Time-dependent markers of comorbidity and prognosis in heart failure patients: transitions across the life course

Claire Alexandra Rushton

Doctor of Philosophy

October 2016

Keele University

Keele University

Declaration Part 1. To be bound in the thesis

SUBMISSION OF THESIS FOR A RESEARCH DEGREE

Part I. DECLARATION by the candidate for a research degree. To be bound in the thesis

Degree for which thesis being submitted: Doctor of Philosophy

Title of thesis: Time-dependent markers of comorbidity and prognosis in heart failure patients: transitions across the life course

This thesis contains confidential information and is subject to the protocol set down for the submission and examination of such a thesis.

NO [please delete as appropriate; if YES the box in Part II should be completed]

Date of submission 19/07/2016 Original registration date 02/02/2012
(Date of submission must comply with Regulation 2D)

Name of candidate Claire Alexandra Rushton

Research Institute Science and Technology in Medicine

Name of Lead Supervisor Professor Umesh T Kadam

I certify that:

(a) The thesis being submitted for examination is my own account of my own research
(b) My research has been conducted ethically. Where relevant a letter from the approving body confirming that ethical approval has been given has been bound in the thesis as an Annex
(c) The data and results presented are the genuine data and results actually obtained by me during the conduct of the research
(d) Where I have drawn on the work, ideas and results of others this has been appropriately acknowledged in the thesis
(e) Where any collaboration has taken place with one or more other researchers, I have included within an 'Acknowledgments' section in the thesis a clear statement of their contributions, in line with the relevant statement in the Code of Practice (see Note overleaf).
(f) The greater portion of the work described in the thesis has been undertaken subsequent to my registration for the higher degree for which I am submitting for examination
(g) Where part of the work described in the thesis has previously been incorporated in another thesis submitted by me for a higher degree (if any), this has been identified and acknowledged in the thesis
(h) The thesis submitted is within the required word limit as specified in the Regulations

Total words in submitted thesis (including text and footnotes, but excluding references and appendices) 98,144

Signature of candidate…… Date …19/07/2016……

Note

Extract from Code of Practice: If the research degree is set within a broader programme of work involving a group of investigators – particularly if this programme of work predates the candidate's registration – the candidate should provide an explicit statement (in an 'Acknowledgments' section) of the respective roles of the candidate and these other individuals in relevant aspects of the work reported in the thesis. For example, it should make clear, where relevant, the candidate's role in designing the study, developing data collection instruments, collecting primary data, analysing such data, and formulating conclusions from the analysis. Others involved in these aspects of the research should be named, and their contributions relative to that of the candidate should be specified (*this does not apply to the ordinary supervision, only if the supervisor or supervisory team has had greater*

Declaration

I am a Lecturer in Nursing at the School of Nursing and Midwifery at Keele University and an honorary cardiology nurse at the University Hospitals of North Midlands. I was supported by a National Institute Health Research (NIHR) doctoral research fellowship to complete a three year full time PhD within the Health Services Research Unit at Keele University.

The initial research ideas were formed from my time spent with patients in practice and student nurses and in discussion with Professor Umesh T. Kadam. The full PhD plan, interpretation and discussion of the findings in the thesis are my own. This thesis contains analysis of the Clinical Practice Research Datalink Database (CPRD) which was extracted using a CPRD licence funded by Institute of Primary Care Sciences at Keele University. The PhD was funded by my NIHR doctoral award (NIHR-DRF-2012-05-288). I conducted my own systematic review and meta analyses with support from my supervisory team. I developed my own conceptual framework for including comorbidity in prognosis which I applied in this thesis. For the CPRD analyses, I performed all the data cleaning, organising and linkage and planned and conducted my own analyses.

I wrote all chapters with overview support from my supervisory team (Professor Umesh T Kadam and Professor Peter Jones).

Abstract

Heart failure (HF) disease carries a poor prognosis despite optimisation of cardiovascular (CVD) treatments. Non-CVD comorbid diseases are common and known to influence the HF clinical course. These comorbidities change in severity over time from new onset yet, only static measures of comorbidity have been included in prognosis. A major gap in the management and prognosis of HF is how non-CVD comorbidity severity and longitudinal change influences individual risk.

A systematic review (SR) and two phase observational study were conducted in the general HF population to test the hypothesis that increasing severity and change of non-CVD comorbidity would be associated with worse outcomes. The SR showed that the three most common non-CVD comorbidities included in prognosis were diabetes mellitus (DM), chronic obstructive pulmonary disease (COPD) and chronic kidney disease (CKD). Hospital admission outcome studies were limited and there were no studies on quality of life. With the exception of hospital based renal studies, comorbidity severity and change evidence was scarce.

The observational studies used a case-control study nested within the UK Clinical Practice Research Datalink database (2002- 2014), of 50,114 incident HF patients. Using risk set sampling, multiple controls were matched to cases on follow-up time. A framework for measuring recent comorbidity severity and change was devised using drug or physiological indicators for DM, COPD and CKD measured in two time-windows prior to the match date. Conditional logistic regression was used to estimate adjusted odds ratios for all-cause hospital admission and mortality.

The observational study findings were that all three comorbidities were common and associated with both outcomes. Severe and worsening comorbid disease was also common and independently and significantly associated with increased risk of hospital admission and mortality. These dynamic measures of non-CVD comorbidity significantly improved HF prognostic models which has important implications for HF management and prognosis.

Table of Contents

List of Electronic Appendices

E-Appendix A: Tables

List of tables

List of figures

Acknowledgements

I would first like to thank my lead supervisor, Professor Umesh Kadam for his insight and supporting my PhD. His epidemiological, clinical and academic experience has guided me to the completion of this thesis. In addition his clear leadership has encouraged me to think independently which has enhanced my current and future potential as a researcher. I would also like to thank Professor Peter Jones for his consistent support and from whom I have gained a better statistical understanding. Additionally I would like to thank Lucy Teese and acknowledge her frequent support in the application of the statistical approaches used in my work.

I would like to thank Professor Karl Moons from the Cochrane Prognosis Methods group for reviewing my systematic review protocol and the Clinical Practice Research Datalink team for their initial extraction of the data used in the main analyses but also their ongoing support as questions have arisen.

In addition to the main supervisory team, I would like to thank Pauline Walsh as Head of the School of Nursing at Keele for her support and continued mentorship during my research.

My personal acknowledgements go to my mum, dad and children, Lewis and Ruby for their unwavering patience, understanding and support.

Context setting

Chapter 1 Introduction

This chapter will introduce the broader principles that underpin the thesis which include epidemiology, disease and health, comorbidity and prognosis. Current conceptual thinking on the key terms will be explored before application to the context of heart failure disease in Chapter 2.

1.1 Concepts of epidemiology

1.1.1 Definition

Epidemiology is based on the assumption that the occurrence of disease does not happen randomly but that there are factors, identifiable through systematic enquiry, which cause and prevent disease(1). Epidemiology is defined as 'the study of the distribution and determinants of disease frequency(2) and the science of epidemiology is focused on the investigation and description of these fundamental concepts.

1.1.2 Background

The notion that the development and occurrence of disease may be related to factors that are external to the person, such as the environment, dates back to the beginning of medicine and can be found in the works of Hippocrates in 5BC(3) but it wasn't until the 17th century that the importance of using routinely collected data to quantify disease was recognised. John Graunt used the weekly 'Bills of Mortality' to summarize natural observations of births and deaths, trends in disease, categorisation of disease and population growth. The recognition of the value of routinely collected data by Graunt has been said to form the basis of modern epidemiology(1). Many years later John Snow in characterising the cholera epidemic of 1853-1854 provided one of the first attempts to establish cause and effect when he determined that the water supply was responsible for the outbreak and spread of cholera(4). The epidemiological approaches used by Snow are still in use today but until the early nineteenth century were almost solely applied to outbreaks of infectious diseases. Over the past century the patterns of mortality have changed with non-communicable diseases

becoming the greatest threat to public health(5). Modern epidemiological approaches have developed in sophistication to account for the broader and more complex epidemic of chronic diseases.

1.1.3 Core concepts: frequency, distribution and determinants

The *frequency* of disease occurrence is determined within specific populations. Studying and comparing populations allows for the identification and investigation of patterns of disease that may be characterised by personal attributes e.g. age, or demographic characteristics e.g. geographical areas or time periods. Measures of disease frequency include (i) disease prevalence; or the number of people with a particular disease (new and existing cases) divided by the total number in the population at a given time (point prevalence) or time period (period prevalence) and (ii) disease incidence; or the number of new disease cases over a specified time(6). Within a population of interest the occurrence of disease will not usually be evenly distributed and the investigation of the uneven *distribution* allows for the development of hypotheses about possible causal and preventive factors or 'exposures' which can then be tested. The focus of the investigation here is on the person, place and time(7):

- Person: who within the population is contracting the disease?
- Place: where are the people that are contracting the disease?
- Time: when are they contracting the disease?

The enquiry into possible causal or preventive exposures is often precipitated by prior knowledge or experience of potential important factors generated by basic science, clinical experience or biological knowledge.

Testing hypotheses about possible causal or preventive exposures allows for the investigation of the true *determinants* of disease. In order to demonstrate that a true possible relationship between an exposure and the occurrence of disease exists, the group of individuals with the exposure are compared with an appropriate unexposed group(8).

The investigation of a determinant or *cause* of disease is central to epidemiological research. True *cause and effect* requires that there is a direct relationship between an exposure and an outcome that cannot be explained by any other factors. This does not mean that any individual factor will be fully responsible for an outcome or event. Instead there are usually combinations of causal components, that commonly include genetic and environmental factors and that may be separated by time, which *interact* to bring about an outcome(9). The induction time to disease which refers to the time between causal action and disease onset, is therefore specific to each component rather than the disease itself(10). As multiple components may be necessary to cause the event, the *strength* of a component cause is more a function of the number of cases in which the causal component plays a role. This means that the strength of cause (or subsequent effect) is neither strong nor weak but depends on the role the factor plays in its interaction with the other causal components and may change from population to population and over time if the distribution of the multiple causal components changes.

Rothman explains this phenomenon by way of a Causal Pie Model(2) by the use of three pies shown in Figure 1.1. Each pie has the potential to cause a specific disease but requires all of its constituent parts to be a sufficient cause. It can be seen that for each sufficient cause (pie) the constituent parts are equally important to the individual experiencing the disease. If one part was missing the disease would not occur. However constituent A is part of all sufficient causes and so at a population level is responsible for a higher burden of the disease and would yield a higher strength of cause.

Biological interaction, also referred to as causal interaction(11)*,* occurs where some cases require the presence of two variables together. The first pie in Figure 1.1 requires both A and B for causation. This does not mean that A depends on B as A also contributes independently of B to the other two pies. Instead, for some cases, both A and B are required. A simple way of measuring whether there is biological interaction between A and B is to subtract the individual effect of A, the individual effect of B and the effect of having neither A nor B from the joints effects of having A and B.

Biological interaction = Effect AB – Effect A – Effect B – Effect neither A nor B

1.1.4 Population-based versus clinical epidemiology

A *population* refers to a group of individuals who share a characteristic such as a geographical area but can also relate to more specific characteristics such as employment, year of birth, gender or schools. *Population-based epidemiology* refers to the scientific enquiry into the determinants of disease within a population which is often done by comparing those in the population who develop the disease with those who don't. Alternatively the characteristic that a population shares may be the disease itself and here scientific enquiry will focus on determining how to accurately diagnose the disease, predict its natural or clinical course or the best intervention or treatment(12). This *clinical epidemiology* concept, often referred to as health services research or prognosis research, shares the same methodological approaches to population epidemiology(13). In this thesis, this clinical epidemiology approach is used to determine the importance of other co-existing chronic diseases for predicting the clinical course of people with an index disease.

1.1.5 Importance of epidemiology for current healthcare

The increasing number of people with non-communicable or chronic diseases in developed countries over the past century(5,14) requires sophisticated epidemiological investigation in order to address important questions relating to the cause, progression, management and prevention of disease. Primary and secondary prevention of chronic disease has become a public health priority demanding deeper understanding of causal factors. Prevention strategies in chronic disease management have focused primarily on risk factors associated with a specific disease for example cardiovascular aetiology or risk factors such as smoking or hypertension in heart failure, but important questions remain about how factors that are external to the specific disease may influence both the onset and progression of disease. One clear example is where people experience more than one chronic disease at the same time and public health policy needs to understand how different diseases influence each other and which combinations of diseases are most influential on health deterioration.

The longer latency period between a sufficient cause and the clinical presentation of a chronic disease means that causal associations are more difficult to infer(15). Similar challenges within chronic disease populations occur where the period of time between the exposure to factors that influence the progression of the disease and the manifestation of this progression in the outcomes of the disease can be lengthy. Many different exposures may confuse the relationship between a specific exposure and outcome and possible causation may be less obvious and difficult to detect. A further complication of the longer latency is that the exposure itself may change in intensity or duration from its initial inception. In order to clearly understand the association between the exposure and the outcome, the quality and quantity of exposure needs to be characterised.

Whilst true mechanisms of association may be challenging to prove, carefully designed epidemiological approaches can provide important information for public health policy that is immediate and reliable(16). Unlike basic research conducted in a laboratory, epidemiology focuses on what is happening in human populations and can ascertain the effects of exposures in the real world context.

1.1.6 Application to nursing practice

One of the first pioneers of epidemiology after John Snow was Florence Nightingale. As an accomplished epidemiologist and statistician Nightingale made major contributions to public health through her nursing leadership. Through meticulous scrutiny of carefully collected data, Nightingale observed the uneven distribution of mortalities during the Crimean war, UK military and rural India and was able to identify key associations with sanitation, nutrition and medical interventions(9).

Today, epidemiological evidence underpins public and preventative healthcare policy and legislation that impacts directly on nursing practice. The role of epidemiology in identifying and measuring health states and testing interventions to improve them also has wide ranging implications for nursing. However, despite the importance of epidemiology for nursing practice and the early promise from Nightingale's work, epidemiology as a scientific research discipline has since predominated in medicine rather than nursing.

Whilst nurses may apply epidemiological approaches in their studies, explicit acknowledgement in the literature is lacking. It has been argued that a major barrier to epidemiology for nurses has been the close alignment of epidemiology with the biomedical research model(17). Alongside the move away from biomedical and positivist models of health in the nursing discipline, nursing research moved away from quantitative approaches focused on disease and towards qualitative approaches aimed at understanding health more broadly and from the individual perspective(18). The early alignment of epidemiology with the quantification of disease within populations created obvious conflicts with the interests of nursing research.

However, the development of clinical epidemiology in the era of chronic disease means that the nursing profession is well placed to contribute to and gain from the science of epidemiology. The study of disease causation indicates that there is rarely a single causal agent. Social epidemiology is now adding to the evidence of wider determinants including social inequalities and poverty(19). The uncertainty inherent in true casual relations, the complex and multiple factors that interact to determine outcomes and the range of possible outcomes that are of interest to patients with a specific disease, means that nursing with its roots in holism and a humanistic perspective is well placed to embrace epidemiology. This broader perspective on the determinants and outcomes of disease in clinical populations favours an integrated approach where medical and nursing models can easily complement each other to develop epidemiological science.

1.2 Concepts of disease and health

Epidemiology has traditionally been concerned with the biological causation of disease within populations which resulted in the concept of health being aligned with the state of 'non-disease'. The development of clinical epidemiology with its focus on the onset, progression and outcomes of people with a chronic disease requires a broader perspective both in the consideration of the possible determinants and selection of outcomes which define the health of patients. Before clinical epidemiology questions can be subjected to scientific scrutiny, a clear definition of health is required.

1.2.1 Defining disease

The term 'disease' is multifaceted and challenging to define. In its simplest form disease is used to indicate illness or conditions that impair normal functioning and that have a specific set of signs and symptoms or pathology(20). What constitutes as 'normal functioning' is however context dependent and changes across cultures and time(21) and at a minimum could be argued is a product of evolution rather than a 'natural state'(22). Whilst the term disease is diverse and includes illnesses and conditions with varying duration, 'chronic' diseases are defined by their long duration, slow progression and incurable nature and account for 63% of annual deaths globally(23).

1.2.2 Measuring disease

In its simplest form chronic disease can be measured in individuals by its presence or absence or '*diagnostic status*'. However, the impact of diseases on populations and individuals will vary for different diseases and provides a notion of '*between disease'* severity(24). Individuals with a specific disease will also vary in their experience of the same disease. The progressive nature of chronic disease from new onset to more severe disease and to death means that '*within disease'* the stage and severity needs careful consideration. Severity then can be measured at different levels which include the population (between diseases), the disease (within disease) and the person (experience).

Measuring disease severity within populations: The severity staging of diseases according to their population impact can be applied to biological systems e.g. to hypertension or ischemic heart disease within the broader definition of cardiovascular disease or between systems e.g. a cardiovascular disease compared to respiratory disease. Morbidity indexes attempt to weight diseases by severity as determined by outcomes or characteristics of the disease(25). The Charlson index was one of the first morbidity indexes to recognise that in summing up the impact of multiple coexisting diseases that both the number and severity of the diseases were important. Charlson weighted different diseases according to their relative risk of mortality and grouped them into four severity groupings(26). In the Kadam morbidity index(27), morbidities were categorised using broader definitions of severity characterised by general practitioners through their experience of routine consultations. This index demonstrated that increasing degrees of severity in relation to four morbidity

dimensions (chronicity, time-course, health-care-use, patient impact) were associated with worse physical health. Other work has demonstrated the cumulative effect of overall severity of disease combinations on physical health and symptoms(28,29). This staging of diseases in populations provides useful insight into the overall health needs of patient groups and for planning of health care delivery. However, the caveat to this approach is that within disease groups there will be variation of the patient experience, health and intensity of interventions required.

Measuring severity and change within diseases: Following the reduction in deaths caused by infections, accidents and childbirth, the trajectory of illnesses have changed from sudden and critical to slow and progressive(30) and patients' experience of a disease will vary according to the stage or severity of the disease. '*Within disease'* severity is important in determining the possible outcomes or trajectory of disease. Three distinct illness trajectories have been described for people with three different types of chronic diseases (Figure 1.2) which provide an external measure of disease severity(31). The first trajectory of slow progression and then a clear terminal phase (graph A) is usually associated with cancer where as in end organ failure disease (graph B) there is a gradual decline punctuated by acute episodes of exacerbation requiring intensive treatment and hospital admission. These episodes occur with increasing frequency as the failure progresses. Death can occur at any point and is often less predictable. The final trajectory of steady and progressive decline (graph C) is associated with fragility or dementia.

Another external measure of disease severity is indicated by the processes of care received. Chronic diseases develop along different aetiological pathways and patients will usually present first in general practice with a trajectory of increasing symptoms leading to an initial diagnosis and on-going care(32). More detailed diagnostic or acute care may occur in hospital settings if the disease onset is more sudden or severe or where there is an exacerbation of symptoms during the disease course. Finally end of life care may be a combination of hospital and community care until death.

The heterogeneous nature of chronic disease development with multiple potential causes and varied progression means that disease severity at different points in the trajectory can vary widely across individuals.

Whilst disease and care trajectories provide the average severity path of populations with the disease, internal and external measures of disease severity in individuals provides more detailed information. Internally a range of tests are used in the diagnosis and measurement of individual disease severity and externally, medical interventions such as prescribed medications are common and are likely to change with increasing symptoms as the disease progresses. Whilst these provide potential mechanisms for measuring severity, for the same intervention or physiological measure there will still be variation in the '*patient experience'* which is itself influenced by a range of multidimensional factors.

Measuring severity of disease by patient experience: The manifestation of disease in individuals, which changes and progresses from its onset, will be both determined by and influence a range of multiple and complex individual factors. Concepts of health are multiple and attempt to define the combination of factors by which disease is expressed.

1.2.3 Defining health

Biomedical model to biopsychosocial model of health: Definitions of health have evolved from simple reductionist definitions to multidimensional patient models, yet there is still no clear consensus. Engel argued that the biomedical model which reduced health and its deviation 'disease' to measurable somatic variables, reinforced mind-body dualism and rather than remain a scientific model for investigation, it became a cultural perspective about disease in Western society(33). This was to the detriment of the practice of medicine which aims to reconcile the experience of illness as expressed by patients. Here Engel refers to biochemical defects as being one component of the diagnostic criteria for disease but as inadequate as an explanation for the manifestations of disease. How illness is experienced and reported by patients, its causes and remedies, requires a broader perspective, one which considers psychological, social, and cultural factors.

Extended definitions of health: The World Health Organisation defines health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity"(34). Whilst this reflects the biopsychosocial model purported by Engel it also sets a qualifying standard on health that has been criticised

as failing to distinguish health from happiness and making the achievement of health an unattainable goal(35). Broader definitions still, refer to other more specific determinants of health such as Bircher's definition "health is a dynamic state of wellbeing characterized by a physical, mental and social potential, which satisfies the demands of a life commensurate with age, culture, and personal responsibility. If the potential is insufficient to satisfy these demands the state is disease"(36: p.336). In the term disease he includes malady, illness, ill-health and sickness. Saracci takes an intermediate approach and defines health as "a condition of well-being free of disease or infirmity and a basic and universal human right"(35: p.1409).

These definitions are framed within a concept of the human being as a person whose experience of disequilibrium whether physical, mental or both are presented through a multidimensional lens. This lens is a kaleidoscope of real life contexts including psychological, cultural, spiritual and social influences. The resulting image is the patient's health. This closely reflects 'holism' and the philosophical and practical approach that is embedded within nursing practice in managing a patient's response to health, illness, frailty, and disability(37).

Definition of health for thesis: Core elements of the emerging definitions of health are biological, psychological and social dimensions and the notion that more than one factor impacts on a person's experience and presentation of illness. For scientific enquiry definitions need to be clear, measurable and reproducible. In the context of epidemiology Saracci's definition of health as a state of well-being, free of disease or infirmity, recognises the holistic and subjective nature of health whilst providing a mechanism for measuring health with appropriate indicators such as mortality, morbidity and health related quality of life.

1.2.4 Measuring health related quality of life

Health related quality of life (HR-QoL) includes a range of multi-dimensional measures that summarise the patient's perception of their health experience. These often include assessment of symptoms, ability to perform daily tasks, physical limitations and well as the impact of their illness on behaviours such as self-care and psychological well-being including resilience, illness perception or anxiety(38). General HR-QoL measures overall health status using generic indicators of health which is important for assessing and

managing patients whereas disease specific HR-QoL focuses on the health status attributed to a specific disease and can help to guide the clinical management of the disease.

1.3 Concepts of Comorbidity

Extending the epidemiological approaches from infections and non-communicable diseases is the most recent focus on the presence of multiple diseases occurring at the same time in ageing populations which has become a global healthcare priority(39). These diseases together will determine the overall health trajectory of the patient. Individually or in combination, additional diseases may also influence the progression or trajectory of a patient with a specific disease.

1.3.1 Definition

The terms 'multimorbidity' and 'comorbidity' have been applied interchangeably although they are distinct terms. Multimorbidity describes the co-occurrence of two or more chronic or acute diseases or medical conditions within one person at any one time(40). The illnesses are not prioritised and each may have an impact on the other. The importance for health and health care is on how the illnesses combine to have an overall impact on the patient and their healthcare management. This differs from comorbidity where the focus of interest is on one illness or disease and at how additional illnesses or diseases impact on the presentation, progression, management and experience of the patient with the 'index' disease. Alvan Feinstein in 1976 defined comorbidity as "any distinct additional clinical entity that has existed or that may occur during the clinical course of a patient who has the index disease under study"(41: p.456-457).

In multimorbidity or comorbidity the choice of health 'entities' included in the definition has been categorised by number and type. Morbidity 'type' has included acute and chronic diseases versus only chronic diseases, somatic disorders with or without psychiatric and social disorders and physician diagnosed or patient reported problems(42). Once 'type' of morbidity has been decided some studies have limited the number included to specific morbidities. Broader definitions may include risk factors, conditions, complications or treatments or the wider determinants of health including biological, psychological, social and cultural factors(40).

The broader definition of comorbidity reflects closely the bio-psycho-socio-cultural model of health. Fried(43) argues that to fully understand the interplay or interactions of different diseases or illnesses that the definition should be opened up further still to include multiple physiologic or pathophysiological levels such as impairments, biomediators or subclinical diseases and not simply focus on manifested diseases. However, whilst this has clear importance in understanding causal mechanisms and the pathophysiological and aetiological interactions between specific diseases it is important that the definition of comorbidity is clear and not loosely defined. To an extent the scientific definition of comorbidity will much depend on the focus of the enquiry. The definition may be multi-dimensional including biological, psychological, cultural and social factors or focused on clinical morbidities (symptoms, illnesses and diseases)(27). As with health, this thesis will bring a focus to the definition of comorbidity and will define comorbidity as:

'any clinical chronic disease in addition to the index disease, in this case HF'.

The disease definition of comorbidity as opposed to the broader morbidity definitions provides a mechanism in which to measure comorbidity severity and its change over time following a structured framework that can be applied across diseases.

1.3.2 Impact and importance of comorbidity

The impact of comorbidity or multimorbidity can be measured at various levels from the patient to health services to society more generally and in various settings from the general population to specialist settings.

Patient: Patients with multimorbidity receive fragmented care from different specialists(44) and polypharmacy, adverse drug reactions and inconsistent monitoring are common(45). Self-care is challenging for patients with multimorbidity due to difficulties in symptom recognition, lifestyle modifications and drug adherence(46-49). Multimorbidity and comorbidity have been associated with increased mortality risk, greater symptom burden, limited physical and psychological functioning and reduced quality of life for patients(50-52).

Healthcare systems and economics: People with multimorbidity have increased healthcare utilisation and associated costs(53) and are by far the greatest users of healthcare resources both in primary and secondary care. This healthcare burden increases with the number of morbidities and in the United States 65% of all healthcare utilisation is by people with multiple chronic conditions and two thirds of healthcare expenditure is on people with 5 or more conditions(54). Healthcare systems have been configured along disease specific pathways and policies which cause challenges for the management of people with multiple coexisting diseases. Adherence to individual practice guidelines for each of several different diseases can lead to a high number of drugs prescribed and conflicting management plans and lifestyle advice for patients(55). This can leave patients with a heavy treatment burden and lead to adverse interactions between regimes(56). The increasing burden of multiple chronic diseases for patients then, reflects a growing need for redesign of healthcare service delivery and new quality indicators and interventions that focus on patients with comorbidity.

Public health: It is estimated that 23% of the population have multimorbidity, rising to up to 80% in those over 80 years(51). The number of co-existing morbidities also increases with age with one Swedish study finding an average of 17 coexisting chronic health problems in nursing home residents(57). By 2018 the Department of Health(58) forecasts a rise in the number of patients with multimorbidity to 2.9 million (53% increase over 10 years) and multimorbidity is set to become an international public health priority(39).

1.3.3 Classification

Co-occurring morbidities have been categorised in a number of ways. Simple categorisations include those that occur at the same time regardless of the underlying reason. Cluster morbidities include those that occur together at a higher rate than expected purely by chance. The ratio of observed co-occurrence is higher than that expected by multiplying the risks of each morbidity(59). These morbidities show an association but are not known to have a causal relationship. Causal morbidities lie along the same pathophysiological pathway. These may be disease specific where one morbidity is dependent on another(40).

Comorbidities can also be classified by their time of occurrence. A comorbidity may be a cause or consequence of an index disease or may be unrelated. It may develop before the index disease and be prevalent at the time of diagnosis of the index disease or develop as an incident case during the life-course of the index disease. This temporal relationship may influence the effect of the comorbidity on the index disease or the outcomes of the person with the index disease.

1.3.4 Measuring comorbidity

Comorbidity can be measured in multiple ways from simple counts to index scores to more complex criteria which is usually determined by the purpose of investigation. Early work of Feinstein and Kaplan developed criteria for classifying people with diabetes comorbidity by type (vascular or non-vascular) and severity based on physical function specific to each comorbidity(60). Later Mary Charlson developed a comorbidity Index(26) which scored burden according to the sum and severity of a specified number of diseases. Severity was determined by mortality risk and the index enabled the control of sicker individuals in prospective clinical trials. This index was then adapted by Deyo et al(61) for use in clinical administrative datasets. More focused definitions classify comorbidity by an event such as Elixhauser's definition which defines comorbidity as an additional clinical condition that exists before a hospital admission, that is not related to the primary cause of admission but which is likely to affect outcomes(62) or by Kadam severity criteria within different dimensions of common morbidity in the general population(27).

Index scores of comorbidity can be useful as the number and overall severity of combined diseases is known to influence outcomes including the general health of the patient(25).This also provides important information at the population level for use by health providers for case management, health policy and resource allocation. However, similar to 'between disease' severity, this index score approach to severity of combined diseases does not provide individual or detailed information about the 'within disease' severity of the specific comorbid diseases. Chronic disease comorbidity will change from new onset to more severe disease which will alter its influence on the patient with an index disease. This more detailed information on comorbidity severity has the potential to guide clinical care through modification of the comorbid disease as it progresses using targeted strategies.

1.4 Concepts of prognosis

Central to clinical epidemiology is the accurate prediction of the clinical course of disease. Given the long duration of chronic disease, the factors that are important in prediction and the outcomes that might indicate its clinical course and impact on the patient are multiple and varied. An important consideration in this *prognosis* of chronic disease is the high prevalence of comorbidity and how this influences the outcomes of the patient with the index disease.

1.4.1 Definition

Prognosis is the prediction of the probable course and outcome of specific diseases in populations, groups and individuals. Prognosis research investigates the relationship between future outcomes (endpoints) among people with a given baseline health state (start point) in order to improve health(63). The most credible source of information for prognosis is past clinical observations on similar patients to predict what will happen. Crucial to this is how the clinical observations were made and how they are then interpreted(64).

1.4.2 Importance of prognosis

The prognosis concept has increased in importance in recent years as more people are living with one or multiple chronic diseases. Health care policy makers and providers as well as patients have a vested interest in the outcomes of disease and the efficacy of treatments and care aimed at improving those outcomes. The scope of prognosis research is broad. At a public health level, policy makers require information on the burden of different diseases in order to allocate resources for primary and secondary prevention. The variation in a specific outcome across geographical and clinical contexts for a given disease also facilitates the investigation of systems and quality of care(65).

At a healthcare level, new health technologies for screening, assessment and management may be difficult to test in a randomised controlled trial and prognosis provides a method for investigating their benefit over standard tests or treatments. Where randomised trials are feasible, prognosis provides a useful rationale for testing the effectiveness of new treatments, a good indication of likely events to inform sample size

calculations and a baseline risk from which to translate relative risks back to an absolute scale(63). Clinically, knowledge of the key factors that increase the risk of poor outcomes within disease groups is also important to enable clinicians to modify those factors through targeted interventions or to target groups for treatments. For patients, tailored prognostic information is important so that they can make informed decisions about treatments and care and share decision-making with clinicians as their disease progresses and changes(66,67).

1.4.3 Prognostic research

Four 'PROGRESS' themes of prognosis research have been proposed for methodological development which are (i) overall prognosis (ii) prognostic factor research (iii) prognostic model research and (iv) stratified medicine.

Overall prognosis; describing risk in populations: In the broadest sense (PROGRESS theme 1) 'overall prognosis' research is descriptive and seeks to determine the average risk of outcomes for a given disease in relation to different diagnostic or treatment practices(63). Populations of people within a disease group may be described by different settings or interventions. The quality of healthcare delivery from screening, investigation, diagnosis, therapeutic treatments and care relies on good overall prognostic information(68). The risk or rate of a particular outcome may be measured in a group as a whole or in a sub population who share demographic or clinical characteristics. Part of the remit of overall prognosis research is to understand the variations in average risk between individuals or groups across clinical or geographical settings.

Prognostic factor research; explaining risk in groups: For individual patients the average risk of an outcome is of less importance than the variability in the average and where they fit into the range of risk observed(69). 'Prognostic factor' research (theme 2) refines the overall average prognosis by identifying the risk of an outcome in subgroups of people that share a particular factor. Prognostic factor research may focus on identifying one factor or a range of individual factors that each varies the average risk of outcomes(70). These factors may include biomarkers and clinical, social, environmental, psychological and behavioural factors. Classification of prognostic factors have been more simply defined as a triad of interacting causes (i)

environment e.g. socio-economic, health care organisations or climate (ii) host: e.g. demographic, behavioural, psychosocial, premorbid biologic and (iii) disease e.g. imaging, pathophysiologic, genomic, proteomic, metabolomics(66) If the prognostic factor causes or partly causes the outcome and is modifiable, it can provide an important target for interventions to improve the outcome.

Prognostic factors have a range of uses from refining the diagnosis of diseases, informing treatment recommendations by prioritising those with the worst prognosis with a greater potential health gain or monitoring of disease progression and severity, identification of useful potential components for prognostic models and identification of potential differential treatment responses across groups(70). Prognostic factor research usually starts with exploratory studies based on biological plausibility and once a factor is identified continues with replication of the associations before assessment of the predictive value of the factor.

Prognostic model research; predicting risk in individuals: Prognostic model research (theme 3) *combines* multiple prognostic factors in an attempt to define risk at the individual patient level. Prognostic models, also referred to as prognostic indexes, rules or prediction models, combine the risk associated with a number of factors to provide an overall risk of a specific endpoint over a specified time(71). This risk estimate may be converted to an absolute risk which takes into account the baseline risk or be left as a relative risk or a risk score. Whilst a large number of prognostic models have been developed(72,73) their use in clinical practice has been extremely limited(74). This is partly due to their complexity(75), lack of validation(74), use of novel factors which are not translated into routine practice and lack of inclusion of routinely collected data(71). The potential of a model to influence or modify care will in part rely on its ability to accurately predict an outcome and part on whether the factors included in the model are accessible in practice and amenable to intervention to reduce risk.

Stratified medicine; explaining risks and benefits of interventions: The fourth theme of prognosis research 'stratified medicine' focuses on tailoring treatments to individuals or groups with the most potential clinical benefit or least potential harm(76). Treatment with a uniform effect across two groups may be targeted at the higher risk group who will have the largest absolute benefit. Treatment may also be stratified where it has a

non-uniform effect and a particular patient factor is associated with a change in the treatment effect. Prognosis is important to stratified medicine in order to improve outcomes and reduce harm, develop interventions based on known patient factors, evaluate efficacy in drug trials and to determine the effectiveness of the stratified medicine approach.

1.4.3 Key prognosis concepts

Start points: Prognosis is concerned with predicting the risk of an outcome that is associated with a number of exposures compared to the risk of the same outcome in those without the exposures(64). It is important that the overall prognosis in the unexposed group is as close to those people as possible in whom future predictions will be made(9). A uniform time point in the disease progression should ideally be selected to facilitate this, which means that the disease status in the exposed group and unexposed groups will be comparable. Many studies use a prevalent cohort where there will be wide variation in the 'average' patient and predictive estimates are based on the average overall prognosis.

Endpoints: A study endpoint will be determined by the outcome of interest. Prognosis often focuses on a clear end point such as mortality and much less on other important outcomes such as hospital admissions or quality of life(63). It is becoming increasingly recognised that prognosis needs to address questions that are important to both patients, clinicians and policy makers, which includes patient reported outcome measures (PROMs)(64). The inclusion of PROMs in prognosis is a key indication within international position statements on palliative care(77) but is currently under-investigated(63). In terms of understanding chronic disease progression from new onset to more severe disease, prognosis research needs to identify the factors that influence a range of outcomes across the life course of a disease from patient reported health to hospital admission and to death. Another key consideration is whether the focus is on disease specific or generic outcomes. Disease specific outcomes provide understanding about how factors may be causally associated with the index disease or its progression and guide clinical management of the disease whereas generic outcomes provide important information for guiding the overall clinical management of patients.

Time: The association between exposures and outcomes can be measured at a specified future time point which relies on all patients being followed up for the same time or measured as an average risk per unit of time where patients contribute varied time to their follow-up(78). The specified follow-up time will in part determine the prognosis information used. The time chosen may be short, medium or longer term and will be decided by the research question and life course of the index disease. Longer term prognosis approaches may include static exposures that accumulate risk over time such as age, gender or the presence of comorbidity. Short term prognostic approaches may include some of the cumulative risk factors but in addition will include prognostic factors that identify immediate high risk or unstable states which may include dynamic factors such as physiological markers, recent hospital admission or the current severity of or change in the index disease or the comorbid disease.

Longer term prognosis approaches provide important information that can help plan long term care or public health interventions whereas shorter term prognosis can act as a useful trigger to clinical intervention. Current prognosis approaches that focus on shorter term outcomes have often been hospital based where patients have been admitted due to deterioration in disease and there is a risk of imminent readmission or death. Prognostic approaches for identifying similar instability within the general practice population need to include dynamic exposures that indicate current high risk or instability prior to events but here the evidence is limited.

1.4.4 Prognosis and comorbidity

The growing number of people with comorbidity and its influence on outcomes requires specific prognostic consideration but the present evidence on comorbidity has serious limitations. Prognosis studies are limited to the presence or absence of comorbid disease which ignores the severity and dynamic nature of comorbidity that develops and changes over time. Similar to an index disease which varies in severity from its onset, comorbidity also presents with varying severity and then progresses at different rates in individuals. This comorbidity may have developed many years before the onset of the index disease or subsequent to it. Current prognosis approaches have included the static measure of the presence or absence of comorbidity at baseline but have not yet taken account of the dynamic nature of comorbidity, its severity or change which may be important to individual risk stratification and for the development of comorbidity interventions(79-81).

The dynamic status of comorbidity also provides one mechanism for identifying patients with higher risk of imminent outcomes through an increase or recent change in the comorbidity severity.

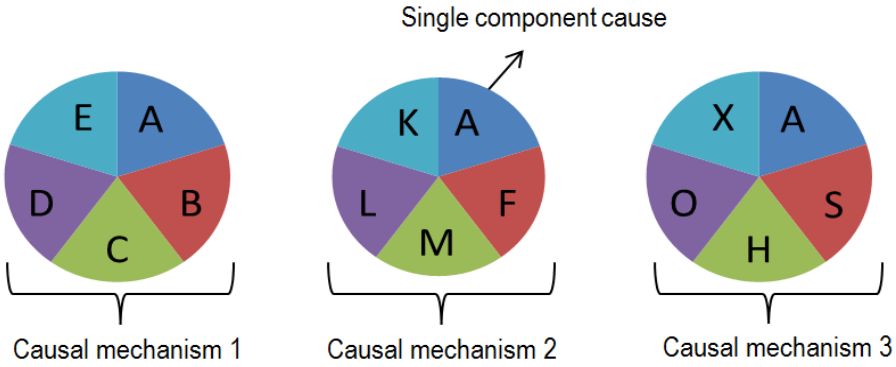
The relationship between comorbid disease severity and patient outcomes is further complicated in the context of multiple conditions where the interrelations between comorbid diseases may result in an unexpected combined effect. As discussed in Section 1.1.3 in the context of biological interaction, two exposures may each have an independent effect on a given outcome (i.e. the two comorbid effects add together) or a combined synergistic (greater than expected) or antagonistic (less than expected) effect(11). This is an important consideration when considering the effects of different chronic disease comorbidities on the outcomes of an index disease.

1.5 Summary

This chapter has discussed the concepts of epidemiology, disease and health, comorbidity, prognosis and the importance of these to patients, healthcare systems and population-level policy. The key questions identified are how to bring the concepts that apply to individual chronic disease clinical epidemiology into the concept of comorbidity clinical epidemiology and prognosis for a given index disease.

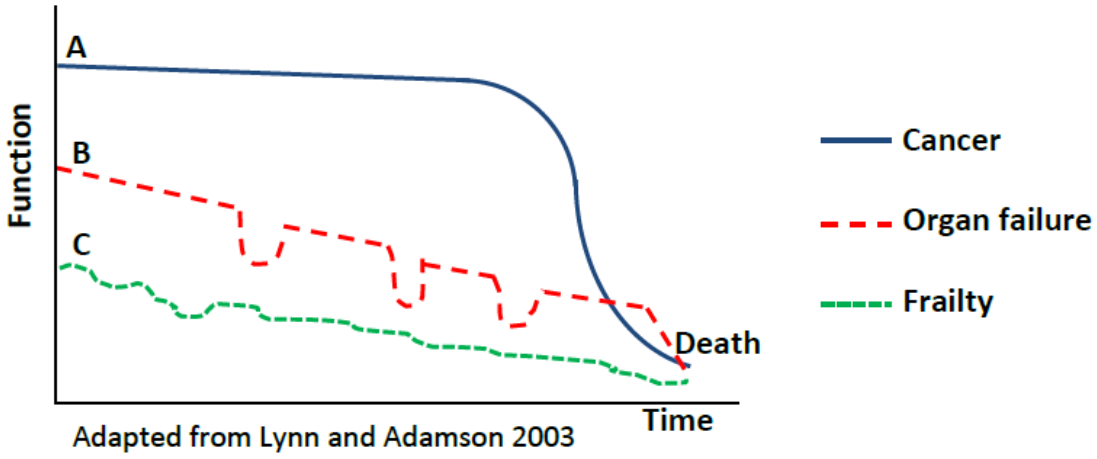
Figures

Figure 1.1 Causal Pie Model(2)



- The causal mechanisms illustrated in pies 1, 2 and 3 each describe a sufficient cause for the same disease.
- Each pie requires all of their constituent parts to be a sufficient cause.
- Pie 1 requires components A,B,C,D and E together for the disease to occur
- Some components, for example component B, are not always required for the same disease to occur and B does not appear in Pies 2 or 3.
- However for some cases of the disease A and B are required to be present together. This is referred to as biological interaction between component A and B.
- Constituent A is part of all sufficient causes and so at a population level is responsible for a higher burden of the disease and would yield a higher strength of cause.

Figure 1.2 Chronic disease trajectories(31).



Chapter 2 Context setting

This chapter applies the broader concepts explored in chapter 1 to the heart failure (HF) population and introduces the current literature on HF, comorbidity and prognosis. The prevalence of HF and comorbidity and their combined impact on public health, healthcare economics, patients and clinical care are addressed. The current evidence on prognosis research in HF and the inclusion of comorbidity will be introduced before providing a summary of the key gaps that underpin the rationale for the thesis and introducing a new framework for the inclusion of comorbidity in HF prognosis in Chapter 3.

2.1 Heart failure

2.1.1 Definition

HF is a complex syndrome, caused by a structural or functional disorder that impairs the ability of the heart to meet the body's physiological demand(82). This triggers a multi-system syndrome that includes abnormal haemodynamic and neurohormonal function, metabolism, energetics and inflammatory activation. As heart function deteriorates patients present with key symptoms of shortness of breath, fluid overload and limitations in activity.

Aetiology: The most common cause of HF is ischaemic heart disease (IHD) which accounts for 70% of cases(83). Coronary risk factors then (such as diabetes, smoking, hypertension) are also associated with the development of HF(84). Other causes cover a range of diverse aetiologies such as hypertension, thyroid disease, cardiomyopathy or valvular heart disease (85). High output HF also occurs when a normal heart can no longer meet abnormal physiological demand(86).

Mechanisms: HF has been associated with two primary mechanisms; left ventricular systolic dysfunction (LVSD) and diastolic dysfunction often characterised respectively as HF with reduced ejection fraction (HFrEF) and HF with preserved ejection fraction (HFpEF). A more recent perspective is that these are two extremes in a spectrum of overlapping phenotypes(87) with diastolic and systolic dysfunction commonly co-

occurring(88). Systolic dysfunction occurs when there is reduced ability of the left ventricle to pump enough blood forward to meet the body's metabolic oxygen demands. Damage to myocardial cells and subsequent myocardial remodelling can lead to poor contractility, eccentric ventricular dilatation and electrical instability. The ventricular dilatation initially compensates for the reduction in cardiac output by increasing end diastolic volume albeit usually with a reduced ejection fraction (the ventricle pumps a lower percentage of a larger volume of blood). Increased end diastolic and systolic volume causes an increase in left atrial and pulmonary venous pressures and post-capillary pulmonary hypertension. Right HF ensues which is associated with signs and symptoms of systemic venous hypertension. Progression to reduced stroke volume and cardiac output following reduced ejection fraction usually occurs(89).

Diastolic dysfunction which has been mostly associated with HFpEF is more common in people who are older, female and have a higher body mass index, hypertension and atrial fibrillation (AF)(90,91). Diastolic dysfunction is caused by the inability of the ventricle to relax(92). Common causes of this diastolic failure are hypertension, coronary artery disease and cardiomyopathy. In primary diastolic dysfunction, initial compensatory mechanisms will maintain the cardiac output by increasing the filling pressure of the left ventricle (end diastolic volume). This in turn leads to increased left atrial, pulmonary and eventually right ventricular pressure and symptomatic HF. Most patients have an increased myocardial wall thickness (concentric hypertrophy) and an increased left atrial size. Eventually as compensatory mechanisms fail, the cardiac output will reduce whilst the left ventricular ejection fraction is maintained (the ventricle is pumping a high percentage of a lower volume of blood). Other mechanisms include abnormalities of the valves, pericardium, endocardium and cardiac conduction(93).

2.1.2 Pathophysiology

Given the complex and often multifactorial and overlapping mechanisms inherent in HF it is difficult to distinguish between systolic and diastolic HF clinically. Whatever the underlying mechanism, HF is characterised by cardiac muscle, skeletal muscle and renal dysfunction and a final common pathway that includes activation of the sympathetic and the renin-angiotensin-aldosterone system (RAAS). These neurohormonal processes that support the heart in normal physiological circumstances play a central role in

the development and progression of HF(84). The RAAS, triggered by hypoperfusion of the kidneys, leads to increased renin, angiotensin II and aldosterone. Collectively these hormones cause vasoconstriction, retention of sodium and water and excretion of potassium, endothelial dysfunction and myocardial fibrosis(82). Sympathetic stimulation causes further vasoconstriction and retention of salt and water. The catecholamines produced contribute to cardiac myocyte apoptosis, hypertrophy and focal myocardial necrosis(84). Reduced baroreceptor function and parasympathetic tone leads to abnormal autonomic modulation of the sinus node.

2.1.3 Classification

Two of the most frequent classifications of patients with HF are via the ejection fraction and symptom severity.

Ejection fraction: Traditionally HF has been classified by the primary mechanisms of systolic or diastolic HF. As previously mentioned (Section 2.1.1) the challenge with this classification is that these impairments usually coexist. Also whilst systolic and diastolic failures are largely associated with reduced and preserved ejection fraction respectively, the relationship is not exclusive. Sensitive imaging may detect mild systolic dysfunction in preserved ejection fraction HF and diastolic dysfunction that coexists with systolic dysfunction is more likely to have reduced ejection fraction. These factors together with the existence of other mechanisms in HF has led to the classification moving away from systolic function to one that is based on whether the ejection fraction is affected i.e. HF with reduced ejection fraction (HFrEF) and HF with preserved ejection fraction (HFpEF)(93,94). Whilst HFrEF forms the largest part of diagnosed HF, epidemiological studies demonstrate that HFpEF is as common as HFrEF and has similar risk of mortality(91,95). Clinically, whilst the primary HF mechanism is associated with the choice of evidence based pharmacological treatments, the aetiology of HF and the common syndrome that ensues is the critical factor in treatment choice.

Symptoms: The HF syndrome is also grouped by symptoms using the New York Heart Association (NYHA) classification(93)(see Figure 2.1). This system, which stages HF patients according to symptom-related physical functional ability, applies to all HF patients and provides a classification system that can be applied to all patients across the disease trajectory.

2.1.4 Diagnosis

Diagnosis of HF is made using a combination of symptoms (e.g. shortness of breath and fatigue), signs (e.g. ankle swelling, pulmonary oedema) and objective evidence of cardiac abnormality (e.g. ventricular function assessment using echocardiogram)(93). One of the challenges of accurate diagnosis in HF is the non-specific nature of the presenting symptoms and signs and this is often worse in women, the elderly and those with comorbidities such as obesity or chronic obstructive pulmonary disease (COPD)(96). For example breathlessness, oedema and fatigue are common generic symptoms which are sensitive but not specific to HF. Physical signs such as raised jugular pressure or a third heart sound may also not be present in less advanced disease and where present, difficult to determine(97). Most are related to fluid overload and may be absent in patients treated with diuretics(93). A critical part of diagnosis is to rule out a broad range of alternative explanations for the cause or exacerbation of symptoms.

Diagnostic criteria will be applied in practice according to the clinician and availability of diagnostic tests. Whilst current guidelines recommend the use of brain or b-type natriuretic peptide (BNP), N-Terminal pro-BNP (NT-proBNP) and echocardiogram(93,94), together with electrocardiogram (ECG), chest x-ray and blood profile, the clinical diagnosis of the prevalent HF population will vary according to the clinical setting of the diagnosis and time the diagnosis was made.

Specialist setting versus general healthcare settings: Ready access to physiological testing within specialist settings means that the diagnosis of HF can be aided by the use of BNP testing, echocardiogram and interpretation by a cardiologist. Patients often present first in general practice(98) with symptoms that develop gradually over time and HF diagnosis has tended towards using more broad criteria(99-101) such as signs and symptoms.

Diagnostic timing: Since 2003, the use of echocardiogram to confirm diagnosis has been widely recommended in national evidence(102) and was introduced into the Quality Outcomes Framework (QoF) for general practitioners (GP) in the UK in 2006(103). The QoF initiated a reward system to GPs for the provision

of quality care and helped to standardise medical practices including the use of echocardiogram for HF diagnosis and registers which also record optimal drug treatment.

The setting and timing of HF diagnosis then is an important consideration for epidemiological research. Whilst different population settings carry the stage of disease trajectory and severity, the definition of the disease itself may vary across settings and within settings as a function of time.

2.1.5 Heart failure disease frequency

Global figures show that there were 5.7 million new cases of HF in 2004 with 3.1 million of those cases from Europe(104) and in Britain the incidence rate has been reported as 1% per annum in women and 1.7% in men aged 85years and over with a median diagnosis age of 76 years(101). Incidence rate increases with age and recent UK statistics show a rise from 0.07% per year for men aged 55-64yrs to 0.3% for those aged 75yrs and over. For women the incidence rate increased per year from 0.03% for women aged 55-64 years to 0.2% for those aged 75years and over(105). In America the rate was higher with 2.2% per annum for women and 2.7% for men aged 80-89 years(106).

The prevalence of HF also increases with age with figures below 3% in the younger age groups (age 55-64 years) and above 12% in the 85 years and over group(100,107). Most recent UK statistics show a prevalence rate of 13.1% in men aged 75years or over and 11.9% women(108). Overall approximately 2% of people in the Western world have HF(109,110).

2.1.6 Impact of heart failure

The demographic of HF has changed over the past two decades with the average HF patient now older with more comorbidities and increased polypharmacy(111). The overall prognosis of HF patients can be determined at the population level or within sub populations that are defined by their severity. Overall HF is a serious, life limiting disease where 14-30% of patients die in the first six months of diagnosis and 38% within a year(83,101). In the ECHOES cohort, the 5-year survival rate of those with prevalent HF was only 58% compared to 93% in the general population(112). About half of all patients experience daily symptoms

including breathlessness and fatigue(113) and up to half of patients remain symptomatic despite optimisation of treatments during hospital admission(114). Patients suffer problems with walking, activities of daily living, self-care, pain and depression. These symptoms are exacerbated with increasing age(115) and consequently the quality of life of patients with HF is poor(116).

Among older people HF is one of the most common causes of consultation in general practice(117) and accounts for 2% of all hospital inpatient bed days, with over 61,000 hospital admissions between 2011-2012 in the UK(118). With an estimated 5% of UK deaths attributed to HF the annual cost to the NHS has been approximated to be over £625 million(119) which equates to 2% of the total NHS budget. The largest proportion of spending is consequent to protracted lengths of hospital stay and frequent readmissions. This figure is likely to rise by fifty percent over the next 25 years as a consequence of an ageing population, improved HF treatments and improved survival following acute coronary events(94).

2.1.7 Heart failure severity and sub populations

As with chronic disease generally, severity of HF disease can be measured at the population level between different diseases, within the disease or by HF related health measures. HF falls at the end of the cardiovascular disease (CVD) spectrum and is associated with worse outcomes than other less severe CVDs such as ischemic heart disease (IHD) or hypertension(29). Between non-cardiovascular diseases, HF remains one of the most severe chronic diseases with a higher mortality rate than many common cancers(120).

Within the HF population, individuals will have a variation in symptoms, outcomes and levels of treatments indicating variation in the severity of HF.

As previously discussed (Section 1.1.3) the absolute effect of an exposure will be determined by its frequency in causal mechanisms of the outcome studied. However the *relative* effect of the comorbidity exposure will depend on the baseline risk of the outcome in the comparator 'unexposed' group. This in turn will be determined by the HF severity in population studied. Where there is no biological interaction between the exposure and HF severity (the absolute effect of the exposure is uniform across different HF severities) the

relative effect of the exposure will reduce as the risk in the baseline group increases(9). Where a relative exposure effect remains the same or increases as the HF baseline risk increases, biological interaction between the exposure and HF severity is indicated(9). This leads to the hypothesis that the relative effects of an exposure will differ across different HF severity groups. This is an important perspective in order to interpret relative measures of exposure risk and to prioritise the targeting of interventions aimed at the index disease or the exposure. To more clearly understand the independent effect of an exposure, the severity of the underlying HF in the exposed and unexposed groups being compared needs to be similar or accounted for in the analyses. Severity of HF disease can be measured by external factors such as settings or stage of disease (such as incident versus prevalent cases) or internal factors such as BNP or ejection fraction which have varying degrees of precision.

2.2 Heart failure comorbidity

Most HF patients have additional conditions or diseases which have been found to impact on their health, their risk of being admitted to hospital and earlier death. An important consideration for prognosis is understanding which comorbid diseases are important for which outcomes in HF, whether the effect of comorbid disease is influenced by HF severity, how the severity of comorbid disease and its progression influences outcomes in HF and how different comorbid diseases might interact to influence the progression of the disease and experience of the patient with the disease. This information can then be used to predict and furthermore prevent future poor outcomes.

2.2.1 Measurement and type of comorbidities

Current understanding on the prevalence, interrelations and consequences of comorbidity in HF is determined by the definition and measurements used so far in scientific enquiry. A systematic review of comorbidity measurement in cardiovascular disease studies identified inconsistent terminology and measurement of comorbidity across 27 CVD studies which remained unchanged over time (1965-2009). Each individual study used its own definition of the comorbidity concept. A few studies (n=5) used an index score to measure comorbidity burden, mostly the Charlson Comorbidity index(26). Most studies used the presence or absence

of comorbidities but did not provide rationale or a conceptual framework for the selection of comorbidities included(121).

In HF studies the focus so far has been on the underlying causative cardiovascular disease (CVD) comorbidities such as hypertension and IHD which tend to fall along a spectrum of CVD status(122). These comorbidities have direct pathophysiological links and are associated with the progression and severity of the HF syndrome(123,124). Within an ageing population, the prevalence of non-CVD comorbidities in HF have increased and now predominates over CVD comorbidities in older patients over 85 years(125). The importance of non-CVD comorbidities on HF mortality outcomes is also becoming well recognised and may be a consequence of shared risk factors or common pathophysiology with the progression of one disease having a direct or indirect effect on the progression of a co-existing disease such as HF(126). Despite the growing recognition of the importance of non-CVD comorbidities in HF, their management and monitoring is lacking in routine care and they feature far less in prognosis assessment of HF than other cardiovascular comorbidities (see Section 2.3). This may be due, in part, to the evidence gap on how their severity and progression over the course of HF influences outcomes for the patient (Section 2.2.3.2) which is the focus of this thesis.

2.2.2 Frequency of heart failure comorbidities

Older age in HF increases the chances of developing a range of both cardiovascular and non-cardiovascular comorbidities. A range of non-cardiovascular comorbid diseases have been found in HF patients including chronic kidney disease (CKD), chronic obstructive pulmonary disease (COPD), dementia, diabetes mellitus (DM), osteoarthritis and osteoporosis(126,127) and older patients have been found to have in excess of five non-CVD comorbidities(128,129). Prevalence of different comorbidities and patterns of their co-occurrence are similar for HFrEF and HFpEF with a slightly higher burden and prevalence in HFpEF(130,131). Prevalence does more markedly increase with the severity of HF disease(129,131) with higher levels in hospital settings versus community settings(132).

Many studies have focused on individual cardiovascular comorbidities in HF but fewer studies have investigated multiple non-cardiovascular comorbidities(129-131,133). A multi-regional European study of

3,226 acute 'de novo' or chronic HF patients (outpatients, mean age 66 years)(131) showed a high prevalence of non-cardiac comorbidities with the highest prevalence encountered for hypertension (58%), CKD (41%) and DM (29%). In another large cross-sectional study of 122,630 Medicare HF patients (>65years)(129) similar high prevalence was found (hypertension 55%, DM 31%, COPD 26%). The prevalence of comorbidity in part depends on the measurement used with self-reported or clinically diagnosed comorbidities indicating higher levels than figures based on clinical administration data(111) (see Table 2.1).

More recently there has been a focus on comorbidities that commonly co-occur in HF populations. These clusters are identified when the observed prevalence of two or more HF comorbidities combined is greater than that expected by chance. In a study of 417,477 adult HF hospitalisations, comorbidity profile groups were identified including; few and diffuse comorbidities (46.6%), metabolic comorbidities (19%), endocrine and hematologic comorbidities (29%) and vascular comorbidities (5%)(134). Another study of 23,435 hospital or outpatient HF patients identified that certain pairs of comorbidities commonly coexist at a greater rate than expected. DM and hypertension occurred at a rate of 25% more than expected, hypertension and dyslipidemia in 21% and hypertension, dyslipidemia and visual impairment in 56% more than expected(130). Understanding which comorbidities co-occur is important to guide clinical decision making and guidelines. What is less well studied is the interaction or impact of co-occurring comorbidities on prognosis outcomes in HF.

2.2.3 Impact of heart failure comorbidities

Comorbidity for the HF patient can lead to altered HF symptoms and delayed diagnosis, complicated and conflicting treatment regimens and poor tolerance or reduced prescription of evidenced based pharmacotherapy(135) including fewer-blockers and angiotensin converting enzyme inhibitors (ACEi)(131). Patients have been found to have difficulty understanding their symptoms and may attribute worsening HF to another co-morbid disease or simple ageing(50,136) which can delay health seeking behaviour.

2.2.3.1 Current HF comorbidity prognosis evidence

Non-CVD comorbidity has been shown to influence the clinical course of HF and associated outcomes to vary. However, the relative impact of comorbidities on outcomes in HF will be influenced by which diseases are present, the severity of the underlying HF and comorbid diseases and the outcome measured. In the few studies that have investigated a range of non-CVD comorbidities and outcomes in HF the focus has been on the presence or absence of comorbidity identified by clinicians, patient self-report, or administration data yet, the severity of the comorbidity is rarely taken onto account.

Comorbidities in HF are significantly associated with increased risk of mortality outcome. In 3,226 HF outpatients followed for a year in Europe(131), a significant association was found between non-CVD comorbidities and increased mortality risk, a finding shared with a previous cohort study of 9,442 Veterans followed for a two year time-period(133). In this latter study, significant associations with increased mortality in HF were found for CKD, COPD stroke, liver disease, diabetes, cancer, anaemia and dementia. In a combined hospital and outpatient setting a cross sectional study of 122,630 HF patients found respiratory disease, CKD, COPD, depression and DM to be significantly associated with increased risk of mortality(129). Hospital admission risk similarly increases in the presence of HF comorbidity with significant associations found for CKD and DM(129,131), respiratory disease, CKD, Alzheimer's disease, COPD, asthma, hypertension, depression and stroke(129). Depression, DM and respiratory disease are also associated with reduced physical function and quality of life in HF patients (137).

Whilst '*within comorbidity disease'* severity is rarely taken into account in prognosis studies, the notion of 'severity' has been investigated in terms of the number and type of disease. The number of comorbidities has been found to be important for HF outcomes including increased symptom burden(52) and when more than three comorbidities are compared to no comorbidities, increased risk of mortality and hospital admission(131). Index comorbidity scores which take the number and severity of multiple comorbidities into account have also shown significant associations with mortality. Relative mortality risk significantly increased when HF patients with more than 4 points on the Charlson Comorbidity Index were compared to those with 1-2 points(138).

Influence of HF severity: As the severity of the underlying HF increases the relative effect of the comorbidity will reduce, unless biological interaction is present. Biological interaction between the comorbid disease and HF severity may counter the increased baseline risk and maintain the relative effect of the comorbidity as the HF progresses. In a cohort study of 18,322 Medicare beneficiaries with HF, Ahluwalia et al compared the more severe HF group (defined by those with a first hospital admission) to the less severe HF group (no hospital admission)(132). The more severe group had a higher burden of CVD and non-CVD comorbidities which had significant associations with mortality with CKD, COPD and dementia conveying the highest risk. This demonstrates the importance of comorbidity even in the most severe HF disease. However, whilst still significant in the more severe group, the relative risk associated with myocardial infarction, lung cancer, CKD, dementia, COPD, stroke and arthritis was significantly lower than in the less severe group. There was no difference in the comorbidity effects on mortality across HF severity groups for IHD, other cancers and DM indicating possible interaction between these comorbidities and the HF severity. Another study of 8,507 hospitalised HF patients found that more severe HF, this time defined by older age (above 85 years of age), compared to younger age groups, were associated with lower relative comorbidity effects for CKD and DM(139).

Influence of HF mechanisms: The question of whether comorbidity effects differ across HF groups defined by mechanisms has generated considerable research interest. Prevalence of different comorbidities are slightly higher in HFpEF than HFrEF(140) and patients with HFpEF tend to die more of non-CVD comorbidities (60%) than those with HFrEF (36%)(141). In a study of 98 hospitalised patients with HFpEF, diabetes and renal dysfunction were significantly associated with outcomes (death and/or HF admission) and yet cardiac parameters showed no association(142). This raises the question of whether the effects of non-CVD comorbidities are greater in the HFpEF group. Where this has been tested, with the exception of COPD, there was no significant difference in effects of a range of comorbidities on mortality across the two ejection fraction groups(133).

2.2.3.2 New evidence on HF comorbidity and prognosis

Prior evidence on cardiovascular comorbidity in HF such as hypertension or ischemic heart disease has demonstrated that the severity and progression of the cardiovascular aetiology is associated with worse outcomes in HF. Examples include significant associations between the level of left ventricular dysfunction and HF admissions(143) and between cardiovascular events and mortality in HF patients(123,144). There is also developing evidence that the severity of the non-CVD comorbid disease and its change overtime increases its absolute effect on outcomes and may be an important consideration for prognosis.

Comorbidity severity evidence: Two prior HF systematic reviews have demonstrated the importance of renal dysfunction severity for mortality. In the first review, eleven studies stratified renal dysfunction by severity as measured by levels of estimated glomerular rate (eGFR). Unadjusted associations showed an estimated 33% increase in all-cause mortality risk from 'any' to moderate/severe renal impairment. A linear relationship was also reported between increasing serum creatinine, decreasing eGFR and all-cause mortality(145). The second review included 57 prior studies and found, in adjusted analyses, that moderate renal impairment was associated with a 59% increase in mortality risk compared to those with normal renal function. This increased to a 117% increase in risk associated with severe impairment(146).

Few other non-CVD comorbid diseases have been investigated for severity and outcomes in HF. In 348 HF patients, the severity stage of COPD measured by spirometry was significantly associated with mortality with increasing risk from lower to higher severity groups(147). This finding was consistent with a study of 184 HF patients with comorbid COPD where stage 3 COPD severity compared to stage 1 was associated with a 220% increase in mortality risk(148). Diabetes severity staged by treatment type has been found to be an important prognosis factor in other cardiovascular studies with increased mortality risk for myocardial patients found from no medication to oral hypoglycaemic to insulin(149,150).

Comorbidity severity change evidence: Current interest in 'cardiorenal' syndrome which focuses on the temporal relationship between renal and cardiac function has demonstrated a close link between worsening renal function and reduced survival in patients with HF(151-156). In two further renal HF systematic reviews

worsening renal function (WRF) was shown to be associated with a combined estimated 62%-95% increase in mortality risk and 30% increase in hospital admission risk(146,157). This effect was similar in outpatients and hospitalised patients and the associated risk increased with the severity of the change in renal function. Whilst renal severity and change may act as a marker for worsening HF, the dose-response relationship together with the strength and consistency of the renal evidence suggests that a causal relationship may exist. This causal relationship has biological plausibility given the common pathophysiological pathways in cardiac and renal disease(158).

Broader evidence that change in a comorbid disease or condition may precipitate or contribute to outcomes in HF is provided by studies that investigate the cause of hospital admissions and death. One study of 27,477 admitted HF patients in Scotland found that HF was the principal diagnosis in only 42% of all hospital discharges or deaths(159). An estimated 11·8% of discharges and deaths were associated with COPD, 8·3% had either chronic or acute renal failure, 5·3% had had a stroke and 15·4% had atrial fibrillation. More recently a large national sample of admissions in America showed that the number of admissions with HF as the principal diagnosis had reduced by 10% over a decade compared to the same size increase in admissions with HF as a secondary diagnosis. These studies indicate that change in comorbid states (indicated by need for, and cause of admission) contribute to all-cause hospital admissions and mortality in HF.

Other work has shown a link between acute non-CVD comorbid events and HF specific hospital admissions or death. In 25,090 HF patients in Quebec, acute gout episodes (defined by hospital admissions for gout within 60 days of the event) were associated with an estimated 177% increase in risk compared to those with no acute gout(160). Interestingly the effect of longstanding gout (present at baseline) had less effect on HF outcomes (estimated 63% increase in risk). This evidence demonstrates a potential link between a change in comorbidity status prior to an event and change in HF status.

The collective evidence on severity and change generates the hypothesis that weighting of non-CVD comorbidities by their severity and its change over time may provide better risk stratification for identifying patients with the worst prognostic outcomes. The dynamic nature of comorbidity severity and its potential for

prognosis is important for three main reasons. First it provides a mechanism for identifying HF patients whose prognosis is changing so that interventions can be triggered. Second, comorbidity is routinely monitored and recorded in general practice. Third, comorbidities are potentially modifiable by targeted interventions so that risk can be reduced.

2.2.3.4 Importance for Healthcare systems and Economics

Healthcare utilization increases for HF patients in the presence of comorbidity(129,161) and the cost of hospital admission increases. Out of 122,630 HF patients above 65years, patients with more than 5 comorbidities were responsible for 81% of all preventable hospital admissions(129). As well as higher users of healthcare, HF patients with comorbidities have longer lengths of stay(162) and higher cost per admission. This has been shown to differ by common co-occurring groups of comorbidities with a study of 417,477 adult HF hospitalisations, showing that patients with metabolic comorbidities cost 10.2% more, endocrine and hematologic comorbidities cost 16.7% more and those with vascular comorbidities cost 21.4% more than patients with few or diffuse comorbidities(134).

2.2.3.3 Importance for Public Health

HF policy and national guidelines have to date focused on the management of HF aetiology, pathophysiology, cardiovascular comorbidities and symptom management. The majority of evidence for the benefits of pharmacological treatments in HF is derived from large clinical trials that have selected patients on the basis of LVSD and HFrEF(87) and excluded patients with complex non-CVD comorbidities. These trial populations do not reflect the general population of HF patients with common comorbidity(163). The application of these guidelines can be challenging as there is little evidence of potential risks and benefits in patients with non-CVD comorbidities or on how to reconcile optimal treatment in the presence of co-existing chronic disease, the management of which may conflict with that of HF. HF is a clear example of where guidelines are focused on disease specific targets and quality indicators. Whilst latest guidelines have begun to recognize the co-existence of a few common non-CVD conditions in HF(93), the growing prevalence and importance of non-CVD comorbidity has yet to be recognized within government policy. Better understanding of which non-CVD

comorbidities are important and how the severity and progression of non-CVD comorbidity influences HF outcomes is required for health resource prioritisation, better prevention strategies and to develop and test interventions that target the comorbid disease at the population level.

2.3 Prognostication in heart failure

Prognostication in HF is challenging for clinicians due to its complex disease trajectory (see Figure 1.2, trajectory B) and this is further exacerbated by the high prevalence of comorbidity. Following initial diagnosis and management, patients can have a period of relative functional stability, before a gradual and chronic decline in health status, interrupted by acute episodic exacerbation of symptoms and frequent hospital admissions. Sudden death at any point is common(77). The variation in prognosis for individuals with HF is poorly predicted by clinicians(164) or patients(165) and this unpredictable nature has led to the reluctance of some health professionals to discuss prognosis with patients. Subsequently patients' preferences in clinical care have been poorly represented in healthcare decisions(166-168).

Prognosis is important in HF in order to equip patients and clinicians with the information required for decision making and to identify high risk groups for targeting interventions. The trajectory of severity in HF means that interventions escalate from the prescription of pharmacotherapy to implantation of devices to invasive monitoring and intravenous diuretics. The most severe patients will be considered for a left ventricular assist device (LVAD), continuous intravenous inotropes or heart transplant(169). Each intervention has an associated mortality benefit tempered by associated side effects, costs and limited resources. The prognosis of the HF patient is important not just to trigger but also to tailor interventions according to risks and benefits and evaluate interventions across different groups.

Current HF prognostic tools to assist clinicians commonly rely on complex and invasive clinical data and biometrics(170,171), are often developed using selected hospital based(172-175) and clinical trial patient populations(80,175-179), focus mainly on mortality and do not relate to the usual general HF patient with non-CVD comorbidity. Whilst there is a range of evidence across the 4 themes of prognosis research this

discussion will be restricted to prognostic factor and prognostic model evidence which are the focus of this thesis.

2.3.1 Heart failure prognostic factors

In HF, there are a range of different factors that have been associated with outcomes. They are indicated in E-Appendix A1 in accordance with the previously defined triangle of interacting causes: environment, host and disease(66). There are a range of factors in each of the three domains and whilst the strength of evidence varies across different factors the most common and consistent factors are presented(180,181).

A prior review of prognostic factors focusing on host and disease characteristics for predicting hospital readmission in HF included 112 studies and summarized 26 different factors(182). Most studies were single site studies that focused on one factor whilst adjusting for others. The most frequent host factors included were age (81%), gender (71%), ethnicity (35%), CVD comorbidities (hypertension 41%, coronary artery disease 34%, AF 27%) and DM (46%). Common HF disease factors were EF (56%), NYHA (35%), blood urea nitrogen or creatinine (45%), sodium (25%), BNP (21%) and haemoglobin (19%).

2.3.2 Heart failure prognostic models

Two systematic reviews have summarised six HF prognostic models for hospital readmission(182,183). Five of the models included only patient (host and disease) level characteristics and the only common factors that predicted readmission across the models were history of diabetes and a history of prior hospital admission. Only one model used a validation cohort and there was minimal agreement across all six models of the factors that predict readmission. Only two reported discrimination for hospital readmission risk which was poor.

Three systematic reviews included prognostic models for mortality and readmission comparing the number, type and predictive power of the variables used and the models constructed. The first systematic review of prognostic models for mortality was restricted to ambulatory HF patients and included 20 models in 34 studies(184). Only 5 models were validated in an independent cohort and were the focus of the review (see

summary table in E-Appendix A2). The models showed poor to moderate discrimination in the validation cohorts who were mostly younger, male and with a low ejection fraction. The most widely validated and used models (HF Survival Score (HFSS)(185) and Seattle HF Model (SHFM))(177) were derived from populations predating 1994 which may explain their lower level of discrimination in more contemporary cohorts. Both models focus on cardiac parameters and neither model includes non-CVD comorbidity. In the 15 non validated models half had focused on specific populations defined by ejection fraction.

The second and third systematic reviews were similarly focused on predictive models for mortality and/or readmission in HF. The second review identified 117 different models in 55 papers for both mortality and/or hospital admission(186). Most of the models were focused on hospital settings for the outcome of mortality. The number of variables included in the models ranged from 1 to 65 with a total of 249 different variables used, the most common being age (64%), gender (44%), systolic blood pressure (42%), sodium (38%), diabetes (34%) and creatinine (31%). Models showed moderate to poor discrimination in specific populations for hospital admission outcomes, hospital admission or mortality and only slightly higher discrimination for mortality only. Models with higher discrimination identified through meta-regression were those produced in prospective cohort/registries, using medical records data and using more variables(186).

The third systematic review(187) identified 64 main models with an additional 50 modifications of the main models. Again, these models were focused on mortality in hospital settings, over half were American and less than a third of the models were based on routine clinical data. The number of variables included ranged from 3 to 314 with a median of 9. In addition to sodium; age, renal function, blood pressure, ejection fraction, sex, BNP, NYHA , diabetes, body mass index, gender and exercise capacity emerged as the strongest predictors. Model discrimination was again higher for the mortality outcome with shorter follow-up.

The two most promising HF prognostic models for ambulatory patients with comorbidity are the MAGGIC model(188) and the 3C-HF model(189). The MAGGIC model, although based on mainly trial patients, does include a range of comorbidities and was well calibrated for mortality. The 3C-HF model was based on hospital discharge patients and whilst it includes two non-CVD comorbidities it mostly focused on HF severity

indicators and short term mortality following hospital admission. The biggest gap in these models, like most available prognostic models, is that they mainly include static measurement of baseline measures including comorbidity to predict future outcomes.

Whilst dynamic measures that indicate the level or severity of exposure have been included in prognostic models the focus has been on cardiovascular exposures such as functional status(190), NT-proBNP(79), NYHA(179), ejection fraction(191,192) or risk factors for example blood pressure and body mass index(188). More contemporary approaches have begun to consider change in CVD exposure using multiple measures in prognosis which have been shown to be favourable compared to single measures. These have included time-series measurement of cytokines which were found to significantly improve the discrimination of 1 year mortality in advanced HF over baseline measurements(193), change in cardiac symptoms over two weeks which had substantial value for predicting hospital admission(80) and change in BNP(79). In another study a dynamic risk score was calculated daily in HF patients with an implantable device using diagnostic parameters that fluctuate. The peak risk category over the month was strongly associated with 30 day hospital admissions although the event rate was low(81).

The discrimination of HF prognostic models has been found to be better for HF-specific mortality than all-cause mortality(187) which may be due to the inclusion of more cardiac specific exposures. Discrimination also improves when dynamic measures of change relating to these CVD exposures are included(79,80,193). Despite the growing evidence on the importance of non-CVD comorbidity severity and change, with the exception of renal severity, dynamic measures of comorbidity severity or change have yet to be considered in HF prognostic models. This is important as non-CVD comorbidity is common in the general population of HF patients, is routinely monitored and recorded, changes in severity as the disease progresses and may act as an important and modifiable prognostic factor for both HF specific and generic outcomes.

2.4 Summary

In summary, HF is a serious chronic disease with poor outcomes. The risk of poor outcomes varies considerably between different groups and individuals. Prognosis is important to identify those at the most risk so that interventions can be optimized and information clearly communicated to patients. Non-CVD Comorbidities are extremely common in HF and whilst evidence has shown their importance for mortality there is an important gap in terms of how their severity that changes over time influences their association with the range of outcomes in HF. Current prognostic models have been developed on selected hospital based patients for the outcome of mortality and not for the general population of HF patients with comorbidity. Ambulatory models include mainly selected trial patients where comorbidity may be excluded. Where comorbidities are considered they are included as a static measure which ignores the reality of their change over time that may be important to risk assessment. Inclusion of the dynamic nature of exposure such as non-CVD comorbidity has the potential for identifying ambulatory HF patients whose risk is changing so that timely interventions can be triggered to improve outcomes.

Tables

Table 2.1 Comorbidity prevalence in HF

	Europe(131) Out Patients (OP) Mean 66years, N=3226	US(130) Hospital and OP ≥65years, N=23435		US(133) Ambulatory, Veterans (95% men) Mean 70yrs, N=9442		US(129) Hospital and OP ≥65years, N=122,630	US(111) Ambulatory ≥65years, N=534
Comorbidities	All (%)	HFpEF(%)	HFrEF(%)	HFpEF(%)	HFrEF(%)	All(%)	All(%)
Mean no.		4.5	4.4	4	3.5		>5
CKD	41			48.8	51.9		45.9
Stroke	11	23	22.4	21	21.3	3	19.4
Diabetes	29	19.3	19.1	44.5	40	31	38.3
COPD	15			33.9	26.6	26	30.9
Lung disease		32.6	28.9				
Hypertension	58	58.5	59.5	70.5	62	55	73.3
Anaemia	29	59.8	52.1	33.2	28.4		22.2
AF	40	28.2	22.1	35	35.4		
CHD		16.1	22.4	27.1	40.4	51	
Depression		17.7	17.2			8	
Arthritis						16	62

Data are number patients (%). HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction. CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; AF, atrial fibrillation; CHD, chronic heart disease

Figures

Figure 2.1 New York Heart Association (NYHA) classification(93)

Class I	No limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, dyspnoea (shortness of breath).
Class II	Slight limitation of physical activity. Comfortable at rest. Ordinary physical activity results in fatigue, palpitation, dyspnoea (shortness of breath).
Class III	Marked limitation of physical activity. Comfortable at rest. Less than ordinary activity causes fatigue, palpitation or dyspnoea.
Class IV	Unable to carry out physical activity without discomfort. Symptoms of heart failure at rest. If any physical activity is undertaken, discomfort increases.

Chapter 3 Conceptual framework and thesis overview

This chapter summarises the key concepts relating to HF comorbidity and prognosis using the evidence cited in Chapter 1 and 2. The evidence gaps in relation to the concepts are highlighted and a new conceptual framework for the inclusion of comorbidity within prognosis is presented.

3.1 Summary of evidence and key gaps

The number of people with chronic disease such as HF is set to rise in an ageing population. HF is one chronic disease that carries a high mortality risk similar to many common cancers and yet the trajectory of disease for those with HF is less clear. This means that people with HF tend to live with uncertainty and clinicians are unable to plan timely interventions to provide optimal care for those with HF. Comorbidity also increases with older age, further complicating the HF disease course and is known to have an important influence on prognostic outcomes.

HF prognostic research has so far focused on the risk of limited outcomes such as mortality using baseline measurement and static prognostic factors such as the presence or absence of comorbid diseases. This may not be the best approach for exposures that change over time where this change may alter the risk associated with the exposure. Exploratory prognostic factor studies that do not take this into account will result in estimates of risk that are based on an 'average' of this change leading to poor individual risk prediction. Non-CVD comorbidity in HF predominates over CVD comorbidity but has been less well investigated or included in HF prognosis. Investigation of non-CVD comorbidity severity and change in HF prognosis will provide the potential for better risk stratification of more sensitive outcomes such as transitions across the specialist care interfaces and changes in health related quality of life.

Key gap - comorbidity severity: Severity of HF disease has been clearly indicated as a risk factor for a range of outcomes and is included in prognosis models as biochemical or physiological markers or symptoms

classification. Evidence, mainly in hospital populations, on renal disease comorbidity and limited evidence on diabetes and chronic obstructive pulmonary disease comorbidity has demonstrated an increased risk of death in HF with increased severity of the comorbid disease (see Section 2.2.3.2). With the exception of renal status as measured by kidney function the evidence on non-CVD comorbidity severity has not been translated into prognosis and provides the potential for more sensitive risk stratification. A key question is:

- *What is the effect of non-CVD comorbidities on outcomes in HF stratified by their severity?*

Key gap - comorbidity severity change: A key epidemiological challenge is how the exposure should be defined. The traditional approach to prognosis is to investigate one or more factors in a model at baseline. These factors are then tested to predict the occurrence and level of risk of an event at a future point in time or overtime. However the baseline model approach does not take account of exposures that develop or change over time (time-dependent exposures) or exposure effects that vary over time (time-varying effects)(194,195). This omission is important and particularly in chronic disease populations where the disease and associated clinical factors develop and change overtime. Age is a classic example of a factor that changes overtime but is taken account of by the follow-up time included in cohort analysis. Other factors such as chronic disease comorbidity that may *develop after baseline* or *change over time* once developed may have important significance for outcomes but are not currently considered. Two unanswered and important questions are:

- *What are the effects of comorbidities that develop before or after index HF on hospital admissions and death in HF?*

- *Does the effect of comorbidities on outcomes in HF differ by their timing of development before or after index HF?*

One approach to measure exposure that changes overtime that can be accommodated in analytical models is to use a cumulative approach that sums up the exposure or exposure time over the follow up period. One example may be the total daily dose of a drug and the number of days exposed(196). These approaches

require routinely recorded repeated measures of exposure and this can be challenging when using long term follow-up data or in quantifying *change* in exposure as opposed to quantity. A limitation of cumulative approaches is that they sacrifice information about potentially important patterns of exposure change that may have epidemiological importance. This issue of change over time generates the hypothesis that change in comorbidity severity that occurs in the recent time before an event may be an important prognostic factor in HF. A key question is:

- *What is the effect of comorbidities on outcomes in HF stratified by recent change in their severity?*

Key gap - comorbidity interaction: An important consideration for the inclusion of comorbidity into prognostic models is where people experience several different comorbidities simultaneously. Multiple comorbidities may have a combined effect that is additional to their independent effects(29) but this is not well studied within the HF comorbidity evidence. Common comorbidities that share pathophysiological links may provide important combined prognostic information that further stratifies risk for individuals with multiple comorbidities. A key question is:

- *Is the effect of two comorbid diseases on outcomes in HF different from the sum of their independent effects?*

Key gap - prognosis outcomes: Available prognostic models in HF often focus on the risk of mortality. This means that the available tools have limited use in identifying HF patients' change across the disease life course from stability to instability, to hospital admission and to death. Investigation of hospital admissions and death as two separate outcomes provides the natural comparison to use these HF trajectory points as indicators of increasing severity of HF status. It is hypothesised that the factors that predict change in HF status as indicated by hospital admission and death would also predict change in HF patient reported health related quality of life. A key question is:

- *Does the effect of comorbidities differ by the outcome investigated in HF patients?*

Key gap - prognostic models: Current prognostic models in HF are often developed in specific select populations such as those defined by HF phenotypes (HF with reduced or preserved ejection fraction) or hospital cohorts. There are limited models developed for use in the general practice population and where comorbidity is included, it is defined by the presence or absence of disease. A key question is;

- *Do comorbidity severity and change indicators improve a predefined HF prognostic model?*

3.2 Conceptual framework

The preceding chapters and key gaps identified in the evidence suggests that there needs to be a clear framework for incorporating comorbidity into new prognostic models with the potential future application in clinical practice. A new conceptual framework for comorbidity and prognosis will now be proposed in response to the key gaps identified. The overall components within the framework for the index disease and comorbidity are status, timing, severity, severity change and disease interactions to influence the prognostic model based outcomes over the life course (Figure 3.1).

Status: The simplest approach to incorporating comorbidity in prognosis is where the presence or absence of chronic disease at the point of measurement, usually baseline, is included. Whilst all people will have the index disease which is the focus of enquiry, other specific chronic diseases are included which may or may not be present with the index disease. The comorbid diseases selected for inclusion should be based on their common prevalence in the people with the index disease and importance for the outcomes of interest.

Timing: A key choice in the measurement of the index disease is whether it is incident or prevalent. By choosing incident disease it allows the comorbidity to be measured according to the life course of the index disease. Of interest is whether comorbid disease that develops prior to the index disease (prevalent comorbidity), that may be associated with the development of the index disease or unrelated, has a different

prognostic effect than comorbidity that develops after the onset of the index disease (incident comorbidity). Incident comorbidity may be a consequence of the index disease through pathogenic links or unrelated.

Comorbidity severity: Severity of both the index disease and comorbid disease can be measured by indicators that are external to the disease such as disease duration, pharmacological interventions or healthcare use such as recent hospital admission for the disease. The HF index disease duration for an incident cohort will be determined by the follow-up time. Internal factors will include physiological markers of disease status such as blood test indicators, for example biochemical markers, or functional test indicators such as ejection fraction or peak exercise oxygen concentration in HF. To an extent the measures selected will depend on the epidemiological design and available data. Administrative clinical databases now provide a continuous record of patients with chronic disease as they present in practice and as their diseases develop over time. Information including diseases, drugs and tests are routinely recorded by general practitioners and provide an appropriate mechanism for measuring comorbidity severity which can be incorporated into prognostic approaches.

Comorbidity severity change: Change in comorbidity severity indicators over the progression of the index disease or at specific points in the index disease trajectory may provide important prognostic information for the occurrence of a future outcome. Again the indicators selected will depend on their availability. Many chronic diseases are subject to continuous monitoring as part of quality indicators and care(197). This means that whilst the frequency of severity measures may differ across different diseases and measures there are usually multiple records for each patient which provides the mechanism for measuring change over time. Drug prescriptions that can provide an indication of disease severity are one of the most consistent and frequently recorded measures.

Interaction: Interaction, which is how two diseases combine to influence outcomes, should be considered to estimate risk for people with multiple comorbidities. This measure of interaction can be assessed on the disease status or in relation to the severity measures of different diseases. For example there may be an interaction between two diseases when they are at their most severe. Interaction between the comorbidity and

the index disease is also of importance and particularly to understand the influence of comorbidities on the progression of the index disease itself (as opposed to the patient outcomes). This approach requires a cohort of patients with and without the index disease(29). Another approach to investigating the interaction between an index disease and comorbidity is to investigate the influence of the comorbidity on index disease-specific outcomes where these are available.

Life course outcomes: Typically prognostic models focus on mortality outcome. Yet, in understanding chronic disease progression and change, different outcomes are of interest. The trajectories outlined in Chapter 1 suggest that disease severity changes over time (whether the index or the comorbidity) so the principles of life-course outcomes(198,199) need to be applied. These outcomes can range from those of importance at the new onset of disease e.g. patient reported quality of life, to more severe disease e.g. hospital admission and finally to death.

Prognostic model: In addition to comorbidities, prognostic models need to include the range of multidimensional factors that are known to influence risk in individuals. These factors should represent the wider definitions of health and comorbidity and include a range of factors from the environment, host and disease. These factors will be guided by prior evidence and clinical experience of the determinants of health and outcomes for the patients with the index disease (Figure3.1).

3.3 Aims and Objectives

The overall aim of this PhD thesis was to investigate the influence of non-CVD comorbidities on prognostic outcomes in HF. It investigated the influence of other diseases on the likelihood of hospital admission and death and developed new comorbidity prognostic factors for inclusion in prognostic models for transitions from general practice to hospital care populations and to death. The Clinical Practice Research Datalink (CPRD) database was used as this provided HF and other chronic disease routinely collected clinical data in the general practice population and also linkage to the hospital admissions and mortality.

The **specific objectives** in three phases were:

Phase 1: To systematically review evidence on current prognostic studies in the general HF population and identify measures of comorbidities that have been shown to influence HF outcomes.

Phase 2: To develop measures of comorbidity severity and change using routinely collected clinical and healthcare data (from CPRD) that predict risk of mortality in the HF general practice population and investigate their contribution to prognostic models.

Phase 3: To develop measures of heart failure comorbidity severity and change using routinely collected clinical and healthcare data (from CPRD) that predict risk of unplanned hospital admissions in the HF general practice population and investigate their contribution to prognostic models.

3.4 Specific hypotheses

This thesis was focused on three main aims. Firstly, understanding which comorbidity measures are associated with different outcomes in the general HF population; quality of life, hospital admission and death. Secondly, understanding how the *recent* severity of comorbid disease and recent change in severity of comorbid disease influence prognosis for two different outcomes in the general practice HF population; hospital admission and death. Thirdly on understanding how chronic disease comorbidities should be included in HF prognostic models for the general practice population.

Three main hypotheses were formulated:

I. Chronic disease comorbidities are significantly associated with hospital admissions and death in the general practice HF population.

II. Chronic disease comorbidity severity and recent change in comorbidity severity are significantly associated with hospital admissions and death in the general practice HF population.

III. Chronic disease comorbidity severity and recent severity change factors will improve the fit of HF prognostic models for the outcomes of hospital admissions and death.

3.6 Thesis overview and summary

The chapters are based on (i) an overview of concepts and identifying key questions (**chapters 1 to 4**); (ii) systematic review (**chapter 5**) and analyses based on the Clinical Practice Research Datalink (CPRD) data (**chapters 6 to 12**), and the final discussions (**chapter 13 to 14**).

Chapter 1 Introduction and key concepts; the principles of epidemiology, disease and health, comorbidity and prognosis that underpin the thesis are introduced.

Chapter 2 Context setting; this chapter applies the broader principles outlined in the first chapter to the context of HF.

Chapter 3 Conceptual framework and thesis overview; a new conceptual framework for the inclusion of comorbidity in prognosis of chronic disease is presented. The rationale for the thesis is stated and the aims, objectives and key hypotheses that underpin the design and analyses of the three phases of the study are framed.

Chapter 4 Provides discussion and justification of the epidemiological approaches used to perform the investigations with a key focus on nested case-control design used in the CPRD analyses.

Chapter 5 presents the systematic review that synthesises the key available evidence on non-CVD comorbidity and prognosis in a general population of HF patients. The review identifies the key non-CVD

comorbidity prognostic factors and models for potential outcomes; patient reported quality of life, all-cause hospital admissions and all-cause mortality.

Chapter 6 presents the CPRD framework for measuring severity and change using routinely collected data. It is the first of 6 chapters relating to Phase 2 and 3 which investigate which non-CVD comorbidity measures (disease, severity and change) are associated with hospital admissions and death in a general practice population of HF and what the contribution of these measures is to a prognostic model.

Chapter 7 describes the CPRD cohort of HF patients by study factors and the non-CVD comorbidity measures for all-cause mortality outcome. This chapter identifies the potential confounding factors to be considered in the investigation of the prognostic associations with mortality.

Chapter 8 describes the CPRD sub-cohort of HF patients by study factors and the non-CVD comorbidity measures for first all-cause hospital admission outcome. This chapter identifies the potential confounding factors to be considered in the investigation of the prognostic associations with hospital admission.

Chapter 9 uses the nested case-control design to investigate the strength of association between the non-CVD comorbidity measures with mortality. Unadjusted associations are first presented followed by stratification by key potential confounders and then adjusted associations.

Chapter 10 uses the nested case-control design to investigate the strength of association between the non-CVD comorbidity measures with first hospital admission in the CPRD subset of patients that had linked hospital data. Unadjusted associations are first presented followed by stratification by key potential confounders and then adjusted associations.

Chapter 11 tests the interaction between comorbidity measures for the outcomes of hospital admissions and mortality. Statistical and biological interactions are tested for the comorbid disease pairs.

Chapter 12 tests the contribution of the non-CVD comorbidity prognostic measures to pre-specified HF prognostic models for mortality and hospital admission.

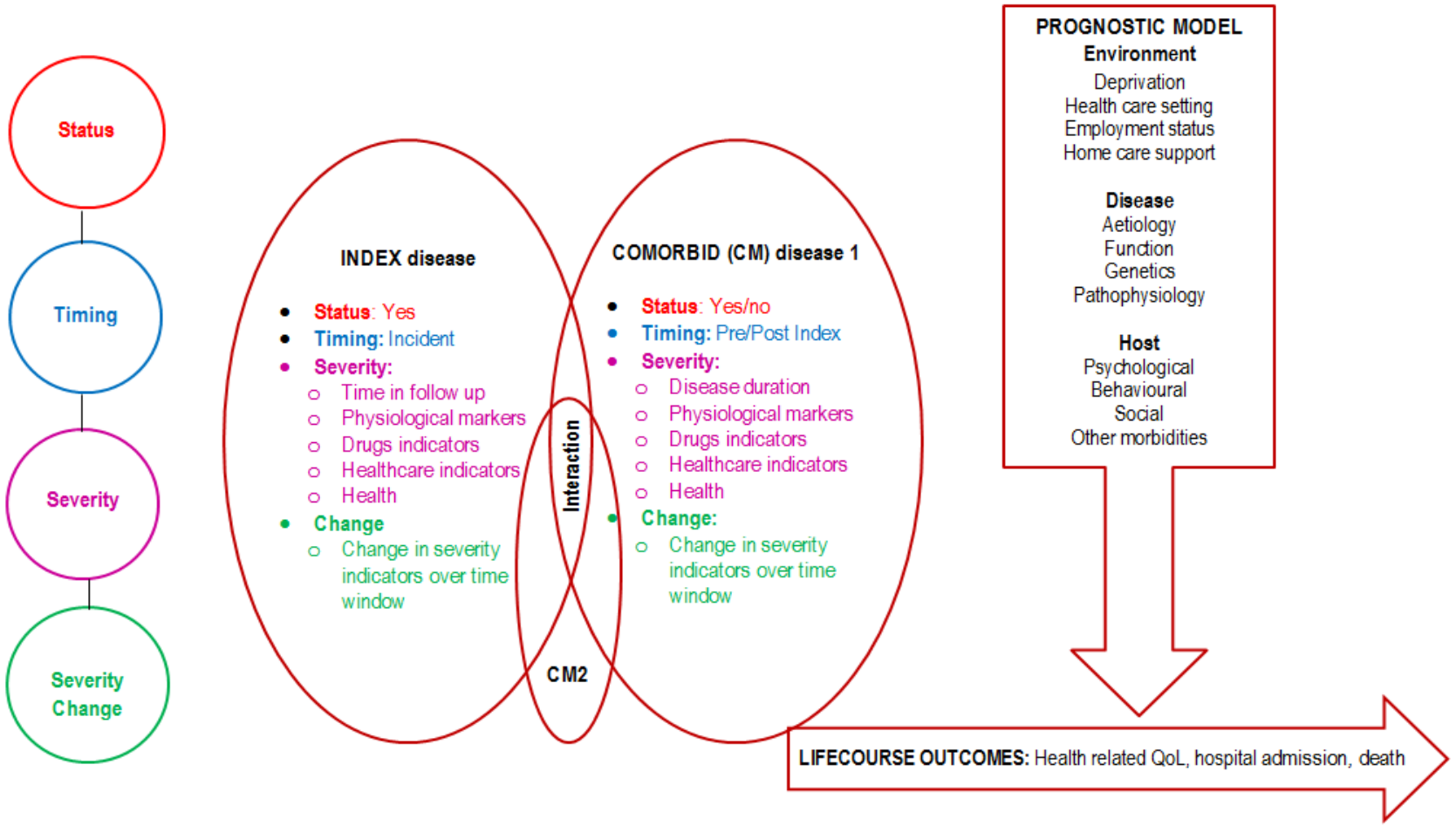
Chapter 13 presents the overall discussion relating to the CPRD studies in Chapters 7-12.

Chapter 14 provides the final summary of the thesis.

This chapter has summarised the key gaps in the evidence on HF comorbidity and prognosis. A new conceptual framework for the inclusion of comorbidity in prognosis has been presented and used to frame the thesis aims, objectives, hypotheses and subsequent chapter plan. The overall components within the framework for the index disease and comorbidity are status, timing, severity, severity change and disease interactions to influence the prognostic model based on outcomes over the disease life course.

Figures

Figure 3.1: A comorbidity prognostic model



Chapter 4 Epidemiological and Statistical Methodology

4.1 Theoretical underpinnings

This chapter covers the general components of epidemiological and statistical methods from key basic concepts to the more advanced methods and analyses used in the subsequent chapters. The purpose is to illustrate the development of the thesis framework, which included a rigorous systematic review as well as using a national clinical database to answer key prognosis questions in the HF population.

