptoms and development of coronary heart disease events: the Italian longitudinal study on aging. *J.Gerontol.A Biol.Sci.Med.Sci.*, 60, (1) 85-92

McMunn A., Hyde M, Janevic M, & Kumari M 2003, *In Health, wealth and lifestyles of the older population in England: ELSA 2002*, J. Banks et al., eds. London: The Institute for Fiscal Study, pp. 207-230.

Mendes de Leon, C.F., Krumholz, H.M., Vaccarino, V., Williams, C.S., Glass, T.A., Berkman, L.F., & Kas, S.V. 1998. A population-based perspective of changes in health-related quality of life after myocardial infarction in older men and women. *J.Clin.Epidemiol.*, 51, (7) 609-616

Mendes de Leon, C.F., Dilillo, V., Czajkowski, S., Norten, J., Schaefer, J., Catellier, D., & Blumenthal, J.A. 2001. Psychosocial characteristics after acute myocardial infarction: the ENRICHD pilot study. Enhancing Recovery in Coronary Heart Disease. *J.Cardiopulm.Rehabil.*, 21, (6) 353-362

Merkin, S.S., Cavanaugh, K., Longenecker, J.C., Fink, N.E., Levey, A.S., & Powe, N.R. 2007. Agreement of self-reported comorbid conditions with medical and physician reports varied by disease among end-stage renal disease patients. *J.Clin.Epidemiol.*, 60, (6) 634-642

Molenberghs, G. & Kenward, M.G. 2007. Missing data in clinical studies. Chichester: John Wiley & Sons

Musil, C.M., Warner, C.B., Yobas, P.K., & Jones, S.L. 2002. A comparison of imputation techniques for handling missing data. *Western Journal of Nursing Research*, 24, (7) 815-829

Muthén, L.K., & Muthén, B. O. 2007. Mplus user's guide (5th ed.). Los Angeles, CA: Muthén & Muthén.

Naqvi, T.Z., Naqvi, S.S., & Merz, C.N. 2005. Gender differences in the link between depression and cardiovascular disease. *Psychosom.Med.*, 67 Suppl 1, S15-S18

Netuveli, G., Wiggins, R.D., Hildon, Z., Montgomery, S.M., & Blane, D. 2006. Quality of life at older ages: evidence from the English longitudinal study of aging (wave 1). *Journal of Epidemiology and Community Health*, 60, (4) 357-363

Nevalainen, J., Kenward, M.G., & Virtanen, S.M. 2009. Missing values in longitudinal dietary data: a multiple imputation approach based on a fully conditional specification. *Stat.Med.*, 28, (29) 3657-3669

Newman, D.A. 2003. Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*, 6, (3) 328-362

Norris, C.M., Saunders, L.D., Ghali, W.A., Brant, R., Galbraith, P.D., Graham, M., Faris, P., Dzavik, V., & Knudtson, M.L. 2004. Health-related quality of life outcomes of patients with coronary artery disease treated with cardiac surgery, percutaneous coronary intervention or medical management. *Can.J.Cardiol.*, 20, (12) 1259-1266

Norris, C.M., Hegadoren, K., & Pilote, L. 2007. Depression symptoms have a greater impact on the 1-year health-related quality of life outcomes of women post-myocardial infarction compared to men. *Eur.J.Cardiovasc.Nurs.*, 6, (2) 92-98

Okura, Y., Urban, L.H., Mahoney, D.W., Jacobsen, S.J., & Rodeheffer, R.J. 2004. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J.Clin.Epidemiol.*, 57, (10) 1096-1103

Orsini, N., Bellocco, R., Bottai, M., Wolk, A., & Greenland, S. 2008. A tool for deterministic and probabilistic sensitivity analysis of epidemiologic studies. *Stata Journal*, 8, (1) 29-48

Parashar, S., Rumsfeld, J.S., Spertus, J.A., Reid, K.J., Wenger, N.K., Krumholz, H.M., Amin, A., Weintraub, W.S., Lichtman, J., Dawood, N., & Vaccarino, V. 2006. Time course of depression and outcome of myocardial infarction. *Arch.Intern.Med.*, 166, (18) 2035-2043

Penninx, B.W., Geerlings, S.W., Deeg, D.J., van Eijk, J.T., van Tilburg, W., & Beekman, A.T. 1999. Minor and major depression and the risk of death in older persons. *Arch.Gen.Psychiatry*, 56, (10) 889-895