Epidemiology is an empirical science based on *inductionism* that relies on a sample of observations from the real world, in order to infer generalisable conclusions. Its converse *deductionism* uses a top down approach to draw definitive conclusions based on logical reasoning(9). Providing the premises from which the conclusion is drawn are true, the conclusion will be true, a philosophy that underpins mathematical reasoning(200). Epidemiology however deals with truths that have to be *induced* from observations in the real world that are often incomplete and fallible(201). Epidemiologists identify patterns that exist within data to generate hypotheses or generalisations about nature. Skeptics of inductionism criticise the assumption that past observations would hold true in the future and that observations are reliant on senses that are subject to error(9). A response to this argument can be found in *refutationism* which holds that induced generalisations can be corroborated or refuted by further observations but that a refuting observation will carry more weight than any supporting observation(202). This leads to the subjection of a hypothesis to rigorous testing and reformulation as well the acceptance that scientific knowledge is no more than a body of currently unrefuted hypotheses.

4.1.1 Causal criteria

In the assertion of causality a number of criteria have been proposed to examine relationships that strengthen inference (E-Appendix A3). Arguably there is no definitive criteria with which to prove causation and the

criteria in a list may be more than required or not sufficient to determine causation. The criteria do provide a useful approach to supporting inference made from observations but the inability to satisfy most criteria does not necessarily rule out causation(203).

Epidemiology uses statistical methods to summarise, examine associations and test hypotheses within a collection of data for a study sample, in order to infer generalisable conclusions(13). The methods used will depend on a number of factors including the measures of association sort and the study designs used to obtain them.

4.2 Risk and Measures of association

There are a number of different measures of association that can be calculated to determine the 'effect' of an exposure on a given outcome. The choice of effect measure will be determined by the data available and the measure of outcome used in the exposed and unexposed groups, for example, a risk or a rate. The different effect measures each compare the risks or rates of an outcome between the exposed and unexposed groups. This comparison may take the form of an absolute measure e.g. the 'difference' in risks or rates or a relative measure e.g. the risk or rate in the exposed group as a 'proportion' of the risk or rate in the unexposed group. It is first necessary then to calculate the risk or rate of outcome in the two groups.

4.2.1 Measuring risk

Risk refers to the proportion of new cases of an outcome, over a time period, in an outcome free population (also called cumulative incidence or incidence proportion)(13).

$$\frac{\text{The number of new cases of an outcome in a time period}}{\text{The number of people followed for the time period}}$$

However when calculated with the baseline 'at risk' population as the denominator this can artificially reduce the risk estimate of a specific outcome if participants are either lost to follow up or experience another competing event that prevents the outcome occurring. These lost participants cannot appear in the numerator

but still appear in the denominator thus reducing the risk of the outcome(9). If they had remained in follow-up (i.e. had not experienced competing event X or been lost to follow-up) then their risk of outcome Y would be greater.

Incidence rate: One approach to accounting for competing events or loss to follow-up in the analysis is to replace the number of individuals in the denominator with the person-years of time in follow-up. This gives an incidence rate(204):

$$\frac{\text{The number of new cases of an outcome in a time period}}{\text{Total time experienced for the subjects followed [person years]}}$$

This way each person is censored at the point of loss. The risk can then be approximated as(9):

$$\text{Risk} \approx \text{rate x time}$$

This approximation only holds true where the risk is not large (e.g. the risk is <20%)(9). The risk is calculated on the number of people in follow-up over the total follow-up time. Rates are an instantaneous measure calculated on units of time which mean that the population at risk changes at different time points. There will be decay of the number at risk, as participants experience the outcome, for each subsequent time point. Using a risk (approximated from a rate) to calculate the expected number of outcomes for a population, where the follow-up is long or the risk high, would lead to an inflated predicted number.

This method of dealing with loss to follow-up and competing events by using time in follow-up deals with the following question 'given that competing event X has not occurred, what is the risk of outcome Y'? In relation to competing risks this method is sometimes referred to as cause-specific competing risks(205).

Competing risks and prognosis: Removing participants who have a competing event in the scenario above is useful when trying to determine true causality as all remaining participants are at risk of the outcome

investigated. However, it is useful and often of interest to patients and clinicians, to know what the effect of an exposure (e.g. a risk factor or an intervention) is on an outcome in the real world setting where other events can and do occur. Another way of dealing with competing risks (sub-distribution competing risks(206)) deals with an alternative question; 'given that event X may have occurred what is the risk of outcome Y'? For this approach, participants lost to follow up are censored at the point of loss but participants experiencing a competing event X that makes Y impossible (e.g. a death by another cause) are only censored at the study end. Leaving individuals in the follow-up in this way gives a true risk of Y in the context that other things (X) may happen. As the denominator is now larger the overall risk of Y is reduced.

Risk and Survival analysis: When observing risk over a time period where the incidence rate changes, then survival analysis can be used to analyse the time to event. Survival analysis splits the follow up period into time intervals over which the incidence rate is assumed to be constant(207). Survival proportions are calculated at each point and the numbers of events are accumulated across the time-periods. These survival proportions are calculated by subtracting the incidence proportion (number of new cases of the outcome / *number at risk of the outcome*) from one. This method can include the *number at risk of the outcome* as (i) everyone at baseline without the event (no loss to follow up or competing event considered) or (ii) removing the loss to follow up *and* other competing events (cause-specific competing risks) or (iii) censoring loss to follow-up but leaving in the participants with competing events until study end (sub distribution competing risks). In this way survival analysis provides a method of calculating risk directly whilst dealing with loss to follow up rather than calculating incidence rates. Whilst incidence rates use time in follow-up as the denominator, survival analysis removes half of the loss to follow-up (over the time interval) from the at risk denominator to account for an even spread of loss over the time period(9). This is the method applied in Kaplan-Meier or Cox proportional hazard regression modelling(64) (Discussed below in Section 4.4).

4.2.2 Measuring causal effects

As well as investigating outcome occurrence, epidemiology attempts to measure the association between an exposure and an outcome in order to infer causation at a population level. The ideal approach to compare outcome risk between groups with and without the exposure under investigation would be to compare people

with clones of themselves at the same point in time, one exposed and one not. That way all other risk factors, except for the exposure investigated, are balanced.

Other epidemiological approaches try to achieve groups that, apart from their exposure status, are as comparable as possible in their susceptibility to the outcome e.g. randomised controlled trials attempt to achieve comparability by random allocation of participants to an intervention in large enough samples to ensure all known and unknown characteristics are evenly distributed(208). Observational methods such as cohort or case-control design select participants using methods that ensure that the exposure distribution in the sample reflects the source population and is not distorted by biased procedures. The balance of other factors between the exposed and unexposed groups that may be associated with the outcome, can be achieved through the design (i.e. matching or purposive sampling) or analysis (i.e. stratification or adjustment)(64).

Once two groups can be assumed to have the same risk of an outcome outside of their exposure status, comparison of the risk or rates of outcome between the exposed and unexposed groups can be made. This can be done in a number of ways as follows:

Difference in means: Where an outcome is continuous and the exposure is dichotomous then a difference in means is calculated between the exposed and unexposed groups(13). This is an absolute measure and so provides an indication of the magnitude of excess risk relating to the exposure.

Risk difference (RD) (Attributable risk) and incidence rate difference (attributable incidence rate): These are other absolute measures of effect used when the exposure and outcome are dichotomous. The risk or rate in the unexposed group is subtracted from the risk or rate in the exposed group. The difference is the risk or rate that is associated with the exposure(78). Alternatively a *population attributable risk* can be calculated, which is the difference between the overall risk of an outcome in a population (incidence proportion) and the risk in the unexposed group. This measure is a function of the exposure risk *and* the amount of exposure in a population. For example, if the risk associated with smoking was the same in two countries but one country

had a much higher number of smokers, the smoking attributable risk would be the same but the population attributable risk would be greater in the country with more smokers.

Risk ratio (RR) or Incidence Rate ratio (IRR): When both the exposure and the outcome are dichotomous then a ratio measure can also be calculated which is the risk in the exposed group (R_1) as a proportion of the risk in the unexposed group (R_0)(209). Similarly an incidence rate ratio is the incidence rate in the exposed group (IR_1) as a proportion of the incidence rate in the unexposed group (IR_0). Rather than an absolute measure of risk difference, this measure provides information on how much more likely the exposed group is to experience the outcome than the unexposed group.

$$\text{Risk ratio} = \frac{R1}{Ro} \qquad \text{Incidence rate ratio} = \frac{IR1}{IRo}$$

The measure does not however reflect the prevalence of the outcome in the two groups and two exposures with the same relative risk may have very different risk differences(9). For example, if in group A, there was a 20% risk of an outcome in the exposed group compared to a 10% risk in the unexposed this would lead to a risk ratio of 2.0 but a risk difference of 10%. Conversely if in group B, there was a 2% risk of outcome in the exposed group compared to a 1% risk in the unexposed group, the risk ratio would still be 2.0 but the risk difference would be much lower at 1%. In the same way that risks can be approximated from rates where risks are small, the risk ratio will be close to the rate ratio under the same circumstances.

Relative effect: Unlike the risk difference measure of effect, the risk and rate ratio do not indicate the magnitude of risk within the exposed group that is due to the exposure (as opposed to other causal factors also experienced by the unexposed group). The risk attributable to the exposure (risk or rate difference) can be expressed as a proportion of the risk or rate in the unexposed group. This measure provides the percentage increase or decrease in risk that is directly related to the exposure (as opposed to the 'exposed group' in risk or rate ratios) compared to the risk or rate in the unexposed group and is obtained by either (using risk ratio as an example):

$$\frac{\text{RD}}{\text{Ro}} \quad \text{or} \quad \text{RR} - 1$$

This means the percentage increase in risk in the exposed group that is due to the exposure can be calculated e.g. a RR of 2.5 would equate to (2.5 – 1 = 1.5) a 150% increase in risk due to the exposure. A RR of 1.4 = 40% increase in risk and a RR of 0.8 = 20% reduction in risk(9).

Attributable Fraction: The attributable risk of the exposure (RD) can also be expressed as a fraction of the total risk in the exposed group using(78):

$$\frac{\text{RD}}{\text{R1}} \quad \text{or}$$

$$\frac{\text{RR} - 1 \text{ (risk proportion attributable to the exposure)}}{\text{RR (risk in all exposed out of all unexposed)}}$$

Odds ratio: An alternative to the risk ratio is the odds ratio. This is the ratio of the odds of the event in the exposed to the unexposed group(78). The odds of the event are calculated by dividing the number of individuals with the event by the number of individuals without the event within a group. This is different to a risk calculation where the whole group is included in the denominator. However where the number of events is low (<20%) the odds will be similar to the risk. This is because the small number of events means that the size of the denominators for the odds and risk calculations would be similar. The odds ratio will be greater than the risk ratio where the numbers of events is high (9). The odds ratio can be used to calculate the ratio of the odds of exposure in different groups, an approach used in case control studies, discussed next.

4.3 Epidemiological designs

Observational approaches such as cohort, case-control, cross-sectional studies and case cross over studies deal with the observation and measurement of exposure and outcomes as they naturally occur(208). The method chosen is often a combination of pragmatism and possibility given the research questions asked and

the data available. In order to investigate the association between chronic disease comorbidity and outcomes in HF, a number of approaches were considered before selecting a cohort with a nested case-control approach, the rationale is provided in the following sections.

4.3.1 Cohort design

A cohort study is the key design for following of a group of individuals over a period of time. The group is subdivided into sub-groups which are usually classified by their exposure status at study entry and should be otherwise similar in their risk of developing the event. Epidemiological experiments are one type of cohort study where the exposure is assigned(210). In other observational cohort studies the exposure is observed as it naturally occurs and measured. The total cohort, regardless of their exposure status, makes up the *population at risk*. This population should each be at risk of the outcome at the start of follow-up.

Calculating risk in a cohort

The numerator: This is counted as the number of events in the group of interest. For outcomes such as hospital admission, the event can occur more than once to each patient in follow-up and a number of options exist for analysis:

(i) the first event only is counted in the numerator which is preferable when it is difficult to distinguish between a new event and an exacerbation of a previous event or when subsequent events may have different precipitating factors(209). Where first events only are counted, participants are removed from the at risk population after experiencing the event,

(ii) each event can be counted as a separate event. This way the estimate of an exposure effect would apply to any of the events or

(iii) to investigate each event separately e.g. first event, second event, third event.

The denominator: In most cohorts incidence rates are used to calculate a rate ratio (for the exposed to unexposed). To calculate the incidence rates for the exposed and unexposed groups, the 'at risk' person time in follow up is used for the denominator as this allows adjustment for loss to follow-up and competing risks. The incidence rate ratio can be used as an estimate of the risk ratio where the risk is small or the follow-up

short(9). Risk can also be calculated directly using the number of people at risk in the exposed and unexposed groups as the denominator where there is no loss to follow-up or by using survival analysis and segments of follow-up time.

Classifying exposure: The assignment of participants to the exposed or unexposed 'at risk' person time in follow-up should be done at study baseline but there are two important considerations; induction time and exposure status change. Following exposure, 'sufficient causation' for an event to occur requires a number of complementary causal factors to interact for example environmental and biological factors(9). Only then will the person be at risk of the outcome. Based on this phenomenon, some analyses will only contribute the 'at risk' person time in the exposed group that commences after the induction time. Similar to cause-specific competing risks, this approach seeks to determine the true causation as a measure of effect between the exposure and the outcome. Similarly an adaption of this approach is to add the induction time to the 'at risk' person time in the unexposed group as any event occurring during this time will be unrelated to the exposure. An alternative approach is to ignore the induction time which is more aligned to prognosis where the time to event from a given point, in the real life context, is of interest. Each approach will give a different rate ratio and will be determined by the questions framed.

The presence or absence of exposure status may change during the follow-up period and the 'at risk' person time may then switch and contribute to a different group i.e. from unexposed person time to exposed person time or vice versa. In historical cohort studies where all exposure status is known for the follow-up in advance of the start of the study, it is important to classify the exposure status appropriately. If an individual that became exposed during follow-up was classified as exposed for the total duration of the follow up this leads to 'immortal time'(211). Immortal time the participant is contributing to the time 'at risk' in the exposed group where they could not have experienced the event (they were not exposed during this time and had they experienced the event they could not have become exposed). This results in an underestimation of effect in the exposed group and an overestimation of effect in the unexposed group(212).

In addition to the presence or absence of exposure that can change during follow-up, for some exposures such as chronic diseases, the quantity or 'severity' of the exposure may change. The new exposure 'severity' or 'change' that has occurred may alter the association between the exposure and the outcome. As stated earlier in Chapter 3 one approach to measure exposure that changes overtime is a cumulative approach that sums up the exposure or exposure time over the follow up period but this approach is limited by the frequency of routinely recorded measures of the exposure in clinical databases. Also this approach does not easily capture the *change* in exposure as opposed to its quantity or the timing of the exposure severity and change that might be of epidemiological importance such as in the current time-period leading up to an event. Other more advanced statistical methods have been developed to accommodate exposures than change over time such as Cox proportional hazard regression modelling(64) but this becomes more complicated where the change is frequent or continuous or the number of events is high (see Section 4.4).

4.3.2 Case-control design

Case-control studies compare the exposure status of all cases within a defined population sample over a time period to a *sample* of controls in the population at risk. Case-control samples are nested within a population or *cohort* and the control sample should reflect as closely as possible the exposure distribution in the source population from where it was drawn(213).

Cumulative case-control sampling: In a cumulative case-control design, controls are selected at the end of follow-up from the participants that did not experience the event. Exposure status is then measured in the cases and controls at the beginning of follow-up(214). Case-control studies estimate odds ratios (OR) as a measure of risk. In Section 4.2.2 it was reported that an OR will not approximate the risk ratio when the risk of an event is high. Furthermore bias might be introduced in case-control studies where controls are selected on the basis of their non-case status representing the number of people in follow-up who did not experience the event of interest. Unlike the denominator population using a cohort approach, this source population does not include the exposure experience of the case group. Providing that the event is rare, then this missing information will be minimal and the odds ratio will be a good approximation of the risk ratio. However where the exposure is related to the outcome and the outcome is common, the denominator that excludes this

information can lead to overestimation of the risk ratio(215). This issue was an important consideration for the investigation of comorbidity exposure in HF due to the high risk of outcomes(209) and a worked example is given below.

If the exposure distribution of a cohort is considered, some of the baseline cohort will go on to experience the event and some won't. If half of the baseline were exposed, one may hypothesise that the proportion of exposure, once the events have occurred, would be greater in the case group. An odds ratio of exposure is calculated using $\frac{a/b}{c/d}$ where:

a	number of exposed people in the case group
b	number of unexposed people in the case group
c	number of exposed people in the control sample
d	number of unexposed people in the control sample

This can also be written as a cross product fraction $\frac{ad}{bc}$

The control sample $c\,and\,d$ should reflect the same proportion of exposure as in the source population. In cumulative sampling where the control group is analysed after the events have occurred, the exposure distribution in the control group will differ to the source population as the exposure experience of the cases has been removed. This means that proportionally, in a given control sample, d (number of unexposed people) will be more and c (number of exposed people) will be less than the source population. This will overestimate the risk ratio.

Case-Cohort sampling: In case-cohort sampling the controls are sampled from the source population at baseline (regardless of their future event status). Given that the exposure distribution in the control group using this approach does include those that go on to become cases, the *'rare event assumption'* required in

cumulative sampling is not required in case-cohort sampling(216). Here, the odds ratio provides a good estimation of the risk ratio(217). Controls can be used more than once and may later also become cases, which closely reflects survival analysis where nearly all subjects are used more than once, or cohort analysis where every subject contributes to the denominator time until they experience the event. The control sample in a case cohort design represents the exposure proportion of the source population.

Both cumulative and case-cohort sampling approaches to case-control studies mirror a closed cohort study where the denominator is sampled from the people at the start of follow-up and both measure baseline exposure status. This creates a challenge for studies where there is loss to follow-up in the study or the exposure status changes over the duration of the follow-up.

Risk set sampling: Where the time in study may vary across individuals or where exposure status may change during follow-up, risk set sampling (also called *density sampling*) provides a mechanism where the *rate ratio* can be estimated using a case-control design(218). The aim of risk set sampling is to produce a control sample that mirrors the distribution of *exposure time* in the source population in terms of 'at risk' exposed and 'at risk' unexposed person time. With this method, the odds ratio will be a good estimation of the incidence rate ratio as opposed to the risk ratio. Consider the following(9):

a	number of cases in the exposed group
b	number of cases in the unexposed group
PT_1	'at risk' person time in the exposed group
PT_0	'at risk' person time in the unexposed group
c	Number of people exposed in the control sample (should reflect the same proportion of exposure time as in the source population)
d	Number of people unexposed in the control sample (should reflect the same proportion of unexposed time as in the source population)
I_1	Incidence rate in the exposed group

I_0 Incidence rate in the unexposed group

Incidence rate ratio is calculated by

$$\frac{IR1}{IRo}$$

which can be equally written as

$$\frac{a\ /PT1}{b\ /PT0}$$

or its cross product fraction $\frac{a}{b} \times \frac{PT0}{PT1}$

If density sampling has been performed properly then $\frac{PT0}{PT1}$ should be equivalent to $\frac{d}{c}$

And so $\frac{IR1}{IR0} \approx \frac{ad}{bc}$

The approximation will be subject to sampling error and reduced statistical precision but this can be reduced by increasing the ratio of controls to cases. Key to this approach is the sampling of controls that represent the *distribution of exposure time* in the source population.

Because the exposure distribution in the cohort will change overtime as cases are accrued, the approach is to match controls at the same point of follow-up time that the case was identified. Using this approach controls will be selected randomly from those subjects who have not experienced the event on the match date. Similar to the case-cohort approach this means that controls can be used more than once and later they may become a case. Where risk set sampling differs is that the inclusion of a control and its exposure contribution now depends on their length of follow-up and exposure status on the match date(214). A control with longer follow-up time has more chance of being selected as a control and having their exposure included. This way the

proportion of exposure in the control sample will reflect the length of follow-up and exposure time of the control sample. A control's future potential as a case also means that the control sample mirrors the exposure time in the source population at the time the case event occurs producing an unbiased estimate of a rate ratio(219).

4.4 Epidemiological and statistical methods in CPRD

4.4.1 Epidemiological method

The main thesis questions were answered using a case-control design that was nested within a defined CPRD cohort. Participants were followed up from baseline which was their first HF consultation code applied in their general practice clinical record over a 10 year time-period. Due to the varying time in follow-up of the sample and the expected change in the exposure distribution in the source population during follow-up as the cases were accrued, 'risk set' sampling of controls that were matched to cases on follow-up time, was chosen (Figure 6.10 for diagram). This approach means that the exposure distribution in the control group reflected the exposure time in the source population at the same point in follow-up as the case occurred giving an unbiased estimate of the rate ratio. Due to the long follow-up (of up to 12 years) controls were also matched on calendar time to reduce the influence of clinical management changes during the length of the study (detailed further in Chapter 6).

The focus of the thesis investigation was on comorbidity severity and the change that occurs prior to an outcome event. The risk set sampling approach provided a mechanism where comorbidity change could be measured longitudinally over a specific time-windows prior to an event and compared to a control group that remained 'at risk', at the same point of follow-up as the cases. The comorbidity severity and change occurring during the course of HF disease in the source population could be captured by measurement in the control group at the same point in time as each case occurred. The potential for multiple selections of the same control does not necessarily duplicate exposure or confounder time as these factors may also change during the follow-up time(9).

This approach is analogous to Cox proportional hazards regression which accommodates time-dependent exposures and yields similar estimates of effect but less efficiently(220). Cox regression is a form of survival analysis (see Section 4.2.1) which uses the time of each event to calculate survival proportions and covariates are compared at each of these time points (221). This means that covariates that change over time can be easily accommodated. Using the counting approach to Cox regression analysis, each change in exposure status is included as a new observation for the length of time it remains unchanged. This results in multiple observations for each subject corresponding to the different exposure changes and all are included in the analysis. The resulting effect estimate adjusts for the level and duration of exposure(222). The challenge with this approach for comorbidity measured using physiological biomarkers or drugs is that these changes are very frequent. Time dependent markers such as biomarkers that change continuously cannot be easily accommodated using this approach. Instead, their time dependency needs to be based on their current value measured prior to an event(221).

An alternative approach referred to as the 'programming statements approach' to Cox regression analysis measures the exposure at the point of each event thus creating multiple exposure statuses corresponding to different time points for the same subject. All of the different exposure measures (corresponding to each event time point) are included in the model as separate covariates which becomes challenging when there are a lot of events(221).

In this thesis the risk set sampling with nested case control approach was selected for three key reasons;

(i) the thesis studies included a large sample with over 26,0000 events and investigated measures of comorbidity severity that normally change frequently over time
(i) the approach selected closely mirrors Cox regression that includes time-dependent exposures, but is more efficient where there are a large number of events and accommodates exposures that experience frequent change
(ii) the approach allows the measurement of exposure severity and change prior to an event to test the study hypotheses

(iii) the approach produces unbiased estimates of rate ratios by comparing exposures for the case and control at the time point in follow-up of each event.

Unless the distribution of the comorbidity exposure and its change over time is known, using current information prior to an event in this approach means that it cannot be used directly for prognosis. The strength of association for the different comorbidity exposures does however yield important information and generate hypotheses about the potential importance of comorbidity severity and change for prognosis which can be further tested in a prognostic model using different approaches(see Chapter 14).

The two routinely available outcomes investigated in CPRD were all-cause mortality and first all-cause hospital admission. For all-cause mortality there were no competing events as any events other than death would not prevent death occurring. However, for the hospital admissions outcome, death itself becomes a competing event. Where this competing event is at least as prevalent as the main outcome it can produce biased estimates of effect(223,224). In this study, for first hospital admission, the competing risk of death was less than the risk of the main event and so was not considered further.

Matching on follow-up time also facilitated the comparison of cases and controls that would be of similar HF severity defined by disease duration. The severity of the HF disease is associated with the outcome of mortality and hospital admission and also with the relative effect of the comorbidity exposure. Matching on follow-up time provided the mechanism for comparing the exposure distribution of cases and controls that were of similar HF disease severity.

4.4.2 Statistical methods

Regression models combine a number of variables into an equation that i) describes the best fitting line for the data given the observations and ii) looks at each variable independently(225). In simple linear regression a straight line is fitted that best describes the relationship between one or more independent variables and a continuous dependent variable. Consider:

Y	dependent variable
a	intercept
β	coefficient
X	independent variable
k	1,2,3,4…….
e	error

Then a linear equation would be:

$$\mathrm{Y} = \mathrm{a} + \beta 1\mathrm{X}1 + \cdots + \beta \mathrm{k}\mathrm{X}\mathrm{k} + e$$

The intercept describes the outcome variable where all independent variables are set to zero (the baseline risk when all exposure variables are not present) and the β is the coefficient which estimates the absolute risk difference between the exposed and unexposed group for each unit of the variable X(9).

4.4.2.1 Logistic regression

Where the outcome is dichotomous such as the event of death or hospital admission it is expressed as a risk or a probability of the event occurring and so the range of predicted values of Y needs to fall between 0 and 1. As probabilities cannot lie outside of 0 and 1 and the risk of an event usually changes very little at the high and low extremes of any singular or combined independent variables, the relationship is usually nonlinear and instead takes a 'S' or sigmoid shape (Figure 4.2)(226).

A linear equation to predict the risk of a dichotomous outcome would produce estimated values of Y that are outside of the range of 0 and 1. In order to derive a linear equation for a logistic curve the Y variable (probability of the event) must be transformed into a value that can take any value from – infinity (∞) to + ∞. This requires a two stage transformation(226). Firstly the risk (R) must be converted to an odds by dividing it by its own complement R/(1-R). This transforms the range to positive values of 0 to ∞ (when converted back using odds/(1+odds) the value is then below 1). Secondly the logarithm of the odds needs to be taken which transforms the positive odds to a value between - ∞ to + ∞ (when converted back by exponentiation this will

places the predicted value on the positive scale again). The two steps together = ln [R/(1-R)] produces the logit. Logistic regression can now produce an equation to predict the logit (Y):

$$\text{logit (Y)} = \ln\left[\frac{\text{R}}{1-\text{R}}\right] = \text{a} + \beta 1\text{X1} + \cdots + \beta\text{kXk}$$

or

$$\text{Odds (Y)} \left(\frac{\text{R}}{1-\text{R}}\right) = \exp\left(\text{a} + \beta 1\text{X1} + \cdots + \beta\text{kXk}\right)$$

β in a logistic regression model is the estimated increase in the logged odds of the outcome per unit increase in the value of the independent variable X. When the β is exponentiated it is no longer an absolute measure of risk but a relative measure which is an odds ratio (OR).

The risk (Y) can be back transformed by

$$\text{Y} = \frac{e^{\text{a}+\beta\text{x}+..+\beta\text{k}}}{1+e^{\text{a}+\beta\text{x}+..+\beta\text{k}}} = \frac{1}{1+e^{-(\text{a}+\beta\text{x}+..+\beta\text{k})}}$$

4.4.2.2 Conditional logistic regression

Where cases are matched to controls on any factor (such as follow-up time), they form a matched group. Conditional logistic regression is very similar to logistic regression but the analysis is performed at the individual group level. The positive event within each group is selected (in this study each group had one case matched to multiple controls). The likelihood of the data is therefore dependent on conditional probabilities. This is the probability of the observed pattern of the case and controls within a group conditional on their being one case(227). Only discordant pairs are included in the analysis. If any variables within the group are constant (same for the case and the controls) they are dropped from the analysis and remain not estimated. One constant term in each group is the intercept and so this is not estimated in a conditional logistic regression. Therefore the equivalent conditional logistic equation to standard logistic regression would be

$$\text{logit (Y)} = \ln\left[\frac{\text{R}}{1-\text{R}}\right] \text{is proportional to } \beta 1\text{X1} + \cdots + \beta\text{kXk}$$

or $$\text{Odds(Y) are proportional to exp } (\beta 1\text{X1} + \cdots + \beta \text{kXk})$$

Whereas the probability from a standard logistic equation uses

$$\text{P(Y)} = \frac{e^{a+\beta x+..+\beta k}}{1+e^{a+\beta x+..+\beta k}}$$

a conditional model where a group has one case and one control asks:

$$\text{P}(1,0 \mid \text{one positive outcome})$$

or

$$\frac{\text{P}(1,0)}{\text{P}(1,0)+\text{P}(0,1)}$$

where P is the probability, 1 indicates the case and 0 the control. The probabilities in the above equation represent the possible combinations of the case and control within a group that has two subjects and only one case. Transformation back to a probability then uses

$$\text{P}(1,0 \mid \text{one positive outcome}) = \frac{\text{e}^{\beta \text{x}+..+\beta \text{k}}(1,0)}{\text{e}^{\beta \text{x}+..+\beta \text{k}}(1,0)+\text{e}^{\beta \text{x}+..+\beta \text{k}}(0,1)}$$

This equation seeks the probability of the observed case (numerator) given the probability of the different scenarios (observation 1 or observation 2 being the case)(227).

4.4.2.3 Model assumptions

Logistic regression does not carry many of the underlying assumptions of linear regression models(228). However there are a few assumptions listed below which will be further considered in Chapter 6.

- No important variables are omitted (model selection, Section 6.3.4, 6.3.6, 6.4.2, 6.4.3).
- No extraneous variables are included (model selection, Section 6.3.4, 6.3.6).
- The independent variables are measured without error (measurement bias, Section 6.3.4, 6.3.6)
- The logit of the dependent variable has a linear relationship with the parameters of the continuous independent variables (linearity, Section 6.4.2, 9.2, 10.2) .
- The independent variables are not linear combinations of each other (collinearity Section 6.4.3, 9.3, 10.3).
- Variables are independent (additivity assumption, Section 6.4.2, Chapter 11).

4.5 Chance, bias and confounding

4.5.1 Chance

As well as establishing the presence and size of an association, statistical testing also allows the investigation of chance through the measurement of variability in the data. Confidence intervals are used to indicate the precision of point estimates of unknown population values or parameters from a sample from that population(208). The degree of confidence is defined as the probability, which depends on the estimate's (assumed) probability distribution, that the interval contains the true value of the parameter of interest(229). Often we wish to estimate the difference in means of two independent populations in which case the interval would then be approximately the difference in sample means ± two standard errors of the difference, assuming normality and 95% confidence. This level is purely based on statistical variability and assumes that all systematic error including bias and confounding are completely removed or controlled.

The other common statistical test is the *P* value. A P value or significance level is the probability of observing a more extreme value than an observed estimate of a parameter when the null hypothesis is true. A P value below .05 is usually chosen to indicate a departure from the null hypothesis.. This value should be interpreted with caution. The *P* value is a function of both the effect estimate size and the study sample size. This means that a study with a big effect size but low precision (wide confidence interval) may be found statistically non-significant(230). A P value of 0.06 indicates that there is a 6% probability of a value at least as extreme as the

observed value if the null hypothesis (RR=1) was true and this, in most studies, would be too high a probability in order to reject the null hypothesis. However this value also indicates that the variable could be worthy of further investigation since it approaches significance. A larger sample size may lead to a statistically significant effect size. This is in contrast to a small and clinically insignificant effect that is found to be statistically significant through larger sample size). It would be counterintuitive to give more weight or importance to the second scenario in considerations of clinical importance(9).

4.5.2 Bias

Bias relates to the systematic error that is introduced in the collecting or analysis of data(64) and is an important consideration for the accuracy of findings. Two main biases related to observational studies are the selection of the study sample and measurement. Selection bias occurs when the participants entering into a study differ in some way to the general population of potential participants with the same inclusion criteria(209). This can be through the self-selection of volunteers who may be healthier than non-volunteers or conversely have more concerns about their health than non-volunteers. Similarly investigator-led selection can also introduce similar bias, for example in a case control study where control participants are not chosen independently of the exposure.

Information collected about or from participants can result in bias due to misclassification which may be differential (related to other variables such as the exposure or the outcome) or non-differential (unrelated to other factors). One example of differential bias may be where the recall of exposure is greater in cases than non-cases or where the outcome is more likely to be observed or diagnosed in the exposed group due to better monitoring(204). In situations where misclassification is non-differential and the exposure is dichotomous then the exposure effect will be diluted and the resultant bias in the exposure effect estimate will be towards the null value. Where there are two or more exposure categories, non-differential misclassification leads to convergence of the effects across categories(9).

In this study, selection bias was minimised by using a national sample of heart failure patients with random sampling of controls from the risk sets of the cases. Steps taken to reduce misclassification bias will be detailed in Chapter 6.

4.5.3 Confounding

Confounding is caused where there is a factor, other than the exposure of interest, which is associated with the exposure and separately with the outcome. A confounder has a causal effect on the outcome or is a proxy marker for another causal factor and may be unevenly distributed across exposure groups(13). A confounder should not be an effect of the exposure itself (i.e. on the causal pathway between the exposure and effect)(1). Unless a confounding factor is taken into account it may lead to an incorrect inference on the association between the exposure and the outcome. Restriction of participants to one category of the confounder (e.g. specific age) or matching participants on a confounding factor can limit the confounding through selected study design(207).

When matching is done on the basis of the cases in a case-control study it can adversely influence the exposure distribution in the control sample. A confounding variable is by definition associated with the exposure and the outcome. By matching on this variable the control group is forced to mirror the cases on a factor that is also associated with the exposure of interest i.e. increase the exposure proportion in the control group which will artificially reduce the effect of the exposure on the outcome. For this reason, matching in this study was restricted to the time in follow-up and calendar time as opposed to any potential confounding factor(231). Potential confounders were investigated and then adjusted within the regression models (see Chapter 6).

4.6 Clinical data sources

There are a range of clinical data sources which vary in nature, content and detail. The choice of data source for use in epidemiological research will be determined by the questions asked and the ability of the data to represent the source population from where it is drawn. For this reason data that has a clear denominator

population is an important consideration for epidemiological research. Other common considerations across different data sources are accuracy, consistency, completeness, timeliness, precision and access(232). There are four main sources of clinical care data: national records, registers, surveys and electronic databases.

4.6.1 National records, disease registers and surveys

Countries including the UK keep a national archive of all births, marriages and deaths(232) which provide important information on population demographics including cause of death. Whilst accurate, this data lacks the detail and frequency of recording necessary to study disease. Other national data records related to broader determinants of health include scores of deprivation related to small geographical areas such as the Index of Multiple Deprivation (IMD) which was used in this study where available through linkage with the CPRD (see Chapter 6). Disease registers can be an excellent source of information on people with a specific risk profile, disease, condition or intervention which can be linked using patient identifiers to other health data(233). One of the advantages of disease registers is that they are based on a concerted effort to identify all people within a defined denominator population who share a particular health characteristic rather than relying on health care contacts. However limitations of this data source are that registries are usually district or regionally based and there is heterogeneity across registers in terms of quality and purpose(234). Surveys, such as the National Census(235) and the Health Survey for England(236), collect self-reported information from the public and are a means of collecting socio-demographic as well as health information from a targeted denominator population. An advantage of surveys is that they capture a wider population than those who consult for ill health providing important information about the true burden of disease in populations. Limitations with self-reported health surveys include the smaller samples and subjectivity across patients.

4.6.2 Electronic databases and the Clinical Practice Research Datalink (CPRD)

Healthcare data is routinely recorded in clinical practice during patient contacts and can be a rich source of information on disease, interventions and outcomes. As the largest provider of ongoing care in the UK general practice captures most of the reported health problems in the population. The introduction of a GP UK contract in 2004 which established the Quality Outcomes Framework (QoF) has meant that the recording of clinical information and completeness of data has improved for specific conditions such as HF and renal

disease. This QoF data is collected centrally although it is not accessible directly and targets a limited set of conditions(197).

One of the largest longitudinal captures of general practice data which was used in this thesis, is the Clinical Practice Research Datalink (CPRD). This database is the world's largest electronic database of anonymised records for primary care(237). CPRD has, since 1987 collected data from voluntary GP practices across the UK to create anonymised, longitudinal medical records of registered patients' demographic information (e.g. age, sex, weight), clinical events (medical diagnoses), referrals to specialists and secondary care settings, prescriptions issued in primary care, diagnostic testing, lifestyle information (e.g. smoking and alcohol status) and all other types of care administered as part of routine general practice.

The main primary care database held by CPRD is known as GOLD which covers approximately 8.8% of the UK population, including practices in England, Northern Ireland, Scotland and Wales. As of September 2014 there were 684 GP practices and 13.58 million research quality patients in GOLD, of which 5.69 million are active (still alive and registered with the GP practice)(238). The CPRD sample is therefore representative of the UK general population, with comparable age and sex distributions as those reported by the UK National Population Census.

The CPRD has established linkages to a number of other datasets, including Hospital Episode Statistics (HES), Hospital Treatment Insights (HTI), cancer registry data, Index of Multiple Deprivation (IMD) and mortality data from the Office for National Statistics (ONS). Linkage of CPRD GOLD data with other patient level datasets is only available for English practices that have consented to participate in the linkage scheme. These linkages cover approximately 75% of the contributing practices in England, or roughly 58% of all practices in the database.

4.7 Summary

This chapter has presented an overview of key epidemiological and statistical methods. The thesis focus is on the investigation of comorbidity severity and change that occurs prior to an event in HF which poses epidemiological and statistical challenges. Different methods were considered in order to illustrate the development of and rationale for the methods chosen to investigate the thesis hypotheses. The CPRD cohort with a nested case control design, using risk set sampling, will be further detailed in Chapter 6 with specific application to the thesis.

Figures

Figure 4.1 **Cohort with nested case control approach using risk set sampling**

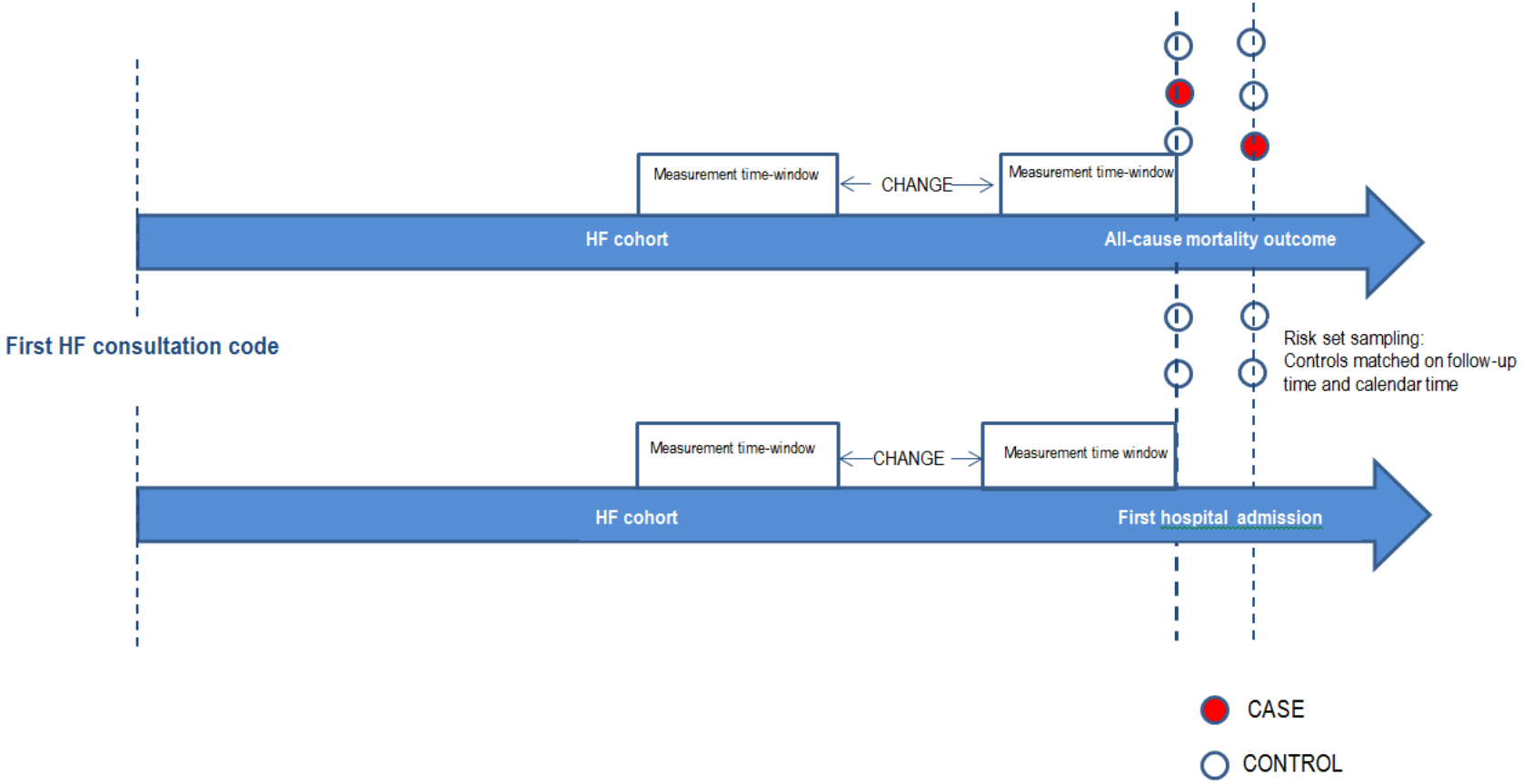
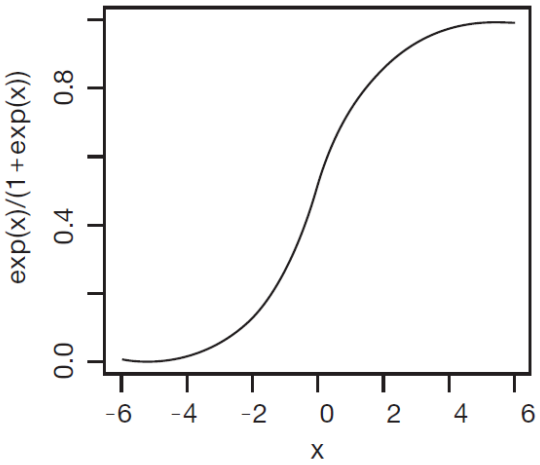


Figure 4.2 Logistic curve



Taken from Park (2013)(228)

Phase 1: Systematic review

Chapter 5 Non-cardiovascular comorbidity and prognosis in HF populations: A systematic review and meta-analysis

The association between different non-cardiovascular (CVD) chronic disease comorbidities and outcomes in HF has been the focus of a number of studies over recent years but these studies have mostly included *selected* HF patients and there is no systematic framework for the inclusion of comorbidities in prognosis studies. This chapter describes a systematic review to examine which non-CVD comorbidities have been included in general HF population prognosis studies and estimate the magnitude of association between non-CVD comorbidities and different outcomes in HF(239). The influence of comorbidity severity and how this has been included in prognostic models for HF is also described. Inclusion of non-CVD comorbidity in prognosis provides the potential for developing prognostic factors that are sensitive to changing severity of an index disease such as HF and to predict key transitions in health and healthcare.

5.1 Introduction

The evidence focus stated in Chapter 2 has been on cardiovascular comorbidities in selected populations despite non-CVD comorbidities now predominating in the older general HF population. The impact of chronic disease comorbidity on patient outcomes depends on a number of factors which include (i) the disease, (ii) the severity status of the disease and (iii) the outcome investigated. Prognostic models that aim to provide individual risk for HF patients currently focus on mortality outcomes(73,174,240) and whilst they have included one or two non-CVD chronic disease comorbidities, no systematic approach has been developed for clinical practice and severity status of comorbid disease is rarely studied.

A systematic review and synthesis of current evidence provides the opportunity to compare the influence of different chronic disease comorbidities for different outcomes which lays the basis for understanding how non-

CVD comorbidity may be used to predict key life course changes with disease progression in health, unplanned hospital admission and mortality outcomes over the disease course.

5.2 Aims and Objectives

Aims: The aims of this systematic review was to examine the current evidence on the influence of clinical non-CVD comorbidities on outcomes in HF and on current prognostic models for HF that incorporate non-CVD comorbidities to predict outcomes for unselected HF patients in community and hospital settings.

Objectives: The three specific objectives were;

(i) To determine which clinical non-CVD comorbidities influence the HF outcomes of patient reported health status measures, hospital admissions or mortality.

(ii) To investigate if and how non-CVD comorbidity severity and change in severity status has been investigated in HF prognosis studies.

(iii) To determine how non-CVD comorbidities have been included in current HF prognostic models, by chronic disease type, severity or status change.

5.2.1 Hypotheses

In addition to identifying the current evidence and measuring the association between non-CVD comorbidity in HF and outcomes, three hypotheses were postulated:

(i) The risk estimates for the association between non-CVD comorbidities and outcomes in HF will differ by disease and chosen outcome.

(i) Increased non-CVD comorbidity severity is associated with worse HF outcomes.

(ii) Increasing non-CVD comorbidity severity change is associated with poorer HF outcomes.

5.3 Methods

In total medical databases of published and unpublished studies were searched and studies were included that focused on the HF population, the prognosis outcomes of self-reported health status, hospital admission or mortality and which included a chronic disease non-CVD comorbidity measure. Selection was performed within the review team and data extraction was undertaken using a predefined template. Quality assessment was performed by two reviewers and inter-rater agreement was tested using Cohen's kappa coefficient. The risk estimates for different chronic diseases were described for different outcomes and meta-analysis was performed where appropriate. Finally the inclusion of comorbidity in prognostic models for the different outcomes was described.

It has been recognised over recent years that prognosis studies vary in terms of methodology, quality and reporting(241). For studies focused on a particular disease there can be much variation in clinical definitions, settings and selection of exposures and outcomes. Systematic reviews of prognosis studies that do not take these factors into consideration can lead to collective bias, misrepresentation of the effects of exposures or conflicting results between multiple reviews with the same objective (observed in reviews of cancer and stroke amongst other diseases)(242). As a number of similar prognosis studies accumulate it is important to identify all the available studies to provide an overall assessment of risk. However this is complex given the variation in the reporting of studies (identification) and variation in methods used (overall assessment of risk)(243). These issues were carefully considered in a protocol for the systematic review which was registered with PROSPERO (protocol no.CRD42013003605; www.crd.york.ac.uk/prospero/) (E-Appendix D1) and was reviewed by a member of the Cochrane Prognosis Methods Group (E-Appendix E1). Initial steps taken to improve the identification and methodological consistency of studies were the use of a validated search strategy for prognosis studies, quality appraisal of each study by two reviewers and careful consideration of the studies used in any meta-analysis. The Meta-analysis Of Observational Studies in Epidemiology (MOOSE) guidelines(244) were used for the reporting of the review.

5.3.1 Inclusion and exclusion criteria

The review inclusion and exclusion criteria for selecting studies (Table 5.1) were developed in relation to key components using the PICOS acronym (population, intervention or exposure, comparators, outcomes and study design)(245). The comparator component was not relevant to this review. Given the variation possible within each of the components, clearly defined criteria were required to allow for the comparison of the comorbidity effect estimates across the studies included and meta-analysis.

Population: The study samples were adults aged 18 years and over with HF. HF in younger groups is more likely to be associated with a rare cause such as a congenital heart defect(246) which would require specific consideration for prognosis.

HF definition: The sample included any person with a clinical diagnosis of HF, but this varied according to the study settings and time of investigation. Current European diagnosis recommendations include the combination of symptoms (e.g. dyspnoea, limited exercise tolerance, fatigue) signs (e.g. peripheral or pulmonary oedema, jugular venous distension) and objective evidence of HF (e.g. impaired ejection fraction or ventricular dysfunction)(93). The use of echocardiogram to provide objective evidence has only been introduced into the Quality Outcomes Framework for general practitioners since 2006 in the UK(247). Inclusion in this review was on the basis that the HF population was not selected on the basis of a specific aetiology or manifestation of HF but instead included a contemporary mix of patients with HF reflective of the general population.

Settings: Samples from three different broad settings were included. Firstly hospital based samples which included both de novo HF and chronic HF admissions (primary cause of admission or discharge). Secondly, community based samples which included general population, general practice and outpatient clinics. Thirdly randomised controlled trial populations which are often hospital based or include both hospital and community patients.

Setting exclusions: Excluded from the review were studies that selected the HF populations by causes e.g. cardiomyopathy or myocardial infarction, by procedure e.g. heart transplant or device or by hospital sub setting e.g. intensive care. Some research studies on HF have broadly selected groups such as patients with a preserved or reduced ejection fraction(248-251) and these were excluded to focus on the general population of HF patients with both preserved and reduced ejection fraction(252) and to aid the comparability and synthesis of effect estimates.

Exposures - comorbidity definitions: The focus was on comorbidity exposure defined as any non-cardiovascular chronic disease. Chronic diseases were based on clinical diagnosis or patient self-report with inclusion of disease severity. Acceptable comorbidity indicators included in the screening process were:

(i) chronic disease type e.g. diabetes

(ii) chronic disease severity measure e.g. fasting glucose or glycosylated haemoglobin (HbA1c) for diabetes(253). For physiological indicators of disease severity, only those used in current diagnostic guidelines were included.

(iii) chronic disease severity change measure e.g. changes to drug prescription or dose, change in blood test indicator (such as HbA1c) or frequency of healthcare episodes.

In an initial scoping of evidence, two systematic reviews on renal function and renal function change and outcomes in HF were identified(145,157). This review included renal studies published after the prior reviews (May 2005) together with relevant studies from the prior reviews. Whilst CKD is defined as kidney damage with or without reduction in estimated glomerular filtration rate (eGFR) or a reduction in eGFR for greater than three months(254), the latter criteria is more commonly used to describe CKD and relates to stage 3 or more out of five stages of CKD. Most studies that include renal comorbidity use eGFR level as a marker of dysfunction rather than a formal clinical diagnosis of chronic kidney disease.

Exposure exclusions: Cardiovascular comorbidities are usually associated with the aetiology or consequences of HF (e.g. ischemic heart disease or hypertension) and whilst important to prognosis they

have been well described in the literature(143,144,255-257), and hence excluded from this review. The focus of the review was on chronic disease comorbidity so conditions such as anaemia or depression were excluded as were novel biomarkers or physiological indicators that are not routinely used or not specific to a chronic disease status. Examples include liver function tests, cystatin-c, urea, blood glucose, blood urea nitrogen (BUN) and oxygen saturations.

Outcomes: The primary outcomes included in the review were health related quality of life (HR-QoL), all-cause hospital admissions or readmissions and all-cause mortality. HR-QoL outcomes included general measures such as the Short-form 12 or 36 or HF specific measures such as the Kansas City Cardiomyopathy Questionnaire (KCCQ). All-cause mortality and hospital admissions were chosen as opposed to specific causes for comparability of studies and to reflect the study population with HF and comorbidity.

Outcome exclusions: Composite outcomes such as 'hospital admission *or* death' were excluded with the exception of 'death *or* urgent heart transplant' as this emergency procedure prevents inevitable death in HF and is often counted as a death event in HF studies(189,258,259).

Study designs: Observational studies with more than 30 days follow up were included. These constituted cohort studies and secondary analyses of randomised controlled trials (RCTs). RCTs were included only where both trial arms were used for the prognosis study whether or not the intervention was accounted for in the analysis. This approach was appropriate as the controlled distribution of the intervention and the often negligible effect of the intervention on the effect of the exposure(225).

Three types of prognosis studies were included:

(i) Prognostic factor study (chronic disease focus). These were any HF studies that focused on the association between a non-cardiovascular chronic disease comorbidity and an outcome. Studies that did not focus on a non-CVD chronic disease e.g. a pharmacological intervention study and only adjusted for a chronic disease comorbidity were excluded.

(ii) Prognostic factor study (general). These were any studies that investigated a range of potential prognostic factors and included a non-CVD chronic disease comorbidity to be an independent prognostic factor.

(iii) Prognostic model study. These were any studies that *combined* a number of prognostic factors to provide an estimate of individual risk in HF and which included a chronic disease comorbidity prognostic factor. The primary study rather than validation or impact studies were included. Where studies included a derivation and validation cohort the derivation cohort was included.

Study design exclusions: Excluded study designs included cross sectional studies and RCTs where inclusion criteria did not reflect the inclusion criteria in the review. Given the scope of the review it was decided to limit papers to those published in English language only. Whilst this may introduce some bias through authors only publishing in English journals when results show a significant effect, investigation of two meta-analyses has demonstrated that the removal of non-English studies has not affected the overall results(260). Studies that were only published in abstract form were excluded if a full report could not be identified.

5.3.2 Search strategy

A search strategy was designed that included database selection, development of search terms, search validation and additional searches.

5.3.2.1 Literature Databases

The main bibliographic databases used to identify the studies of interest were MEDLINE, EMBASE and CINAHL. MEDLINE is produced by the United States National Library of Medline and indexes over 5,200 journals in 37 languages and EMBASE, a European database, indexes over 4,800 journals in 30 languages(261). Both generate a similar number of relevant references and whilst there is some overlap depending on the topic, it has been recommended that both databases be used for a comprehensive search(261,262). This may be of particular importance where meta-analysis is considered as some bias in the

reporting of effects has been found across the two databases with EMBASE including proportionally fewer studies with smaller effects(263). CINAHL (Cumulative Index to Nursing and Allied Health Literature) which indexes journals that are orientated to nursing and allied health was searched in addition to capture any studies relating to HR-QoL that may be missed by the other databases. Each database has its own search interface and related syntax depending on the service provider. For this search the interface was determined by validated prognosis search strings for MEDLINE and EMBASE and were as follows; (i) MEDLINE: PubMed (http://www.ncbi.nlm.nih.gov/pubmed), (ii) EMBASE: OvidSP (http://ovidsp.ovid.com/), and (iii) CINAHL: EBSCO (http://www.ebscohost.com/).

Other databases: To improve the retrieval of any prior systematic reviews in relation to the topic, DARE (Database of Abstract of Reviews of Effects) and CDSR (Cochrane Database of Systematic Reviews) were also searched. CENTRAL (The Cochrane Central Register of Controlled Trials) was searched to indicate any prognosis studies that were performed as a secondary analysis of a trial.

5.3.2.2 Developing search terms

Search strings were developed for three of the four components of the inclusion criteria: population, outcome and study type. A decision not to include the comorbidity exposure in the search strategy was made due to the breadth of possibilities when considering all chronic diseases and their indicators. Also for prognostic factor studies (general) and prognostic model studies, chronic disease terms would be unlikely to be listed in free text or controlled vocabulary where they were not the focus of the study.

HF population search string: A validated search string for HF was identified(264) for use in MEDLINE using OvidSP. The HF search string used a combination of controlled vocabulary words and free text headings (E-Appendix E2) which retrieved 98% of records validated using two different methods. Precision was quite low and 25% of all records retrieved using the strings in MEDLINE were not relevant to the HF population(264). When this population string was used in conjunction with the prognosis and outcomes string it resulted in over 23,000 citations in EMBASE alone (E-Appendix E3). Initial random check of 500 articles demonstrated that only 29% (n=145) of the articles were focused on the HF population. As HF falls at the end of a spectrum of

cardiovascular disease it was often indexed as a controlled vocabulary term in studies where HF was an exposure or an outcome rather than the focus of the study. Following this initial search a HF string focused to the title of articles was then designed.

HF title search: The controlled vocabulary terms and their subheadings from each database were reviewed to develop a list of possible descriptors of the population that might appear in an article title. This list was then reviewed and added to by two HF clinical experts (E-Appendix E4). The search was rerun using the new population string focused to the title and then 1% of the initially excluded articles were checked for appropriate exclusion. None of the excluded studies checked were relevant to the review.

Prognosis string: Validated prognosis search strings were identified for retrieval of prognosis studies in MEDLINE and EMBASE. Prognosis search filters have been found to be less sensitive for prognostic factor studies than prognostic model studies so search strings that optimized the capture of both types of study were selected. An updated version of the Haynes broad filter was chosen for MEDLINE (E-Appendix E5a) as this yielded a sensitivity of 0.84 for prognostic factor studies and 0.9 for prognostic model studies when validated against the included studies in four prior systematic reviews(265). Before the update low specificity was found for both factor and model studies (0.67) and this remained after the update reported as a high 'number need to read' (NNR). The NNR reflects the number of articles that need to be read before a good quality article is likely to be identified. The NNR was 1010 for factor studies and 208 for model studies. The high sensitivity but low specificity means that most good quality studies would be captured by the search but the efficiency of the search would be low leading to a high number of articles requiring screening out.

For EMBASE a number of different combinations of terms had been previously tested for the retrieval of prognosis studies. The chosen string(266) (E-Appendix E5b) was one that again optimized sensitivity (99%) over specificity (51%). This search string was adapted for CINAHL using the appropriate syntax for EBSCO interface (E-Appendix E5c).

Outcomes string: Free text words for each of the three outcomes included in the review were selected using common descriptors and thesaurus terms (E-Appendix E6).

5.3.2.3 Validating the search

All three search strings were combined using the Boolean operator 'AND'. The resulting combined unique citations were cross checked to make sure that they included a selection of known comorbidity prognosis articles. Four of the fifteen known studies were not included in the Medline citations but were included in EMBASE. Free text words that were common to all of the MEDLINE non retrieved articles and which were not part of the prognosis string in MEDLINE were risk, HR, hazard. These terms were then added to the MEDLINE prognosis string as free text words. When the MEDLINE citations were reran then this included all known articles. Also it was noted that the outcomes 'admission' and 'readmission' were often termed hospitalization or rehospitalisation in American studies. These words were added to the outcomes string.

The final search strategy was ran (E-Appendix A4) and two systematic reviews of prognostic factors(182) and prognostic models(183) for readmissions in HF were used to check the validity of the current review. The combined 112 prognostic factor studies and 6 prognostic model studies across the two prior reviews were screened against the current review inclusion criteria. Only one factor study and two model studies met the inclusion criteria for this review and had been identified by the current search. Most of the factor studies in the prior review were focused on a non-chronic disease prognostic factor or an outcome of less than 30-day follow-up.

As the inclusion criteria for the current review was applied during the screening stage and not the search stage, half of the prognostic factor studies (n=56) from the prior review that did not meet the inclusion criteria for the current review were again cross checked against the citations identified by the current search and 51 of the studies had been identified. Of the 5 not identified, 4 did not have a HF term in the title and none were focused on the HF population.

5.3.2.4 Additional searches

Grey literature: To reduce publication bias, online databases were searched for unpublished studies. These included HMIC (Health Management Information Consortium) for unpublished studies, Index to Scientific and Technical proceedings, Zetoc and the Conference Papers Index for conference abstracts and ETHOS, ProQuest and Networked Digital Library of Theses and Dissertations for any related thesis. Studies only in abstract form that had not been reported in full elsewhere were excluded.

*Other searches: A*ll reference lists from included studies were reviewed for any missed relevant studies and key authors in the field were contacted. These included Professor John McMurray (Scotland), Professor Martin Cowie (London) and Professor Greg Fonarow (America). Citation searching was performed on some of the key included articles to increase sensitivity(267) using Web of Science, an online scientific citation indexing service. The Cochrane Prognostic Method's Group website was also searched to identify any prognosis studies or prognosis reviews relevant to the topic. Hand searching of key journals included the European Journal of HF, Journal of Cardiac Failure, HF Reviews and the Journal of Comorbidity. Zetoc was also used to set up an email alert and RSS feed of the journal's contents.

5.3.3 Screening of articles

This stage was performed by a total of three reviewers comprising topic and methodological expertise (CAR, LD, UTK). Initial screening was performed by the first reviewer (CAR) based on title and where necessary abstract, to exclude duplications and studies that were not suitable and these were cross checked (1% of articles) by a second reviewer (UTK). This initial screen focused on two main inclusion criteria (i) a prognostic factor study that focused on a non-CVD chronic disease or (ii) a general prognostic factor or prognostic model study that reported a non-CVD chronic disease factor as independently associated with an outcome. The first reviewer then selected articles based on abstracts that met the inclusion criteria of the review. Two second reviewers (LD, UTK), who were blinded to the decision of the first reviewer, screened half of the abstracts each and any disagreement between the first and second reviewer was arbitrated by the remaining second reviewer. The final set of papers was read in full by the first reviewer (CAR) and second reviewer (UTK) separately and any disagreements were resolved by consensus.

5.3.4 Data extraction

Methods, exposures and outcomes: A data extraction form was designed in relation to the review's aims and objectives. The form needed to facilitate the collection of information relating to the key elements of the review for analysis and interpretation of the findings. The form included criteria in relation to the study source, eligibility, methods, participants, exposure, outcomes, results and analysis (E-Appendix A5). It was decided to collect the information needed to assess risk of bias on a separate form to simplify the process (see next section 5.3.5). A structured approach to data extraction was performed by two reviewers (CAR, UTK) separately and any discrepancies were again resolved through consensus.

Extraction of the results: Where summary variables were reported by sub groups, whole group results were calculated. For the participant variables of age and ejection fraction the sub group means and standard deviations were combined using a formula to calculate the combined mean and pooled standard deviation (E-Appendix A6a) which provides an *approximation* to the true standard deviation of the total group. An alternative, more detailed formula that takes the difference in group means into account could have been used (E-Appendix A6b) however this was not deemed necessary given the small difference in the means and standard deviations of the subgroups within the included studies(268). Where the effect estimates were presented by sub groups and the whole group, the whole group effect estimates were extracted. The only sub group effect estimates that were extracted were those presented by age or gender.

5.3.5 Quality appraisal

Systematic reviews of prognosis studies need to appraise the methodology of the included studies to assess the risk of bias that is associated with the study design, conduct, analysis and interpretation(269). Failure to do this can lead to misinterpretation of the collective evidence. Collectively, included studies also need to be appraised for reporting and publication bias particularly where meta-analysis is considered (see section 5.3.6). Risk of bias assessment was performed in this review for each included study after the initial selection. It was used to include only studies with overall low or moderate risk in the meta-analyses and to perform sensitivity analysis on the combined effects, removing studies that had one or more individual domains at high risk. The overall and specific domain risk assessment was also included in the description of the included studies.

Quality appraisal Assessment tool: The Quality in Prognosis Studies Tool (QUIPS) tool was identified for the review which is a tool constructed by a team of epidemiological, clinical and statistical experts and based on 163 prior reviews of prognosis studies(270). The tool has six domains relating to areas of potential study bias in prognosis studies (E-Appendix A7a). Each domain has a number of prompt questions to facilitate the reviewer making a scientific judgement of high, moderate or low risk of bias for each domain. The reviewer then makes an overall study judgement of risk using the six domains.

The QUIPs assessment attempts to go beyond the determination of whether or not a study was clearly reported or conducted well, an approach used in other quality assessments(270) as these studies may still be open to bias. The Cochrane review risk assessment categorises risk as low, high or unknown which allows for the allocation of a risk category to studies with non-reported information. Using the QUIPs tool, where information was missing or information was presented but the level of risk was unclear, the moderate risk category was used.

A challenge with assessing the overall risk of a study across the six domains is that assessments are subject to personal interpretation of the importance or weight given to the different domains. Attempts to overcome this challenge for this review were made in two ways. Firstly a component approach to assessment that considered the specific QUIPs criteria for each domain was conducted and a focused set of objective criteria were selected for reporting to aid transparency of the overall risk assessment. The QUIPs tool provides 38 different criteria for consideration across the 6 domains which can be time resource intensive to report in any detail (taking up to an hour per paper for experienced trained assessors)(270). Whilst these criteria were used in the assessment of risk, selected specific objective measures were used for reporting. These measures were selected from the QUIPs domain criteria as well as some additional criteria pertaining to internal and external validity used in other prognosis reviews(271). The main supplementary criteria related to statistical analysis where there was less detail given in the QUIPs tool (E-Appendix A7b). As the review included prognostic model studies as well as prognostic factor studies additional criteria for models were also added. Secondly, whilst domain and overall study risk were assessed, to reduce subjectivity, CAR and UTK

performed the quality assessment and inter-rater agreement in the overall study risk assessment was measured using Cohen's Kappa coefficient.

5.3.6 Synthesis of findings

Overview: Meta-analysis was performed on the sub-set of prognostic factor (chronic disease focus) studies with effect estimates that could be combined. Quantitative synthesis was limited to these studies for three main reasons. Firstly these studies were more likely to appropriately consider confounding and provide unadjusted and adjusted comorbidity effect estimates. Secondly because the effect of comorbidity on an outcome was the focus of the studies, the measure of comorbidity exposure status was commonly more detailed which aided comparison of the studies. Lastly, whilst reporting bias can still exist in focused studies, this is more likely for individual exposures in general prognostic factor studies. These studies commonly group a large number of factors together for analysis and only those found to have statistically significant independent effects are usually reported by the study authors(272). For this reason the prognostic factor studies (general) in this review were selected on the basis that a chronic disease exposure was reported as independently significant in a final model. The purpose of this approach was to investigate the scope of studies that had found comorbidities to be important to specific outcomes whilst at the same time acknowledging that a number of studies that had not found a significant and independent effect of chronic disease would have been missed. Combining these studies would therefore not have been appropriate. Where heterogeneity in the observational study designs, population and analyses prevented meta-analysis, a descriptive approach was used for the synthesis.

The studies were summarised using the following hierarchy (1) chronic disease (1.1) prognostic outcome (1.1a) prognostic factor studies (chronic disease focus) including meta-analysis where appropriate (1.1b) prognostic factor studies (general), and (1.1c) prognostic model studies.

5.3.6.1 Meta-analysis

The framework for the synthesis of the prognostic factor (chronic disease focus) studies to assess the strength of association between a given factor and a specified outcome was as follows;

- *What is the direction of the effect?*
- *What is the size of the effect?*
- *Is the effect consistent across studies?*
- *What is the strength of evidence of the effect?*

Meta-analysis is the statistical combining of results from two or more separate studies(273). Performed appropriately meta-analysis can (i) increase the power to detect a real effect if it exists (ii) improve precision and reduce the variation in the overall effect by including more information and (iii) allow for the investigation of the consistency of chronic disease exposure effects across populations and exposure definitions that may differ slightly across studies. However care has to be taken not to combine studies that are so diverse that the meta-analysis becomes meaningless or have serious publication or reporting bias which would provide an inappropriate summary(272). Steps taken to address these considerations will be described through the different steps of analysis. Inclusion criteria for the studies used in meta-analysis are summarised in Table 5.2.

- **What is the direction and size of the effect?**

Meta-analysis uses effect estimates from each individual study to produce a summary pooled effect estimate which is a weighted average of the exposure effects estimated in the individual studies. The following formula was used(274):

$$\text{Weighted average} = \frac{\sum \text{estimate X weight}}{\sum \text{weight}}$$

The effect measure chosen for meta-analysis in the review was the adjusted Hazard Ratio as this was most commonly reported in the included studies. In a fixed-effect meta-analysis the weight given to the effect estimates of individual studies uses the inverse-variance method(275). Using this approach more weight is given to the effect estimates with less variance (or smaller standard errors). This way, larger studies that tend to have small standard errors are given more weight and therefore contribute more information to the pooled estimate of effect(272). Given that the chosen effect measure was a ratio measure the natural logarithm of the estimate was included in the analysis which makes the confidence intervals appear symmetrical around the study mean(268) but then displayed on the original scale.

Variation: Due to variation in the exposure effects across individual studies, random-effects meta-analysis using the DerSimonian and Laird method(276) was used in this study to combine maximally adjusted hazard ratios (HR) using Metan in Stata version 13. This method assumes that the studies are not estimating the same exposure effects but instead are estimating different exposure effects that are related and follow a distribution across studies. The conventional choice of distribution is the normal distribution with smaller and larger exposure effects evenly distributed around the mean although this cannot be easily validated(274). The pooled summary effect in a random-effects meta-analysis reflects the centre of this distribution and the confidence interval, the variability around this 'centre' point.

A Tau^2 estimate was calculated to give an estimate of the between study variance which was then square rooted to give an estimate of the standard deviation in the underlying effects. Tau X 2 above and below the pooled estimate gives an approximation of the range of effects across 95% of the studies(274). As the pooled estimate was calculated on the natural logarithm scale then the range needed to be exponentiated to give the interval on the original scale. The weighting of individual study effect estimates in a random-effects meta-analysis takes account of the variation in effects by adjusting the standard errors of the individual studies with Tau^2.

$$\text{weight}_i = \left(\frac{1}{\text{weighti}} + \text{Tau}^2\right)^{-1}$$

The standard error of the pooled exposure effect was then used to derive a confidence interval and a *P* value.

- **Is the effect consistent across studies?**

Heterogeneity is a term used to describe the difference in study effect estimates and where present, often reflects both clinical and methodological diversity of the studies(243). Testing for heterogeneity provides a measure of whether the variation in the effect estimates of individual studies reflects random variation (or chance) or true difference in the study effects. It is common in prognosis studies for the population definition (in this case HF) and the exposure measure (chronic disease comorbidity) to differ across studies(272). This is only detrimental to the review if a specific definition of these factors is of interest. However if the object of the review is to ascertain the 'average' effect of having a chronic disease on the 'average' HF patient then this is not a problem(274). This review included HF patients that were not characterised by any specific characteristics and chronic disease exposures that could be routinely identified in practice. This means that the presence of disease as defined by routine practice could be investigated for its effect on the average HF patient.

Measuring heterogeneity: Two approaches for testing heterogeneity were used in this review: Cochran's Q test and I^2. The Cochran's Q tests whether the observed variation in effects of individual studies could be due to chance. A low *P* value of the test can be interpreted as evidence of heterogeneity across studies. However due to the low power commonly observed in meta-analysis resulting from small sample sizes or few studies, a non-significant *P* value should not be taken as evidence of no heterogeneity. A *P* value level of 0.1 is often used for this reason(273).

The Cochran's Q test provides a method to detect heterogeneity however it is usually understood that heterogeneity will always exist particularly across observational studies(272). An alternative approach to assessing heterogeneity and the method used alongside the Cochran's Q test in this review was to test the *level* of heterogeneity present using the I^2 statistic. The I^2 test was used to estimate the percentage of the variability in effect estimates that was due to heterogeneity as opposed to sampling error using this formula(278):

$$I^2 = \left(\frac{Q - df}{Q}\right) \text{ x } 100\%$$

where Q= Cochran's Q and df = degrees of freedom which is calculated as the number of included study effects – 1.

Whilst there is no clear consensus on what I^2 value may indicate important heterogeneity, a value of 40% or below is usually considered a low level of heterogeneity(274).

- ***What is the strength of evidence of the effect?***

Supporting evidence for the effect was investigated through exploration of bias introduced by explained heterogeneity or reporting bias. Random-effects meta-analysis should only be used to combine studies where heterogeneity cannot be readily explained(279) and the study effects are randomly distributed around the mean. Three approaches for exploring heterogeneity were considered for this review: sub-group analysis, meta-regression and sensitivity analysis. Sub-group analysis stratifies the total population (either at study or individual patient data level) by characteristics that may modify the effect. The sub-groups' effects are then compared to identify any significant difference between them(275). Random-effects meta-regression was also considered which allows multiple potential effect modifiers to be investigated together. This approach investigates the influence of the factor on the pooled effect estimate (which becomes the outcome in this case). Both approaches require a-priori selection of potential effect modifiers to prevent potential bias in selecting factors once study results are known. Sub-group analysis requires that individual studies have analysed data at the sub-group level which is often not consistent across studies. Meta-regression requires adequate numbers of studies and typically ten studies are recommended for each potential modifier investigated(280). Given these limitations it was decided to perform sensitivity analysis to explore heterogeneity. This form of exploration repeats the initial meta-analysis of the included studies but replaces certain decisions in the review process with alternative decisions to investigate whether the pooled effect remains unchanged by those decisions(281). Demonstrating the robustness of the pooled effect in this way

provides credibility to the findings of the review. Sensitivity analysis was performed for three a-priori characteristics included in the study protocol; setting, population definition and exposure definition and a post-hoc characteristic which was identified during the review which was risk of bias level.

In addition, Galbraith plots were used to further explore studies that may contribute to heterogeneity. In this plot each estimate is divided by its standard error (y axis) and this is plotted against the inverse of its standard error (x axis). Larger studies move to the right of the plot. Two reference lines are plotted which are 2 units above and below the pooled effect estimate line. These mark out a region where 95% of the study effects will lie. Studies that fall outside of these lines have confidence intervals that do not contain the pooled effect and can indicate heterogeneity(282).

Reporting bias: Reporting bias occurs when the dissemination of research findings are influenced by the size, significance or direction of the study results(267,275). Attempts to reduce reporting bias were made through using multiple resources for the review(Section 5.3.2). For each meta-analysis conducted where there were sufficient studies (>5) a funnel plot was performed to display small study effects. These plotted the standard error of individual studies on the y-axis against the hazard ratio on the x-axis, both on the logarithmic scale. The logarithmic scale meant that equal but opposing effect sizes such as 0.5 and 2 would be equidistant from 1(283). Where no publication bias exists the funnel plot will display an inverted funnel with an even broad spread of small study effects at the base and larger studies congregated more towards the estimated pooled effect at the top. Visual inspection was performed to explore asymmetry(284).

An Egger's test was performed which is based on the Galbraith plot. This test looks for a correlation between the study effect estimate and the weight given to the study (which was derived from the standard error). The Egger's test regresses the effect estimate divided by its standard error on the inverse of the standard error and tests the null hypothesis that the intercept is equal to 0 in the population. A low *P* value is evidence of publication bias. This test also usually has low power given the low number of studies usually included and so 0.1 is usually used as a reference point for evidence of bias(284).

Pooled effects from the random-effects meta-analyses were also compared to the pooled effects from fixed-effects meta-analyses. If the pooled effect was greater in the random-effects analysis (which gives more weight to the smaller studies) then this provided some evidence of small study effects. If this could not be rationalised by the sample size reasonably influencing the effect size (e.g. as might be the case if smaller studies had higher risk patients), then a decision to exclude the smaller studies had to be made. An alternative approach would have been to use fixed-effects meta-analysis which gives less weight to smaller studies, but has the disadvantage of not taking the variance in effects across studies into account(274).

5.3.6.2 Descriptive synthesis

Prognostic factor studies (general): The descriptive synthesis for the prognostic factor studies (general) used the following framework:

- Which chronic disease comorbidity factors have been found to be significantly associated with which outcomes?
- What is the size and direction of effects of the chronic disease in these studies?

Questions of whether effect estimates were consistent across studies and the strength of evidence of the effects were not explored given the selective nature of the studies that was based on statistically significant effects. Some discussion on the risk of bias in the included studies was however provided.

Prognostic model studies: The main question addressed for the prognostic model studies focused on establishing how chronic disease had been included in prognostic models and specifically whether the inclusion was via comorbid disease type, severity or severity change.

5.4 Results

Results will be reported as numbers (n), percentages (%), mean (standard deviation; SD) or median [interquartile range; IQR]. Effect estimates will be reported as Hazard Ratio (HR) with a 95% confidence interval (95% CI).

5.4.1 Identified prognosis articles on HF

A total of 10,331 unique articles were identified from the three databases and after title screening by the lead reviewer (CAR), 417 were retained. 1% of all the original articles excluded were checked by the second reviewer (UTK) for appropriate exclusion (n=99). Abstract screening was performed by three reviewers and consensus was reached for the exclusion of 327 articles. This resulted in 90 articles remaining for full article review. During the full article review by two reviewers (CAR, UTK), 27 articles selected in the abstract screen were found not to meet the inclusion/exclusion criteria and were subsequently excluded. Full consensus for the excluded articles was achieved by the two reviewers. This left 63 articles from the database searches to be included in the review. Four additional articles were identified from citation and reference searches. A further article was sourced from the one of the experts contacted resulting in a final set of 68 articles to be included in the review (Figure 5.1).

5.4.2 Overall characteristics of the studies

There were 34 prognostic factor (chronic disease focus) studies (E-Appendix A8) including 4 chronic diseases, 22 prognostic factor (general) studies (E-Appendix A9) including 7 chronic diseases and 12 prognostic model studies (E-Appendix A10) including 7 chronic diseases (summary provided in Table 5.3). The 68 studies were based in 16 countries covering 4 continents (North America, 35%, South America, 3%, Europe, 43%, Asia, 15%, multiple 4%). All studies used a cohort design; 26% (n=18) were secondary analysis of randomised controlled trials (RCTs), 60% (n=41) were hospital based, 12% (n=9) were hospital and community based and 2% (n=1) were community based. 90% of studies (n=60) had 1 or more years of follow-up and 35% (n=24) had 5 or more years of follow-up. Mortality outcome was investigated in 93% (n=63)

studies and hospital admissions in 16% (n=11) studies. There were no studies investigating HR-QoL outcomes.

5.4.3 Quality of the overall studies

Overall risk of bias was low in 34% (n=23) studies, moderate in 63% (n=43) studies and high in 3% (n=2) studies (E-Appendix A11). Agreement was good between the first (CAR) and second (UTK) reviewers with a Kappa Coefficient of 0.86 (E-Appendix E7 for worked example). Out of the 43 moderate risk studies, 6 had one of the six risk domains assessed as high risk, 7 studies had 2 domains at high risk and one study had 3 domains at high risk. 29 of the studies had no domains at high risk. The risk domains with the highest proportion of low risk scores were outcome measurement and study attrition (Table 5.4). In the outcome domain, 95% of the studies achieving a low risk score investigated mortality outcome. In the study attrition domain 91% of studies achieving a low risk score had less than 10% loss to follow up. Risk domains with the highest proportion of high risk scores were study confounding and statistical analysis and reporting. In the study confounding domain 63% of studies achieving a high risk score did not report what confounders were included in the analysis. Of the 65% of studies achieving a moderate risk score all reported confounders considered but only 9% provided any rationale or detail about their selection. In the statistical analysis domain 91% of studies achieving a high risk score did not report unadjusted effects or consider interactions between exposures and none of the studies using proportional hazards referred to testing model assumptions.

5.4.4 Characteristics of prognostic factor (chronic disease focus) studies

- ***Diabetes Mellitus (DM) studies***

The 11 studies focusing on DM(285-295), had a combined sample of 138,953 HF patients from 6 countries (North America n=4, South America n=1, Europe n=5, Multiple=1). HF was identified by clinical records in one study, administration codes (admission or discharge) in 4 studies (2 were verified by clinical records) and clinical diagnosis in 6 studies (4 included signs, symptoms and objective evidence and 2 unspecified). The combined sample was from 6 RCT, 1 hospital/community and 4 hospital samples. Mean ages of the samples ranged from 50 (SD 11) years to 77 (12) years. The samples showed the proportion of males was 46-73%,

systolic HF range was 47 to 87%, mean ejection fraction (EF) 27 (14) to 47(13)% and the prevalence range for New York Heart Association (NYHA) Class grade 3 and 4 was 36% to 84%.

Of the 138,953 combined sample of HF patients, 22,790 (16%) had DM. Prevalence of DM in individual studies ranged from 13-47%. Mean follow up was from 6 months to 5 years. 3 studies investigated the effect of DM by gender and one study by age. Two studies investigated DM severity by glycaemia (glycosylated haemoglobin (HbA1c), fasting glucose) and two studies stratified the sample by DM treatments. All studies investigated all-cause mortality and only one study investigated hospital admissions

Quality of DM studies: There were 64% (n=7) of studies that were assessed as moderate risk using the QUIPs tool and 36% (n=4) of studies that were low risk. No studies had any of the six QUIPs domains scored at high risk. All studies described the sample by key characteristics and gave some detail of the baseline population. Of the total 73% had <10% loss to follow-up (27% not reported), all described the chronic disease exposure and only one study used imputation methods for missing data, the outcome and confounders were described in all studies but no study provided a rationale for confounders selected in any detail. Only 36% of studies provided unadjusted associations, 64% examined for interactions but only 18% reported on whether proportional hazards assumptions had been checked.

- ***Chronic obstructive pulmonary disease (COPD) studies***

The 5 studies focusing on COPD(181,296-299) had a combined sample of 7,121 HF patients from 5 European countries. HF was identified by hospital discharge administration codes in 2 studies and clinical diagnosis in 3 studies (all using signs, symptoms and objective evidence). The combined sample was from 1 RCT, 1 hospital or community and 3 hospital samples. Ages of the samples ranged from mean 70 (SD 12) to median 80 [IQR 75-87] years. In the study samples the proportion of males was 50% to 71%, systolic HF prevalence range was 25% to 83%, mean EF was 33(12) to 50(16)% and the prevalence range for NYHA Class 3 and 4 was 52 to 95%.

Of the 7,121 combined sample of HF patients, 1,384 (19%) had COPD. Mean follow up was from 2.9 (SD 2.1) years to 4.5 (range 2.9-5.5) years. Only one study investigated COPD severity (reduction in Forced Expiratory Volume in one second [FEV_1]). All studies investigated all-cause mortality.

Quality of COPD studies: There were 60% (n=3) of studies that were assessed as moderate risk using the QUIPs tool and 40% (n=2) of studies that were assessed as low risk. One study had one domain scored at high risk and one study had two domains (study participation and prognostic factor measurement) scored at high risk. All studies described the sample by key characteristics and gave some detail of the baseline population. Of the total, 80% had <10% loss to follow-up (10% not reported), all described the chronic disease exposure, no studies used imputation methods for missing data, the outcome and confounders were described in all studies but no study provided a rationale for confounders selected in any detail. All studies provided unadjusted associations but only 20% examined for interactions or reported on whether proportional hazards assumptions had been checked.

- ***Renal dysfunction (RD) studies***

The 17 studies focusing on RD(155,156,300-314) had a combined sample of 102,638 HF patients from 10 countries (North America n=7, South America n=1, Asia n=4 and Europe n=5). HF was identified by administration codes (discharge) in 2 studies, clinical diagnosis in 14 studies (11 included signs, symptoms and objective evidence and 3 unspecified) and was unspecified in 1 study. The combined sample was from 5 RCTs, 1 hospital/community and 11 hospital samples. Ages of the samples ranged from mean 62 (SD 15) to median 80 years. In the samples, the proportion of males was 43 to 76%, systolic HF prevalence was 36% to 82%, mean EF was 27(12) to 44(16)% and the prevalence range for NYHA Class grade 3 and 4 was 32% to 84%.

All renal studies used blood measurements to define RD. Mean follow up was from 6 months to 6.5 years. Out of the 17 renal studies, 8 studies investigated 'any' RD and 6 of these categorised RD by eGFR severity level. 2 of the 8 studies also investigated increments of eGFR by mls/min. Any RD was defined as all categories of renal function with eGFR of <60ml/min. Mild RD was defined as eGFR 60-89ml/min, moderate

RD as eGFR 30-59ml/min and severe RD as eGFR <30ml/min compared to the highest category(254). For prevalence of RD see Table 5.5.

There were two additional studies that also investigated increments of eGFR by mls/min and the remaining 7 studies investigated change in renal severity. Moderate change in severity was defined as an increase in creatinine (Cr) of ≥0.3mg/dL from hospital baseline to study defined end point (4 out of the 5 studies including this definition used discharge as the end point) and severe increase in severity by an increase in Cr of ≥0.5mg/dL. This classification was based on previous studies(157,315). Out of the 7 studies investigating change in renal severity, 5 studies used change in Cr as a measure, 1 study used a reduction in eGFR and 1 study investigated improvement (increase) in eGFR. Two studies stratified samples by timing or duration of severity change (Table 5.6).

All 17 studies investigated all-cause mortality and 4 studies investigated hospital admissions.

Quality of renal studies: There were 6% (n=1) of studies that were assessed as high risk using the QUIPs tool, 65% (n=11) studies were moderate risk and 29% (n=5) studies were low risk. Two studies had one domain (study confounding, statistical analysis and reporting) scored at high risk and two studies had three domains (study participation/prognostic factor measurement, study confounding, statistical analysis and reporting). All studies described the sample by key characteristics and 94% gave some detail of the baseline population. Out of the total, 47% had <10% loss to follow-up (41% not reported), all described the chronic disease exposure, 18% studies used imputation methods for missing data, the outcome was described in all studies and 88% studies described the confounders considered and 24% provided a rationale for the confounders selected. Of the total, 50% studies provided unadjusted associations and adjusted associations, 53% examined for interactions and 24% reported on whether proportional hazards assumptions had been checked.

- ***Other non-CVD comorbidity (Chronic Disease Focus) studies***

There was one study focusing on rheumatoid arthritis (RA)(316) which had a sample of 955 de novo HF patients from a hospital setting in North America. HF was identified clinical diagnosis using signs, symptoms and objective evidence. Mean age of the sample was 77 years (SD 12). The proportion of males was 45%, 51% had systolic HF, mean EF was 44%.103 (11%) had RA. RA was identified from clinical diagnosis and patients were followed up for 1 year to investigate all-cause mortality. The study was of moderate risk of bias using the QUIPs tool with two domains out of the possible six scored at high risk (study confounding and statistical analysis and reporting).

5.4.5 Characteristics of prognostic factor (general) studies

From the 22 prognostic factor (general) studies(132,139,317-335), sample sizes ranged from 181 to 62,330 HF patients from 12 countries (North America n=7, Asia n=6 and Europe n=9). HF was identified by clinical records in 1 study, administration codes (discharge) in 6 studies (1 was verified by clinical records) and clinical diagnosis in 15 studies (all included signs, symptoms and objective evidence). The samples were from 1 RCT, 17 hospital samples, 3 hospital/community and 1 community sample. 3 samples were drawn from the same hospital cohort and 6 samples were de novo HF. Mean ages of the samples ranged from 66(SD12) to 86(5) years. In samples the proportion of males was 40% to 71%, prevalence of systolic HF ranged from 36% to 79%, mean EF was 35(14) to 54(15)% and prevalence of NYHA Class grade 3 and 4 was 12% to 96%.

The 22 studies included 7 different non-CVD comorbidities. Prevalence of the chronic diseases in the different studies ranged from 1% (renal failure) to 70% (RD) (Table 5.7). Chronic diseases were identified from clinical diagnosis (n=1), clinical records (n=9), blood test (n=6), a combination of clinical records, prescriptions, blood tests or self-report (n=7), administration codes (admission or discharge) (n=2), and not specified (n=3). Mean follow up was from 6 months to 6 years. Only one study stratified the sample by age and one by gender and two studies focused specifically on women. All-cause mortality was investigated in 19 studies and 4 studies investigated hospital admissions. A chronic disease severity indicator was included in 4 renal studies and 1 renal study included a severity change measure.

Quality of the prognostic factor (general) studies: There were 5% (n=1) studies assessed at high risk of bias, 64% (n=14) studies were moderate risk and 32% (n=7) studies were low risk based on the QUIPS tool. Two studies had one domain (prognostic factor measurement, statistical analysis and reporting) scored at high risk, five studies had three domains (study participation/prognostic factor measurement/study confounding and statistical analysis and reporting) and one study had three domains (study participation, study confounding and statistical analysis and reporting) scored at high risk. All studies described the sample by key characteristics and gave some detail of the baseline population. Most of the studies (77%) had less than 10% loss to follow-up (but 23% of the studies did not report this figure). Of the total, 68% described the chronic disease exposure, 9% studies used imputation methods for missing data, the outcome was described in all studies and 86% studies described the predictors considered. Only 14% provided a rationale for the predictors selected and 27% studies provided unadjusted associations as well as adjusted associations. Some of the studies (23%) had examined interactions between HF and comorbid conditions and 29% of applicable studies reported on whether proportional hazards assumptions had been checked.

5.4.6 Characteristics of prognostic model studies

From the 12 prognostic model studies(175,179,188,189,258,259,336-341) sample sizes ranged from 152-198,640 HF patients from 2 continents (North America n=6 and Europe n=6). HF was identified by administration codes (discharge) in 4 studies (1 was verified by clinical records) and clinical diagnosis in 7 studies (all included signs, symptoms and objective evidence). One study used individual patient data from 30 different cohort studies. The samples were from 5 RCTs, 5 hospital samples and 2 hospital/community samples. Mean ages of the samples ranged from 66(SD 11) to 79(6) years. In the samples, the proportion of males was 34% to 98%, prevalence of systolic HF was 44% to 90%, mean EF was 33(9) to 43(14)% and prevalence of NYHA class 3 to 4 was 25% to 55%.

The 12 model studies included 7 different chronic diseases. Prevalence of the chronic diseases in the different studies ranged from 1% (liver cirrhosis) to 36% (DM) (see Table 5.7). Chronic diseases were identified from clinical diagnosis (n=1), clinical records (n=10), blood test (n=2) and a combination of clinical records and prescriptions (n=1). Mean follow up was from 2.4 months to 5.2 years. Most studies (n=10) had investigated

all-cause mortality and 2 studies all-cause hospital admissions. Only 4 studies had used a measure of chronic disease severity.

Quality of the prognostic model studies: There were 58% (n=7) of studies that were assessed as moderate risk using the QUIPs tool and 42% (n=5) of studies that were low risk. One study had one domain (study participation) scored at high risk. All studies described the sample by key characteristics and gave some detail of the baseline population. Of the total 58% studies had <10% loss to follow-up (42% not reported), 50% described the chronic disease exposure and 42% studies used imputation methods for missing data. The outcome and the predictors considered were described for all studies. Of the total, 42% studies provided a rationale for the predictors selected and 33% provided unadjusted associations as well as the final adjusted model. Most (67%) studies examined for interactions and 43% of applicable studies reported on whether proportional hazards assumptions had been checked. Model reduction was achieved in 42% of studies by backwards or forwards selection, 25% by stepwise selection, 25% by the significance level in univariate associations and 1 by another reduction technique. The majority of studies (83%, n=10) reported more than 10 events per predictor included and measures of internal validity, model performance and calibration. Testing of external validity was only reported in 25% of the studies. All models derived were used to create a clinical risk score.

5.4.7 Associations between non-CVD comorbidities and outcomes in HF

The chronic disease comorbidity studies will now be summarised for each chronic disease by *prognostic outcome* structured by (i) prognostic factor studies (chronic disease focus) including meta-analysis where appropriate (ii) prognostic factor studies (general) and (iii) prognostic model studies.

5.4.7.1 Diabetes comorbidity in heart failure

All-cause mortality

- *Prognostic factor (chronic disease focus) studies:*

Nine studies reported all-cause mortality rates associated with comorbid DM status (presence or absence) adjusted for socio-demographic and clinical covariates for mean follow-up ≥6 months (range 6 months to 7.5 years). The study samples were derived from RCT and hospital settings with only one study including community patients(294). Out of 139,761 HF patients, 21,921 (16%) had DM (range 13-47%). Over the follow up period 63% of patients without DM died compared to 62% of those with DM. This translated into a combined adjusted mortality risk of Hazard Ratio (HR) 1.34 (95% confidence interval [CI] 1.24-1.46) (Figure 5.2) using random effects meta-analysis and the test of overall association was significant (p=<0.0001).

Heterogeneity: The Cochran's Q test indicated significant heterogeneity (p=0.008). The I^2 value indicated that 61% of the variability in effect estimates was due to heterogeneity as opposed to sampling error. Tau^2 was used to provide an approximation that the range of associations across most of the studies (95%) was Hazard Ratios from 1.14 to 1.58 (E-Appendix C1). The included studies were similar in terms of their settings, population definition, exposure definition and risk of bias. One study included community patients(294) and one study included more chronically severe patients (NYHA stage 3-4 within past month)(287). Following removal of the first study(294) heterogeneity remained significant (p=0.004, I^2=66%) (E-Appendix C2). Replacing the first study and removal of the second study(287) reduced the level of heterogeneity and it became non-significant (p=0.15, I^2=35%). This did not alter the combined effect estimate which remained elevated and significant, HR 1.30 (1.21-1.39) (E-Appendix C3). A Galbraith plot also identified Gustafsson et al (2004) as outside of the 95% range of study effects (E-Appendix B1).

Publication bias: The pooled effect estimate from the random effects meta-analysis was slightly larger than the pooled effect from the fixed effects meta-analysis (HR 1.27) which indicated small study effects (Figure 5.2). A funnel plot revealed possible publication bias with two of the smallest studies having the biggest effect estimates (E-Appendix B2). An Egger's test was performed (p=0.119) (E-Appendix C4) which, given the low power of the test, provided some evidence of publication bias. However, removal of the two smaller studies(285, 289) in a sensitivity analysis made little change to the pooled effect estimate which became HR 1.31(1.22-1.41) p=<0.0001 (E-Appendix C5) and Eggers's test (p=0.43) (E-Appendix C6) indicating no significant evidence of publication bias. See E-Appendix A12 for sensitivity analysis summary.

Sub group analysis: Three studies had investigated all-cause mortality risk associated with DM stratified by gender. Out of 64,697 women and 61,462 men 16% and 15% respectively had DM. Over the follow-up (4.8 to 8.5 years) 74% of diabetic women versus 69% of non-diabetic women and 67% of diabetic men versus 65% of non-diabetic men died. This resulted in a combined HR of 1.52 (95%CI 1.22-1.89) for women and 1.27 (1.17-1.38) for men (Figure 5.3). Whilst there was some overlap in the confidence intervals of the combined effect estimates by gender each individual study found a significant sex-DM interaction with women worse off than men. Two studies stratified the groups by age with one finding significant sex-DM interaction to be present in age ≥65yrs only (p=0.005)(292) and the other also finding significant interaction in the younger (<65yrs) and older age groups (65-74yrs) (p=0.005). Within gender groups significant age-DM interaction was found with <65yrs worse off (p=0.001)(291).

Diabetes comorbidity severity: One study included a DM severity indicator by treatment type(293) and two studies included a measure of glycaemia to investigate the associated risk of all-cause mortality(290). In the first study, out of 400 HF patients from a hospital cohort, diabetics were stratified into three groups by diabetes treatment type and a group of undiagnosed DM. The latter group was defined as any subject with multiple raised fasting plasma glucose measures in their outpatient's notes but no formal diagnosis of DM recorded. These groups were compared to the no DM group. Whilst non-significant effects were found for oral and diet treated DM the effects increased and became significant in the undiagnosed (not treated, HR 1.69) and insulin treated group (HR 2.11) (Figure 5.4).

The second study (N=2412) had investigated the level of glycosylated haemoglobin (HbA1c) (290). In this study, 38% of the HF sample had DM but the measure was taken on all patients regardless of their DM status and the risk per 1% higher HbA1c within the RCT study sample was HR1.14 (1.06-1.23). There was a significant interaction between DM status and HbA1c with those with no history of DM experiencing worse effects per 1% rise in HbA1c (p=.008). The third study (N=456, RCT cohort) found no associated risk of DM (not reported) but an increased risk associated with glycaemia ≤ 5.5mmol/L of HR 1.45 (1.09-1.69) (288).

- *Prognostic factor (general) studies and prognostic models*

The range of effect estimates for the risk of all-cause mortality associated with DM were HR 1.16 to 3.19 in the prognostic factor (general) studies (E-Appendix B3) and HR 1.34 to 2.37 in the prognostic model studies (E-Appendix B4). DM was included in all 14 prognostic factor (general) studies by type. Only three studies included community patients(132,259,322). A significant interaction was found between age and DM in one study(139) with younger age groups worse off (p=0.014).

Three of the six prognostic models included a severity indicator. The first included DM stratified by insulin prescription(179) (HR 1.8) and 'other' (HR 1.5) and the second two studies(189,259) included DM with target organ damage. None of the prognostic factor or model studies included an indicator of severity change.

All-cause hospital admissions

- *Prognostic factor (chronic disease focus) studies:*

Only one study had investigated the effect of DM on all-cause hospital admissions within a RCT cohort and found an associated risk of HR 1.28 (1.19-1.38)(292). Again women (HR 1.49) were worse off than men (HR 1.21).

- *Prognostic factor (general) studies and prognostic models*

The range of effect estimates for the risk of all-cause hospital admissions associated with DM were HR 1.13 to 1.53 in the prognostic factor (general) studies and HR 1.17 in the prognostic model study (E-Appendix B5). Two studies had included community patients (discharged from hospital)(318,320). One study reported an interaction between age and DM with younger age groups worse off (p=0.014)(139). DM was included in all 5 studies by type.

5.4.7.2 Chronic Obstructive Pulmonary Disease (COPD) comorbidity in heart failure

All-cause Mortality

- Prognostic factor (chronic disease focus) studies

Five studies reported all-cause mortality rates adjusted for socio-demographic and clinical covariates for mean follow-up ≥1 year (range 1-8yrs). Out of 7,121 HF patients, 1309 (18%) had COPD (range 17-35%). Only one study had included community patients(73). Over the follow up period, in the 3 studies providing crude mortality data 38% of patients without COPD died compared to 48% of those with COPD. This translated into a combined adjusted all-cause mortality risk of HR 1.39 (95% CI 1.21-1.6) (Figure 5.5) using random effects meta-analysis. The test of overall effect was significant (p=<0.0001).

Heterogeneity: The Cochran's Q test indicated low evidence of heterogeneity (p=0.17) and I^2 was 38%. Tau^2 provided an approximation that the range of effects across 95% of the studies was HR1.27 to 1.54 (E-Appendix C7). Whilst the test for heterogeneity was non-significant, random effects meta-analysis was used due to the presence of some heterogeneity indicated by the I^2 test. The included studies were similar in terms of their settings and exposure definition. One study had a lower proportion of systolic HF patients (25%) than the other studies and two risk of bias domains assessed as high(181). This study was removed in a sensitivity analysis (E-Appendix C8) and the effect estimate remained elevated and significant, HR 1.41 (1.18-1.68). A Galbraith plot supported the low level of heterogeneity with no studies falling outside of the 95% range of study effects (E-Appendix B6).

Publication bias: The pooled effect from the random effects meta-analysis was similar to the pooled effect from the fixed effects meta-analysis (HR 1.36) which indicated no evidence of small study effects (Figure 5.5). The funnel plot shows that most studies (n=4) had similar sample sizes with one larger study contributing a smaller effect estimate (E-Appendix B7). An Egger's test was performed (E-Appendix C9) which provided only limited evidence of publication bias (p=0.118).

COPD comorbidity severity: One study (N=532) investigated the association between COPD severity and all-cause mortality risk using forced expiratory volume over a second (FEV_1)(296). Whilst the risk (unadjusted HR

1.26) was non-significant in the moderate severity group (FEV_1 50-79%) this became significant and elevated (HR 1.68) in the severe group (FEV_1 <49%) (Figure 5.6**).** Increasing FEV_1 (per 10% of predicted) was protective, adjusted HR 0.86 (0.8-0.92).

- Prognostic factor (general) studies and prognostic models

The range of effect estimates for the risk of all-cause mortality associated with COPD were HR 1.24-1.7 in the prognostic factor (general) studies and HR 1.23-1.6 in the prognostic model studies (E-Appendix B8**).** Two of the studies included community patients(132,259) one of which compared community patients to hospital patients and found a significant difference between the associated risks(132). For the community sample the risk was HR 1.7(1.58-1.82) and for hospital patients HR 1.24(1.19-1.31). COPD was included in all ten prognostic factor (general) and prognostic model studies by type.

All-cause hospital admissions

- *Prognostic factor (general) studies and prognostic models*

No studies had focused on the association between COPD and hospital admissions. Only one prognostic factor (general) study and one prognostic model study included COPD to investigate hospital admissions (HR 1.47 and OR 1.14 respectively, Figure 5.7). Both studies which included community patients included COPD by type.

5.4.7.3 Renal dysfunction (RD) comorbidity in heart failure

All-cause Mortality

- Prognostic factor (chronic disease focus) studies

Seven studies reported all-cause mortality rates adjusted for socio-demographic and clinical covariates for mean follow-up >1year (median follow up range 20 to 38months). Only one study had included community patients(313). Out of 69,520 HF patients, 28,596 (41%) had 'any' RD (range 36 to 70%). Over the follow up period 42% of patients without RD died compared to 51% of those with RD. This translated into a combined

adjusted mortality risk of Hazard Ratio (HR) 1.52 (95% CI 1.34-1.71) using random effects meta-analysis (Figure 5.8). The test of overall effect was significant (p=<0.0001).

Heterogeneity: The Cochran's Q test indicated significant heterogeneity (p=<0.0001) and I^2 was 88%. Tau^2 was used to provide an approximation that the range of effects across 95% of the studies was HR 1.2 to 1.6 (E-Appendix C10). The included studies were similar in terms of their settings and exposure definition. One study had higher proportion of males (76%) and a lower mean EF (32%)(301) and another had an older population (mean age, 76 years) with a lower proportion of males (43%)(306) compared to the other studies. Following removal of the first study(301) heterogeneity remained significant (p=<0.0001, I^2=83%) (E-Appendix C11). Replacing the first study and removal of the second study(306) had little effect on the heterogeneity (p=<0.0001, I^2=81%) (E-Appendix C12). Removal of both studies removed the heterogeneity (p=0.825, I^2=0%). This increased the combined effect to HR 1.62 (1.59-1.66) (E-Appendix C13). A Galbraith plot identified three studies including the two removed in sensitivity analysis as falling outside of the 95% range of study effects (E-Appendix B9). See E-Appendix A13 for sensitivity analysis summary.

Publication bias: The pooled effect from the random effects meta-analysis was slightly smaller than the pooled effect from the fixed effects meta-analysis (HR 1.58) which indicated no small study effects (Figure 5.7). A funnel plot (E-Appendix B10) and Egger's test (p=0.56, E-Appendix C14) revealed no evidence of publication bias. One further study found eGFR >53ml/min compared to <35ml/min to be protective (HR 0.65)(302).

Renal dysfunction comorbidity severity: Five studies reported adjusted all-cause mortality rates stratified by severity groups. Out of 64,257 HF patients, 24,349 (38%) had moderate RD and 3784 (6%) had severe RD. Over the follow up period 42% of patients without RD died compared to 48% of those with moderate RD and 63% of those with severe RD. Random effects meta-analysis was performed, stratified by severity group (Figure 5.9). This resulted in a combined adjusted mortality risk of HR 1.01 (0.84-1.22) in the mild severity group, HR 1.21 (1.18-1.24) in the moderate group to HR 2.01 (1.60-2.52) in the severe group using random effects meta-analysis.

Meta-regression was used to plot the upper eGFR limit for each study defined severity category (x-axis) against their associated hazard ratio (y-axis) and there was a dose response relationship between the effect estimates and reducing eGFR upper limit (Figure 5.10). Three hospital studies(302,308,312) and one RCT study(304) investigated all-cause mortality risk by mls/min/m^2 eGFR (Figure 5.11). Risk reduced with increasing eGFR (above dashed line) and increased with reducing eGFR (below dashed line).

Renal dysfunction comorbidity severity change: Five studies investigated change in renal function from hospital admission baseline to discharge or study defined endpoint. Random effects meta-analysis was performed, stratified by severity of renal function change (Figure 5.12). The combined adjusted hazard ratios were significant and increased from 1.53 (1.09-2.14) in the moderate change group to 2.29 (1.63-3.21) in the severe change group. One study found a significant interaction between baseline creatinine and worsening renal function with the higher than median baseline creatinine group being worse off(311).

One further study investigated a monthly percentage reduction in eGFR of >=1% (measured in the 1 year either side of hospital admission) and found an associated adjusted mortality risk of HR 3.6 (2.2-5.7)(307). Another study investigated improved renal function (≥20% increase eGFR) during hospital admission and found an elevated adjusted mortality risk of HR1.3 (1.1,1.7)(314).

- *Prognostic factor (general) studies and prognostic models*

The range of effect estimates for the risk of all-cause mortality associated with RD were HR 1.35 to 2.27 in the prognostic factor (general) studies (E-Appendix B11) and HR 1.37 to 5.22 in the prognostic model studies (E-Appendix B12). Only two studies included community patients(132,259). One study compared the associated risk of severe renal dysfunction in older and younger groups and found a significant difference between the risks(139). The associated mortality risk was HR 1.36 (1.13-1.63) in the older age group and HR 2.21 (2.02-2.43) in the younger age group. Renal dysfunction was included in 4 prognostic factor (general) studies and one model study by a severity indicator and in one factor study by severity change.

All-cause hospital admissions

- Prognostic factor (chronic disease focus) studies

One study(301) focused on the association between different severities of RD and risk of all-cause hospital admission. Over the study follow-up 66% of patients without RD experienced a hospital admission compared to 68% of those with moderate RD and 73% of those with severe RD. Adjusted risk increased with severity group (Table 5.8). Another study(302) found that an eGFR of >53ml/min compared to <35ml/min had an adjusted protective effect of HR 0.77 (0.56-1.06).

In hospital change in renal function of Creatinine (Cr) ≥3mg/dL was found in one study to be associated with increased risk of readmission(156) (HR 1.5) and insignificant in another study(155). All studies were hospital or RCT based.

- Prognostic factor (general) studies and prognostic models

RD was associated with increased risk of hospital admission in one prognostic factor (general) study (HR 1.32) and one prognostic model study (OR 1.09) (E-Appendix B13)(338). Both studies included community patients(318,338).

5.4.7.4 Other non-CVD comorbidities and heart failure

Twelve prognostic factor (general) and model studies included 5 additional other non-CVD comorbidities: arthritis, dementia, cancer, lung disease and liver disease. All chronic diseases were included by type (Table 5.9 for the adjusted risk estimates). Only one study, including three diseases had investigated hospital admissions(338)

5.4.7.5 Comparing comorbidity effects

Meta-analysis of the association of the three main non-CVD diseases; DM, COPD and RD, were compared for the mortality outcome (Figure 5.13) and all three comorbidities had similar adjusted and significant effects on mortality in HF. The effect estimates were HR 1.3 for DM, 1.4 for COPD and 1.5 for RD.

5.4.8 Discussion

Summary of findings: Non-CVD comorbidities are common and are strongly and independently associated with poorer outcomes in HF. Comorbidity prognosis studies for non-selected HF populations have to date focused on hospital settings and the outcome of mortality, but their importance in other HF settings and for hospital admissions has been demonstrated in general HF prognosis studies. The impact of comorbidities on mortality, where most of the evidence lies, appears similar across chronic diseases but there is some evidence to suggest that it differs across different HF populations. Differentiation of risk within a population for a specific comorbidity is evident once severity of the comorbid disease is taken into account with clear examples for DM, COPD and RD. Comorbidity severity change is also predictive of higher mortality and increased risk of hospital admission, but here the evidence is limited to RD comorbidity in hospitalised HF populations. Prognostic models for HF often include an indicator of HF disease severity but have yet to take account of comorbidity severity or of how the impact of comorbidity may in part depend on the severity of the index disease.

The scope of comorbidity: Comorbidities were highly prevalent within the included studies with up to at least a third of HF patients experiencing DM, COPD or RD. Other diseases with high prevalence such as arthritis and dementia featured far less and there were no studies that included other prevalent diseases such as osteoarthritis(129). The estimates of disease prevalence varied across the studies and depended on the diagnostic criteria applied as well as the measurement tools used. Spirometry use in COPD diagnosis reveals higher prevalence than self-report with levels of 40% reported in studies using spirometry(342) and less than half of those diagnosed on spirometry, self-reporting COPD(343). DM identified by anti-diabetic drug use misses those controlled by diet and hospital based HF populations reported higher prevalence of DM(293). The most accurate identification of comorbidity is probably in renal disease where blood test indicators are commonly used in diagnosis. However even in renal disease the cut off value for diagnosis often differs across studies and accounts for some of the variation identified.

Outcomes: The outcome of mortality predominated in the current prognosis studies in 93% of studies. Only 16% of studies had included hospital admission and 5 of these focused on a chronic disease exposure,

representing only two diseases; DM (n=1) and RD (n=4). This is surprising given the high number of hospital admissions associated with HF and the association between comorbidities and hospital admissions found in cross sectional studies(129). In a large incident HF population 83% of patients experienced at least one admission and 54% experienced three or more admissions in 5 years of follow up and non-cardiac comorbidities accounted for 62% of those admissions(320).

Remarkably, among the 68 studies included in the review there was no study that investigated the association between non-CVD comorbidities in non-selected HF and HR-QoL outcomes. This was a surprising finding given the low HR-QoL inherent in the HF population that worsens as the disease progresses(116). HR-QoL is itself associated with hospital admissions and mortality(344-346) and so provides a clear indicator of disease progression. In cross sectional studies(136,347) and selected HF populations comorbidity has been found to influence QoL and future QoL(348,349) but its impact in contemporary general populations of HF patients has yet to be investigated.

Settings: Only one study from the 68 included was set solely within a community setting. Eight further studies included community patients. There is evidence that the community population of HF patients differs significantly from hospital based populations which represent a more severe group with poorer prognosis(350-352) and much higher prevalence of chronic disease comorbidities(132). This has implications for the investigation of comorbidity exposure where there is evidence of a differential exposure effect in community and hospital settings with a significantly greater relative risk of exposure from comorbidities in the community setting(132,322). This is an important consideration for the development of prognostic models for the general population of HF patients particularly at earlier stages of disease and relating to quality of life and hospital admissions.

Comorbidity as a prognostic factor: Consistent predictors of mortality and hospital admissions in this review were comorbid DM, COPD and RD. The strength of effect was similar across the different diseases ranging from a 34% increase risk of mortality for DM to 52% increase risk for RD. The magnitude of effect associated with the individual comorbidities is important for four main reasons. First, it is well recognised that the latency

period between chronic disease comorbidity exposure and outcome is likely to be long and in HF often occurring many years prior to the onset of the index disease. Second, the exposure status is not static and will change over the progression of the index disease making precise measurement difficult. Third, the long latency period and imprecise and fluctuating exposure effect is likely to lead to small to moderate exposure effects which may be difficult to detect without careful epidemiological and statistical approaches(1). Fourth, HF is a severe index disease which is common in older people. These factors carry a high risk of mortality and hospital admission and so any significant relative exposure effect is likely to have an important impact on the absolute risk difference compared to non-exposed groups. The long latency period between exposure and outcome and the dynamic status of comorbidity exposure leads to the hypothesis that prognosis approaches that capture a *change* in exposure status may result potentially in bigger exposure effects.

The association between HF and different chronic disease comorbidities is complex with mechanisms that include shared risk factors, pathophysiological links, diagnostic and management conflicts or merely by chance as a function of age.

Comorbid diabetes mellitus: In this review findings show that DM was strongly and independently associated with mortality (34% higher risk) and hospital admission (28% higher risk). DM shares many of the risk factors of cardiovascular disease which is a common comorbidity(353) and the Framingham study showed an increase of cardiovascular risk by up to 3 times in diabetics compared to non-diabetics(354). HF also presents earlier and at higher rates in type 2 diabetes(355).

Hyper-glycaemia is known to activate neurohormonal and metabolic pathways which trigger a number of processes including vascular inflammation, endothelial dysfunction and oxidative stress in turn leading to; early and widespread atherosclerosis(356), apoptosis and fibrosis in the heart and kidney(357), reduced left ventricular function(358) and insulin resistance(359). Hypertension, diabetic nephropathy and cardiomyopathy are common in DM(360). Cardiovascular symptoms may be atypical or non-existent in the diabetic patient(361) and so diagnosis and management can be delayed leading to more advanced disease.

The gender difference in DM effect was significant in this review with women almost 30% more at risk of mortality and hospital admission than men. Women with DM tended to be older and with more severe heart disease than men(292) but these factors did not influence mortality in women compared to men in the non-diabetic group. However given the potential of this interaction in the older age groups it is likely to be, in part, due to the smaller DM effect in men which did not increase with age. The increased effect in older women may be a function of age, delayed and more extensive presentation of cardiovascular disease in diabetic women(362), poorer management of risk factors(363), increased insulin resistance and cardiovascular disease progression in post-menopausal women(364) or higher prevalence of preserved ejection fraction.

Comorbid COPD: COPD has been found, in previous studies, to be clearly associated with cardiovascular disease morbidity and mortality(365-367) however it has not been well explored in the context of HF and whilst many studies have included COPD as a factor only five studies focused on this association in this review. Pooling of these similar studies reported the strong and independent association between COPD and mortality (39% increased risk) and hospital admission (up to 47% increase risk) in the non-selected HF population.

It is well recognised that COPD causes chronic hypoxia and hypercapnia, neurohormonal activation and low grade systemic inflammation which in turn is associated with myocardial loading, deranged metabolic processes, atherosclerosis and cardiovascular events(368-370). Right ventricular dysfunction is common with or without pulmonary hypertension in COPD and is associated with increased mortality(371). COPD can mask the common diagnostic signs and symptoms of HF which leads to a later diagnosis and delayed management(372). Investigations such as echocardiogram and BNP can be unsatisfactory and HF can be easily missed where pulmonary vascular remodelling can hide pulmonary oedema and chest hyperinflation reduces the cardiothoracic ratio(96,373). The symptom of breathlessness is shared by the two diseases and there is no unique feature that differentiates them(374). Once HF diagnosis is made, challenges exist in the application of often conflicting treatment regimens with the clear example of the under-prescribing of beta-blockers(298) which amounted to less than 10% of patients in one study(299).

HF can also be misdiagnosed in COPD as jugular venous distention, ankle oedema, and hepatomegaly that are each indicative of right ventricular failure can also be caused by lung hyperinflation with hepatic displacement found in COPD. This misdiagnosis of HF can often be seen in the overrepresentation of COPD in HF with preserved ejection fraction(375). Comorbid COPD has been found to be significantly associated with mortality in this review where the proportion of HFpEF was low(297), where ejection fraction was accounted for in the analysis(299) and where the interaction between COPD and ejection fraction was specifically tested(296). Where studies include a proportion of misdiagnosed HF the effect estimate is likely to be an underestimate of the true effect.

Comorbid renal dysfunction: RD has been shown to be associated with adverse outcomes in HF previously in studies largely focused on selected HF populations(376-380). This review was able to combine studies including a more general population of HF patients encountered in the 'real world' of clinical practice, where RD was also significantly associated with mortality (52% increased risk) and hospital admissions (77% increased risk in severe disease).

However, the temporal relationship between the two diseases is not clear. Shared risk factors including age, hypertension, low high-density lipoprotein cholesterol and DM which result in atherosclerosis of renal disease(381) mean that RD and HF have a bi-directional relationship. However whilst this may explain some of the increased risk associated with RD, adjustment of these factors in analysis did not diminish the risk which remained significant and elevated(155,156,301,303,306,380). RD has also been considered as a marker of haemodynamic status and severity of the underlying cardiac disease(382). However adjustment for HF severity in a number of studies included in this review did not support this(156,301,310).

Whilst the precise mechanism of the temporal evolution of the cardio-renal syndrome is not known, what is clear is that once present, both conditions serve to aggravate each other. In chronic kidney disease, hypertension, arteriosclerosis and anaemia lead to myocardial pressure and volume overload in turn leading to hypertrophy and ventricular dilatation(383) and in HF poor cardiac output leads to reduced renal perfusion. Both diseases trigger neuro-hormonal mechanisms and their harmful counterparts; inflammation, endothelial

dysfunction and altered metabolic processes, all of which have been indicated as powerful mediators in cardio-renal syndrome(384). Management of HF can also be challenging in RD particularly in relation to diuretics(385) and ACE inhibitors(386,387) and patients are less likely to be prescribed first line drugs(303,306,388).

HF severity: The effect of the comorbidity exposure will be influenced by both the severity of the index disease and the associated comorbidity. This will depend to some extent on the mechanisms that link the diseases together with completely unrelated comorbidities competing with the risks of the index disease for specific outcomes and interrelated comorbidities that share pathophysiological processes and potentially modifying the risk of the index disease. Unlike cancer studies where the risk of comorbidity diminishes as the cancer progresses(389), in HF, comorbidity is found to have a reduced but still important effect in more advanced disease.

The importance of the severity of the index disease for the risk of comorbidity exposure was evident for comorbid DM studies. The mortality risk of DM was significantly higher in the younger groups than the older groups(291). Follow-up in this study was 5 years, so the risk of death from the index disease becomes higher as the group ages, thus diminishing the *relative* risk exerted by the DM. This DM-age interaction was not evident in the Digitalis Investigation Group trial(292), but this may be in part explained by the lack of adjustment for ejection fraction or HF severity in the former study. The increased effect of DM in the younger group was likely to be a function of less severe disease in this group which strengthens the hypothesis that index disease severity contributes to the relative impact of DM on mortality. Higher ejection fraction has in another study been associated with a greater DM impact(390) and again, this could be a function of reduced baseline risk in this group yielding higher DM relative risk or alternatively a function of competing risk, with DM having lesser effect in the more severe patients with greater systolic dysfunction. DM has been found not to contribute significant risk to mortality in advanced HF with very low ejection fraction(391).

Further examples of the impact of HF severity include the differential effect of comorbid COPD in different HF populations. The relative impact of comorbid COPD for mortality was found to be significantly worse in the

community setting than the hospital setting(132) which may reflect the less severe HF group. In comorbid renal disease there was a significant difference in the exposure effect between younger and older groups with the greatest effect of renal dysfunction in the younger group(139) and also in the community and hospital setting with worse exposure effects in the community setting(132).

Comorbidity severity: Comorbidity severity should be an important factor in prognosis studies. In this review the potential importance of disease severity for each of the three main comorbidities was identified. DM severity was included in this review in 3 focused studies. Firstly treatment type was found to differentiate DM risk with patients prescribed insulin having a more than 50% increase in risk of mortality than patients prescribed oral medication or diet control(293). Increased risk in type 1 DM was also included in a prognostic model for mortality in HF with 30% more risk in the insulin group than the non-insulin DM group(179). Although evidence on DM severity by treatment in HF is limited it has been found in other cardiovascular studies(149,150) and the insulin group is likely to represent those with either more severe metabolic effects or advanced DM disease.

Mortality risk was also increased in the undiagnosed diabetes group(293) which might relate to the development of insulin resistance as HF progresses, which presents both a hyperglycaemic state and a proxy marker for HF severity(392). This evidence is important as DM is undiagnosed in 1 in 10 stable HF patients(393), but effects of comorbid DM in the diet controlled group have been non-significant as a result of adjustment by other associated risk factors.

Higher chronic glucose levels measured by each 1% rise in HbA1c was associated with increased mortality risk in HF and this was in all patients regardless of their DM status with stronger effects noted in the non-diabetic group(290). This group may reflect undiagnosed DM with the associated insulin resistance and increased HF severity and of interest was the increased mortality risk from lower glycaemia suggesting a U shape relationship between glucose level and risk of mortality. Postulated mechanisms include myocardial energy starvation leading to a dependence on free fatty acid uptake(288) or aggressive glucose lowering drug

therapy which may be hazardous to the HF patient. Increased risk of cardiovascular events from intensive therapy to target normal glycated haemoglobin levels has been found in type 2 DM(290).

These findings suggest that both DM treatment and glycaemic levels stratify the prognostic risk in HF and further investigation into the appropriate cut points for increased mortality risk given the U-shaped relationship with HbA1c is required. Three prognostic models for mortality included the severity of DM albeit two of these were limited to the inclusion of more severe DM by clinical complications(189,259).

In comorbid COPD, risk of death increased as FEV_1 decreased. Effect estimates were non-significant in the moderate COPD severity group but became significant and elevated in the severe group. Whilst the investigation of severity in COPD was limited to one HF study, this link has been found in previous selected populations (147,148) and in cardiovascular disease(394). This evidence suggests that inclusion of 'any COPD' in HF prognosis studies could lead to underestimation of effect in those who have severe COPD and overestimation in those who have less severe disease. A challenge with using FEV_1 for severity assessment is that it is influenced by the HF status. One study showed that lung volume is reduced as a function of HF disease severity(395). Pulmonary oedema causes constriction of the airways and bronchial hyper responsiveness(396) which means that FEV_1 may be partly reflective of congestion rather than obstruction(296). Assessment of COPD should include both FEV_1 to FVC ratio which is independent of HF(397) and FEV_1 assessment once patients are euvolaemic. COPD severity in this systematic review was recorded when patients did not have pulmonary congestion(296). FEV_1 was more predictive than FVC or FEV_1/FVC ratio and independent of ejection fraction. Given the prognostic importance of HF severity, FEV_1 which may represent elements of HF and COPD severity appears a useful prognostic factor and provides a mechanism to determine severity change over time.

This systematic review identified renal severity for a combined sample of 64,257 unselected HF patients. Over a third of HF patients had moderate RD (eGFR 30-59ml/min) and 6% had severe disease (<30ml/min). The mortality risk of any RD (<60ml/min) became significantly differentiated by severity, with higher estimates for more severe RD. This was independent of age, gender, other comorbidities and HF severity. A linear

relationship was observed between eGFR and mortality risk, a finding shown in other studies(304,308,312), but this requires further investigation. eGFR is often categorised and becomes non-significant at different cut points, which in this review was at eGFR ≥60-89ml/min, but the actual level is unknown. Renal severity was also found to differentiate the risk of hospital admission, with the most severe group having a 60% greater risk of admission than the moderate severity group(301). Given the importance of comorbid RD severity it was surprising that only one mortality prognostic model had included it as a factor.

This systematic review selected eGFR as a routinely collected measure of RD. Estimated GFR equations use serum creatinine to estimate true GFR, adjusting for other factors. The Modification of Diet in Renal Disease (MDRD) adjusts for significant non-renal influences such as age, sex, race, and body size and is preferred to other equations such as the Cockcroft-Gault formula(398,399). The renal studies included in this meta-analysis all used the MDRD equation to estimate GFR. Whilst one of the most accurate formulas, it is important to note that this measure has several drawbacks including its tendency to underestimate GFR in the higher ranges (>90ml/min)(400,401). As with other creatinine based equations it also has a tendency to overestimate eGFR (and thus underestimate RD) in more severe disease and older age due to falling creatinine levels that can result from reduced muscle metabolism and increased active tubule secretion(302). Whilst the mean age and ejection fraction was similar across studies, NYHA (class 3 or 4) was more varied and it can't be ruled out that RD may have been underestimated in those with more severe HF disease by symptoms. Whilst this might overestimate the risk associated with less severe renal dysfunction (due to the spurious inclusion of more severe dysfunction in this category) there was still a significant separation between the severity groups and a clear dose-response relationship between renal dysfunction and mortality risk was shown.

Severity change: The importance of comorbidity severity change in HF was limited to comorbid RD studies, where it was shown to have a strong and significant effect on both mortality and hospital admission. The more severe the degree of renal function change the higher the associated mortality risk with a 53% increased risk in the moderate change group (Creatinine of ≥0.3mg/dL increase) rising to 129% in the severe change group (Creatinine of ≥0.5mg/dL increase). This 50% rise in prognostic risk in moderate severity RD change was also

found for the outcome of hospital admissions(156).These effects are larger than the exposure effect estimates of the respective severity groups demonstrating that both disease severity and disease progression are important considerations for risk assessment. Most studies use serial serum creatinine measurements as a marker of change but, due to the exponential relationship between serum creatinine and GFR, this can lead to change from lower baseline creatinine values having a more marked effect on GFR than change from higher baseline values. For this reason change in eGFR is recommended as a measure of renal disease progression(157), with a decline in eGFR of >5ml/min over a year being clinically significant. Also baseline renal function needs to be considered in the assessment of the change. One study found a significant interaction between baseline creatinine and worsening renal function with the higher than median baseline creatinine group being worse off (311). The potential mechanism between baseline RD severity and severity change needs further investigation and within other comorbid diseases for HF populations.

5.4.10 Strengths and limitations

This is the first review to investigate the impact of non-CVD comorbidity and severity on different outcomes, in an unselected HF population. The study samples included a broad range of people with preserved or reduced ejection fraction. All studies were conducted after the introduction of modern HF treatments and so applicable to current practice. Multiple comprehensive databases were searched using validated search strategies as well as grey literature, reference and citation searches and expert contact to identify all good quality studies and the search was validated with previous systematic reviews. Clearly defined inclusion criteria were used and detailed data extraction and quality appraisal of each study was performed to allow for appropriate meta and sensitivity analysis where indicated.

The population HF definition varied across studies and included the use of administration codes which may be subject to misclassification. Most studies did however use a clinical diagnosis of HF (up to 82% of renal studies) based on a combination of signs, symptoms and objective evidence and studies based on self-report or diuretic prescription were not included. Only non-selected HF samples were included but there was a high proportion of hospital-based studies so generalisability to other populations is limited. This however does mean that those studies included in meta-analysis were more homogenous and it is likely from the few

community studies included that the effect estimates maybe even greater in community settings. Comorbid chronic disease exposure was also variably defined and often relied on large administrative databases or patient self-report where the diagnosis and severity of comorbidity cannot be well characterised. This was evident in some of the review studies that identified undiagnosed comorbid groups within their HF samples. The prevalence of DM(293) and COPD(296) increased where subjects underwent physiological testing for the comorbidity. Inclusion of these potentially higher risk subjects in the analyses with a misdiagnosis of no comorbidity is likely to diminish the comorbidity effect estimates(9). However the risk estimates for the different comorbid diseases were significant and there were severity examples for each disease based on physiological markers for comparison. The variation in definitions of both the index HF and comorbid diseases reflects real life practice, where prognosis occurs.

Inherent in systematic reviews of observational studies is the possibility of publication bias. Non English language studies were excluded, which could potentially mean that neutral or negative studies might have been missed(402). The prognostic factor (general) and prognostic model studies were selected on the basis that a non-CVD comorbidity was significantly and independently associated with an outcome. This approach was chosen apriori to scope the studies where comorbidity was included in multivariable models but will mean that studies were excluded where the exposure was found not to be significant in adjusted models. However, these studies were not included in meta-analysis and publication bias was tested and accounted for in each of the main meta-analyses conducted.

Another challenge for meta-analysis of observational studies is the heterogeneity across studies in terms of design, populations, exposures and outcomes. The systematic review methods clearly defined inclusion criteria and data extraction were used to allow comparability of studies and the appropriate selection of studies to combine in meta-analysis. Sensitivity analysis was performed where required. Heterogeneity was reduced for the three main meta-analyses by exclusion of studies and any change to the effect estimates was reported. Any residual unexplained heterogeneity was accounted for by random effects meta-analysis. The risk estimates obtained take account of the variation within and between studies and reflect the average risk for the unselected HF population.

The ability to accurately account for confounding is limited in meta-analysis of observational studies without access to individual patient data. For this reason only the adjusted effect estimates from the chronic disease focus studies were combined, which were more likely to include appropriate confounders. Only the confounders reported in the final adjusted model were extracted in the systematic review and so apparent gaps in adjustment could reflect those factors found not to be significant in individual studies. Whilst unmeasured confounding cannot be ruled out, the effect estimates were consistent across the studies.

Risk of bias was assessed using a recommended and objective instrument(270) which has six domains in order to assess individual and overall risk. Overall assessment can be subjective but this was reduced by a high level of agreement between two reviewers. All included studies in meta-analyses were moderate or low risk and bias was accounted for where necessary in the analyses.

Tables

Table 5.1 Inclusion criteria for articles in the final systematic review

Inclusion Criteria	Specification
Population	Clinical diagnosis of de novo or chronic HF Unselected by cause or ejection fraction Aged 18 years or older Hospital or community based
Exposure	Chronic disease comorbidity or indicator of chronic disease severity Renal studies to include eGFR or creatinine change (after May 2005) Diabetes studies to include HbA1c or fasting glucose COPD studies to include FEV_1
Outcome	All-cause mortality All-cause admission (including readmissions) Health status measure (generic or specific)
Study	Observational studies with more than 30 days follow-up Prognostic factor study (chronic disease focus or general factors that include a chronic disease) Prognostic model study English language

eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; FEV_1, forced expiratory volume in 1 second; pp, percent predicted

Table 5.2 Inclusion criteria for studies used in meta-analysis in the review

Criteria	Description
Study type	Prognostic factor (chronic disease focus)
Chronic disease	Clinically defined or patient reported
Outcome	Mortality, hospital admissions
Population	Clinically defined HF
Follow-up	> 3 months
Sample size	>100 patients
Effect estimate	Adjusted hazard ratios
Risk of bias	Moderate or low

Table 5.3 Chronic diseases included in the systematic review

Chronic disease	Prognostic factor (chronic disease focus)	Prognostic factor (general)	Prognostic model
Diabetes	11	17	7
COPD	5	6	6
Renal dysfunction	17	11	5
Arthritis	1 (Rheumatoid)	2	
Cancer		7	3
Dementia		2	3
Other lung disease		1	1
Liver disease			3

COPD, chronic obstructive pulmonary disease

Table 5.4 Proportion of risk score categories allocated for individual risk domains

Risk domain	Low	Moderate	High
Study participation	35%	59%	6%
Study attrition	69%	28%	3%
Prognostic factor measurement	21%	74%	6%
Outcome Measurement	81%	19%	
Study confounding	23%	65%	12%
Statistical analysis and reporting	29%	55%	16%

Table 5.5 Prevalence of renal dysfunction by severity level

Renal dysfunction (RD) severity level	Studies	HF sample	RD number	Prevalence
Any	8	74,873	36,785 (49%)	39-79%
Mild	1	2,680	1,137 (43%)	43%
Moderate	6	73,968	26,441 (36%)	20-67%
Severe	6	73,968	4,201 (6%)	2-33%

Any RD is defined by eGFR <60ml/min; mild by eGFR 60-89ml/min; moderate by eGFR 30-59ml/min and severe by eGFR <30ml/min compared to the highest study defined category.

Table 5.6 Prevalence of renal severity change

Renal dysfunction (RD) severity change level	Studies	HF sample	RD change number	Prevalence
Moderate	3	21,743	3823 (18%)	11-21%
Severe	2	1,033	239 (23%)	22-25%

Moderate severity change was defined by an increase in creatinine (Cr) of ≥0.3mg/dL from hospital baseline to study defined end point and severe by an increase in Cr of ≥0.5mg/dL.

Table 5.7 Prevalence of chronic diseases in prognostic factor (general studies)

Chronic disease exposure	Prognostic factor (general studies)		Prognostic model studies	
	Studies	Prevalence range	Studies	Prevalence
Diabetes	17	3* to 61%	7	14 to 36%
COPD	6	20 to 47%	6	10 to 31%
Renal dysfunction	11	1[ł] to 70%	5	8 to 25%
Arthritis	2	5* to 29%		
Cancer	7	2 to12%	3	2 to 9%
Dementia	2	10 to 23%	3	5 to 9%
Other lung disease	1	7%	1	9%
Liver disease			3	1 to 3%

COPD, Chronic Obstructive Pulmonary Disease. *Studies with low prevalence used hospital admission codes to define chronic disease. [ł] Low prevalence defined by renal failure

Table 5.8 Renal dysfunction stratified by severity and all-cause hospital admission

Severity group	Risk of hospital admission (adjusted HR)
eGFR <60ml/min	1.18 (1.08-1.29)
eGFR 30-59ml/min	1.16 (1.06-1.27)
<30 ml/min	1.77 (1.16-2.69)

eGFR, estimated glomerular filtration rate

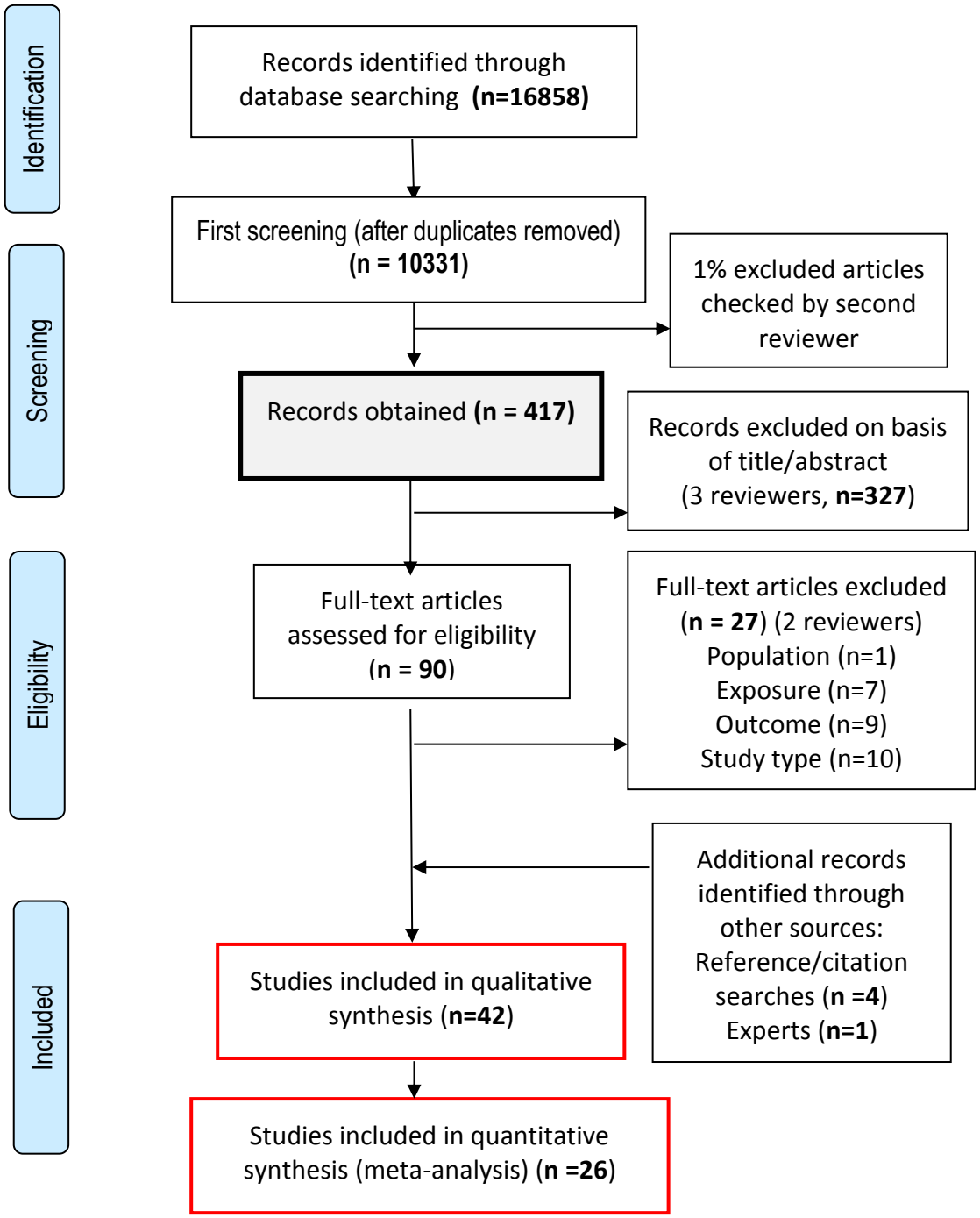
Table 5.9 Other diseases included in prognostic factor (general) and model studies

Disease	Mortality	Hospital admissions
Rheumatoid arthritis (Chronic disease focus) (n=1)	HR 1.89	
Arthritis (n=2)	HR 0.87-1.16	
Dementia (n=5)	OR 2.0, HR 1.44 to 2.02	OR 1.1
Cancer; Lung (n=1)	HR 1.86 to 3.58	
Cancer; Colorectal (n=1)	HR 1.39	
Cancer; Endometrial (n=1)	HR 2.11	
Cancer; metastatic (n=2)	LogOR 4.36, OR1.22	
Cancer; any (n=7)	HR 1.44-2.97, OR 1.85-3.02	OR 1.22
Lung Disease (n=2)	HR 1.37-1.58	
Liver disease (n=3)	HR1.98, OR 5.8	OR 1.29

Figures

Figure 5.1 Screening process

PRISMA Flow Diagram Comorbidity and prognosis in HF populations: A systematic review



From: Moher et al(403).

Figure 5.2 Combined adjusted associations between comorbid DM in HF and all-cause mortality

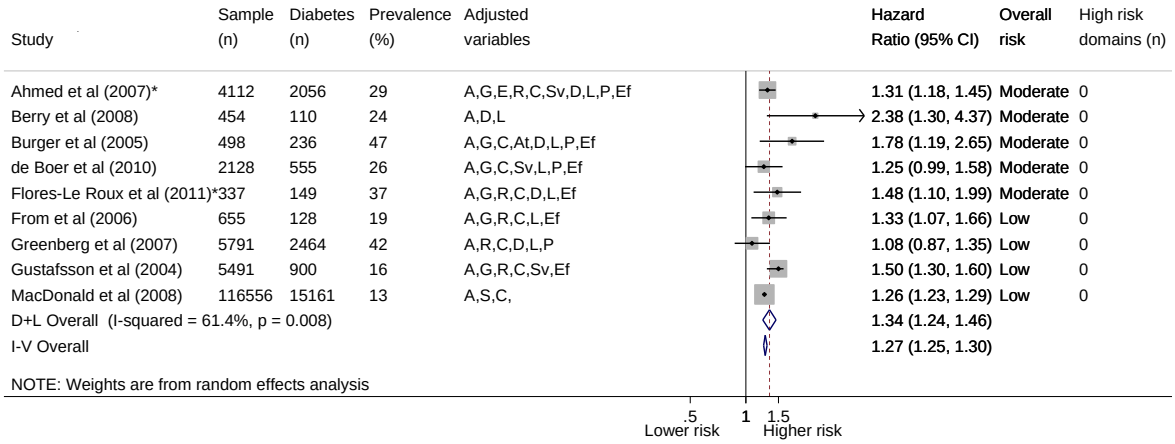
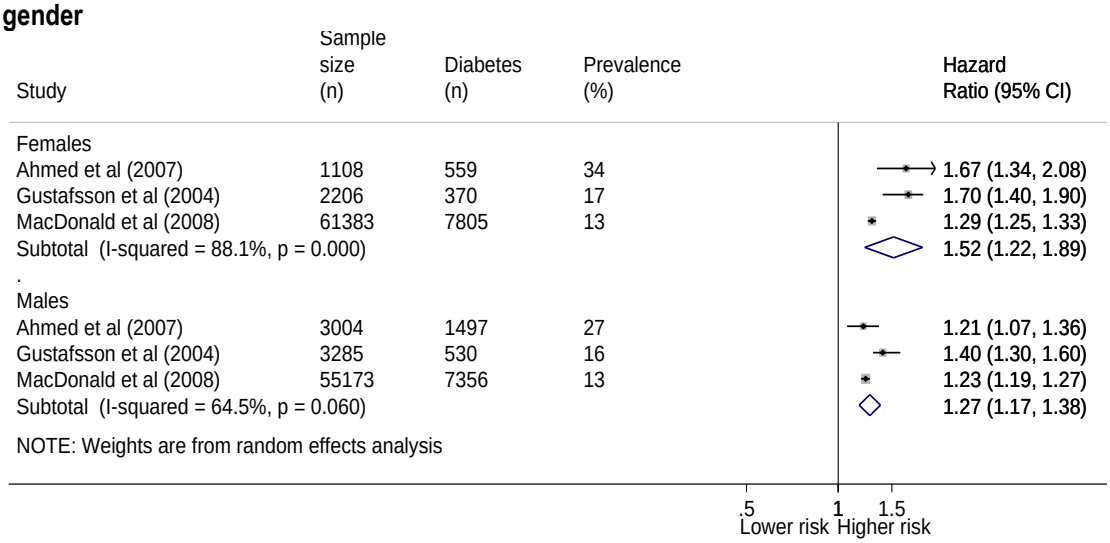


Fig 5.2 * the prevalence for Ahmed et al (2007) is based on the baseline sample before matching on DM status for the analysis. There were 7,788 subjects in the baseline sample and 2,218 (29%) of these had diabetes. Flores-LeRoux had a baseline sample of 400 subjects resulting in a 37% prevalence of clinical DM. The undiagnosed DM group (n=63) were excluded from the current analysis. Adjusted variables: age(A), gender(G), ethnicity(E), social(S), risk factors(R), comorbidities(C), aetiology(At), HF severity(Sv), drugs(D), laboratory(L), physical(P), ejection fraction(Ef)

Figure 5.3 Combined adjusted associations between comorbid DM in HF and all-cause mortality by gender



* the prevalence for Ahmed et al (2007) is based on the baseline sample before matching on diabetes status for the analysis. There were 1,926 women and 5,862 men in the baseline sample and 650 (34%) women and 1,568 (27%) men had diabetes.

Figure 5.4 Comorbid DM severity stratified by treatment type in HF and all-cause mortality

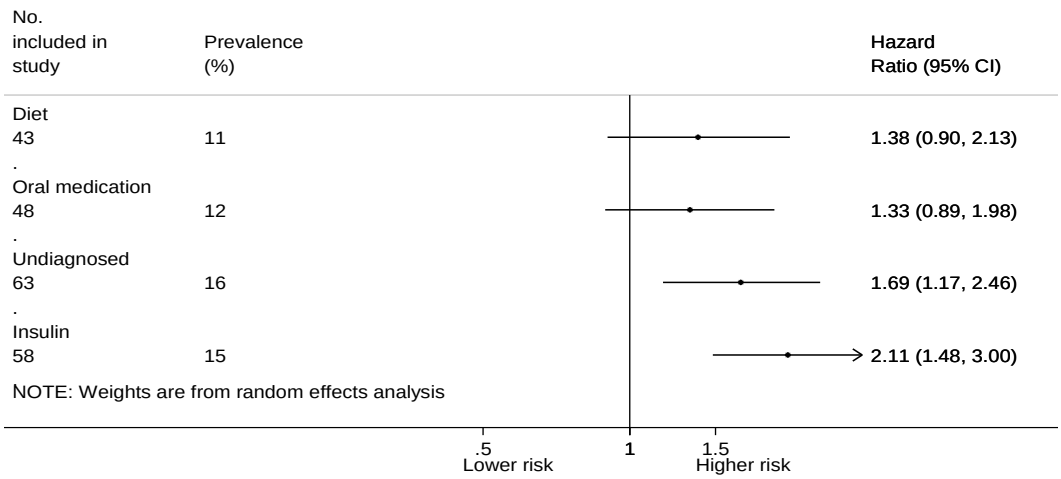


Figure 5.5 Combined adjusted associations between comorbid COPD in HF and all-cause mortality

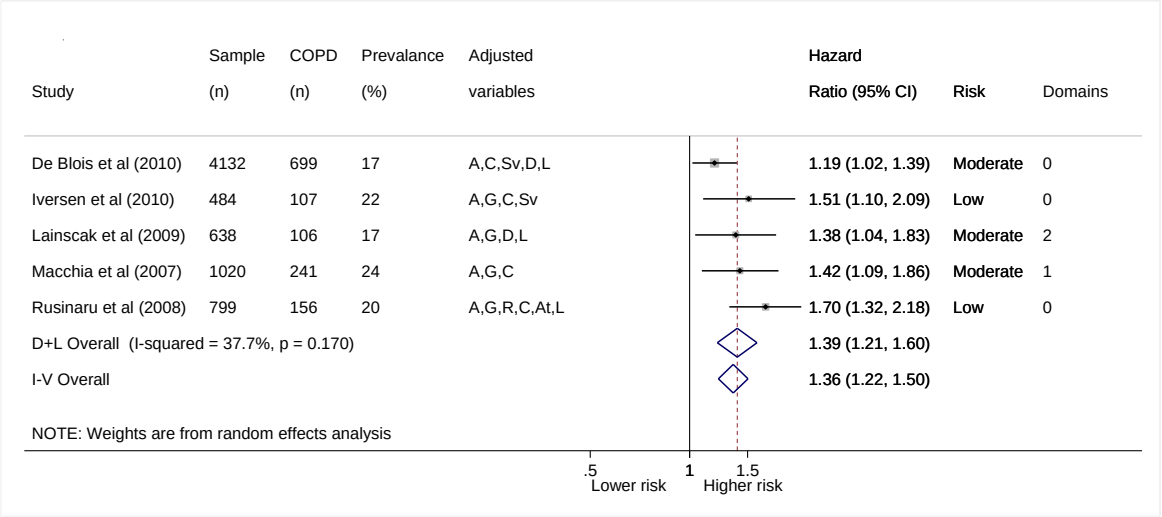


Fig 5.5 Adjusted variables: age(A), gender(G), ethnicity(E), social(S), risk factors(R), comorbidities(C), aetiology(At), HF severity(Sv), drugs(D), laboratory(L), physical(P), ejection fraction(Ef)

Figure 5.6 Comorbid COPD stratified by FEV1 in HF and all-cause mortality

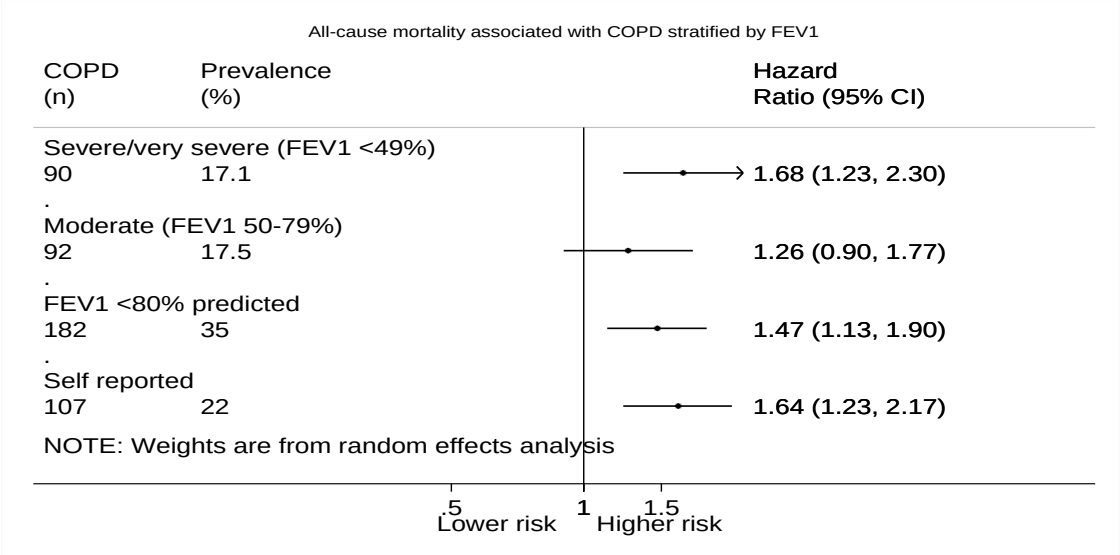


Figure 5.6 there were 532 subjects in the baseline sample. Only 484 had self-reported information. The reference group for FEV groups was FEV≥80% predicted (n=350). The reference group for the self-reported COPD was the non-self-reported COPD group (n=377).

Figure 5.7 Comorbid COPD in HF and all-cause hospital admission

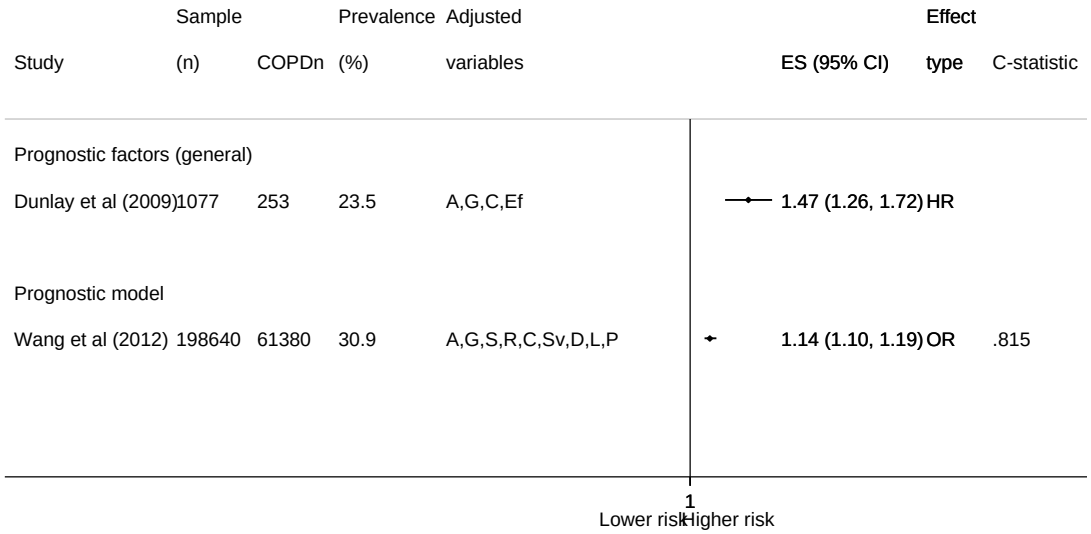


Fig 5.7 Adjusted variables: age(A), gender(G), ethnicity(E), social(S), risk factors(R), comorbidities(C), aetiology(At), HF severity(Sv), drugs(D), laboratory(L), physical(P), ejection fraction(Ef)

Figure 5.8 Combined adjusted associations between Renal Dysfunction and all-cause mortality

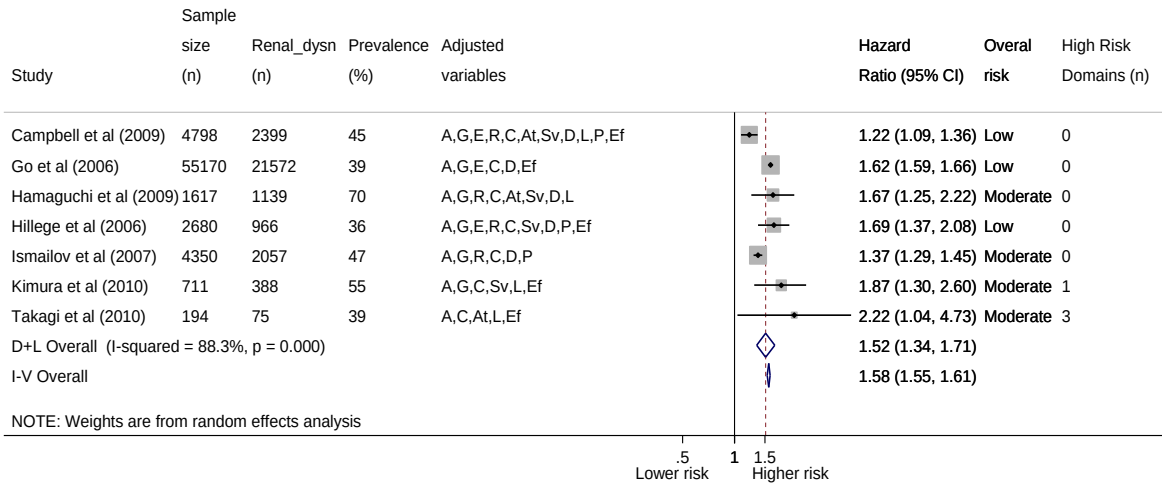
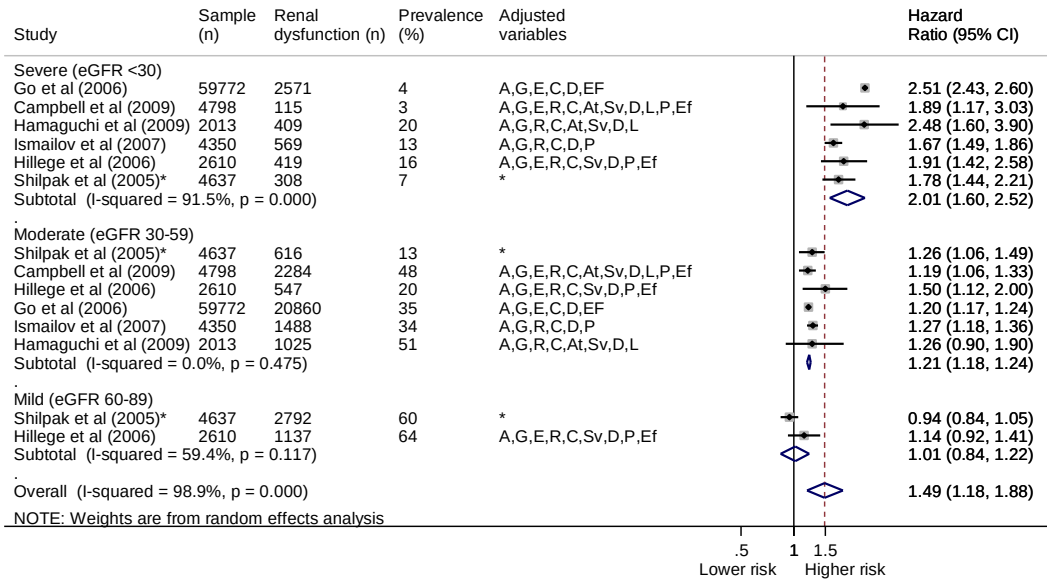


Fig 5.8 Adjusted variables: age(A), gender(G), ethnicity(E), social(S), risk factors(R), comorbidities(C), aetiology(At), HF severity(Sv), drugs(D), laboratory(L), physical(P), ejection fraction(Ef)

Figure 5.9 Combined associations between Renal Dysfunction stratified by severity and all-cause mortality



Test of association between the severity groups and study effect estimates (p<0.001)**

Fig 5.9 Adjusted variables: age(A), gender(G), ethnicity(E), social(S), risk factors(R), comorbidities(C), aetiology(At), HF severity(Sv), drugs(D), laboratory(L), physical(P), ejection fraction(Ef) *from prior systematic review(145) **Test of association between the severity subgroups and the study effect estimates was performed using random effects meta-regression with Monte Carlo permutations to calculate the *P* value. The reference group for the studies was eGFR >60. The two studies plot that investigated mild eGFR (60-89) used eGFR≥90 as a reference group.

Figure 5.10 All-cause mortality in HF by study defined upper eGFR severity category limit

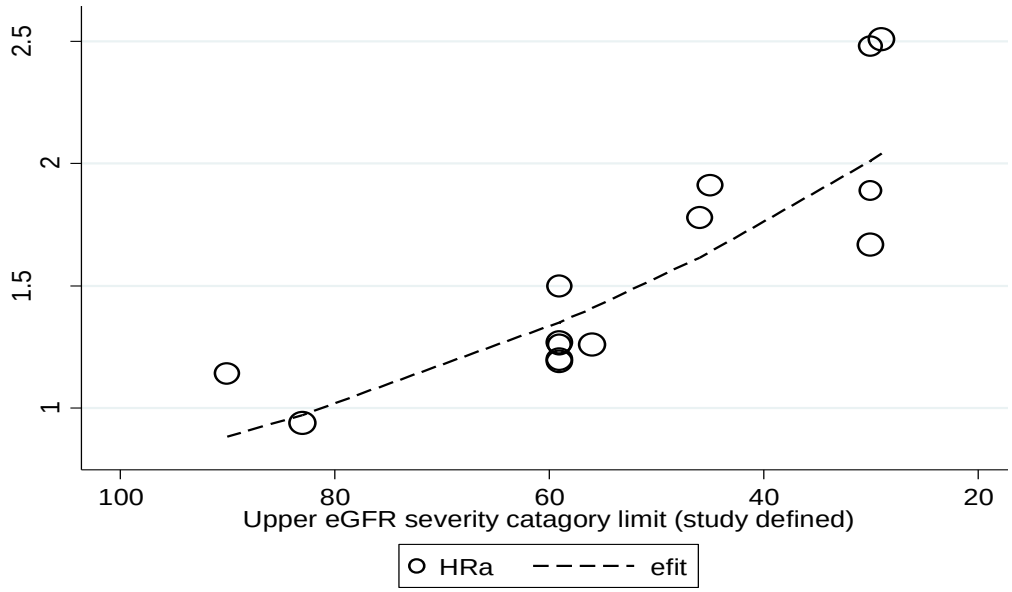


Figure 5.10 The reference group for the studies was eGFR >60. The two studies at the bottom left of the plot that investigated mild eGFR (60-89) used eGFR≥90 as a reference group.

Figure 5.11 All-cause mortality risk in HF by eGFR mls/min

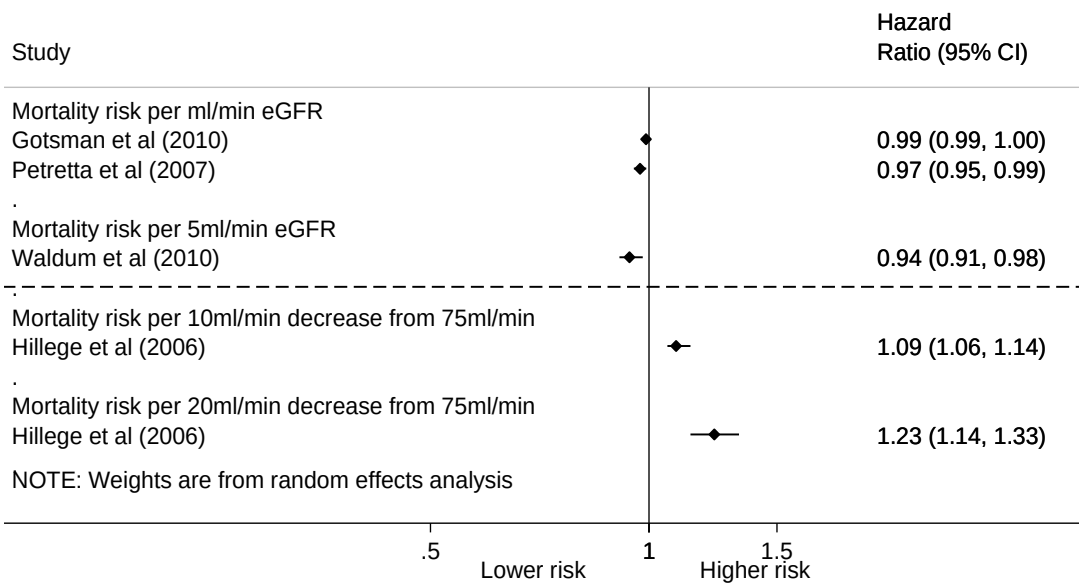


Figure 5.12 Combined associations between renal function change stratified by severity and all-cause mortality

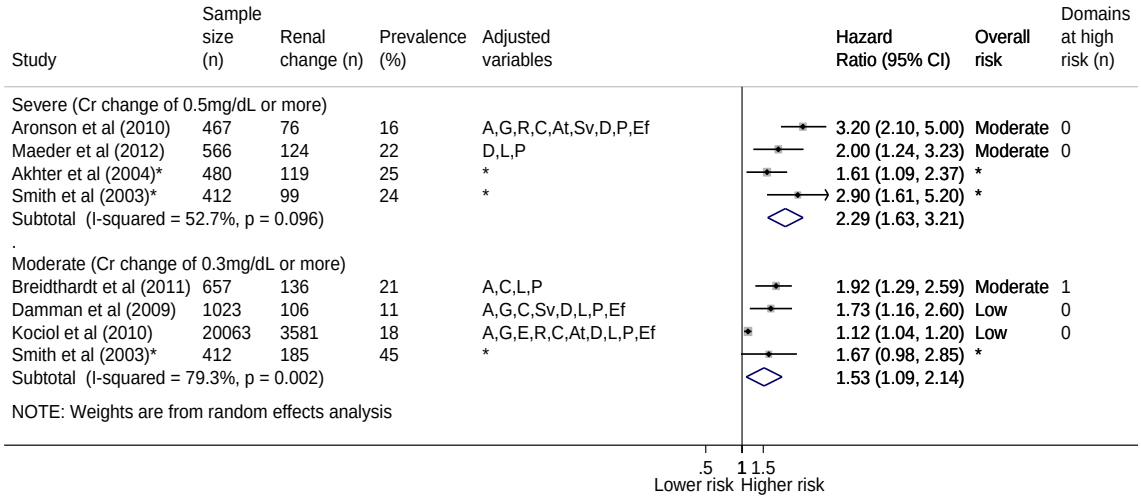
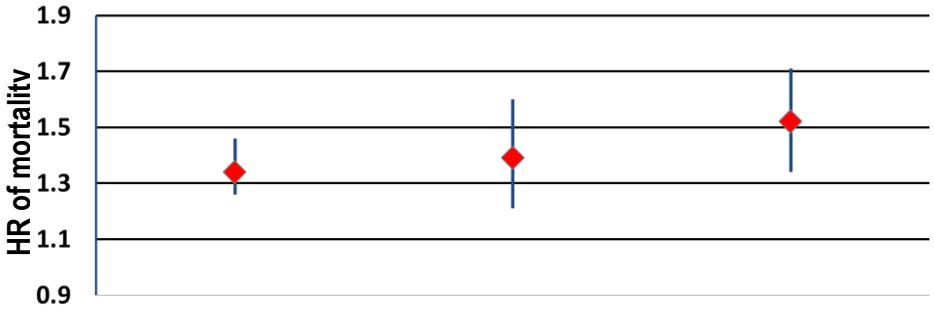


Fig 5.12 Adjusted variables: age(A), gender(G), ethnicity(E), social(S), risk factors(R), comorbidities(C), aetiology(At), HF severity(Sv), drugs(D), laboratory(L), physical(P), ejection fraction(Ef) *From previous systematic review(157). Renal function change was measured as the difference between the admission creatinine and the study defined end point.

Figure 5.13 Meta-analysis summary of the association of the three main diseases with mortality



Comorbidity	Diabetes	COPD	Renal dysfunction (eGFR<40%)
Studies (n)	9	5	7
Sample (n)	135,402	7,121	69,520
Prevalence (%)	16	18	41

Phase 2 and 3: Database studies

Chapter 6 Time-dependent comorbidity severity and change and outcomes: introduction and methods

This chapter will introduce the series of Clinical Practice Research Datalink (CPRD) studies used to investigate the main hypotheses focusing on the 3 specific common non-cardiovascular disease (CVD) comorbidities that were identified from the systematic review (SR): diabetes mellitus (DM), chronic obstructive pulmonary disease (COPD) and Chronic Kidney Disease (CKD). The study objectives and hypotheses will be presented followed by a detailed discussion of the methods including sample selection, the framework for measuring comorbidity severity and change using routinely collected data, outcome and covariate measurement and the statistical analysis framework.

6.1 Introduction

From the 68 heart failure (HF) prognosis studies in the systematic review (Chapter 5) that included non-CVD comorbidities, only one had focused on a community setting. Most studies had also investigated mortality rather than other outcomes that span the life course of HF disease which are important to patients and policy makers such as health-related quality-of-life (HR-QoL) or hospital admissions. For the outcome of hospital admission, the few HF comorbidity prognosis studies had only focused on the single comorbid example of RD, and there has been little structured investigation of comorbidity status and definitions in relation to both hospital admission and mortality. The CPRD provided easily accessible national population data for linking HF, comorbidity and mortality and hospital admission data through Hospital Episodes Statistics and mortality data through Office of National Statistics data. So for the purposes of the thesis, all following analyses in Chapters 7 to 10 focus on comorbid DM, COPD and RD in HF and their effect on the outcomes of first hospital admission after HF diagnosis and mortality. The outcome of HRQoL therefore lies outside the scope of the thesis, but initial analysis for this outcome in a separate analysis is presented as a supplementary published paper (E-Appendix E18).

6.2 Main study objectives, hypotheses and questions

The main study which focuses on Phases 2 and 3 of the thesis is in three parts and uses the example of DM, COPD and CKD comorbidity in HF.

Part 1: Non-CVD comorbidity measures in an incident HF general practice cohort sample by all-cause mortality and first hospital admission.

Part 1 of the CPRD analyses forms Chapters 7 and 8 and the following objectives, hypotheses and questions are addressed:

***Chapter 7 objectives**: HF patient characteristics by all-cause mortality*

(i) To describe the *baseline* characteristics and comorbidity status for three non-CVD comorbidities in a national cohort sample of incident HF patients aged 40 years and over for a 10-year time period. Characteristics and comorbidity measures will be described in the total sample and in those who died compared with those who remained alive over a maximum 12-year follow up period.

(ii) To compare the *time-dependent* characteristics and comorbidity status, severity and change measures between cases and controls in a cohort sample of incident HF patients where cases were matched to controls on calendar and follow-up time.

***Chapter 8 objectives**: HF patient characteristics by first hospital admission*

(i) To describe the *baseline* characteristics and comorbidity status for three non-CVD comorbidities of a sub-sample of the cohort of incident HF patients aged 40 years and over, that were linked to Hospital Episode Statistics (HES) in a 10-year time period. Characteristics and comorbidity measures will be described in the HF sample and in those who had a first hospital admission compared with those who did not over a maximum 12-year follow up period.

(ii) To compare the *time-dependent* characteristics and comorbidity status, severity and change measures between cases and controls in a sub-sample cohort of HF patients with linked HES data, where cases were matched to controls on calendar and follow-up time.

Questions addressed in Chapters 7 and 8 were:

(i) What is the prevalence of COPD, DM and CKD in the incident HF cohort at baseline?

(ii) Is the prevalence of COPD, DM and CKD at baseline significantly different in HF patients with the two primary outcomes of all-cause mortality and first hospital admission and those without these outcomes?

(iii) What is the prevalence of COPD, DM and CKD defined by diagnostic status, severity and recent change measured before the two primary outcomes compared to those without these outcomes measured at the same time?

Hypotheses tested in Chapters 7 and 8 were:

(i) Baseline comorbid status COPD, DM and CKD are associated with increased all-cause mortality or increased risk of first hospital admission in HF.

(ii) Time-dependent measures of recent comorbidity status, severity and change are associated with increased all-cause mortality or increased risk of first hospital admission in HF.

Part 2: Non-CVD comorbidity prognostic factors in an incident HF general practice cohort sample: strength of associations with all-cause mortality and first hospital admission

Part 2 of the CPRD analyses in the same general practice HF population aged 40 years and over forms three chapters as follows:

- strength of association with all-cause mortality (Chapter 9),
- strength of association with first hospital admission (Chapter 10) and
- strength of association and interactions (Chapter 11).

Chapter 9, 10 and 11 objectives were:

(i) To investigate the potential confounding factors for the unadjusted associations between the comorbidity measures and all-cause mortality and first hospital admissions identified in Part 1.

(ii) To investigate whether time-dependent non-CVD comorbidity status, severity and recent change exposures independently and significantly increase the risk of all-cause mortality and first hospital admission compared to non-comorbid groups.

(iii) To investigate whether the strength of associations or '*estimate of effects'* (hereafter referred to as 'effects') of non-CVD comorbidities on all-cause mortality or first hospital admission in HF are significantly stratified by measures of comorbidity severity or change.

(iv) To investigate potential first order interactions between pairs of the 3 main comorbidities and between the comorbidities and other key study variables.

Chapter 9, 10 and 11 questions addressed were:

Confounding:

(i) What HF patient characteristics are significantly different in those with and without each of the comorbidities?

(ii) Are the unadjusted effects of comorbidity on all-cause mortality and first hospital admission altered or modified by the adjustment by any patient characteristic?

Associations:

(iii) What is the effect of comorbidities on all-cause mortality and first hospital admission that develop before or after index HF?

(iv) Does the effect of comorbidity that develops before HF differ from the effect of comorbidity that develops after HF on all-cause mortality and first hospital admission?

(v) Does recent comorbidity severity and severity change significantly and independently increase the risk of all-cause mortality and first hospital admission in HF?

(vi) Are the effects of comorbidities on all-cause mortality and first hospital admission in HF significantly stratified by measures of their severity and recent change?

(vii) Do the effects of comorbidity prognostic factors on all-cause mortality in HF differ from their effects on first hospital admission?

Interactions:

(viii) Is there evidence of statistical interaction between two comorbid diseases?

(ix) Is the effect of two comorbid diseases on outcomes in HF different than the sum of their independent effects?

(x) Is the effect of comorbidity status on outcomes in HF modified by HF severity groups defined by age?

(xi) Is the effect of comorbidity status on outcomes in HF modified by other key patient characteristics?

Chapter 9, 10 and 11 hypotheses tested were:

(i) Time-dependent measures of non-CVD comorbidity status, severity and change will independently and significantly increase the risk of all-cause mortality and first hospital admission compared to non-comorbid groups.

(ii) The independent effect of HF comorbidity status on all-cause mortality and first hospital admission compared to non-comorbid HF groups will be significantly stratified by measures of comorbidity severity and recent change.

(iii) The effect of having two non-CVD comorbidities in HF at the same time will differ from the sum of their independent effects on all-cause mortality and first hospital admission.

(iv) The effects of non-CVD comorbidities on all-cause mortality and first hospital admission will significantly differ by HF severity groups defined by age.

Part 3: Non-CVD comorbidity prognostic models; comparison of multivariable models for all-cause mortality and hospital admission using different comorbidity measures.

Part 3 focuses on Phase 3 with the following objective (Chapter 12):

To investigate the contribution of non-CVD comorbidity severity and recent change measures to pre-defined multivariable models for all-cause mortality or first hospital admission in the general HF population.

The question addressed was:

Does the statistical fit of a pre-specified multivariable model that includes comorbidity diagnostic status measures improve when these measures are replaced by comorbidity severity and change measures?

The hypothesis tested was:

Replacing comorbidity diagnostic status measures with chronic disease comorbidity severity and severity change measures will improve the fit of pre-specified HF multivariable models for the outcomes of all-cause mortality and first hospital admission.

6.3 Methods

This section describes in detail the study design, data, setting and sample selection, the comorbidity exposure measures, the outcome measures and the covariate measures used in each of the three analytic parts.

Part 1 – Descriptive studies of the:

a) baseline cohort of HF general practice population patients in the mortality sample and the hospital admission sub-sample, and

b) nested case-control samples of the baseline populations to test the hypotheses that

(i) baseline status COPD, DM and CKD are associated with all-cause mortality and first hospital admission and

(ii) time-dependent measures of comorbidity status, comorbidity severity and change are associated with all-cause mortality and first hospital admission.

Part 2 – Nested case-control studies to investigate the main hypotheses that the effects of time-dependent measures of comorbidities on mortality and first hospital admission are significantly stratified by measures of their severity and change.

Part 3 – Nested case-control studies to investigate the hypothesis that replacing comorbidity diagnostic status measures with time-dependent measures of comorbidity severity or change would improve the fit of HF multivariable models for mortality and first hospital admission outcomes.

6.3.1 Study design and setting

Using a nested case-control approach to analysis, a historical cohort of incident HF patients from the CPRD was used to investigate the effect of non-CVD comorbidity severity and severity change measures on all-cause mortality and first hospital admission compared to non-comorbid HF groups. The nested case control study design used risk set sampling of controls to estimate unbiased rate ratios for the comorbidity exposures. On the case index date, controls were randomly sampled from the HF cohort that had not yet experienced the outcome and remained at risk of the event. Controls were also matched within one month of the case HF index (baseline) date. The study design is discussed in detail in Chapter 4.3.2 and shown in Figure 4.1.

CPRD clinical data (GOLD) (discussed in detail in Chapter 4) was extracted from the database of all eligible patients who had had a first consultation code for HF between 1st January 2002 and 1st March 2012 from the most recent data capture available on January 1st 2014. Subjects were followed until their date of death from any cause or their date of first hospital admission (depending on the outcome under study), the date they transferred out of the practice (TOD), the date the practice stopped contributing data to the CPRD or end of the study period (January 1st 2014), whichever occurred first. Linked datasets were also extracted on hospital episode statistics (HES), mortality data from the Office for National Statistics (ONS) and Index Multiple Deprivation (IMD), 2007. Access to the anonymised CPRD data was provided under Keele licence and following Independent Scientific Advisory Committee (ISAC) approval (see E-Appendix D2 for the CPRD scientific protocol).

6.3.2 Study sample

Patients aged 40 years and over were eligible for inclusion if they had a first HF consultation code applied to their general practice clinical or referral records between 1st January 2002 and 1st March 2012. The age criterion was chosen as earlier presentations of HF are unusual and indicate more rare causes such as genetic disorders(246). Clinical data is coded within the CPRD using the Read code classification which is a hierarchy using three tiers; diagnostic, process and medication codes. The code set used to select the HF sample was based on diagnostic codes from Read chapter G58(404). An additional search was done using the CPRD medical code dictionary browser for any additional codes related to the clinical terms 'ventricular', 'cardiac' or 'heart', in combination with 'failure'. All but one process code ('HF confirmed') was eliminated from this latter search as they represented ongoing care or symptoms in a prevalent cohort rather than the index date of HF. The code set was validated by HF specialists and against previous literature(405,406) (E-Appendix A14 for the selection code set).

Exclusion criteria were applied to the initial sample of 79,629 HF patients. First, patients were removed if their practice had less than 3 years of up to standard (UTS) clinical data (CPRD-defined minimum quality standards for data recording based on completeness, consistency and plausibility) or where the patient's current registration date (CRD) was less than 3 years prior to the HF index date. Second, patients were removed on the basis of implausible data recording. This constituted those where their CRD, UTS, TOD or HF index date was greater than one month after their death date or where their CPRD and Office National statistics death date differed by greater than 3 months. Third, patients with unexplained gaps in their clinical record were excluded. These comprised those who had a CRD or UTS date after their index HF date, their TOD preceded their HF index date or there were recorded data gaps that were not explained by the difference between their first and CRD. Fourth, patients that appeared twice under different patient id numbers were removed. These patients move from one CPRD practice to another and are given a new patient ID number. They were identified through their unique hospital identifier (Figure 6.1 and E-Appendix A15 for detail of the exclusions). There were 50,114 HF patients in the study sample following this process for the mortality outcome and a sub-sample of 30,061 patients linked to HES and used for the hospital admission outcome.

6.3.3 Matched sampling

In order to measure time-dependent exposures of comorbidity severity and severity change, over 12-years of follow-up, matched samples were created from the two baseline samples (mortality sample and HES linked sub-sample).

Mortality: For each mortality case, 4 controls were randomly sampled from the cohort members in the risk sets defined by the case, after matching on HF index date (+/- 1 month) and duration of follow-up. This number of controls is adequate to produce similar parameter estimates to that obtained using full cohort data in a Cox regression model and provides 80% of the statistical efficiency of using an infinite number of controls(220). Added benefit from a greater number of controls is only likely where the exposure status of controls is likely to be low or there are very few matched sets(407). The risk sets defined by the case included all individuals still at risk of the outcome on the date of the case event. Controls were therefore alive, active in practice and event free on the match date. The match date for the controls was defined by the same duration of follow-up to the case. Using this approach controls are eligible to be selected multiple times as a control and later as a case, approximating the situation in Cox-regression where every case is compared to all controls in its risk-set(214). This matching process resulted in 133,645 observations comprising 26,729 cases and 106,916 controls. Over 70% of subjects were used less than 4 times in the analyses (E-Appendix E8 for the frequency of subject use in the matched set and E-Appendix A16 for comparison of the pre and post matched sample).

Hospital admission: For each first hospital admission case, initially 8 controls were randomly selected among the cohort members in the risk sets defined by the case, after matching on HF index date (+/- 1 month) and duration of follow-up. Controls were alive, active in practice, event free and still at risk of first hospital admission on the date of the case event. Again, the date resulting in the same duration of follow-up for the cases and controls defined the match date for the controls. Due to the high number of first hospital admission events within 3 and 6 months of the HF index date, controls were only retained in the match sets if they did not experience an event within three months of their selection as a control. This approach is analogous with the wash-out period in a case-crossover study(408) and provided a suitable period where cases were considered different from the controls on their event status, to allow for the comparison of comorbidity

exposures. The matching to 8 controls allowed for the removal of unsuitable controls, leaving each case matched to a varying number of up to 4 controls. This matching resulted in 110,789 observations comprising 24,339 cases and 86,916 controls. Over 70% of subjects were used up to 4 times only in the analysis (E-Appendix E8 for the frequency of subject use in the matched set and E-Appendix A17 for comparison of the pre and post matched sample).

6.3.4 Measure of comorbidity exposures

Three non-CVD comorbidities identified by the systematic review were selected to test the study hypotheses on the basis of (i) their high prevalence (COPD 18%, DM 16%, RD 41%) in the general practice population of HF(239) and (ii) their routine monitoring using physiological measures or their treatment with regular drug prescriptions in order to measure comorbidity severity or change over time.

The comorbidity cohort selections were based on clinical codes or physiological measures (CKD) or a combination of clinical codes and drug measures (COPD and DM). For each comorbid disease, subjects were selected using search strategies that were validated by appropriate clinical experts and prior research validation where available. Differentiation was made between comorbidities that had developed before or after HF. The 'before HF' comorbidity cohorts were identified through a combination of diagnostic and process Read codes that indicated either a comorbidity diagnosis consultation (comorbidity index event) or an ongoing care consultation (comorbidity prevalence event) applied in the 3 years of UTS clinical data prior to their HF index date. The after HF comorbidity cohorts were identified by a first consultation 'index' code that indicated a comorbidity diagnosis event (using diagnostic codes and a limited set of process codes that were specific to a comorbidity index event) applied subsequent to their HF index date. Those with a comorbidity prevalence event code applied after HF but no comorbidity prevalence or index event code applied in the 3 years of UTS prior to HF in their clinical record were excluded from the comorbidity cohorts. These were rare but meant that the placement of the comorbidity index date could not be verified and given the routine recording and monitoring of chronic diseases, including these subjects would increase the risk of misclassification bias(409).

The measures for each of the three comorbid diseases will now be discussed. The extraction and measurement of the clinical data using time-dependent windows followed a uniform approach for all three comorbid diseases and will be detailed for both outcomes at the end.

6.3.4.1 Chronic obstructive airways disease

Three categories of COPD measures were included; diagnostic status, severity and severity change.

COPD diagnostic status: The COPD 'before HF' cohort was based on a combination of at least one COPD prevalence or index event code applied in the three year time-window before the HF index date and at least one COPD related drug prescription in the same time window. The 'after HF' cohort was based on those (i) without a COPD prevalence or index event code applied in the 3-year before HF, (ii) with at least one COPD index event code applied at any time after HF index date but before the match date, AND (iii) with at least one COPD related drug prescription in the same time window after HF. A COPD related drug prescription was defined as any COPD related drug indicated in the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines(410).

There is no single diagnostic test for COPD which relies on clinical judgement (history, symptoms, frequency of recent exacerbations) combined with the presence of persistent airflow obstruction using spirometry(410). However diagnosis is complicated by non-specific symptoms such as breathlessness in COPD in the presence of HF and spirometry is not recorded for up to a third of patients with both diseases(411). Where recorded, spirometry can also be misleading in HF where restriction caused by cardiomegaly or pulmonary congestion can produce a false obstructive pattern. These issues combined can lead to both over and under diagnosis of COPD.

Use of clinical codes for discriminating between those with and without COPD or for the positive prediction of COPD has been validated in two recent studies. The first study was based on a combination of terms (chronic bronchitis, emphysema, chronic obstructive asthma and chronic airway obstruction) which excluded 'none specified bronchitis' as a non-specific term. This study found that using multiple ICD-9 codes applied in the

clinical record made little difference to the discriminative ability of the codes to correctly identify COPD however the use of a combination of ICD-9 codes with prescribed COPD related drugs improved discrimination(412).

A more recent study to validate different approaches to identifying subjects with COPD within the CPRD found that the positive predictive value (PPV) of using a COPD specific code set yielded a PPV of 86.5%. When drugs or spirometry were added to the COPD codes, the PPV of COPD did not improve(413). However there was improvement in the PPV with the addition of drugs or spirometry to clinical codes for those with more mild COPD disease defined by forced expiration volume in one second [FEV_1] level. The validated set of codes used, excluded bronchitis codes and included only a narrow set of COPD specific codes which promotes specificity but not sensitivity.

COPD was included in the Quality and Outcomes Framework (QOF) in 2004 in England and Wales which introduced more codes to identify COPD and promoted the recording of spirometry data. Following the introduction of QOF and the NICE guideline for COPD the prevalence of recorded COPD increased by 14.4% from 2003 to 2005 and the presence of recorded spirometry data in people with COPD increased from 18% to 62% over the same time-frame(411). The positive predictive value of using COPD codes to identify COPD also improved after 2008. However non-specific bronchitis codes and symptom codes to record COPD were more likely to be used before 2008 rather than after 2008 which required consideration in the present study. A further consideration for the present study was the reduction in diagnostic accuracy of COPD when using clinical codes combined with spirometry in those who also had cardiovascular disease(413).

COPD search strategy: A search strategy based on broad terms was used to identify all possible relevant COPD codes applied in the patients clinical or referral record (E-Appendix E9a). The results of the search strategy were reviewed by a respiratory consultant and cross-referenced with the validated codes in the previous CPRD study and the codes included in QOF. None specified or acute bronchitis codes were excluded to reduce misclassification bias. Out of a possible 230 codes, 44 prevalence codes and 54 index codes were agreed, which comprised diagnostic and process codes for COPD. A total of 47 codes were

included in QOF and 28 had been previously validated. The remaining new codes, not previously validated or included in QOF were specific to COPD (E-Appendix E10).

For the 'before and after' HF COPD cohorts it was decided to combine at least one clinical code with at least one COPD drug prescription in the same time period. The intention was to reduce the risk of misclassification bias but optimise the number of COPD subjects available for analysis. This was chosen over spirometry for validation of the COPD code given the relatively low recording of spirometry data in COPD patients even after 2005(411,413). However, by not including spirometry in the definition, the ability to stage COPD according to GOLD stages was not possible for all patients with COPD included in the analysis.

CPRD product codes for drugs prescribed were identified using the database browser. A search identified any COPD related drug as defined by the GOLD guidelines(410) using appropriate British National Formulary (BNF)(414) codes (Table 6.1). Additional searches using truncated drug names identified those in the CPRD that are not linked to BNF codes, yielding the most relevant COPD drug codes available.

COPD severity using physiological data: GOLD guidelines recommend the use of FEV_1 to measure the severity of lung function in COPD. This is used when there is evidence of obstruction defined by a FEV_1/FVC (forced vital capacity) ratio of <70%. Four severity categories of FEV_1 used to guide clinical care(410) (Table 6.2) were included in this study. The FEV_1/FVC ratio could not be used as it is not routinely recorded. FEV_1 is measured as a percentage of the predicted value based on gender, height and age. A complication in HF is that a reduction in FEV_1 of up to 20% predicted is frequently observed independently of COPD leading to an overestimation of the severity of obstruction in HF patients with concomitant COPD(342,343). Up to quarter of HF patients and a half of severe HF patients without COPD have an FEV_1 of <80%(415). Whilst this creates potential difficulty in pin pointing the underlying cause of reduced FEV_1, the lack of alternative measures mean that the GOLD classification was applied in the HF population.

COPD severity using drug data: An alternative measure of COPD severity using drugs prescribed was also developed. Drugs are routinely and automatically recorded through electronic prescribing and so provided the

opportunity to stage all COPD patients based on drug measures. GOLD guidelines recommend a step up approach to prescribing in COPD that increases with the patient's symptoms and risk of exacerbations(410). This staging ranges from short acting anticholinergics or beta2-antagonists as required to mono, dual or triple therapy using a combination of long acting anticholinergics, beta2-antagonists and inhaled corticosteroids. An additional, most severe stage was added which included oral steroids and long-term oxygen therapy. These latter two drugs are most commonly used following acute exacerbations or in most severe disease and do not form part of routine prescribing. The severity category for each COPD patient was identified by at least one prescription of a combination of different drugs in a specified time-window prior to the match date (Section 6.3.4.5). To do this, all drugs included in the original drug code set were given a number according to their drug type and algorithms for identifying the correct drug severity category within a specified time window prior to the match date were developed (Table 6.3 for the COPD drug severity framework and E-Appendix A18 for the CPRD drug severity extraction algorithms). Three drug-based severity classifications were defined as comorbidity measures (Table 6.4).

COPD severity change using physiological data: Two important questions for defining the measurement of severity change were how much and over how long? There appeared to be no prior studies which had measured COPD change by FEV_1 status in HF populations but there were other definitions. In the Lung Health Study rapid decline in smokers was defined as a loss of >3% FEV_1 over a year but this group did not necessarily have COPD or HF(416). The mean rate of absolute decline in this study was 4.1% of predicted FEV_1. In a group of patients with idiopathic pulmonary fibrosis with or without emphysema a relative decline in FEV_1 of 5,10,15 and 20% over a year was predictive of mortality(417). As there is no clear definition of COPD severity change, for the thesis analyses a range of new measures were defined (Table 6.5).

It was hypothesised that recent comorbidity severity change that occurred prior to a hospital admission would occur over a shorter time period, so change in these measures over 6 months was used for this outcome (Section 6.3.4.5).

COPD severity change using drug data: Two classifications of COPD drug severity change were considered (Table 6.6). In the first classification, for simplicity, it was decided to collapse the possible number of drug change categories into two mutually exclusive discrete categories. This also allowed for the comparison of those with COPD whose medication remained stable or improved with those with worsening drug severity indicated by a step up in drug therapy. The second classification focused on a change in COPD drugs, prior to an outcome, to the most severe drug groups.

6.3.4.2 Diabetes mellitus

Three categories of DM comorbidity measures were included: diagnostic status, severity and severity change.

Diabetes diagnostic status: The DM 'before HF cohort' was based on at least one DM prevalence or index event code or at least one DM prescription applied in the three year time-window before the HF index date. The 'after HF cohort' was based on those (i) without a DM prevalence or index event code or prescription applied in their three years before the HF index date AND (ii) at least one DM index event code or prescription applied at any time after HF index date but before the match date. A DM related prescription was defined as any DM related drug for blood glucose control indicated in the National Institute for Health and Care Excellence Guidelines(418,419).

Diabetes mellitus (DM) is a progressive disease characterised by chronic hyperglycaemia and disturbances of carbohydrate, fat and protein metabolism caused by defects in insulin production or action or both(420). In Type 1 diabetes, which is normally diagnosed in younger age groups, the pancreas is no longer able to produce insulin and requires insulin replacement. In type 2 diabetes, which accounts for approximately 85% of diabetes in England, not enough insulin is produced and there may also be resistance to any insulin produced. This is usually diagnosed in older age groups, is less obvious to detect than type 1 diabetes and treatment ranges from strict diet and weight control to medicines for glucose control(421). The diagnosis of diabetes is made through laboratory glucose measurements. This includes a single measure in conjunction with classic symptoms or a repeated measure(418). The presence of a raised fasting glucose or raised glycated haemoglobin (HbA1c) at any time in the blood is indicative of diabetes(420). These tests provide a

specific and objective diagnostic measure that is not reliant on clinical judgement or heterogeneous signs and symptoms, thus reducing the risk of misclassification bias.

Diabetes search strategy: A search strategy was devised based on the Read code 'Diabetes Mellitus' (C10) as well as a broad search for any diabetes related term to identify all possible relevant codes applied in the patients clinical or referral record (E-Appendix E9b). The results of the search strategy were reviewed by a general practitioner and cross referenced with other previous CPRD search strategies(422) and the codes included in QOF. The current strategy was combined by a search on all product codes for diabetes related drugs to identify drug defined diabetes subjects. Out of a possible 646 codes, 346 index codes and 208 prevalence codes were agreed, these comprised of diagnostic and process codes for DM and 125 were included in QOF and all were used in prior studies (E-Appendix E11).

CPRD product codes for DM drugs were identified using the CPRD product browser and a search was performed using Chapter 0601 of the British National Formulary (BNF). Using the product browser an additional search was performed using truncated drug names of some of the main drugs identified through the first search to yield the most CPRD product codes available.

Diabetes severity: There is no standard measure of severity used for DM. Studies have previously used diabetes related complications as a measure of severity which has indicated an increased risk of death(189,259). This was not used in this study given the varied and complex code set required to identify this group and the use of time-windows within the present study to measure severity and change meant that the codes that are not routinely repeated in clinical coding might be missed. So two approaches to define severity were a physiological measure using glycosylated haemoglobin (HbA1c) and a drug based measure.

Diabetes severity using physiological data: HbA1c reflects the blood glucose levels over the preceding 2 to 3 months and is used as an index of mean glycaemia(420) for diagnosis, monitoring and targeting of treatment for diabetes. Whilst targets should be individualised for patients, national guidelines state levels below 7.5% and 6.5% for adults with type 1 and type 2 diabetes respectively are appropriate targets for risk reduction,

whilst risk of hypoglycaemia may increase if levels reach 6.1% or lower(418,419). This measure is preferred over blood glucose as it does not require the patient to fast and is not as susceptible to daily fluctuations. Although HbA1c is more a marker of glucose stability rather than DM severity, it has been found to have similar associations to the development of retinopathy as glucose(420). In the general population and in people with DM poor glycaemic control has been associated with adverse outcomes including microvascular and neuropathic complications(423,424) and new onset HF(425). In HF patients an increased risk of death with rising HbA1c level has been found which was worse in non-diabetic than diabetic HF subjects(290).

A challenge for using HbA1c to determine risk of mortality is the U shaped relationship that has been found between them, with both high and low levels of HbA1c associated with increased risk. In patients with type 2 diabetes, risk of HF development was increased at levels of HbA1c <6% as well >10%(426). In HF a U-shape relationship has also been identified(288) with one study of unselected HF veterans finding the highest risk of mortality in the outer most quintiles (<6.4% and >9%)(427). In HF some studies have found the U shape to shift to the right with a higher threshold for HbA1c. In a non-selected HF group of ambulatory veterans with diabetes the lowest risk of mortality was found in those with only modest glucose control (quintile 7.1%-7.8%)(427). In a small sample of patients with advanced systolic HF and diabetes HbA1c >7.0 was protective with lowest risk of mortality risk in the third quartile (7.8-8.9%)(428). This threshold in HF has not been found in other studies, a group of HF patients undergoing revascularisation surgery were found to have the lowest risk at HbA1c levels between 5.8 and 6.2%(429).

Due to the lack of consensus over the level of HbA1c associated with the lowest risk of adverse outcomes in HF, HbA1c levels were defined by 6 categories that included the guideline level of 6.5-7.5% (Table 6.7). Sub-analysis was also performed using deciles of HbA1c to allow a more precise assessment of the lowest risk group and shape of association.

Diabetes severity using drug data: The profile of people with type 1 and type 2 diabetes differs with the latter group generally having more comorbidity such as obesity and a later onset. Type 2 diabetes is managed through a scale of treatment from good diet control to oral medications, which may be supplemented by

insulin for more intense glucose control. People with type 1 diabetes usually have the disease for longer and require intensive insulin interventions from the outset(421). Studies have shown that in the general population of people with type 2 diabetes(430) and in people with HF and diabetes(293) the prescription of insulin is associated with higher risk of adverse outcomes and may represent more severe metabolic disturbances and advanced disease.

A simple drug severity classification was used to differentiate the sample by type of diabetes (Table 6.8). The two oral hypoglycaemic groups were included to differentiate between people with type 2 diabetes who may have more progressive disease requiring supplemental insulin therapy(430).

Diabetes severity change using physiological data: Change in HbA1c has been measured in a prior study. In a CPRD cohort of people with type 2 diabetes, HbA1c change of >1% increase or decrease was compared to <1% change over a year before death(431). Both groups were associated with increased mortality risk and this was most marked in the HbA1c group where levels decreased. In the thesis, the approach was taken to compare patients with >1% increase or decrease with patients who had a change of 1% or less over 1 year or 6 months prior to death and hospital admission.

Diabetes severity change using drug data: For the measurement of diabetes drug severity change, an increase or a decrease in at least one drug category compared to no drug category change over specified time-windows (Section 6.3.4.5). For this change measure the two oral hypoglycaemic drug categories were collapsed into a single category (oral +/- insulin) to create three distinct groups.

6.3.4.3 Chronic kidney disease (CKD)

Three categories of comorbidity measures were included: diagnostic status, severity and severity change.