Peyre, H., Leplege, A., & Coste, J. 2011. Missing data methods for dealing with missing items in quality of life questionnaires. A comparison by simulation of personal mean score, full information maximum likelihood, multiple imputation, and hot deck techniques applied to the SF-36 in the French 2003 decennial health survey. *Quality of Life Research*, 20, (2) 287-300

Polk, D.M. & Naqvi, T.Z. 2005. Cardiovascular disease in women: sex differences in presentation, risk factors, and evaluation. *Curr.Cardiol.Rep.*, 7, (3) 166-172

Popper, K. R. 2002. *Conjectures and refutations: the growth of scientific knowledge.* Routledge 2nd edition.

Power, M., Quinn, K., & Schmidt, S. 2005. Development of the WHOQOL-Old module. *Quality of Life Research*, 14, (10) 2197-2214

Radloff, L.S. 1977. The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385-401

Reiter, J.P., Raghunathan T.E., & Kinney S.K. 2006. The Importance of Modeling the Sampling Design in Multiple Imputation for Missing Data. *Survey Methodology*, 32, (2) 143-149

Richardson, L.G. 2003. Psychosocial issues in patients with congestive heart failure. *Prog.Cardiovasc.Nurs.*, 18, (1) 19-27

Robins, J.M. & Wang, N.S. 2000. Inference for imputation estimators. *Biometrika*, 87, (1) 113-124

Royston, P. 2005. Multiple imputation of missing values: update. *Stata Journal*, 5, (2) 188-201

Royston, P. 2007. Multiple imputation of missing values: further update of ice, with an emphasis on interval censoring. *Stata Journal*, 7, (4) 445-464

Royston, P., Carlin, J.B., & White, I.R. 2009. Multiple imputation of missing values: New features for mim. *Stata Journal* , 9, (2) 252-264

Royston, P. 2009. Multiple imputation of missing values: Further update of ice, with an emphasis an categorical variables. *Stata Journal*, 9, (3) 466-477

Rozanski, A., Blumenthal, J.A., & Kaplan, J. 1999. Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. *Circulation*, 99, (16) 2192-2217

Rubin, D.B. 1976. Inference and Missing Data. *Biometrika*, 63, (3) 581-590

Rubin, D.B. 1987. *Multiple imputation for nonresponse in surveys*. Wiley.

Rubin, D.B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical Association*, 91, (434) 473-489

Schafer, J.L. 1997. Analysis of incomplete multivariate data. Monographs on Statistics and Applied Probability 72, Chapman and Hall/CRC

Schafer, J.L. and Olsen, M.K. 1998 Multiple imputation for multivariate missing-data problems: a data analyst's perspective. *Multivariate Behavioral Research,* 33, 545-571

Schafer, J. L. & Graham, J. W. 2002. Missing data: our view of the state of the art. *Psychological Methods*, 7, 147-177

Scheffer, J. 2002. Dealing with missing data. *Res. Lett. Inf. Math. Sci.* 3, 153-160. http://www.massey.ac.nz/~wwiims/research/letters/ accessed on August 2009

Schleifer, S.J., Macari-Hinson, M.M., Coyle, D.A., Slater, W.R., Kahn, M., Gorlin, R., & Zucker, H.D. 1989. The nature and course of depression following myocardial infarction. *Arch.Intern.Med.*, 149, (8) 1785-1789

Schoevers, R.A., Geerlings, M.I., Beekman, A.T., Penninx, B.W., Deeg, D.J., Jonker, C., & van Tilburg, W. 2000. Association of depression and gender with mortality in old age. Results from the Amsterdam Study of the Elderly (AMSTEL). *Br.J.Psychiatry*, 177, 336-342

Scholes, S., Taylor, R., Cheshire, H., Cox, K., Lessof, C. 2009. *Technical report (ELSA wave 3): living in the 21st century: older people in England.* http://www.ifs.org.uk/elsa/report06/w2_tech.pdf accessed August 2009

Schumacker, R.E. & Lomax, R.G 1996. *A beginner's guide to structural equation modelling.* Mahwah, N.J, L. Erlbaum Associates.

Shin, J.H. 2009. Application of repeated-measures analysis of variance and hierarchical linear model in nursing research. *Nurs.Res.*, 58, (3) 211-217

Sinharay, S., Stern, H.S., & Russell, D. 2001. The use of multiple imputation for the analysis of missing data. *Psychol.Methods*, 6, (4) 317-329

Smith, H.J., Taylor, R., & Mitchell, A. 2000. A comparison of four quality of life instruments in cardiac patients: SF-36, QLI, QLMI, and SEIQoL. *Heart*, 84, (4) 390-394

Sproston, K. & Primatesta, P. 2004. *Health Survey for England 2003* London, The Stationery Office.

Stafford, M., McMunn, A., Zaninotto, P., & Nazroo, J. 2011. Positive and negative exchanges in social relationships as predictors of depression: evidence from the English Longitudinal Study of Aging. *J.Aging Health*, 23, (4) 607-628

Stansfeld, S.A., Fuhrer, R., Shipley, M.J., & Marmot, M.G. 2002. Psychological distress as a risk factor for coronary heart disease in the Whitehall II Study. *Int.J.Epidemiol.*, 31, (1) 248-255

Steffick D.E. (2000). *Documentation of affective functioning measures in the Health and Retirement Study*. (HRS/AHEAD Documentation. Report DR-005). Survey Research Center, University of Michigan: Ann Arbor, MI <http://hrsonline.isr.umich.edu/docs/userg/dr-005.pdf>. Accessed 8 October 2008.

Stek, M.L., Gussekloo, J., Beekman, A.T.F., van Tilburg, W., & Westendorp, R.G.J. 2004. Prevalence, correlates and recognition of depression in the oldest old: the Leiden 85-plus study. *Journal of Affective Disorders*, 78, (3) 193-200

Stern, M.J., Pascale, L., & Ackerman, A. 1977. Life adjustment postmyocardial infarction: determining predictive variables. *Arch.Intern.Med.*, 137, (12) 1680-1685

Sterne, J.A., White, I.R., Carlin, J.B., Spratt, M., Royston, P., Kenward, M.G., Wood, A.M., & Carpenter, J.R. 2009. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, 338, b2393

Stramba, B.M. & Priori, S.G. 2006. [Current strategies to diminish the impact of cardiovascular diseases in women]. *Rev.Esp.Cardiol.*, 59, (11) 1190-1193

Strike, P.C. & Steptoe, A. 2002. Depression, stress, and the heart. *Heart*, 88, (5) 441-443

Stuart-Shor, E.M., Buselli, E.F., Carroll, D.L., & Forman, D.E. 2003. Are psychosocial factors associated with the pathogenesis and consequences of cardiovascular disease in the elderly? *J.Cardiovasc.Nurs.*, 18, (3) 169-183

Taylor, R., Conway, L., Calderwood, L., Lessof, C., Cheshire, H., Cox, K., Scholes, S. 2007. *Technical report (ELSA wave 1): health, wealth and lifestyles of the older population in England http://www.ifs.org.uk/elsa/report03/w1_tech.pdf accessed August 2009*

Thompson, D.R. & Yu, C.M. 2003. Quality of life in patients with coronary heart disease-I: assessment tools. *Health Qual.Life Outcomes.*, 1, 42

Turner-Bowker, D.M., Bartley P.J., & Ware, J.E., Jr. 2002. *SF-36® Health Survey & "SF" Bibliography*, Third (1988-2000) ed. Lincoln (RI): QualityMetric Incorporated.

van Buuren, S., Boshuizen, H.C., & Knook, D.L. 1999. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat.Med.*, 18, (6) 681-694

van Buuren, S. 2007. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat.Methods Med.Res.*, 16, (3) 219-242

van Gool, C.H., Kempen, G.I., Penninx, B.W., Deeg, D.J., Beekman, A.T., & van Eijk, J.T. 2003. Relationship between changes in depressive symptoms and unhealthy lifestyles in late middle aged and older persons: results from the Longitudinal Aging Study Amsterdam. *Age Ageing*, 32, (1) 81-87

van Jaarsveld, C.H., Sanderman, R., Miedema, I., Ranchor, A.V., & Kempen, G.I. 2001. Changes in health-related quality of life in older patients with acute myocardial infarction or congestive heart failure: a prospective study. *J.Am.Geriatr.Soc.*, 49, (8) 1052-1058

Vargas-Chanes, D., Decker, P.A., Shroeder, D.R., Offord, K.P. 2003. An introduction to multiple imputation methods: Handling missing data with SAS V 8.2. Tehnical Report #67. http://mayoresearch.mayo.edu/mayo/research/biostat/upload/67.pdf accessed on September 2009

Victor, C. R. S. T. 2005, "Social isolation and loneliness," *In Understanding quality of life in old age*, A. Walker, ed., Maidenhead ; New York : Open University Press, pp. 110-116.