CKD diagnostic status: The CKD 'before HF cohort' was based on at least one CKD prevalence or index code in the three year time-window before the HF index date. The 'after HF cohort' was based on those (i) without a CKD prevalence or index code applied in their three year before HF and (ii) with at least one CKD

index code applied at any time after HF index date but before the match date. A second approach used the physiological definition of estimated glomerular filtration rate (eGFR) of <60ml/min/m2 to indicate renal dysfunction. This was identified by the most recent measure before the match date but within a maximum 3-year time window for mortality and 6-months for hospital admission. As the eGFR measure can change over time, renal dysfunction definition of CKD was not used to place the timing of comorbid disease in relation to the HF index date.

CKD search strategy: CKD describes abnormal kidney function and/or structure(432) and, due to its close pathophysiological role in HF, is a key comorbidity for this disease group(433). CKD can develop through primary renal disease or as a consequence of HF and both result in similar functional changes to the kidney with glomerular and tubular damage; processes which are intensified when both occur together(433). The diagnosis of CKD is made through the identification of persistent kidney abnormalities or eGFR <60ml/min/1.73m^2 over 3 months(434).

Identification of CKD in administrative data using clinical codes can be challenging given the variety of clinical terms used across a number of Read code chapters to describe a wide range of presentations. The alternative approach using eGFR for case definition also has challenges. Not everyone has serum creatinine measures and whilst defining CKD by medical record likely underestimates the number of cases, using eGFR as a gold standard may overestimate CKD cases by misclassifying older people with normal renal function change or people with acute kidney injury(435).

A broad search strategy for identifying CKD by medical codes was devised based on key words that might indicate kidney disease and specific codes used in the Quality Outcomes Framework (QoF) to indicate kidney disease (E-Appendix E9c). The results of the search strategy were reviewed by a renal registrar and cross referenced with the codes included in QoF. Out of a possible 1449 codes, 120 index codes and 56 prevalence codes were agreed. Codes related to both diagnostic and process codes for CKD and 28 were included in QoF. Any terms that that were not definitive of CKD such as screening or investigation terms or those that could be acute or chronic terms such as renal failure or impairment were excluded (E-Appendix E12).

CKD severity using physiological data: CKD defined by eGFR <60mL/min per 1.73m2 has been found to double the risk of mortality in HF(239,436). CKD is further classified by categories recommended by Kidney Disease: Improving Global Outcomes (KDIGO)(432,434) which were used in this study to measure severity (Table 6.9). The first stage in the guidelines (≥90ml/min/m2) was subdivided into two stages to identify a 'highest' eGFR group which has been found to carry an increased risk of mortality in the general population(434) above 90ml/min/m2 (437). Various reasons put forward include the inadequacies of the eGFR formula at low serum creatinine levels(437) or other causes of low creatinine giving a false account of kidney function such as loss of muscle mass in frailty or the dilutional effects of HF(438).

CKD severity change using physiological data: The most common measure used for detecting renal change over time has been an increase in serum creatinine of ≥0.3mg/dL, usually measured during acute HF admission(146). There is some evidence that change in creatinine, in a more stable state, 6 months after discharge, may have a stronger influence on mortality than when measured during admission(156).

Creatinine is used to calculate eGFR which in turn is used as an estimate of GFR. An alternative to using creatinine is to measure change in eGFR directly. As well as being a better indicator of GFR in a stable population, eGFR change has also been considered as less dependent on baseline renal function than creatinine change(439). Creatinine has an exponential relationship with eGFR. This means that for the same change in serum creatinine there might be a small or large change in eGFR depending on a low or high baseline eGFR respectively. As the rate of change of eGFR is linear across all baseline categories of eGFR(440), using creatinine change means that those with low baseline eGFR (in whom a smaller reduction in eGFR will reach the threshold of creatinine change) are more likely to experience worsening renal failure (WRF) defined by this measure.

There has been conflicting findings as to the association between baseline renal function and the occurrence of WRF with studies showing that both low and high baseline renal functions are associated with subsequent WRF depending on the measure of change used. Whilst some studies have shown an association between

high baseline eGFR and occurrence of WRF measured by eGFR change(439,441,442), two reviews of studies using creatinine change(146) or a mixture of eGFR change and creatinine change(440) have found an association with low baseline function, which is likely due to a faster rate of decline in these groups.

Creatinine change may not be a precise measure of eGFR change, but the importance for prognosis of a large change in eGFR at the higher baseline values and a small change in eGFR at low baseline values may be similar and has been demonstrated by the consistent effect of WRF using serum creatinine change on outcomes that is independent of baseline eGFR(146). This suggests that the threshold for the effect of WRF on mortality is lower in more severe renal disease. If this is true, relative change measures (which take some account of the proportion of renal function loss) may be a better approach.

A change in eGFR was selected over serum creatinine due to its better estimation of GFR in a stable population and the potential for measuring relative change over time. Two indices of change in eGFR have been used previously in general population studies, the absolute difference and percentage change in eGFR, although guidelines have traditionally used absolute change(443).

The thesis definition was an absolute measure of change and a relative measure (to better adjust for the baseline eGFR value) (Table 6.10). The measures included in both classifications are indicated in current clinical guidelines to represent accelerated change in eGFR status over 12 months(432).

6.3.4.4 Extraction and measurement of physiological comorbidity data

The extraction and cleaning of the physiological data was performed by identification of all possible values using the CPRD 'Entity' and Medcodes, conversion of differing units of measure to a standard common measure and removal of biologically implausible extreme data. For detail of the steps taken for each comorbid disease including formula for transformations and plausible ranges see E-Appendix E13.

HbA1c was converted from mmol/mol to % because, although new guidelines now advocate mmol/mol(444), % has been the most common unit reported and used at the time of the data recording. FEV_1 was mostly

reported in litres and so was converted to percent predicted (pp) using a formula that included height and age. Whilst height data was available for most patients, missing data was imputed with the average height by age and gender from the Health Survey for England(445).

eGFR is not reported directly in CPRD and so was converted from creatinine using the simplified 4-variable Modification of Diet in Renal Disease (MDRD) formula. This equation adjusts for significant non-renal influences such as age, sex, race, and body size, is preferred to other equations such as the Cockcroft-Gault formula(399) and has been validated in the HF population(398). The ethnicity component of the formula could not be included given the scarcity of this data in the CPRD. However the prevalence of black people within the UK is less than 3%(446) and given the underestimation of GFR in this group the associated risk would be biased towards the null value.

6.3.4.5 Measurement time windows for comorbid exposure

Measuring comorbidity severity and change using routinely collected physiological data can be restricted by missing data. Whilst the timing for measurement of severity and change will be guided by the research question, missing data will be increased where the time windows for measurement are narrow and/or the timings between repeated measures are short. The key thesis focus was whether in HF, *'recent'* comorbidity severity or change influences outcomes. The hypothesis was that important comorbidity severity change that occurs prior to death may happen over a longer time (i.e. over one year) than that which occurs before hospital admission (over 6 months).

Physiological measurement time-windows: For all physiological measures of comorbidity severity the closest value to the match date was used but with a maximum of 3 years for mortality and 6 months for hospital admission. A previous measure was then taken between 6 months and 3 years before for mortality and between 1 month and 1 year before for hospital admission (Figure 6.2a). Change over 1 year for mortality outcome and over 6 months for hospital admission outcome was then calculated as follows:

For the absolute change measures:

$$\left[\text{most recent measure} - \text{prior measure}\right] \text{X} \frac{\text{Outcome time interval}* \text{ (days)}}{\text{Measure interval (days)}}$$

For the relative change measures (eGFR):

$$\frac{\text{Absolute change}}{\text{Prior measure}} \text{ X } 100$$

*Outcome time interval was 365 days for mortality and 182.5 days for hospital admission outcome

Drug measurement time-windows: Drug measures are routinely and automatically recorded in the CPRD through electronic prescribing systems. This means that the measurement time-windows could be more precise but needed to be wide enough to capture drug prescriptions that may be repeated less regularly. The '*recent drug severity measure'* was defined for both outcomes as at least one prescription in a 4-month time-window before the match date. The prior 4-month time-window was placed at 12 months to 16 months before the match date for mortality and 6 months to 10 months for hospital admission (Figure 6.2b).

6.3.5 Outcome measures

Mortality: Mortality outcome was defined as death by any cause recorded in the patients CPRD record. There are a number of different entry types within the CPRD that indicate a death event and each has an associated date. This means that there may be multiple records and dates for any one patient. The study date of death was derived using a CPRD verified algorithm which takes the earliest of the patient transfer out date (with reason 'death'), first statement of death Read code or date of death/record added in the death administration area of the CPRD.

Hospital admission: For the subset of practices that are part of the CPRD linkage scheme, Hospital Episodes Statistics (HES) data were available. The data available through CPRD includes all admitted inpatient care episodes to English NHS practices including private patients. Discharge date and all ICD-10 codes associated with the patient discharge are provided free of charge. For all patients that were eligible for linkage the

hospital admission outcome was defined as first admission for any cause and the first date of discharge after (but not including) the HF index date was used.

6.3.6 Covariate measures

All covariates included were based on previous evidence on the outcomes in HF identified in the thesis systematic review(239) and other research (180,182,447), as well on their clinical relevance and availability in routinely collected clinical data.

Person and socio-demographics: The person and socio-demographic data available within CPRD includes year of birth, sex and index of multiple deprivation (IMD) score (2007) of the study subjects at entry to the CPRD practice. Age was based on current age on the match date. The IMD score combines seven weighted indicators which cover economic, health, social and housing domains into a single deprivation score(448). Within CPRD the IMD score which is based on English lower super output area is only provided for patients belonging to linked practices with a valid postcode which applied to approximately 60% of study patients. The score was ranked into quintiles ranging the lowest deprivation (quintile 1) to the highest (quintile 5).

Anthropometric data and risk factors: Information was available on body mass index (BMI), smoking, alcohol, cholesterol, haemoglobin and blood pressure (BP). Smoking and alcohol were categorised into three distinct categories (current, previous or never) and all other measures were retained as continuous variables. All measures were converted to a standard measure where necessary, biologically implausible values were removed (E-Appendix E13) and then the most recent value to the match date was extracted.

Drug data: Exposure definition of HF drugs was defined as at least one prescription in a 4-month time window before the match date. Drug information was extracted for angiotensin converting enzyme inhibitors, angiotensin II receptor blockers, beta-blockers and diuretics. Again a search on key words relating to the drugs was also performed to optimise the inclusion of product codes that had not been categorised by a BNF code.

Imputation of data: Missing data were present for the comorbid exposure and covariate measures (E-Appendix A19). Single imputation using chained equations was used to impute values for the covariate data using ICE in Stata version 13. This was the chosen approach for two main reasons. Firstly to maximise the number of patients retained in the analyses to aid the comparison of comorbidity prognostic models and secondly to reduce the risk of selection bias associated with full case analysis where the missing data varies for the cases and controls(225). Regression imputation was selected as the missing values were not completely at random and for some measures were related to the outcome (E-Appendix A20-21). Where this is the case, full case analysis or use of a missing indicator can lead to bias(449). Single rather than multiple imputation of the main confounders was chosen so as not to over complicate the analysis by combining multiple data sets. For the chained equations all predictors were added to the models and logistic, polyotomous or linear regression was used to predict the missing values. Imputation was not performed for IMD for the mortality outcome due to the high level of missing data. Imputation was performed on the matched sample after the investigation of potential confounders and checking the model assumptions (linearity, collinearity) but before the regression modelling. The mean sample values of the pre and post imputed matched data were compared to check that they were similar (E-Appendix A22-23).

6.4 Statistical analysis

6.4.1 Part 1: Descriptive studies

Continuous data variables are first investigated for normality using histograms. Quantile plots were also used to graph quantiles of the continuous variable against quantiles (Q-Q plots) of a normal distribution. Rather than focusing on the centre of a distribution this approach identifies whether any irregularities are in the tails of the distribution(450). Data are then presented as means and standard deviations (SD) for normally distributed continuous data, whilst skewed continuous data is presented as medians with interquartile ranges [IQR]. Dichotomous data are presented as counts and percent prevalence. Data tables are stratified according to the outcome status. Significant difference between groups was determined by parametric tests using independent samples t-test, non-parametric tests using the Wilcoxon rank-sum test and categorical variables between groups using the Chi-square test. Due to the high power to detect very small differences with large samples,

the *P* value was considered significant if ≤0.01. Following matching, due to the large number of observations generated and the dependence of the observations created by the inclusion of repeated controls, the groups were compared descriptively only and absolute differences in units were reported to indicate the magnitude of difference. No adjustment for multiple testing was used.

Description of the samples were carried out in four steps: (i) overall baseline patient characteristics for both cohorts by all-cause mortality and first hospital admission sub-sample, (ii) baseline characteristics for sub-cohorts with different follow-up time-periods to give a more accurate comparison of survivors and non-survivors with varying lengths of follow-up, (iii) two calendar cohort periods that pre and post-dated the introduction of HF into the Quality and Outcomes framework (QOF) (April 2006), to investigate any 'time effect' caused by the introduction of new guidelines and associated clinical practices, and (iv) time-dependent general and comorbidity characteristics of the matched samples stratified by outcomes.

6.4.2 Part 2: Hypothesis testing and strength of associations

There were five stages to the main analyses using the matched data: investigation of confounding, testing for linearity, identifying correlations, regression modelling and testing for interactions.

(i) Potential confounders were investigated in four steps. First, was to identify the confounders from the systematic review. Second, patient characteristics were compared in the comorbid versus non comorbid groups to identify whether the potential confounder was different across the groups. Third, for each of the comorbidities, the unadjusted comorbid effect on the two outcomes was investigated using conditional logistic regression and then observed in different strata of the potential confounder. A 10% difference in the comorbidity exposure effect in any strata of the confounder was deemed important. Fourth, each potential confounder was entered individually, into a conditional logistic model with each comorbidity exposure to observe a difference of 10% in the odds ratio to indicate confounding(9).

(ii) Testing for linearity between the continuous independent variables and the logit of outcome was performed by Likelihood ratio tests. The shapes of the association were also observed using fitted line plots and Eccles plots.

Likelihood ratio tests were performed to compare predictive models with a continuous variable included in its most simple form to the same models with a continuous variable included together with a quadratic and cubic extensions or log terms. A *P* value of ≤0.01 was considered to be significant to reject the null hypothesis that the simpler model (nested model) was a better fit.

Eccles plots graph the mean predicted value of the outcome against deciles of a continuous variable. The eclplot includes a continuous variable in the model and then plots the predicted values of the outcome over the same variable in categories. This means that the variable can be added in different forms to the model (continuous, continuous with quadratic or cubic extensions or as a log transformed variable) and fully adjusted for other covariates(451). Lowess lines were then overlaid on the eclpolt. Lowess provides a locally weighted scatterplot smoothing where a separate weighted regression is performed for every point in the data using the data point (most weight) and a few of the nearby data points (least weight). Lowess therefore follows the data more closely and can be used to observe the calibration between the predictive models (including different expressions of the continuous variable) and the observed data.

Following these steps, for any non-linear continuous variables it was decided to use quadratic extensions rather than any higher order term or more complex transformations. This was to avoid overfitting and also for ease of clinical interpretation(225).

(iii) Multi-collinearity was investigated for all variables to be added to models. Where variables had a correlation coefficient of >0.5, the most clinically relevant one was selected. Where continuous variables were included with quadratic extensions they were centered at their mean to remove collinearity between the variable and its quadratic term(452).

(iv) Strength of associations: Conditional logistic regression (discussed in detail in Chapter 4.4.2) was performed for each comorbidity measure to estimate rate ratios for all-cause-mortality and first hospital admission. Unadjusted associations were followed by stepped adjustment of confounders according to their potential importance indicated by the confounder investigation. These included variables related to the person (age, gender), socio-demographic factors (deprivation), comorbidities (remainder of COPD, DM and CKD), lifestyle factors (alcohol, smoking), risk factors (blood pressure, cholesterol, BMI, haemoglobin) and HF medication (angiotensin converting enzyme inhibitor, angiotensin-II receptor blocker, beta-blocker, diuretic). The mortality models did not include deprivation given the level of missing data in the overall cohort but a sensitivity analysis with the deprivation variable was performed. The hospital admission models were further adjusted for recent hospital admission within three months.

As eGFR is routinely measured and collected on all HF patients the renal change comorbidity exposure was adjusted with the most recent eGFR measure and separately the prior eGFR measure used in the change calculation. Other comorbidity change measures are only routinely recorded in comorbidity groups so these measures were compared to non-comorbid reference groups and could not be adjusted for baseline comorbid status.

Stratified associations: For each comorbidity effect on outcomes, significant stratification by categories of comorbidity severity and change were investigated by observing the separation of the confidence intervals across strata. Where confidence intervals overlapped a linear test of trend was performed on the adjusted associations using likelihood ratio tests.

(v) *Interactions*: *Statistical interactions* between (a) pairs of the three comorbidities (DM, COPD, CKD), (b) comorbidities and HF severity indicated by age as a proxy measure, and (c) comorbidities and specific patient characteristics indicated in the investigation of confounding, were investigated. This was done by observing the unadjusted effects of the comorbidities on outcomes within strata of a second variable of interest. Separation of the confidence intervals around effects was used to indicate *potential* interaction (the effect of the comorbidity is modified by the second variable). These associations are unadjusted and the relative

effects are partly determined by the baseline risk in the strata groups. Where the baseline risk differs, the relative effect of the comorbidity will differ across strata even if the absolute effect of the comorbidity is the same. Using the approach, the absolute risk associated with the comorbidity across strata, cannot be directly compared(9).

These interactions were further investigated by adding interaction terms to conditional logistic models and performing likelihood ratio tests between the fuller and simpler models ($p \leq 0.01$). As conditional logistic regression is an exponential model, a significant statistical interaction indicated a departure from multiplicity. This approach indicates whether a more flexible model which includes an interaction term has better fit. Again, even where statistical interaction exists, whether the absolute risk associated with a comorbidity differs due to its interaction with another variable, cannot be investigated using this approach(9).

Biological interaction refers to the risk associated with an exposure on an outcome being, in part, dependent on the presence of two variables (Section 1.1.3). This is a theoretical concept which refers to a deviation from additivity of two risk factors when considered together. In biological interaction, the combined effect of two exposures is more or less than the sum of their separate effects. This means that these factors act together in a causal mechanism (11). Whilst this departure from additivity can be easily assessed in a linear model which is an additive model, when using a multiplicative model (e.g. logistic regression), specific analysis is required.

To test biological interaction between two comorbidities in the conditional logistic regression model three steps were taken as recommended by Rothman(453). First step, dummy indicator variables were created for the presence or absence of three of the four possible disease combinations for each pair of comorbidities. In the DM and COPD example this was VAR_{01}: DM=0, COPD=1; VAR_{10}: DM=1, COPD=0; VAR_{11}: DM=1, COPD=1. The fourth possible category (DM=0, COPD=0) served as reference category (Table 6.11). Second step, conditional logistic regression was then performed for the dummy variables for each disease pair in turn, adjusted for all other confounders.

Three measures of biological interaction were considered(11):

(1) The relative excess risk due to the interaction (RERI) is the difference between the expected risk associated with two diseases combined and the observed risk of the individual diseases:

$$RERI = RR_{11} – RR_{01} – RR_{10} +1$$

(2) Attributable proportion due to interaction (AP) is the proportion of risk that is due to interaction among persons with both diseases:

$$AP = RERI / RR_{11}$$

(3) Synergy index (SI) is the excess risk from joint exposure in the presence of interaction relative to the risk from exposure without interaction:

$$SI = \frac{RR11 - 1}{(RR01 - 1) + (RR10 - 1)}$$

Biological interaction is indicated where RERI and/or AP ≠ 0 and/or synergy index ≠ 1. For multivariable models that have a number of covariates the synergy index is recommended over the other measures (as this remains stable across strata of covariates) and was the main approach taken. Third and final step, using a covariance matrix generated by the model, confidence intervals were estimated(454).

6.4.3 Part 3: Model building; testing different comorbidity measures

Pre-specified models for each outcome were developed using all non-comorbidity covariates (the core model) in order to compare models with the inclusion of the different comorbidity measures. A pre-specified model was chosen from the available measures in the CPRD on the basis that the covariate was known to influence outcomes through prior evidence(180,182,447) and was clinically relevant and routinely collected. This also avoided using an approach that relies on statistical testing such as stepwise selection which can lead to bias and overfitting(225,455-457). The core model included age, gender, alcohol, smoking, systolic blood pressure, cholesterol, BMI, haemoglobin, HF medication and quadratic extensions of the continuous variables where indicated. Deprivation was not available for all subjects in the mortality sample and was non-significant in the

multivariable models for first hospital admission and so was not included in the core model. Prior hospital admission was added to the core models for hospital admission outcome. All continuous variables were centred about their mean prior to their inclusion to remove collinearity where their quadratic terms were also included and for better interpretation of the final models (Chapter 12).

Each set of measures for the three comorbidities were tested separately in the core model. A comorbidity was added to the core model first by its status (presence or absence indicator) (model 1). Then the comorbidity status measure was exchanged for a comorbidity severity measure (drugs; model 2a *or* physiological measure; model 2b). Finally the comorbidity severity measure was exchanged for a severity change measure (drugs; model 3a *or* physiological measure; model 3b). The remaining two comorbidities were added to the each core model by their status measure (DM, COPD and eGFR<60 for CKD). See Figure 6.3 for an example of the models for diabetes mellitus.

Conditional logistic regression was performed for each model and then discrimination and tests of model fit were used to compare models. Models were first created using the full data followed by a restricted set of data. Due to varying amounts of missing data in the comorbidity measures, in the second step, models were restricted to the data which had complete observations for all the models to be compared. This meant that a set of models could then be compared using the same data with the same number of observations.

A final step compared a set of models with all three comorbidities added by (i) status, (ii) exchanged for comorbidity severity measures and lastly (ii) exchanged for severity change measures.

Calibration and discrimination: Calibration using the Hosmer Lemeshow test could not be performed following conditional logistic regression due to the model intercept not being estimated. The assumption is that without the baseline risk the predicted risk would be less than the observed risk. Discrimination used the area under the receiver operating characteristic (ROC) curve also called the concordance index (C-index). This index investigates the probability that for a randomly selected case and control the predicted risk of the case would be higher. This is a less sensitive measure for prediction modelling where the magnitude of risk within groups

is more important than the presence of risk. ROC also performs particularly poorly at evaluating the contribution of predictors to a model or where the range of risk within a population is narrow(458). For this reason other measures of model fit were used in addition to the ROC.

Goodness of fit measures: Log-Likelihood and likelihood ratio tests ($p \leq 0.01$) were used to compare nested models. The null model was the comorbidity status model compared to its alternative more flexible model where the comorbidity was stratified into groups of severity or change. Akaike and Bayesian information criterion (AIC and BIC) which also take account of the complexity of the models were also used as measures of model quality. Both BIC and AIC penalize for the number of parameters in the model and can be used to compare non-nested models. Difference of more than 2 in the AIC or BIC between models indicates positive difference and favour of the model with the smallest AIC or BIC(459). Finally McFadden's pseudo R^2 was used to assess model fit. Unlike R^2 in linear regression, McFadden's pseudo R^2 is not based on how well the models explains the variance in the data but is an index of the likelihood ratio with 0.2 to 0.4 indicating excellent fit(460) (for summary of measures see E-Appendix A24).

6.5 Chapter summary

This chapter has detailed the methods and analyses that were applied in the main CPRD studies described in the following six chapters. DM, COPD and CKD comorbidities were selected from the systematic review and different measures of comorbidity severity and change using routinely collected drug and physiological data were defined. Part 1 includes the descriptive studies of the baseline and matched HF samples for the two outcomes; mortality (Chapter 7) and first hospital admission (Chapter 8). Part 2 focuses on the comorbidity prognostic factor studies for mortality (Chapter 9), hospital admission (Chapter 10) and interactions (Chapter 11). Part 3 focuses on the comorbidity prognostic model study (Chapter 12). Chapters 7 to 12 will present the main results before detailed discussion in Chapter 13.

Tables

Table 6.1 COPD-related drugs

Drug class	BNF section
Beta2-antagonists: SHORT	3.1.1.1
Beta2-antagonists: LONG	3.1.1.1
Anti-cholinergics: SHORT	3.1.2
Anti-cholinergics: LONG	3.1.2
Methylxanthines	3.1.3
Compound inhalers (SHORT x2)	3.1.4
Inhaled steroids	3.2
Inhaled steroids + beta2-antagonist:SHORT	3.2
Inhaled steroids + beta2-antagonist:LONG	3.2
Oral steroids – prednisolone	6.3.2
Oral steroids – methylprednisolone	6.3.2
Oxygen	3.6

BNF; British National formulary

Table 6.2 COPD severity stages based on GOLD criteria

Severity Stages	Description	Forced expiration volume in 1 second (FEV_1)
GOLD 1	Mild	$FEV_1 \geq 80\%$ predicted
GOLD 2	Moderate	$50\% \leq FEV_1 < 80\%$ predicted
GOLD 3	Severe	$30\% \leq FEV_1 < 50\%$ predicted
GOLD 4	Very severe	$FEV_1 < 30\%$ predicted

From GOLD guidelines(410)

Table 6.3 COPD drug severity measure framework

Severity Group	Guideline Recommended First Choice	Severity category description
A Short acting	Short-acting anticholinergic prn OR Short-acting beta2-agonist prn	Any short acting bronchodilators
B Monotherapy	Long-acting anticholinergic OR Long-acting beta2-agonist OR Inhaled corticosteroid	Any one long acting bronchodilator on its own or Methylxanthines monotherapy (without other step up therapy) or inhaled steroid on its own +/- short acting inhalers
C Dual therapy	Inhaled corticosteroid AND [long-acting beta2-agonist OR long-acting anticholinergic] OR Both long-acting beta2-agonist AND long-acting anticholinergic AND no inhaled corticosteroid	Any long acting bronchodilator and an inhaled steroid (separate or combined). Methylxanthines may replace one of the bronchodilators +/- short acting inhalers OR Both long acting bronchodilators without inhaled steroid +/- short acting inhalers
D Triple therapy	Inhaled corticosteroid AND long-acting beta2-agonist AND long-acting anticholinergic	Both long acting bronchodilators and an inhaled steroid. Methylxanthines may replace one of the bronchodilators or be additional to both +/- short acting inhalers
E1 Additional therapy E2 Additional therapy	On oral steroids On prescribed oxygen therapy	May be in addition to any of the drugs A-D

Table 6.4 COPD drug severity classifications

Severity classification A	Severity classification B	Severity classification C
1 No drugs in time window	1 No drugs in time window	1 No steroids or oxygen
2 Short term inhalers only (A)	2 On inhalers only (A-D)	2 On oral steroids but no oxygen (E1)
3 Monotherapy (B)	3 On oral steroids but no oxygen (E1)	3 On oxygen (E2)
4 Dual therapy (C)	4 On oxygen (E2)	
5 Triple therapy (D)		
6 Steroid (E1)		
7 Oxygen (E2)		

Severity categories A-E devised using GOLD guidelines(410).

Table 6.5 COPD physiological severity change measures

COPD FEV_1% severity change classification for all-cause mortality/ hospital admission	
1	An absolute decline of 5% in FEV_1 pp over a year/ 6 months
2	An absolute decline of 10% in FEV_1 pp over a year/6 months
3	A decrease in at least one GOLD stage over a year/6 months

FEV_1, forced expiratory volume in 1 second; pp, percent predicted

Table 6.6 COPD drug severity change classifications

Severity classification A	Severity classification B
1 Drug category same or better	1 No new steroid or oxygen
2 Drug category worse	2 New steroid but no new oxygen
	3 New onto oxygen

Table 6.7 Diabetes physiological severity stages

Level	HbA1c (Glycated haemoglobin)
1	<5.5%
2	5.5-6.4%
3	6.5-7.5%
4	7.6-8.5%
5	8.6-9.5%
6	>9.5%

Table 6.8 Diabetes drug severity classification

Level	Diabetes drug severity classification
1	No drugs in time window
2	Oral hypoglycaemic drugs only
3	Oral hypoglycaemic drugs and insulin combined
4	Insulin only

Table 6.9 Chronic kidney disease physiological severity stages

CKD Category	eGFR (ml/min/m2)
1+	>105
1	90-105
2	60-89
3a	45-59
3b	30-44
4	15-29
5	<15

Adopted in guidelines(434,461). eGFR, estimated glomerular filtration rate

Table 6.10 Chronic kidney disease physiological severity change measures

Severity class 1 (absolute change)	Severity class 2 (relative change)
0-5 mls decrease in eGFR	0-5% decrease in eGFR
6-15 mls decrease in eGFR	6-25% decrease in eGFR
>15 mls decrease in eGFR	>25% decrease in eGFR
Any increase	Any increase

eGFR; estimated glomerular filtration rate (using MDRD equation)

Table 6.11 Definition of dummy variables for different comorbidity exposure combinations

Exposure levels	VAR $_{01}$	VAR $_{10}$	VAR $_{11}$
DM =0, COPD = 0 (reference)	0	0	0
DM =0, COPD = 1	1	0	0
DM =1, COPD = 0	0	1	0
DM =1, COPD = 1	0	0	1

Figures

Figure 6.1 Flow chart of the patient selection process

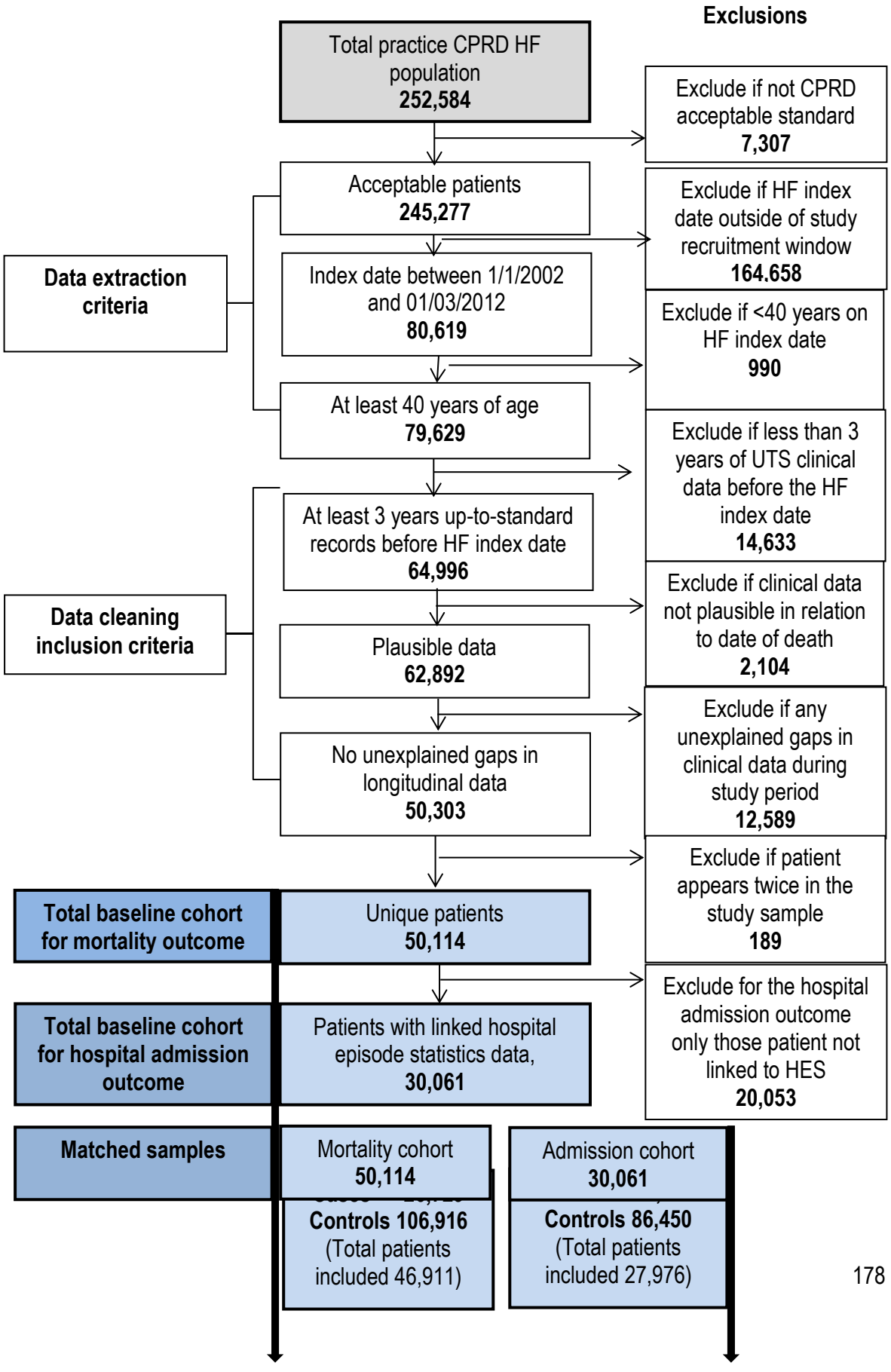


Figure 6.2a Time-windows for comorbidity severity change using physiological measures

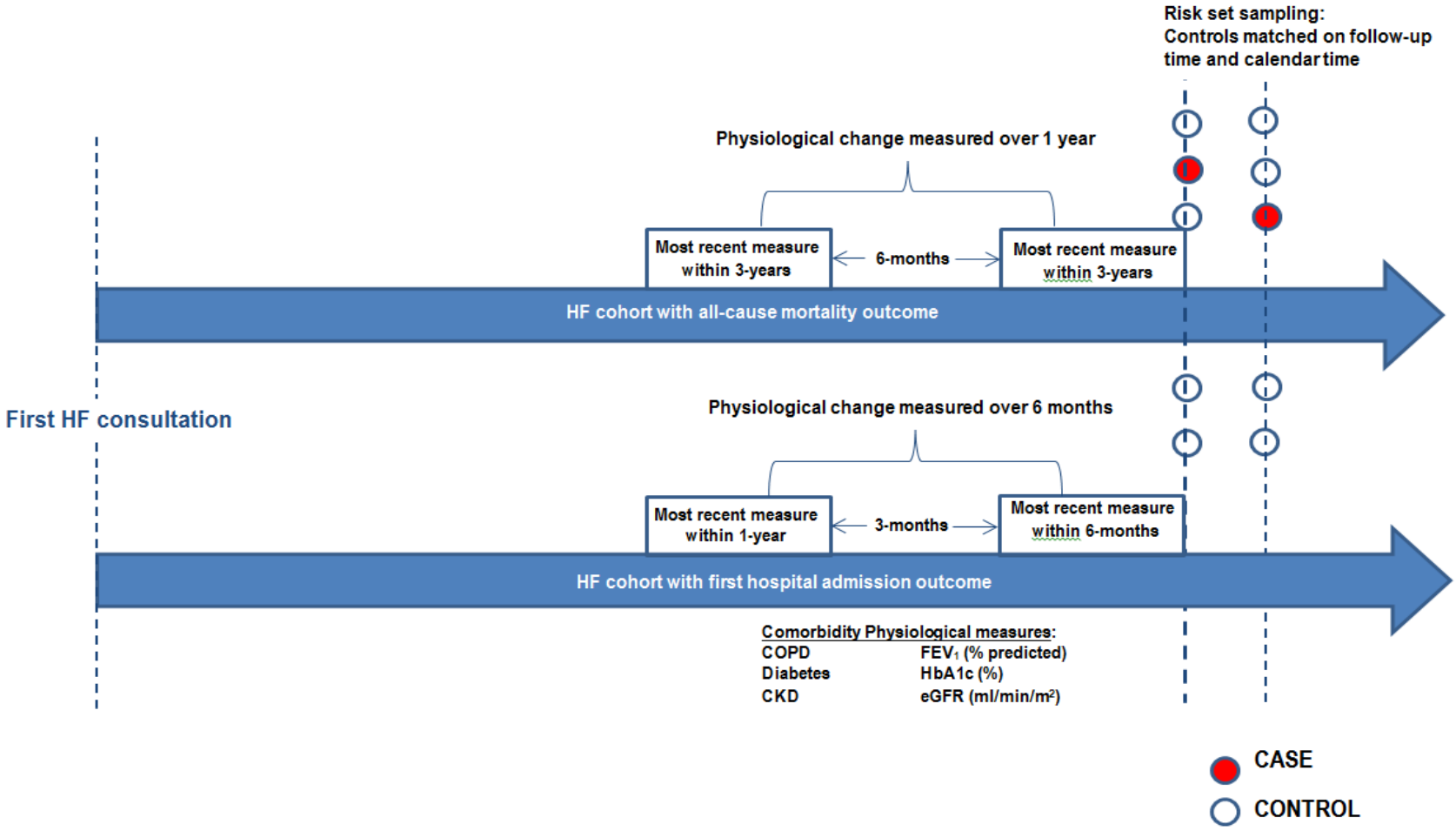


Figure 6.2b Time-windows for comorbidity severity change using drug measures

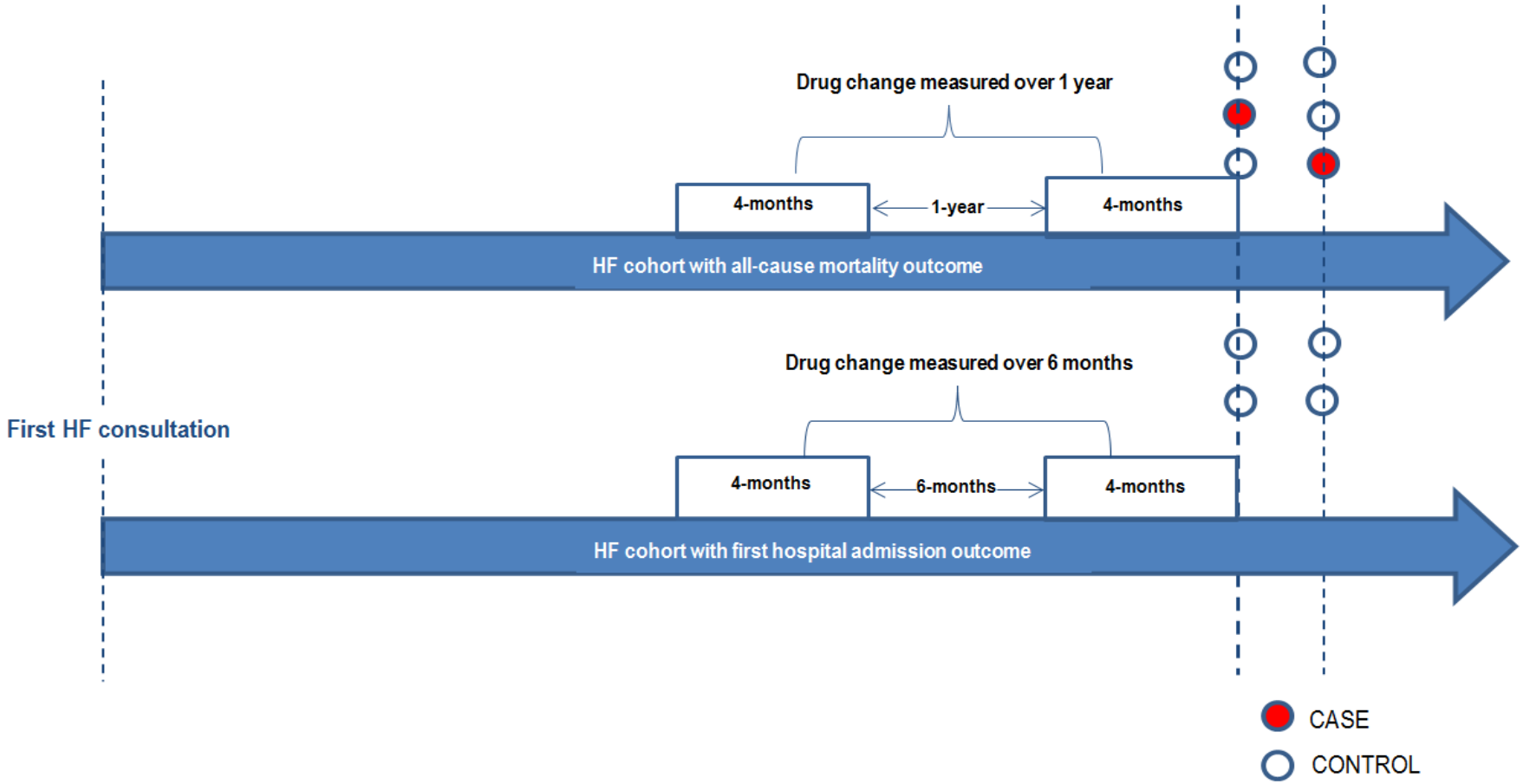
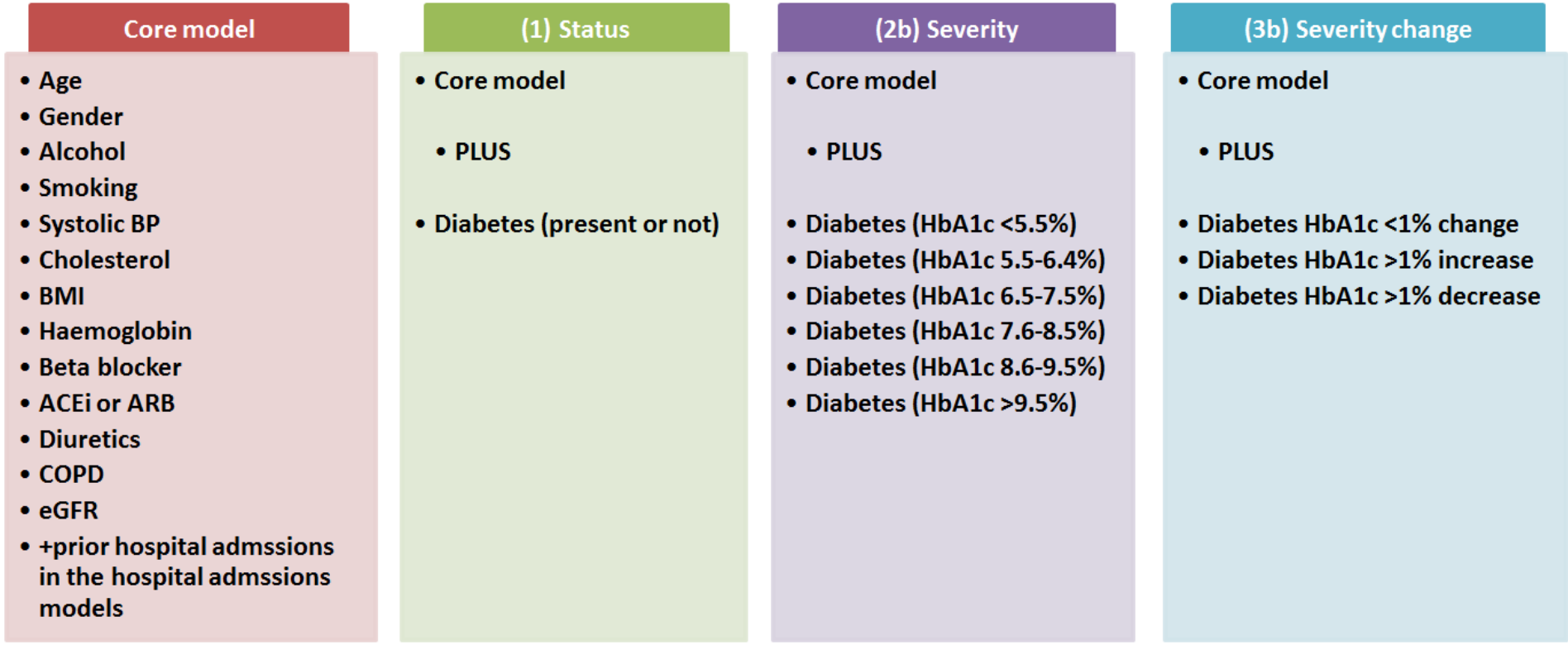


Figure 6.3 Models for diabetes mellitus

.



Chapter 7 Non-CVD comorbidity prognostic factors in an incident HF general practice cohort sample by all-cause mortality

This chapter presents the results of the main incident HF CPRD cohort aged 40 years and over, with a first HF consultation between 1st January 2002 and 1st March 2012 followed up until January 2014. The chapter describes the HF population as a whole and investigates the unadjusted associations between non-CVD comorbidities at baseline and as they develop over time and all-cause mortality, but also the other patient, social, lifestyle, risk and medication factors that may influence this outcome. Baseline characteristics are described in the sample using the available data including missing values. Unless stated all baseline associations were significant at a level of $p<0.01$. For the matched sample, the patient numbers reported for the cases and controls refers to the number of observations generated by the matching process. Following imputation, differences in time-dependent characteristics between the cases and controls are described. Results are presented as number (%), mean (SD) or median [Interquartile range; IQR].

7.1 Study sample: baseline characteristics of total sample by mortality outcome

CPRD sample and follow-up: There were 50,114 subjects with a first HF consultation code between 1st January 2002 and 1st March 2012. Follow-up ranged from 0 to 12 years (median 2.57 [IQR 0.81-4.96] years and mean 3.25 (SD 2.86) years). The risk of death expressed as a percentage of the total cohort at baseline decreased exponentially with each year of follow-up time, shown in Figure 7.1 (dark purple), with most events occurring in the shorter follow-up time. There were 26,729 (53.3%) deaths from any cause over the follow-up, with the risk of death in the first 3-months (13.3%), first year (23.3%) and within 5 years (46.4%). Figure 7.1 also shows the cumulative risk of death per year of follow-up (light purple). Following the initial year, the risk of death in the remaining survivors per year (with the number of deaths expressed as a percentage of the remaining survivors in each year of follow-up) remained stable at 8-11% per year shown in Figure 7.2.

7.1.1 Patient characteristics at baseline in the HF population sample by mortality outcome

Baseline age at HF diagnosis was negatively skewed within the sample with a higher number of older participants; mean 76.9 (SD 10.9) years and median 78 [IQR 71-85] years (E-Appendix E14 for all continuous variable histograms and Q-Q plots). Over the total follow-up time, compared to the survivors, non-survivors were more likely to be older, female and more deprived (Table 7.1). Age was significantly higher in those that died (81 [74-87] years) than those who remained alive (75 [66-82] years). The sample was 47.1% female with an increase in the proportion of women in the non-survivors compared to survivors (48.5% versus 45.4%). In the non-survivors there were more current smokers (15% v. 13.3%) but fewer alcohol drinkers (68.9% v. 73%).

Body mass index (BMI) was positively skewed with more participants with lower BMI. Cholesterol and estimated glomerular filtration rate (eGFR) showed slight positive skewness in the histograms but mean and median values were similar and the Q-Q plots showed that the central parts of their distributions were normal. Haemoglobin and blood pressure were approximately normally distributed. At baseline the non-survivors, compared to the survivors, had lower diastolic blood pressure (mean 75.8 (SD11.9) v. 77.7 (2) mmHg), BMI (median 26.2 [IQR 23.1-30] v. 27.9 [24.7-32] Kg/m^2), cholesterol (4.6 (1.2) v. 4.7 (1.2) mmol/L) and haemoglobin (12.7 (1.9) v. 13.4 (1.8) g/dL).

Of the total HF sample, 32.7% were prescribed beta-blockers at baseline, 55.3% were prescribed an angiotensin converting enzyme inhibitor (ACEi) or an angiotensin receptor blocker (ARB) and 64% were prescribed diuretics. Non-survivors compared to survivors, were significantly less likely to be prescribed beta-blockers (29.7% v. 36.1%), ACEi *or* ARB (50.9% v. 60.2) and more likely to be prescribed diuretics (72.2% v. 54.5%).

7.1.2 Comorbidity exposures at baseline in the HF population sample by mortality

At HF index date, 5,848 (11.7%) of the sample had chronic obstructive pulmonary disease (COPD), 10,533 (21%) had diabetes mellitus (DM) and 7,621 (15.2%) had chronic kidney disease (CKD). When CKD was defined by an eGFR <60 ml/min/m^2, prevalence at baseline increased to 20,084 (49.8%). Prevalence of all the

comorbidities were significantly higher in the non-survivors compared to survivors; COPD (13.6% v. 9.4%), DM (22% v. 19.9%) and CKD (diagnostic: 16% v. 14.7%, eGFR: 57% v. 41%) (E-Appendix B14). Non-survivors were also more likely to have ≥2 co-existing comorbidities (21.6% v. 14.3%). The most prevalent disease combination was DM and CKD (eGFR defined) (10.3%) and this predominated in the non-survivors (12% v. 8.4%) (E-Appendix B15).

In the comorbid DM sample the median glycated haemoglobin (HbA1c) at baseline was 7.1 [IQR 6.4-8.1] % which was not associated with mortality (*P*=0.89). In the COPD sample, the median forced expiration volume in 1 second (FEV_1) was 53.3 [39-69.6] percent predicted (pp) which was significantly lower in the non-survivors than survivors (52.8 [38.2-69.2] v. 59 [46-72] pp). eGFR was measured in 80.5% of the baseline HF sample and was mean 61.0 (20.3) ml/min/m^2. This level was significantly lower in non-survivors than survivors (57.4 (20.7) v. 64.9 (19.3) ml/min/m^2).

7.2 Study sample: baseline characteristics by mortality outcome of HF sub-samples with (i) different follow up periods and (ii) different calendar time-periods

The total HF sample will now be reported (i) in 3 subgroups defined by their follow-up time and (ii) in two sub-groups according to whether their diagnosis date was before or after April 2006 when HF was introduced into the Quality Outcomes Framework (QOF) in the UK. This was done firstly to investigate whether the unadjusted associations between baseline factors and mortality in the total sample were confounded by variations in the exposure time between the non-survivors and survivors with different follow-up times within the sample. Secondly, it was to investigate whether changes to HF clinical practice as a consequence of new guidelines or practices during the total follow-up time influences this association.

7.2.1 HF sub-samples with different follow-up periods: patient characteristics at baseline by mortality outcome

The cohort was sub-grouped into three overlapping HF samples with different follow-up periods: (i) at least up to 1 year (whole sample), (ii) between 1 and 5 years and (iii) ≥5 years. Non-survivors and survivors at the end of each time-period were compared (Table 7.2). The unadjusted associations between the baseline characteristics and mortality were estimated for the total follow-up time and by three different lengths of follow-up (see E-Appendix A25). As the three lengths of follow-up had different HF sample sizes it was important to observe the magnitude of exposure difference as well as the significance. The smaller sample in the longer follow-up period would reduce the power to detect significant difference between the groups. Both the significance and magnitude of the difference in exposure status between non-survivors and survivors differed across the three follow-up periods for some exposures.

Older age remained consistently associated with mortality across all 3 follow-up periods but the median baseline age of the non-survivors reduced from 83 years to 78 years from the shorter to the longer follow-up. Proportion of the most deprived groups and smokers remained higher in the non-survivors than survivors across the 3 follow-up periods but significance was lost in the longest follow-up groups for deprivation and in the longest 2 follow-up periods for smoking. The magnitude of difference in the proportion of women between survivors and non-survivors decreased over time (4.7% to 2.4%) and whilst remaining higher in the non-survivors in all three follow-up periods, became non-significant in the longest follow-up period. Lower diastolic blood pressure, BMI, cholesterol and haemoglobin were significantly associated with mortality in each of the 3 follow-up periods (with the exception of cholesterol which became non-significant in the longest follow-up period), but the magnitude of exposure difference between survivors and non-survivors reduced over time for each risk factor. Lower systolic blood pressure (SBP) was associated with mortality in the shortest follow-up period but the association was between higher SBP and mortality in the two longer follow up periods.

7.2.2 HF sub-samples with different follow-up periods: comorbidity exposures at baseline by mortality outcome

All three baseline comorbidities were associated with increased mortality. The magnitude of difference in comorbidity prevalence between non-survivors and survivors remained consistent and significant across all 3 follow-up periods for COPD and CKD (eGFR<60ml/min/m^2) with higher prevalence in non-survivors. Diagnostic CKD prevalence in the group with the longest follow-up period was much lower compared to the total cohort (8% v. 15.2%) and its association with mortality reversed in the longest follow-up period with higher prevalence in survivors. The magnitude of difference in prevalence of DM between non-survivors and survivors was small in the shortest follow-up period (0.1%) but became larger in the longer 2 follow-up periods (3.5-3.6%). In the DM group, there was minimal difference in median HbA1c across survivors and non survivors (0.1%) and whilst significant in the larger shortest follow-up group, this lost significance in the two longer follow-up periods. The magnitude of difference in FEV_1 (in the COPD group) and eGFR between the survivors and non-survivors was consistent across all 3-follow-up periods but FEV_1 lost its significance in the longest follow-up group.

7.2.3 HF sub-samples with different temporal periods: patient characteristics at baseline by mortality outcome

There were 22,331 (44.6%) participants who had a first diagnosis of HF before April 2006 (approximately halfway through the cohort time) and 27,763 (55.4%) after (E-Appendix A25). With the exception of baseline age and deprivation, all patient, lifestyle, risk and medication factors were significantly different across the two time-periods. After April 2006, there were a smaller proportion of women with HF, fewer smokers and drinkers; systolic and diastolic blood pressure (DBP), cholesterol and haemoglobin were significantly lower and BMI was significantly higher. Less people were prescribed diuretics and more people were prescribed beta-blockers, ACEi and ARB.

7.2.4 HF sub-samples with different temporal periods: comorbidity exposures at baseline by mortality outcome

Prevalence of COPD and DM comorbidities in HF significantly increased after April 2006 by 2.9% and 4.3% respectively. Diagnostic CKD increased from 2.9% in the period before April 2006 to 25.1% after April 2006, an increase of 22%. CKD defined by eGFR <60ml/min/m^2 decreased from 55% to 46.5% over the same time-period.

7.3 Study sample: Time-dependent characteristics of the matched HF sample by mortality outcome

From the total HF population sample, all 26,729 all-cause mortality cases were matched to four controls. Controls were live, active in practice and remained at risk of death on the case index date. Controls were matched to cases within 1-month of their HF diagnosis study-in date and on duration of follow-up. This resulted in 133,645 observations including 106,916 controls. The match date for the controls was the date that corresponded with the same amount of follow-up time as the case. Person, socio-demographic factors, lifestyle factors and risk factors were measured as the most proximal to the match date during follow-up. Drug and comorbidity exposures were measured in specific study time-windows as outlined in chapter 6 (Section 6.3.4.5).

7.3.1 Patient characteristics in the matched HF population sample by mortality

Median current age on the match date for the mortality cases was 81 [IQR 74-87] years which was higher than controls (76 [68-82] years). Cases were also more likely to be female and more deprived (Table 7.4). The matched sample was 46.2% female with an increase in the proportion of women in the cases compared to controls (48.5% versus 45.6%). The prevalence of current smokers between cases and controls was similar (11.7% versus 11.3%) but there were more current alcohol drinkers in controls (68.9% v. 73%).

Cases compared to controls, had lower most recent systolic blood pressure (mean 126.9 (SD 22.3) v. 132.5 (19.9) mmHg), diastolic blood pressure (71.1 (12) v. 73.9 (11.1) mmHg), BMI (median 25.4 [22.1-29.3] v. 27.3 [24-31.3] Kg/m^2), cholesterol (4.4 (1.2) v. 4.5 (1.2) mmol/L) and haemoglobin (12.3 (2) v. 13.1 (1.8) g/dL).

Of the total sample, 55.5% were prescribed beta-blockers on the match date, 70.7% were prescribed an ACEi or an ARB and 77.3% were prescribed diuretics. Cases compared to controls, were less likely to be prescribed beta-blockers (45.5% v. 58%), ACEi *or* ARB (56.4% v. 74.3%) and more likely to be prescribed diuretics (80.7% v. 76.4%).

7.3.2 Comorbidity exposures in the matched HF population sample by mortality

Comorbid COPD group characteristics: COPD was present in 18,478 (13.8%) HF patients before the match date. Prevalence was higher in cases (17.3%) than controls (13%). Compared to HF patients without COPD, those with HF-COPD were more likely to be younger and deprived but less likely to be female. The COPD group had lower blood pressure and cholesterol, but higher BMI and haemoglobin than the non-COPD HF group. In the HF-COPD group, 20.7% were current smokers compared to only 9.7% in the non-COPD HF group. Compared to the non-COPD HF group, the HF-COPD group were less likely to be prescribed beta-blockers (36.7% v. 58.6%) or ACEi/ARB (68.9% v. 71.1%) but more likely to be prescribed diuretics (80.9% v. 76.7%). Prevalence of concomitant diabetes and CKD were less likely in the HF-COPD group who had a higher mean eGFR (61.0 (SD 22.4) v. 57.3 (20.9) ml/min/m^2) (E-Appendix A27).

COPD status: Of the 18,478 HF patients comorbid with COPD, 3,995 (21.6%) had developed the comorbidity after HF diagnosis. COPD that developed before or after HF was associated with mortality but the magnitude of difference between cases and controls was greater for the pre-HF than the post-HF COPD group (3.5% v. 0.9%) (Table 7.4 and Figure 7.3).

COPD severity: There were 8,515 (53.9%) HF-COPD patients who had had at least one recent FEV_1 recorded within the 3-year measurement window before the match date (median 295 [IQR 137-524] days). In the comorbid group FEV_1 was lower in the cases than controls (median 51.9 [IQR 38-69] v. 55.2 [41-71] pp) and most were in GOLD severity stage 2 (43.7%) and 41.3% were in stages 3-4 (see Figure 7.4). HF-COPD cases were more likely than controls to be in the most severe GOLD severity groups (stage 3 or 4; 46.3% v. 38.7% respectively).

Of the comorbid group, 12.8% were not on any COPD-related drugs at the time of matching which was similar for cases and controls. Prescribing of at least one inhaler therapy was lower in the case group compared to controls (49.4% v. 59.4% respectively). Cases were less likely to be on mono, dual or triple inhaler therapy than controls but more likely to be prescribed steroids (30.6% v. 24% respectively) or oxygen therapy (8% v. 3.6%).

COPD severity change: In the HF-COPD group, 4,882 (25.3%) had had a second FEV_1 (pp) recorded prior to the most recent measure (median 462 [IQR 360-663] days). Over a year, before the match date, 19.1% had experienced a worsening of at least one GOLD stage which was a higher in cases (21.8%) than controls (17.5%). Mortality cases were also more likely than controls to experience a 5% or 10% decrease in FEV_1 (pp) over the same time period. In the HF-COPD group, 1,425 (30%) of those with two FEV_1 measures experienced a 5% decrease in FEV_1 and 710 (15%) experienced a 10% decrease in FEV_1. These decreases predominated in mortality cases over controls (10% FEV_1 loss; 16.9% v. 14.3%; 5% FEV1 loss; 33.3% v. 28.8%).

Of the HF-COPD group, 3,399 (18.4%) were newly prescribed oral steroids or oxygen therapy compared to the year before. New prescription of these drugs was more prevalent in cases than controls (steroids; 19.2% v. 14.4%; oxygen; 5% v. 2% respectively).

Comorbid DM group characteristics: DM was present in 31,962 (23.9%) of HF patients before to the match date. Prevalence was higher in mortality cases (25.1%) than controls (23.6%). Comorbid HF-DM patients were more likely to be younger and deprived but less likely to be female than the non-DM group. The DM group had lower cholesterol and haemoglobin but higher BMI than the non-DM group (median 29.2 [IQR 25.5-33.7] v. 26.3 [23.1-30] kg/m^2). There were fewer current smokers and alcohol drinkers in the comorbid HF-DM group compared to the non-DM group. The HF-DM group were more likely than the non-DM group to be prescribed beta-blockers (61.5% v. 53.7%), ACEi/ARB (77.9% v. 68.5%) and diuretics (82.5% v. 75.7%). Prevalence of concomitant COPD was less likely in the HF-DM group but concomitant CKD was more likely

with a mean eGFR of 55.6 (SD 22.4) compared to 58.6 (20.7) ml/min/m^2 in the non-DM HF group (E-Appendix A28).

DM status: Of the 31,962 HF patients with DM, 4,787 (15.0%) developed DM after HF diagnosis. DM that developed before HF was higher in the case group (22% in cases v. 19.9% in controls). These figures were opposite for DM that developed after HF, with prevalence higher in controls than cases (3.7% v. 3.1%) (Table 7.5 and Figure 7.5).

DM severity: Of the comorbid group, 28,865 (90.3%) had had at least one recent HbA1c recorded before the match date (median 114 [IQR 52-206] days). Around 36% of the HF-DM patients had had a recent HbA1c level between 6.5-7.5%. Median HbA1c was significantly lower in the cases than controls (7 [IQR 6.3-8] v. 7.1 [6.4-8.1] %), and lower HbA1c categories (<6.5%) but not higher HbA1c categories (≥6.5%) were more prevalent in cases than controls.

Of the case group, 26.4% were not on any DM-related drugs at the time of matching and these patients were more prevalent in the mortality cases (30.4%) than the controls (25.4%). Most HF-DM patients were on oral medication (57%) +/- insulin, which was less prevalent in the case group compared to controls (50.3% v. 58.8%). Of the comorbid group, 5,061 (16.6%) were prescribed insulin only which was associated with mortality in 19.3% of the cases and 15.8% of controls.

DM severity change: There were 26,174 (81.9%) HF-DM patients who had had a second HbA1c recorded before the most recent measure (median 298 [IQR 230-379] days). Over a year prior to the match date, 16.8% of HF-DM with two HbA1c measures experienced an absolute increase of >1% HbA1c and 19.3% experienced an absolute decrease of >1% HbA1c. A greater than1% decrease in HbA1c was more prevalent in cases (22%) than controls (18.6%). Within the HF-DM group, 14% of patients experienced a change in drug category in the year before their match date. Of the cases 83% had stable drug regimens compared to 86.7% of the controls but a decrease in drug category was more prevalent in the cases (7.3%) than controls (3.3%).

Most (81%) of patients who had a decrease of at least one drug category were on no DM drugs at the time of matching.

Comorbid CKD group characteristics: Due to the low prevalence of CKD defined by diagnostic codes in patients entering the cohort before April 2006, CKD was defined by eGFR <60ml/min/m^2 which was recorded in 119,615 (89.5%) of the total HF cohort before the match date (median 101 [IQR 37-223] days). CKD was present in 66,301 (55.4%) of HF patients. Prevalence was higher in cases (66.2%) than controls (52.7%). Compared to HF patients without CKD, those with HF-CKD were more likely to be older (74 v. 80 years) and female (53% v. 38%) and less deprived. The comorbid group had lower haemoglobin, SBP and BMI than the non-comorbid group. There were fewer current smokers and alcohol drinkers in the HF-CKD group compared to the non-CKD group. Compared to the non-comorbid group, the comorbid group were less likely to be prescribed ACEi/ARB (74.7% v. 70.9%), more likely to be prescribed diuretics (84.9% v. 70.3%) and equally likely to be prescribed beta-blockers. Prevalence of concomitant COPD was less likely in the HF-CKD group than the non-CKD group, but concomitant DM was more likely (*P*<0.001) (see E-Appendix A29).

CKD severity: CKD severity categories based on eGFR level in the HF-cohort (Figure 7.6) showed different associations with mortality. The lowest three eGFR severity categories (<45ml/min/m2) were more prevalent in the cases than controls (40.7% v. 24.1%), whereas the highest three eGFR categories (45-105ml/min/m^2), were more prevalent in controls (73.8%) than cases (57%) (Table 7.6). eGFR >105 ml/min/m^2 was slightly more prevalent in cases than controls (2.3% v. 2.1%).

CKD severity change: Of the HF cohort, 104,884 (78.5%) had had a second eGFR measure before the most recent measure (median 315 [IQR 232-420] days). Of the cohort, 12.6% had experienced >15mls decrease in eGFR and 13.5% experienced >25% decrease in eGFR, over a year before their match date. An increase in eGFR was experienced in approximately 41% of patients over the same time frame. Decrease in eGFR but not an increase was associated with mortality. Mortality cases had 6.9% higher prevalence of decreasing eGFR (mls) and 10.5% higher prevalence with >25% decrease in eGFR than controls.

7.4 Chapter summary

The characteristics of people with HF in the general practice population have changed over the last decade and despite more prescribed medications and improved risk profiles in the more recent years, a quarter of people with HF still die in the first year following diagnosis and about 10% per year after that. The unadjusted associations between some the patient characteristics and comorbidity exposures measured at baseline and death differed according to the length of follow-up time. The indications are that the effects of exposure change over time (time-varying effects) or the exposures themselves change (time-varying exposures). DM, COPD and CKD were associated with death in each of the different follow-up periods but there was less difference in prevalence between the cases and controls for comorbidities that developed after HF incident, than before HF (for COPD, DM). All measures of comorbidity severity and longitudinal change that were recorded in the recent time-windows to death differed between cases and controls. These findings will be discussed in detail in Chapter 13.

Tables

Table 7.1 Baseline characteristics of the HF sample by mortality outcome over maximum 12 year follow-up

Factors and exposures	All (n=50,144)	Dead (n=26,729)	Alive (n=23,385)	*P*
Person and socio-demographic factors				
Age, years	78[71-85]	81[74-87]	75[66-82]	<0.001
Women	23,595 (47.1)	12,974(48.5)	10,621(45.4)	<0.001
IMD quintile				<0.001
1	5,846 (19.5)	3,064(18.7)	2,782(20.5)	
2	6,952 (23.2)	3,708(22.6)	3,244(23.9)	
3	6,340 (21.1)	3,570(21.8)	2,770(20.4)	
4	6,235 (20.8)	3,427(20.9)	2,808(20.6)	
5	4,613 (15.4)	2,613(16)	2,000(14.7)	
Anthropometric and clinical factors				
BMI (Kg./m^2)	27[23.8-31]	26.2[23.1-30]	27.9[24.7-32]	<0.001
Cholesterol (mmol/L)	4.7 ±1.2	4.6±1.2	4.7±1.2	<0.001
Hb (g/dL)	13.1 ±1.9	12.7±1.9	13.4±1.8	<0.001
Systolic BP (mmHg)	137.3 ±21.5	137.2`±22.1	137.2±20.8	0.48
Diastolic BP (mmHg)	76.7 ±12	75.8±11.9	`77.7±12	<0.001
Lifestyle factors				
Smoking status				<0.001
yes	6,841 (14.2)	3,798(15)	3,043(13.3)	
No	22,618 (46.9)	11,898(46.8)	10,720(47)	
Ex	18,757 (38.9)	9,703(38.2)	9,054(40)	
Alcohol status				<0.001
Yes	31,228 (70.9)	15,822(68.9)	15,406(73)	
No	11,093 (25.2)	6,229(27.1)	4,864.0(23)	
Ex	1,746 (4.0)	915(4)	831(3.9)	
Drug factors				
Diuretics	32,071(64.0)	19,311(72.2)	12,760(54.5)	<0.001
Beta blocker	16,382 (32.7)	7,933(29.7)	8,449(36.1)	<0.001
ACEi	22,582 (45.1)	11,229(42)	11,353(48.6)	<0.001
ARB	6,118 (12.2)	2,849(10.7)	3,269(14)	<0.001
Comorbidity exposures				
Diabetes Mellitus (DM)	10,533 (21.0)	5,883(22.0)	4,650(19.9)	<0.001
HbA1c (%)	7.1[6.4-8.1]	7.1[6.4-8.1]	7.1[6.4-8]	0.89
COPD	5,848 (11.7)	3,640(13.6)	2,208(9.4)	<0.001
FEV_1 (pp)	53.3[39-69.6]	52.8[38.2-69.2]	59[46-72]	<0.001
Renal disease (medical code)	7,621 (15.2)	4,194(16)	3,427(14.7)	<0.001
Renal disease (eGFR<60)	20,084 (49.8)	12,132(57)	7,952(41)	<0.001
eGFR (ml/min/1.73m^2)	61.0 ±20.3	57.4±20.7	64.9±19.3	<0.001
Number of listed comorbidities				<0.001
0	13565 (33.6)	5,623(26.7)	7,942(41.3)	
1	19471 (48.3)	10,916(51.7)	8,555(44)	
2	6,795 (16.8)	4,224(20)	2,571(13.4)	
3	515 (1.3)	339(1.6)	176(0.9)	

Data are number patients (%) or mean± standard deviation or median[IQR]. IMD, index multiple deprivation (1=least deprived, 5= most deprived); BMI, body mass index; Hb, haemoglobin; BP, blood pressure; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; HbA1c, glycated haemoglobin; COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 second; pp, percent predicted; eGFR, estimated glomerular filtration rate. Number of comorbidities is based on those listed (COPD, DM, CKD defined by eGFR). HbA1c and FEV_1 were measured in the respective DM and COPD groups only. eGFR was measured in all HF patients.

Table 7.2 Death events per year of follow-up

Follow-up year	Number entering follow-up year	Died during follow-up year n(%)	Censored during follow-up year n(%)	Alive at end of follow-up year
1	50114	11678(23.2)	2298(4.6)	36138
2	36138	4269(11.8)	2494(6.9)	29375
3	29375	3247(11.1)	3851(13.1)	22277
4	22277	2370(10.6)	3183(14.3)	16724
5	16724	1710(10.2)	2687(16.1)	12327
6	12327	1231(10)	2192(17.8)	8904
7	8904	848(9.5)	1771(19.9)	6285
8	6285	633(10.1)	1493(23.8)	1459
9	4159	384(9.2)	1147(27.6)	2628
10	2628	221(8.4)	1029(39.2)	1378
11	1378	116(8.4)	754(54.7)	508
12	508	22(4.3)	486(95.7)	486

Censored data included subjects who (i) belonged to a practice that stopped contributing data to the CPRD during the follow-up year (ii) were transferred out of the practice (iii) reached the study end.

Table 7.3 Time-matched general characteristics of the matched sample by mortality outcome over maximum 12 years follow-up

Patient characteristics	All (n=133,645)	Cases (n=26,729)	Controls (n=106,916)	Missing n(%)
Person and socio-demographic factors				
Age, years	77[IQR 69-83]	81[74-87]	76[68-82]	-
Women	61,732(46.2)	12,974(48.5)	48,758(45.6)	-
IMD quintile				54262(40.6)
1	15,908(20.0)	3,064(18.7)	12,844(20.4)	
2	18,089(22.8)	3,708(22.6)	14,381(22.8)	
3	16,666(21)	3,570(21.8)	13,096(20.8)	
4	16,451(20.7)	3,427(20.9)	13,024(20.7)	
5	12,269(15.5)	2,613(16.0)	9,656(15.3)	
Anthropometric and clinical factors				
BMI (Kg/m2)	26.9[23.5-31]	25.4[22.1-29.3]	27.3[24-31.3]	-
Cholesterol (mmol/L)	4.5±1.2	4.4±1.2	4.5±1.2	-
Hb (g/dL)	13.0±1.9	12.3±2.0	13.1±1.8	-
Systolic BP (mmHg)	131.4±20.5	126.9±22.3	132.5±19.9	-
Diastolic BP (mmHg)	73.3±11.4	71.1±12	73.9±11.1	-
Lifestyle factors				
Smoking status				-
yes	15,002(11.2)	3,127 (11.7)	11,875 (11.1)	-
No	61,936(46.3)	12,573 (47.0)	49,363 (46.2)	-
Ex	56,707(42.4)	11,029(41.3)	45,678(42.7)	-
Alcohol status				-
Yes	92,438(69.2)	17,740(66.4)	74,698(69.9)	-
No	34,755(26.0)	7,611(28.5)	27,144(25.4)	-
Ex	6,452(4.8)	1,378(5.2)	5,074(4.8)	-
Drug factors				
Diuretics	103,283(77.3)	21,574(80.7)	81,709(76.4)	-
Beta blocker	74,221(55.5)	12,171(45.5)	62,050(58)	-
ACEi	74,373(55.7)	12,207(45.7)	62,166(58.1)	-
ARB	22,753(17.0)	3,170(11.9)	19,583(18.3)	-
ACEi or ARB	94,547(70.7)	15,079(56.4)	79,468(74.3)	-
Beta blocker AND (ACEi or ARB)	58,632(43.9)	7,981(29.9)	50,651(47.4)	-
Comorbidity exposures				
Diabetes	31,962(23.9)	6714(25.1)	25,248(23.6)	-
COPD	18,478(13.8)	4,630(17.3)	13,848(13.0)	-
Renal disease (eGFR <60)	66301(55.4)	15827(66.2)	50474(52.7)	14030(10.5)
Number of comorbidities				14030(10.5)
0	34607(28.9)	4766(19.9)	29841(31.2)	
1	58711(49.1)	12512(52.3)	46199(48.3)	
2	24025(20.1	5964(24.9)	18061(18.9)	
3	2272(1.9)	673(2.8)	1599(1.7)	

Data are number patients (%) or mean±standard deviation or median[IQR]. IMD, index multiple deprivation (1=least deprived, 5=most deprived); BMI, body mass index; Hb, haemoglobin; BP, blood pressure; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate. Number of comorbidities is based on those listed (COPD, DM, CKD defined by eGFR.

Table 7.4 Time-dependent COPD comorbidity measures of the matched sample by mortality outcome

Comorbidity measures	HF All (n=133,645)	HF cases (n=26,729)	HF Controls (n=106,916)	Missing n(%)
COPD diagnostic code (n=18,478)	18,478(13.8)	4,630(17.3)	13,848(13.0)	
Diagnosis before HF (78.4%)	14,483(10.8)	3,640(13.6)	10843(10.1)	-
Diagnosis after HF (21.6%)	3,995(3.0)	990(3.7)	3005(2.8)	-
COPD severity in COPD group only	All COPD (n=18,478)	Cases (n= 4,630)	Controls (n=13,848)	
FEV_1(percent predicted)	54.1[IQR 39.7-70.0]	51.9 [37.7-68.5]	55.2 [40.6-70.6]	9,963(53.9)
COPD GOLD severity stage				9,963(53.9)
1: FEV_1 ≥80% normal	1284(15.1)	396(13.9)	888(15.7)	
2: FEV_1 50-79% normal	3721(43.7)	1140(39.9)	2581(45.6)	
3: FEV_1 30-49% normal	2679(31.5)	989(34.6)	1690(29.9)	
4: FEV_1 <30% normal	831(9.8)	334(11.7)	497(8.8)	
COPD Drug severity				-
No drugs	2,359(12.77)	561 (12.12)	1,798(12.98)	
Short term inhalers only	1624(8.79)	380 (8.21)	1244(8.89)	
Monotherapy	2417(13.08)	542 (11.71)	1875(13.54)	
Dual therapy	3329(18.02)	664 (14.34)	2665(19.24)	
Triple therapy	3140(16.99)	699 (15.0)	2441(17.63)	
Any inhalers only	10,510(56.9%)	2,285(49.4)	8,225(59.39)	
No steroids or oxygen	12,869(69.6)	2846(61.5)	10,023(72.4)	
Oral steroids but no oxygen	4,744(25.7)	1415(30.6)	3329(24.0)	
On oxygen	865(4.7)	369(8.0)	496(3.6)	
COPD severity change in COPD group only				
Gold stage same or better	3790 (81)	1,326(78.2)	2,464(82.5)	13796(74.7)
Gold stage worse (at least one stage)	892(19.1)	369(21.8)	523(17.5)	
10% change				
<10% change	3,341(71.4)	1191(70.3)	2150(72.0)	13796(74.7)
≥10% increase (better)	631(13.5)	218(12.9)	413(13.8)	
≥10% decrease (worse)	710(15.2)	286(16.9)	423(14.2)	
5% change				
<5% change	2,091(44.7)	739(43.6)	1,352(45.2)	
≥5% increase (better)	1,166(24.9)	392(23.1)	774(25.9)	
≥5% decrease (worse)	1,425(30.4)	564(33.3)	861(28.8)	

COPD severity change (Drugs)				
Drug category same or better	10,908 (59.0)	2,703 (58.4)	8,205 (59.3)	
Drug category worse	7, 570 (41.0)	1,927 (41.6)	5,643 (40.8)	
No new steroids or oxygen	15,079(81.6)	3508(75.8)	11571(83.6))	
New on steroids but no new oxygen	2,885(15.6)	889(19.2)	1996(14.4)	-
New on oxygen	514(2.8)	233(5.0)	281(2.0)	-

COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 second; pp, percent predicted. Most recently recorded FEV_1 (pp) before the match date was used to measure COPD severity within a maximum of a 3-yr time window. % change for FEV_1 was calculated as the absolute change in pp over one year using the most recent measure and an earlier FEV_1 measure recorded up to a maximum of 3-years before the most recent measure. The FEV1 change was adjusted for the time interval. Mono, dual or triple therapy; one, two or three of respectively: long acting beta2-antagonist, long acting cholinergic, methylxanthines, inhaled steroid either individually or in combination inhalers.

Table 7.5 Time-dependent DM comorbidity exposures of the matched sample by mortality outcome

Comorbidity measures	HF All (n=133,645)	HF cases (n=26,729)	HF Controls (n=106,916)	Missing n(%)
Diabetes (DM) diagnostic code (n=31,962)	31, 962 (23.9)	6714(25.1)	25,248(23.6)	-
Diagnosis before HF (85%)	27,175(20.3)	5,883(22)	21,292(19.9)	-
Diagnosis after HF (15%)	4,787(3.6)	831(3.1)	3,956(3.7)	-
Diabetes severity (HbA1c) in DM group only	All Diabetes (n=31,962)	DM Cases (n= 6,714)	DM Controls (n=25,248)	
HbA1c (%)	7.1[IQR 6.4-8.1]	7 [6.3-8]	7.1[6.4-8.1]	3097(9.7)
<5.5%	1275(4.4)	357(6.0)	918(4.0)	
5.5-6.4%	6621(22.9)	1476(24.7)	5145(22.5)	
6.5-7.5%	10379(36.0)	2056(34.4)	8323(36.4)	
7.6-8.5%	5122(17.7)	1010 (16.9)	4112(18.0)	
8.6-9.5%	2581(8.9)	520(8.7)	2061(9)	
>9.5%	2887(10)	554(9.3)	2333(10.2)	
Diabetes drug severity				
1: None	8,452(26.4)	2 041(30.4)	6,411(25.4)	-
2: Oral only	14,700 (46.0)	2,835 (42.2)	11,865 (47.0)	
2: Oral (+Insulin)	3,518(11.0)	545(8.1)	2,973(11.8)	-
3: Insulin only	5,292(16.6)	1 293(19.3)	3,999(15.8)	-
Diabetes severity change (HbA1c) in DM group only				
<1% change	16,706(63.8)	3380(61.8)	13326(64.4)	5788(18.1)
>1% increase	4,407(16.8)	889(16.2)	3518(17)	
>1% decrease	5,061(19.3)	1205(22.0)	3856(18.6)	
Diabetes severity change (Drugs)				-
No drug category change	27,466(85.9)	5575(83.0)	21891(86.7)	-
Increase in drug category	3,177(9.9)	647(9.6)	2530(10.0)	-
Decrease in drug category	1,319(4.1)	492(7.3)	827(3.3)	-

HbA1c, glycated haemoglobin; Most recently recorded Hba1c (%) prior to the match date was used to measure DM severity within a maximum of a 3-yr time window. % change for HbA1c was calculated as the absolute change in % over one year using the most recent measure and an earlier HbA1c measure recorded up to a maximum of 3-years prior to the most recent measure. The HbA1c change was adjusted for the time interval.

Table 7.6 Time dependent CKD comorbidity measures of the matched sample by mortality outcome

CKD measures	HF all (n=133,645)	Cases (n= 26,729)	Controls (n=106,91)	Missing n(%)
Renal disease diagnosis				
eGFR <60	66301(55.4)	15827(66.2)	50474(52.7)	14030(10.5)
Renal severity (eGFR)				
eGFR mL/min/1.73m^2	57.8±21.2	52±22.9	59.3±20.5	14030(10.5)
Renal severity by eGFR stage*				14030(10.5)
1: >105	2,591(2.2)	558(2.3)	2033(2.1)	
1: 90-105	6,505(5.4)	1028(4.3)	5477(5.7)	
2: 60-89	44,218(37.0)	6502(27.2)	37716(39.4)	
3A: 45-59	33,510(28.0)	6091(25.5)	27419(28.7)	
3B: 30-44	23,130(19.3)	5890(24.6)	17240(18.0)	
4: 15-29	8,252(6.9)	3159(13.2)	5093(5.3)	
5: <15	1,409(1.2)	687(2.9)	722(0.8)	
eGFR severity change				28761(21.5)
Classification 1 (Absolute change)				
0mls-5mls decrease (reference group)	26,056(24.8)	4522(21.3)		
>15mls decrease	13,238(12.6)	3848(18.1)	21534(25.8)	
6mls to 15mls decrease	22,817(21.8)	4690(22.0)	9390(11.2)	
any increase	42,733(40.8)	8218(38.6)	18127(21.7)	
Classification 2 (Percentage change)				34555(41.3)
0-5% decrease (reference group)	15,244(14.5)	2262(10.6)		
>25% decrease	14,119(13.5)	4639(21.8)		
6-25% decrease	31,729(30.3)	5954(28.0)	12982(15.5)	
Any % increase	43,792(41.8)	8423(39.6)	9480(11.3)	

eGFR, estimated glomerular filtration rate. * National Kidney Foundation Kidney Disease Outcomes Quality Initiative (KDOQI) guidelines. >105 group added to stage one due to the prior evidence of increased risk in high eGFR
% change was calculated as the absolute change in eGFR over one year and the relative change (difference in eGFR over a year as a proportion of the first eGFR measure).

Figures

Figure 7.1 All-cause mortality as a percentage of the total HF survivors in each year of follow-up

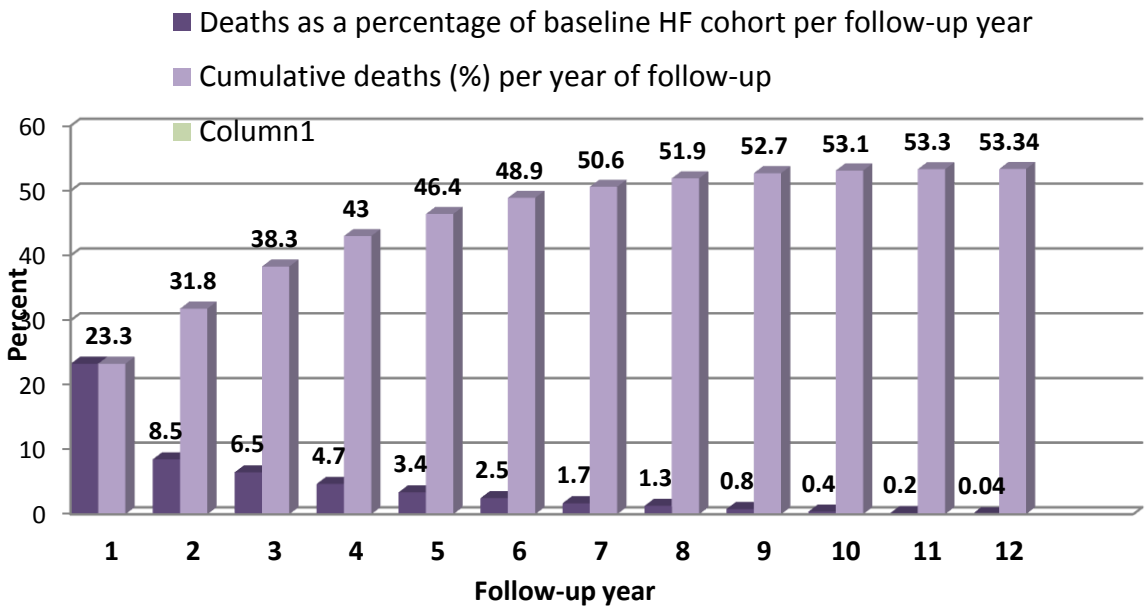


Figure 7.1: The dark purple bars show the percentage of the baseline cohort who died in each year of follow up. The denominator in this figure is the 50,114 subjects in the baseline cohort. The light purple bars show the cumulative percentage of the baseline cohort who died over the 12 years of follow-up.

Figure 7.2 All-cause mortality as a percentage of remaining HF survivors per follow-up year

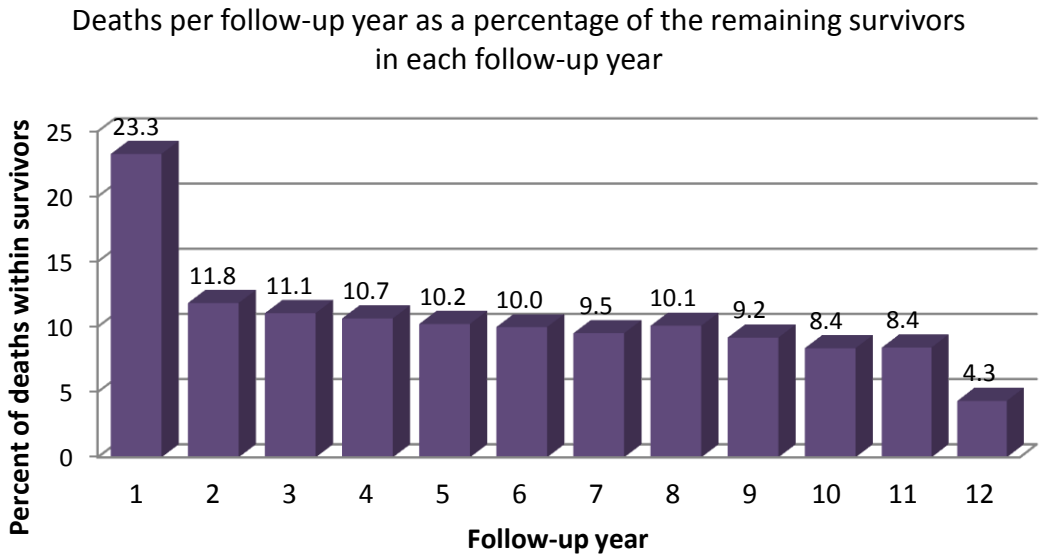


Figure 7.2: The purple bars show the percentage of deaths in the remaining survivors who entered each year of follow up. In year 12 most subjects were censored due to the practice not contributing data to the final CPRD data capture (see Table 7.2).

Figure 7.3 Comorbid COPD diagnosed pre and post HF by mortality outcome

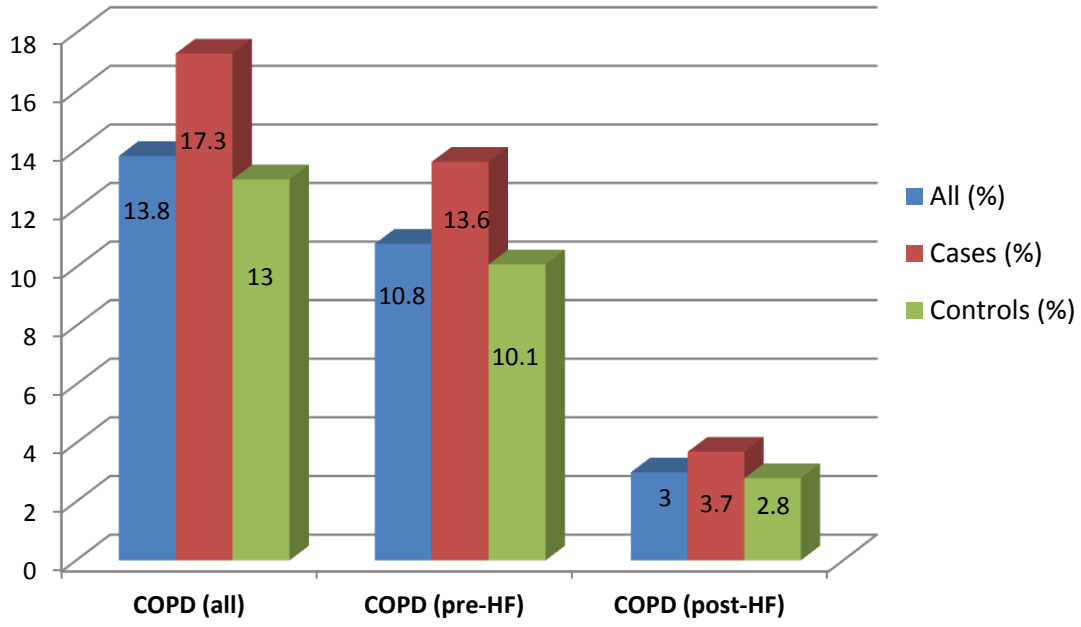


Figure 7.4 Comorbid COPD GOLD severity stages by mortality

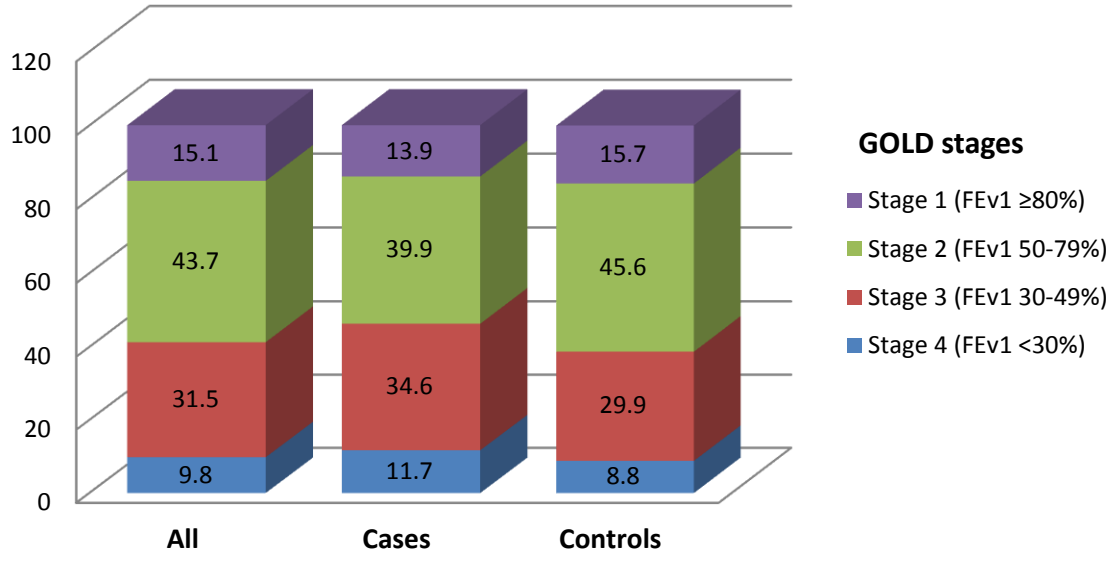


Figure 7.5 Comorbid DM diagnosed pre and post HF by mortality outcome

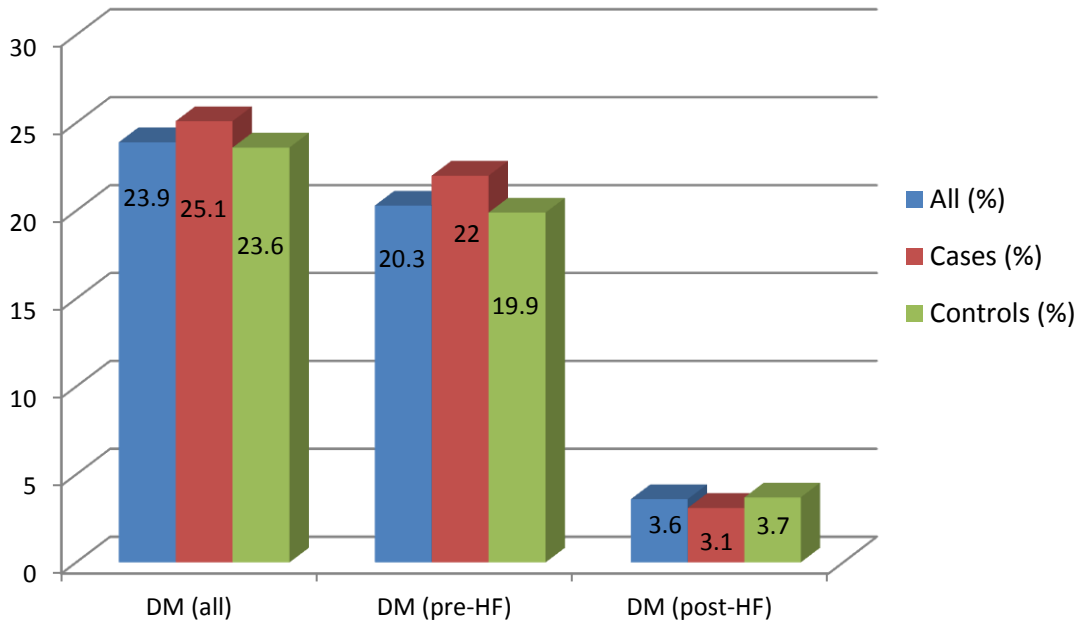
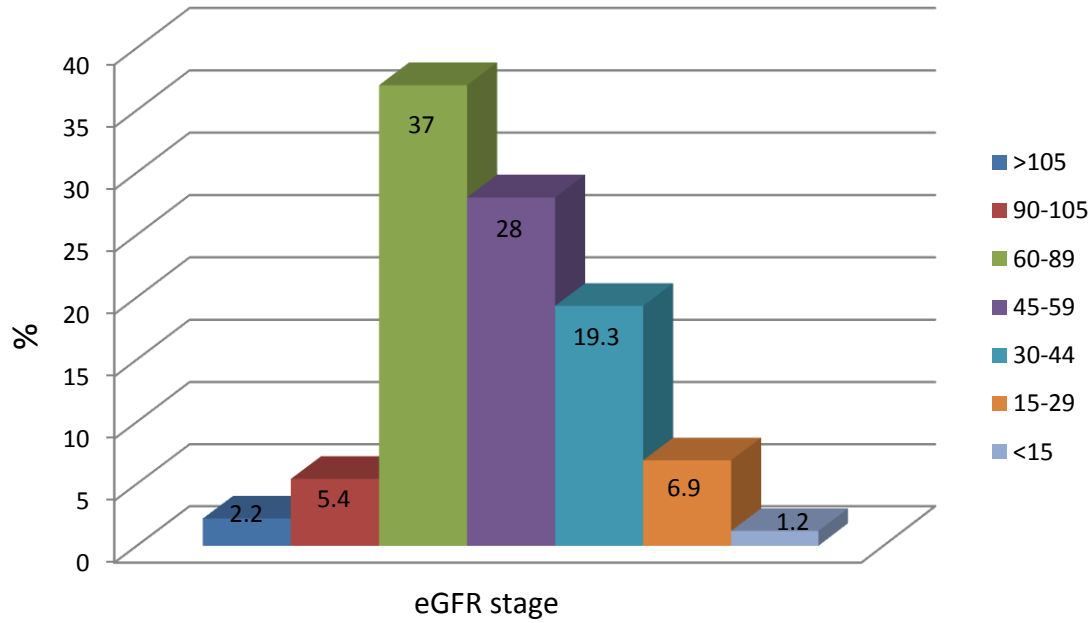


Figure 7.6 Comorbid renal stages based on eGFR in the total HF-cohort



Chapter 8 Non-CVD comorbidity prognostic factors in an incident HF general practice cohort sample by first hospital admission

This chapter presents the results of a sub-sample of the main incident HF CPRD cohort linked to hospital episodes statistics (HES) data. The HF sample linked to hospital admissions (HF-HA) was taken from the main sample of HF patients aged 40 years and over, with a first HF consultation between 1st January 2002 and 1st March 2012 followed up until January 2014. This chapter describes the HF-HA sample and investigates the unadjusted associations between non-cardiovascular comorbidities at baseline and as they develop over time and first hospital admission, but also the other patient, social, lifestyle, risk and medication factors that may influence this outcome. Baseline characteristics are described in the sample using the available data including missing values. Unless stated all baseline associations were significant at a level of $p<0.01$. For the matched sample, the patient numbers reported for the cases and controls refers to the number of observations generated by the matching process. Following imputation, differences in time-dependent characteristics between the cases and controls are described. Results are presented as number (%), mean (SD) or median [Interquartile range; IQR].

8.1 Study sample: baseline characteristics of total HF-HA sample by first hospital admission outcome

Sample and follow-up: There were 30,061 participants with a first HF consultation code between 1st January 2002 and 1st March 2012 who had linked HES data. Follow-up ranged from 0 to 12 years (median 82 [IQR 12-435] days). The risk of first hospital admission following HF diagnosis expressed as a percentage of the total cohort at baseline decreased rapidly with each month of follow-up time, shown in Figure 8.1 (dark purple), with most events occurring in the shorter follow-up time. There were 24,538 (81.6%) first hospital admissions from any cause over the 12-year follow-up, with the risk of admission in the first 1-month (33%), first 3 months (46%) and within 1 year (64.6%). Figure 8.1 also shows the cumulative risk of first admission per month of

follow-up (light purple). The risk of first admission in the remaining HF patients not admitted in each month of follow-up also decreased steeply with time shown in Figure 8.2.

8.1.1 Patient characteristics at baseline in the HF-HA population sample by first hospital admission

Baseline age at HF diagnosis was negatively skewed with a higher number of older participants; mean 77.3 (SD 10.9) years and median 79 [IQR 71-85] years (E-Appendix E15 for all continuous variable histograms and Q-Q plots). Over the total follow-up time, compared to those not-admitted, people admitted were similar in age and gender but significantly more deprived (see Table 8.1). The sample was 47.1% female with a decreased proportion of women in those admitted compared to those not-admitted (46.8% v. 48.5%) although this difference was non-significant. Of those admitted, 39.9% had experienced a previous admission within 3-months of their HF index date, compared to 35% of those not admitted and there were more current smokers (13.5% versus 12.6%).

Body mass index (BMI) was positively skewed with more participants with lower BMI. Haemoglobin, cholesterol and blood pressure were approximately normally distributed (E-Appendix E15). At baseline those admitted, compared to those not admitted, had higher systolic blood pressure (137.6 (21.5) v. 136.4(21) mmHg) and lower haemoglobin (13.0 (1.9) v. 13.2 (1.8) g/dL). All other associations were non-significant. Median BMI in the HF-HA sample was 26.9 [IQR 23.7-30.8] Kg/m2 and mean cholesterol 4.7 (1.2) mmol/ml.

Of the total sample, 32.2% were prescribed beta-blockers at baseline, 55.7% were prescribed an Angiotensin-converting-enzyme inhibitor (ACEi) or an Angiotensin receptor blocker (ARB) and 64.4% were prescribed diuretics. HF patients admitted compared to those not admitted, were significantly more likely to be prescribed diuretics (65.6% v. 59.2%). Other baseline medications were similar in those who had been admitted and not admitted.

8.1.2 Comorbidity exposures at baseline in the HF-HA sample by first hospital admission

At HF index date, 3,504 (11.7%) of the sample had chronic obstructive pulmonary disease (COPD), 6,298 (21%) had diabetes mellitus (DM) and 4,510 (15.0%) had chronic kidney disease (CKD). When CKD was defined by an estimated glomerular filtration rate (eGFR) <60 ml/min/m^2, prevalence at baseline increased to 12,687 (50.1%). With the exception of diagnostic CKD, prevalence of all comorbidities was significantly higher in the patients experiencing first hospital admission compared to those not admitted; COPD (12.1% v. 9.7%), DM (21.8% v. 17.3%) and CKD (eGFR: 51.4% v. 44.5%) (E-Appendix B16). Diagnostic CKD was non-significant (14.9% v. 15.4%). HF patients admitted were also more likely to have ≥2 coexisting comorbidities (18.8% v. 13.3%). The most prevalent disease combination was DM and CKD (eGFR defined) (12.7%) and this predominated in those who were admitted (13.3% v. 9.7%) (E-Appendix B17).

In the comorbid DM sample the median glycated haemoglobin (HbA1c) at baseline was 7.1 [IQR 6.4-8.1] % which was higher in those with an admission (7.1 [6.4-8.1] v. 6.9 [6.3-7.8] %). In the comorbid COPD, sample the median forced expiration volume in 1 second (FEV1) was 53.5 [38.8-70.7] percent predicted (pp) which was significantly lower in those admitted than not (53.3 [38.6-70.6] v. 54.8 [41-72] pp). eGFR was measured in 84.3% of the baseline HF sub-sample and was mean 60.8 (20.3) ml/min/m^2. This level was significantly lower in those admitted than not (60.2 (20.1) v. 64.5 (20.6) ml/min/m^2).

8.2 Study sample: baseline characteristics by first hospital admission outcome of HF sub-samples with different follow up periods

The total sample will now be reported in 3 overlapping subgroups defined by their follow-up time. As for the mortality outcome, this was done to investigate whether the unadjusted associations between baseline factors and first hospital admission in the total sample were confounded by variations in the exposure time between those admitted and those not admitted (resulting from different follow-up times within the sample). Given the significant associations between the baseline factors with the timing of the HF index date for mortality (before or after April 2006) it was not necessary to repeat this investigation for the hospital admission sub-sample.

8.2.1 HF sub-samples with different follow-up periods: patient characteristics at baseline by hospital admission outcome

The cohort was grouped into three overlapping HF-HA sub-samples with different follow-up periods (i) at least up to 1 month (whole sample), (ii) between 1 month and 1 year and (iii) ≥1 year. HF patients with their first admission and those without were compared at the end of each follow-up period (see Table 8.2). The unadjusted associations between the baseline characteristics and first hospital admission outcome were observed over the total follow-up time and by three follow-up time-periods (see E-Appendix A30). Both the significance and magnitude of the difference in exposure status between those admitted and not admitted differed across the three follow-up periods for some exposures.

Age and prescription of beta-blockers had a significant association with hospital admission in the first and longest follow-up periods but there was a lack of association with medium term follow-up. Females and higher deprivation were only significantly associated with hospital admission in the medium term follow-up (1 month to 1 year). Other baseline factors varied in the nature of their associations with first hospital admission. Lower mean baseline SBP and cholesterol were associated with admission in the medium term follow-up period but this reverted to higher mean baseline SBP and cholesterol levels in the longer follow-up period. There was no association between these factors and admission in 1-month follow-up period. Prescription of diuretics was significantly associated with lower likelihood of admission in the 1-month follow-up period but with higher admission in the two longer follow-up periods. The magnitude of the difference in exposure status between those admitted and not admitted became less over time for prior hospital admission within three months and for haemoglobin, BMI, alcohol drinkers and prescription of ACEi. These factors had significant associations with first hospital admission in the 1-month follow-up period but most lost significance over longer follow-up time.

8.2.2 HF sub-samples with different follow-up periods: comorbidity exposures at baseline by hospital admission outcome

All three baseline comorbidities were significantly associated with increased first hospital admission. The magnitude of difference in comorbidity exposure for COPD and DM between those admitted and not admitted

reduced with longer follow-up and became non-significant.in the longest follow-up period. Diagnostic CKD was associated with admission in the follow-up periods up to 1 year, but switched to non-admission for the longest follow-up period. CKD defined by eGFR was consistently associated with admission and the magnitude of difference in CKD prevalence between those admitted and not admitted more than doubled in the longer follow-up period from 4% to 10% difference (49.6% admitted v. 39.6% non-admitted in follow-up > 1 year).

In the comorbid DM group, HbA1c level was similar in those admitted and not admitted within the first month of follow-up, but higher HbA1c level became significantly associated with admission in the 12-month follow-up period (median 7.2 [IQR 6.5-8.2] % in those admitted v. 6.9 [6.3-7.7] % in those not-admitted). FEV_1 in the comorbid COPD group (54.1 [39.6-71] pp) was similar in those with and without admission. eGFR was consistently and significantly lower in those admitted than those not admitted across all three follow-up periods.

8.3 Study sample: Time-dependent characteristics of the matched HF-HA sample by first hospital admission outcome

From the total HF-HA population sample, 24,339 all-cause first hospital admission cases over the total follow-up period were matched to up to four controls. Controls were live, active in practice, remained at risk of first hospital admission on the case index date and were not themselves admitted for the next 3-months. Controls were matched to cases within 1-month of their HF diagnosis study-in date and on duration of follow-up. This resulted in 110,789 observations including 86,450 controls. The match date for the controls was the date that corresponded with the same amount of follow-up time as the case. Person, socio-demographic factors, lifestyle factors and risk factors were measured as the most proximal to the match date during follow-up. Drug and comorbidity exposures were measured in specific study time-windows as outlined in chapter 6 (Section 6.3.4.5).

8.3.1 Patient characteristics in the matched HF population sample by first hospital admission

Median current age on the match date for the hospital admission cases was 79 [IQR 72-85] years which was higher than in the controls (78 [70-84] years). Cases were less likely to be female and more deprived (Table 8.3). The matched sample was 48.6% female with a decrease in the proportion of women in the cases compared to controls (46.8% versus 49.1%). The differences in prevalence of current smokers and alcohol drinkers between cases and controls were significant. The proportion of smokers in the case group was 12.4% compared to 10.4% in the controls. Current alcohol drinkers were less prevalent in the cases than controls (71.9% v. 74.2% respectively). Cases were five times more likely to have had a prior hospital admission in the previous 3 months than controls (20.9% v. 4.1% respectively).

Cases compared to controls, had lower recent systolic blood pressure (mean 134.1 (21.5) v. 136 (19.8) mmHg), diastolic blood pressure (74.7 (12) v. 75.9 (11) mmHg), BMI (median 26.7 [IQR 23.5-30.6] v. 27.2 [23.9-31.2] Kg/m^2), cholesterol (4.66 (1.2) v. 4.73 (1.2) mmol/L) and haemoglobin (12.9 (1.9) v. 13.4 (1.6) g/dL).

Of the total sample, 40.1% were prescribed beta-blockers on the match date, 72.6% were prescribed an ACEi or an ARB and 72.1% were prescribed diuretics. Cases compared to controls, were less likely to be prescribed beta-blockers (36.5% v. 41.2%), ACEi *or* ARB (64.3% v. 74.9%) or diuretics (69.9% v. 72.7% respectively).

8.3.2 Comorbidity factors in the matched HF-HA population sample by hospital admission

Comorbid COPD group characteristics: COPD was present in 11,903 (10.7%) HF patients before the match date, and prevalence was higher in cases (13.3%) than controls (10.3%). Compared to HF patients without COPD, those with HF-COPD were more likely to be younger and more deprived but less likely to be female. The HF-COPD group had lower blood pressure, BMI and higher haemoglobin than the non-COPD HF group. In the HF-COPD group, 21.3% were current smokers compared to only 9.6% in the non-COPD HF group. Compared to the non-COPD HF group, the HF-COPD group were less likely to be prescribed beta-blockers

(20.1% v. 42.6%) but more likely to be prescribed diuretics (75.8% v. 71.6%). Prevalence of concomitant diabetes and CKD were less likely in the HF-COPD group who had a higher mean eGFR (62.0 (SD 20) v. 59.2 (19.3) ml/min/m^2) (E-Appendix A31).

COPD status: Of the 11,903 HF patients with COPD, 1,060 (8.9%) developed COPD after HF diagnosis. COPD that developed before or after HF was more prevalent in the hospital admission group but the magnitude of difference between cases and controls was greater for the pre-HF than the post-HF COPD group (3.0% v. 0.3%) (Table 8.4 and Figure 8.3).

COPD severity: Only a small proportion of HF-COPD patients had a FEV_1 recorded within 6 months of admission (1,872; 15.7%). Within the HF-COPD group, median FEV_1 was 50pp [38-69] which was similar in the cases and controls. HF-COPD categorised into GOLD severity stages was not associated with hospital admission. Of the HF-COPD patients, 23.7% had been prescribed oral steroids before admission which was higher in the cases (28.6%) than controls (21.8%). Cases were also more likely than controls to be prescribed oxygen therapy prior to admission (5.1%% v. 2.1% respectively) (Figure 8.4).

COPD severity change: There were only 653 (5.5%) HF-COPD patients who had a second FEV_1 (pp) recorded in the severity measurement time-windows. Whilst there were a higher proportion of cases than controls experiencing a worsening of GOLD stage (21.8% v. 16.7%) over the 6-months prior to the match date this difference was non-significant. There were 1,171 (9.8%) HF-COPD patients with newly prescribed oral steroids or oxygen therapy in the 4-months before hospital admission. New prescription of these drugs were lower in cases than controls for steroids (8.8% v. 9.5%) and higher for oxygen (0.8% v. 0.5%), but these associations were non-significant.

Comorbid DM group characteristics: Diabetes mellitus was present in 21,291 (19.2%) of HF patients prior to the match date which was higher in cases (22.9%) than controls (18.2%). Compared to HF patients without DM, comorbid patients were more likely to be younger and more deprived but less likely to be female. The comorbid group had lower cholesterol and haemoglobin, but higher BMI than the HF group without DM.

There were fewer current smokers and alcohol drinkers in the HF-DM group compared to the non-comorbid group. Compared to the non-comorbid group, the HF-DM group were more likely to be prescribed beta-blockers (45.9% v. 38.8%), ACEi/ARB (82.1% v. 70.3%) and diuretics (76.5% v. 71%). Prevalence of concomitant COPD was lower in the HF-DM group but concomitant CKD was more likely with a mean eGFR of 58.6 (SD 20.7) compared to 59.7 (19) ml/min/m^2 in the non-DM HF group (E-Appendix A32).

DM status: Of the 21,291 HF-DM group, 1,498 (7%) developed DM after HF index diagnosis. DM that developed before HF was more prevalent in the hospital admission group with 21.8% in the cases v. 16.8% in the controls. DM that developed after HF however was more prevalent in controls (1.4% v. 1.1%) (Table 8.5).

DM severity: Of the comorbid group, 14,578 (68.5%) had had at least one recent HbA1c recorded (median 73 [IQR 34-118] days). Around 38% of the HF-DM patients had had a most recent HbA1c between 6.5-7.5%. Median HbA1c was similar in the cases and controls (7.2 [6.4-8.2] v. 7.1 [6.4-8.1] %) but the lowest HbA1c category (<5.5%) and highest category (≥9.5%) were more prevalent in cases than controls.
Of the HF-DM group, 27.2% were not on any DM drugs at the time of matching and these patients were less prevalent in the admitted cases (24.9%) than the controls (28.1%). Most HF-DM patients were on oral medication (60%) +/- insulin, which was similar in the cases and controls (60.3% v. 59.9%). Of the comorbid group, 2,713 (12.7%) were prescribed insulin only which was a first hospital admission in 14.8% of the cases and 12.0% of controls.

DM severity change: There were 12,242 (57.5%) HF-DM patients who had had a second HbA1c recorded before the most recent measure (161 [IQR 101-212] days). Over 6-months prior to the match date, 12.7% of HF-DM with two HbA1c measures experienced an absolute increase of >1% HbA1c and 17.2% experienced an absolute decrease of >1% HbA1c. Both categories of HbA1c change were more prevalent in the cases than controls (P<0.05). Within the HF-DM group, 7.9% of patients experienced a change in drug category in the 6-months prior to their match date. Stable drug regimens were high in both cases and controls (90.6% v. 92.6% respectively), but more cases had an increase in drug category (6.4% v. 5.8%) or a decrease in drug category (3% v. 1.6%).

Comorbid CKD group characteristics: Due to the low prevalence of CKD defined by diagnostic codes in patients entering the cohort before April 2006, CKD was defined by eGFR <60ml/min/m^2 which was recorded in 72,399 (65.3%) of the total HF-HA cohort before the match date (median 47 [IQR 18-94] days). CKD was present in 37,784 (52.2%) of HF patients. Prevalence was higher in cases (57.9%) than controls (56%). Compared to HF patients without CKD, those with HF-CKD were more likely to be older (75 v. 81 years) and female (39.1% v. 55.7%). The comorbid group had lower haemoglobin and BMI and higher SBP than the non-comorbid group. There were fewer current smokers and alcohol drinkers in the HF-CKD group compared to the non-CKD HF group. Compared to the non-comorbid group, the comorbid group was less likely to be prescribed ACEi/ARB (77.7% v. 75.9%), more likely to be prescribed diuretics (67.8% v. 81.4%) and equally likely to be prescribed beta-blockers. Prevalence of concomitant COPD lower in the HF-CKD group than the non-CKD HF group, but concomitant DM was more likely (E-Appendix A33).

CKD severity: CKD severity categories based on eGFR level in the HF-cohort (Figure 8.6) showed different associations with first hospital admission outcome. The lowest three eGFR severity categories (<45ml/min/m2) were more prevalent in the cases (29.1%) than controls (20.7%), whereas the higher three eGFR categories (45-105ml/min/m^2), were more prevalent in controls (77.8%) than cases (69.2%). eGFR >105 ml/min/m^2 was slightly more prevalent in cases than controls (1.8% v. 1.7%).

CKD severity change: Two eGFR measures were available for 54,772 (49.4%) of the HF-HA sub-sample. The earlier of two measures was recorded 110 days [IQR 59-199] before the most recent measure. Of the cohort, 12.6% had experienced >15mls decrease in eGFR and 16.4% a >25% decrease in eGFR over the 6 months before their match date. An increase in eGFR was experienced in approximately 41.5% of patients over the same time frame. Decrease in eGFR but not increase was associated with first hospital admission. Cases had 4.6% higher prevalence of decreasing eGFR (>15mls) and 5.8% higher prevalence of >25% decrease in eGFR than controls.

8.4 Chapter summary

A third of people with HF in the general practice population are admitted to hospital in the first month following diagnosis and two thirds within a year. The unadjusted associations between baseline patient characteristics and comorbidity exposures and first hospital admission after HF differed by the follow-up time periods. Over the total follow-up time these unadjusted associations differed according to whether they were measured at baseline or before the event. The unadjusted associations in the shorter follow-up periods were most similar to the time-dependent associations measured closest to the event. There was a bigger difference in the prevalence of a previous hospital admission within 3 months between the cases and the controls when measured prior to match date rather than at baseline. Other patient characteristics measured before the match date showed associations with first hospital admission that were similar to mortality. All three selected comorbidities in HF were more prevalent in the hospital admission group than the controls but there was less difference for comorbidities that developed after HF than before HF. Comorbidity severity defined by physiological and drug measures differed between cases and controls and measures of recent severity change differed across the two groups for comorbid diabetes and CKD. These findings will be discussed in detail in Chapter 13.

Tables

Table 8.1 Baseline characteristics of the sample by first hospital admission outcome over total follow-up

Factors and exposures	All (n=30,061)	First hospital admission (n=24,538)	Not admitted (n=5,523)	*P*
Person and socio-demographic factors				
Age, years	79[IQR 71-85]	79[71-85]	79[70-86]	0.292
Women	14,163 (47.1)	11,485(46.8)	2,678(48.5)	0.024
IMD quintile				<0.01
1	5,844 (19.5)	4,698(19.2)	1,146(20.8)	
2	6,946 (23.2)	5,640(23.0)	1,306(23.7)	
3	6,340 (21.2)	5,181(21.1)	1,159(21.0)	
4	6,232 (20.8)	5,090(20.7)	1,142(20.7)	
5	4,612 (15.4)	3,859(15.7)	753(13.63)	
Anthropometric and clinical factors				
BMI (Kg./m^2)	26.9[IQR 23.7-30.8]	26.9[23.7-30.8]	27.0[23.8-31]	0.298
Cholesterol (mmol/L)	4.7 ±1.2	4.7±1.2	4.7±1.2	0.881
Hb (g/dL)	13.1 ±1.9	13.0±1.9	13.2±1.8	<0.001
Systolic BP (mmHg)	137.4 ±21.4	137.6`±21.5	136.4 ±21.0	<0.001
Diastolic BP (mmHg)	76.8 ±12	76.7±12.1	`76.9±11.6	0.503
Prior Hospital admission				<0.001
<3 months	11,719 (39)	9,779 (39.9)	1,940 (35.1)	
3-6 months	2,471 (8.2)	2,064 (8.4)	407 (7.4)	
>6months - 1 year	2,711 (9.0)	2,257 (9.2)	454 (8.2)	
Lifestyle factors				
Smoking status				<0.001
Yes	3,866 (13.3)	3,198(13.5)	668(12.6)	
No	13,614 (47.0)	10,963(46.3)	2,651(49.9)	
Ex	11,504 (39.7)	9,511(40.2)	1,993(37.5)	
Alcohol status				0.927
Yes	19,296 (72.8)	15,790(72.8)	3,506(72.9)	
No	6,152 (23.2)	5,043(23.2)	1,109(23.1)	
Ex	1,060 (4.0)	864(4)	196(4.1)	

Drug factors				
Diuretics	19,359(64.4)	16,089(65.6)	3,270(59.2)	<0.001
Beta blocker	9,683 (32.2)	7,877(32.1)	1,806(32.7)	0.390
ACEi	13,589 (45.2)	11,044(45.0)	2,545(46.1)	0.148
ARB	3,779 (12.6)	3,091(12.6)	688(12.5)	0.777
Comorbidity exposures				
Diabetes	6,298 (21.0)	5,345(21.8)	953 (17.26)	<0.001
HbA1c (%) (in diabetes only)	7.1[IQR 6.4-8.1]	7.1[IQR 6.4-8.1]	6.9[IQR 6.3-7.8]	<0.01
COPD	3,504 (11.7)	2,968(12.1)	536(9.7)	<0.001
FEV_1 (pp) (in COPD only)	53.5[IQR 38.8-70.7]	53.3[38.6-70.6]	54.8[IQR 41-72]	0.464
CKD (medical code)	4,510 (15.0)	3,662(14.9)	848(15.4)	0.419
CKD (eGFR<60)	12,687 (50.1)	10,616(51.4)	2,070(44.5)	<0.001
eGFR (ml/min/1.73m^2)	60.8 ±20.3	60.2±20.1	64.5±20.6	<0.001
Number of comorbidities				<0.001
0	8,518 (33.6)	6,629(32.1)	1,889(40.6)	
1	12,290 (48.5)	10,161(49.2)	2,129(45.7)	
2	4,202 (16.6)	3,601(17.4)	601(12.9)	
3	322 (1.3)	283(1.4)	176(0.9)	

Data are number patients (%) or mean± standard deviation or median[IQR]. IMD, index multiple deprivation (1=least deprived, 5=most deprived); BMI, body mass index; Hb, haemoglobin; BP, blood pressure; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; HbA1c, glycated haemoglobin; COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 second; pp, percent predicted; eGFR, estimated glomerular filtration rate. Number of comorbidities is based on those listed (COPD, DM, CKD defined by eGFR). HbA1c and FEV_1 were measured in the respective DM and COPD groups only. eGFR was measured in all HF patients

Table 8.2 First hospital admission by short, medium and long follow-up time periods

Follow-up month	Number entering follow-up month (n)	Admitted during follow-up month (%)	Censored during follow-up month (%)	Not admitted at end of follow-up month (n)
1	30061	9801(32.6)	1395(4.6)	18865
2	18865	2471(13.1)	170(0.9)	16224
3	16224	1573(9.7)	140(0.9)	14511
4	14511	1147(7.9)	119(0.8)	12245
5	13245	819(6.2)	91(0.7)	12335
6	12335	771(6.3)	80(0.6)	11484
7	11484	600(5.2)	69(0.6)	10815
8	10815	563(5.2)	68(0.6)	10184
9	10184	480(4.7)	82(0.8)	9622
10	9622	443(4.6)	58(0.6)	9121
11	9121	417(4.6)	64(0.7)	8704
12	8704	349(4.0)	50(0.6)	8223
13	8223	301(3.7)	52(0.6)	7870
14	7870	275(3.5)	51(0.6)	7544
15	7544	241(3.2)	60(0.8)	7243
16	7243	221(3.1)	50(0.7)	6972
17	6972	245(3.5)	53(0.8)	6674
18	6674	202(3.0)	40(0.6)	6432
≥19	6432	3619(56.3)	2831(44)	6190

Censored data included subjects who (i) belonged to a practice that stopped contributing data to the CPRD during the follow-up month (ii) were transferred out of the practice (iii) died or (iv) reached the study end.

Table 8.3 Time-matched characteristics of the matched sample by first hospital admission outcome over maximum 12 years follow-up

Factors and exposures	All (n=110,789)	Cases (n= 24,339)	Controls (n=86,450)	Missing n(%)
Person and socio-demographic factors				
Current age, years	78[IQR 70-84]	79[72-85]	78[70-84]	-
Women	53,804(48.6)	11,388(46.8)	42,416(49.1)	-
IMD quintile				-
1	22,567(20.4)	4,676(19.2)	17,891(20.7)	
2	26,602(24.0)	5,612(23.1)	20,990(24.3)	
3	23,003(20.8)	5,155(21.2)	17,848(20.6)	
4	22,579(20.4)	5,060(20.8)	17,519(20.3)	
5	16,038(14.5)	3,836(15.8)	12,202(14.1)	
Anthropometric and clinical factors				
BMI (Kg/m2)	27.2[23.9-31.2]	26.7[23.5-30.6]	27.3[24.1-31.3]	-
Cholesterol (mmol/L)	4.7±1.2	4.66±1.2	4.73±1.2	-
Hb (g/dL)	13.3±1.7	12.9±1.9	13.4±1.6	-
Systolic BP (mmHg)	135.6±20.2	134.1±21.5	136.0±19.8	-
Diastolic BP (mmHg)	75.7±11.2	74.7±12.0	75.9±11.0	-
Prior Hospital admission				-
<3 months	8,663 (7.8)	5,085(20.9)	3,578(4.1)	
3-6 months	8,387 (7.6)	2,509(10.3)	5,878(6.8)	
>6months to 1 year	12,930 (11.7)	2,918(12.0)	10,012(11.6)	
Lifestyle factors				
Smoking status				-
yes	12,028(10.9)	3,025 (12.4)	9,003 (10.4)	
No	53,310(48.1)	11,240(46.2)	42,070 (48.7)	
Ex	45,451(41.0)	10,074(41.4)	35,377(40.9)	
Alcohol status				-
Yes	81,676(73.7)	17,500(71.9)	64,176(74.2)	
No	24,783(22.4)	5,779(23.7)	19,004(22.0)	

Ex	4,330(3.9)	1,060(4.4)	3,270(3.8)	
Drug factors				
Diuretics	79,860(72.1)	17,023(69.9)	62,837(72.7)	-
Beta blocker	44,467(40.1)	8,893(36.5)	35,574(41.2)	-
ACEi	63,907(57.7)	12,477(51.3)	51,430(59.5)	-
ARB	19,324(17.4)	3,722(15.3)	15,602(18.1)	-
ACEi or ARB	80,420(72.6)	15,645(64.3)	64,755(74.9)	-
Comorbidity exposures				
Diabetes	21,291(19.2)	5,577(22.9)	15,714(18.2)	-
COPD	11,903(10.7)	3,230(13.3)	8,673(10.3)	-
Renal disease (eGFR <60)	37,784(52.2)	9,199(57.9)	28,585(50.6)	38,390(34.7)
Number of comorbidities				38,390(34.7)
0	23,607(32.6)	4,240(26.7)	19,367(34.3)	
1	36,190(50.0)	8,011(50.4)	28,179(49.9)	
2	11,687(16.1)	3,350(21.1)	8,337(14.8)	
3	915(1.3)	291(1.8)	624(1.1)	

Data are number patients (%) or mean± standard deviation or median[IQR]. IMD, index multiple deprivation (1=least deprived, 5=most deprived); BMI, body mass index; Hb, haemoglobin; BP, blood pressure; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; HbA1c, glycated haemoglobin; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate. Number of comorbidities is based on those listed (COPD, DM, CKD defined by eGFR),

Table 8.4 Time-dependent COPD comorbidity exposures of the matched sample by first hospital admission outcome

Comorbidity measures	All HF-HA (n=110,789)	Cases (n= 24,339)	Controls (n=86,450)	Missing n(%)
COPD diagnostic code (n=11,903)	11,903(10.7)	3,230(13.3)	8,673(10.3)	
Diagnosis before HF (91.1%)	10,843(9.8)	2,946(12.1)	7,897(9.1)	-
Diagnosis after HF (8.9%)	1,060(1.0)	284(1.2)	776(0.9)	-
COPD FEV¹ severity (COPD group only)	All HF-COPD (n=11,903)	Cases HF-COPD (n= 3,230)	Controls HF-COPD (n=8,673)	
FEV_1(percent predicted)	50[36.6-65.0]	51.6[37.0-66.0]	49.7[36.5-64.6]	10,031(84.3)
COPD GOLD severity stage				10,031(84.3)
1: FEV_1 ≥80% normal	187(10.0)	53(9.7)	134(10.1)	
2: FEV_1 50-79% normal	758(40.5)	232(42.4)	526(39.7)	
3: FEV_1 30-49% normal	666(35.6)	190(34.7)	476(35.9)	
4: FEV_1 <30% normal	261(13.9)	72(13.2)	189(14.3)	
COPD Drug severity				-
No steroids or oxygen (ref)	8,743(73.5)	2,143(66.4)	6,600(76.1)	-
Oral steroids but no oxygen	2,817(23.7)	923(28.6)	1,894(21.8)	-
On oxygen	343(2.9)	164(5.1)	179(2.1)	-
COPD FEV¹ severity change (COPD group only)				
Gold stage same or better (ref)	534 (81.8)	151(78.2)	383(83.3)	11,250(94.5)
Gold stage worse (at least one stage)	119(18.2)	42(21.8)	77(16.7)	
10% change				
<10% change (ref)	379(58.0)	116(60.1)	263(57.2)	11,250(94.5)
≥10% increase (better)	126(19.3)	31(16.1)	95(20.7)	
≥10% decrease (worse)	148(22.7)	46(23.8)	102(22.2)	
5% change				11,250(94.5)
<5% change (ref)	253(38.7)	78(40.4)	175(38.0)	
≥5% increase (better)	175(26.8)	52(26.9)	123(26.7)	
≥5% decrease (worse)	225(34.5)	63(32.6)	162(35.2)	
COPD severity change (Drugs)				-
No new steroids or oxygen (ref)	10,732(90.2)	2,920(90.4)	7,812(90.1)	
New on steroids but no new oxygen	1,107(9.3)	285(8.8)	822(9.5)	-
New on oxygen	64(0.5)	25(0.8)	39(0.5)	-

Data are number patients (%) or mean± standard deviation or median[IQR]. First hospital admission is the first all-cause admission following the incident HF date. COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in 1 second; pp, percent predicted; ref, reference group. Most recently recorded FEV1 (pp) before the match date was used to measure COPD severity within a maximum of a 6-month time window. % change for FEV1 was calculated as the absolute change in pp over 6-months using the most recent measure and an earlier FEV1 measure recorded up to a maximum of 1-year before the most recent measure. The FEV1 change was adjusted for the time interval.

Table 8.5 Time-dependent DM comorbidity measures of the matched sample by first hospital admission outcome

Comorbidity exposure measures	All HF-HA (n=110,789)	Cases HF-HA (n= 24,339)	Controls HF-HA (n=86,450)	Missing n(%)
Diabetes status (n=21,291)	21,291(19.2)	5,577(22.9)	15,714(18.2)	
Diagnosis before HF (93%)	19,793(17.9)	5,300(21.8)	14,493(16.8)	-
Diagnosis after HF (7%)	1,498(1.4)	277(1.1)	1,221(1.4)	-
Diabetes HbA1c severity (DM group only)	All HF-DM (n=21,291)	HF-DM Cases (n= 5,577)	HF-DM Controls (n=15,714)	
HbA1c (%)	7.1[IQR 6.4-8.1]	7.2 [6.4-8.2]	7.1[6.4-8.1]	6,713(31.5)
<5.5%	424(2.9)	137(3.8)	287(2.6)	
5.5-6.4%	3,307(22.7)	803(22.2)	2,504(22.9)	
6.5-7.5%	5,461(37.5)	1,282(35.4)	4,179(38.1)	
7.6-8.5%	2,771(19.0)	673(18.6)	2,098(19.2)	
8.6-9.5%	1,302(8.9)	328(9.1)	974(8.9)	
>9.5%	1,313(9.0)	398(11.0)	915(8.4)	
Diabetes drug severity				
1: None	5,799(27.2)	1,390(24.9)	4,409(28.1)	-
2: Any oral (+/-Insulin)	12,779(60.0)	3,364(60.3)	9,415(59.9)	-
3: Insulin only	2,713(12.7)	823(14.8)	1,890(12.0)	-
Diabetes HbA1c severity change (DM group only)				9,049(42.5)
<1% change (ref)	8,592(70.2)	2,056(67.7)	6,536(71.0)	
>1% increase	1,551(12.7)	409(13.5)	1,142(12.4)	
>1% decrease	2,099(17.2)	570(18.8)	1,529(16.6)	
Diabetes drugs severity change				-
No drug category change	19,607(92.1)	5,052(90.6)	14,555(92.6)	-
Increase in drug category	1,261(5.9)	356(6.4)	905(5.8)	-
Decrease in drug category	423(2.0)	169(3.0)	254(1.6)	-

Data are number patients (%) or mean± standard deviation or median[IQR**].** First hospital admission is the first all-cause admission following the incident HF date. HbA1c, glycated haemoglobin; ref, reference group. Most recently recorded HbA1c (%) prior to the match date was used to measure diabetes severity within a maximum of a 6-month time window. HbA1c change was calculated as the absolute change in HbA1c over 6-months using the most recent measure and an earlier HbA1c measure recorded up to a maximum of 1-year before the most recent measure. HbA1c change was adjusted for the time interval.

Table 8.6 Time-dependent CKD comorbidity exposures of the matched sample by first hospital admission outcome

Comorbidity measures	All (n=110,789)	Cases (n= 24,339)	Controls (n=86,450)	Missing n(%)
Renal disease diagnosis				
eGFR <60	37,784(52.2)	9,199(57.9)	28,585(50.6)	38,390(34.7)
Renal severity (eGFR)				
eGFR mL/min/1.73m^2	59.5±19.4	56.5±20.9	60.3±18.9	38,390(34.7))
Renal severity by eGFR stage*				38,390(34.7)
1: >105	1,248(1.7)	281(1.8)	967(1.7)	
1: 90-105	3,633(5.0)	778(4.9)	2,855(5.1)	
2: 60-89	29,734(41.1)	5,634(35.5)	24,100(42.7)	
3A: 45-59	21,497(29.7)	4574(28.8)	16,923(30.0)	
3B: 30-44	12,616(17.4)	3,182(20.0)	9,434(16.7)	
4: 15-29	3,372(4.7)	1,243(7.8)	2,129(3.8)	
5: <15	299(0.4)	200(1.3)	99(0.2)	
eGFR severity change				56,017(50.6)
Absolute change				
0mls to 5mls decrease (ref)	13,020(23.8)	2,691(21.9)	10,329(24.3)	
>15mls decrease	8,856(16.2)	2,431(19.7)	6,425(15.1)	
6mls to 15mls decrease	10,161(18.6)	2,307(18.7)	7,854(18.5)	
any increase	22,735(41.5)	4,888(39.7)	17,847(42.0)	
Percentage change				
0-5% decrease (ref)	8,148(14.9)	1,540(12.5)	6,608(15.6)	
>25% decrease	8,972 (16.4)	2,578(20.9)	6,394(15.1)	
6-25% decrease	14,548(26.6)	3,233(26.3)	11,315(26.7)	
Any % increase	23,104(42.2)	4,966(40.3)	18,138(42.7)	

Data are number patients (%) or mean± standard deviation. First hospital admission is the first all-cause admission following the incident HF date. eGFR, estimated glomerular filtration rate; ref, reference group; ref, reference group. * National Kidney Foundation Kidney Disease Outcomes Quality Initiative (KDOQI) guidelines. >105 group added to stage one due to the prior evidence of increased risk in high eGFR. % change was calculated as the absolute change in eGFR over 6-months and the relative change (difference in eGFR over 6-months as a proportion of the first eGFR measure)

Figures

Figure 8.1 First hospital admissions as a percentage of the total HF-HA sample at baseline in each month of follow-up

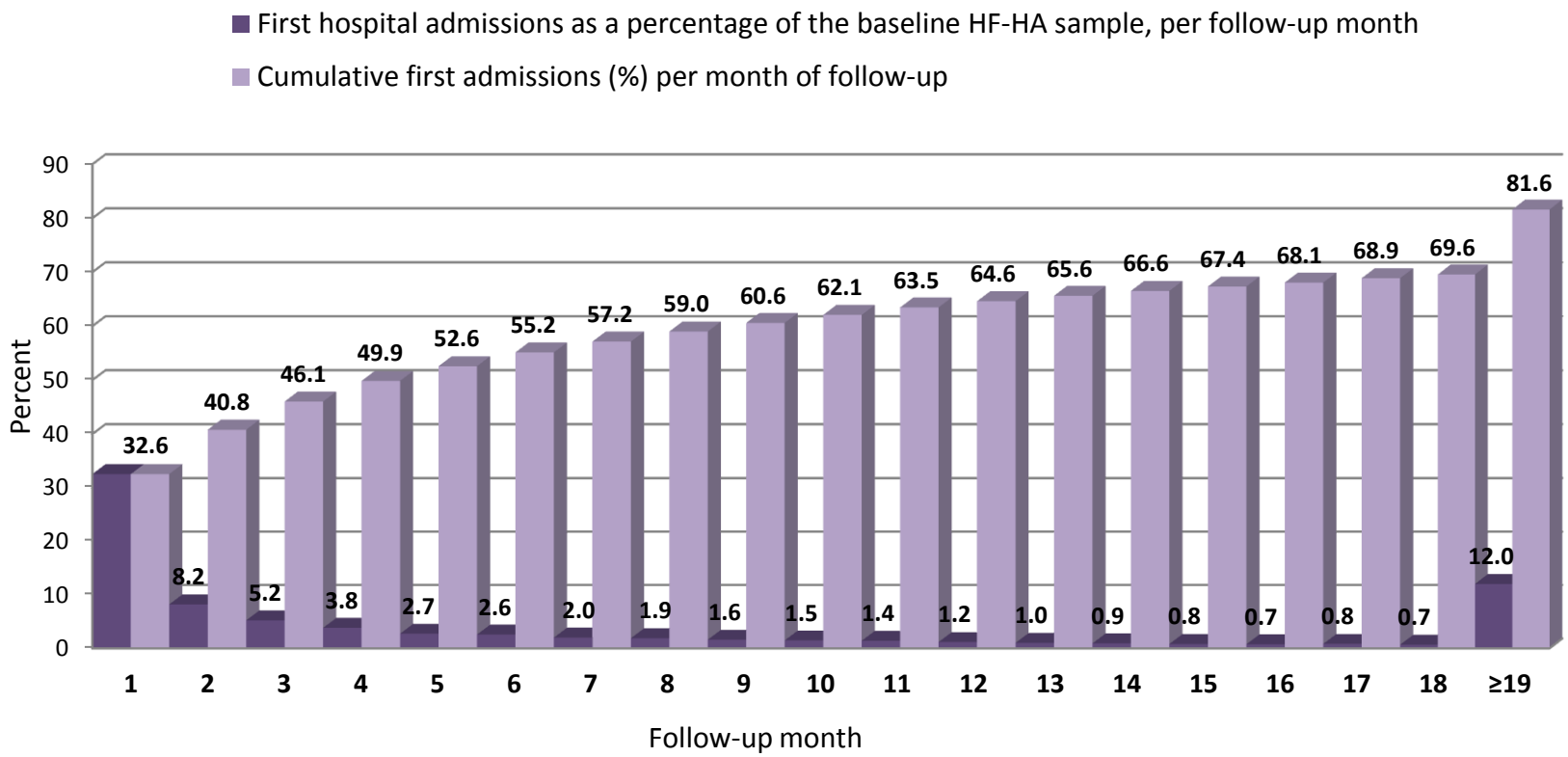


Figure 8.1: The dark purple bars show the percentage of the baseline cohort who were admitted in each month of follow up. The denominator in this figure is the 30,061 subjects in the baseline cohort. The light purple bars show the cumulative percentage of the baseline cohort who were admitted over the 12 years of follow-up.

Figure 8.2 First hospital admissions as a percentage of the remaining HF patients not admitted, per follow-up month

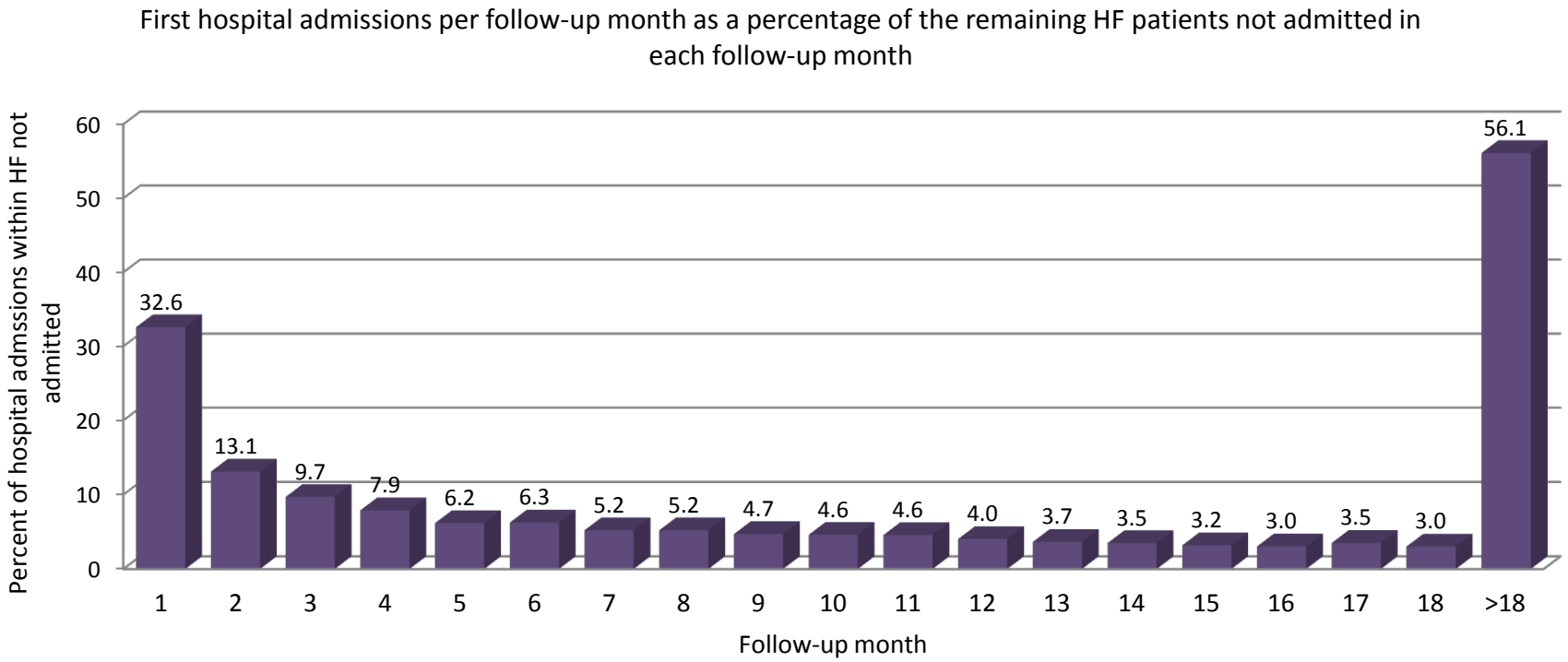


Figure 8.2: The purple bars show the percentage of remaining subjects who were admitted who entered each month of follow up.

Figure 8.3 Comorbid COPD diagnosed pre and post HF by first hospital admission outcome

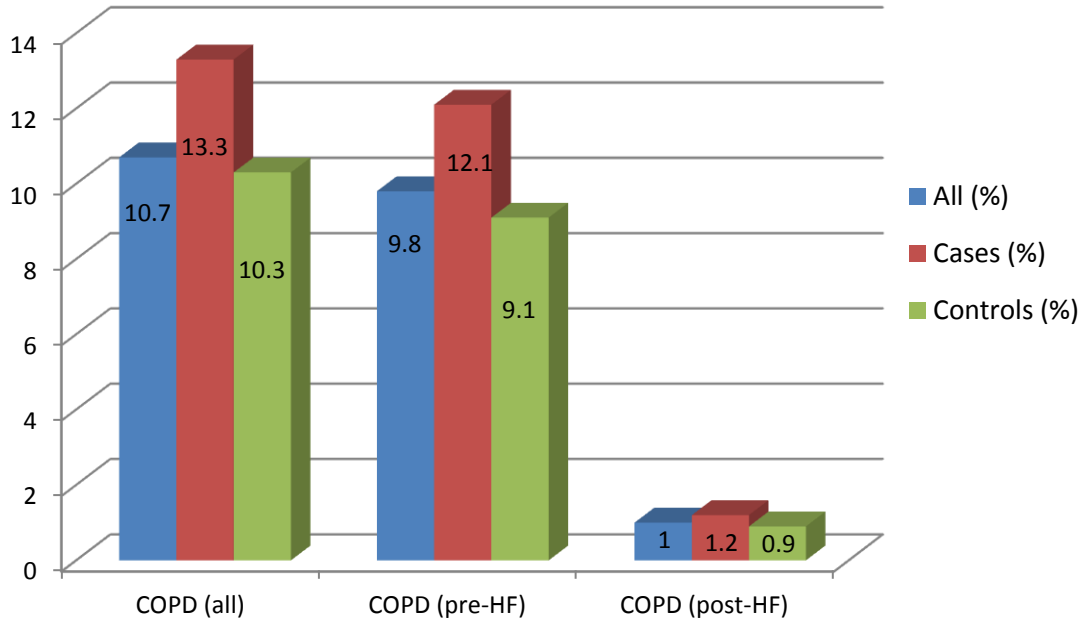


Figure 8.4 Comorbid COPD group drug severity stages

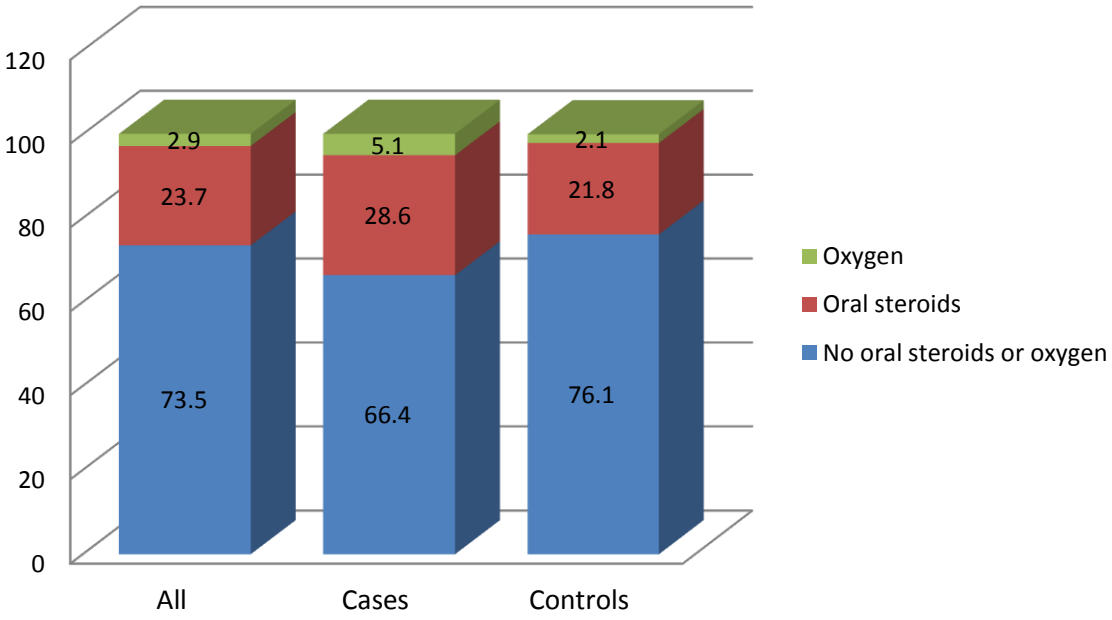


Figure 8.5 Comorbid DM diagnosed pre and post HF by first hospital admission outcome

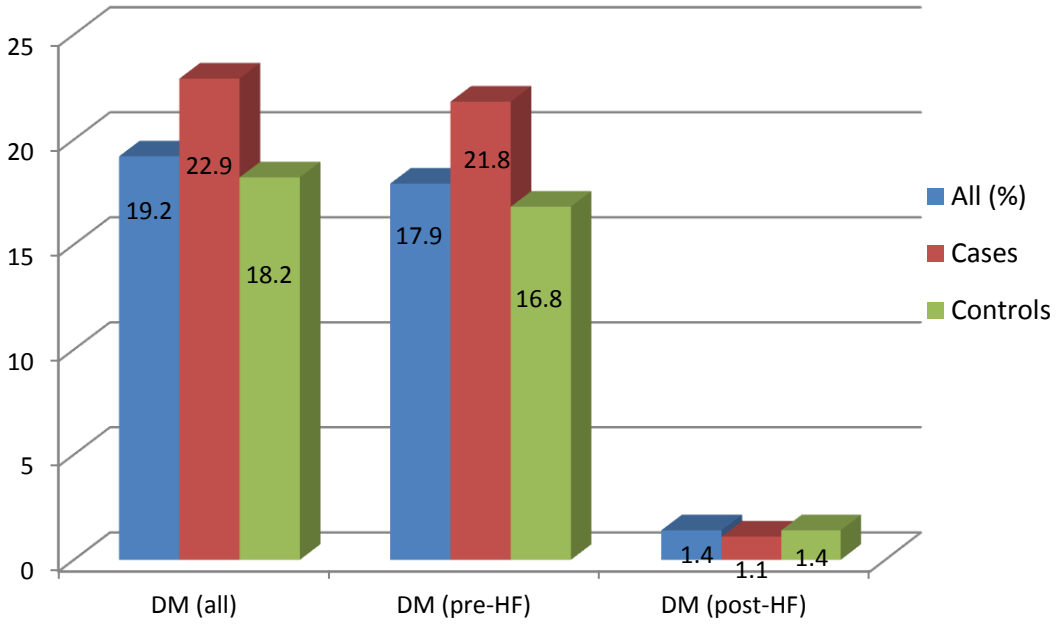
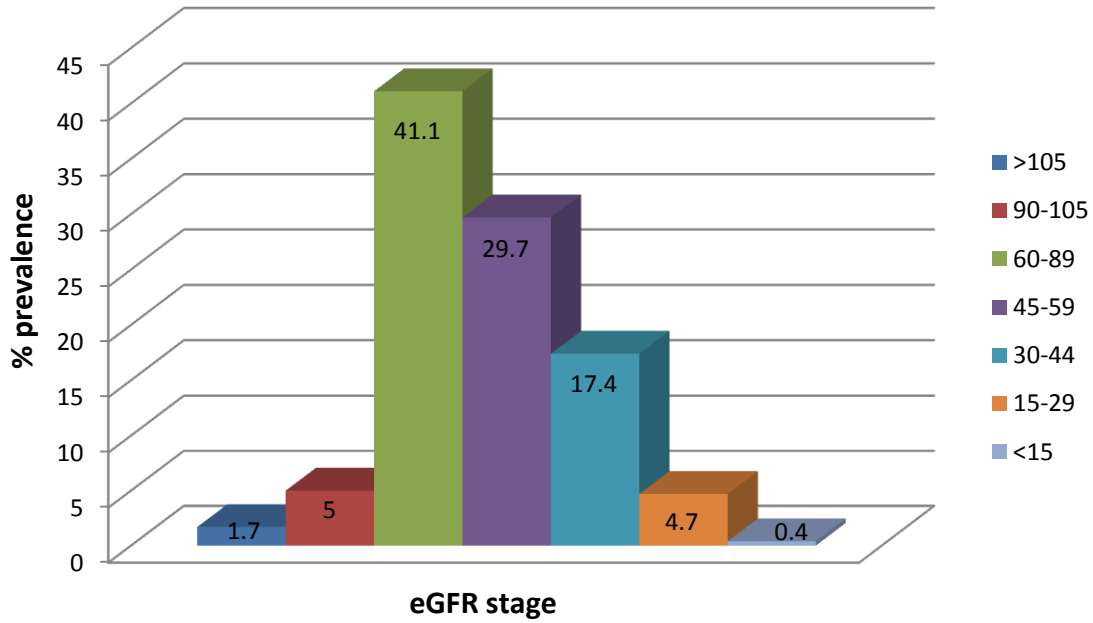


Figure 8.6 Comorbid renal stages based on eGFR in the HF-HA cohort



Chapter 9 Non-CVD comorbidity prognostic factors in an incident HF general practice population: strength of associations with all-cause mortality.

This chapter follows on from the descriptive findings (Chapter 7) of the main incident HF CPRD cohort and the outcome of all-cause mortality. This chapter presents the results of the investigation of the strength of associations between the selected non-CVD comorbidities (COPD, DM and CKD) in HF and all-cause mortality. The effects of time-dependent exposures of the three comorbidities on all-cause mortality are firstly presented by their status 'present or not' followed by their stratification by categories of severity and change. All effect estimates will be presented as odds ratios (OR) with 95% confidence intervals (CI).

This chapter presents the findings of the five stages to analysis:

(i) Confounding investigation

(ii) Test of model assumptions: linearity

(iii) Test of model assumptions: collinearity

(iv) Adjusted associations of the comorbidities with mortality

(v) Comorbidity effects stratified by categories of severity and change

9.1 Confounding investigations

There were four steps to the investigation of confounding. The potential confounders for each of DM, COPD and CKD that were identified in all four steps are summarised in Table 9.1. The first step was to list all confounders that were identified for each of the comorbidities from previous HF evidence included in the systematic review (step 1, Table key; SR). The second step was to identify from the available CPRD data, the potential confounders which differed between the comorbid and the non-comorbid groups in Chapter 7 (step 2, Table key *). In step 3, each comorbid DM, COPD, CKD effect in HF was now compared in strata of these

potential confounders to investigate whether their strength of association with mortality differed across strata (step 3, Table key #). Finally, the three comorbidity effects are then adjusted for each available potential confounder in turn, to further investigate confounding (step 4, Table key ł). A 10% difference in the comorbidity effect in steps 3 and 4 was used to indicate confounding(9).

Comorbidity effects within strata of the potential confounders (step 3): The unadjusted associations between the potential confounders in HF and all-cause mortality are reported in Table 9.2. The potential confounders that had the strongest unadjusted associations with mortality were age, body mass index (BMI), cholesterol, haemoglobin and cardiovascular medications. When the unadjusted associations between the three comorbidities and mortality were observed within strata of potential confounders, it was these factors (with the exception of cholesterol) that resulted in a >10% change to the odds ratio of the comorbidity (E-Appendix A34). The relative effects all three comorbidities were strengthened in the youngest age groups and in the groups prescribed beta-blockers for HF-DM and HF-CKD. The relative effects were reduced in the groups with lower BMI or not prescribed ACEi or ARB for HF-COPD and HF-CKD. The effect of HF-CKD was reduced in the group with lower haemoglobin level or with combined HF-COPD.

Comorbidity associations with all-cause mortality adjusted by each potential confounder (step 4): Age was a consistent confounder across all three comorbidities (COPD, DM, CKD) (E-Appendix A35). In the HF-DM and HF-COPD groups (which were younger than their non-DM or non-COPD HF counterparts), adjustment for age strengthened the comorbidity associations with mortality. This was the reverse in the older HF-CKD group where its association with mortality was reduced by the adjustment of age (OR 1.77 reduced to 1.34).

Absence of beta-blockers was strongly associated with mortality (OR 1.75; 95% CI 1.70-1.80) and was observed more in the HF-COPD group (Chapter 7). Adjustment of HF-COPD by prescription of beta-blockers reduced its effect on mortality. BMI which also had a strong association with mortality (OR 0.94 per increase Kg/m^2; 0.94-0.95). BMI was higher in the HF-DM group than the non-DM HF group (median 29.2 [IQR 25.5-33.7] v. 26.3 [23.1-30] kg/m^2). Adjustment of the HF-DM effect by BMI strengthened its association with mortality (OR 1.09 increased to 1.32).

Haemoglobin had one of the strongest unadjusted effects on mortality (OR 0.78 per increase g/dL; 0.77-0.78). The effect estimates on mortality of HF-DM and HF-CKD (which both had a lower mean haemoglobin than the non-DM and non-CKD groups respectively), were reduced following adjustment by haemoglobin (DM, OR 1.09 reduced to 0.98; CKD, OR 1.77 reduced to 1.42).

Summary of the confounders for each of the three comorbidities: All of the available covariates from CPRD were identified as potential confounders for each of the three comorbidities through the four steps of confounding investigation. The covariates were used to adjust each comorbidity exposure in order of their potential importance as follows (summarised Table 9.3):

a. Factors identified in step 3 (10% change of the comorbidity effect in any strata of the potential confounder) *and/or* step 4 (10% change in the comorbidity effect when adjusted by the potential confounder),

b. Previous evidence of confounding identified in the systematic review studies (step 1) *AND* a difference in the prevalence of the potential confounder between the comorbid disease compared to the non-comorbid group in the CPRD analyses (step 2),

c. One of step 1 *OR* step 2,

d. Sub-analysis adjusted for deprivation.

9.2 Test of model assumptions: linearity

There were 7 continuous covariates to include in the adjusted models, which were age, BMI, cholesterol, haemoglobin, systolic and diastolic blood pressure and estimated glomerular filtration rate (eGFR). The fit of the covariate as a linear term was investigated by using likelihood ratio tests. The fit of different expressions of

the covariate was also observed using fitted line plots and Eccles plots. For each continuous variable, fully adjusted models using the covariates in Table 9.2 were derived and then quadratic and cubic terms for the continuous variable incrementally added. The models were compared using likelihood ratio tests. Fitted line plots were then used to compare a Lowess line, which closely follows the observed data, with linear and quadratic extension fitted lines of the association between the continuous variable and predicted probabilities of mortality. Finally Eccles plots were used to graph deciles of each covariate against predicted probabilities using fully adjusted models with the continuous covariate included and with the additional quadratic and cubic extensions. These plots included Lowess lines to compare the predicted models with the observed data. All linearity tests for the 7 continuous covariates are shown in E-Appendix E16 and summarised in E-Appendix A36 with two examples detailed below.

- Age

Likelihood ratio tests showed that addition of a quadratic extension to age significantly improved the model fit with a small increase in log-likelihood of 53.26, (Table 9.4). Further addition of a cubic term made no significant difference. The fitted line plots showed a straight line for the largest distribution of age covering the inter-quartile range (71-85 years) (Figure 9.1). Eccles plots showed good fit of the adjusted models with (i) current age and with (ii) current age plus the quadratic term (age^2) (Figure 9.2). Given that there was minimal improvement in model fit by addition of a quadratic extension and the largest distribution of age range had a linear association with mortality, the final modelling decision was to include age in its simple form assuming linearity.

- eGFR

Likelihood ratio tests showed that addition of a quadratic extension to eGFR significantly improved the model fit with a substantial increase in log-likelihood of 385.06 (Table 9.5). Further addition of a cubic term also made a significant difference but with only a small change to log-likelihood of 9.53. The fitted line plot showed a curvilinear association between eGFR and mortality risk with an upward slope starting at approximately eGFR 100ml/min/m^2. The fitted line using a quadratic extension followed the observed line closely (Figure

9.3). Eccles plots showed good fit of the adjusted models with eGFR plus the quadratic term ($eGFR^2$) (Figure 9.4).

Given that there was a substantial improvement in model fit by adding a quadratic extension and only a slight improvement by adding a cubic extension, the final modelling decision was to include eGFR with a quadratic extension ($eGFR^2$) in the adjusted models.

9.3 Test of model assumptions: collinearity

Most covariate pairs had a correlation coefficient of $r<0.25$ The strongest correlations were systolic and diastolic blood pressures ($r=0.55$) and CKD and eGFR ($r=0.8$) (correlation matrix E-Appendix C15). Systolic blood pressure and not diastolic blood pressure was selected as a covariate based on previous evidence as a predictor of outcomes in cardiovascular disease(462) and eGFR as a continuous measure for adjusting in the DM and COPD models. All continuous variables were centred at their means (Age 77 years; cholesterol 5 mmol/ml; Hb 13 g/dL; BMI 28 $Kg./m^2$; systolic BP 131 mmHg; eGFR 59 $ml/min/m^2$) which removed the collinearity between the variable with a respective quadratic term (E-Appendix C16).

9.4 Adjusted associations between DM, COPD, CKD comorbidity in HF and mortality

9.4.1 Diabetes Mellitus

In the incident HF population, the unadjusted association between comorbid DM and all-cause mortality was OR 1.09 (95% CI 1.06-1.12). This estimate was most influenced by adjustment for the first set of confounders: age, eGFR, BMI, Hb and no beta-blocker (OR 1.26; 1.21-1.30) (Table 9.6). Adjustment of the remaining confounders (OR 1.28; 1.23-1.33) or deprivation did not diminish these estimates (OR 1.26; 1.19-1.33).

9.4.1.1 Timing of DM comorbidity

When DM was stratified by prevalent DM at the time of HF diagnosis (pre-HF DM) or incident DM developing after HF diagnosis (post-HF DM), there was a difference in the associations with mortality. The adjusted association for pre-HF DM was OR 1.31 (1.26-1.37) which was not influenced by deprivation. The adjusted association for post-HF DM was 1.13 (1.03-1.23) (Figure 9.5). This association became non-significant when adjusted for deprivation in the sub-analysis (1.08; 0.95-1.23).

9.4.1.2 Comorbid Diabetes Mellitus severity

- **Physiological severity measure**

The HF group with comorbid DM was stratified into groups by glycated haemoglobin (HbA1c) levels. This categorisation used the closest available measure preceding the match date (median 114 [IQR 52-206] days) comparing the HF group with and without diabetes.

Shape of association: There was a curvilinear relationship between categories of HbA1c in the HF-DM group and their respective predicted probabilities of mortality observed (Figure 9.6) using a margins plot. A margins plot graphs categories of a continuous variable against predicted risk whilst fixing other covariate effects at zero. The DM category associated with the lowest risk was HbA1c 6.5%-7.5%. When HbA1c was categorised into deciles, the curved U shape curve remained with the lowest risk decile being HbA1c 5.9%-6.2% (Figure 9.7).

Strength of associations: Fully adjusted associations between categories of HbA1c in the HF-DM group and mortality, compared to the no DM HF group are shown in Table 9.6. The association between HbA1c category 6.5%-7.5% and mortality was OR 1.19 (95% CI 1.12-1.26). From this lowest risk level, mortality increased with both decreasing and increasing HbA1c categories. The estimate for the lowest HbA1c category (<5.5%) was OR 1.34 (1.16-1.54) and for the highest HbA1c category (>9.5%) was 1.45 (1.30-1.62) (Figure 9.8. In the fully adjusted model, the linear tests for trend in the reducing categories from '6.5%-7.5%' was non-significant but was significant in the increasing categories from 6.5%-7.5%.

The highest mortality risk categories were HbA1c 8.6%-9.5% (adjusted OR 1.49 (1.34-1.67)) and HbA1c >9.5% (adjusted 1.45; (1.30-1.62)). The association was also significant for the lowest risk category of HbA1c 6.5%-7.5% (1.19; 1.12-1.26), but notably the confidence intervals did not over-lap with higher HbA1c levels indicating significant stratified DM effects by severity (Figure 9.8).

- **Drug severity measure**

The HF-DM group were stratified into groups based on DM-related prescribed medications in a four month time-window before the match date. The groups were: (i) DM but no related medications, (ii) DM and oral hypoglycaemics, (iii) DM and oral hypogycaemics plus insulin, (iv) DM and insulin only. These groups were compared to the HF sample without diabetes.

Strength of associations: When adjusted for all confounders (Table 9.6), all HF-DM drug groups had a significant association with mortality with the largest effects in the 'no medications group' (OR 1.37; 1.29-1.46) and the 'insulin only group' (OR 1.56; 1.44-1.68). Sub-analysis with adjustment for deprivation made little difference to the effect estimates.

The HF-DM group with prescribed 'oral hypoglycaemic drugs' had similar adjusted associations with mortality to the HF-DM group with 'oral hypoglycaemic drugs plus insulin' (OR 1.16 (95% CI 1.10-1.22) and 1.18 (1.07-1.32) respectively). The largest effects estimates for 'no medications' and 'insulin only' shown above had confidence intervals that did not overlap with the estimate for the DM with 'oral hypoglycaemic drugs +/- insulin' (OR 1.17; 1.11, 1.22) indicating stratified effects of comorbid DM as measured by drugs prescribed (Figure 9.9).

9.4.1.3 DM severity change

- **Physiological severity change measure**

The HF-DM group were stratified into three groups according to change in their glycated haemoglobin (HbA1c) level over 12-months before the match date. Change definition was based on the closest available

HbA1c measure preceding the match date and a previous measure before the most recent measure (median 298 [IQR 230-379] days). Change over one-year was calculated and compared to the HF sample without DM.

Strengths of associations: All three categories of HbA1c change in HF were significantly associated with mortality in the adjusted models (Table 9.6). Those with less than 1% change or a >1% increase in HbA1c had similar adjusted associations with OR 1.22 (1.16-1.28) and 1.28 (1.17-1.40) respectively. The association for a >1% reduction in HbA1c had the strongest association with mortality with OR 1.49 (1.37-1.61). When the first two groups with a similar estimate of effects were combined (<1% change or >1% increase), the confidence interval of the effect estimate of the group with a >1% decrease in HbA1c did not overlap with the combined group (OR 1.23; 1.18-1.29) (Figure 9.10).

Drug severity change measure: The DM group were stratified into three groups based on change in their prescribed DM-related drugs over a year before the match date. This measure used the DM drugs prescribed in a 4-month time-window before death compared to a 4-month time-window a year before death. The two oral hypoglycaemic drug categories were combined for this measurement resulting in three drug categories; 'no drugs', 'oral hypoglycaemic +/- insulin' and 'insulin only'. Change in drug category over one year was categorised into 'no change', an 'increase in at least one drug category' and a 'decrease in at least one drug category'. These groups were compared to the HF sample without diabetes.

Strength of associations: Fully adjusted associations between drug category change and mortality showed the highest risk was in the 'decreased at least one drug category' group with OR 2.30 (2.01-2.63). The adjusted associations for 'no change' and 'increased at least one drug category' were similar with OR 1.23 (1.19-1.29) and 1.33 (1.20-1.48) respectively.

9.4.2 Chronic obstructive pulmonary disease

The adjusted association between HF-COPD and all-cause mortality compared to non-comorbid group was OR 1.47 (95% CI 1.41-1.54). Adjustment was by age, no beta-blocker, BMI, eGFR, no ACEi/ARB and Hb

(Table 9.7) and the estimate reduced to OR 1.35 (1.29, 1.41) in the fully adjusted model (Table 9.7). Sub-analysis adjusting for deprivation made little difference.

9.4.2.1 Timing of COPD comorbidity

When COPD was stratified by prevalent COPD at the time of HF diagnosis (pre-HF COPD) or incident COPD developing after HF diagnosis (post-HF COPD), there was a difference in the associations with mortality. The association between pre-HF COPD and mortality was 1.32 (1.25-1.38) in the fully adjusted model. The adjusted association between post-HF COPD and mortality was increased following the same adjustment, with an OR 1.46 (1.34-1.60). This association was not further influenced by deprivation (Figure 9.5).

9.4.2.2 Comorbid COPD severity

- **COPD Physiological severity measure**

The HF-COPD group was stratified into four groups according to their forced expiration volume in 1 second (FEV_1) levels defined by GOLD guidelines(410). This measure used the closest available measure preceding the match date (median 295 [IQR 137-524] days) comparing the HF group with and without COPD.

Shape of association: The shape of the adjusted associations between the GOLD stages and their respective predicted probabilities of mortality appeared linear for the last three stages (Figure 9.11) using a margins plot.

Strength of associations: Fully adjusted associations between the GOLD stages and mortality, compared to no COPD HF are shown in Table 9.7. All GOLD stages showed higher risk estimates for mortality than the total HF-COPD group which included patients with and without recorded FEV_1. The lowest risk category was FEV_1 ≥80% with an estimate of OR 1.73 (1.50-1.99). Mortality risk increased with GOLD stages from stage 1 (FEV_1 ≥80%) to stage 4 (<30%) which had a relative risk of OR 3.14 (2.65-3.73) shown in Figure 9.12. Sub-analysis with adjustment for deprivation made little difference to the risk estimates.

The linear test for trend showed a significant adjusted association between the GOLD stages in HF and mortality. The highest risk categories were FEV_1 30-49% (adjusted OR 2.31; 2.09-2.54) and FEV_1 <30%

normal (OR 3.14; 2.65-3.73). Both risk estimates had confidence intervals that did not overlap with the confidence interval of the lower two risk categories of FEV_1 ≥80% (OR 1.73; 1.50-1.99) and FEV_1 50-79% (1.76; 1.61-1.91) (Figure 9.12). There was more than doubling of the risk of mortality in the higher risk HF-COPD groups compared the lower risk COPD groups.

- **COPD drug severity measure**

The HF-COPD group was stratified into seven groups according to their prescribed COPD-related medications in a 4-month time-window before the match date. The 7 groups were: (i) COPD and no medications, (ii) COPD and short terms inhalers only, (iii) COPD and monotherapy, (iv) COPD and dual therapy, (v) COPD and triple therapy, (vi) COPD and prescribed oral steroids and no prescribed oxygen, and (vii) COPD and prescribed oxygen. Mono, dual or triple therapy related to the prescription of one, two or three of long acting beta2-antagonist, long acting cholinergic, methylxanthines and inhaled steroids, either individually or in combination inhalers. These groups were compared to the HF sample without COPD.

Strength of associations: Fully adjusted associations between the COPD drug severity groups compared to non-comorbid groups and mortality are shown in Table 9.7. With the exception of 'COPD and triple therapy', the associations between COPD inhaler therapies and mortality were non-significant. In the second drug severity classification, combining the inhaler therapies into one group ('any inhalers') showed a relative mortality risk of OR 1.09 (1.03-1.16). The COPD drug groups with the largest effect estimates were 'no drugs' (OR 1.28; 1.14-1.43), 'COPD and oral steroids but no prescribed oxygen' (1.83; 1.69-1.97) and 'prescribed oxygen' (2.94; 2.47-3.50). In the third classification, when the lowest risk groups ('no drugs' and 'any inhalers') were combined into a 'no prescribed steroids or oxygen' group, the effect estimate shown was OR 1.13 (1.07-1.19). The confidence intervals of the highest drug severity groups did not overlap with each other or the confidence intervals of either of the lower severity groups indicating significant stratification of comorbid COPD effects by drugs prescribed (Figure 9.13).

9.4.2.3 COPD severity change

- **Physiological severity change measure**

The HF-COPD group were stratified firstly into two groups according to (i) improvement in GOLD stage or stable GOLD stage and (ii) worsening of GOLD stage over the one year before the match date. Secondly the comorbid group were stratified into three groups according to whether they had a (i) <10% change in $FEV_{1,}$ (ii) ≥10% increase in FEV_1 (better) or (iii) ≥10% decrease FEV_1 (worse) over the one year before the match date. Finally the COPD group was stratified into the same groups but using a 5% change measure. These measures used the closest available FEV_1 measure before the match date and a previous measure (median 462 [IQR 360-663] days). Change over one year was calculated and these groups were compared to the HF sample without COPD.

Strength of associations: The fully adjusted associations of all HF-COPD physiological change measures compared to no COPD HF are shown in Table 9.7. The group that had a stable or improved GOLD stage had an adjusted association with mortality of OR 2.15 (1.97- 2.34). There was an increase in risk of 55% for the COPD group with worsening of GOLD stage with an OR 2.70 (2.30-3.17) Figure 9.14). Using the 10% change severity classification, the adjusted risk in the HF-COPD group with <10% change was OR 2.18 (1.99-2.38) which was similar to the group with ≥10% improvement in FEV_1 with OR 2.22 (1.82-2.70). The highest mortality risk was in the comorbid COPD group with ≥10% worsening of FEV_1 with OR 2.60 (2.17, 3.11). The linear test of trend in the fully adjusted models was not significant. Further adjustment with deprivation strengthened the associations of the different severity change measures weakly.

- **Drug severity change measure**

The HF-COPD group were stratified into three groups based on change in their prescribed COPD drugs over 12 months before the match date. This measure used the COPD drugs prescribed in a 4 month time-window before death compared to a 4 month time-window, a year prior to death. Change in drug category over one year was categorised into (i) 'drug category same or better' or (ii) 'drug category worse'. A second classification used three categories of (i) no new steroids or oxygen, (ii) new on steroids but no new oxygen and (iii) new on oxygen. These groups were compared to the HF sample without COPD.

Strength of associations: Using the first classification of drug severity change, the fully adjusted association between 'drug category same or better' (OR 1.34; 1.27-1.42) was similar to 'drug category worse' (OR 1.36; 1.27-1.45) (Table 9.7). The second drug classification showed increasing risk from the first to third category. 'No new steroids or oxygen' resulted in an adjusted effect estimate of OR 1.22 (1.16-1.28) which increased to OR 1.84 (1.67-1.28) for the 'new on steroids but no new oxygen' group and 3.41 (2.71- 4.29) for the 'new on oxygen' group (Figure 9.15). These groups were non-significant for a linear trend in the adjusted model but the confidence intervals of the three groups did not overlap indicating that the effect of COPD on mortality was stratified by drug severity change.

9.4.3 Chronic kidney disease

The unadjusted association between HF-CKD and all-cause mortality compared to no CKD HF was OR 1.77 (95% CI 1.72-1.82), which reduced when adjusted for age, gender, COPD and Hb to 1.21 (1.17-1.26). This estimate was not influenced by further adjustment for the remaining confounders (1.22; 1.18-1.26) or by adjustment for deprivation in the sub-analysis (Table 9.8).

9.4.3.1 CKD severity

- **Physiological severity measure**

The HF sample was stratified into groups based on the eGFR level. This categorisation used the closest available measure before the match date (median 101 [IQR 37-223] days) and these groups were compared to the baseline eGFR group 60-89 ml/min/m^2.

Shape of association: The shape of the adjusted associations between categories of eGFR and their respective predicted probabilities of mortality was curved, shown in Figure 9.16 using a margins plot. The lowest risk eGFR category was 60-89 ml/min/m^2. When eGFR was categorised into deciles the curved U shape remained and the lowest risk deciles were 7 and 8 corresponding to eGFR of 62-74.9 ml/min/m^2 (Figure 9.17).

Strength of associations: Mortality risk was associated with all eGFR categories above and below the reference group. The lowest risk category was eGFR 90-105 with an unadjusted OR 1.09 (1.01-1.17) and this increased to 1.59 (1.44-1.76) in the next higher group (eGFR >105). Risk also increased with each category below eGFR 90-105 to a maximum of an unadjusted OR 5.53 (4.94- 6.18) in the lowest eGFR group (<15 ml/min/m^2). The same trend was observed in the fully adjusted associations (Table 9). Risk was lowest compared to the reference group in the eGFR 45-59 group (1.04; 1.00-1.09) and increased above this to 1.64 (1.47-1.83) in the highest eGFR >105 group and to 3.26 (2.87-3.69) in the lowest eGFR <15 group. Further adjustment for deprivation made little difference to these estimates.

In the fully adjusted model, the tests for linear trend in the reducing and increasing eGFR categories from eGFR 45-59 were non-significant. The increasing and decreasing eGFR categories from the lowest risk group (eGFR 45-59) had confidence intervals that did not overlap with one another, which indicated stratified effects of CKD defined by recent eGFR severity levels on mortality outcome (Figure 9.17).

9.4.3.2 CKD severity change

- **Physiological severity change measure**

The HF sample was stratified into three groups based on change in their eGFR level over one year before the match date. This change definition used the closest available eGFR before the match date and a previous eGFR before the most recent measure (315 days [IQR 232-420] days). Absolute change was defined as: (a) 0-5mls decrease – reference group, minor decline, (b) any increase, (c) 6-15mls decrease – moderate decline, and (d) >15mls decrease – severe decline. Percentage change was defined as: (a) 0-5% decrease – reference group, minor decline, (b) any increase, (c) 6-25% decrease – moderate decline, and (d) >25% decrease – severe decline. Change over one year was calculated.

Strengths of associations: All categories of absolute and percentage change in eGFR were associated with mortality when compared to the reference group of minimal decline in the adjusted models. HF patients with any increase or moderate decline in eGFR had similar associations compared to the minimal decline reference group (Table 9.8). The highest risk was observed in the most severe change groups with eGFR

>15mls group (OR 1.83; 1.73-1.94) and >25% decrease (2.14; 2.00-2.28). Adjusting for deprivation had minimal influence on these associations. The confidence intervals of these estimates did not overlap, indicating the stratified effects of comorbid CKD by CKD change on mortality outcome in HF (Figure 9.20).

Adjustment for eGFR: Firstly, the influence of the start and end eGFR used in the calculation of change, on the estimates of ORs of the eGFR change was investigated by observing the associations within sub-groups according to their starting eGFR (<60 and ≥60 ml/min/m^2). Starting eGFR was the first and earliest measure in the change definition. The change estimates were also observed within same strata of the end eGFR measure. This was to investigate whether the estimates of change were :(i) influenced by the starting eGFR and whether this influence was less for the percentage change measure which already took an account of the starting point and (ii) depended on the end eGFR which might influence a greater effect estimate in the <60mls sub-group. Secondly, the estimates for the change measures in the adjusted model were further adjusted, separately for the start and end eGFR in order to investigate the independent associations of change in eGFR (Table 9.8).

Change stratified by baseline start and end eGFR: The effect estimates for eGFR moderate and severe deterioration were larger in those with a starting eGFR of <60ml/min/m^2, for both the absolute and percentage change measures. The biggest difference between the groups with eGFR starting point <60 and ≥60 was observed in the absolute change group where the confidence intervals for the adjusted severe change effect estimates did not overlap between the ≥60 group (adjusted OR 1.80; 1.61- 2.01) and the <60 group (2.27; 2.05-2.52). Increase in eGFR had a stronger effect in the higher baseline group than the lower baseline group for the absolute but not the percentage change measure (Table 9.9). Using the end eGFR measure, the mortality effect estimates for all eGFR decline measures were greater in the <60 group than the ≥60 group. For both the absolute and percentage measures of the severest change category, the difference in risk between the ≥60 and <60 end eGFR group was 57% and 70% respectively and the confidence intervals did not overlap, which indicated the stratified comorbid effects of CKD severity change by the end eGFR level. The effect estimate associated with an increase in eGFR was again more influenced by end eGFR for the absolute change measure having smaller effects in the <60 eGFR group.

Change adjusted by baseline start and end eGFR: In the adjusted model without the baseline eGFR adjustment, the most severe category of change using the absolute (OR 1.83; 1.73-1.94) and percentage change (2.14; 2.00-2.28) measures differed in strength and confidence intervals that did not overlap (Table 9.8). Following adjustment for the starting eGFR these estimates became similar in strength with absolute change OR 2.08 (1.96-2.21) and percentage change 2.10 (1.96-2.24) (Table 9.10). Adjustment for the end eGFR, also resulted in similar effect estimates with absolute change OR 1.78 (1.67-1.88) and percentage change 1.77 (1.66-1.90).

9.5 Chapter summary

The three comorbidities DM, COPD and CKD were significantly and independently associated with all-cause mortality in the non-selected general practice population of HF. These associations were not explained by confounders based on routinely collected clinical data. Whilst the comorbidities had similar strength of associations with mortality, there were differences according to whether their onset was before or after the HF index date. When each of the comorbidities was stratified by categories of recent severity and severity change during the course of their HF, there were significant adjusted associations with increased mortality. The associations between increasing severity categories of HbA1c or eGFR and mortality were curvi-linear. The strongest risk estimates were in the higher HbA1c and lower eGFR categories levels. The HF-COPD group with FEV_1 measures had a higher risk of mortality than the total comorbid COPD group and this risk increased with decreasing categories of FEV_1. A recent decrease in HbA1c, eGFR or FEV_1 over one year were all significantly associated with mortality. Adjustment for CKD severity change using the start or end eGFR measure diminished the associations but they remained significant and independent demonstrating the additional prognostic importance of comorbidity severity change to comorbidity severity alone. As eGFR is the only severity measure routinely collected in all HF patients this adjustment was not possible for HbA1c or FEV_1 change. Comorbidity severity and change measures using prescribed drugs also provided stratified comorbidity effects for both DM and COPD on mortality outcomes. These findings will be discussed in detail in Chapter 13.

Tables

Table 9.1 Summary of confounders indicated in the four steps

Potential confounder	Diabetes	COPD	CKD
Person and socio-demographic factors			
Age	SR*# ł	SR*# ł	SR*# ł
Gender	SR*	*	SR*#
White	SR		SR
Deprivation	*	*	*
HF factors			
Left ventricular ejection fraction	SR	H	SR
Ischaemic aetiology	SR		SR
Other aetiology			SR
Left ventricular systolic dysfunction	SR		
Brain Naturetic Peptide			SR
New York Heart Association class	SR		SR
Cardiovascular Drugs			
Beta-blocker	SR*#	SR *ł	SR
ACEi	SR*	*#	SR *
ARB	SR*	*	SR *
Statin	SR		SR
Diuretics	SR*	H	SR
Digoxin	SR		
Spironolactone	SR	H	SR
Comorbidities			
Renal disease	SR*	*ł	-
Diabetes	-	SR*	SR*
COPD	SR*	-	*#
Vascular disease			SR
Cerebrovascular accident	SR		SR
Hypertension	SR		SR
Charlson index	SR		
Atrial fibrillation	SR		SR
Anthropometric and Clinical observations			
Systolic blood pressure	SR	*	SR*
Diastolic blood pressure	*	*	*
Heart rate	SR	SR	SR
Oedema	SR		
Dyspnoea	SR		
Elevated jugular venous pressure	SR		
Sinus rhythm		SR	
Body mass index	SR*ł	SR*ł	SR*
Cholesterol	SR*	*	SR*
Glucose level	SR		
Blood Urea Nitrogen (mg/dL)	SR		SR
Potassium			SR
Haemoglobin	*ł	*ł	SR*ł
Creatinine	SR	SR	SR
Estimated glomerular filtration rate	*ł	*ł	SR*
Lifestyle factors			
Smoking	SR*	SR*	SR*

Alcohol status	*	*	*

Red text, available factors extracted from CPRD. Additional factors were those indicated by the systematic review. SR, Systematic review (step 1),* difference identified between comorbid and non-comorbid group (step 2), # ≥10% change in comorbidity effect observed in at least one strata of the potential confounder (step 3) ł association identified by ≥10% change in OR of comorbid exposure by adjustment of potential confounder (step 4).

Table 9.2 Unadjusted estimates of association between the potential confounders and all-cause mortality

	All-cause mortality OR (95% CI)
Person and socio-demographic factors	
Age per year	1.06 (1.05-1.06)
Women	1.12 (1.10-1.16)
IMD quintile*	
1	1.0
2	1.08 (1.02-1.14)
3	1.15 (1.08-1.22)
4	1.10 (1.04-1.17)
5	1.15 (1.07-1.22)
Anthropometric and clinical factors	
BMI (per Kg/m2)	0.94 (0.94-0.95)
Normal	1.0
Underweight	1.81 (1.70-1.92)
Overweight	0.66 (0.64-0.68)
Obese	0.51 (0.49-0.53)
Cholesterol (per mmol/L)	0.93 (0.92-0.94)
Haemoglobin (per g/dL)	0.78 (0.77-0.78)
Systolic BP (per mmHg)	0.99 (0.99-0.99)
Diastolic BP (per mmHg)	0.98 (0.98-0.98)
Lifestyle factors	
Smoking status	
No	1.0
Ex	0.95 (0.92-0.98)
Yes	1.03 (0.99,1.08)ł
Alcohol status	
No	1.0
Ex	0.97 (0.91-1.03)ł
Yes	0.85 (0.82-0.87)
Drug factors	
Not on beta-blocker	1.75 (1.70-1.80)
Not on ACEi	1.66 (1.62-1.71)
Not on ARB	1.68 (1.61-1.75)
Diuretic	1.30 (1.25-1.34)
Comorbidity exposures	
Diabetes	1.09 (1.05-1.12)
COPD	1.41 (1.36-1.46)
eGFR <60ml/min)	1.77 (1.71-1.82)
eGFR (per ml/min/m2)	0.98 (0.98-0.98)

*IMD in 79,383 patients only. IMD, index multiple deprivation (1=least deprived, 5=most deprived); BMI, body mass index; BP, blood pressure; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate.

Table 9.3 Summary of the steps of adjustment for each comorbidity

Comorbidity	Step 1	+ Step 2	+ Step 3	+ Step 4
COPD	Age, beta-blocker, BMI, eGFR or renal disease, ACEi/ARB, Hb	DM, smoking	Gender, diuretics, blood pressure, cholesterol, alcohol	Deprivation
DM	Age, eGFR, BMI, Hb, beta-blocker	Gender, ACEi/ARB, diuretics, COPD, cholesterol, smoking	Alcohol, blood pressure	Deprivation
CKD	Age, Hb, gender, COPD	DM, ACEi/ARB, systolic/diastolic, BMI, smoking, eGFR, cholesterol	beta-blocker, diuretics, alcohol	Deprivation

BMI, body mass index; Hb, haemoglobin; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; HbA1c, glycated haemoglobin; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; COPD, chronic obstructive pulmonary disease

Table 9.4 Liklihood ratio tests for current age

Multivariable model	Log likelihood	LR test
a Age	-30016.98	
b Age+Age^2	-29963.72	<0.001 a nested in b
c Age+Age^2+Age^3	-29974.32	0.370 b nested in c

Table 9.5 Likelihood ratio tests for eGFR

Multivariable	Log likelihood	LR test
a eGFR	-29917.303	
b eGFR+eGFR^2	-29532.247	<0.0001 a nested in b
c eGFR+eGFR^2+eGFR^3	-29522.713	<0.0001 b nested in c

Table 9.6 Associations between diabetes exposures and all-cause mortality in HF

Diabetes exposures measures	OR (95% CI) Unadjusted	OR (95% CI) Adjusted[1]	OR (95% CI) Adjusted[2]	OR (95% CI) Adjusted[3]	OR (95% CI) Sub-analysis[4]
Diabetes status *110,505 observations					
No diabetes (ref)	1.0	1.0	1.0	1.0	1.0
Diabetes anytime	1.09 (1.06-1.12)	1.26 (1.21-1.30)	1.24 (1.19-1.29)	1.28 (1.23-1.33)	1.26 (1.19-1.33)
No diabetes (ref)	1	1	1	1	1
Diagnosis before HF	1.14 (1.10-1.18)	1.28 (1.23-1.33)	1.26 (1.21-1.31)	1.31 (1.26-1.37)	1.29 (1.22-1.37)
Diagnosis after HF	0.85 (0.79-0.93)	1.15 (1.05,-1.25)	1.12 (1.03-1.22)	1.13 (1.03-1.23)	1.08 (0.95-1.23)
Diabetes defined by HbA1c severity *106,790 observations					
No diabetes (ref)	1.0	1.0	1.0	1.0	1.0
<5.5%	1.55 (1.37-1.76)	1.34 (1.16-1.53)	1.25 (1.09-1.45)	1.34 (1.16-1.54)	1.30 (1.05-1.60)
5.5-6.4%	1.18 (1.11-1.26)	1.24 (1.16-1.33)	1.22 (1.14-1.30)	1.24 (1.16-1.33)	1.26 (1.14-1.39)
6.5-7.5%	1.01 (0.96-1.06)	1.16 (1.10-1.23)	1.14 (1.08-1.21)	1.19 (1.12-1.26)	1.21 (1.12-1.31)
7.6-8.5%	1.01 (0.94-1.09)	1.23 (1.13-1.33)	1.22 (1.12-1.31)	1.27 (1.17-1.37)	1.18 (1.05-1.33)
8.6-9.5%	1.06 (0.96-1.17)	1.44 (1.29-1.60)	1.42 (1.27-1.59)	1.49 (1.34-1.67)	1.44 (1.23-1.69)
>9.5%	0.97 (0.88-1.07)	1.42 (1.28-1.58)	1.39 (1.25-1.55)	1.45 (1.30-1.62)	1.31 (1.12-1.54)
Diabetes defined by drugs severity *110,505 observations					
No diabetes (ref)	1.0	1.0	1.0	1.0	1.0
1: None	1.32 (1.25-1.40)	1.37 (1.29-1.46)	1.34 (1.26-1.43)	1.37 (1.29-1.46)	1.34 (1.23-1.47)
2: Oral only	0.96 (0.93-1.01)	1.13 (1.08-1.19)	1.12 (1.07-1.18)	1.16 (1.10-1.22)	1.15 (1.07-1.24)
3: Oral + Insulin	0.76 (0.69-0.84)	1.14 (1.03-1.27)	1.12 (1.01-1.25)	1.18 (1.07-1.32)	1.13 (0.97-1.32)
2/3 combined: (Oral +/- insulin)	0.93 (0.89-0.97)	1.13 (1.08-1.19)	1.12 (1.07-1.18)	1.17 (1.11-1.22)	1.15 (1.07-1.23)
3: Insulin only	1.35 (1.26-1.44)	1.50 (1.39-1.62)	1.48 (1.37-1.59)	1.56 (1.44-1.68)	1.52 (1.36-1.70)
Diabetes defined by HbA1c severity change *102,990 observations					
No diabetes (ref)	1.0	1.0	1.0	1.0	1.0
<1% change	1.04 (1.00-1.08)[†]	1.19 (1.13-1.24	1.17 (1.11-1.23)	1.22 (1.16-1.28)	1.21 (1.13-1.29)
>1% increase	1.03 (0.95-1.11)[†]	1.26 (1.16-1.38)	1.24 (1.13-1.35)	1.28 (1.17-1.40)	1.21 (1.07-1.38)
>1% decrease	1.30 (1.21-1.39)	1.48 (1.37-1.60)	1.44 (1.33-1.55)	1.49 (1.37-1.61)	1.50 (1.34-1.69)
Diabetes defined by drugs severity change *110,505 observations					
No diabetes (ref)	1.0	1.0	1.0	1.0	1.0
No drug category change	1.04 (1.01-1.08)	1.20 (1.15-1.25)	1.19 (1.14-1.24)	1.23 (1.19-1.29)	1.23 (1.16-1.30)
Increase in drug category	1.06 (0.96-1.16)[†]	1.37 (1.24-1.52)	1.32 (1.19-1.46)	1.33 (1.20-1.48)	1.20 (1.04-1.40)

Decrease in drug category	2.55 (2.27-2.87)	2.43 (2.14-2.77)	2.26 (1.98-2.58)	2.30 (2.01-2.63)	2.18 (1.79-2.66)

Ref, reference group.* each unadjusted and adjusted measure was based on the same number of observations. There was complete data for comorbidity status and all confounders with the exception of eGFR. All associations excluded those without eGFR. For the comorbidity severity exposures, associations excluded the comorbid groups without the severity exposure or that had eGFR missing. Ref, reference group.