Wang, H.X., Mittleman, M.A., & Orth-Gomer, K. 2005. Influence of social support on progression of coronary artery disease in women. Social Science & Medicine, 60, (3) 599-607

Webb, E., Blane, D., McMunn, A., Netuveli, G. Proximal predictors of change in quality of life at older ages. *Journal of Epidemiology and Community Health* 2011 Jun;65(6):542-7.

Wells, K.B., Stewart, A., Hays, R.D., Burnam, M.A., Rogers, W., Daniels, M., Berry, S., Greenfield, S., & Ware, J. 1989. The functioning and well-being of depressed patients. Results from the Medical Outcomes Study. *JAMA*, 262, (7) 914-919

Wenger, N.K. 2002. Clinical characteristics of coronary heart disease in women: emphasis on gender differences. *Cardiovasc.Res.*, 53, (3) 558-567

Westin, L., Carlsson, R., Erhardt, L., Cantor-Graae, E., & McNeil, T. 1999. Differences in quality of life in men and women with ischemic heart disease. A prospective controlled study. *Scand.Cardiovasc.J.*, 33, (3) 160-165

White, I.R. & Carlin, J.B. 2010. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat.Med.*, 29, (28) 2920-2931

Wiggins, R. D., & Sacker, A. 2001. A comparative evaluation of strategies to handle missing data in the context of structural equation modelling: A user's perspective. In G. Marcoulides & I.Moustaki eds., Latent variable and latent structure models. New Jersey: Lawrence Erlbaum Associates, pp.105–120.

Wiggins, R.D., Netuveli, G., Hyde, M., Higgs, P., & Blane, D. 2008. The evaluation of a self-enumerated scale of quality of life (CASP-19) in the context of research on ageing: A combination of exploratory and confirmatory approaches. *Social Indicators Research*, 89, (1) 61-77

Wiklund, I., Herlitz, J., & Hjalmarson, A. 1989. Quality of life five years after myocardial infarction. *Eur.Heart J.*, 10, (5) 464-472

Wiklund, I., Herlitz, J., Johansson, S., Bengtson, A., Karlson, B.W., & Persson, N.G. 1993. Subjective symptoms and well-being differ in women and men after myocardial infarction. *Eur.Heart J.*, 14, (10) 1315-1319

William A. 1977, "Measuring the quality of life of the elderly," *In Public economics and the quality of life*, Evans A. & Wingo L., eds., Baltimore: John Hopkins University Press, pp. 13-27.

Wotke, W. 2000: Longitudinal and multigroup modeling with missing data. I T.D. Little, K.U. Schnabel & J. Baumert (eds): Modeling longitudinal and multilevel data. Practical issues, applied approaches, and specific examples. New Jersey: Lawrence Erlbaum.

Yamagishi, K., Ikeda, A., Iso, H., Inoue, M., & Tsugane, S. 2009. Self-reported stroke and myocardial infarction had adequate sensitivity in a population-based prospective study JPHC (Japan Public Health Center)-based Prospective Study. *J.Clin.Epidemiol.*, 62, (6) 667-673

Young, MA, Fogg, LF, Scheftner, WA, Kelle, MB, Fawcett, JA.1990. Sex differences in the lifetime prevalence of depression: does varying the diagnostic criteria reduce the female/male ratio? *J Affect Disord*. 18:187-192.

Zaninotto, P., Pierce, M., Breeze, E., de, O.C., & Kumari, M. 2010. BMI and waist circumference as predictors of well-being in older adults: findings from the English Longitudinal Study of Ageing. *Obesity (Silver.Spring)*, 18, (10) 1981-1987

Ziegelstein, R.C., Fauerbach, J.A., Stevens, S.S., Romanelli, J., Richter, D.P., & Bush, D.E. 2000. Patients with depression are less likely to follow recommendations to reduce cardiac risk during recovery from a myocardial infarction. *Arch.Intern.Med.*, 160, (12) 1818-1823

Ziegelstein, R.C. 2001. Depression in patients recovering from a myocardial infarction. *JAMA*, 286, (13) 1621-1627