[1]Adjusted for age, eGFR, eGFR^2, BMI, BMI^2, Hb, Hb^2, beta-blocker

[2]Adjusted further for gender, ACEi or ARB, Diuretic, COPD, cholesterol and smoking

[3] Adjusted further for alcohol, systolic, systolic^2

[4] Adjusted further for deprivation (44% missing data)

Table 9.7 Associations between comorbid COPD in HF and all-cause mortality

COPD exposures	Unadjusted OR (95% CI)	Adjusted[1] OR (95% CI)	Adjusted[2] OR (95% CI)	Adjusted[3] OR (95% CI)	Sub-analysis[4] OR (95% CI)
COPD status					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
COPD anytime	1.41 (1.36-1.47)	1.47 (1.41-1.54)	1.40 (1.34-1.46)	1.35 (1.29-1.41)	1.32 (1.23-1.40)
No COPD (ref)	1	1	1	1	1
COPD diagnosis before HF	1.43 (1.37-1.49)	1.45 (1.38-1.52)	1.37 (1.30-1.44)	1.32 (1.25-1.38)	1.31 (1.22-1.41)
COPD diagnosis after HF	1.38 (1.28-1.50)	1.58 (1.45-1.71)	1.50 (1.38-1.63)	1.46 (1.34-1.60)	1.34 (1.18-1.52)
COPD FEV1 severity					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
1: FEV_1 ≥80% normal	1.91 (1.69-2.17)	1.91 (1.66-2.19)	1.81 (1.57-2.07)	1.73 (1.50-1.99)	1.71 (1.40-2.08)
2: FEV_1 50-79% normal	1.87 (1.73-2.01)	1.97 (1.81-2.13)	1.86 (1.71-2.02)	1.76 (1.61-1.91)	1.76 (1.73-2.01)
3: FEV_1 30-49% normal	2.53 (2.32-2.75)	2.55 (2.32-2.80)	2.43 (2.21-2.68)	2.31 (2.09-2.54)	2.48 (2.32-2.75)
4: FEV_1 <30% normal	2.94 (2.52-3.42)	3.51 (2.97,4.15)	3.35 (2.83-3.96)	3.14 (2.65-3.73)	3.21 (2.52-3.42)
COPD drugs severity					
Classification 1					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
No drugs	1.33 (1.20-1.47)	1.37 (1.23-1.53)	1.29 (1.15-1.44)	1.28 (1.14-1.43)	1.23 (1.05-1.45)
Short term inhalers only	1.29 (1.13-1.46)	1.24 (1.08-1.42)	1.15 (1.00-1.31)	1.12 (0.97-1.29)	1.12 (0.92-1.37)
Monotherapy	1.20 (1.08-1.33)	1.20 (1.07-1.34)	1.12 (1.00-1.26)	1.07 (0.95,1. 20)	0.99 (0.84-1.17)
Dual therapy	1.05 (0.96-1.16)	1.10 (1.00-1.21)	1.05 (0.95-1.16)	1.02 (0.92-1.12)	1.01 (0.88-1.16)
Triple therapy	1.21 (1.11-1.32)	1.31 (1.19-1.44)	1.24 (1.13-1.37)	1.18 (1.07-1.31)	1.26 (1.09-1.45)
Oral steroids but no oxygen	1.85 (1.73-1.98)	1.97 (1.83-2.12)	1.89 (1.75-2.03)	1.83 (1.69-1.97)	1.78 (1.59-1.99)
On oxygen	3.20 (2.74-3.74)	3.31 (2.79-3.93)	3.22 (2.72-3.82)	2.94 (2.47-3.50)	2.18 (2.07-3.82)
Classification 2					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
No drugs	1.33 (1.20-1.47)	1.37 (1.23-1.53)	1.29 (1.14-1.89)	1.28 (1.14-1.43)	1.23 (1.04-1.45)
Inhalers only	1.17 (1.11-1.23)	1.20 (1.14-1.27)	1.14 (1.07-1.21)	1.09 (1.03-1.16)	1.09 (1.00-1.19)
Oral steroids but no oxygen	1.85 (1.73-1.98)	1.97 (1.83-2.12)	1.88 (1.75-2.03)	1.83 (1.69-1.97)	1.78 (1.59-1.99)
On oxygen	3.20 (2.74-3.74)	3.31 (2.79-3.93)	3.22 (2.72-3.83)	2.94 (2.47-3.50)	2.18 (2.07-3.82)
Classification 3					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
No steroids or oxygen	1.19 (1.14-1.25)	1.23 (1.17-1.30)	1.17 (1.11-1.23)	1.13 (1.07-1.19)	1.11 (1.03-1.20)
Oral steroids but no oxygen	1.79 (1.68-1.91)	1.97 (1.83-2.12)	1.88 (1.75-2.03)	1.83 (1.69-1.97)	1.78 (1.59-1.99)

On oxygen	3.14 (2.74-3.60)	3.31 (2.79-3.93)	3.23 (2.72-3.83)	2.94 (2.47-3.51)	2.81 (2.07-3.82)
COPD FEV1 severity change					
Classification 1: GOLD stage					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
Gold stage same or better	2.33 (2.17-2.51)	2.41 (2.22-2.62)	2.29 (2.11-2.49)	2.15 (1.97-2.34)	2.14 (1.89-2.42)
Gold stage worse (at least one stage)	3.07 (2.66-3.55)	3.05 (2.61-3.57)	2.87 (2.45-3.37)	2.70 (2.30-3.17)	3.07 (2.39-3.94)
Classification 2: 10% change					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
<10% change	2.41 (2.23-2.61)	2.46 (2.26-2.68)	2.33 (2.14-2.54)	2.18 (1.99-2.38)	2.20 (1.93-2.51)
≥10% increase (better)	2.36 (1.98-2.82)	2.44 (2.02-2.96)	2.34 (1.93-2.83)	2.22 (1.82-2.70)	2.29 (1.70-3.08)
≥10% decrease (worse)	2.81 (2.39-3.30)	2.92 (2.46-3.48)	2.74 (2.30-3.26)	2.60 (2.17-3.11)	2.73 (2.09-3.57)
Classification 3: 5% change					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
<5% change	2.36 (2.14-2.60)	2.46 (2.21-2.74)	2.32 (2.09-2.59)	2.20 (1.97-2.45)	2.22 (1.89-2.62)
≥5% increase (better)	2.27 (1.99-2.59)	2.35 (2.03-2.71)	2.24 (1.94-2.59)	2.10 (1.82-2.44)	2.24 (1.80-2.79)
≥5% decrease (worse)	2.80 (2.49-3.14)	2.78 (2.45-3.15)	2.63 (2.32-2.98)	2.44 (2.14-2.78)	2.43 (2.00-2.95)
COPD drugs severity change					
Classification 1					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
Drug category same or better	1.41 (1.34-1.48)	1.45 (1.38-1.53)	1.39 (1.31-1.46)	1.34 (1.27-1.42)	1.35 (1.24-1.46)
Drug category worse	1.43 (1.35-1.51)	1.51 (1.41-1.60)	1.42 (1.33-1.51)	1.36 (1.27-1.45)	1.27 (1.16-1.40)
Classification 2					
No COPD (ref)	1.0	1.0	1.0	1.0	1.0
No new steroids or oxygen	1.29 (1.23-1.34)	1.33 (1.27-1.40)	1.27 (1.21-1.33)	1.22 (1.16-1.28)	1.22 (1.13-1.31)
New on steroids but no new oxygen	1.92 (1.76-2.09)	2.00 (1.82-2.19)	1.90 (1.73-2.09)	1.84 (1.67-2.03)	1.75 (1.52-2.02)
New on oxygen	3.63 (2.96-4.47)	3.94 (3.15-4.92)	3.83 (3.06-4.80)	3.41 (2.71-4.29)	2.88 (1.94-4.27)

COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 second; pp, percent predicted; ref, reference group. Most recently recorded FEV_1 (pp) prior to the match date was used to measure COPD severity within a maximum of a 3-yr time window. % change for FEV_1 was calculated as the absolute change in pp over one year using the most recent measure (<3yrs) and an earlier FEV_1 measure (<3 years prior to the most recent measure). The FEV_1 change was adjusted for the time interval. Mono, dual or triple therapy; one, two or three of respectively: long acting beta2-antagonist, long acting cholinergic, methylxanthines, inhaled steroid either individually or in combination inhalers *each unadjusted and adjusted measure based on the same number of observations (excluding those with the comorbidity exposure or eGFR confounder missing)

[1]Adjusted for age, no beta-blocker, BMI BMI2, eGFR eGFR2 no ACEi or ARB, Hb Hb2

[2]Adjusted further for diabetes, smoking

[3]Adjusted further for gender, diuretics, systolic, systolic2, cholesterol, alcohol

[4] Adjusted further for deprivation (44% missing data)

Table 9.8 Associations between CKD exposures in HF and all-cause mortality

CKD exposures measures	Unadjusted OR (95% CI)	Adjusted[1] OR (95% CI)	Adjusted[2] OR (95% CI)	Adjusted[3] OR (95% CI)	Sub-analysis[4] OR (95% CI)
Renal **disease diagnosis (eGFR <60)**					
eGFR ≥60 (ref)	1.0	1.0	1.0	1.0	1.0
eGFR <60	1.77 (1.72-1.82)	1.21 (1.17-1.26)	1.21 (1.17-1.25)	1.22 (1.18-1.26)	1.20 (1.14-1.26)
Renal severity (eGFR)					
eGFR mL/min/1.73m^2					
60-89 (ref)	1.0	1.0	1.0	1.0	1.0
>105	1.59 (1.44-1.76)	1.95 (1.76-2.17)	1.68 (1.50-1.87)	1.64 (1.47-1.83)	1.78 (1.52-2.09)
90-105	1.09 (1.01-1.17)	1.31 (1.21-1.41)	1.22 (1.13-1.32)	1.20 (1.11-1.30)	1.18 (1.06-1.33)
45-59	1.30 (1.25-1.35)	1.03 (0.99-1.08)	1.04 (1.00-1.09)	1.04 (1.00-1.09)	1.01 (0.95-1.07)
30-44	1.99 (1.91-2.07)	1.33 (1.27-1.39)	1.30 (1.24-1.36)	1.30 (1.25-1.37)	1.31 (1.23-1.40)
15-29	3.60 (3.41-3.80)	2.16 (2.04-2.29)	2.02 (1.91-2.15)	2.05 (1.93-2.18)	2.07 (1.90-2.27)
<15	5.53 (4.94-6.18)	3.70 (3.28-4.17)	3.15 (2.78-3.56)	3.26 (2.87-3.69)	3.29 (2.72-3.97)
eGFR severity change					
Classification 1					
Absolute change					
0-5mls decrease (ref)	1.0	1.0	1.0	1.0	1.0
Any increase	1.13 (1.09-1.18)	1.19 (1.05-1.06)	1.14 (1.09-1.20)	1.14 (1.09-1.20)	1.13 (1.06-1.21)
6mls to 15mls decrease	1.24 (1.18-1.30)	1.22 (1.16-1.28)	1.19 (1.13-1.25)	1.18 (1.12-1.25)	1.17 (1.09-1.26)
>15mls decrease	1.98 (1.88-2.09)	1.97 (1.87-2.09)	1.85 (1.75-1.96)	1.83 (1.73-1.94)	1.93 (1.77-2.10)
Classification 2					
Percentage change					
0-5% decrease (ref)	1.0	1.0	1.0	1.0	1.0
Any % increase	1.37 (1.30-1.44)	1.31 (1.24-1.39)	1.25 (1.18-1.32)	1.25 (1.18-1.32)	1.23 (1.13-1.33)
6-25% decrease	1.33 (1.26-1.40)	1.23 (1.16-1.30)	1.20 (1.13-1.27)	1.19 (1.12-1.26)	1.17 (1.08-1.27)
>25% decrease	2.87 (2.70-3.04)	2.34 (2.20-2.50)	2.14 (2.01-2.29)	2.14 (2.00-2.28)	2.22 (2.02-2.45)

* each unadjusted and adjusted measure based on the same number of observations (excluding those with the comorbidity exposure or eGFR confounder missing). eGFR, estimated glomerular filtration rate; ref, reference group.

[1]Adjusted for Age, Hb, Hb2, gender, COPD

[2]Adjusted further for Diabetes, no ACEi or ARB, systolic, systolic2, BMI BMI2, smoking, cholesterol

[3]Adjusted further for no beta-blocker, diuretics, alcohol

[4] Adjusted further for deprivation (44% missing data)

Table 9.9 Change in eGFR effect estimates with mortality stratified by start baseline and end levels

eGFR severity change	Start eGFR ≥60 OR (95% CI)	Start eGFR <60 OR (95% CI)	End eGFR ≥60 OR (95% CI)	End eGFR <60 OR (95% CI)
Absolute change				
0mls to 5mls decrease (ref)	1.0	1.0	1.0	1.0
Any increase	1.21 (1.09, 1.34)	1.07 (1.01-1.14)	1.32 (1.19-1.47)	1.05 (0.98-1.12)
6mls to 15mls decrease	1.13 (1.02-1.27)	1.30 (1.21-1.41)	1.10 (0.96-1.26)	1.19 (1.11-1.27)
>15mls decrease	1.80 (1.61-2.01)	2.27 (2.05-2.52)	1.40 (1.20-1.64)	1.97 (1.82-2.13)
Percentage change				
0-5% decrease (ref)	1.0	1.0	1.0	1.0
Any % increase	1.21 (1.08-1.35)	1.24 (1.13-1.35)	1.29 (1.15-1.45)	1.19 (1.09-1.29)
6-25% decrease	1.13 (1.01-1.27)	1.30 (1.18-1.42)	1.06 (0.93-1.21)	1.21 (1.12-1.32)
>25% decrease	2.01 (1.76-2.28)	2.34 (2.11-2.59)	1.48 (1.20-1.81)	2.18 (1.99-2.38)

Ref, reference group. All associations were fully adjusted by all remaining covariates. eGFR, estimated glomerular filtration rate (ml/min/m^2). Change was calculated over a year before the match date using the most recent value up to a maximum of 3 years and a prior value between 6 months and 3 years.

Table 9.10 Change effect estimates with mortality adjusted for most recent and prior eGFR

eGFR severity change	Adjusted[3] OR (95% CI)	Adjusted for start eGFR and $eGFR^2$ OR (95% CI)	Adjusted for end eGFR and $eGFR^2$ OR (95% CI)
Absolute change			
0mls to 5mls decrease (ref)	1.0	1.0	1.0
Any increase	1.14 (1.09-1.20)	1.16 (1.11-1.21)	1.24 (1.18-1.30)
6mls to 15mls decrease	1.18 (1.12-1.25)	1.25 (1.19-1.32)	1.18 (1.12-1.24)
>15mls decrease	1.83 (1.73-1.94)	2.08 (1.96-2.21)	1.78 (1.67-1.88)
Percentage change			
0-5% decrease (ref)	1.0	1.0	1.0
Any % increase	1.25 (1.18-1.32)	1.21 (1.14-1.28)	1.24 (1.17-1.31)
6-25% decrease	1.19 (1.12-1.26)	1.19 (1.12-1.26)	1.14 (1.08-1.22)
>25% decrease	2.14 (2.00-2.28)	2.10 (1.96-2.24)	1.77 (1.66-1.90)

eGFR, estimated glomerular filtration rate (ml/min/m^2); ref, reference group. All associations were fully adjusted by all remaining covariates and the start or end eGFR. Due to the non-linear association between eGFR and mortality $eGFR^2$ was also used in the models. Change was calculated over a year before the match date using the most recent value up to a maximum of 3 years and an earlier value between 6 months and 3 years before.

Figures

Figure 9.1 Calibration plots of observed versus predicted fitted effects for current age and mortality

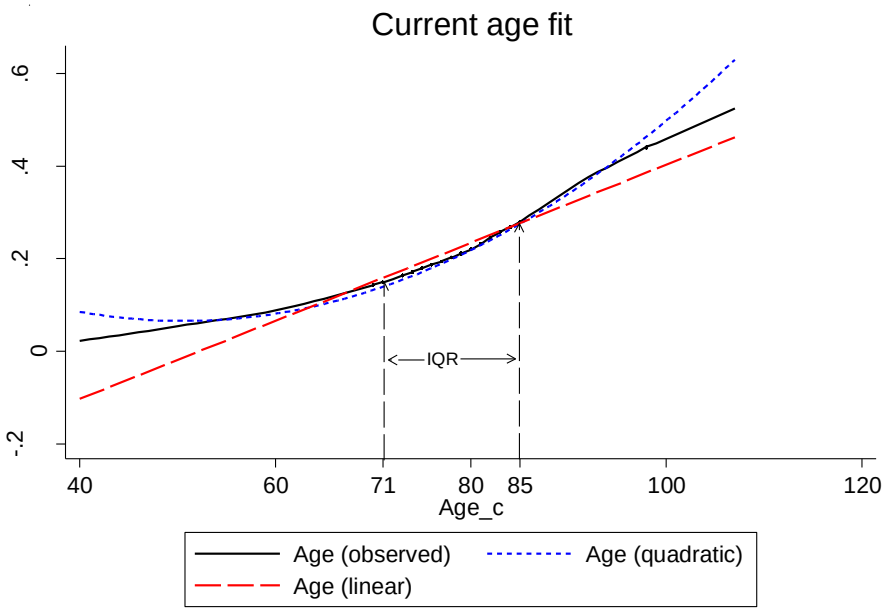
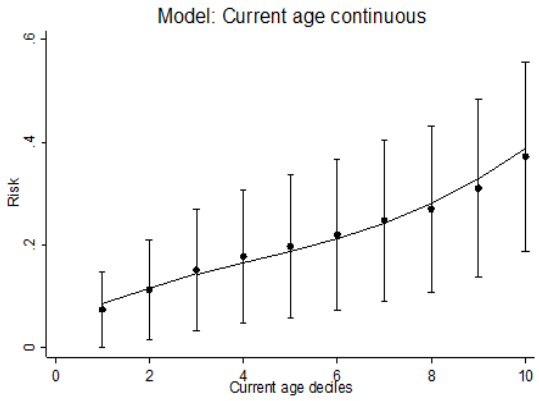


Figure 9.2 Eccles plots of current age and mortality



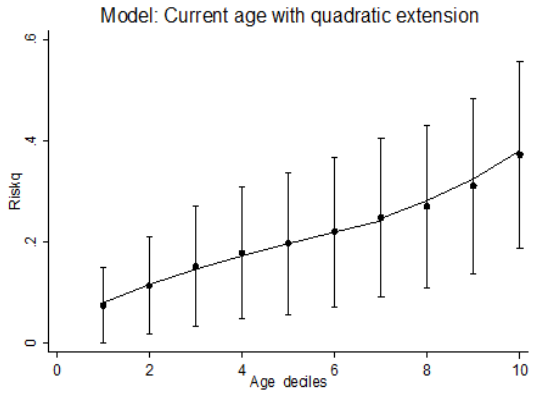


Figure 9.3 Calibration plots of observed versus predicted fitted effects for eGFR

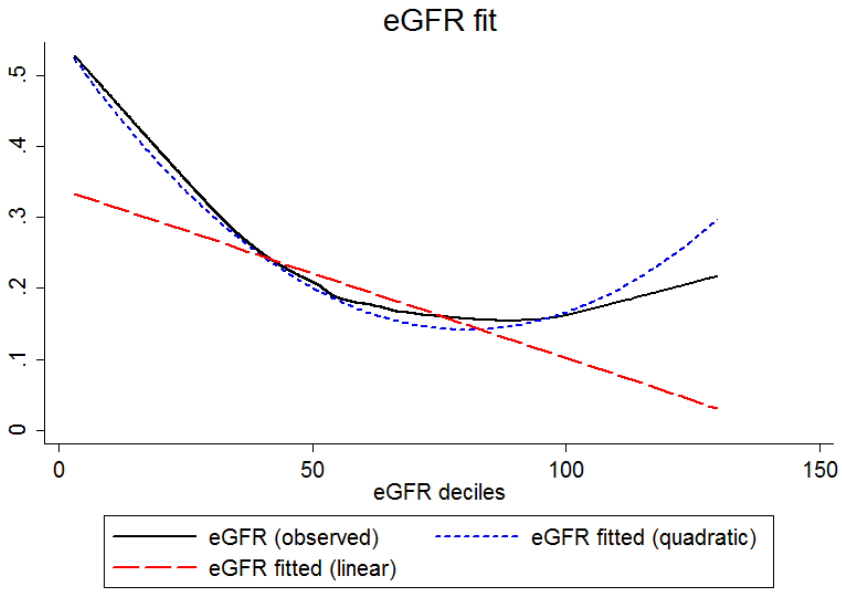


Figure 9.4 Eccles plots of eGFR and mortality

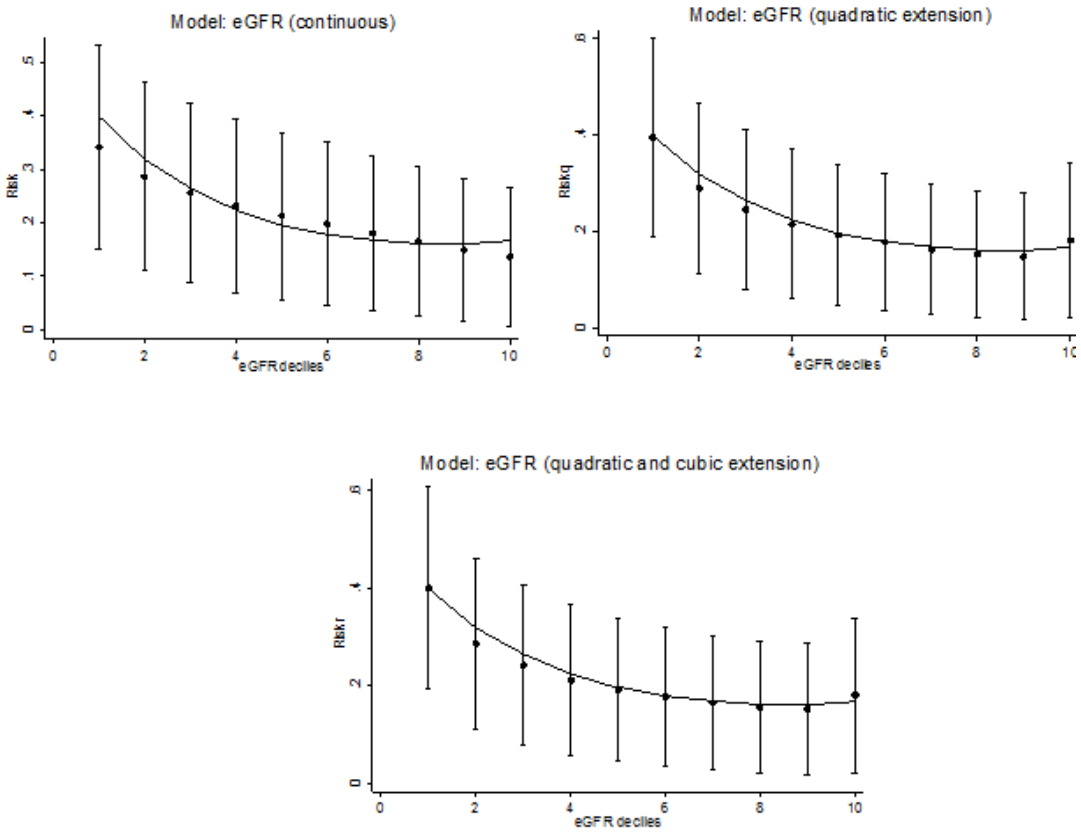


Figure 9.5 Adjusted associations of comorbidities with all-cause mortality

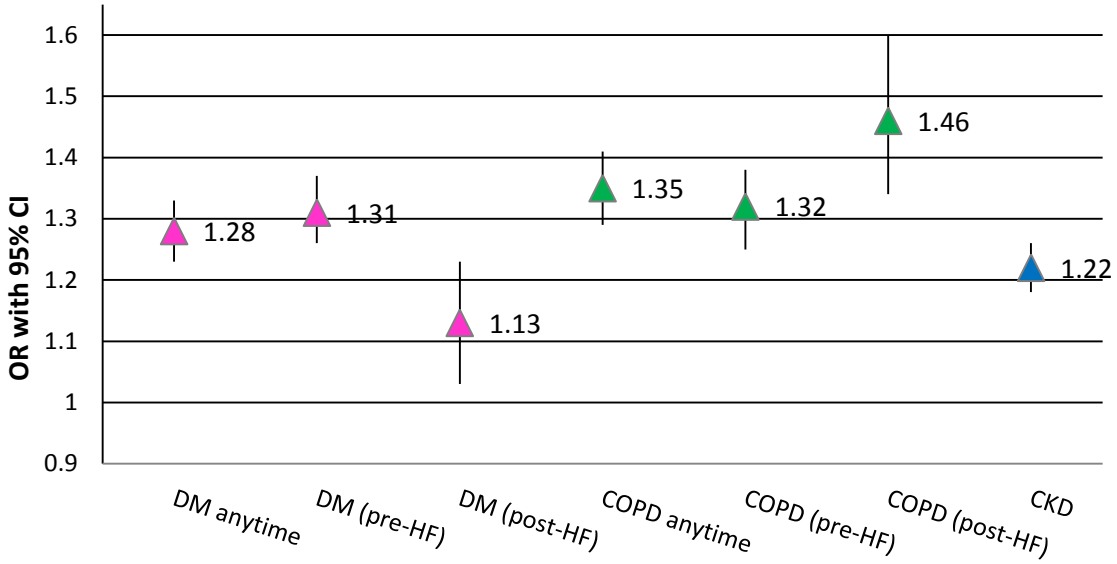


Figure 9.6 Margins plot of HbA1c categories and mortality

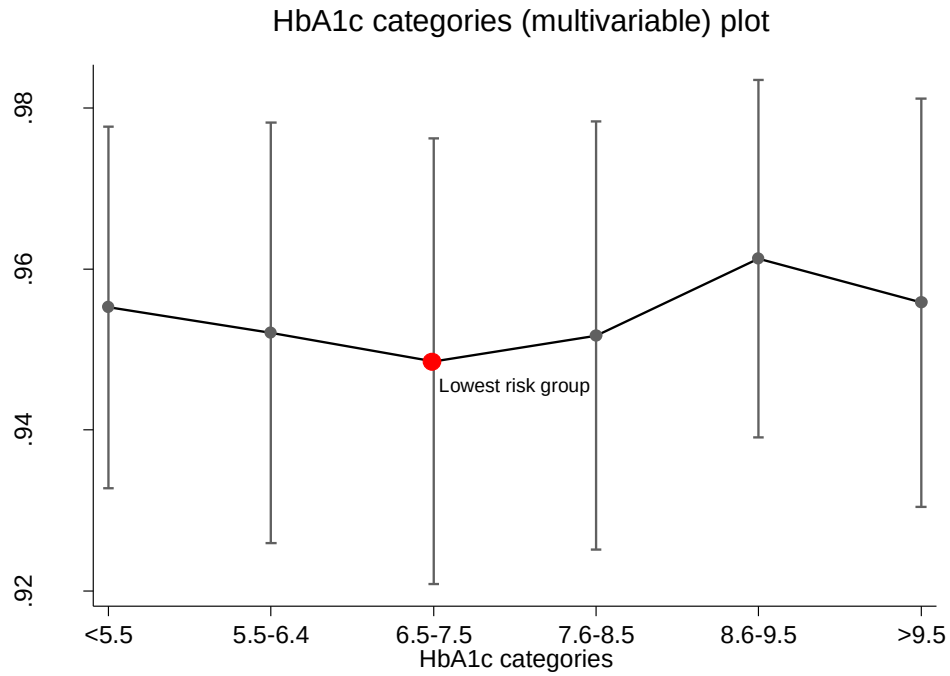


Figure 9.7 Margins plot of HbA1c deciles and mortality

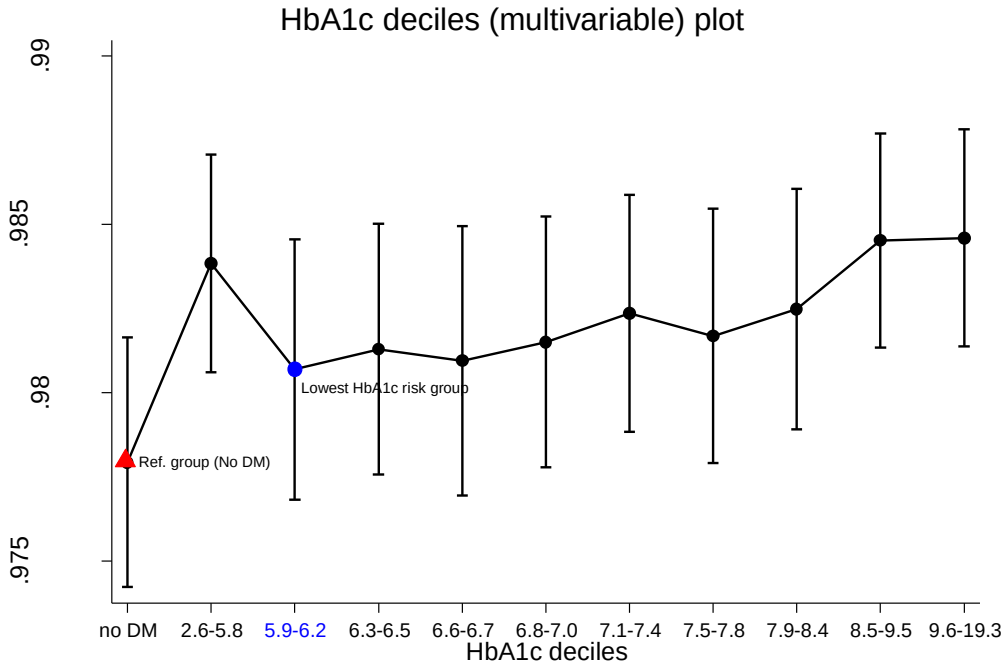


Figure 9.8 Adjusted associations of HbA1c categories and mortality

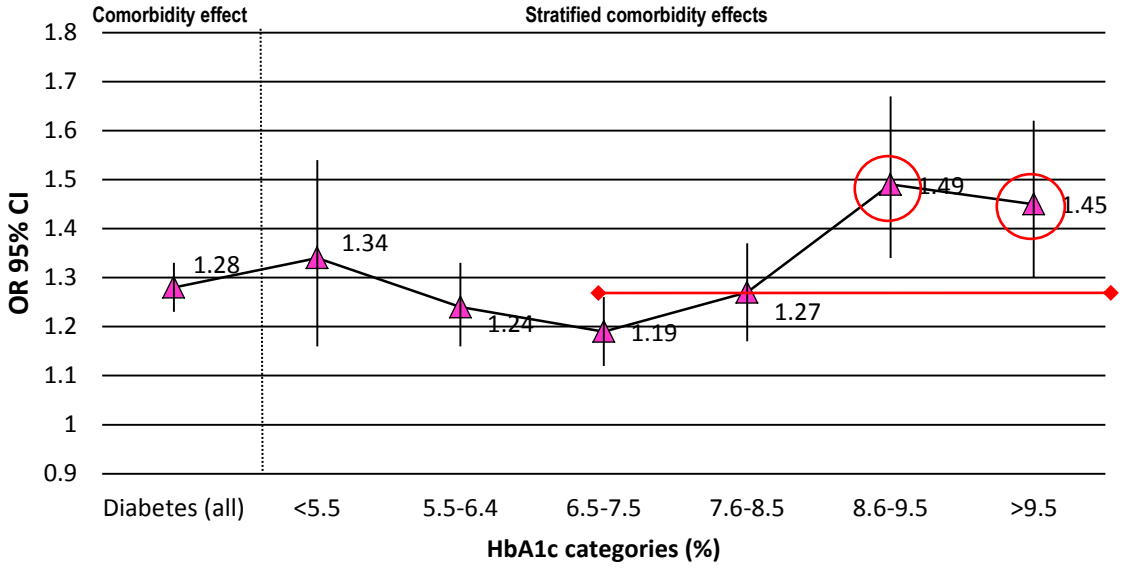


Figure 9.9 Adjusted associations of DM drug severity categories and mortality

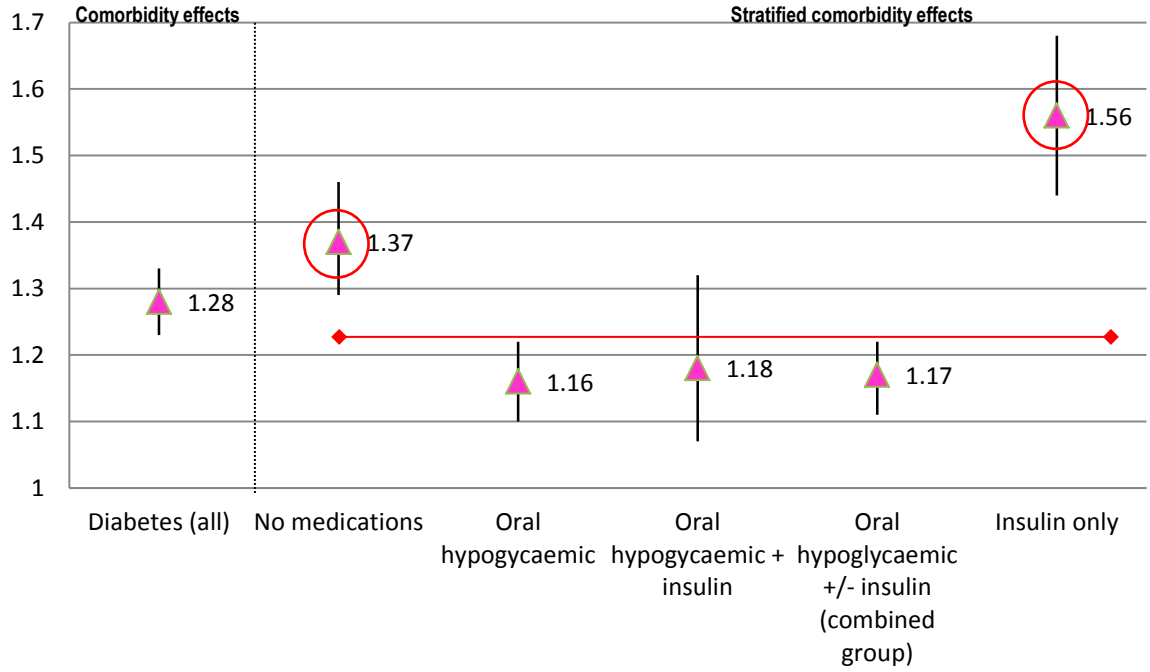


Figure 9.10 Adjusted associations of categories of DM and HbA1c change and mortality

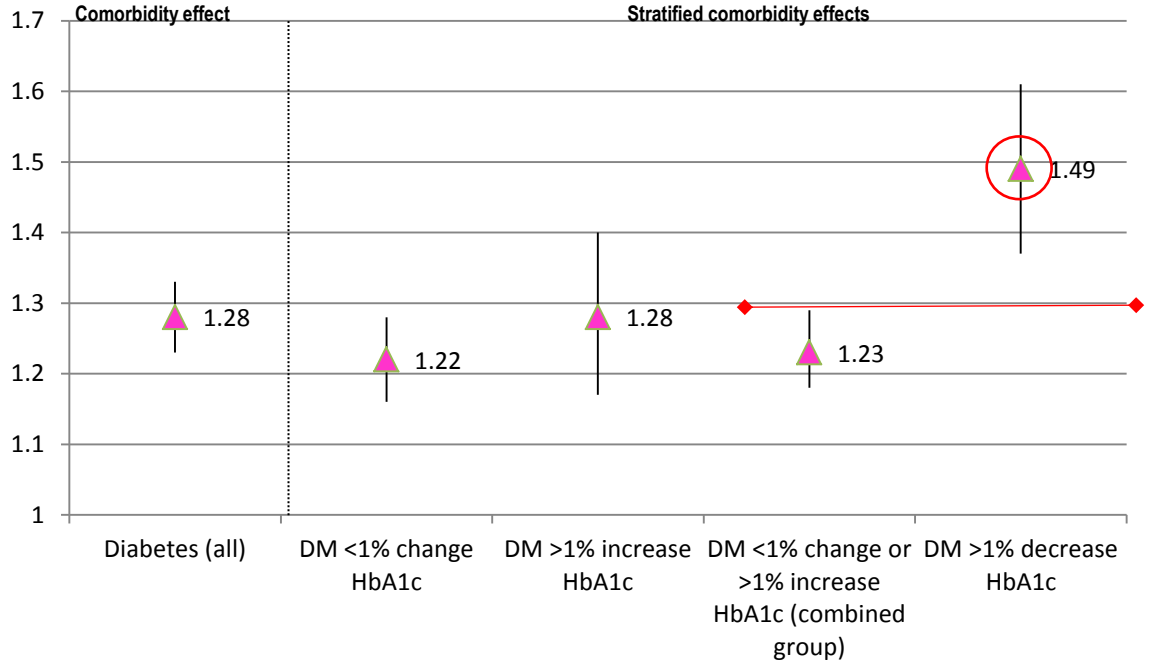


Figure 9.11 Margins plot of COPD GOLD stages and mortality

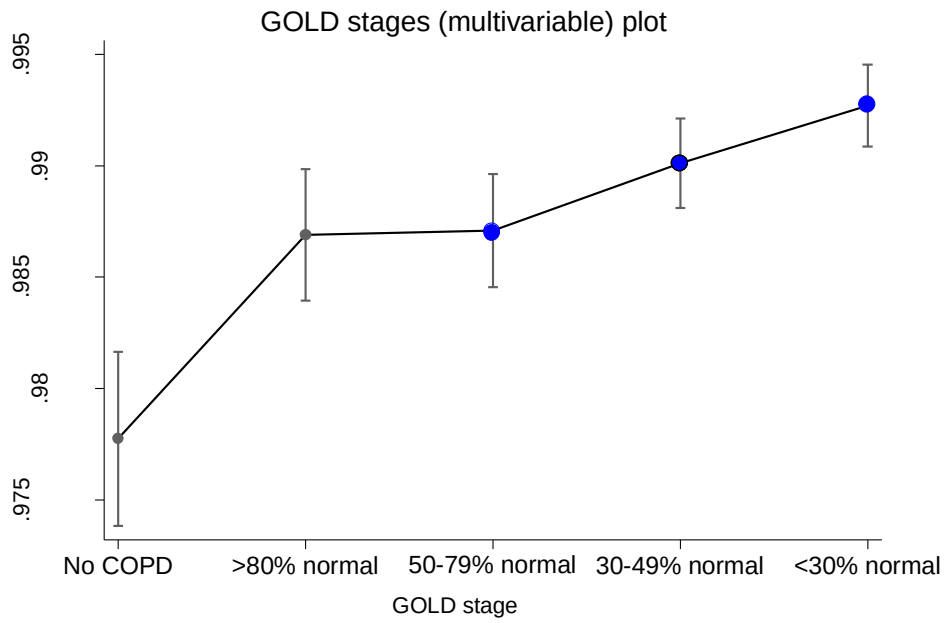


Figure 9.12 Adjusted associations of COPD GOLD stages and mortality

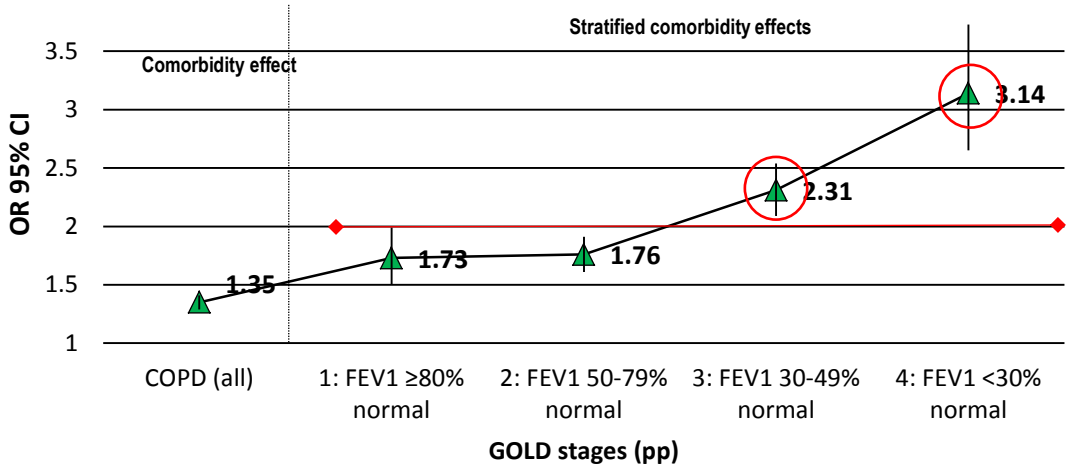


Figure 9.13 Adjusted associations of COPD drug severity categories and mortality

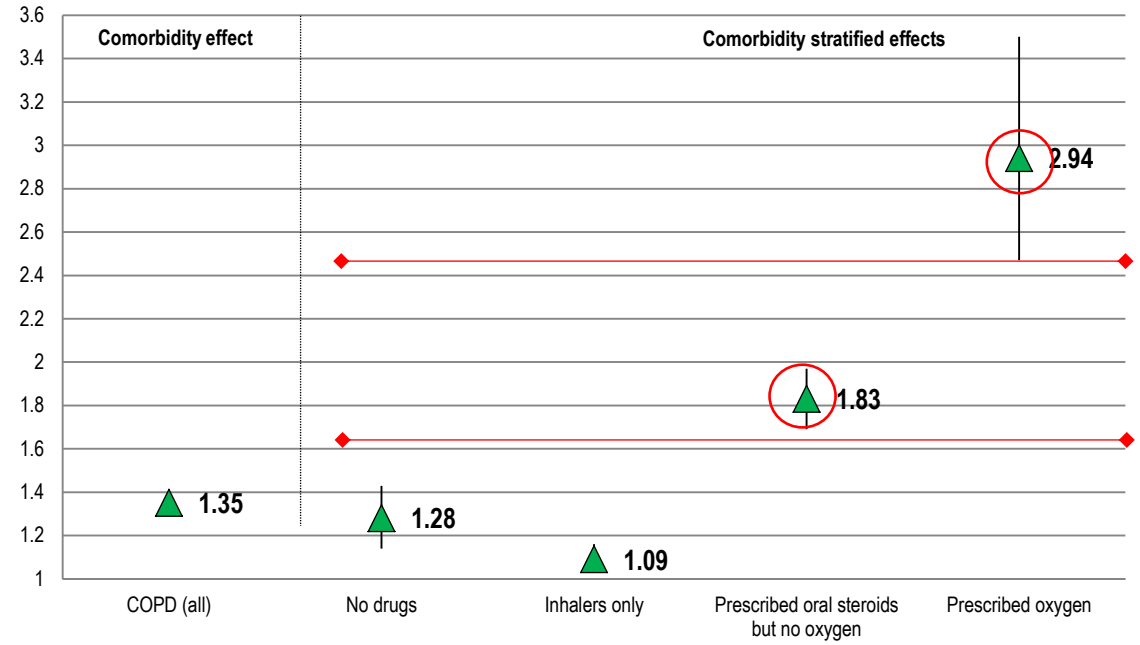


Figure 9.14 Adjusted associations of categories of COPD and FEV_1 change and mortality

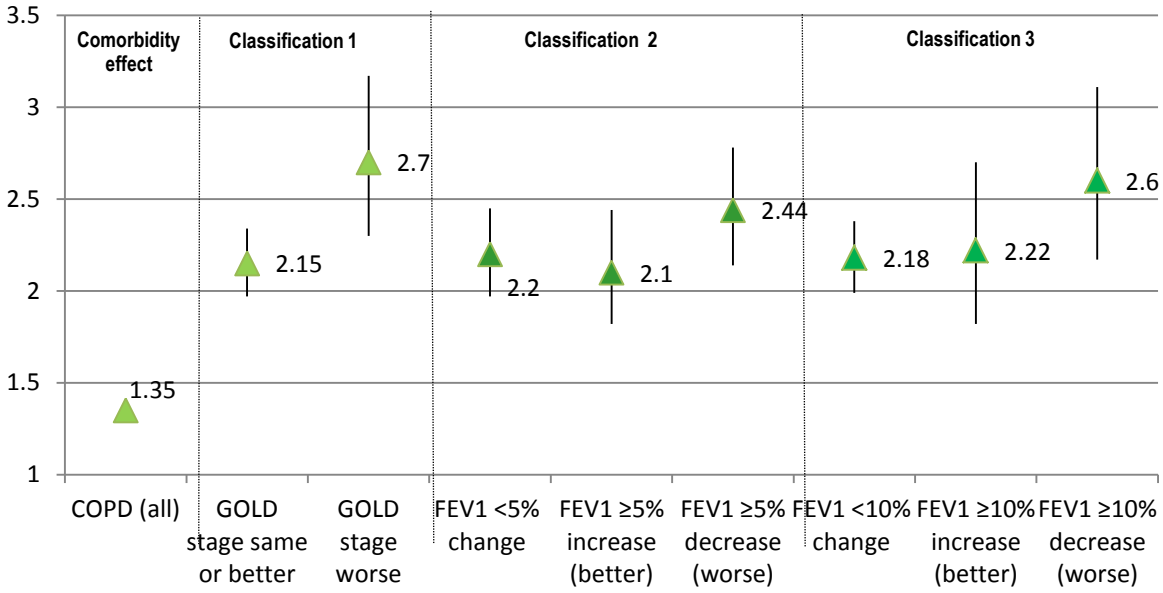


Figure 9.15 Adjusted associations of categories of COPD drug severity change and mortality

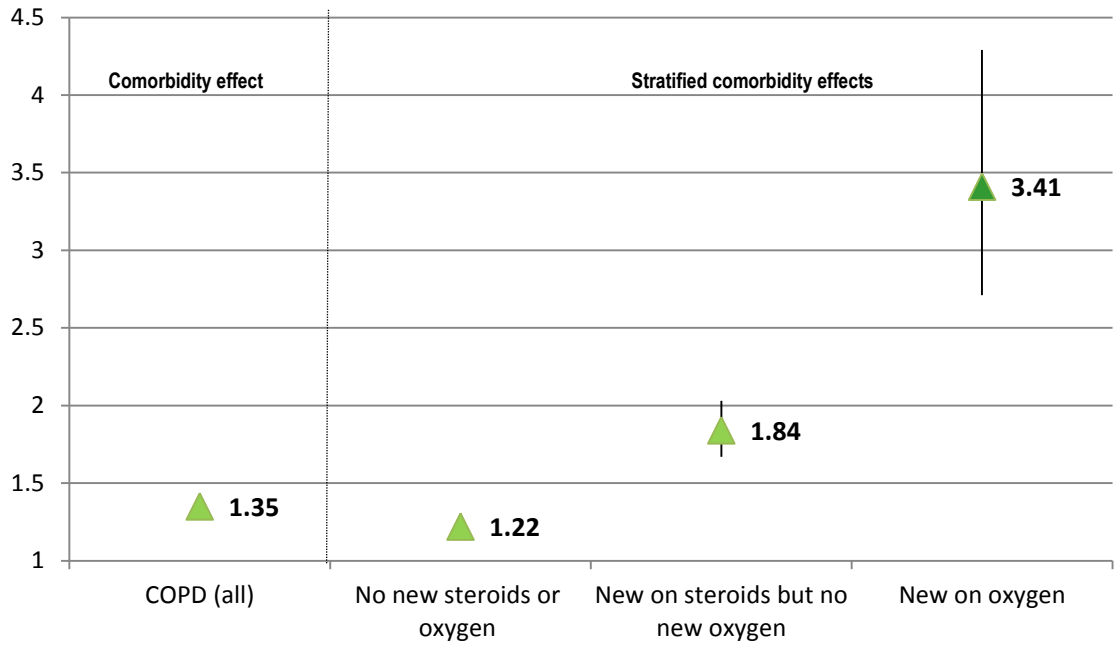


Figure 9.16 Margins plot of eGFR severity categories and mortality

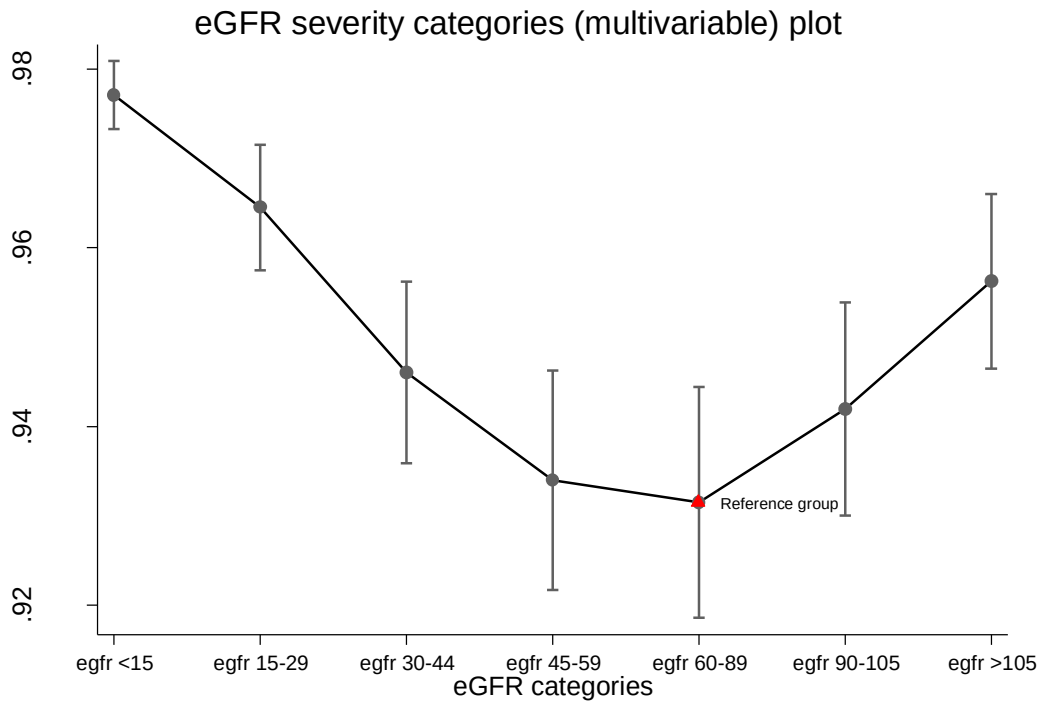


Figure 9.17 Margins plot of eGFR deciles and mortality

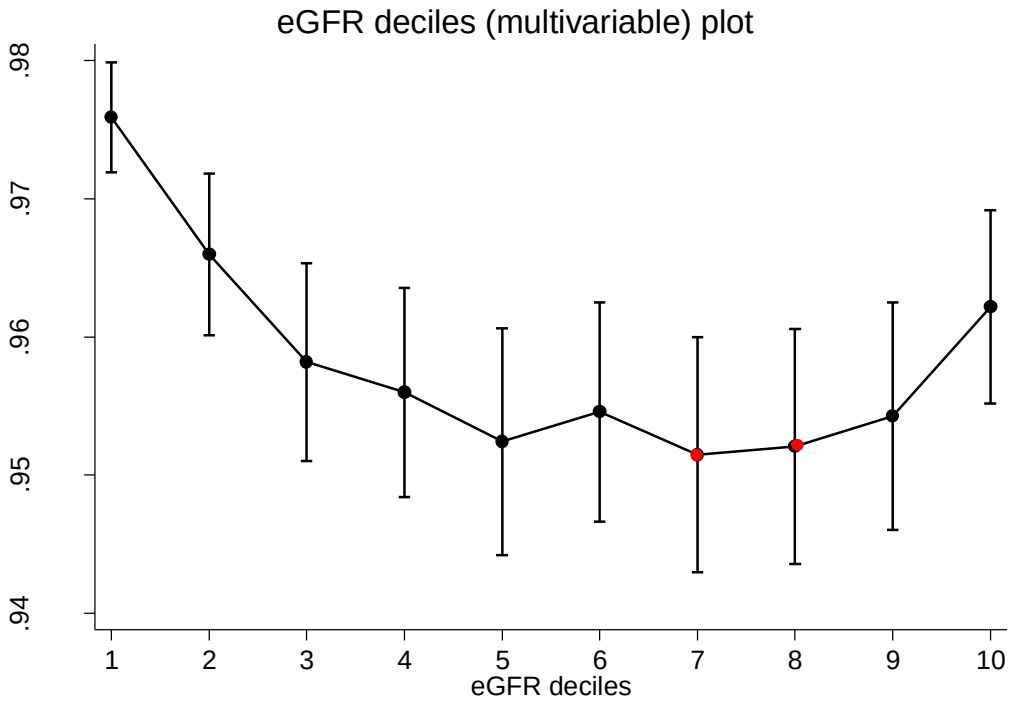


Figure 9.19 Adjusted associations of eGFR categories and mortality

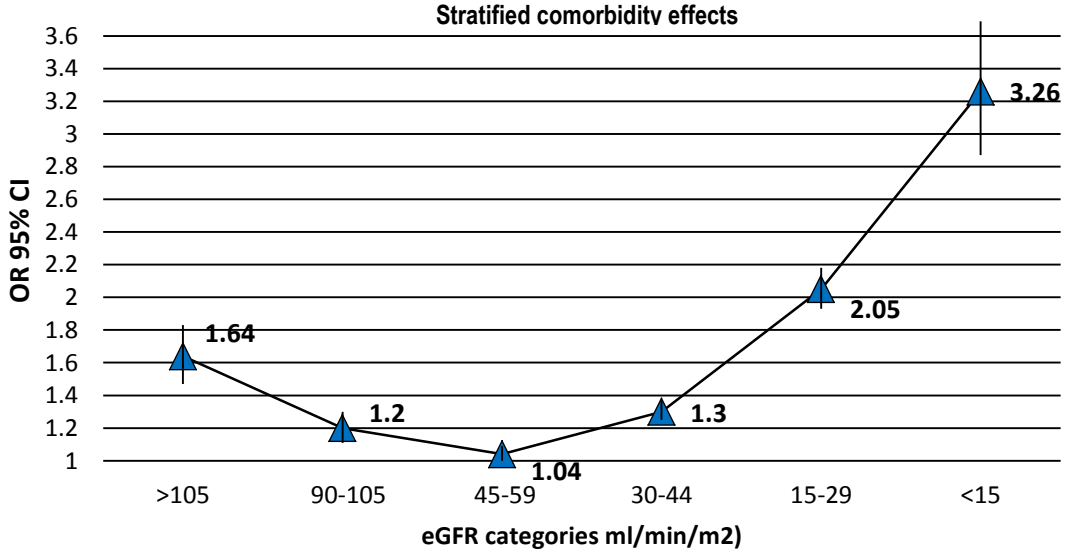
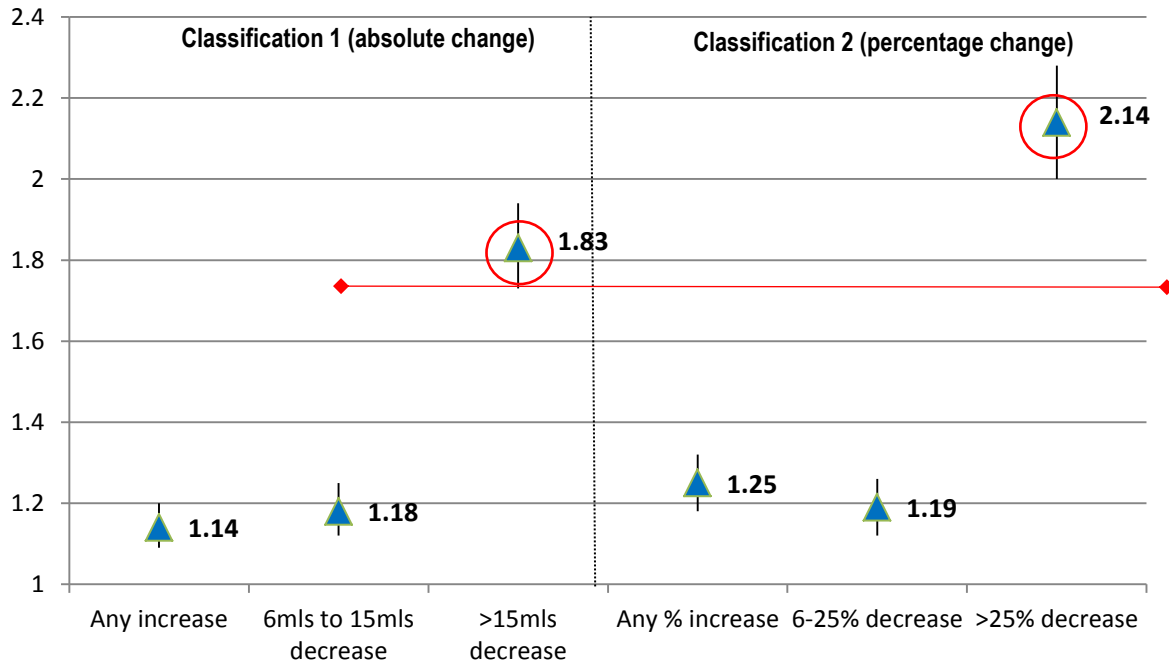


Figure 9.20 Adjusted associations of categories of eGFR change and mortality



Chapter 10 Non-CVD comorbidity prognostic factors in an incident HF general practice population: strength of associations with first hospital admission

This chapter follows on from the descriptive findings (Chapter 8) of the main incident HF CPRD cohort linked to hospital admissions data (HF-HA) and the outcome of first hospital admission. This chapter investigates the comorbidity measures that were found to have the strongest associations with mortality, for their association with first hospital admission. The chapter presents the results of the investigation of the strength of associations between the selected non-CVD comorbidities (COPD, DM and CKD) in HF and first hospital admission. The effects of time-dependent exposures of the three comorbidities on first hospital admission are firstly presented by their status 'present or not' followed by their stratification by categories of severity and change. All effect estimates will be presented as odds ratios (OR) with 95% confidence intervals (CI).

This chapter presents the findings of the five stages to analysis:

(vi) Confounding investigation

(vii) Test of model assumptions: linearity

(viii) Test of model assumptions: collinearity

(ix) Adjusted associations of the comorbidities with first hospital admission

(x) Comorbidity effects stratified by categories of severity and change

10.1 Confounding investigations

The four steps used to investigate confounding in the associations between the three selected comorbidities and mortality in HF were also applied for the outcome of first hospital admission and summarised in Table 10.1. Given that only 2 studies in the systematic review had focused on the association between the three

comorbidities and hospital admission outcome the evidence for mortality was used to indicate potential confounders for hospital admission (step 1; Table key, SR). Also shown in the table are the potential confounders available from the CPRD data extracted, which differed between the comorbid group and the non-comorbid groups in Chapter 8 (step 2; Table key, *). In step 3, each comorbid DM, COPD, CKD effect in HF was now compared in strata of these potential confounders to investigate whether their strength of association with first hospital admission differed across strata (step 3, Table key #). Finally, the three comorbidity effects were then adjusted for each available potential confounder in turn, to further investigate confounding (step 4, Table key ŧ). A 10% difference in the comorbidity effect in steps 3 and 4 was used to indicate confounding(9).

Comorbidity effects within strata of the potential confounders (step 3): The unadjusted associations between the potential confounders and first hospital admission can be seen in Table 10.2. The potential confounders that had the strongest unadjusted associations with first hospital admission were deprivation, prior hospital admission, cholesterol, haemoglobin, smoking and cardiovascular medications. When the unadjusted associations between the three comorbidities and first hospital admission were observed within strata of potential confounders, it was these factors (with the exception of cholesterol, ACEi and beta-blocker) together with age and BMI that resulted in a >10% change to the odds ratio (E-Appendix A37). The relative effect of the comorbidities were strengthened across most age strata for DM and COPD and reduced in all age strata for CKD, with the exception of the youngest age group. The comorbidity effects varied across strata of deprivation and prior hospital admission for all three comorbidities. The effect of DM was stronger in the lower haemoglobin group than the higher haemoglobin group and the effect of CKD was reduced when observed in both strata of haemoglobin. The comorbidity effect was mainly increased in strata of BMI for DM and reduced for CKD and increased in one strata of diuretics for DM and CKD. DM and CKD effects were reduced in the non-smoking strata.

Comorbidity associations with first hospital admission adjusted by each potential confounder (step 4): Prior admission, which had the strongest unadjusted association with first hospital admission after HF diagnosis, was a consistent confounder across all three comorbidities (COPD, DM, CKD) (E-Appendix A38). Each

comorbidity effect was reduced with the adjustment of prior admission across most severity categories with the most marked reduction in the most severe category of the comorbidities.

Haemoglobin also had one of the strongest unadjusted effects on first hospital admission (OR 0.84 per increase g/dL; 95% CI 0.84-0.86). The effect estimates on hospital admission of DM and CKD (which both had a lower mean haemoglobin than the non-DM and non-CKD groups respectively), were reduced following adjustment by haemoglobin (DM, OR 1.33 reduced to 1.26; CKD, OR 1.34 reduced to 1.15. The effect estimates on first admission of the comorbid COPD group with a higher mean haemoglobin than the non-COPD group, were increased following adjustment by haemoglobin (OR 1.36 increased to 1.43 and OR 1.28 increased to 1.49 in the most severe COPD group).

Adjustment of COPD and DM by CKD status had opposing results. The effect of HF-DM, a group with a higher prevalence of CKD than the non DM HF group (Chapter 8) was reduced following adjustment of CKD. The significant effect of HF-COPD, a group with a lower prevalence of CKD than the non COPD HF group was increased following adjustment of CKD.

Summary of the confounders for each of the three comorbidities: All of the available covariates from CPRD were identified as potential confounders for each of the three comorbidities through the four steps of confounding investigation. The same stages of adjustment in Chapter 9 were applied to the comorbidity and hospital admission analyses as follows (Table 10.3):

a) Those identified in step 3 (10% change of the comorbidity effect in any strata of the potential confounder) *OR* step 4 (10% change in the comorbidity effect when adjusted by the potential confounder,

b) Previous evidence of confounding identified in the systematic review studies (step 1) *AND* a difference in the prevalence of the potential confounder in the comorbid disease group compared to the non-comorbid group in the CPRD analyses (step 2),

c) One of step 1 *OR* step 2,

10.2 Test of model assumptions: linearity

The 6 continuous covariates used to adjust the comorbidity measures for the mortality outcome (Chapter 9) were investigated for their linear association with first hospital admission using likelihood ratio tests. The fit of models using different functions of the covariate was also studied using fitted line plots and Eccles plots. The covariates were age, BMI, cholesterol, haemoglobin, systolic blood pressure and eGFR.

For each continuous variable, fully adjusted models with all covariates were derived (Table 10.2) and then quadratic and cubic terms for the continuous variable incrementally added. The models were compared using likelihood ratio tests. Fitted line plots were then used to compare a Lowess line, which closely follows the observed data, with linear and quadratic extension fitted lines of the association between the continuous variable and predicted probabilities of hospital admission. Finally Eccles plots were used to graph deciles of each covariate against predicted probabilities using fully adjusted models with the continuous covariate included and with the additional quadratic and cubic extensions. These plots included Lowess lines to compare the predicted models with the observed data. All linearity tests for the continuous covariates are shown in E-Appendix E17 and summarised in E-Appendix A39 with two examples detailed below.

- Body Mass Index (BMI)

Likelihood ratio tests showed that addition of a quadratic extension to BMI significantly improved the model fit with a small increase in log-likelihood of 8.48 (Table 10.4). Further addition of a cubic term made no significant difference. The fitted line plots showed BMI and risk of hospital admission appeared to show a linear association which increased quite steeply from the line at the highest BMI values approximately >40 kg/m^2 (Figure 10.1). However, there were only few observations of a BMI ≥40 (0.04%) and ≥50 kg/m^2 (0.005%). The interquartile range for BMI is also shown in the figure (23.9 to 31.2 kg/m^2). Eccles plots showed good fit of the adjusted models with (i) BMI and with (ii) BMI plus the quadratic extension (BMI^2) (Figure 10.2). Given that there was minimal improvement in model fit by adding a quadratic extension and the

largest proportion of BMI followed a linear association with the probability of hospital admission, the final modelling decision was to include BMI in its simple linear form.

- Haemoglobin

Likelihood ratio tests showed that addition of a quadratic term of Hb significantly improved the model fit with an increase in log-likelihood of 122.32 (Table 10.5). Further addition of a cubic term made no significant difference. The fitted lines showed that the line with the quadratic extension followed the observed line (Figure 10.3). Eccles plots showed good fit of the adjusted models with Hb plus the quadratic term (Hb^2) (Figure 10.4). Given that there was improvement in model fit by adding a quadratic extension to Hb (Hb^2), the final modelling decision was to include the quadratic terms in the adjusted models.

10.3 Test of model assumptions: collinearity

Most covariates were weakly correlated ($r<0.35$). The strongest correlations again were CKD and eGFR ($r=0.79$) (E-Appendix C17). eGFR continuous was selected for adjustment of DM and COPD measures over CKD, which used a dichotomised eGFR definition. All continuous variables were centered at their means (Age 77 years; cholesterol 5 mmol/ml; Hb 13 g/dL; BMI 28 Kg./m^2; systolic BP 136 mmHg; eGFR 59 ml/min/m^2) which removed the collinearity between the variables with a respective squared term (E-Appendix C18).

10.4 Adjusted associations between DM, COPD, CKD comorbidity in HF and first hospital admission

The comorbidity measures were investigated in detail for the outcome of mortality in Chapter 9. For the current investigation of the sub-sample of the HF cohort that were linked to hospital data, the main comorbidity severity and change exposures that were found to have the strongest associations with mortality, are now investigated for the outcome of first hospital admission.

10.4.1 Diabetes Mellitus

In the incident HF-HA population, the unadjusted association between DM and first hospital admission, compared to no diabetes was OR 1.30 (95% CI 1.25-1.36). This estimate was most influenced by adjustment for the first set of confounders: age, diuretic, prior admission, BMI, Hb and eGFR (adjusted OR 1.23; 1.17-1.30) (Table 10.9). Adjustment of the remaining confounders made little difference to this estimate (OR 1.24; 1.18-1.31) (Table 10.6) which was similar to the association between DM-HF and mortality (OR 1.28; 1.23-1.33).

10.4.1.1 Timing of DM comorbidity

When DM was stratified by prevalent DM at the time of HF diagnosis (pre-HF DM) or incident DM developing after HF diagnosis (post-HF DM), there was a difference in the associations with first hospital admission. The adjusted association for pre-HF DM was OR 1.27 (95% CI 1.20-1.34). The association for post-HF DM was non-significant following full adjustment (OR 0.93; 0.78, 1.10). This pattern of association between pre-HF DM, post HF DM and admission (Figure 10.5) was similar to that found with mortality.

10.4.1.2 Comorbid Diabetes Mellitus severity

- **Physiological severity measure**

HF group with comorbid DM was stratified into groups by glycated haemoglobin (HbA1c) levels. This categorisation used the closest available measure preceding the match date (median 73 [IQR 34-118] days) comparing the HF group with and without diabetes.

Shape of association: The shape of the adjusted associations between categories of HbA1c in the HF-DM group and their respective predicted probabilities of first hospital admission was slightly curved, which is shown in Figure 10.6 using a margins plot which fixes other covariate effects at zero. The lowest risk HF-DM category was HbA1c 5.5%-6.4%. To observe the shape of association more closely, HbA1c was categorised into deciles. There was a trend of increasing risk with higher HbA1c deciles with the highest risk in the 10th decile. There was a distinct group with lower risk of admission in decile 6 (7.1%-7.4%) with the next lowest risk group being decile 2 (5.9%-6.3%) (Figure 10.7).

Strength of associations: Fully adjusted associations between categories of HbA1c in the HF-DM group and first hospital admission, compared to the no DM HF group are shown in Table 10.6. The lowest risk category was HbA1c 5.5%-6.4% (OR 1.12; 1.01-1.25). From this lowest risk category, risk of hospital admission increased with increasing HbA1c level categories. Risk increased in the highest HbA1c category (>9.5%) with an OR 1.64 (1.40-1.93) (Figure 10.8). This was a higher risk than for admission (OR 1.45; 1.30-1.62). The lowest category of HbA1c was non-significant for hospital admission (OR 1.19; 0.90-1.59).

The adjusted associations showed a significant trend for increasing risk of admission from the second HbA1c group (5.5-6.4%) to the highest Hba1c group (>9.5%). The highest risk category of HbA1c (>9.5%) confidence intervals did not overlap with those of the lowest risk categories (Figure 10.8). This translated to an increase in the risk of admission of 64% in this group compared to the non-comorbid group. Increased risk in the HF-DM group with guideline normal HbA1c range (6.5-7.5%), compared to the non-comorbid group was 15%.

- **Drug severity measure**

The HF-DM group were stratified into groups based on DM-related prescribed medications in a four month time-window before the match date. The groups were: (i) DM and no medications, (ii) DM and oral hypoglycaemics plus or minus insulin, (iii) DM and insulin only. These groups were compared to the HF sample without diabetes.

Strength of associations: The adjusted associations (Table 10.6) with the largest effects were the 'oral hyoglycaemics +/- insulin OR 1.33 (95% CI 1.25-1.42) and 'insulin only group' (1.41; 1.25-1.59) (Figure 10.9). The confidence intervals did not overlap with the 'no medications group' (OR 1.00; 0.91-1.09) indicating the stratified comorbid effects of DM drug severity on increased first hospital admission.

10.4.1.3 DM severity change

- **Physiological severity change measure**

The HF-DM group were stratified into three groups according to change in their glycated haemoglobin (HbA1c) level over 6-months before the match date. Change definition was based on the closest available HbA1c measure preceding the match date and a previous measure before the most recent measure (median 161 [IQR 101-212] days). Change over 6 months was calculated and compared to the HF sample without DM.

Strengths of associations: All three categories of HbA1c change in HF were significantly associated with first admission in the adjusted models (Table 10.6) with increasing risk from the <1% change group (OR 1.17; 1.08-1.26) to the >1% decrease group (1.34; 1.18-1.53). There was a significant trend for increasing risk in the adjusted models, but the confidence intervals of all three groups overlapped (Figure 10.10).

Drug severity change measure: The DM group were stratified into three groups based on change in their prescribed DM drug over six months before the match date. This measure used the DM drugs prescribed in a 4-month time-window before first hospital admission compared to a 4-month time-window six months before first admission. Change in drug category over 6 months was categorised into 'no change', an 'increase in at least one drug category' and a 'decrease in at least one drug category'. These groups were compared to the HF sample without diabetes.

Strength of associations: Fully adjusted associations between drug category change and first hospital admission showed the highest risk was in the 'decreased at least one drug category' group with OR 1.60 (1.18-2.17). The adjusted associations for 'no change' were OR 1.22 (1.16-1.29) which strengthened to 1.40 (1.19-1.66) in the 'increased at least one drug category' group. There was a significant linear trend of increasing adjusted risk from the no change to the increase drug category group to the decrease drug category group.

10.4.2 Chronic obstructive pulmonary disease

The adjusted association between HF-COPD and first hospital admission compared to non-comorbid group was OR 1.32 (1.23-1.41) (Table 10.7) which was similar to the association between HF-COPD and mortality (OR 1.35; 1.29-1.41).

10.4.2.1 Timing of COPD comorbidity

When COPD was stratified by prevalent COPD at the time of HF diagnosis (pre-HF COPD) or incident COPD developing after HF diagnosis (post-HF COPD), there was a difference in the associations with first hospital admission. The adjusted association with pre-HF COPD was OR 1.31 (1.22-1.41) which was similar to the association with mortality (OR 1.32; 1.25-1.38). The adjusted association between post-HF COPD and first hospital admission was OR 1.40 (1.13-1.73) (Figure 10.5), which was also similar to the association with mortality (OR 1.46; 1.34, 1.60).

10.4.2.2 Comorbid COPD severity

Due to the low numbers of recently recorded FEV_1 measures in the HF-HA cohort, comorbid COPD severity was restricted to drug based measures

- **COPD drug severity measure**

The HF-COPD group was stratified into three groups according to their prescribed COPD medications in a 4-month time-window before the match date. The groups were: (i) COPD and no oral steroids or prescribed oxygen, (ii) COPD and oral steroids but no prescribed oxygen, and (iii) COPD and prescribed oxygen. These groups were compared to the HF sample without COPD.

Strength of associations: Fully adjusted associations between the HF-COPD drug groups compared to no COPD-HF and hospital admission are shown in Table 10.7. Risk of hospital admission increased from the first category of 'no oral steroids or prescribed oxygen' (OR 1.17; 1.09-1.27) to the 'oral steroids but no prescribed oxygen' group (1.73; 1.53-1.96) to the highest risk in the 'prescribed oxygen group' (2.42; 1.66-3.52). There was a significant trend of increasing risk from the first to the third drug category in the in the fully adjusted models. The confidence intervals of the two higher risk groups did not overlap with the confidence intervals of the lowest risk group indicating significant stratification of comorbid COPD effects by drugs prescribed (Figure 10.11).

10.4.2.3 COPD severity change

- **Drug severity change measure**

The HF-COPD group were stratified into three groups based on change in their prescribed COPD drugs over 6 months before the match date. This measure used the COPD drugs prescribed in a 4 month time-window before first hospital admission compared to a 4 month time-window, six months before admission. Change in drug category over 6 months was categorised into (i) no new steroids or oxygen, (ii) new on steroids but no new oxygen and (iii) new on oxygen. These groups were compared to the HF sample without COPD.

Strength of associations: The risks associated with the first two drug change categories were similar (OR 1.31; 1.22-1.41) and OR 1.35 (1.11-1.64) respectively) (Figure 10.12), but the risk estimate for the 'new on prescribed oxygen' group was strongest but non-significant (OR 2.02; 0.93, 4.39).

10.4.3 Chronic kidney disease

The unadjusted association between HF-CKD and first hospital admission compared to no CKD HF was OR 1.34 (95% CI 1.29-1.40), which diminished considerably when adjusted for age, gender, diuretics, diabetes, prior admission, BMI, smoking, and Hb to 1.15 (1.10-1.21). This estimate was not influenced by further adjustment for the remaining confounders (Table 10.8). This association was weaker than HF-CKD and mortality (OR 1.22; 1.18, 1.26).

10.4.3.1 CKD severity

- **Physiological severity measure**

The HF sample was stratified into groups based on the eGFR level. This categorisation used the closest available measure preceding the match date (47 [IQR 18-94 days]) and these groups were compared to the baseline eGFR group 60-89 ml/min/m^2.

Shape of association: The shape of the adjusted associations between eGFR categories and their respective predicted probabilities of first hospital admission was curved, with the risk of admission falling steeply from the lowest eGFR group (<15 ml/min/m^2). This is shown in Figure 10.13 using a margins plot. The lowest risk group was eGFR 60-89. Whilst the risk of admission begins to increase in the higher eGFR groups, this curve is only slight and did not show the same U shape as for the mortality outcome. When eGFR was categorised into deciles the curve was similar to the eGFR categories but with a steeper increase in risk from the 9^{th} to 10^{th} decile (Figure 10.14). The lowest risk deciles were 7-9 corresponding to eGFR of 63.3 to 84.9 ml/min/m^2 .

Strength of associations: Unadjusted associations between categories of eGFR and hospital admission compared to the reference group of eGFR 60-89 are shown in Table 10.8. Risk was associated with all categories above and below the reference group. The lowest risk category was eGFR 90-105 (OR 1.15; 1.05-1.26) and this OR increased only slightly to the next and highest group (eGFR >105). Risk also increased with each category below eGFR 90-105 to a maximum of OR 9.17 (6.78-12.4) in the lowest eGFR group (<15). The same trend was observed in the fully adjusted associations (Table 10.8). Risk was lowest compared to the reference group in the eGFR 90-105 group (1.09; 0.99-1.21) which was non-significant. Risk of admission increased from the 45-59 eGFR group (1.12; 1.05-1.17) to the lowest <15 eGFR group (4.45; 3.19-6.21).

In the fully adjusted model, the linear test for trend was non-significant (Figure 10.15). The risk associations of the two lowest eGFR groups had confidence intervals that did not overlap with the confidence intervals of the other eGFR categories, which indicated stratified effects of CKD defined by eGFR severity levels on hospital admission outcome.

10.4.3.2 CKD severity change

- **Physiological severity change measure**

The HF sample was stratified into three groups based on change in their eGFR level over 6 months before the match date. This measure used the closest available eGFR measure before the match date and a previous eGFR before the most recent measure (median 110 [IQR 59-199] days). Absolute change was defined as: (a) 0-5mls decrease – reference group, minor decline, (b) any increase, (c) 6-15mls decrease – moderate

decline, and (d) >15mls decrease – severe decline. Percentage change was defined as: (a) 0-5% decrease – reference group, minor decline, (b) any increase, (c) 6-25% decrease – moderate decline, and (d) >25% decrease – severe decline. Change over 6 months was calculated.

Strengths of associations: All categories of absolute and percentage change in eGFR were associated with first hospital admission when compared to the reference group of minimal decline in the adjusted models. HF patients with any increase or moderate decline in eGFR had similar associations compared to the minimal decline reference groups in the adjusted models (Table 10.8). These associations were non-significant using the absolute change measure. The highest risk was observed in the most severe change groups in both classifications; >15mls group (OR 1.37; 1.26-1.49) and >25% decrease (OR 1.49; 1.36-1.63). The confidence intervals of these groups did not overlap with the confidence intervals of the other categories of change, indicating the stratified effects of comorbid CKD by CKD change in HF on hospital admission outcome (Figure 10.16).

Adjustment for eGFR: Firstly, the influence of the start or end eGFR on the risk estimates of the different change measures was investigated by observing the associations within sub-groups according to the starting eGFR (<60 and ≥60 ml/min/m^2). Starting eGFR was the first and earliest measure in the change definition. The change estimates were also observed within the same strata of the end eGFR measure. This was to investigate whether the effect estimates of eGFR change were: (i) influenced by the starting eGFR and whether this influence was less for the percentage change measure that already takes some account of the starting point and (ii) depended on the end eGFR point meaning that the influence of decline might be greater in the <60mls sub-group. Secondly, the estimates for the change measures were further adjusted, for the starting and separately for the end eGFR in order to investigate the independent association of change in eGFR (Table 10.8).

eGFR change stratified by baseline start and end eGFR: The effect estimates for eGFR increase and decline were similar in those with a starting eGFR of <60ml and ≥60 ml/min/m^2 for both the absolute and percentage change measure (Table 10.9). Using the end eGFR measure, the effect estimates for all change measures

were similar between the <60ml and ≥60 ml/min/m^2 groups for the absolute measure. For the percentage change measure, decline in eGFR had a greater risk estimate in the <60 group than the ≥60 group

eGFR change adjusted by baseline start and end eGFR: Adjustment of the absolute and percentage eGFR change measures with the start eGFR made little difference to the effect estimates for hospital admission (Table 10.10). Adjustment of the same measure with the end eGFR made little difference to the absolute change measure but reduced the most severe percentage change measure marginally..

10.4 Chapter summary

The three comorbidities DM, COPD and CKD were significantly and independently associated with first hospital admission in the non-selected general practice population of HF. The strengths of association were similar to the estimates for mortality and were not explained by confounders based on routinely collected clinical data. When each of the comorbidities were stratified by categories of recent severity, there were significant associations with first hospital admission. The associations between categories of HbA1c or eGFR and hospital admission were less curved than when observed for mortality and there was a trend of increasing risk with higher HbA1c and lower eGFR level. Comorbidity severity measures using prescribed drugs also provided significant stratified comorbidity effects for both DM and COPD. The associations between comorbidity severity change and hospital admission were varied. There were significant associations between HbA1c change and increased hospital admission and between the drug severity change measures for DM and COPD and increased hospital admission, but the categories of change for each of the comorbidities were not significantly different to each other. Accelerated or severe eGFR decline over 6 months was associated with increased risk of hospital admission compared to minimal decline. Adjustment for baseline eGFR measures made little difference to these associations. These findings will be discussed in detail in Chapter 13.

Tables

Table 10.1 Summary of confounders indicated in the four steps

Potential confounder	Diabetes	COPD	CKD (eGFR <60)
Age	SR *#	SR *#	SR *#
Gender	SR *	*	SR *#
White ethnicity	SR		SR
Deprivation	*#	*#	*#
HF factors			
Left ventricular ejection fraction	SR	SR	SR
Ischaemic aetiology	SR		SR
Other aetiology			SR
Left ventricular systolic dysfunction	SR		
Brain Naturetic Peptide			SR
New York Heart Association class	SR		SR
Cardiovascular Drugs			
Beta-blocker	SR *	SR *#	SR
ACEi	SR *	#	SR *
ARB	SR *	#ł	SR *
Statin	SR		SR
Diuretics	SR *#	SR *	SR *#
Digoxin	SR		
Spirolactone	SR	SR	SR
Comorbidities			
Renal disease	SR *ł	*ł	-
Diabetes	-	SR *#	SR *#
COPD	SR *	-	*
Vascular disease			SR
Cerebrovascular accident	SR		SR
Hypertension	SR		SR
Charlson index	SR		
Atrial fibrillation	SR		SR
Anthropometric and Clinical observations			
Prior admission	*#ł	*#ł	SR *#ł
Systolic blood pressure	SR *	*	SR*
Diastolic blood pressure	*	*	*
Heart rate	SR	SR	
Edema	SR		
Dyspnoea	SR		
Elevated jugular venous pressure	SR		
Sinus rhythm		SR	SR
Body mass index	SR *#	SR *	SR *#
Cholesterol	*	SR	SR*
Glucose level	SR		SR
Blood Urea Nitrogen (mg/dL)	SR		SR
Potassium			SR
Haemoglobin	*#	*ł	SR *#ł
Creatinine	SR	SR	SR
Estimated glomerular filtration rate	*ł	*ł	SR *
Lifestyle factors			
Smoking	*	SR *	SR *#
Alcohol status	*	*	*

Red text; available factors extracted from CPRD. SR; Systematic review evidence (step 1),* difference identified between comorbid and non-comorbid group (step 2), # ≥10% change in comorbidity effect observed in at least one strata of the potential confounder (step 3) łassociation identified by ≥10% change in OR of comorbid exposure by adjustment of potential confounder (step 4).

Table 10.2 Unadjusted associations between confounders and first hospital admission

Factors	OR (95% CI)
Person and socio-demographic factors	
Age per year	1.01 (1.01, 1.01)
Women	0.91 (0.89-0.94)
IMD quintile*	
1	1.0
2	1.02 (0.97-1.06)
3	1.10 (1.05-1.15)
4	1.10 (1.05-1.15)
5	1.19 (1.14-1.25)
Anthropometric and clinical factors	
Prior Hospital admission	
3 months	7.84 (7.43-8.29)
6 months	2.39 (2.26-2.53)
1 year	1.52 (1.45-1.59)
BMI (per Kg/m2)	0.98 (0.98-0.99)
Cholesterol (per mmol/L)	0.95 (0.93-0.96)
Haemoglobin (per g/dL)	0.84 (0.84-0.86)
Systolic BP (per mmHg)	0.99 (0.99-1.00)
Diastolic BP (per mmHg)	0.99 (0.99-0.99)
Lifestyle factors	
Smoking status	
No	1.0
Ex	1.06 (1.03-1.10)
Yes	1.28 (1.20,1.32)
Alcohol status	
No	1.0
Ex	1.06 (0.91-1.03)
Yes	0.90 (0.87-0.93)
Drug factors	
Not on beta-blocker	1.21 (1.17-1.24)
Not on ACEi	1.36 (1.33-1.41)
Not on ARB	1.20 (1.15-1.25)
Diuretic	0.90 (0.87-0.93)
Comorbidity exposures	
Diabetes	1.33 (1.28-1.37)
COPD	1.36 (1.31-1.43)
eGFR <60ml/min)	1.34 (1.29-1.40)
eGFR (per ml/min/m2)	0.99 (0.99-0.99)

First hospital admission is the first all-cause admission following the incident HF date. *IMD in 79,383 patients only. IMD, index multiple deprivation (1=least deprived, 5=most deprived); BMI, body mass index; BP, blood pressure; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate

Table 10.3 Steps of adjustment for each comorbidity with first hospital admission

Comorbidity	Step 1	+ Step 2	+ Step 3
DM	Age, diuretics, prior admission, BMI, Hb, eGFR, deprivation	Gender, beta-blocker, ACEi, ARB, COPD, systolic BP	Cholesterol, smoking, alcohol
COPD	Age, beta-blocker, ACEi, ARB, diabetes, prior admission, Hb, eGFR, deprivation	Diuretics, BMI, smoking	Gender, systolic BP, alcohol, cholesterol
CKD	Age, gender, diuretics, diabetes, prior admission, BMI, smoking, Hb, deprivation	ACEi, ARB, systolic, cholesterol, eGFR	Beta-blocker, COPD, alcohol

BMI, body mass index; BP, blood pressure; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate

Table 10.4 Likelihood ratio tests for current BMI

Multivariable	Log likelihood	LR test
a BMI	-17245.36	
b BMI+BMI^2	-17236.88	<0.001 a nested in b
c BMI+BMI^2+BMI^3	-17236.79	0.67 b nested in c

Table 10.5 Likelihood ratio tests for current Hb

Multivariable	Log likelihood	LR test
a HB	-17245.36	
b HB+HB^2	-17123.04	<0.001 a nested in b
c HB+HB^2+HB^3	-17122.13	0.179 b nested in c

Table 10.6 Association between comorbid diabetes exposures and first hospital admission in HF

Diabetes exposures	Unadjusted OR (95% CI)	Adjusted[1] OR (95% CI)	Adjusted[2] OR (95% CI)	Adjusted[3] OR (95% CI)
Diabetes status *53,159 observations				
No diabetes (ref)	1.0	1.0	1.0	1.0
Diabetes anytime	1.30 (1.25-1.36)	1.23 (1.17-1.30)	1.25 (1.19-1.32)	1.24 (1.18-1.31)
No diabetes (ref)	1	1	1	1
Diagnosis before HF	1.34 (1.28-1.40)	1.26 (1.19-1.32)	1.28 (1.22-1.35)	1.27 (1.20-1.34)
Diagnosis after HF	0.90 (0.76-1.06)	0.97 (0.81-1.15)	0.94 (0.79-1.12)	0.93 (0.78-1.10)
Diabetes HbA1c severity *48,427 observations				
No diabetes (ref)	1.0	1.0	1.0	1.0
<5.5%	1.89 (1.48-2.43)	1.20 (0.90-1.59)	1.19 (0.90-1.59)	1.18 (0.89-1.57)
5.5-6.4%	1.18 (1.07-1.29)	1.10 (0.99-1.22)	1.12 (1.01-1.25)	1.10 (0.99-1.22)
6.5-7.5%	1.18 (1.09-1.27)	1.15 (1.05-1.25)	1.17 (1.07-1.27)	1.15 (1.06-1.26)
7.6-8.5%	1.21 (1.10-1.35)	1.20 (1.08-1.35)	1.23 (1.10-1.38)	1.23 (1.09-1.38)
8.6-9.5%	1.27 (1.09-1.47)	1.31 (1.12-1.54)	1.30 (1.10-1.53)	1.28 (1.08-1.50)
>9.5%	1.67 (1.44-1.92)	1.62 (1.39-1.90)	1.64 (1.40-1.93)	1.64 (1.39-1.92)
Diabetes drugs severity *53,159 observations				
No diabetes (ref)	1.0	1.0	1.0	1.0
1: None	1.07 (0.98-1.16)	1.00 (0.92-1.10)	1.00 (0.92-1.10)	1.00 (0.91-1.09)
2: Oral (+/- Insulin)	1.36 (1.28-1.43)	1.32 (1.24-1.40)	1.34 (1.26-1.43)	1.33 (1.25-1.42)
3: Insulin only	1.60 (1.44-1.77)	1.37 (1.21-1.54)	1.42 (1.25-1.60)	1.41 (1.25-1.59)
Diabetes HbA1c severity change *45,678 observations				
No diabetes (ref)	1.0	1.0	1.0	1.0
<1% change	1.18 (1.11-1.26)	1.16 (1.08-1.25)	1.18 (1.10-1.27)	1.17 (1.08-1.26)
>1% increase	1.33 (1.17-1.52)	1.24 (1.07-1.44)	1.27 (1.09-1.48)	1.26 (1.09-1.47)
>1% decrease	1.46 (1.30-1.64)	1.36 (1.19-1.56)	1.37 (1.20-1.56)	1.34 (1.18-1.53)
Diabetes drugs severity change *53,159 observations				
No diabetes (ref)	1.0	1.0	1.0	1.0
No drug category change	1.27 (1.22-1.33)	1.22 (1.16-1.28)	1.24 (1.17-1.30)	1.22 (1.16-1.29)
Increase in drug category	1.52 (1.31-1.76)	1.38 (1.17-1.63)	1.42 (1.20-1.69)	1.40 (1.19-1.66)
Decrease in drug category	2.14 (1.70-2.89)	1.58 (1.17-2.13)	1.62 (1.19-2.19)	1.60 (1.18-2.17)

First hospital admission is the first all-cause admission following the incident HF date. Ref, reference group. * each unadjusted and adjusted measure was based on the same number of observations. There was complete data for comorbidity status and all confounders with the exception of eGFR. All associations excluded those without eGFR. For the comorbidity

severity exposures, associations excluded the comorbid groups without the severity exposure or that had eGFR missing.
[1]Adjusted for age, diuretics, prior admission, BMI, Hb Hb[2], eGFR eGFR[2]
[2]Adjusted further for gender, beta-blocker, ACEi or ARB, COPD, systolic bp, systolic bp[2]
[3] Adjusted further for cholesterol, smoking, deprivation, alcohol

Table 10.7 Association between comorbid COPD exposures in HF and first hospital admission

COPD exposures	Unadjusted OR (95% CI)	Adjusted[1] OR (95% CI)	Adjusted[2] OR (95% CI)	Adjusted[3] OR (95% CI)
COPD defined by Diagnostic code *53,159 observations				
No COPD (ref)	1.0	1.0	1.0	1.0
COPD anytime	1.36 (1.29-1.45)	1.36 (1.28-1.46)	1.34 (1.26-1.44)	1.32 (1.22-1.41)
No COPD (ref)	1.0	1.0	1.0	1.0
COPD diagnosis before HF	1.36 (1.28-1.44)	1.36 (1.27-1.45)	1.33 (1.24-1.43)	1.31 (1.22-1.41)
COPD diagnosis after HF	1.44 (1.18-1.76)	1.44 (1.17-1.78)	1.44 (1.17-1.78)	1.40 (1.13-1.73)
COPD defined by severity (drugs) *53,159 observations				
No COPD (ref)	1.0	1.0	1.0	1.0
No steroids or oxygen	1.20 (1.12-1.28)	1.22 (1.13-1.31)	1.20 (1.11-1.29)	1.17 (1.09-1.27)
Oral steroids but no oxygen	1.81 (1.62-2.01)	1.77 (1.57-2.00)	1.76 (1.55-1.98)	1.73 (1.53-1.96)
On oxygen	3.01 (2.16-4.18)	2.59 (1.78-3.77)	2.50 (1.71-3.64)	2.42 (1.66-3.52)
COPD defined by severity change *53,159 observations				
No COPD (ref)	1.0	1.0	1.0	1.0
No new steroids or oxygen	1.36 (1.28-1.44)	1.36 (1.27-1.45)	1.33 (1.24-1.43)	1.31 (1.22-1.41)
New on steroids but no new oxygen	1.39 (1.17-1.66)	1.39 (1.15-1.69)	1.38 (1.13-1.67)	1.35 (1.11-1.64)
New on oxygen	2.46 (1.18-5.12)	2.36 (1.09-5.11)	2.34 (1.08-5.10)	2.02 (0.93-4.39)

First hospital admission is the first all-cause admission following the incident HF date. Ref, reference group. * each unadjusted and adjusted measure was based on the same number of observations. There was complete data for comorbidity status and all confounders with the exception of eGFR. All associations excluded those without eGFR. For the comorbidity severity exposures, associations excluded the comorbid groups without the severity exposure or that had eGFR missing.

[1]Adjusted for Age, beta-blocker, ACEi or ARB, diabetes, prior admission, Hb Hb[2], eGFR eGFR[2]

[2]Adjusted further for diuretic, BMI, smoking

[3] Adjusted further for gender, systolic bp, systolic bp2, deprivation, alcohol, cholesterol

Table 10.8 Associations between CKD exposures in HF and first hospital admission

CKD exposures	Unadjusted OR (95% CI)	Adjusted[1] OR (95% CI)	Adjusted[2] OR (95% CI)	Adjusted[3] OR (95% CI)
CKD status				
eGFR ≥60 (ref)	1.0	1.0	1.0	1.0
eGFR <60	1.34 (1.29, 1.40)	1.15 (1.10-1.21)	1.14 (1.10-1.20)	1.15 (1.10-1.21)
CKD severity				
eGFR mL/min/1.73m²				
60-89 (ref)	1.0	1.0	1.0	1.0
>105	1.19 (1.03-1.38)	1.22 (1.04-1.43)	1.19 (1.02-1.40)	1.17 (0.99-1.37)
90-105	1.15 (1.05-1.26)	1.11 (1.01-1.23)	1.10 (1.00-1.21)	1.09 (0.99-1.21)
45-59	1.16 (1.12-1.22)	1.10 (1.05-1.16)	1.10 (1.04-1.16)	1.12 (1.05-1.17)
30-44	1.41 (1.33-1.49)	1.15 (1.08-1.23)	1.14 (1.07-1.21)	1.15 (1.08-1.22)
15-29	2.50 (2.30-2.73)	1.66 (1.51-1.83)	1.59 (1.44-1.75)	1.61 (1.46-1.77)
<15	9.17 (6.78-12.4)	4.81 (3.46-6.68)	4.43 (3.17-6.17)	4.45 (3.19-6.21)
CKD severity change				
Absolute change				
0mls to 5mls decrease (ref)	1.0	1.0	1.0	1.0
Any increase	1.07 (1.01-1.14)	1.05 (0.98-1.13)	1.05 (0.98-1.12)	1.04 (0.98-1.12)
6mls to 15mls decrease	1.12 (1.04-1.20)	1.07 (0.99-1.16)	1.06 (0.98-1.15)	1.06 (0.97-1.14)
>15mls decrease	1.49 (1.38-1.60)	1.38 (1.27-1.50)	1.38 (1.27-1.50)	1.37 (1.26-1.49)
Percentage change				
0-5% decrease (ref)	1.0	1.0	1.0	1.0
Any % increase	1.21 (1.13-1.30)	1.13 (1.05-1.23)	1.12 (1.03-1.21)	1.11 (1.03-1.21)
6-25% decrease	1.25 (1.15-1.34)	1.17 (1.07-1.27)	1.15 (1.05-1.25)	1.14 (1.05-1.25)
>25% decrease	1.79 (1.65-1.95)	1.51 (1.38-1.66)	1.49 (1.36-1.64)	1.49 (1.36-1.63)

First hospital admission is the first all-cause admission following the incident HF date. Ref, reference group. *each unadjusted and adjusted measure based on the same number of observations (excluding those with the comorbidity exposure or eGFR confounder missing)

[1]Adjusted for age, gender, diuretic, diabetes, prior admission, BMI, smoking, Hb Hb[2]

[2]Adjusted further for ACEi or ARB, systolic bp, systolic bp2, cholesterol

[3]Adjusted further for Beta-blocker, COPD, deprivation, alcohol

Table 10.9 Change estimates with first hospital admission stratified by start baseline and end eGFR

eGFR severity change	Start eGFR		End eGFR	
	≥60	<60	≥60	<60
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Absolute change				
0mls to 5mls decrease (ref)	1.0	1.0	1.0	1.0
Any increase	1.02 (0.89-1.18)	1.02 (0.91-1.14)	1.08 (0.94-1.24)	1.01 (0.90-1.13)
6mls to 15mls decrease	1.01 (0.86-1.18)	1.11 (0.97-1.27)	0.97 (0.81-1.16)	1.05 (0.93-1.19)
>15mls decrease	1.45 (1.24-1.69)	1.32 (1.13-1.53)	1.30 (1.08-1.56)	1.35 (1.19-1.54)
Percentage change				
0-5% decrease (ref)	1.0	1.0	1.0	1.0
Any % increase	1.07 (0.92-1.24)	1.18 (1.02-1.36)	1.10 (0.95-1.28)	1.16 (1.01-1.33)
6-25% decrease	1.13 (0.96-1.32)	1.31 (1.12-1.53)	1.07 (0.90-1.28)	1.25 (1.09-1.44)
>25% decrease	1.55 (1.31-1.84)	1.53 (1.30-1.81)	1.33 (1.08-1.65)	1.55 (1.33-1.80)

All associations were fully adjusted by all remaining covariates. eGFR-estimated glomerular filtration rate (ml/min/m^2); ref, reference group. Change was calculated over 6-months before the match date using the most recent value up to a maximum of 6-months and a prior value between 1 month and 1 year.

Table 10.10 eGFR change effect estimates with first hospital admission adjusted for the start and end eGFR

eGFR severity change	Adjusted[3]	Adjusted for start eGFR and eGFR^2	Adjusted for end eGFR and eGFR^2
Absolute change			
0mls to 5mls decrease (ref)	1	1	1
Any increase	1.04 (0.98-1.12)	1.04 (0.97-1.11)	1.07 (1.00-1.14)
6mls to 15mls decrease	1.06 (0.97-1.14)	1.07 (0.99-1.16)	1.05 (0.97-1.14)
>15mls decrease	1.37 (1.26-1.49)	1.40 (1.29-1.53)	1.36 (1.25-1.47)
Percentage change			
0-5% decrease (ref)	1	1	1
Any % increase	1.11 (1.03-1.21)	1.10 (1.01-1.19)	1.11 (1.02-1.20)
6-25% decrease	1.14 (1.05-1.25)	1.14 (1.05-1.24)	1.13 (1.03-1.23)
>25% decrease	1.49 (1.36-1.63)	1.47 (1.34-1.61)	1.39 (1.27-1.53)

eGFR, estimated glomerular filtration rate (ml/min/m^2); ref, reference group. All associations were fully adjusted by all remaining covariates and the start or end eGFR. Due to the non-linear association between eGFR and hospital admission eGFR^2 was also used in the models. Change was calculated over 6-months before the match date using the most recent value up to a maximum of 6-months and an earlier value between 1 month and 1 year before.

Figures

Figure 10.1 Calibration plots of observed versus predicted fitted effects for BMI

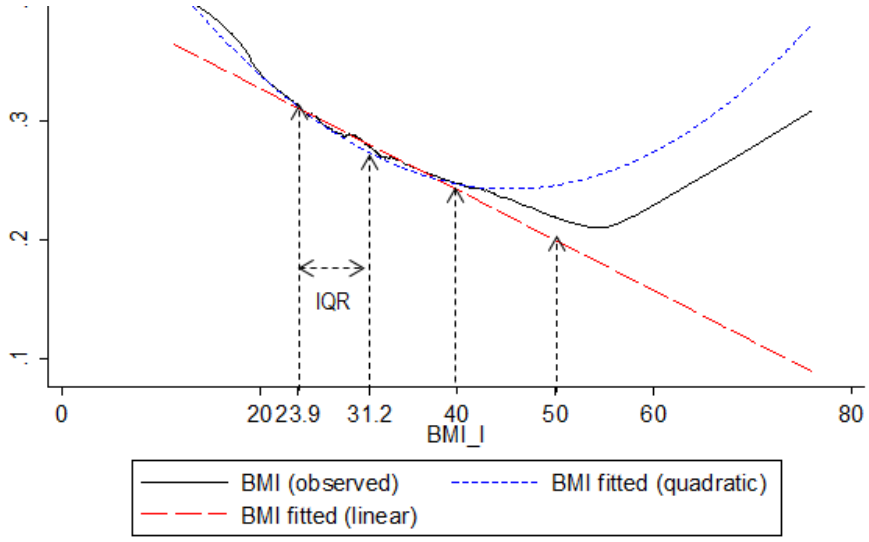


Figure 10.2 Eccles plots of BMI

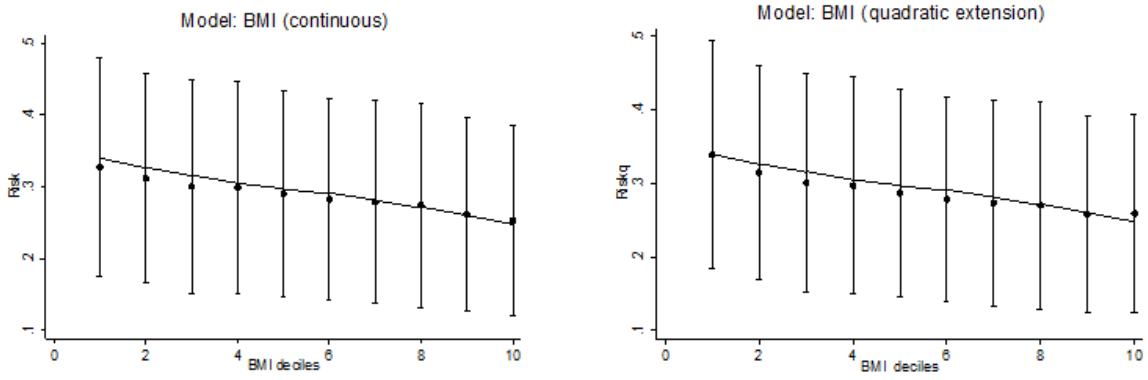


Figure 10.3 Calibration plots of observed versus predicted fitted effects for Haemoglobin

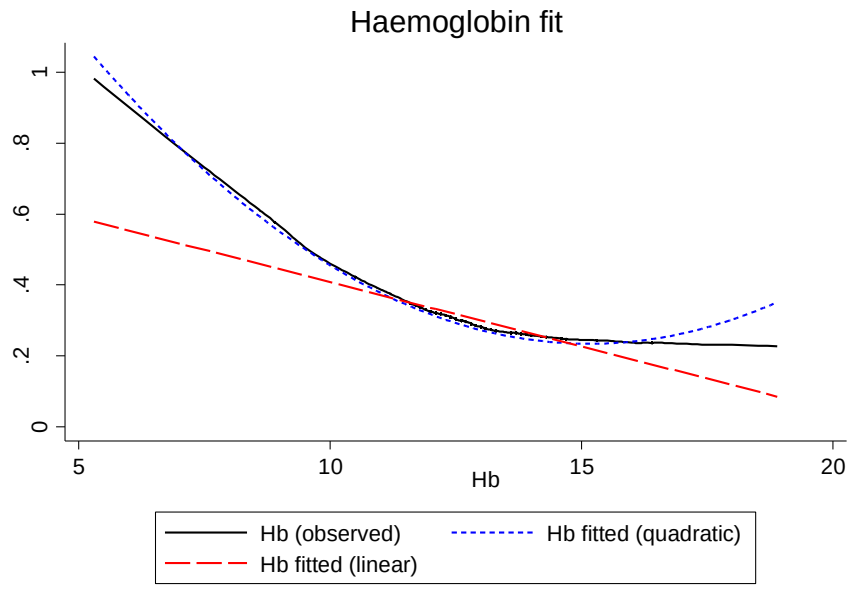


Figure 10.4 Eccles plots of Haemoglobin

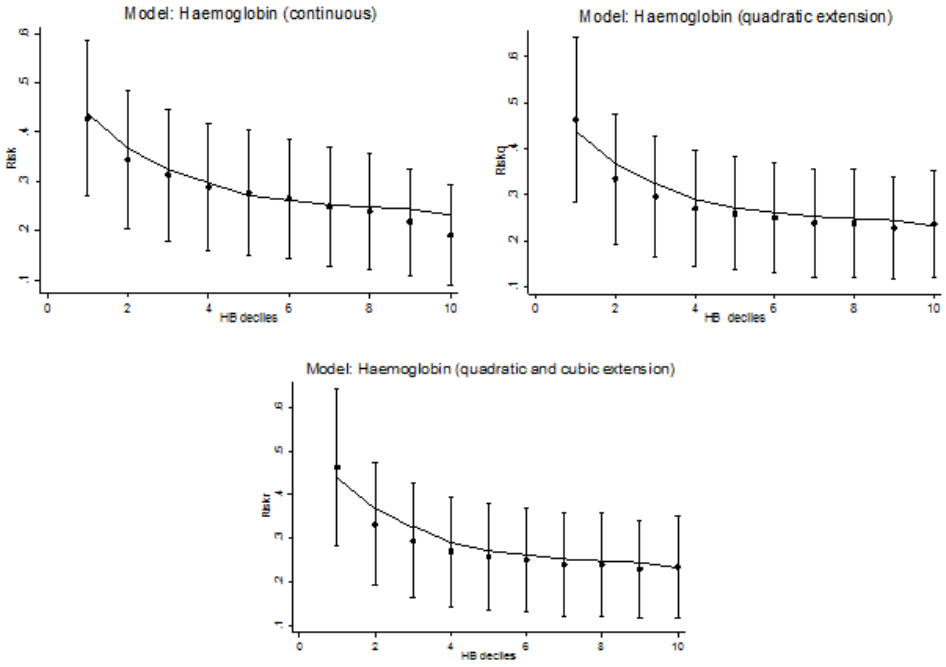


Figure 10.5 Adjusted associations of comorbidities in HF with first hospital admission

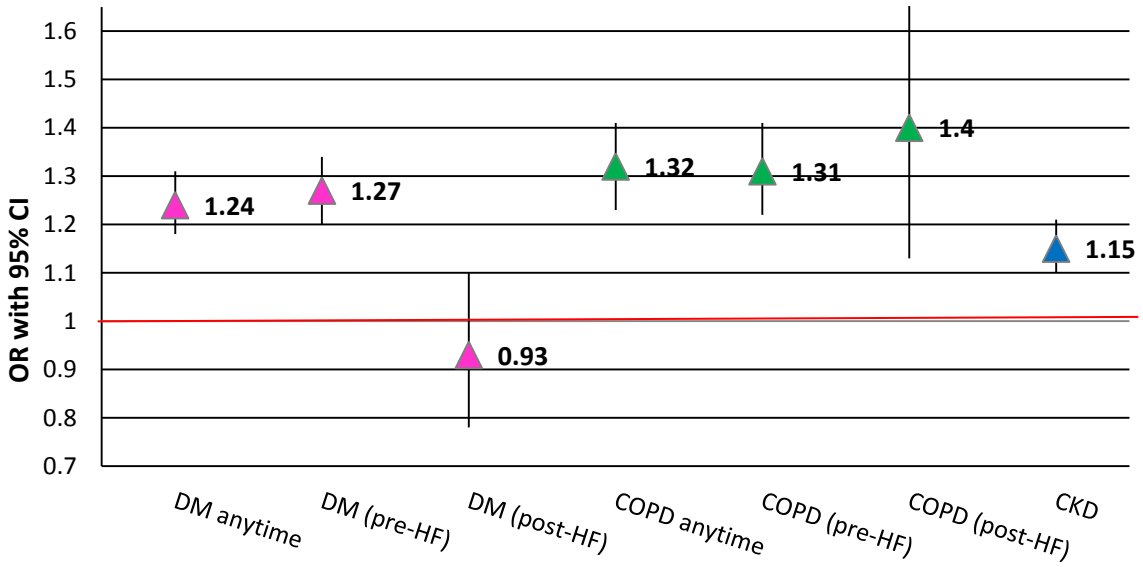


Figure 10.6 Margins plot of HbA1c categories and hospital admission

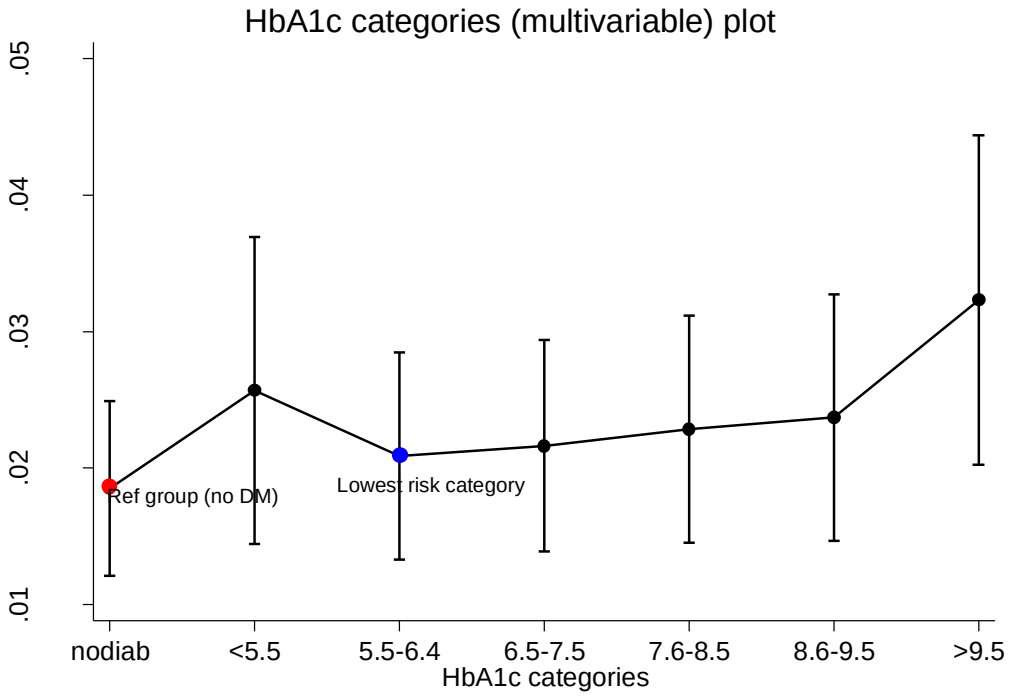


Figure 10.7 Margins plot of HbA1c deciles and first hospital admission

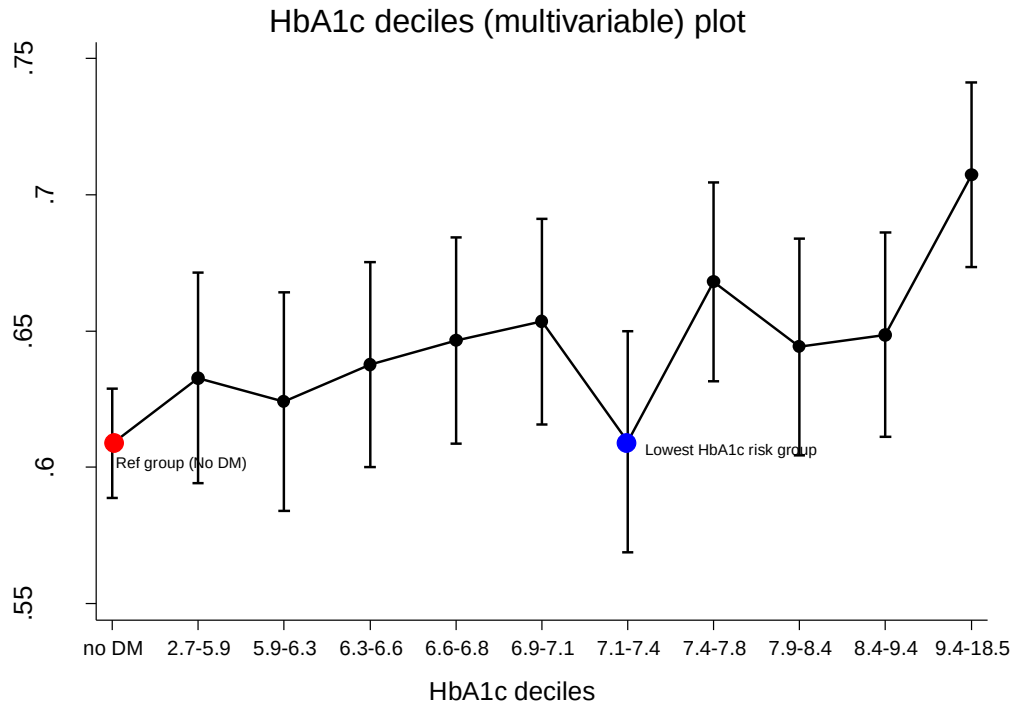


Figure 10.8 Adjusted associations of HbA1c categories and first hospital admission

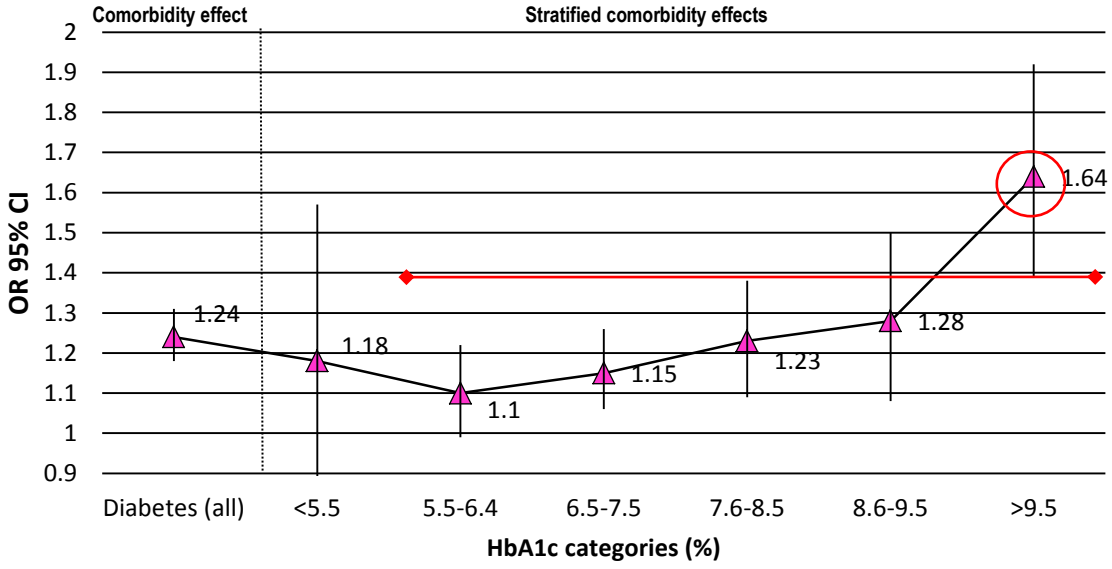


Figure 10.9 Adjusted associations of DM drug severity categories and first hospital admission

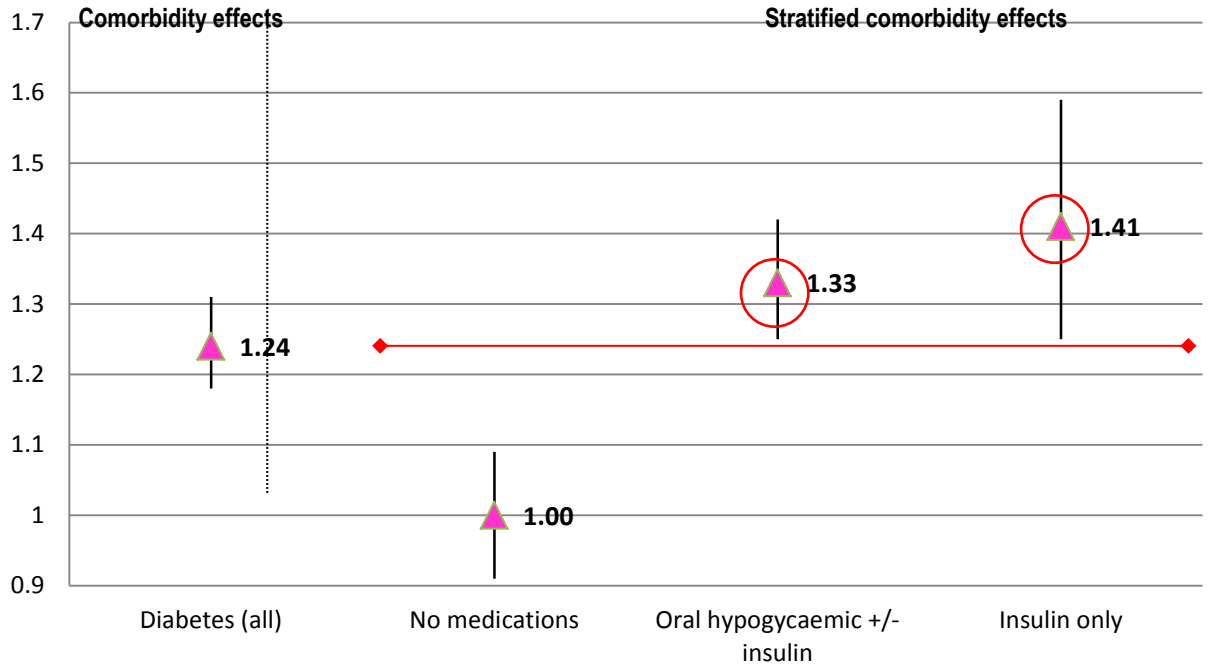


Figure 10.10 Adjusted associations of categories of DM and HbA1c change and hospital admission

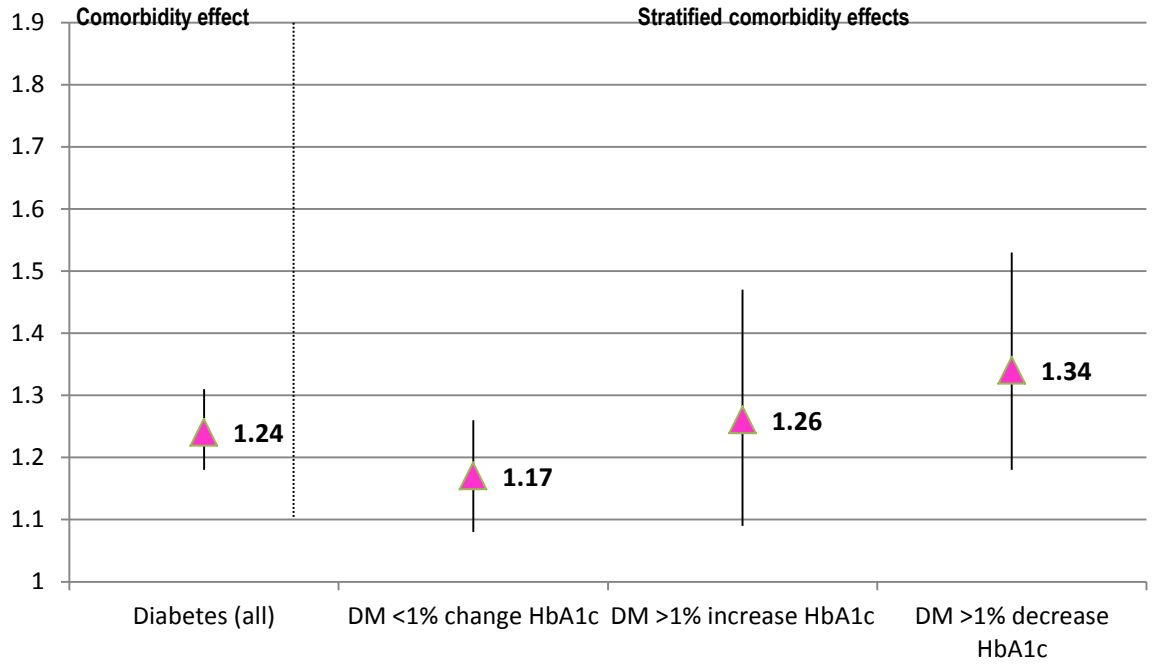


Figure 10.11 Adjusted associations of COPD drug severity categories and hospital admission

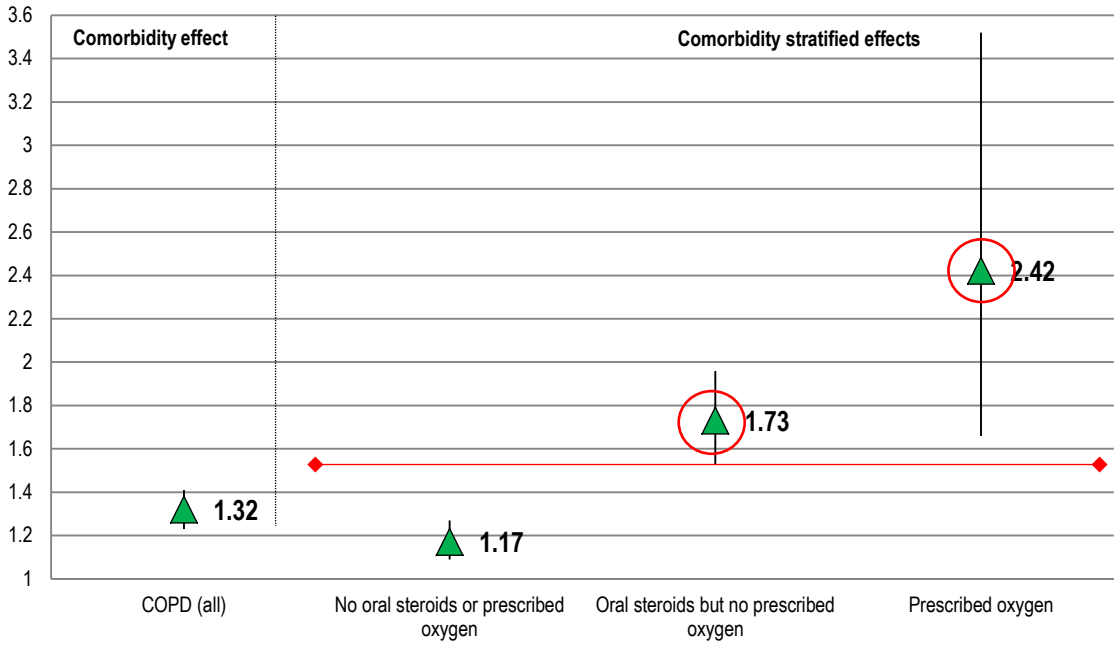


Figure 10.12 Adjusted associations of COPD drug severity change and hospital admission

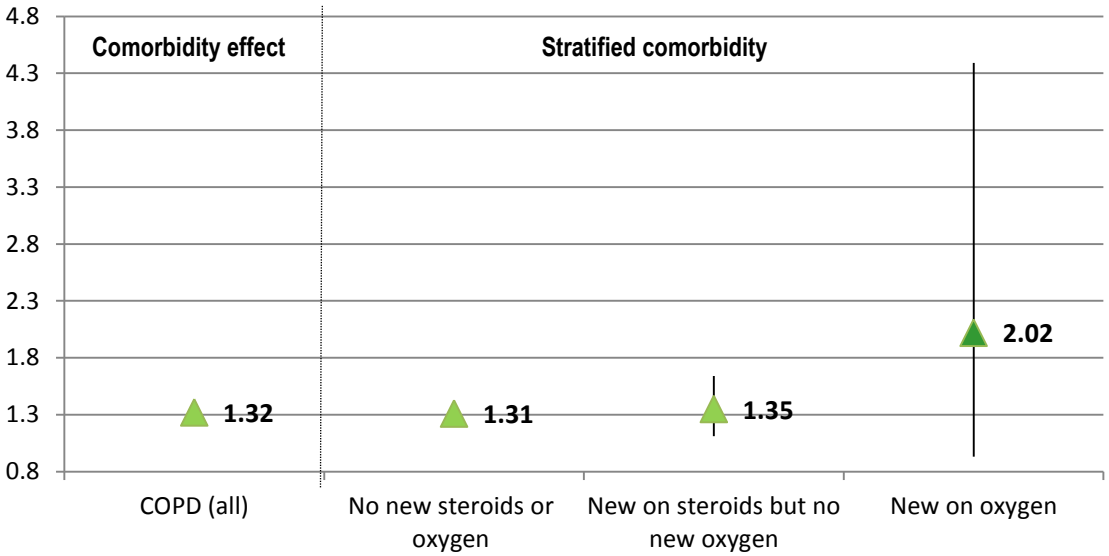


Figure 10.13 Margins plot of eGFR severity categories and first hospital admission

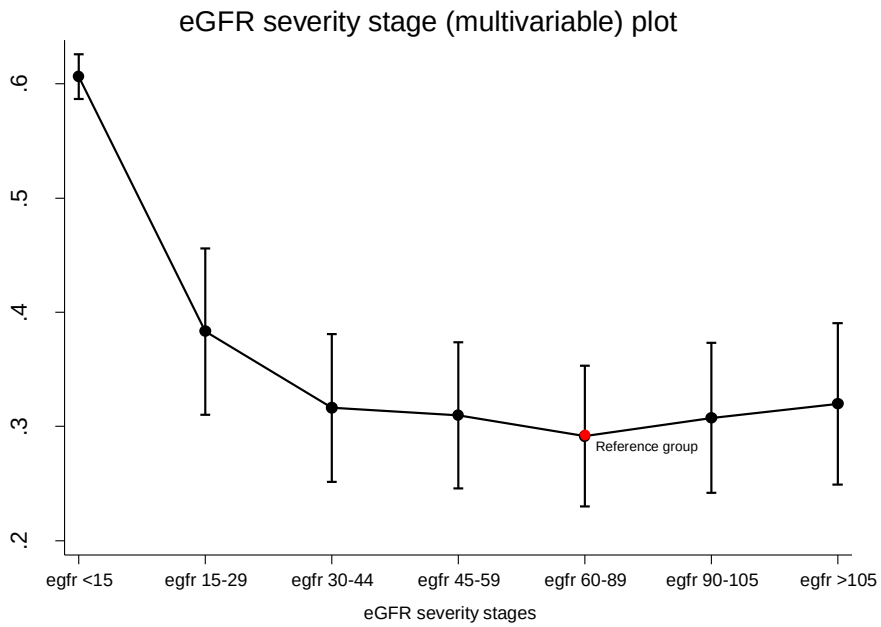


Figure 10.14 Margins plot of eGFR deciles and first hospital admission

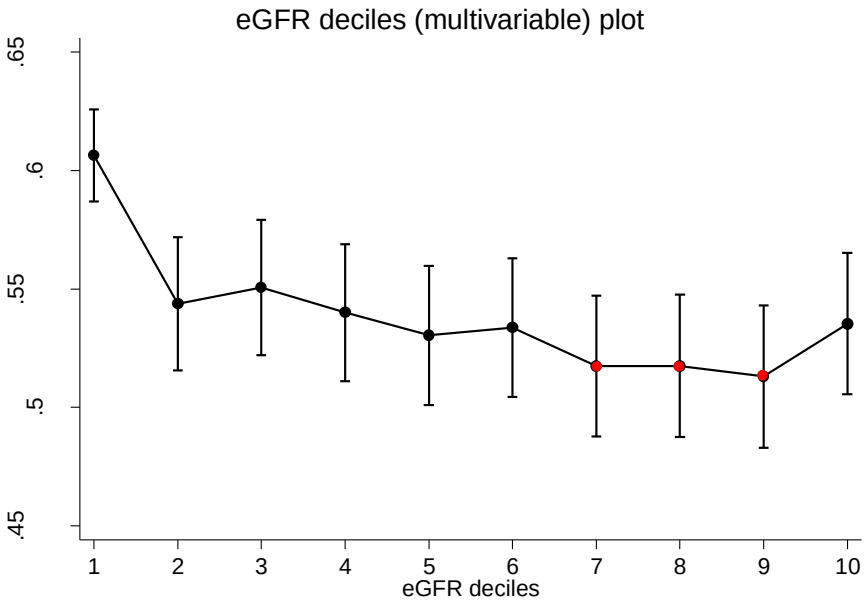


Figure 10.15 Adjusted associations of eGFR categories and first hospital admission

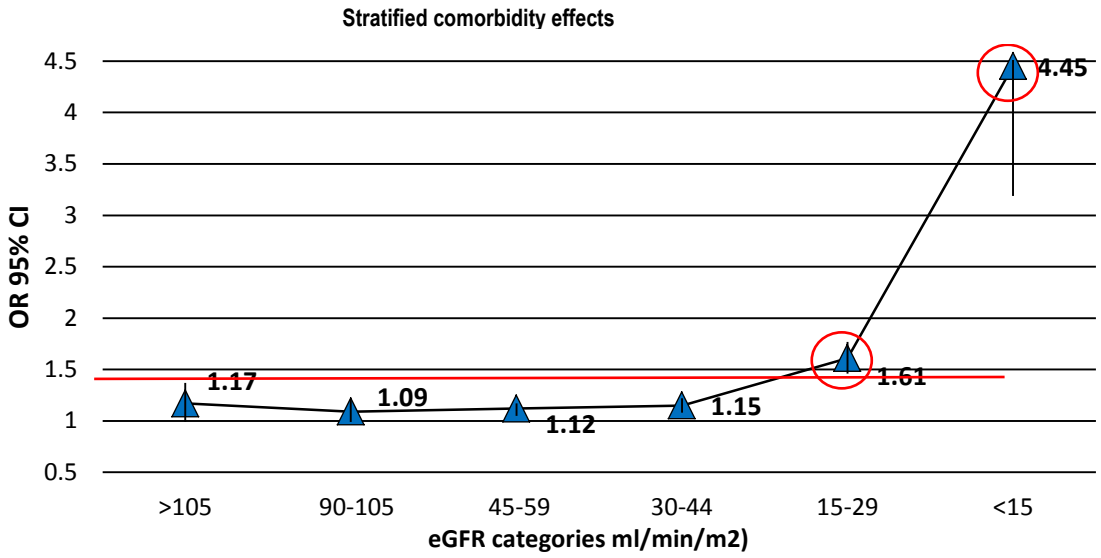
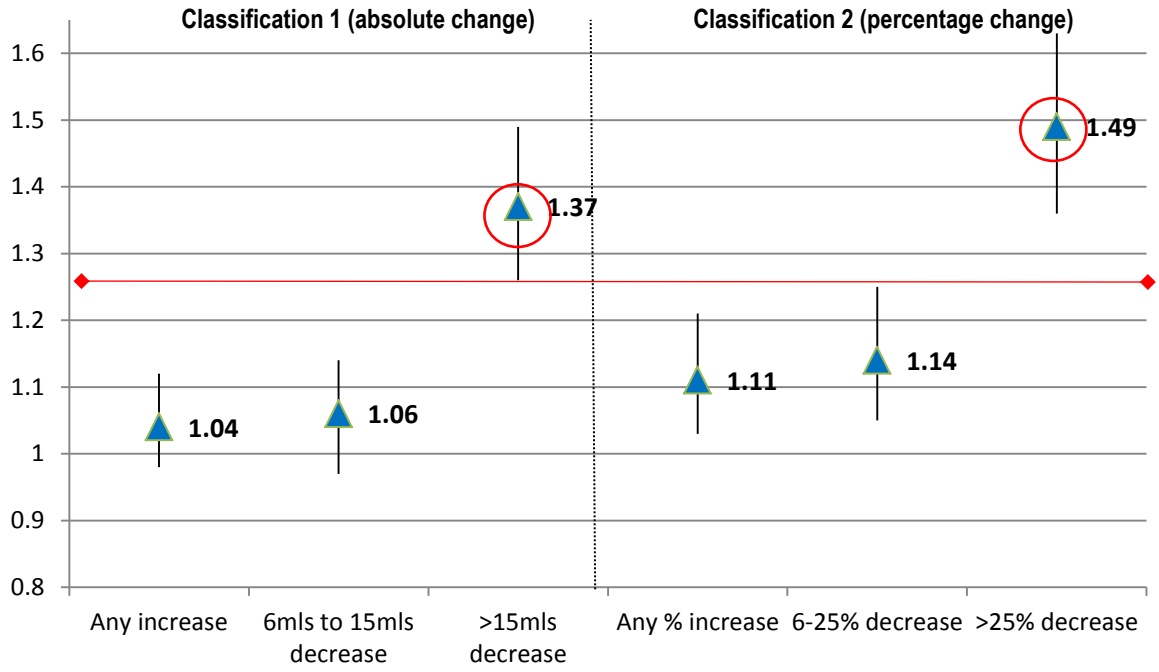


Figure 10.16 Adjusted associations of categories of eGFR change and first hospital admission



Chapter 11 Comorbidity interactions

This chapter investigates first order interactions for the DM, COPD and CKD comorbidities in HF for both mortality and first hospital admission. Interactions were investigated between (i) pairs of the three comorbidities, (ii) each comorbidity and HF severity as defined by age groups and (iii) each comorbidity and any additional patient factors. The patient factors included those that were identified as potential effect modifiers during the investigation of confounding in Chapters 9 and 10. First, the comorbidity effect estimates that differed across strata during the confounding investigations are summarised. Second, likelihood ratio tests are used to compare fully adjusted models with and without interaction terms to identify significant statistical interaction, which was defined as departure from multiplicatively. Third, biological interaction was investigated to investigate departure from additivity.

11.1 Summary of potential interactions

11.1.1 Comorbidity pairs

When the estimate of effect of a comorbidity on both outcomes was observed in groups defined by the presence or absence of an additional comorbidity (i.e. pairs), there was only one disease pair that indicated possible effect modification. The unadjusted association of CKD with mortality was much higher in the group without COPD (OR 1.91; 95% CI 1.85-1.98) than the group with COPD (1.29; 1.13-1.47) (Table 11.1). This indicated that the observed risk of COPD and CKD combined might be lower than expected.

11.1.2 Comorbidity and HF severity

When the effects of each of the comorbidities were observed in quartiles of age, effect estimates were higher in the lower age groups than the higher age groups for mortality and for CKD and first hospital admission. This could indicate (i) that the effect of comorbidity is modified by age (effect modification) which may be due, in part, to competing risks for mortality in older age or (ii) that the absolute effect of the comorbidities is constant

but due to the lower baseline risk in the younger groups, the relative effect of comorbidity in relation to the baseline risk is higher in the younger groups (effect measure modification).

11.1.3 Comorbidity in HF and patient factors

The effect of CKD was higher in men than in women for both outcomes. This gender difference was also found for DM for hospital admission with greater effect estimates of DM for women than men. The effects of COPD and CKD were greater in the high compared to the low haemoglobin groups for mortality and for DM in the high haemoglobin group for hospital admission. Comorbidity effects were higher in the lower risk cardiovascular drug groups (on beta-blocker, on ACEi or ARB, not on diuretics) than the higher risk drug groups for mortality and hospital admission.

11.2 Test of statistical interactions

11.2.1 Comorbidity pairs

There was significant statistical interaction between COPD and CKD for mortality (see Table 11.2) which was less than multiplicative. When an interaction term for having COPD and CKD was included in the model, the likelihood ratio test of the interaction term was significant ($p<0.001$). The effect of having COPD (without CKD) increased to OR 1.52 (1.43, 1.63), of having CKD (without COPD) increased to 1.27 (1.22-1.31) and the added contribution of having both comorbidities was protective OR 0.81 (0.74-0.87). No other comorbidity pairs for either outcome had a significant statistical interaction.

11.2.2 Comorbidity and severity

There was a significant statistical interaction between all three comorbidities and age for mortality which were less than multiplicative. The effect estimates for each of the comorbidities were increased in lower age with a reduction in the comorbidity risk estimate per year of older age (Table 11.2). The biggest influence of age appeared to be for COPD which increased to OR 4.30 (2.90-6.37) following the inclusion of an interaction term with an estimate of 0.98 (0.98-0.99) for each year of older age in those with COPD.

Figure 11.1a shows the observed effect of COPD over age and Figure 11.2b shows the predicted effect of COPD over age. In Figure 11.1a, the observed relative effect of COPD reduces as age increases, which appears to widen out again at the highest values of age. Given the low proportion of observations with Age ≥90 years (7.6%), the second graph based on predicted probabilities does not show this anomaly.

The same reduction in comorbidity effect with age can be seen on Figure 11.2a for CKD and mortality. The observed effect of CKD crosses the line of the no CKD group at older age. The reducing effect of CKD at older age is far more subtle for the predicted effects also shown in Figure 11.2b.

11.2.3 Comorbidity and patient factors

There were significant statistical interactions between gender and COPD for mortality and between gender and CKD for hospital admission (p<0.001). The effect estimate of COPD for mortality was decreased in males over females (11% lower). This was the opposite for CKD and hospital admission where the effect estimate of CKD was increased for males (14% higher, Table 11.2). There was also some evidence that that the effect estimate of DM for hospital admission was decreased in males (10% lower, p<0.05).

The effect estimates for DM and CKD for mortality were reduced significantly in the groups that were not prescribed beta-blockers (p<0.01). There was also some evidence of statistical interaction between haemoglobin and CKD for mortality and haemoglobin and DM for hospital admission. The effect of having both low Hb and CKD was a reduction in risk of mortality of 7% compared to the risk estimates for low Hb and CKD combined (OR for interaction term: 0.93; 0.87-0.99). The effect of having both low Hb and DM was an increase in risk of admission of 12% above the individual risk estimates for both Hb and DM combined (OR for interaction term: 1.12; 1.00-1.22) (p<0.05).

11.3 Test of biological interactions

Biological interaction was investigated between the comorbidity pairs to identify whether the observed effect of having two comorbidities together would differ from the sum of their separate effects. The indication of biological interaction is that where present, the risk of an event will depend at least in part, on the presence of both comorbidities.

Biological interaction was also investigated between each comorbidity and age. Whilst statistical interaction existed between all three comorbidities and age for mortality, this may be due to the difference in baseline risk in different age groups, leading to variation in the relative risks of comorbidity with age. This can be the case where there is no actual difference in the *absolute* risk associated with comorbidity for different ages. Biological interaction was investigated here to explore this association. Lastly biological interaction was also explored between each comorbidity and gender.

Biological interaction, which tests departure from additivity of two factors, cannot be tested directly within logistic or other multiplicative models and requires specific analysis described in detail in Chapter 6 (Section 6.4.2). The Relative Excess Risk due to Interaction (RERI) and synergy index (S) are used to indicate the presence of biological interaction. These refer to the excess risk from joint exposure above the sum of the individual exposures (RERI) and the excess risk from joint exposure in the presence of interaction, relative to the risk from exposure without interaction (S). If there is no excess risk from interaction then the RERI = 0 and the S index = 1.

11.3.1 Comorbidity pairs

The biological interaction measures for each disease pair for both outcomes are summarised in Table 11.3. Worked examples are provided for the disease combinations where interaction was present which were between DM and CKD and between COPD and CKD for mortality.

- **Diabetes and CKD**

DM and CKD were categorised into four mutually exclusive groups of disease combinations as follows; $DM^{+}CKD^{-}$, $DM^{-}CKD^{+}$, $DM^{+}CKD^{+}$ and $DM^{-}CKD^{-}$ (reference group). The expected risk associated with $DM^{+}CKD^{+}$ was calculated by using the equation:

$$\text{OR VAR}_{11}(\text{expected}) = \text{OR VAR}_{10} + \text{OR VAR}_{01} - 1$$

where OR is the odds ratio, VAR_{10} is DM+CKD-, VAR_{01} is DM-CKD+, VAR_{11} is DM+CKD+ and VAR_{00} is DM-CKD- (reference group).

OR 1 occurs where there is no risk associated with a group, relative to the reference group (VAR_{00}) and this equates to the background risk. The -1 in the equation is to account for adding in the background risk (VAR_{00}) twice due to both comorbidity groups in the equation. This resulted in an expected risk of 1.277 + 1.201 – 1 = 1.478.

This was below that observed for DM & CKD which was OR 1.615 (Table 11.3). The relative excess risk due to interaction was calculated as

$$\text{RERI} = \text{OR VAR}_{11} - \text{OR VAR}_{10} - \text{OR VAR}_{01} + 1$$

For DM & CKD this was 1.615 – 1.277 – 1.201 + 1 = 0.137 which had a 95% CI 0.04-0.24.

An increase in absolute risk of 14% above the risk associated with the individual diseases due to interaction between DM and CKD.

The synergy index was calculated as;

$$\frac{\text{OR VAR11} - 1}{(\text{OR VAR10} - 1) + (\text{OR VAR01} - 1)}$$

For DM & CKD this was $\frac{1.615-1}{(1.277-1)+(1.201-1)} = 1.285$ which had a 95% CI 1.05-1.57

This equates to 29% more risk through the presence of interaction relative to if there was no interaction.

Figure 11.3 shows the excess risk associated with the presence of interaction of DM and CKD for mortality. This was not found for these diseases for hospital admission (see Figure 11.4) where the RERI and S were non-significant.

- **COPD and CKD**

COPD and CKD were categorised into four mutually exclusive groups of disease combinations as follows; $COPD^{+}CKD^{-}$, $COPD^{-}CKD^{+}$, $COPD^{+}CKD^{+}$ and $COPD^{-}CKD^{-}$ (reference group). The expected risk associated with $COPD^{+}CKD^{+}$ was calculated by using the equation:

$$\text{OR VAR}_{11}(\text{expected}) = \text{OR VAR}_{10} + \text{OR VAR}_{01} - 1$$

where OR is the odds ratio, VAR_{10} is COPD+CKD-, VAR_{01} is COPD-CKD+, VAR_{11} is COPD+CKD+ and VAR_{00} is COPD-CKD- (reference group).

This resulted in an expected risk of 1.263 + 1.525 – 1 = 1.79. This was higher that observed for COPD&CKD which was OR 1.546 (Table 11.3).

The relative excess risk due to interaction was calculated as:

$$\text{RERI} = \text{OR VAR}_{11} - \text{OR VAR}_{10} - \text{OR VAR}_{01} + 1$$

For COPD & CKD this was 1.546 – 1.263 – 1.525 + 1 = -0.24 which had a 95% CI -0.37 to -0.11. A reduction in absolute risk of 24% below the risk associated with the individual diseases, due to interaction between COPD and CKD. The synergy index was calculated as:

$$\frac{\text{OR VAR11} - 1}{(\text{OR VAR10} - 1) + (\text{OR VAR01} - 1)}$$

For COPD & CKD this was $\frac{1.546-1}{(1.263-1)+(1.525-1)} = 0.694$ which had a 95% CI 0.57- 0.84

This equates to 30% less risk through the presence of interaction relative to if there was no interaction. Figure 11.5 shows the reduced risk associated with the presence of interaction of COPD and CKD for mortality. This was not found for these diseases for hospital admission (Figure 11.6) where there was a slight increased risk associated with this combination but the interaction measures (RERI and S) were insignificant.

11.3.2 Comorbidity and severity

There was no significant statistical interaction between comorbidities and age for hospital admission and this was also the case when tested for biological interaction. None of the measures of biological interaction were significant (Figure 11.7). There was significant sub-additive biological interaction between COPD and age for mortality (S 0.85; 95% CI 0.76-0.95). The effect of COPD was less in older age groups. This was the same direction of effect that was found in the multiplicative model where there was significant statistical interaction between COPD and older age which was less than multiplicative.

DM and older age, which had significant statistical interaction for mortality, did not have significant biological interaction. There was a significant super-additive biological interaction between CKD and older age (S 1.17; 1.05-1.30) which went in the opposite direction than the statistical interaction test between CKD and older age (which was protective). CKD was associated with higher risk in the older age groups.

11.3.3 Comorbidity and patient factors

In the statistical interaction tests, there were 3 significant findings (i) COPD and male gender for mortality was less than multiplicative (ii) CKD and male gender for hospital admission was more than multiplicative and (iii) DM and male gender for hospital admission was protective (10% lower relatives to females). The test of biological interactions between COPD or DM and gender for mortality were insignificant but there was

significant biological interaction between CKD and gender for mortality with males with CKD having a 13% increase in the risk of death relative to females.

The biological interactions between COPD or DM and gender for hospital admission were non-significant. There was a significant biological interaction between CKD and male gender for hospital admission. This association was super-additive with males with CKD having a 18% higher risk for admission than females with CKD (Figure 11.8).

11.4 Chapter summary

There were interactions between CKD and both COPD and DM. For COPD and CKD there was significant statistical and biological interaction that reduced their combined association with mortality when experienced together. Biological but not statistical interaction was present between DM and CKD for mortality with a more than additive effect when these diseases were experienced together.

Statistical interaction but not biological interaction was present for all three comorbidities and age for mortality, reducing the relative effect estimates in the older age groups. Biological interaction was present for two of the comorbidities with age. COPD and older age had a sub-additive effect when both were experienced together whilst CKD and older age had a super-additive effect. This latter interaction was in the opposite direction to the statistical interaction and showed an increase in risk in older people with CKD than would be expected.

COPD was associated with a significant statistical interaction with male gender for mortality and DM and gender for hospital admission. The relative effect of both diseases in males was lower than in females. These interactions were not significant when tested for biological interaction. CKD was associated with a significant statistical interaction with male gender for hospital admission. The relative effect of CKD for hospital admission was increased in males. This combined exposure group was also significant when tested for biological interaction in the same direction but for both outcomes. These results will be discussed in detail in Chapter 13.

Tables

Table 11.1 Stratification of comorbidity effects

Potential confounders	Mortality			Hospital admission		
	Diabetes	**COPD**	**CKD**	**Diabetes**	**COPD**	**CKD**
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Unadjusted	1.09 (1.05-1.12)	1.41 (1.36-1.46)	1.77 (1.72-1.82)	1.33 (1.28-1.37)	1.36 (1.31-1.43)	1.34 (1.29-1.40)
Person and socio-demographic factors						
Age quartile						
1	1.54 (1.38-1.71)	2.03 (1.78-2.30)	2.22 (1.98-2.50)	1.31 (1.19-1.45)	1.40 (1.23-1.59)	1.69 (1.46-1.96)
2	1.32 (1.21-1.44)	1.70 (1.53-1.88)	1.41 (1.29-1.54)	1.36 (1.23-1.50)	1.55 (1.37-1.76)	1.21 (1.08-1.37)
3	1.21 (1.11-1.33)	1.65 (1.49-1.84)	1.36 (1.24-1.50)	1.54 (1.38-1.72)	1.60 (1.40-1.83)	1.18 (1.04-1.34)
4	1.16 (1.06-1.13)	1.17 (1.05-1.31)	1.21 (1.12-1.31)	1.52 (1.35-1.72)	1.41 (1.21-1.64)	1.16 (1.02-1.32)
Male	1.10 (1.05-1.16)	1.47 (1.39-1.55)	1.94 (1.85-2.03)	1.24 (1.18-1.31)	1.36 (1.28-1.46)	1.53 (1.44-1.64)
Female	1.13 (1.07-1.19)	1.36 (1.27-1.46)	1.55 (1.47-1.63)	1.44 (1.35-1.53)	1.38 (1.27-1.50)	1.17 (1.09-1.26)
Anthropometric and clinical factors						
BMI quartile						
1	1.36 (1.24-1.50)	1.24 (1.13-1.36)	1.45 (1.34-1.57)	1.45 (1.28-1.65)	1.45 (1.28-1.64)	1.20 (1.06-1.35)
2	1.34 (1.22-1.47)	1.45 (1.30-1.62)	1.72 (1.57-1.88)	1.50 (1.34-1.69)	1.28 (1.12-1.47)	1.29(1.13-1.46)
3	1.29 (1.17-1.41)	1.37-1.22-1.54)	2.02 (1.83-2.23)	1.20 (1.08-1.33)	1.33 (1.15-1.53)	1.31 (1.15-1.50)
4	1.25 (1.14-1.37)	1.52 (1.35-1.72)	2.19 (1.98-2.43)	1.38 (1.26-1.52)	1.36 (1.19-1.56)	1.41 (1.25-1.60)
Cholesterol high	1.09 (1.03-1.16)	1.39 (1.31-1.48)	1.71 (1.62-1.80)	1.29 (1.20-1.38)	1.40 (1.30-1.51)	1.25 (1.17-1.35)
Cholesterol low	1.04 (0.99-1.09)	1.39 (1.31-1.47)	1.84 (1.75-1.93)	1.33 (1.26-1.40)	1.40 (1.30-1.50)	1.20 (1.13-1.28)
Haemoglobin high	1.10 (1.03-1.17)	1.59 (1.49-1.70)	1.74 (1.65-1.84)	1.14 (1.07-1.21)	1.39 (1.29-1.49)	1.22 (1.13-1.31)
Haemoglobin low	0.96 (0.91-1.00)	1.36 (1.28-1.44)	1.38 (1.31-1.44)	1.14 (1.07-1.21)	1.39 (1.29-1.49)	1.22 (1.13-1.31)
Systolic BP high	1.07 (1.02-1.12)	1.45 (1.36-1.54)	1.67 (1.59-1.76)	1.39 (1.31-1.48)	1.33 (1.23-1.43)	1.31 (1.22-1.41)
Systolic BP low	1.08 (1.03-1.14)	1.34 (1.26-1.42)	1.84 (1.75-1.94)	1.28 (1.21-1.35)	1.37 (1.28-1.47)	1.40(1.32-1.49)
Lifestyle factors						
Smoking yes	1.05 (0.87-1.26)	1.43 (1.20-1.70)	1.83 (1.51-2.22)	1.09 (0.87-1.38)	1.33 (1.08-1.65)	1.04 (0.77-1.41)
Smoking no	1.08 (1.05-1.12)	1.39 (1.33-1.45)	1.78 (1.72-1.84)	1.34 (1.29-1.40)	1.36 (1.29-1.42)	1.34 (1.28-1.39)
Alcohol yes	1.09 (1.05-1.14)	1.43 (1.36-1.50)	1.79 (1.72-1.87)	1.34 (1.28-1.40)	1.37 (1.30-1.45)	1.34 (1.28-1.41)

Alcohol no	1.03 (0.97-1.11)	1.33 (1.22-1.45)	1.66 (1.55-1.78)	1.26 (1.15-1.37)	1.33 (1.18-1.51)	1.30 (1.15-1.46)
Drug factors						
Beta-blocker yes	1.21 (1.15-1.26)	1.37 (1.28-1.47)	2.10 (2.00-2.21)	1.37 (1.28-1.47)	1.37 (1.28-1.47)	1.40 (1.30-1.52)
Beta-blocker no	1.01 (0.96-1.07)	1.15 (1.09-1.22)	1.56 (1.48-1.64)	1.38 (1.31-1.45)	1.15 (1.09-1.22)	1.31 (1.23-1.39)
ACEi or ARB yes	1.18 (1.13-1.23)	1.51 (1.43-1.58)	1.90 (1.82-1.98)	1.40 (1.34-1.46)	1.51 (1.43-1.58)	1.34 (1.28-1.41)
ACEi or ARB no	1.11 (1.03-1.20)	1.14 (1.05-1.23)	1.51 (1.40-1.61)	1.45 (1.31-1.61)	1.14 (1.05-1.23)	1.37 (1.22-1.54)
Diuretic yes	1.02 (0.99-1.06)	1.39 (1.33-1.45)	1.66 (1.60-1.72)	1.31 (1.25-1.36)	1.39 (1.33-1.45)	1.33 (1.27-1.40)
Diuretic no	1.23 (1.10-1.38)	1.31 (1.14-1.49)	1.81 (1.63-2.02)	1.48 (1.34-1.63)	1.31 (1.14-1.49)	1.49 (1.31-1.69)
Comorbidity exposures						
Diabetes	-	1.51 (1.35-1.69)	1.95 (1.78-2.13)	-	1.54 (1.30-1.82)	1.50 (1.31-1.70)
No diabetes	-	1.42 (1.36-1.48)	1.76 (1.69-1.82)	-	1.38 (1.32-1.46)	1.29 (1.22-1.35)
COPD	1.06 (0.92-1.21)	-	1.29 (1.13-1.47)	1.28 (1.03-1.60)	-	1.38 (1.06-1.79)
No COPD	1.10 (1.06-1.13)	-	1.91 (1.85-1.98)	1.34 (1.29-1.39)	-	1.35 (1.29-1.41)
CKD	1.08 (1.03-1.13)	1.29 (1.22-1.37)	-	1.24 (1.14-1.35)	1.43 (1.29-1.58)	-
No CKD	1.05 (0.98-1.13)	1.79 (1.66-1.93)	-	1.32 (1.23-1.42)	1.37 (1.24-1.53)	-

Framed estimates are those that differed across strata with no overlap of confidence intervals. BMI, body mass index; BP, blood pressure; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; COPD, chronic obstructive pulmonary disease; CKD, chronic kidney disease

Table 11.2 Statistical interaction tests

	Unadjusted log likelihood	Interaction *P*	Adjusted log likelihood	Interaction *P*	Independent effects adjusted for interaction OR (95% CI)	Interaction effects
Comorbidity pairs: mortality						
DM and COPD	-42840.51		-29511.79			
DM*COPD	-42840.48	0.811	-29511.29	0.318		
DM and CKD	-35600.98		-29511.79			
DM*CKD	-35600.00	0.162	-29510.92	0.187		
COPD and CKD	-35425.10		-29511.79		COPD 1.52 (1.43-1.63)	COPD*CKD 0.81 (0.74-0.87)
COPD*CKD	-35392.85	<0.001	-29499.30	<0.001	CKD 1.27 (1.22-1.31)	
Comorbidity pairs: hospital admission						
DM and COPD	-36125.46		-15316.33			
DM*COPD	-36125.45	0.93	-15315.58	0.22		
DM and CKD	-18214.91		-15380.32			
DM*CKD	-18212.47	0.03	-15380.26	0.73		
COPD and CKD	-18218.95		-15380.32			
COPD*CKD	-18218.83	0.025	-15380.21	0.64		
Comorbidity and Age: mortality						
DM and Age	-40027.62		-29081.90		DM 2.73 (1.99-3.75)	DM*Age 0.99 (0.99-0.99)
DM*Age	-40011.62	<0.001	-29070.95	<0.001	Age 1.05 (1.05-1.05)	
COPD and Age	-39885.05		-29081.90		COPD 4.30 (2.90-6.37)	COPD*Age 0.98 (0.98-0.99)
COPD*Age	-39848.33	<0.001	-29065.31	<0.001	Age 1.05 (1.05-1.05)	
CKD and Age	-33604.71		-29511.79		CKD 2.40 (1.80-3.21)	CKD*Age 0.99 (0.99-0.99)
CKD*Age	-33582.40	<0.001	-29500.91	<0.001	Age 1.05 (1.05-1.05)	
Comorbidity and Age: HA						
DM and Age	-36078.07		-15316.33			
DM*Age	-36077.01	0.145	-15316.08	0.48		
COPD and Age	-36123.03		-15316.32			

COPD*Age	-36122.58	0.345	-16948.77	0.84		
CKD and Age	-18259.59		-15380.32			
CKD*Age	-18257.07	0.025	-15379.01	0.11		
Comorbidity and patient factors: mortality						
COPD and gender	-42808.57	0.30	-29081.90	<0.01	COPD 1.44 (1.35-1.55)	COPD*Male 0.89 (0.82-0.97)
COPD*gender	-42808.03		-29078.48		Male 1.17 (1.12-1.21)	
CKD and gender	-35602.94		-29511.79			
CKD*gender	-35579.55	<0.001	-29511.79	0.96		
DM and gender	-42966.02	0.82	-29081.90			
DM*gender	-42965.99		-29081.78	0.63		
DM and beta-blocker	-42245.92		-29511.79		DM 1.34 (1.27-1.41)	
DM*beta-blocker	-42233.74	<0.001	-29508.40	<0.01	no BB 1.41(1.36-1.47)	DM*No BB 0.91 (0.85-0.98)
COPD and beta-blocker	-42207.21		-29511.79			
COPD*beta-blocker and interaction	-42196.47	<0.001	-29511.58	0.516		
CKD and beta-blocker	-34949.16		-29511.79		CKD 1.31 (1.27-1.37)	CKD*No BB 0.87 (0.81-0.93
CKD*beta-blocker	-34911.92	<0.001	-29502.82	<0.001	No BB 1.48 (1.40-1.57)	
CKD and Hb (binary; ≤13,>13)						
CKD and Hb	-34565.28		-29994.77			
CKD*Hb	-34540.37	<0.001	-29992.68	0.04	Hb cat(<13)1.72(1.63-1.81) CKD1.31 (1.24-1.38)	Hb binary*CKD 0.93 (0.87-0.99)
Comorbidities and patient factors: HA						
COPD and gender	-36236.43	0.828	-15316.32	0.97		
COPD*gender	-36236.41		-15316.32			
CKD and gender	-18256.667		-15380.32		CKD 1.07 (1.00-1.14)	CKD*Male 1.14 (1.05-1.25)
CKD*gender	-18246.395	<0.001	-15375.77	<0.01	Male 1.13 (1.06-1.21)	
DM and gender	-36209.7	<0.001	-15316.32		DM 1.32 (1.22-1.42)	DM*Male 0.90 (0.81-0.99)
DM*gender	-36202.469		-15314.12	0.04	Male 1.26 (1.19-1.33)	
COPD and beta-blocker	-36200.13		-15316.33			
COPD*beta-blocker	-36199.84	0.450	-15315.86	0.34		
COPD and ACE/ARB	-35809.72		-15316.33			
COPD*ACE/ARB	-35809.68	<0.001	-15316.07	0.47		
Other:						
CKD and Ch	-18255.14		-15383.04			
CKD*Ch	-18251.32	<0.01	-15381.51	0.08		
DM and Hb	-35876.48	<0.001	-15545.89		HB (<13.4) 1.32 (1.25,1.39)	HB binary*DM 1.12 (1.00-1.22)

DM*Hb	-35870.04	-15543.85	0.04	DM 1.20 (1.11-1.30)

*interaction; each adjusted model included age, gender, COPD, eGFR, DM, no beta-blocker, no ACEi or ARB, diuretics, BP, BMI, cholesterol, haemoglobin, smoking and alcohol status; Cat, categories. BMI, body mass index; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; CKD, chronic kidney disease

Table 11.3 Biological interaction measures for the comorbidity pair combinations

Mortality				Hospital admission			
Exposure	**RR**	**Lower CI**	**Upper CI**	**Exposure**	**RR**	**Lower CI**	**Upper CI**
Diabetes	1.277	1.199	1.361	Diabetes	1.241	1.150	1.339
CKD	1.201	1.154	1.250	CKD	1.147	1.088	1.208
Diabetes*CKD	1.615	1.535	1.699	Diabetes*CKD	1.448	1.349	1.554
Measure	Estimate	Lower CI	Upper CI	Measure	Estimate	Lower CI	Upper CI
RERI	0.137	0.037	0.236	RERI	0.060	-0.065	0.186
S	1.285	1.053	1.569	S	1.156	0.847	1.579
Exposure	**RR**	**Lower CI**	**Upper CI**	**Exposure**	**RR**	**Lower CI**	**Upper CI**
Diabetes	1.292	1.239	1.347	Diabetes	1.254	1.187	1.325
COPD	1.362	1.293	1.435	COPD	1.351	1.251	1.459
Diabetes*COPD	1.692	1.558	1.838	Diabetes*COPD	1.542	1.352	1.759
Measure	Estimate	Lower CI	Upper CI	Measure	Estimate	Lower CI	Upper CI
RERI	0.038	-0.115	0.191	RERI	-0.064	-0.289	0.161
S	1.058	0.844	1.325	S	0.895	0.597	1.341
Exposure	**RR**	**Lower CI**	**Upper CI**	**Exposure**	**RR**	**Lower CI**	**Upper CI**
CKD	1.263	1.215	1.312	CKD	1.147	1.092	1.205
COPD	1.525	1.427	1.629	COPD	1.292	1.177	1.419
CKD*COPD	1.546	1.452	1.646	CKD*COPD	1.529	1.389	1.682
Measure	Estimate	Lower CI	Upper CI	Measure	Estimate	Lower CI	Upper CI
RERI	-0.241	-0.369	-0.113	RERI	0.089	-0.089	0.268
S	0.694	0.573	0.840	S	1.204	0.828	1.750

*interaction; CKD, Chronic kidney disease; COPD, chronic obstructive pulmonary disease. Red = significant biological interaction.

Figures

Figure 11.1a Observed COPD and Age interaction for mortality

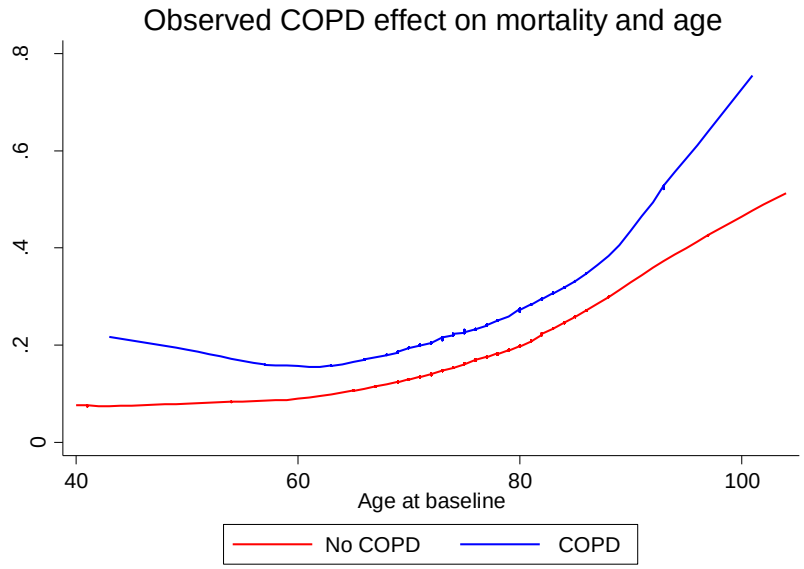


Figure 11.1b Predicted COPD and Age interaction for mortality

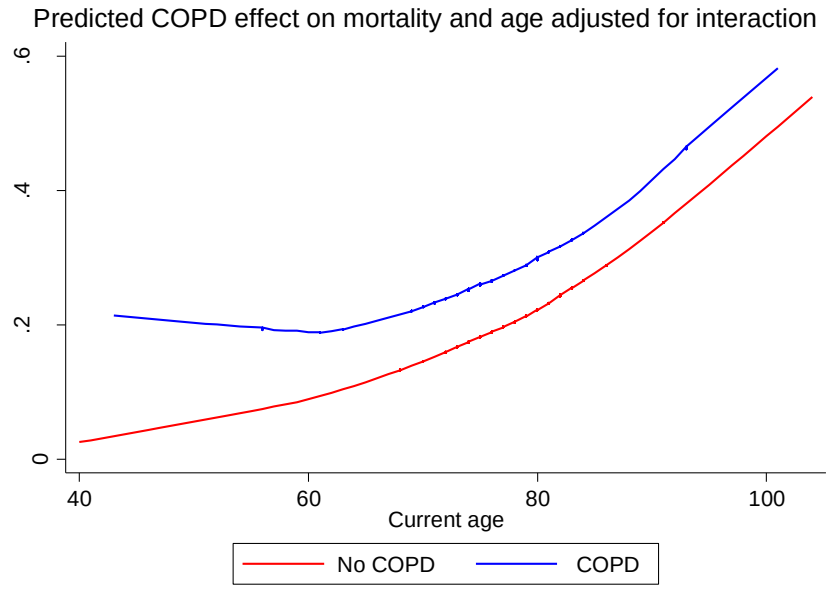


Figure 11.2a Observed CKD and Age interaction for mortality

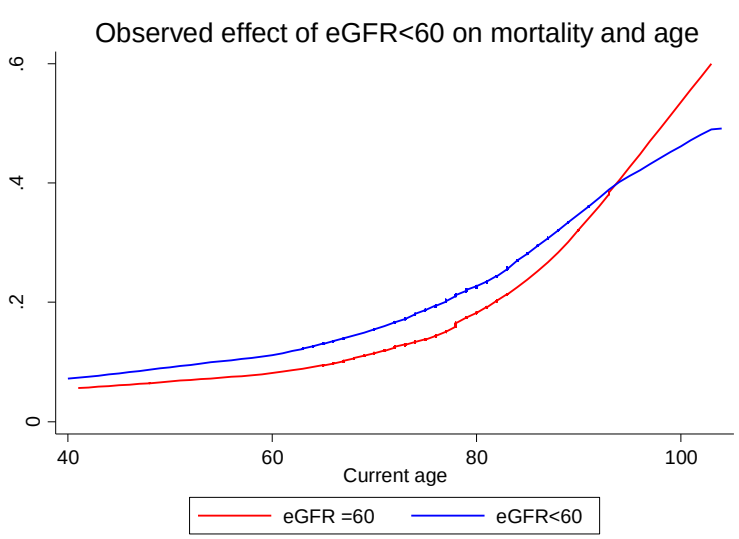


Figure 11.2b Predicted CKD and Age interaction for mortality

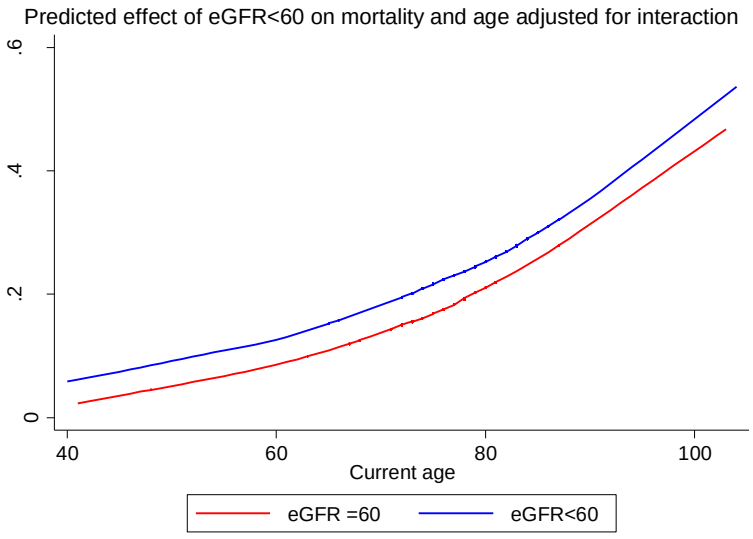


Figure 11.3 Biological DM and CKD interaction for mortality

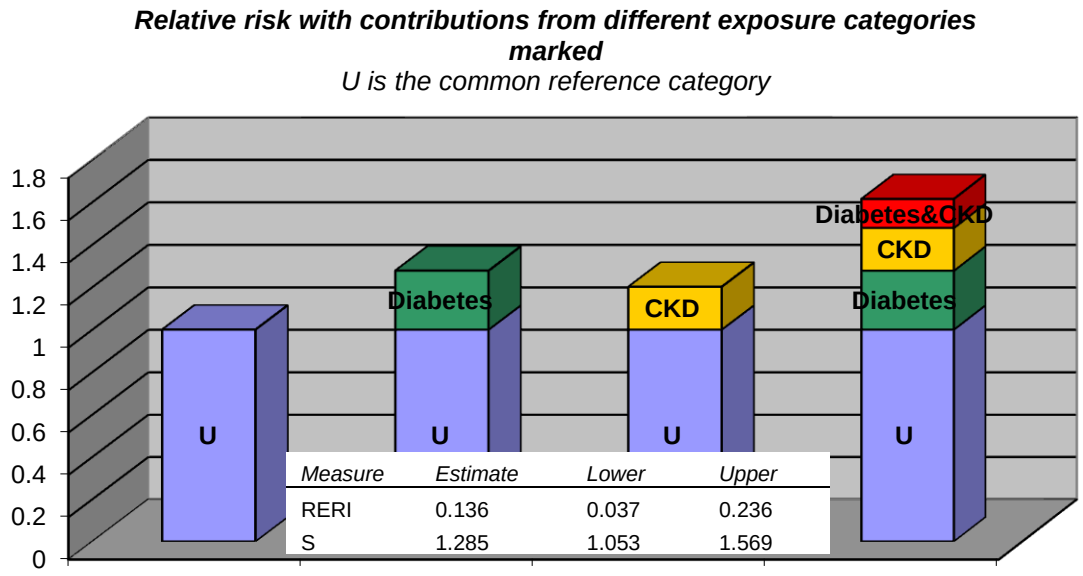


Figure 11.4 Biological DM and CKD interaction for hospital admission

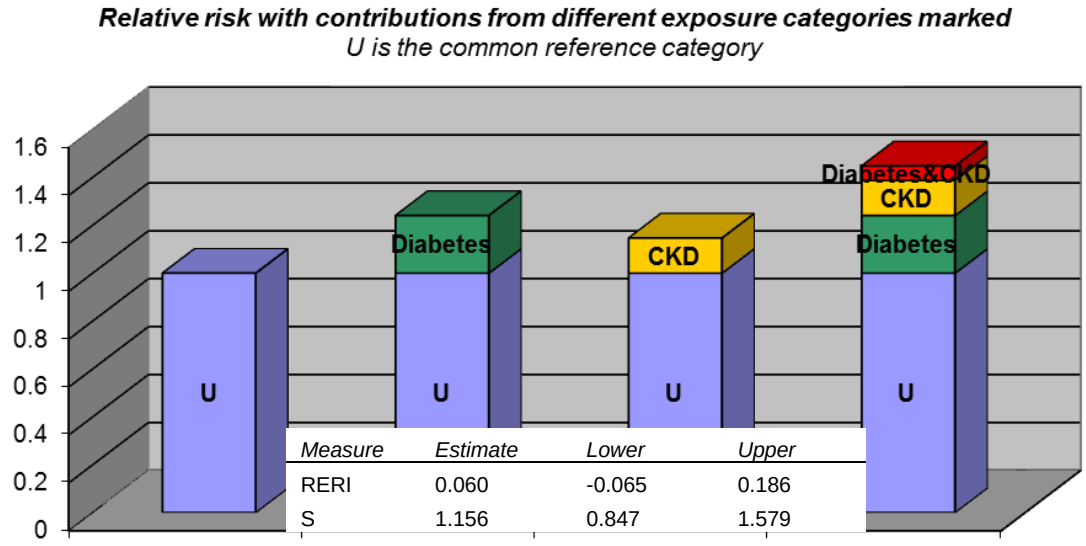


Figure 11.5 Biological COPD and CKD interaction for mortality

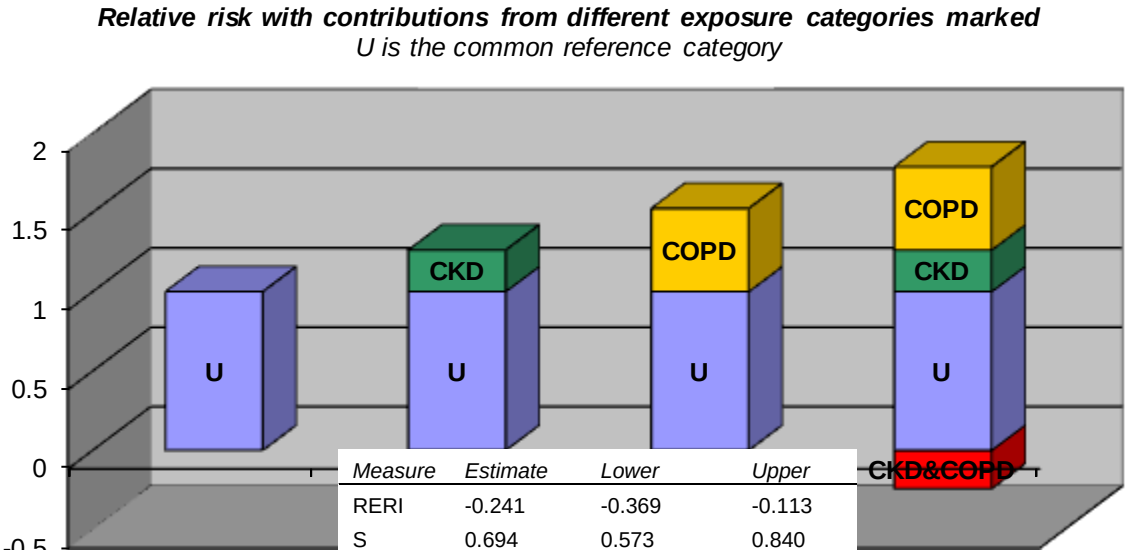


Figure 11.6 Biological COPD and CKD interaction for hospital admission

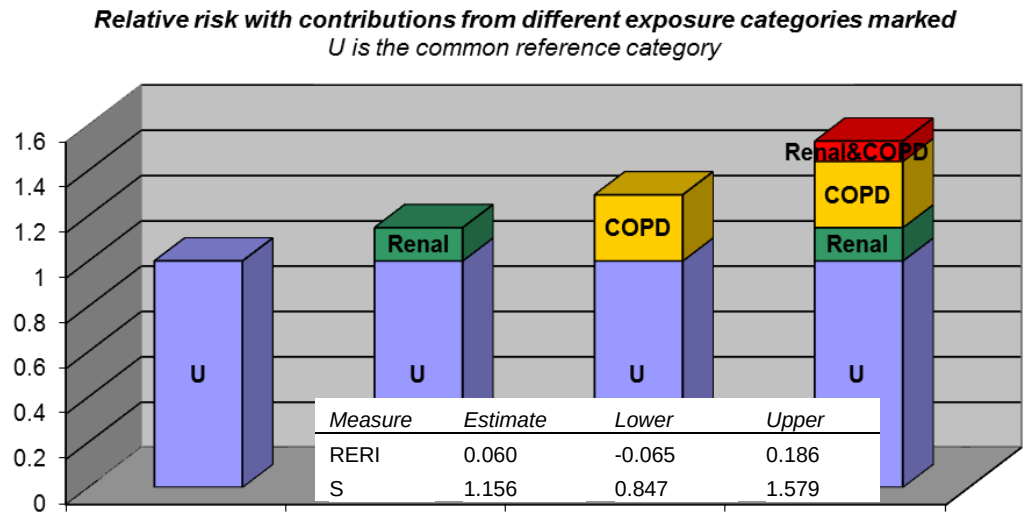
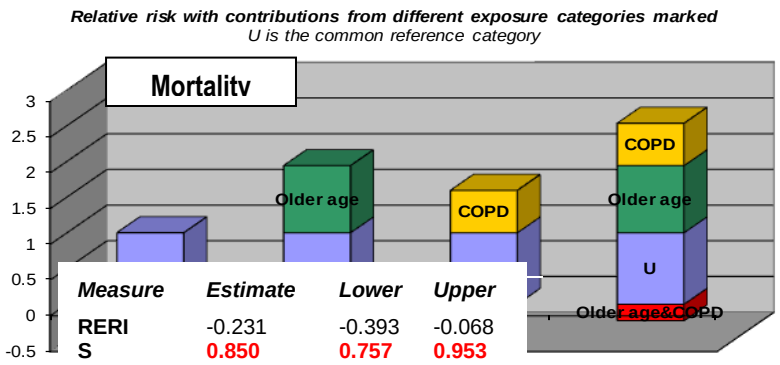
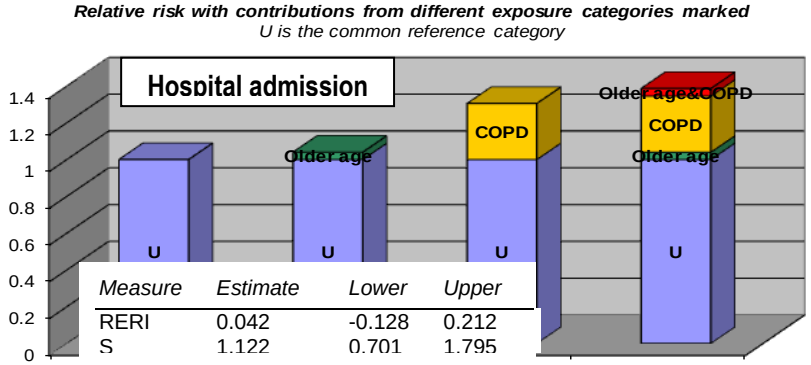


Figure 11.7 Biological comorbidity and age interaction for mortality and hospital admission

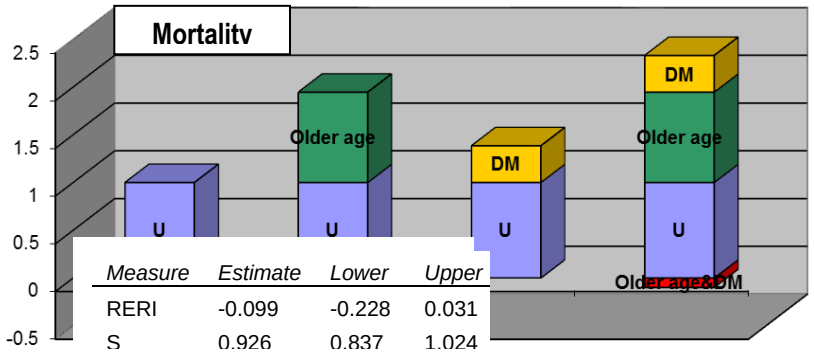
COPD





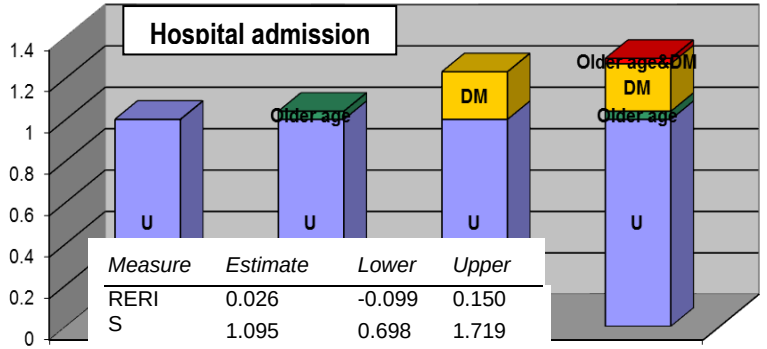
Diabetes

Relative risk with contributions from different exposure categories marked
U is the common reference category



Measure	*Estimate*	*Lower*	*Upper*
RERI	-0.099	-0.228	0.031
S	0.926	0.837	1.024

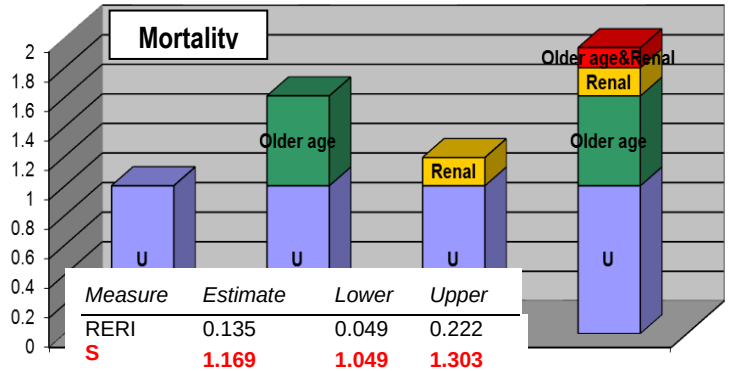
Relative risk with contributions from different exposure categories marked
U is the common reference category



Measure	*Estimate*	*Lower*	*Upper*
RERI	0.026	-0.099	0.150
S	1.095	0.698	1.719

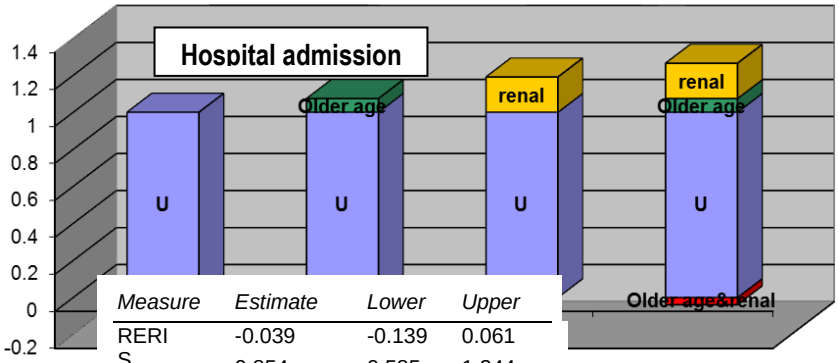
CKD

Relative risk with contributions from different exposure categories marked
U is the common reference category



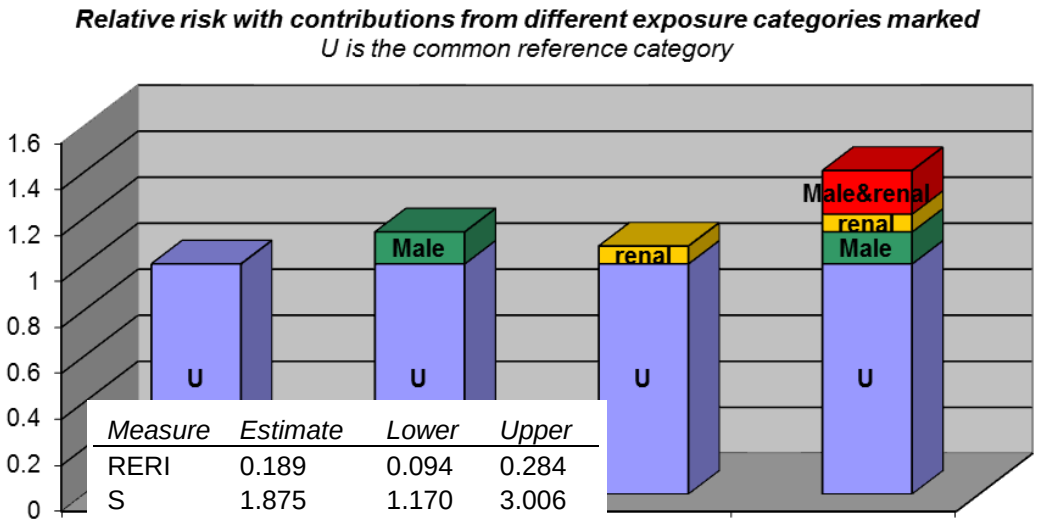
Measure	*Estimate*	*Lower*	*Upper*
RERI	0.135	0.049	0.222
S	**1.169**	**1.049**	**1.303**

Relative risk with contributions from different exposure categories marked
U is the common reference category



Measure	*Estimate*	*Lower*	*Upper*
RERI	-0.039	-0.139	0.061
S	0.854	0.585	1.244

Figure 11.8 Biological CKD and gender interaction for hospital admission



Chapter 12 Model building: testing different comorbidity measures

This chapter investigates and compares the contribution of the different comorbidity measures to pre-specified HF prognostic models for mortality and separately for hospital admission. The work was carried out in three stages: (i) building the core model, (ii) comparing the fit of a set of different models with each of the comorbidity status, severity and severity change measures added and replaced in turn (for DM, COPD, CKD separately) and (iii) comparing the fit of a set of models with comorbidity status, severity and severity change measures for all of the three comorbidities included together.

12.1 Building the core model

First the pre specified model was built for each outcome using all covariates in the adjusted models in Chapters 9 and 10 (termed core model). The core model included age, gender, alcohol, smoking, systolic blood pressure, cholesterol, BMI, haemoglobin, HF medication and quadratic extensions where indicated previously. In addition, prior hospital admission was added to the core models for the hospital admission outcome. All continuous variables were centered about their means prior to their inclusion in the model to remove collinearity with their respective quadratic term. This also meant that the risk estimate generated by the model would be comparable to a patient with the reference category of each of the categorical variables and the average value of the continuous variable.

12.2 Testing measures for each comorbidity

The first approach was to investigate each of the comorbidities separately. A set of models for each of the three comorbidities were constructed. The set of models for COPD example would be:

a) Core model* + COPD (status)

b) Core model* + COPD (severity using drug definition)

c) Core model* + COPD (severity change using drug change definition)

*DM and CKD were added to the core model

First, COPD was added to the core model by presence or absence status (model a). Second, the COPD status measure was exchanged for the COPD severity measure (drugs; model b1 *or* physiological measure; model b2). Lastly the COPD severity measure was exchanged for a severity change measure (drugs; model c1 *or* physiological measure; model c2). The remaining two comorbidities (DM, CKD) were added to each core model by their status measure only. An additional step was included for the CKD models. Given that all HF patients had an eGFR measure, severity and change measures could also be simultaneously included in the final step of the CKD model testing.

Conditional logistic regression was performed for each model and tests of model fit were used to compare models. Tests included the receiver operating curve (ROC), log-likelihood, likelihood ratio test, Akaike and Bayesian information criterion (AIC and BIC) and McFadden's R^2 (Chapter 6.4.3). Models were first created using the full data. Due to varying amounts of missing data in the comorbidity measures, in the second step, models were restricted to the data which had complete observations for all the models, so that they could be compared.

12.3 Testing measures for all comorbidities

In a final step, models were compared using different measures for all three comorbidities simultaneously included by (i) status, (ii) status exchanged for comorbidity severity measures and lastly (ii) severity measures exchanged for severity change measures. An example risk score was developed for the best fitting model for both outcomes.

12.4 HF comorbidity and mortality models

12.4.1 Mortality models with comorbid DM

Using the drug definitions of DM severity and change, there were incremental improvements of the diabetes HF mortality model when the DM status measure was exchanged by the DM severity measure and then the DM severity change measure (Table 12.1a). The most improvement was observed when the status measure was exchanged for the severity change measure (LR test $p<0.001$). There was some improvement in Pseudo R^2, AIC, BIC and log likelihood when the severity measure was exchanged for the severity change measure and the LR test was non-significant for model improvement for the severity model compared to the change model. The same pattern of improvement was observed when the DM severity measures using the physiological definition of DM severity (HbA1c) were compared (Table 12.1b).

When the reduced DM drug severity model was compared to the DM physiological severity model directly using the complete data, the drug measures of severity and change showed improved model fit across all measures compared to the physiological measures.

12.4.2 Mortality model with comorbid COPD

Using the drug definitions of COPD severity, there was improvement of the comorbid COPD mortality model with the replacement of the status measure for the severity measure and the change measure (Table 12.2a). When the status measure was exchanged for the severity measures the AIC and BIC improved by 189 and 170 respectively. The improvement was more modest when the status measure was exchanged for the change measure but both models had a significantly better fit than the status model ($p<0.001$).

When comparing the mortality models using physiological measures of COPD (FEV1), replacement of COPD status by measures of severity showed some improvement in AIC and BIC but this was less than the improvement observed using the drug measures. Log likelihood, ROC and pseudo R^2 also improved (LR test; $p<0.001$). The COPD measure using physiological change was no better than the COPD status model with

similar model fit (Table 12.2b) (LR test; p=0.271). When the reduced COPD drug severity model was compared to the COPD physiological severity model directly (same number of observations), the drug measures of severity and change showed slightly improved model fit across all measures compared to the physiological measures.

12.4.3 Mortality model with comorbid CKD

There was a substantial improvement in model fit when CKD status was replaced by CKD severity measure (eGFR) and the severity change measure (Table 12.3). AIC and BIC improved by 732 and 723 with the replacement of status by severity and by 330 and 312 respectively when the status measure was exchanged for the severity change measure. Pseudo R^2, log likelihood and ROC favoured the severity model. However when baseline eGFR was added to the severity change model (model d) the model had the best overall fit and showed improvement from the severity model (model b) with an improvement of 269 in AIC and 336 in BIC. Following the addition of baseline eGFR into the change model there was incremental improvement from status to severity to change. The LR test for each stage of improvement was significant (p<0.001).

12.5 Mortality model: all comorbidities

A final model comparison was made between the core model with all three comorbidities added by status (status model), followed by severity measures (using drug measures for DM and COPD and eGFR for CKD) and finally severity change measures (using drug measures of change for DM and COPD and eGFR change for CKD) (Table 12.4). Baseline eGFR was then added to this final model.

There was improvement in model fit from the status model to the severity model and from the status model to the change model. The most improvement was observed between the status and the severity model with an improvement of AIC of 937.5, BIC of 890, log likelihood of 474, pseudo R^2 of 0.015 and ROC of 0.01. The severity model had the best model fit until baseline eGFR was added to the change model. This then improved significantly from the severity model (p<0.001).

12.5.1 Example HF comorbidity mortality model

An example model was constructed using all covariates and the measures of best fit for each comorbidity that were observed in the previous investigations. These were: DM drug severity change, COPD drug severity, CKD severity and change. A linear predictor for the risk of mortality was generated by the model. A histogram of this linear predictor showed a range between -2 and 6 with most predictions between 0 and 2 (Figure 12.1). The risk of mortality for the different units of linear predictor are shown in Table 12.5. Taking the centering and quadratic extensions into account the linear predictor for an individual patient would be calculated as set out in Table 12.6. As the continuous predictors were centred at their means, the risk per unit can be interpreted as the risk difference from a HF patient with an average level of haemoglobin, eGFR, systolic BP, cholesterol and BMI and in the reference category of all other predictors. Due to the absence of an intercept, which is inherent in conditional logistic regression, the actual risk cannot be determined from the model. Risk groups are instead constructed based on proportion of risk above the baseline risk.

The C-index for the ROC for the model was 74% (Figure 12.2). This can be interpreted as a 74% chance that in a randomly selected case and control, the case would be given a higher predicted risk than the control(463).

12.6 Heart failure comorbidity and first hospital admission models

12.6.1 Hospital admission model with comorbid DM

Model testing for the DM models was focused on the drug severity measures. This was due to these measures having better fit in the mortality models, the higher percentage of missing physiological data in the narrower measurement time-window for the HF-HA sub-sample and the non-stratified effects for categories of HbA1c change in the adjusted associations with hospital admission (Chapter 10). Using the drug definitions of severity, there was improvement of the diabetes HF hospital admission model with the replacement of the status measure with the severity measure (Table 12.7) (LR test $p<0.001$) and across all measures of model fit (AIC, BIC, Pseudo R^2 and log likelihood); AIC improved by 35 and BIC by 18. The improvement was less for the severity change measure which only had borderline significance for improvement from the status measure ($p=0.05$).

12.6.2 Hospital admission model with comorbid COPD

Using the drug definitions of COPD severity, there was improvement of the hospital admission model with the replacement of the status measure with the severity measure (Table 12.8) (LR test $p<0.001$). AIC improved by 35 and BIC by 17. There was no improvement in model fit by exchanging the status measure for the severity change measure (LR test $p=0.585$) and all measures of model fit were similar.

12.6.3 Hospital admission model with comorbid CKD

There was a substantial improvement in model fit when CKD status was replaced by severity and the severity change measure (Table 12.9). AIC and BIC improved by 82 and 39 with the replacement of status by severity and by 61 and 45 respectively when the status measure was exchanged for the severity change measure. Pseudo R^2, log likelihood and ROC favoured the severity model and this had significantly better fit than the change model ($p<0.001$). However, when baseline eGFR categories were added to the severity change model (model d) the model had the best overall fit and showed improvement from the severity model (model b) with further improvement of 48 in AIC and 23 in BIC. Following the addition of baseline eGFR into the

change model there was incremental improvement from status to severity to change. The LR test for each stage of improvement was significant (p<0.001).

12.7 HF hospital admission model: all comorbidities

A final model comparison was made between the core model with all three comorbidities added by status (status model), followed by severity measures (using drug measures for DM and COPD and categories for CKD) and finally severity change measures (using drug measures of change for DM and COPD and eGFR change for CKD) (Table 12.10). Baseline CKD categories were then added to this final model.

There was improvement in model fit from the status model to the severity model and from the status model to the change model. The most improvement was observed between the status and the severity model with an improvement of AIC of 124, BIC of 48, log-likelihood of 71, pseudo R^2 of 0.01 and ROC of 0.04. The severity model had the best model fit until baseline renal status was added to the change model. This then showed some improvement to the severity model (p=0.02).

12.7.1 Example model

An example model was constructed using all covariates and the measures of best fit for each comorbidity that were observed in the prior investigations. These were; DM drug severity, COPD drug severity, CKD severity and change. A linear predictor for the risk of first hospital admission was generated by the model. A histogram of this linear predictor showed a range between -1 and 5 with most predictions between 0 and 1.5 (Figure 12.3).

The risk of hospital admission for the different units of linear predictor are shown in Table 12.11. As the continuous predictors were centred at their means the risk per linear unit can be interpreted as the percentage risk above that of a HF patient with an average level of haemoglobin, eGFR, systolic BP, cholesterol and BMI and in the reference category of all other predictors. Due to the absence of intercept inherent in conditional

logistic regression the actual risk cannot be determined from the model. Risk groups are instead constructed based on proportion of risk above the baseline risk.

C-index for the ROC for the model was 68% (Figure 12.4). This can be interpreted as a 68% chance that in a randomly selected case and control, the case would be given a higher predicted risk. Taking the centering and quadratic extensions into account the linear predictor for an individual patient would be calculated as set out in Table 12.13.

12.8 Chapter summary

There was significant improvement in model fit for both outcomes when comorbidity status measures were exchanged for severity measures defined by either physiological or drug measures with the latter providing the best overall fit. Comorbidity severity change measures improved mortality model fit when replacing status measures for all comorbidities. This was also true for hospital admission models when including DM drug severity change or eGFR change over their respective status measures. Comorbidity severity change measures improved mortality model fit over the severity measures for DM and CKD but this was only true for CKD in the hospital admission models. The best fitting models used a combination of drug severity and change measures for comorbid DM and COPD and eGFR severity and change for CKD. The C-index for mortality (74%) was better than for hospital admission (68%). These results will be discussed in detail in the next Chapter 13.

Tables

Table 12.1a Mortality models with comorbid DM drug severity measures

Comorbidity measure model *110505 observations	df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a Diabetes Status	18	0.1991	58201.80	58384.44	-29081.90		0.728
b Diabetes Severity	20	0.1998	58154.34	58356.21	-29056.17	$p<0.001^{ab}$	0.729
c Diabetes Severity change	19	0.2001	58128.17	58320.43	-29044.08	$p=1.0^{bc}/p<0.001^{ac}$	0.729
*102990 observations (in order to compare with diabetes physiological severity models)							
a Diabetes Status	18	0.2033	54309.56	54490.86	-27135.78		0.730
b Diabetes Severity	20	0.2040	54254.89	54455.29	-27106.45	$p<0.001^{ab}$	0.730
c Diabetes Severity change	19	0.2042	54248.89	54439.74	-27104.45	$p=1.0^{bc}/ p<0.001^{ac}$	0.731

All models adjusted for age, gender, beta-blocker, ACEi or ARB, diuretics, systolic BP (systolic^2), haemoglobin (Hb^2), BMI (BMI^2), cholesterol, smoking, alcohol, COPD, eGFR. Diabetes status (clinical code or prescription), severity (drug categories in 4-month time-window prior to death; none, any oral, insulin only) and change by recent drug category change over prior 1 year (same or increase, decrease). ab; LR test of model a nested in b; ac; LR test of model a nested in c, bc; LR test of model c nested in b.

Table 12.1b Mortality models with comorbid DM physiological severity measures

Comorbidity measure model *102990 observations	df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a Diabetes Status	18	0.2033	54309.56	54490.86	-27135.78	$p<0.001^{ac}$	0.730
b Diabetes Severity	23	0.2035	54303.68	54532.71	-27127.84	$p<0.01^{ab}$	0.730
c Diabetes Severity change	20	0.2036	54293.51	54493.90	-27125.75	$p=1.00^{cb}$	0.730

All models adjusted for COPD, renal dysfunction (eGFR<60) age, gender, beta-blocker, ACEi or ARB, diuretics, systolic BP (systolic^2), haemoglobin (Hb^2), BMI (BMI^2), cholesterol, smoking, alcohol . Diabetes status (clinical code or prescription), severity (HbA1c categories; most recent within 3 years of death) and change by recent HbA1c change over 1 year prior to death. ac; LR test of model a nested in c; ab; LR test of model a nested in b, cb; LR test of model c nested in b.

12.2a Mortality models with comorbid COPD drug severity measures

Comorbidity measure model *110505 observations	Df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a COPD Status	18	0.199	58201.80	58384.43	-29081.90		0.7284
b COPD Severity	20	0.201	58013.08	58214.95	-28985.54	$p<0.001^{ab}$	0.7301
c COPD Severity change	20	0.200	58083.63	58285.50	-29020.82	$p<0.001^{ac}$	0.7297
*90610 observations (in order to compare to COPD physiological severity models)							
a COPD Status	18	0.211	47965.17	48144.05	-23963.59		0.7331
b COPD Severity	20	0.212	47928.35	48126.05	-23943.17	$p<0.001^{ab}$	0.7337
c COPD Severity change	20	0.212	47935.65	48133.35	-23946.83	$p<0.001^{ac}$	0.7297

All models adjusted for COPD, renal dysfunction (eGFR<60), age, gender, beta-blocker, ACEi or ARB, diuretics, systolic BP (systolic^2), haemoglobin (Hb^2), BMI (BMI^2), cholesterol, smoking, alcohol , COPD status (clinical code and prescription), severity (drug categories in 4-month time-window prior to death; no oral steroids or oxygen, oral steroids but no oxygen, oxygen) and severity change by recent drug category change over 1-year (new onto oral steroids or oxygen) prior to death. ab; LR test of model a nested in b, ac; LR test of model a nested in c.

Table 12.2b Mortality models with comorbid COPD physiological severity measures

Comorbidity measure model *90610 observations	Df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a COPD Status	18	0.211	47965.17	48144.05	-23963.59	$p=0.271^{ac}$	0.7331
b COPD Severity	21	0.212	47930.59	48137.71	-23943.30	$p<0.001^{ab}$	0.7334
c COPD Severity change	20	0.211	47966.56	48164.26	-23962.28	$p<0.001^{cb}$	0.7332

All models adjusted for COPD, renal dysfunction (eGFR<60), age, gender, beta-blocker, ACEi or ARB, diuretics, systolic BP (systolic^2), haemoglobin (Hb^2), BMI (BMI^2), cholesterol, smoking, alcohol. COPD status (clinical code and prescription), severity (FEV1 categories; most recent within 3 years of death) and change by recent FEV1 change over 1 year prior to death. ac; LR test of model a nested in c. ab; LR test of model a nested in b, cb; LR test of model c nested in b.

Table 12.3 Mortality models with comorbid CKD

Comorbidity measure model *90643 observations	df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a Renal Status	17	0.1957	48692.70	48862.17	-24328.35		0.724
b Renal Severity	18	0.2078	47960.74	48139.62	-23961.37	$p<0.001^{ab}$	0.733
c Renal severity change	20	0.2012	48362.38	48550.67	-24161.19	$p=1.0^{bc}/<0.001^{ac}$	0.727
d Model b +c	21	0.2140	47691.71	47803.14	-23776.01	$p<0.001^{bd/cd}$	0.734

All models adjusted for recent age, gender, beta-blocker, ACEi or ARB, diuretics, systolic BP (systolic^2), haemoglobin (Hb^2), BMI (BMI^2), cholesterol, smoking, alcohol, COPD, DM. Renal status (eGFR<60), severity (eGFR categories using most recent eGFR within maximum of 3 years prior to death) and change by categories of eGFR change over 1 year prior to death (same, increase, medium decrease, severe decrease). ab; LR test of model a nested in b, bc; LR test of model b nested in c.

Table 12.4 Mortality models with comorbidities

Comorbidity measure model *90643 observations	df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a Status	18	0.196	48692.70	48862.17	-24328.35		0.724
b Severity	23	0.211	47755.21	47971.75	-23854.60	$p<0.001^{ab}$	0.733
c Severity change	24	0.204	48185.14	48411.09	-24068.57	$p=1^{bc}/<0.001^{ac}$	0.729
d Model c + CKD severity	26	0.217	47431.28	47676.05	-23689.64	$p<0.001^{bc/bd}$	0.736

Status model: COPD (clinical code and prescription) Diabetes (Clinical code or prescription) Renal dysfunction (recent eGFR <60).
Severity model (drugs): Diabetes drug severity measures (none, any oral hypoglycaemic, insulin only) and COPD drug severity measures (no oral steroids or oxygen, oral steroids but no oxygen, oxygen) based on at least one prescription within a 4-month time window before death). Renal dysfunction based on most recent eGFR value within 3-years.
Severity change model (drugs): Diabetes and COPD based on drug severity measures (at least one prescription within a 4 month time window before death compared to at least one prescription in a 4-month time window up to 1-year prior to death). Renal dysfunction change measure based on eGFR (measured as most recent within 3-years and a previous measure between 6-months and 3-years prior to measure 1-year change). All models adjusted for age, gender, beta-blocker, ACEi or ARB, diuretics, systolic BP (systolic^2), haemoglobin (Hb^2), BMI (BMI^2), cholesterol, smoking, alcohol. Severity change model further adjusted for recent eGFR.

Table 12.5 Linear predictor for HF comorbidity mortality model

Linear predictor score	No. observations	No. cases	Risk (% cases)
≤-1	371	5	1
>-1 to 0	7579	278	4
>0 to 1	28531	2,815	10
>1 to 2	32573	7,585	23
>2 to 3	16558	7,200	43
>3 to 4	4279	2,663	62
>4 to 5	679	482	71
>5	73	53	73

Table 12.6 Comorbidity HF mortality risk score

	Score =	Centred value	Increments of the continuous variable		Score
	Age in years	-77	÷5	x	0.24
+	Male				0.13
+	COPD no steroids or prescribed oxygen				0.12
+	COPD on oral steroids but no prescribed oxygen				0.60
+	COPD on prescribed oxygen				1.13
+	Diabetes and stable diabetes drugs over previous year				0.20
+	Diabetes increase in drug stage over previous year				0.27
+	Diabetes reduced drugs over previous year				0.82
+	eGFR (any increase) over previous year				0.21
+	eGFR (6mls to 15mls decrease) over previous year				0.16
+	eGFR (>15mls decrease) over previous year				0.57
-	Current eGFR (ml/min/kg^2)	-58	÷5	x	0.39
+	Current eGFR (ml/min/kg^2)	-58	÷5	x	0.01
+	No beta- blocker				0.31
+	No ACEi or ARB				0.59
+	Diuretics				0.16
-	Systolic BP (mm/hg)	-131	÷5	X	0.08
+	Systolic BP (mm/hg)	-131	÷5	X	0.01
-	Cholesterol (per mmol/L)	-5		X	0.05
-	Haemoglobin	-13		X	0.12
+	Haemoglobin	-13		x	0.23
-	BMI (per kg/m^2)	-28		x	0.03
+	BMI (per kg/m^2)	-28		x	0.002
+	Ex-smoker				0.05
+	Current smoker				0.36
-	Alcohol – current drinker				0.05

Table 12.7 Hospital admission models with DM drug severity measure

Comorbidity measure model *53159 observations	df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a Diabetes Status	18	0.1636	30801.77	30970.50	-15381.89		0.6796
b Diabetes Severity	20	0.1647	30766.47	30952.98	-15362.24	$p<0.001^{ab}$	0.6798
c Diabetes Severity change	20	0.1638	30799.64	30986.15	-15378.82	$p=0.05^{ac}$	0.6801

All models adjusted for recent age, gender, COPD, renal dysfunction (eGFR<60), prior hospital admission, systolic bp (BP^2), beta-blocker, ACEi or ARB, diuretics, haemoglobin (Hb^2), BMI, cholesterol, smoking, alcohol. Diabetes status (clinical code or prescription), severity (drug categories in 4-month time-window prior to admission; none, any oral, insulin only) and change by recent drug category change over prior 6 months. ab; LR test of model a nested in b, ac; LR test of model a nested in c.

Table 12.8 Hospital admission models with COPD drug severity measures

Comorbidity measure model *53159 observations	Df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a COPD Status	18	0.1636	30801.77	30970.51	-15381.89	$p=0.585^{ac}$	0.6796
b COPD Severity	20	0.1647	30766.67	30953.17	-15362.33	$p<0.001^{ab}$	0.6804
c COPD Severity change	19	0.1636	30803.47	30981.09	-15381.74	$p<0.001^{bc}$	0.6796

All models adjusted for recent age, gender, diabetes, renal dysfunction (eGFR<60), prior hospital admission, systolic bp (BP^2), beta- blocker, ACEi or ARB, diuretics, haemoglobin (Hb^2), BMI, cholesterol, smoking, alcohol. . COPD status (clinical code and prescription), severity (drug categories in 4-month time-window prior to admission; no oral steroids or oxygen, oral steroids but no oxygen, oxygen) and severity change by recent drug category change over prior 6 months (new onto oral steroids or oxygen). ab; LR test of model a nested in b, ac; LR test of model a nested in c. bc; LR test of model c nested in b

Table 12.9 Hospital admission models with comorbid CKD

Comorbidity measure model *33907 observations	df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
a Renal Status	18	0.1633	19923.04	20083.23	-9942.52		0.676
b Renal Severity	23	0.1671	19841.52	20043.87	-9896.76	$p<0.001^{ab}$	0.678
c Renal Severity change	21	0.1660	19861.42	20038.48	-9909.71	$p<0.001^{ac/bc}$	0.678
d Model c +renal severity	26	0.1694	19793.34	20020.99	-9869.67	$p<0.001^{bd}$	0.680

All models adjusted for Age, gender, diabetes, COPD, prior hospital admission, beta-blocker, ACEi or ARB, diuretics, BP, haemoglobin, BMI, cholesterol, smoking, alcohol. Renal status (eGFR <60), severity by 7 eGFR categories and change by recent % change over prior 6 months. ab; LR test of model a nested in b, bc; LR test of model b nested in c.

Table 12.10 Hospital admission model: comorbidities

Comorbidity measure model *33907 observations	df	Pseudo R^2	AIC	BIC	Log likelihood	LR test	ROC
Status	18	0.163	19923.03	20083.23	-9942.52		0.676
Severity	27	0.169	19798.92	20035.00	-9871.46	$p<0.001^{ab}$	0.680
Severity change	24	0.166	19863.02	20065.37	-9907.51	$p<0.001^{ac/bc}$	0.677
Model C + renal severity	29	0.170	19795.29	20048.23	-9867.65	$p=0.02^{bd}/p<0.001^{cd}$	0.680

Status measures: COPD (clinical code) Diabetes (Clinical code or prescription) CKD (recent eGFR <60).
Severity measures: Diabetes and COPD based on drug severity measure (at least one prescription within a 4 month time window before admission). CKD based on eGFR within 6months.
Severity change: Diabetes and COPD based on drug severity measures (at least one prescription within a 4 month time window before admission compared to at least one prescription in a 4-month time window up to 6 months prior to admission). CKD based on eGFR (measured within 6months and a previous measure between 1-months and 1-year prior to measure 6 month change).
All models adjusted for age, gender, prior hospital admission, beta-blocker, ACEi or ARB, diuretics, systolic BP (BP^2), haemoglobin (Hb^2), BMI, cholesterol, smoking, alcohol. Severity change model further adjusted for recent eGFR stage.

Table 12.11 Linear predictor for HF comorbidity hospital admission model

Linear predictor score	No. observations	No. cases	Risk (% cases)
≤-0.5	798	121	15
>-0.5 to 0	10646	2,109	20
>0 to 0.5	10610	2,833	27
>0.5 to 1	5413	1,980	37
>1 to 1.5	2231	1,025	46
>1.5 to 2	1325	794	60
>2 to 2.5	1397	952	66
>2.5 to 3	886	646	68
>3 to 3.5	372	314	73
>3.5 to 4	147	133	84
>4 to 4.5	82	73	90

Table 12.12 Comorbidity HF hospital admission risk score

	Score =	Centred value	Increments of the continuous variable		Score
	Age	-77	÷5	x	0.03
+	Male				0.16
+	Prior admission in past 3 months				2.09
+	Prior admission in past 3 months to 6 months				0.89
+	Prior admission in past 6 months to 1 year				0.38
+	COPD not on steroids or oxygen				0.21
+	COPD on oral steroids but no oxygen				0.64
+	COPD and on oxygen				0.86
-	*Diabetes on no drugs				0.001
+	Diabetes on any oral hypoglycaemic				0.25
+	Diabetes on insulin only				0.32
+	*eGFR >105				0.19
+	*eGFR 90-105				0.10
+	*eGFR 45-59				0.07
+	*eGFR 30-44				0.06
+	eGFR 15-29				-0.04
+	eGFR <15				0.008
+	eGFR (any % increase over past 6 months)				0.10
+	eGFR (6-25% decrease over past 6 months)				0.12
+	eGFR (>25% decrease over past 6 months)				0.34
+	No beta-blocker				0.09
+	No ACEi or ARB				0.35
-	*Diuretics				0.03

-	Systolic BP (mm/hg)	-136		0.02
+	Systolic BP^2 (mm/hg)	-136	x	0.004
-	Cholesterol (mmol/L)	-5	x	-.04
-	Haemoglobin	-13	x	0.13
+	$Haemoglobin^2$	-13	x	0.03
-	BMI (kg/m^2)	-28	x	0.008
+	Current smoker			0.11
-	*Alcohol – current drinker			0.02

Figures

Figure 12.1 Histogram of linear predictor from the comorbidity HF mortality model

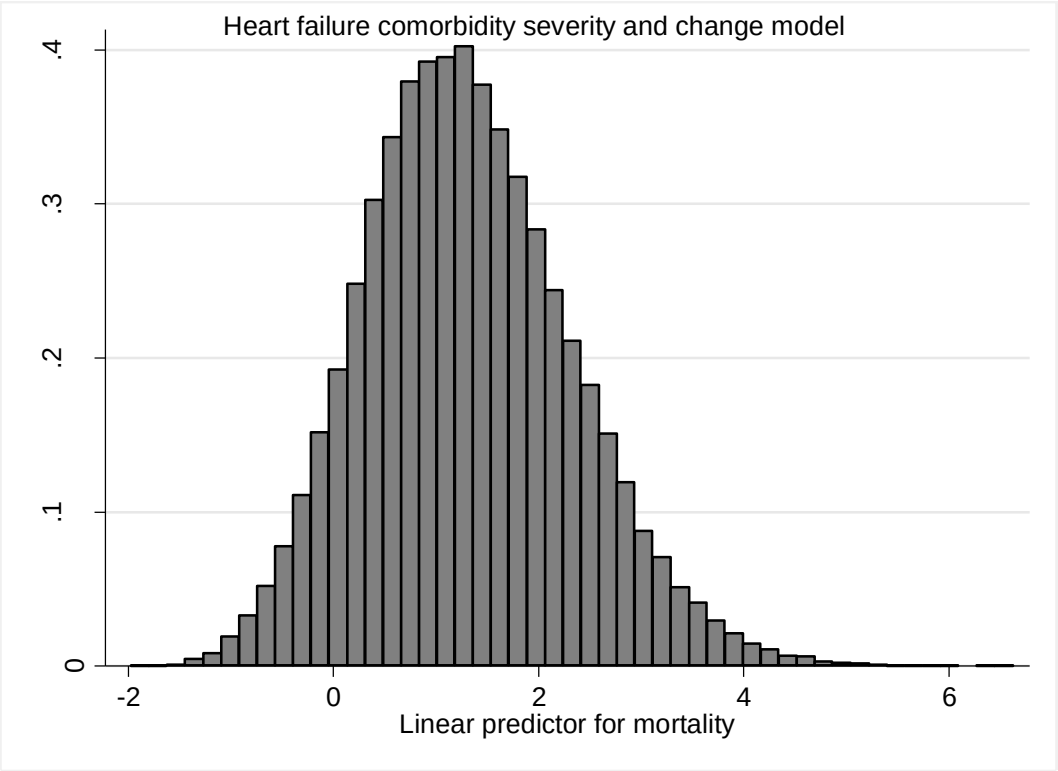


Figure 12.2 Receiver operator curve for HF comorbidity severity and change model.

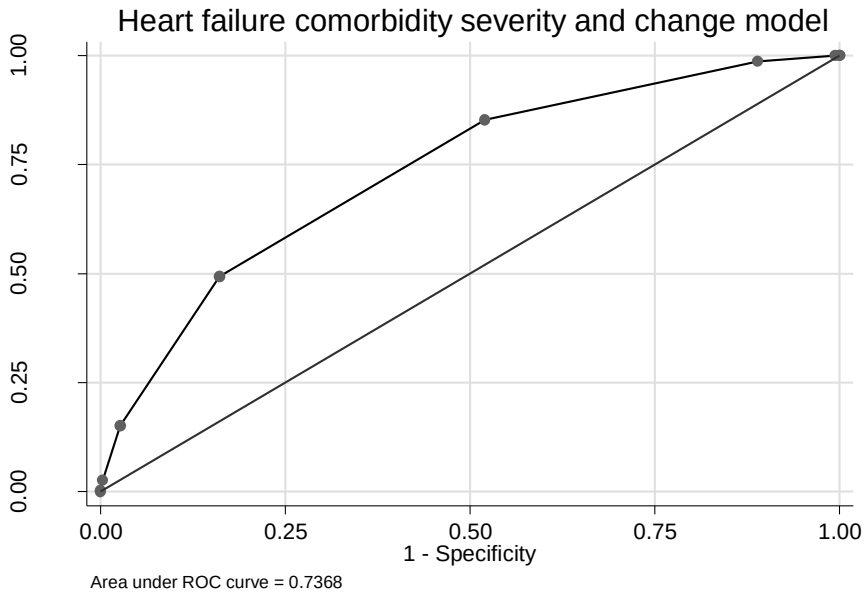


Figure 12.3 Histogram of linear predictor from the comorbidity HF hospital admission model

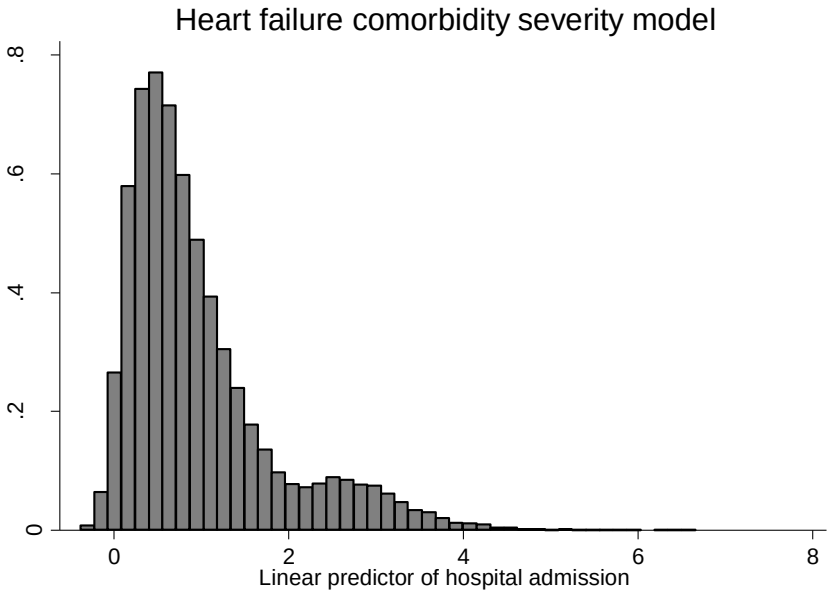
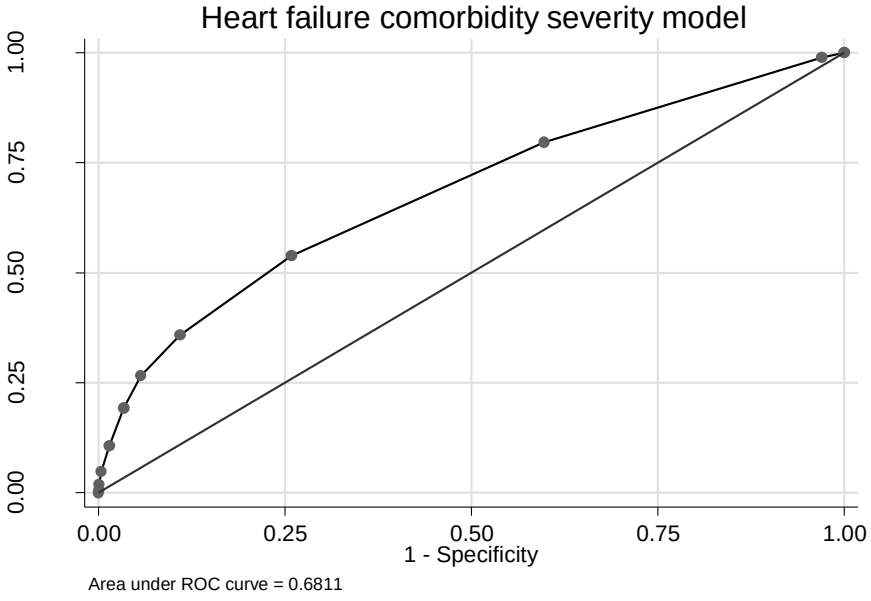


Figure 12.4 Receiver operator curve for HF comorbidity severity and change model.



Chapter 13 CPRD Discussion

This discussion chapter will first summarise the findings from the six CPRD results chapters (Chapters 7 to 12). Further discussion will then be organised via the research questions set out in the methods chapter (Chapter 6) and will finish with a comprehensive consideration of the strengths and limitations of the approaches used with a specific focus on chance, bias and confounding.

13.1 Summary of findings

In an unselected general population incident cohort of HF patients the comorbidities of DM, COPD and CKD were common and associated with the outcomes of mortality and hospital admission. Within these comorbid groups risk ranged by the severity of the comorbid disease during follow-up, which was defined by drug prescriptions or physiological indicators. For each comorbid disease, severity was worse in those experiencing hospital admission or death. Change in the severity of comorbid disease during the HF life course was also common, with a greater prevalence of this change occurring in the imminent time-period prior to the two outcomes.

The strengths of association were similar across all comorbid diseases and for both outcomes with some evidence of stronger associations for COPD than CKD. There was evidence of weaker associations for incident diabetes (occurring after HF diagnosis) than for prevalent diabetes (occurring before HF diagnosis). This association was in the opposite direction for COPD but the difference between prevalent and incident COPD was not significant. For all comorbid diseases the risk associated with both outcomes was significantly differentiated by severity groups defined by drug prescriptions or physiological indicators. There was also significant differentiation in the associations between categories of severity change for each of the comorbidities and mortality. Whilst there was a trend of increasing risk with worsening DM and COPD severity

for hospital admission the categories of change were similar. There was however significant differentiation in the categories of change for CKD.

There was a significant statistical and biological interaction between COPD and CKD with a lower than expected risk estimate for mortality when the two were experienced together. There was a biological interaction between DM and CKD in the opposite direction with a higher than expected combined risk estimate for mortality. When the HF group was stratified by younger and older age there were significant statistical interactions for all three comorbidities with mortality demonstrating a less than multiplicative relationship as a product of the lower relative risk in older age groups. Biological interaction was only present for COPD and age and CKD and age for mortality. There was a lower than expected risk associated with COPD in older age. The association of CKD in older age groups opposed the statistical interaction and demonstrated a higher than expected risk of mortality. Other biological interaction was found for CKD and gender with males having a higher risk of hospital admission and death than females. Whilst there was some evidence of statistical interaction between male gender and DM for hospital admission and between male gender and COPD for mortality this was no evidence of biological interaction.

When the contributions of the different comorbidity measures to prognostic models were tested, there were differences for mortality and hospital admission. For mortality the fit of the models improved when the status measures were exchanged for either the severity or the change measures. This improvement was similar for the comorbid DM severity or change measures by drug and physiological indicators and COPD by drug indicators. When defined by physiological indicators there was better model fit for the COPD severity than change measures which in turn were similar to the status measure. For admission all comorbidity drug severity measures were a better fit than status measures but the change measures did not improve model fit. The exception was CKD where there was an incremental improvement of model fit for both mortality and admission with the replacement of CKD status for severity measures and then the replacement of severity for change measures. For DM and COPD the models using drug severity measures had better overall fit than

those using physiological measures. Overall discrimination was better for the mortality models than the hospital admission models.

13.2 Prevalence of DM, COPD and CKD, severity and change

The prevalence of the three comorbid diseases ranged from 14% for COPD to 24% for DM to 55% for CKD. The largest proportion of DM and COPD comorbidity developed prior to the onset of HF. The prevalence of COPD was lower than in prior studies, which in the systematic review ranged from 17-35% (mean 18%)(239). The review studies were mainly conducted in specialist settings which are more likely to have employed objective measures of detection for COPD such as spirometry, which in turn is known to yield higher prevalence(342). In this study COPD clinical codes were validated by prescriptions but this does not negate undiagnosed COPD in the community setting(343). The prevalence of DM and CKD which have more chance of detection using routinely collected physiological measures were comparable to the mid-range of the review studies (13-47% for DM and 39-79% for CKD). The review studies included mostly hospital patients which tended to have more severe HF patients with a higher prevalence of comorbidity compared to community settings(132).The prevalence of all comorbid diseases was slightly lower in the hospital admission linked sub-sample that had a shorter average follow-up.

In the total mortality cohort, of the COPD patients with FEV_1 data (<50% of the HF-COPD sample), 41% were in the most severe two categories (FEV1 ≤49%). This was comparable to the 49% prevalence of severe COPD disease in a prior but smaller general population HF study(343). Drug severity was categorised for all COPD patients and the most severe groups (oral steroids or prescribed oxygen) had a prevalence of 30%. Most of the DM-HF group had a HbA1c measure (90%) and 19% of these were in the highest two HbA1c categories (HbA1c >8.5%) and 5% in the lowest (<5.5%). Of the DM group, 28% were on insulin which was lower than what has been found in hospital settings (39%)(293). From the 90% of the HF sample who had an eGFR measure, 8% had an eGFR in the lowest two categories (≤29 mls/min/m2) which was comparable to the 6% found in prior general population studies included in the systematic review(239). These prevalence rates were similar in the HF-HA sub sample. Approximately 15-20% of patients within each comorbid group

had experienced a worsening of the severity of the comorbid disease at a given time during follow-up (prior to their match date). Whilst previous evidence on comorbidity severity change in HF has focused on acute renal change in hospital settings the prevalence of this change was similar to that in the CPRD analyses (approximately 22%). This combined evidence shows that the CPRD sample used in this study is comparable to samples in previous studies and whilst the prevalence and severity of comorbidities may be less in the community setting of the general practice HF population than in hospital settings, the figures remain high with a large proportion of HF patients with severe and worsening comorbid disease.

13.3 HF comorbidity associations with mortality and hospital admission

13.3.1 Outcome risk

In this general practice population of incident HF, 23% of patients died within one year of diagnosis and 46% within five years. Following the initial year after diagnosis, HF patients died at an average rate of 10% per year. A third of the HF-HA sub sample had a hospital admission within one month of HF diagnosis and 65% within a year. These figures are similar to other age matched general practice populations and demonstrate the high mortality and morbidity associated with incident HF in the general practice population. Five year mortality in similar HF populations was previously 41% in Rotterdam(322) and 39% in Denmark(352). A more recent study of prevalent European HF outpatients revealed an annual mortality rate of 6% which was less than the 10% found in the present study. The European study(131) had a younger profile with mean age 66 years compared to the 78 years in this current study and included a lower percentage of incident HF patients. In an incident sample of non-selected HF in the outpatient setting in Olmsted County, 74% were hospitalised within a mean of 1.7 years(464) comparable to the 65% within 1 year found in this analysis.

13.3.2 Comorbidity associations

The three comorbidities investigated in this study were strongly and independently associated with all-cause mortality and first hospital admission following HF diagnosis. The adjusted strength of associations were similar across the three comorbidities and for both outcomes. The increase in risk in the comorbid groups relative to the non-comorbid groups ranged for mortality from 22% for CKD to 28% for DM to 35% for COPD

and for hospital admission from 15% for CKD to 25% for DM to 32% for COPD. These associations were comparable to those in the previous meta-analyses for DM (34%) and COPD (39%) for mortality(239). Relative risk increase for CKD was higher in the prior meta-analyses at 52%. The review studies were hospital focused with inclusion of more severe or acute renal disease during admission, which may in turn reflect HF severity status and account for some of this difference compared to the more chronic renal decline in the general HF population.

The relative risk associated with an exposure is determined by the absolute risk generated by the exposure and the baseline risk of the non-exposed population to which it is compared. HF is a serious chronic disease which means that the baseline risk of death and hospital admission within the population is high. Mortality risk in hospital based HF samples has been found to be higher than in community HF samples(465) meaning that relative exposure effects may be smaller in the hospital setting. This has been found previously for individual comorbidities where their relative exposure effects have been found to be greater in community HF samples reflecting the lower baseline risk(132). The importance of the similarity of associations in this general practice population of incident HF compared to the systematic review findings, which were predominantly based in hospital settings, is that it demonstrates the higher baseline risk in an incident cohort of HF patients reflective of the mixed hospital and community based systematic review sample. The comorbid estimates in the present analyses were between those of a prevalent HF community sample and a post hospital admission sample previously reported(132).

13.3.1 Comorbidity Life Course

Previous evidence that uses baseline exposure to predict future outcomes usually includes prevalent comorbidity measures. This may be comorbidity that occurs prior to HF if an incident HF study is conducted or comorbidity that has occurred at any time prior to the study baseline in a prevalent HF study. This is the first study to investigate the relative effect of comorbidity that develops prior to HF compared to that which develops after HF onset for DM and COPD. This was conducted to understand the interrelations between chronic disease comorbidity and outcomes in HF.

Comorbidity may be a cause or a consequence of an index disease or unrelated and this may influence its association with outcomes. In this analysis, prevalent DM had a much stronger association than incident DM where the association was reduced for mortality and abolished for hospital admission. This is likely to be a result of prevalent DM being more severe with a longer duration, resulting in greater cardiovascular complications and more severe HF than incident DM(354,355,360). However, in the baseline descriptive studies, the difference in the prevalence of DM (occurring prior to HF) between survivors and non-survivors was minimal at one year follow-up (0.1%) and increased to 3.6% at >5years follow-up. This indicates that it might not just be the severity of the DM (whose onset may have been many years before HF), but the duration of the DM concurrent with HF. This infers possible interaction between the DM and HF severity which concurs with prior studies that have found DM to have stronger associations in those with coronary heart disease aetiology(294) and LVSD(295). Another potential explanation is that other competing factors driving mortality in the first year after HF diagnosis such as more severe HF disease at onset. The difference in the associations between the prevalent and incident comorbid COPD groups and outcomes were not significantly different, but were stronger for the incident COPD group occurring after HF. This difference for mortality was partly explained by worse deprivation in the COPD group.

13.4 HF comorbidity severity and change associations with mortality and hospital admission

13.4.1 DM drug severity and change

The HF-DM group was stratified into *a priori* severity groups using drug indicators. These were then compared to the HF group without DM for the two outcomes. With the exception of the 'no medications' group for hospital admission, all severity categories were significantly and independently associated with both outcomes. For mortality there was no difference in risk between the oral hypoglycaemic group with or without added insulin, contrary to prior evidence from the general population with type 2 diabetes(430). Risk increased from this combined group, to the 'no medications group' to the highest risk in the insulin only group.

The increased risk in the insulin group has been found in prior hospital studies(179,293). The insulin group are likely to reflect longer duration of DM with more severe metabolic disturbances and increased severity of disease(149,430). These findings show that the insulin group are an important target for risk reduction within the general practice HF population.

For the hospital admission outcome, there was no association with the 'no medications' group compared to the HF group without DM. Risk was significant and increased from the oral hypoglycaemic group to the insulin group. The difference in associations for the no medications group across both outcomes may be a reflection of the time-dependent measurement. No medications in the four months prior to death may be an indication of HF severity where drugs are discontinued in end stage disease(466). When measured earlier in the HF trajectory prior to admission, no medications are more likely to indicate less severe comorbid disease.

This latter hypothesis is supported by the DM drug severity change measures. For both mortality and hospital admission the strongest associations were with the DM group with a recent reduction in drug category. This group carries the highest risk for both mortality and admission and constituted predominantly those with DM prescriptions recently discontinued (81%). The weaker strength of association for hospital admissions than mortality may be reflective of clinical intervention in response to the DM instability and admission avoidance. There was also significant and independent associations between the no drug change DM group and both outcomes and the increased DM drug category group with the latter having stronger associations.

There was significant stratification of effects when the DM-HF group was categorised by drug severity and change measures and for hospital admission by severity. Whilst the drug change measures were not significantly stratified for admission there was a clear trend in increasing risk from the no change group to the increased group to the decreased category group. This is the first study to identify DM drug severity and recent change in routine clinical practice. Instability in diabetes disease indicated by drug category change (better or worse) is as important and feasible indicator of outcomes in the general population, given that these measures are automatically and electronically recorded.

13.4.2 DM physiological severity and change

Despite the common co-existence of DM and HF, previous studies examining the association between HbA1c and outcomes in this population have been limited with only two small studies in the general practice population focused on mortality. The findings provide new evidence that categories of HbA1c severity in DM are independently and significantly associated with mortality and hospital admissions in the general practice HF population. There was a significant and linear trend of increasing risk relative to the non DM HF group from the lowest risk HbA1c category to the highest category of HbA1c for both outcomes. This reached a relative increase in risk of between 45 to 49% for the top two HbA1c categories (HbA1c >8.6%) for mortality and 64% in the top category for hospital admission (HbA1c >9.5%). These groups were significantly different from the lowest risk HbA1c category for mortality and lowest 3 risk categories for hospital admission.

A U shape relationship was identified between HbA1c and mortality but not hospital admission. The lowest risk HbA1c category for mortality reflected the guideline driven target of 6.5-7.5%(418,419) but the lowest risk category for hospital admission was lower at 5.5-6.4%. There was an increase in risk below HbA1c 6.5-7.5% for the mortality outcome but there was no association with hospital admissions in the lowest category of HbA1c (<5.5%). This U shape relationship with mortality has been found in a smaller sample of mostly male veterans(288) that identified higher risk of mortality at levels of below 6.4% and above 9% similar to the present findings, but did not find a U shape relationship with hospital admissions and reported a non-significant relationship of increasing risk between increasing HbA1c quintiles and hospital admission. The thesis findings add to this existing evidence and support a more complex relationship between HbA1c and outcomes in HF patients with diabetes. By studying a large sample of general practice HF population with diabetes, the findings demonstrate that patients in the lower and higher HbA1C categories have a higher mortality than patients with modest glycaemic control (6.5% < HbA1C ≤ 7.5%).

The findings also show that the HbA1c category associated with the lowest risk was higher for mortality than for admissions. This reflects other studies which have found a shift to the right in the U curve in more

advanced disease demonstrating a lower threshold for hypoglycaemia(428). HbA1c in this analysis was the most recent measure prior to death. The higher threshold for low HbA1c for the hospital admission outcome may be in part due to less severe HF. A similar threshold to our study on hospital admissions was found in a less severe HF group undergoing cardiac revascularisation(429). This combined evidence implies that the target for HbA1c control in HF with diabetes needs to be guided by the severity of the HF population with higher targets in more severe disease and with more scope for intensive glucose lowering therapy in less severe HF.

Short term changes in HbA1c have not previously been investigated in HF populations. In this general practice HF population, recent HbA1c change was independently and significantly associated with both outcomes. Importantly whilst the HF-DM group with an increase in HbA1c (>1%) yielded an increase in relative risk of approximately 27%, a decrease in HbA1c (>1%) indicated bigger increases in relative risk (49% for mortality and 34% for hospital admission). This decrease in HbA1c group had a significantly higher risk than the other two change groups for mortality and showed a trend of increased risk for admission. Both change groups were associated with a higher risk than the no change HF-DM group for both outcomes, compared to non-comorbid group. The higher risk in decreased HbA1c may result from a lower HbA1c end point (increasing the associated risk for mortality) or may reflect a higher baseline start HbA1c point (higher risk for both outcomes). Adjustment was not possible for baseline HbA1c which would have helped to further understand this association as HbA1c levels were only available for the DM group. However these findings were supported by a previous study on type 2 diabetes(431). The associations were again stronger in the decreased than the increased HbA1c group and independent of the diabetes reference group, but adjustment for HbA1c level was not possible due to small numbers with HbA1c change.

13.4.3 COPD drug severity and change

This is the first study to use drug based measures of COPD severity to determine risk of poor outcomes in HF. Three different COPD drug severity frameworks were tested for their associations with mortality and hospital admission. The findings were that in the general practice population of incident HF that the level of guideline

driven ‘step-up’ of prescribed inhalers for COPD(410) was not associated with the risk of mortality. The risk estimates for all levels of inhaler therapy were small and only significant for triple therapy compared to no COPD HF. Inhaler therapy predominately is used for symptom control and the poor association with mortality likely reflects the lack of dose-response relationship between all classes of bronchodilators and FEV_1(467-469). However for both outcomes the prescription of oral steroids or oxygen therapy was strongly and independently associated with mortality and hospital admission.

The prescription of oral steroids is known to increase mortality risk in COPD(470). One mechanism postulated relates to weakened muscle strength in prolonged therapy leading to poor pulmonary function(471). Oral steroid use in COPD has also been found to adversely affect the development and progression of other comorbidities including those with cardiovascular aetiologies such as hypertension and diabetes(472) which in HF could contribute to more severe disease. Drugs were measured in a four-month time window and so duration of prescribing could not be measured. However, short term prescribing is usually a result of acute exacerbations of COPD which are also a predictor of mortality(473). Oral steroids in this context are a likely pseudo-marker of more severe COPD disease.

When the COPD group was stratified by a recent increase or decrease in at least one drug severity category both strata were significantly associated with mortality, but these associations were similar to the COPD group overall. This is likely to result from the multiple categories of inhaler therapy which had weak associations with mortality. Using the second drug severity change classification, there were significant, independent and stratified associations for COPD defined by the ‘no new steroids or oxygen’, new prescription of steroids or oxygen with mortality. The new on oral steroids or oxygen group are likely to indicate acute exacerbation in severe COPD or end-stage COPD.

These associations differed for hospital admissions where ‘new oral steroids’ had similar association to the ‘no new steroids or oxygen’ group compared to the no COPD HF group. The prescription of oral steroids whilst indicating a COPD exacerbation may also serve to prevent some hospital admissions in this group. Newly

prescribed oxygen did show a stronger association with admission but this was non-significant and accounted only for a small number of observations (n=64, 0.5%). Most hospital admissions occurred soon after the HF diagnosis. Increasing breathlessness is a shared symptom in patients with HF and COPD(374) and is likely to contribute to the reason for admission following HF diagnosis. This may deter the prescription of oxygen in this group prior to hospital assessment.

13.4.4 COPD physiological severity and change

GOLD guidelines recommend the use of forced expiration volume measured in 1 second (FEV_1) to measure the severity of lung function in COPD(410). Whilst the recording of this measure in practice has historically been low it improved from 18% in 2003 to 62% in 2004 for COPD patients following guideline recommendations(411). The analyses spanned this timeframe and a 50% prevalence of COPD patients with at least one spirometry measure was shown. FEV_1 could not be used for the smaller HF-HA sub sample. Narrower measurement time-windows and shorter average follow-up meant that minimal HF-COPD patients had this measure recorded. There was no previous evidence on the association between lung function in patients with a recorded diagnosis of COPD and mortality in unselected general practice HF populations. A key question for prognosis was whether the risk associated with COPD would differ according to lung function. One prior study in a hospital HF population found that severe COPD (FEV1<49%) was significantly associated with mortality, but the associations for moderate severity were non-significant(296).

The key findings were that there were significant, stratified and independent associations between all COPD-HF GOLD severity categories and higher mortality which increased from the mild group (GOLD stage 1; FEV_1 ≥80% normal) to the severe group (GOLD stage 4; FEV_1 <30%). Of note was the higher risk association in the mildest group than the overall COPD group meaning that the recording of FEV_1 data was, in itself, associated with higher mortality risk. On further examination the FEV_1 group compared to the non-FEV_1 COPD group had a much worse risk profile. They were older, more male (65% versus 58%), more deprived (3% more in most deprived quintile), with a higher prevalence of diabetes and a higher percentage on oral steroids (4%). They had lower mean cholesterol, BMI and haemoglobin albeit there were 3% fewer smokers in

this group. This profile suggests that these risk factors are driving some of the clinical decisions to request spirometry with a focus on higher risk HF-COPD patients. Whilst the associations compared to no COPD HF were adjusted for these risk factors, the higher prevalence and worse risk profile in the FEV_1 COPD group is likely to indicate worse HF severity in this group and might partly explain the higher overall risk in this group compared to the group without FEV_1 measure.

This is the first study to investigate recent change in lung function and mortality in the general practice HF population. Worsening of a GOLD stage over a year yielded a 55% increase in relative risk from the COPD group with 'no change or improved GOLD stage', compared to the no COPD HF group. These two change group associations were independent and significant but not significantly different to each other. There was a trend of increasing risk from the COPD group with no FEV_1 change to the COPD groups with 5% or 10% increase in FEV_1 to the COPD groups with 5% or 10% decrease in FEV_1. Whilst these change groups were not significantly different to each other and may again, in part, indicate worsening HF, they imply a potential important clinical indicator of increased risk. More consistent and routine monitoring with spirometry is required.

13.4.5 CKD physiological severity and change

Previous evidence in the general population of HF has determined that increased severity and change in renal function is an important predictor of HF outcomes. However there was only one study identified through the systematic review that included community patients or new onset HF and only one study that investigated hospital admissions(301). Investigation into change in renal function has so far been limited to the outcome of mortality in hospital populations (with one study including readmissions(474)). These studies have used acute change in creatinine which is susceptible to fluctuations due to patient factors. One prior study did suggest that change in creatinine in a more stable state, 6 months after discharge, was a better predictor of mortality than when measured during admission(156). The thesis findings adds important new evidence on renal function change in the general practice HF population for the outcomes of mortality and first hospital admission. eGFR calculated via the modification in renal disease equation (MDRD) was used to measure this

change which adjusts for age, sex, race and body size and is a more accurate measure in stable populations(398,399).

Two definitions of renal change were used; a relative and an absolute change measure. Previous studies have considered the use of eGFR change as being less dependent on baseline renal function than creatinine change(439). The latter has an exponential relationship with eGFR and for the same change in creatinine there will be a small or large change in eGFR depending on a lower or higher baseline eGFR respectively. This means that using creatinine change to define worsening renal function, those with lower baseline eGFR are more likely to experience this change. The CPRD analyses, showed that whilst lower baseline (start-point) eGFR (<60ml/min/m2) was significantly associated with severe renal decline using the percentage change measure, this switched to higher baseline eGFR using the absolute measure. This potentially reflects the linear rate of change of eGFR found at all baseline levels of eGFR(440). Those with higher baseline eGFR will have greater capacity for a severe absolute loss than those with poorer function and those at lower baseline eGFR are more likely to experience a given percentage loss given that this equates to less absolute loss than in higher baseline eGFR.

A U shape relationship between eGFR and mortality but not hospital admission was shown. The lowest risk category of eGFR for both outcomes was 60-89ml/min/m^2. Compared to this reference group there was an increase in risk of mortality, but not admissions above this category. This was most marked in the eGFR >105 group. The increased risk of mortality in higher eGFR in the general population has been found previously(434) which, like in this study, starts to increase above 90ml/min/m^2(437). This has been reported to reflect the inadequacies of the eGFR formula at low serum creatinine levels(437) or a false account of kidney function due to external causes such as loss of muscle mass in frailty or dilution of creatinine in severe HF(334). This may have been more marked in these analyses due to measurement of eGFR most recent to death in the case group and would also explain the lack of this association with earlier onset HF prior to admission. Risk also increased with every category below the reference group with clear stratification of risk estimates increasing to the highest level in the most severe group (<15ml/min/m2). Associations were similar

for falling eGFR from the reference category for hospital admissions also increased to the highest level in the most severe eGFR group. There was stratification of effects for the lowest two eGFR groups compared to all other eGFR categories. These associations were stronger than those in the prior meta-analysis for mortality(239) and for hospital admissions(475) and this may be indicative of the time-dependent measurement prior to an event having a stronger association with the event.

Recent change in eGFR showed significant independent associations with both outcomes. For mortality, compared to the minimal change group, those with any increase or moderate decrease in eGFR had similar associations using both the absolute and percentage change measures (range OR 1.14-1.25). The risk of mortality associated with increasing eGFR has been previously reported in hospital HF populations(314) and is suggested to be either a pseudo effect of reduced creatinine levels in frailty or a marker of recent and imminent worsening function. The effect estimates for the most severe change in renal function was higher for the percentage change measure than the absolute change measure which was a function of the lower baseline start-point eGFR in the percentage change group. Following adjustment for this measure the effect estimates for both the absolute and relative change measures were the same. These estimates were diminished when adjusted for the end eGFR measure. This demonstrates that both the severity of the change and the resultant eGFR and are important components of risk in the HF-CKD population and can be used individually or better together, to identify patients with worse prognosis.

eGFR change was also independently and significantly associated with first hospital admission following HF diagnosis. The associations were weaker than for mortality and non-significant for moderate decline or increasing eGFR using the absolute change measure. The two severe change groups had the strongest effects and these estimates were not reduced greatly by the adjustment of the start or end-point baseline eGFR or the end. The weaker influence of adjustment may indicate different mechanisms associated with eGFR change before the two outcomes. Most first hospital admissions occurred soon after HF diagnosis. eGFR change prior to admission may indicate HF instability and more acute renal change than the eGFR

change that occurs chronically over time before death. This earlier more acute change and its indication of potential instability may be important for admission in both higher and lower eGFR states.

These findings highlight the importance of the severity of CKD and change in the general practice HF population. Of particular note is the strong association between CKD change and first hospital admission following HF diagnosis that has not previously been investigated. It cannot be ruled out that the strength of associations with both outcomes might be, in part, related to the underlying haemodynamic status of the HF, particularly in the period of time following HF diagnosis and prior to admission or in deteriorating status prior to death(382). Adjustments could not be made for ejection fraction, NYHA or BNP, which are not routinely recorded in the CPRD. However, adjustment for a range of clinical and patient factors which are associated with more severe HF disease including diuretics and systolic blood pressure were made. The dose-response relationship found between categories of eGFR and outcomes in this study and the prior studies that have accounted for HF severity(156,475) all suggest an independent association between renal function and outcomes. eGFR is a well recorded routine measure in HF and can be easily used to target high risk groups in the general practice population for intervention. The difference in associations with CKD severity for both outcomes with a U shape identified for mortality, indicates the need to tailor this assessment according to the severity of the patient with caution given to the interpretation of higher eGFR levels in older frail patients with more severe disease.

13.5 Comorbidity interactions

13.5.1 Inter-relations

There were biological interactions between two of the HF comorbidity disease pairs and mortality; DM-CKD and COPD-CKD. The combination of DM and CKD had an association with mortality that yielded a higher than expected risk estimate. The synergy index, which measures the risk of two factors where there is interaction present as a proportion of the risk if there was no interaction present, was 29% additional risk relative to the pair without interaction. This finding is important for two reasons. Firstly, DM and CKD share pathophysiological pathways that may interact to intensify their progression(476) which increases the risk of

mortality. The intricate relationship between both diseases and HF is likely also to contribute to more severe HF disease. Secondly, this was the most prevalent disease combination in 10.3% of all HF patients, which indicates an important and high risk group for targeting with interventions.

The combination of COPD and CKD had a significant statistical and biological interaction which was in the opposite direction to DM and CKD. When these comorbid diseases were experienced together in HF they exerted a less than expected risk of mortality with a synergy index equating to a 30% reduction in risk relative to if there was no interaction. As previously discussed, both COPD and CKD create a complex interplay of pathological interactions with cardiovascular mechanisms, coupled with responsive processes to haemodynamic compromise in HF. Congestion in COPD and poor cardiac output in CKD can lead to worse physiological measures such as FEV_1 and eGFR which can blur the relationship between the comorbid severity and poor outcomes. The reduction in risk when the two comorbidities coexist supports the hypothesis that both diseases share an indication for HF severity. Prognosis assessment for patients with this comorbid disease combination would require close consideration of their HF disease severity.

13.5.2 HF Severity

Age was used in this study as a pseudo indicator of HF severity, an approach which has been used in prior HF studies and has indicated a differential relative effect of comorbidities across age groups through significant statistical interaction(139). Statistical interaction was found between all three comorbidities and older age which when considered together was less than multiplicative. However this was likely due to the lower baseline risk in the younger age group meaning that the relative risk associated with comorbidity was greater in these groups. When tested for biological interaction, only COPD and older age was significant in the same direction as the statistical interaction. There was also a significant interaction between CKD and age but this went in the opposite direction with a worse mortality risk from CKD in older age.

The less than additive effect of COPD and older age indicates that there are less sufficient causes for mortality that include COPD and older age than expected(477). This may be due to competing risks of other

causes of mortality unrelated to COPD or because of the shared indication of worse HF severity in COPD and older age reducing their overall effect when experienced in tandem. The supra-additive biological interaction between CKD and older age is a likely consequence of a lower threshold for poor renal function in older age. These combined results also indicate the importance of measuring biological interaction to understand causal mechanisms. Statistical interaction considers the need for a product term in a statistical model but does not determine whether the absolute risk differs for any given factor(453). Relying on statistical interaction which is the focus of most previous evidence would not have identified the biological interaction between CKD and DM and would have suggested an opposite association for CKD and older age.

13.5.3 Other factors

There was a significant statistical and biological interaction between male gender and CKD for both outcomes. This might indicate more severe renal disease in men than women or a higher propensity for admission in men than women with renal disease and HF. Previous evidence points to a gender difference in terms of progression to end stage renal disease with women on average being 10 years older(478) and with a slower rate of decline(479). In patients with CKD there was only minimal difference in the prevalence of the most severe renal category (eGFR <15) between men (2.4%) and women (1.9%), but men did have a higher prevalence of severe renal decline (17%) than women (14.5%) and might account for some of the increased risk.

13.6 Comorbidity models

Risk prediction models can assist in identifying individuals at risk of adverse events and to target limited resources to those most at risk. The objective of this thesis was to firstly determine whether the measures of comorbidity severity and change proximal to poor outcomes would be associated with those outcomes. The second objective discussed here was to determine whether the comorbidity measures that had significant associations for mortality and hospital admission would contribute to the fit of a multivariable model for both outcomes. The hypothesis was that comorbidity measures that take account of severity and change over time

may provide better prediction of outcomes. The clinical importance in addition to better prediction for HF patients is that the severity of comorbidity is potentially amenable to modification and can act as a trigger for interventions to reduce risk.

Severity: When the different comorbidity measures were tested in an overall core model including all other covariates there were some similarities and some differences across the two outcomes. For both mortality and hospital admission, for all three comorbidities defined by drug or physiological severity measures (where included) there were significant improvements in model fit when these measures replaced the comorbidity status measures of presence or absence. This is an important consideration for prognosis. Most current prognostic models include comorbidity by status which misses important information for individuals who vary in the severity of their comorbid disease. Inclusion of severity in prognosis would facilitate routine monitoring and targeted intervention to the most at risk groups.

Severity change: With the exception of COPD severity change using physiological measures, comorbidity severity change measures were a significant improvement over status measures for mortality. The model fit however was similar to the models with the comorbidity severity measures. For the hospital admission outcome, with the exception of CKD change, the models with the severity change measures showed no improvement in model fit over the comorbidity status models.

For CKD for both outcomes, the severity model showed the best fit compared to the severity change model but both were a significant improvement on the status model. When baseline renal status was added to the change model there was significant improvement in model fit and this resulted in incrementally better model fit from the renal status model to the renal severity model to the renal change model for both outcomes. This may have been a similar finding for DM and COPD, if it was possible to adjust for baseline comorbidity status in the mortality models and requires further investigation.

The difference in the contribution of the change measures to the hospital admission models compared to the mortality models also requires further investigation. It is possible that with the exception of CKD change, which has a clear indication for worsening HF severity prior to admission, that the high risk of cardiovascular compromise in new onset HF counteracts any additional risk associated with comorbidity change. This generates the hypothesis that comorbidity change in a more stable prevalent HF population may have yielded stronger and significant associations. A further consideration is that any noted change in comorbidity might trigger interventions aimed at hospital admission avoidance which reduces these associations. Where comorbidity change does improve the fit of a statistical model, as was the case in this study for all comorbidities and mortality and for admission with CKD, it provides an important clinical indicator of prognosis change during the life course of HF and comorbidity. Like the comorbidity severity indicator, change that is routinely recorded is a feasible and clearly identifiable indicator of increased risk and a target for intervention.

Overall comorbidity model: When all comorbidities were added by their different measures to the core model simultaneously, the comorbidity severity model had the best fit over the status or change model. Again adding baseline renal status to the change model improved its fit significantly and resulted in the change model having the best fit. The best fitting comorbidity models using mixed measures included DM by drug severity change for mortality and DM drug severity for hospital admissions, COPD by drug severity for both outcomes and CKD by baseline severity and severity change for both outcomes. Model discrimination was better for mortality (C-index 74%) than hospital admission (C-index 68%). Whilst the purpose of this thesis was to test the contribution of different comorbidity measures to a core HF model and not to produce a predictive model, the C-index for both outcomes was better than many prior models which in one meta-analysis showed a C-index of 0.71 for mortality and 0.68 for hospital admission(186).

The C-index has been demonstrated to have an upper limit that is determined by the distribution of risk in the sample. For an average risk of 10-50% with little spread then the upper limit of the C-statistic will be 0.62-0.63(480). The lower C-index for first hospital admissions may be partly due to the narrow distribution of risk within the incident HF population with 33% of patients admitted within a month, 46% within 3 months and most

in a year. This might also explain previous low C-index in prior models to predict hospital admission and readmission(161,164) which is also a high risk outcome for hospitalised patients.

The receiver operating characteristic (ROC) curve, or C-statistic has been criticised for its role in determining the accuracy of a prediction model, despite its constant use as a sole measure in the literature(458). Unlike calibration that determines, through statistics such as Hosmer-Lemeshow(481), how well the model predictions fit the observed data, the ROC merely ranks individuals into high and low risk. This does not indicate at all the range of individual risk within the groups. A model could predict all cases to have a risk of 0.52 and all controls 0.51. This would be very poor at prediction but have perfect discrimination. For this reason other global measures of model fit that don't rely on the ranking of cases have been recommended and were used in this thesis(482).

ROC is also a poor statistic to use in order to choose the best variable from a set of variables to use in a model. A factor that has a strong and significant effect on an outcome may have little influence on the ROC if it doesn't alter the ranking of higher and lower risk. However in terms of prediction it can have a big influence on an individual's overall risk estimate(458). In our analyses the ROC had the least response to the addition of the different comorbidity measures but this does not rule out improved predictive accuracy which was clearly evidenced by the other measures of model fit.

13.7 Strengths and limitations

Strengths and limitations are now discussed focusing on (i) the use of routinely collected data and (ii) specific considerations when using CPRD. Particular attention will be given to chance, bias and confounding.

13.7.1 Using routinely collected data

Health care data that is routinely collected during healthcare activity is becoming increasingly

available and used for health research to answer innovative, efficient and cost-effective research questions(483). There are many advantages of using routinely collected data over other bespoke modes of data collection which include the potential for large representative population based samples, long follow-up, low attrition, low cost and generalisability to real world populations where care is delivered(484). However routinely collected data is for clinical and administration purposes and is not based on any a priori research questions(485). Unlike in a prospective study the data collection cannot be chosen to reduce bias(486). This can create a number of challenges which include the large sample size and spurious correlations(487), inconsistent coding and sampling bias, missing variables, missing data and measurement bias(488). These will now be discussed in turn.

13.7.1.1 Large sample size

One of the key advantages of using routinely collected data is the availability of large samples of observations. All patients who attend routine clinical consultations are captured on electronic databases with drop-out usually only occurring when patients move to a new practice, transfer to a care institution or die which provides the potential for large captures of clinical information for use in research. Large samples mean that detailed analyses of exposures can be performed taking account of multiple confounders. The current investigation was conducted using a large sample from the UK general practice population which allowed detailed analyses of three HF comorbidities and adjustment for a range of covariates. It was possible to investigate exposure effects within sub-groups and to investigate interactions between the exposures and other covariates.

However, due to the high precision in big data, a small and clinically non-significant effect may be statistically significant with a low P value(489). In contrast a low P value in a smaller well designed study indicates a much larger effect size and may have more clinical importance(490). For this reason significance tests have to be viewed with caution when analysing large samples. In this study

the covariates in the multivariable models were pre-specified rather than relying on selection approaches that use statistical tests. Approaches such as stepwise selection when using large samples of observations can lead to the inclusion of clinically unimportant covariates(225). Observed effect sizes, confidence intervals and absolute differences were used where available rather than relying on the P value. Where statistical testing was used, a lower P value (<0.01) to indicate significance was selected. To improve computational efficiency for performing Lowess regression (data plots), a random sample of the data was used.

13.7.1.2 Inconsistent coding and sampling bias

Whilst using big data has the advantage of detailed analyses and generalisability of study findings, one of the key disadvantages is the inability to ascertain with any certainty the clinical status of the subjects included. Routinely collected data includes codes applied by clinicians in practice to record their patient's clinical information during consultations. These codes might relate to symptoms, diagnoses, medical history, treatments, care activity and referrals. Key challenges of using routinely collected data are that the codes applied to any specific disease group or clinical event might vary greatly amongst practitioners, be inaccurately applied or not applied. Coding decisions can also be influenced by health systems that provide financial recompense according to the code applied such as those based on insurance claims(491). Using routinely collected data to identify disease or exposure groups can be biased where there is inconsistent or inaccurate clinical coding.

The Quality and Outcomes framework (QoF) in the UK provides financial reward to GPs linked to the application of specific codes. In QoF diagnostic codes are often linked to appropriate tests and provision of evidence based therapies and so the focus is on more comprehensive and standardised coding, accuracy of diagnosis and improving quality care provision(93). The frequency of disease coding has improved since QoF and the accuracy of coding within the CPRD has been tested

against gold standards such as clinical verification and found to be valid for a range of morbidities(492, 493). However, whilst this increases specificity it is reliant on the selection of an appropriate code set and does not rule out accidental miscoding or negate that there may be subjects with a specific morbidity without a respective code applied. In this study Read codes were used for the selection of the HF sample and identification of the exposed groups(404) and careful consideration of the codes used was required to reduce the risk of misclassification. Any misclassification can lead to bias in the estimation of effects, the direction of which will depend on whether the misclassification is differential or non-differential across cases and controls(9) and will be explored in the following subsections.

Heart failure cohort: When using routinely collected data it is important that the sample is representative of the population from which it is drawn. The CPRD, one of the largest validated sources of routinely collected general practice data globally(237, 492) is representative of the UK general population who are registered with a CPRD general practice(238) but sampling requires careful consideration. An accurate clinical code being applied to a patient's record relies on the patient presenting to their general practice and the appropriate clinical code being applied. The decisions that underlie both factors may vary greatly across individuals(488). The HF cohort was selected from the CRPD on the basis of a first HF consultation code applied over a ten-year time window to March 2012. HF diagnosis relies on the presence of a combination of symptoms, signs and objective evidence of cardiac abnormality(93). HF symptoms in patient consultors are often non-specific which can lead to delayed or inaccurate diagnosis. The use of echocardiogram to aid diagnosis only improved following inclusion in the Quality Outcomes Framework(QoF) for general practitioners in 2006(103) and so the ESC diagnostic criteria for HF(93) could not be applied in this study due to the lack of routine recording of echocardiogram data. This means that the Read code for HF may have been inconsistently applied due to unconfirmed diagnosis and the accuracy of HF

codes may be a function of time. In this study the incident HF sample was selected using READ diagnostic codes for HF and excluding process codes that were indicative of prevalent HF. Whilst the potential for missed or inaccurate coding exists with routinely collected data, the accuracy of the HF diagnosis represented by clinical codes has been found to be 84% (406). Patient data in the CPRD is subjected to stringent quality checks and this study included 3-years of up-to-standard clinical data prior to study entry to validate the HF index date. The clinical code set used was validated by HF specialists and used in previous CPRD studies (405).

The code set used for identification of the HF subjects was based on Chapter G58 of the Read code which focuses on HF clinical diagnosis. A search was also performed for codes outside of chapter G58 that indicated a HF diagnosis. The search included cardiac terms that were combined with the term 'failure'. The inclusion of the term 'failure' was chosen to increase the accuracy of the diagnosis which is a combination of signs, symptoms and impaired ventricular function rather than any individual component. Whilst two of the codes in chapter G58 refer to impaired ventricular function or pulmonary oedema rather than 'failure' per se, these were retained due to their inclusion in the heart failure chapter which means their application is more likely to be a descriptor of HF. Other similar codes and symptom codes outside of chapter G58 were excluded due to the higher potential for reduced precision. Whilst there may be variations in the precision of the HF diagnosis using routinely collected data, the sample reflects the real world of HF in the general practice population which is important for the application of prognosis. This is in contrast to the randomised controlled trials whose use of strict control of internal bias and careful selection of patients, means that generalisability outside of the study sample can be limited. Further validation of the study findings across different HF diagnosis groups would be required.

Cases and controls: Both outcome measures related to objective events that are routinely recorded.

All-cause mortality cases were identified using a CPRD verified algorithm and where available cross checked with Office of National Statistics data. All-cause admission was identified by the first discharge date following the HF incidence date. A control sample should represent the population at risk of becoming a case and be selected independently of the exposure of interest. Controls should differ from cases only on the basis of their event status(494). Controls in the analyses were randomly sampled from risk sets defined by the cases after matching on calendar and follow-up time. Controls could be used more than once and later become a case. In this way the control sample reflected both the population at risk and the exposure time distribution in the entire denominator population. This design, consisting of a nested case-control approach within a well-defined cohort, virtually rules out the potential for selection bias and produces unbiased estimates of the exposure rate ratio(495).

In the HF-HA sub-sample a high number of events occurred in close proximity to the HF index date. To ensure that the controls differed to the cases on event status, controls were used that did not experience the event in the next 3 months after their selection as a control. This is analogous to the wash-out period often employed in case-crossover studies where separation is created between the case and control windows within the same patient to allow for the investigation of exposures that are associated with more insidious case events(408). The aim of this separation of time windows in the case cross-over approach is to ensure that the exposure measurement in the control window is not associated with the comparator case event that is occurring in close proximity. In the thesis, comorbidity exposure status over the six months prior to the first hospital admission was investigated. Whilst, a six month wash-out period could have been used, the control sample would have been too limited in size due to the high event rate within the first six months. This does mean that the comorbidity exposure effects on hospital admission may have been reduced as a result of the controls having an event within the following six-months which creates an overlap between the exposure measurement windows. This creates the hypothesis that investigation of the association

between comorbidity severity and change prior to admission in a prevalent HF cohort (with fewer more widely distributed events) would yield stronger exposure effects.

Exposures: Comorbidity exposure groups were identified using physiological measures (CKD) or a combination of clinical codes and drug measures (DM and COPD). These approaches were validated by clinical experts and prior evidence where available. All patients had 3 years of up to standard data prior to study entry so that prevalent DM or COPD comorbidity exposure occurring before to HF and incident comorbidity exposure occurring after could be identified.

The prevalence of CKD identified by Read codes was very low prior to 2006 which then increased remarkably post 2006. This difference likely indicates the low coding of CKD prior to the Quality Outcomes Framework(QoF) in general practice. Early coding of CKD was reliant on specialist investigations and objective evidence of disease. The QoF introduced the eGFR as a standard diagnostic measure into general practice and encouraged the early diagnosis of CKD using this readily available approach. These factors are likely to have led to the increase of CKD codes beyond this point. Given the clear discrepancy in coding across the two time-points the decision was to use the eGFR measure which is routinely recorded in all HF patients in the analyses. Due to the variation of physiological measures over time, the eGFR measure could not be used to position the timing of the CKD onset.

Whilst the clinical diagnosis of DM and CKD in general practice is based on an objective physiological measure, this is not often the case for COPD. Spirometry is under-utilised(411) and where it is used can be misinterpreted in the context of HF(415). A validated approach for identification of COPD using a combination of clinical codes and prescriptions was used, which has shown good precision in the CPRD(412,413). However within the COPD group, there was a

difference between those with and without an FEV1 measure recorded with the latter having a healthier risk profile and lower mortality risk. This could be due to misclassification of COPD in the absence of spirometry (i.e. the healthier non-FEV1 COPD group did not actually have COPD) or a propensity for using spirometry in more severe or symptomatic HF-COPD patients.

Despite the high precision previously found for identification of COPD in the CPRD using combined clinical and prescription codes, the possibility of misclassification cannot be excluded. In situations where misclassification is non differential and the exposure is dichotomous then the exposure effect will be diluted and the resultant bias will be towards the null value(9). However given that the use of spirometry (with the assumption of improved diagnostic accuracy) appears to be associated with more severe HF-COPD patients it is likely that the misclassification would be differential on the event status with healthier controls subjected to more misclassification (due to the more infrequent spirometry leading to over or under-diagnosis).

Finally, it cannot be ruled out that some of the higher risk in the more severe FEV1 group may also result from conflation of spirometry results in the context of severe HF(342,343) which would increase the relative estimates associated with the COPD status in this group(415). The higher relative risk of mortality in the FEV1 group with mild COPD (GOLD stage 1) than in the general COPD group (with or without FEV1) suggests that the increased risk in this group may in part relate to worse HF severity. However whilst routine use of spirometry including FEV1/FVC ratio is required to test and disentangle these associations(397), the dose-response relationship between reducing GOLD stages and increasing mortality risk does suggest a true association between COPD and mortality.

Confounders: A wide range of routinely collected confounders relating to socio-demographic, risk

factors and drug data were used that were based on evidence in HF prognosis studies and were subjected to detailed investigation. For a specific physiological confounder, the measure was converted from any varied units of measurement to a standard unit and removal of implausible values was performed prior to selection. To account for the time-varying nature of confounders, the most recent available measure to the match date was used. Key challenges for the measurement of confounding when using routinely collected data are measurement error, unmeasured confounding through missing variables and missing data. The CPRD analyses used the most recent measurement of confounders prior to the match date in recognition of the dynamic status of confounders that change over time, but did not include duration of the confounder which may influence its confounding influence on the exposure association. Additionally, measurement error of the confounders in routine practice cannot be ruled out, albeit non-differential measurement error will introduce bias towards the null value(496).

13.7.1.3 Missing variables

Key advantages of prospective studies are the ability to collect all relevant data required to answer the research question and take account of confounding or to randomise the intervention exposure to reduce the bias caused by unmeasured confounding(486). This is not possible in retrospective studies using routine data and lack of information in relation to potential confounders is common(491). Bias will remain where there are unobserved or imperfectly measured factors that influence both exposure and outcomes and it is important to acknowledge the potential confounding caused by missing information in the interpretation of findings.

In this study the covariates were selected using the prior review(239). However there was a key gap in the available data in relation to the research questions. Whilst severity measures were available for the comorbidity exposures these were absent for HF. Adjustments therefore could not be made

for ejection fraction, New York Heart Association (NYHA) class or brain natriuretic peptide (BNP), which are not routinely recorded in the CPRD. This is important because of the close relationship between HF severity and the physiological severity measures for both COPD and CKD. When age was used as a pseudo measure of HF severity there were biological interactions present for both comorbidities which emphasises their interrelationships. The lack of HF severity measures meant that it was not possible to distinguish between the independent influence of the comorbidities on the estimates of effects for the outcomes investigated. Lack of some known confounders in the CPRD such as HF severity means that unmeasured confounding was a possibility and this was acknowledged in the interpretation of the findings.

13.7.1.4 Missing data

An important consideration when using routinely collected data is the varying level of missing data inherent in some clinical variables. Provision of health care by different providers, inconsistency of clinical tests which are performed as needed rather than on a regular schedule and variation in coding of isolated events means that routine data provides only a partial picture of the patient(491). In this study the exposures were focused on chronic disease comorbidities which undergo routine monitoring and standard treatments guided by national policies, which reduces their risk of being missed. Three years of up to standard clinical data was also available prior to study inclusion for each subject reducing the risk of missed recording. The outcome measures were also objective, routinely recorded measures. Hospital admissions are identified through discharge codes which are applied by hospitals to seek financial recompense for the care they deliver and thus increases completeness. For the mortality outcome all deaths are recorded on registration by the Office of National Statistics and this was used, where linkage was available, to validate the CPRD mortality status. However there was varying levels of missing data in the covariates included in the analysis which can potentially lead to bias when associated with case status(9).

The level of missing data ranged from 0.5% (blood pressure) to 18% (cholesterol). Deprivation was available for only 59% of the overall cohort so was only included in a sensitivity analysis for the mortality outcome. Missing data was more prevalent in the case group than the control group. The differential nature of missing data can lead to selection bias(225). In this study, given that the comorbidities were associated with the case group, the differential missingness could have led to lower prevalence of comorbidity than would naturally be found in the denominator population by using complete case analysis.

To minimise this bias, several approaches to handle missing data were considered. Full case analysis, missing indicator or overall mean imputation were not chosen as the missing data were not completely at random and, where this is the case, these approaches can lead to bias(497). Instead confounders were imputed using single chained equations followed by a validity check that the imputed data was not statistically significantly different than the available data. In single imputation, only one estimate is used. In multiple imputation, various estimates are used, reflecting the uncertainty in the estimation of this distribution. Single imputation has been found be better than the other approaches previously mentioned(449) and if the data are missing at random or missing completely at random will result in unbiased, similar direction and magnitude of regression coefficients to multiple imputation(449, 498). However one of the limitations of the single imputation approach used in this study is that by not accounting for the uncertainty in the missing data, it can lead to too small estimated standard errors and confidence intervals(499).

13.7.1.5 Measurement bias; comorbidity severity and change

A priori definitions of comorbidity severity and change based on current clinical guidelines(410,418,419,432) were used supported by available evidence. For DM there is no clear

physiological measure of severity. HbA1c as a salient indicator of glucose stability was used(420) based on evidence of association with cardiovascular outcomes(425) and routine measurement in clinical practice. FEV1 for COPD and eGFR for CKD are evidence based indicators of severity that are included in current guidelines. Whilst the level of missing data was acceptable for DM and CKD (approximately 10%), missing FEV1 data was 50% in the total COPD cohort and more in the hospital admission linked sub-sample which had more narrow measurement time-windows. In addition to the physiological measure, for COPD and DM, a drug severity indicator was used which was complete for the comorbid groups. Given the high level of missing FEV1 data in the hospital admission linked sub-sample, the analyses was restricted to the drug measures for the COPD group.

In CKD, eGFR was estimated using the Modification of Diet in Renal Disease (MDRD) formula. This equation adjusts for significant non-renal influences such as age, sex, race, and body size, is preferred to other equations such as the Cockcroft-Gault formula(399) and has been validated in the HF population(398). Whilst Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) is now the recommended formula for estimation given its higher accuracy in higher eGFR groups, the MDRD was chosen for comparability with prior evidence and preference in older age groups, lower eGFR groups(500) and in HF.

For the severity change measure, two measures were used which were identified in maximum time-windows for each patient and the absolute and relative change between the two measures was standardised over a specified time for each outcome. Other methods were considered such as fitting a regression line to change slopes to calculate the rate of decline for an individual but it was decided that using categories of change based on evidence guidelines was preferable as they are easily identifiable and clinically meaningful.

Adjustment was not possible for the DM or COPD change measures by baseline levels which may have diminished the strength of associations, particularly for DM and mortality where a U shape relationship was observed. The effect of an increase or decrease in HbA1c is likely to be influenced by the start and end HbA1c. Where there are opposing baseline and change effects this could lead to reduced or abolished effects where they are considered alone. This was clearly seen in CKD where adjustment for baseline eGFR status was possible. The adjustment made a clear difference to the associations of eGFR with both outcomes and significantly improved the fit of the multivariable models. Whilst the model fit for the DM and COPD models did not improve when the severity measures were exchanged for severity change measures this could be a product of the lack of adjustment of the severity change measure with severity.

Time-dependent measurement of exposure: Previous studies have been criticised for not including long enough follow up periods to identify the power of individual factors contributing to risk(183) or not taking account of the time-varying nature of exposures. These criticisms were counteracted in two ways. Firstly, the CPRD was a longitudinal HF cohort with up to 12 years follow-up. Secondly, with the recognition that comorbidity exposures and other factors change from disease onset over the course of the HF, severity and change were measured during follow-up and in specific time-windows before an outcome. The time-windows for the physiological measurement of severity were longer for mortality (maximum of 3 years) than for hospital admission (maximum 6 months). This was to allow capture of routinely collected data and the hypothesis that change occurs more chronically and over a longer time prior to death than to hospital admission. That said most measures were recorded within 7 months of mortality.

Future validation is required in a prospective study where exposure change is measured retrospectively at baseline in a prevalent cohort. The hypothesis generated by the CPRD analyses is

that this change should be associated with imminent outcomes (1 year mortality and 6 month hospital admission) in the prevalent general practice HF population. An incident cohort was used, so that comorbidity could be investigated within the life course of HF and associated outcomes. Using a matched analysis on time in follow-up meant that adjustment for the HF duration was possible whilst measuring comorbidity severity and change during follow-up. To test comorbidity severity and change within a prevalent cohort would mean that both these factors would require consideration in the model.

13.7.2 Specific considerations when using CPRD

13.7.2.1 Study design

The purpose of this thesis was to ascertain the importance of recent comorbidity severity and change for hospital admission and death during the heart failure disease course. This was driven by the hypothesis that comorbidities would be more severe and change prior to these outcomes. A nested case-control design using risk set sampling of controls was used for the analyses. The strength of this approach is that it allows the measurement of time-dependent exposures and yields similar estimates of effect to a Cox proportional hazards analysis but more efficiently (220) and has been employed in a number of prior studies(160,431). Increased efficiency was required in this study due to the long follow-up, use of biomarkers to measure severity which change frequently over time and the high number of events (see section 4.4.1).

Using the risk-set sampling approach, the probability of the selection of a control is directly proportional to their person-time contribution to the denominator of the incidence rate. The control series is therefore sampled to represent the person-time distribution of exposure in the source population(9) which provides a valid estimate of the incidence rate ratio(219). The analysis of the nested case-control data was based on the conditional logistic likelihood which is a partial likelihood

similar to Cox's partial likelihood for survival data (501) and is a natural extension of the counting process formulation of Cox's proportional hazards model (section 4.4.1)(502).

However the risk set sampling approach has some limitations. The most recent exposure measurement reflects the exposure time distribution in the source population up to the point of measurement. This means that the comorbidity severity and change experienced by the source population since HF diagnosis and up to the point of analysis is included. The first limitation is that the time-dependent nature of measurement means that these measures could not be used directly for prognosis(221) as prediction can only be based on prior and not future knowledge. Second, the measure of risk generated by this approach compares the rates of event in the exposed and unexposed groups. The rate ratio, similar to a hazard ratio, provides a relative estimate of instantaneous risk over the study period and does not reflect a time-unit of the study or the time to an event. The lack of an intercept using conditional logistic regression also means that absolute risk cannot calculated directly by the model(227) and so survival plots cannot be readily produced. That said, baseline hazard functions can be calculated for case-control studies which are nested within a known cohort and the number of individuals at risk at each failure time is known(503), but this was not done in this study. The focus of this thesis was to ascertain the importance of recent comorbidity severity and change during the course of HF for the outcomes of mortality and hospital admission and not to produce a prognostic model. The significant associations of these measures with both outcomes and the improvement to statistical models following their addition demonstrates their potential importance for prognosis which can now be further tested in a prospective cohort using an non-matched design. The hypothesis is that a prognostic model including these measures should predict 6-month hospital admission and 1-year mortality within a prevalent general practice HF population.

13.7.2.2 Long follow-up

The HF population in CPRD ranged from 2002 to 2012 which means there will be variation in the diagnosis and clinical management of the HF sample. The influence of clinical guidelines and inclusion of HF in the Quality and Outcomes Framework in 2006(103) means that HF management will have changed over this time period. The use of beta blockers and ACEi as the mainstay of HF pharmacological interventions was significantly higher in the HF sample after 2006 as was the prevalence of comorbidity. Use of devices which were not included in this study, have also increased in later years(93). However, matching on calendar time in the analyses means that these time-varying factors were adjusted for with the cases being compared to controls at the same time point. Use of time-dependent measures also means that any developments in clinical or patient characteristics during follow-up will have been captured. In addition the mainline drug treatments were adjusted in the analyses.

13.7.2.3 All-cause outcomes

When selecting appropriate study outcomes, two options exist; disease specific outcomes and generic 'all-cause' outcomes. The choice will depend on the study hypotheses and the availability of data. Disease specific outcomes allow the investigation of the influence of various exposures on the pathophysiological progression of the index disease under study whereas generic outcomes investigate the influence of exposures on the progression of the patient. In this thesis all-cause hospital admissions and mortality were chosen for two main reasons. The life-course experience of people with HF has changed over recent decades. Whereas cardiovascular causes used to be the predominant cause of admission and death in HF, non-cardiovascular comorbid diseases now predominate(125). Given the relative severity of HF which is akin to most common cancers, the assumption might be that other less severe comorbid diseases would only influence these patient outcomes through moderation of the HF severity and would not impact directly. However research

has shown that non-CVD causes have now overtaken cardiovascular causes of both outcomes in the HF patient(159). This means that in terms of understanding the life course of the heart failure patient, these non-CVD comorbidities needed to be captured in the outcomes investigated. For patients this provides important information on how more likely their risk of hospital admission and death are given their multiple diseases compared to having HF alone. Investigation of interaction between the comorbidities also provides important information regarding the summation of risk across diseases.

However, to understand the interaction between the comorbidities and the HF to determine how they influence the progression of the HF, disease specific outcomes would be required. This information is important to identify whether interventions should be targeted at the index disease, the comorbid disease or both and was a limitation of this study. A key challenge to using disease specific outcomes is their availability and accuracy. The CPRD makes available HES data which lists all ICD-10 codes linked to the hospital discharge. These codes are not ordered and don't provide information on the cause of admission. The principal code at discharge can be identified but this requires additional costs which were not available for this study. In England and Wales, the legal requirement to register all deaths means that death registrations with the Office of National Statistics (ONS) provides the most complete data source for mortality statistics. However ONS data is only available under the CPRD linkage scheme which covered only 60% of the study cohort.

An additional consideration when choosing outcomes is the accuracy of the available data. The ability to investigate disease progression by disease specific outcomes rests on the fundamental assumption that the cause of death can be determined accurately. The determination of the underlying cause of death is often ambiguous and the accuracy of death certificates has been found to be low in a number of prior studies(504, 505). Assignment of a specific cause is subject to error particularly when the patient has a number of underlying diseases or health factors. HF itself is

considered a mediator between disease and death and is therefore ill-defined as a cause of death. HF is often applied where the aetiology is not known but this is not consistent and underreporting of circulatory diseases on death certificates is common(506). For this reason coding recommendations from the International Classification of Diseases Tenth Revision (ICD-10) are to apply other plausible heart conditions in place of HF(507).

All-cause mortality outcomes are now advocated in cancer screening trials where bias due to death coding errors has been found in a number of studies(508). The recognition that some cardiac interventions may cause non-cardiac deaths(509) has also seen all-cause mortality being used more in cardiac drug trials(510, 511). Conceptually, in considering an index disease as the cause of death, all deaths that are directly or indirectly caused by the index disease or its treatment should be included(508). An example would be a patient with HF who dies of renal failure secondary to the HF who is then coded with renal failure as the cause of death. Using all-cause mortality means that all indirect and direct deaths caused by HF are captured and there is no bias in the application of the cause of death.

13.7.2.4 Model building

To test the contribution of comorbidity measures to statistical models, a set of pre-specified covariates were used to build the core model. The selected covariates were based on prior evidence identified in the systematic review and through detailed consideration of the exposure associations with and without the covariate included. It has been argued that these approaches are preferable to stepwise selection which relies on statistical testing using arbitrary thresholds for exclusion that can lead to bias and overfitting(225,455-457). One of the limitations in the study and discussed previously, was the missing information on some of the potential predictors indicated by the prior evidence. A key missing variable that was not included in the models was HF severity. Whilst the analysis accounted for duration of HF, severity can vary widely depending on a

number of patient and treatment factors. A range of factors that are associated with HF severity including HF drugs, systolic blood pressure and renal function were included, but better adjustment would include more specific physiological measures of HF severity. Hospital admissions are also a known predictor of mortality in HF and were strongly associated with first hospital admission following HF diagnosis in this study. However the hospital admission data was only available for 60% of the total dataset and so was not included in the mortality analyses but should be considered for future model building. Inclusion of HF severity measures and hospital admissions in both the future testing of the comorbidity severity and change exposures and the development of HF comorbidity severity and change prediction models is required to better elucidate the independent exposure associations and potentially substantially improve prediction.

13.8 Chapter Summary

In the incident HF general practice population, a quarter of patients died within a year and a third had a hospital admission within a month of diagnosis. DM, COPD and CKD were common HF comorbidities and a high proportion of these groups had severe or worsening comorbid disease during their HF disease course. These dynamic severity and change measures of comorbidity were found to have significant and independent associations with both outcomes and improved statistical models in HF. This has important implications for HF management and prognosis which will be discussed in the final Chapter 14.

Chapter 14 Summary and implications

This chapter summarises the key findings from the three thesis phases and within the context of the prior evidence from Chapters 1-4, sets out the implications of the findings for public health, healthcare services and education, nursing practice, clinical care and prognosis research. Finally the future directions of the work are discussed.

14.1 Overview

In three phases, this thesis has synthesised the current evidence and developed new evidence on the influence of non-cardiovascular (CVD) comorbidities on outcomes in the general population of HF patients. In the first phase a systematic review identified 68 prior prognosis studies that focused on non-selected HF patients and included one or more non-CVD comorbidities, but there were important evidence gaps. Previous studies were based mainly in hospital settings and focused on mortality, where non-CVD comorbidities were found to have significant and independent associations. There was little evidence relating to community settings or other outcomes such as hospital admissions or quality of life. Chronic disease comorbidities change in severity as they progress during the life course of an index disease and this may alter their influence on patient outcomes. Whilst some evidence existed on the effects of renal dysfunction severity and change in hospital based HF patients, a key evidence gap was on how comorbidity severity or change influences patient outcomes in the general practice HF population. This information is important for the monitoring and management of comorbidities in the specialist or routine care of the index disease, to identify patients whose prognosis is changing, to target interventions at the highest risk HF groups and to tailor interventions to those with the most potential benefit. Phase two and three included a new set of studies to address these evidence gaps.

In phase two and three, a national UK dataset of new HF patients from the general practice population was used to investigate the influence of dynamic indicators of non-CVD comorbidity severity and change on

mortality (phase 2) and hospital admissions (phase 3) over 12 years of follow-up. A new conceptual framework was developed for the inclusion of comorbidities in the prognosis of an index disease and measures of comorbidity severity and change were devised using routinely collected clinical data and underpinned by current evidence and clinical guidelines. Three comorbidities of DM, COPD and CKD were selected from the systematic review based on their higher prevalence in the HF population and the potential for measuring their severity and change using routinely collected data to test the study hypotheses. This set of studies found that the risk of hospital admission and mortality was high following HF diagnosis and there were similar significant and independent associations between all three comorbidities and both outcomes. An additional finding was that combinations of different comorbidities experienced together, yielded different risk estimates than expected from considering their independent effects.

When each of the three comorbid groups were categorised by their severity measured during follow-up and prior to an event, there was significant differentiation of risk between the severity categories, with higher severity categories associated with the greatest risk for hospital admission and mortality. When comorbidity status measures (presence or absence) were replaced by their severity measures there was significant improvement in the fit of the prognostic models for both outcomes. Recent change in severity over a year was also significantly and independently associated with mortality and provided better model fit over status measures for all comorbidities and over status and severity measures for DM and CKD. Recent change in severity over 6-months using drug measures were significantly associated with first hospital admission but did not significantly improve the fit of the prognostic models. Using a physiological measure of change for CKD, recent severity and change measures incrementally improved model fit for hospital admission.

14.2 General implications

Public health: Chronic diseases such as HF are the predominant cause of death globally. Clinical epidemiology or prognosis seeks to understand the onset, progression, interventions and outcomes of people with a disease and is of key importance for public health policy and primary and secondary prevention. In an ageing society older people often experience more than one chronic disease at the same time and this

comorbidity can lead to poorer health and earlier death and has become a global health care priority requiring a radical redesign of health services, clinical care, education and research.

Prognosis and subsequent public health policy has traditionally focused on the aetiology and risk factors relating to the index disease. In a national cohort of patients from the general practice population where most chronic diseases are routinely managed and recorded, this thesis found that other chronic disease comorbidities are highly prevalent, have significant importance for outcomes in patients with an index disease and require further consideration in prognosis. Understanding which of the most prevalent chronic disease comorbidities are associated with the highest risk in the general population of people with an index disease is of critical importance to public health focused on improving outcomes. Where multiple comorbidities coexist, specific combinations of diseases such as DM and CKD, carry a higher than expected risk for people with an index disease and provide a key target for health interventions aimed at prevention or optimisation of therapies.

Healthcare services: Over the past two decades healthcare systems and delivery have been framed around the management of specific chronic diseases supported by national and international, standardised guidelines and disease-specific policy. This specialism has improved outcomes in patients with a specific disease but for many diseases such as HF, outcomes remain poor. The findings in this thesis demonstrate the importance of other disease comorbidities for outcomes. A key finding was that the progression and increasing severity of chronic disease comorbidities over the course of an index disease is associated with significantly higher risk of poor outcomes. This is important because, whilst chronic disease comorbidities are incurable, increasing severity may itself be amenable to intervention and modification to reduce risk. Better integration of chronic disease management and care is therefore required for the index disease and the comorbidities. Disease specific policy and care delivery needs to take account of other diseases, their routine monitoring and measurement and recommended interventions. Patients require integrated information to optimise their understanding of self-management where multiple diseases coexist and also their potential disease course within the context of their comorbid diseases.

Healthcare Education: In the same way that clinical care is organised around specialism, the review of individual conditions is most often the focus of healthcare education and particularly specialist nurse curricula. This means that specialists might lack knowledge about comorbidities, shared disease mechanisms or how to provide optimal management for the index disease and the comorbidities for a high percentage of patients who experience both. The high prevalence and importance of comorbidities for patient outcomes found in this thesis means that education curricula needs to be redesigned to better prepare clinicians to manage patients with multiple diseases(512).

Nursing practice and heart failure specialisation: The nursing profession has developed over recent years through specialisation with nurses taking on advanced knowledge and skills in the management of a specific disease, condition or intervention. Specialist nurses use these skills to prevent and manage patients' symptoms and to provide interventions in relation to a specific disease, to prevent deterioration. Whilst this specialisation has improved outcomes for patients, the current context of increasing numbers of people with comorbidities means that it can be at the expense of person-centred and holistic care.

This thesis has demonstrated the high prevalence of chronic disease comorbidities in the general practice HF population which leads to earlier hospital admission and death for these patients. Whilst the reasons for this association might be multi-factorial including shared pathophysiological processes and risk factors or conflicts in management and care, it poses new and important challenges for specialist nursing. In the HF example, comorbidity can cause conflicts in treatments, drug interactions and challenges for patient self-care that is managed through multiple different disease specific self-management plans. The inclusion of co-morbidities in HF management is then a key component of holistic nursing care.

Whilst guidelines should help nurses to make decisions in their daily practice, much of the guidance that nurses' provide patients on the key components of self-care behaviour such as lifestyle advice and symptom monitoring and management are not adequately included in guidelines. Comorbidity, which is an important

consideration, has only brief inclusion in this guidance and focuses on medical management. Heart failure specialist nursing has demonstrated excellent advantages for patients through focused nursing care approaches on the management of aetiology, pathophysiology and symptoms, but this thesis indicates that there now needs to be a broader specialist approach which takes account of the growing number of patients with comorbidities.

Prognosis research: Prognosis is important for the management of people with chronic diseases in order to target groups for treatments, assess the efficacy of interventions or differential responses to treatments, to assess different health care practices and to better inform patients for shared decision making. The clinical course of a chronic disease can be lengthy from new onset to more severe disease to death and so the accurate prediction of outcomes requires the simultaneous account of multiple and varied environment, host and disease factors. This thesis has shown that alongside disease specific factors that influence its severity and progression such as aetiology and risk factors related to the index disease, other chronic diseases present at the same time are an important consideration.

Whilst there is growing recognition of the importance of comorbidities for patient outcomes, prognosis studies have so far been limited to the presence or absence of comorbid disease which ignores the dynamic nature of comorbidity that develops in severity and changes over time, often in tandem with the index disease. This thesis has shown that individual risk stratification using comorbidity is a dynamic process that is improved by longitudinal information on the disease severity and its recent change over time. More specifically, recent severity and/or change is a useful indicator of patients whose prognosis is changing and who are at a higher risk of imminent hospital admission or death in the general practice population. This could act as an important trigger for targeting interventions to reduce the risk of poor outcomes.

The comorbidity measures used in this thesis were based on routinely recorded clinical information underpinned by current evidence and clinical guidelines. This means that the measures can be easily applied in practice, are clinically useful in routine care and familiar to clinicians and patients. For some diseases,

routine monitoring using physiological measures is inconsistent and this thesis highlights the importance of regular monitoring of both the index and comorbid disease. The drug measures of comorbidity severity showed significant associations with both outcomes and are automatically and electronically recorded on all patients, providing a consistent disease severity measure. Recent change in drug severity improved prognosis for mortality but not hospital admission. This implies that different factors might be influencing admissions that require further consideration in prognosis. An additional implication for prognosis from this thesis relates to the interaction between comorbidities. Chronic disease comorbidities are currently included independently from one another in prognostic models. The thesis findings that relate to the importance of different combinations of comorbidities, means that they may need joint consideration for more accurate prediction.

14.3 Heart Failure implications

Public health: This thesis has shown that the prognosis of people in the general practice population with a new diagnosis of HF remains very poor despite optimisation of medical interventions over previous years. A quarter of all patients with HF died within a year of diagnosis and half within five years. Most patients also experienced at least one admission into hospital and a third of these were within a month of diagnosis. Non-CVD comorbidities were highly prevalent with half of all HF patients having at least one of the three comorbidities (DM, COPD, CKD) investigated. Despite the high risk associated with a new diagnosis of heart failure, these prevalent comorbid diseases increased risk of admissions and mortality significantly. Current public health policy focuses on the primary and secondary prevention of cardiovascular disease. The increasing number of people with CVD and other comorbidities means that these approaches need to be extended to identify and target shared risk factors for the most common CVD and comorbidity combinations. Further improvements in the poor prognosis in end-stage CVD require a broader approach to prevention as well as diagnosis, investigations and interventions that take account of non-CVD comorbidities.

Clinical care: The monitoring and management of non-CVD comorbidities needs to be included in routine HF care. Central to current HF clinical guidelines is the optimisation of drug therapies as well as assessment for interventions such as devices or surgery. The importance of the non-CVD comorbidities for patient outcomes

in this thesis implies that optimisation of treatments needs to include the comorbidity as well as the HF. The lack of guidance on comorbidities often inherent in disease specific guidelines means that better integration of specialist services may be required to deliver optimal care across diseases.

HF Prognosis: The trajectory of HF is complicated with wide variation in the experience of patients following diagnosis. Current prognosis tools have limitations including their use of invasive and complex clinical data and biometrics, their focus on select populations and use of static or single prognostic indicators measured at baseline. Whilst some dynamic cardiovascular prognostic indicators are included in HF prognostic models such as BMI, ejection fraction and blood pressure these are all based on a single measure and do not include indicators of change. This thesis has shown that dynamic measures of the severity of non-CVD comorbidities and their recent change that occurs during the HF disease course has significant associations with both mortality and hospital admissions and improves prognostic models for both outcomes. These comorbidity factors based on routinely collected clinical data and clinical guidelines provide the potential for a simple and clinically useful prognostic model for identifying HF patients whose risk is changing and require more intensive support. This is turn has implications for HF interventions that so far, have largely been tested in randomised controlled trials that have excluded patients with comorbidity. Optimal pharmacotherapy for this patient group remains poorly defined and there is also limited evidence of the efficacy of interventions that are considered in higher risk HF groups such as the implantation of devices or heart transplant.

14.4 Future developments

14.4.1 HF comorbidity and prognosis

Prognostic factors: This thesis demonstrated that dynamic measures of non-CVD comorbidities are important prognostic factors for hospital admission and mortality in an incident general practice HF population. There were interactions between some of the comorbidity pairs which indicate possible shared aetiology, risk factors or pathophysiological processes or a shared indication for an unmeasured factor. Further investigation is required to understand these mechanisms. The interaction between HF and the comorbidities is also important in order to inform the development of possible interventions which could be aimed at the HF, the

comorbidity or the combined diseases. Future work should consider these interactions which would require a general practice population with and without HF for the analyses. Other approaches within the HF cohort would include the investigation of CVD versus non-CVD outcomes or the interaction between the comorbidity severity and HF severity on outcomes. Age was used as a pseudo measure of severity and showed a significant biological interaction between CKD and older age in the opposite direction than expected when observing statistical interaction. The indication is that the effect of CKD is modified in more severe HF which is likely given the close pathophysiological pathways. These relationships require further detailed investigation using more refined HF severity measures.

Prognostic models: This thesis has shown that prognostic models are significantly improved by the inclusion of dynamic and time-dependent measures of non-CVD comorbidities. The CPRD is a comprehensive national dataset that includes longitudinal and routinely recorded clinical data on chronic disease management. Whilst HF related drug data and risk factors are available, physiological or blood test data to measure HF severity is not yet included. This means that the final models focused on severity and change for the comorbidities but not the HF. The final prognostic models using these comorbidity measures were comparable or better than prior HF models that included HF severity measures. This thesis also used a nested case-control design with risk set sampling of controls which allowed the efficient measurement of time-dependent severity and change. Whilst this approach yields unbiased estimates of rate ratios, absolute risk cannot be calculated due to the lack of an intercept. Model development now requires the addition of dynamic and time-dependent measures of HF severity as well as the comorbidities and testing within a non-matched cohort. With the addition of a variable that indicates time since HF diagnosis, this model should predict 6-month hospital admission and 1-year mortality within a prevalent general practice HF population. This model development will require the linkage of specialist and general practice data.

14.4.2 Interventions

HF prognosis studies have included comorbidities but there is still no clear evidence of tailored therapeutic interventions to reduce the individual risk of HF patients with comorbidity. Further epidemiological enquiry is required to investigate the complex relationships between comorbidities and outcomes in HF so that big data studies can be translated to the development of interventions. Therapeutic trials need to be developed to investigate the efficacy and efficiency of HF interventions in patients with comorbidity so that optimal therapy can be tailored to the patient. Pharmacological treatments and optimal doses need to be investigated for any differential effect across comorbid groups and new drugs that target shared mechanisms may need to be developed. Stratified medicine prognosis approaches are also required so that interventions such as devices in HF can be targeted to those with the most potential benefit and least harm. Whilst these interventions are often targeted at the higher risk HF groups there is limited information on whether high risk that is due to comorbidity would be mitigated by these approaches. Health service interventions also require testing to investigate their effectiveness in improving the joint optimisation of treatment for combined diseases. The integration of services and specialist care and of patient self-management approaches all require investigation.

14.4.3 Outcomes

A key gap identified in this thesis is the lack of prognosis studies that included patient centred outcomes such as health related quality of life (HR-QoL). This is an important gap given the poor HR-QoL found in HF populations which deteriorates as the disease progresses and which is itself associated with poor outcomes. The multi-factorial components of HR-QoL mean that is potentially amenable to modification through targeted interventions which makes it a clinically important outcome. The inclusion of patient important outcomes in prognosis is a key indication within international position statements and provides a mechanism to identify patients earlier whose prognosis is changing.

Given the high prevalence and importance of non-CVD comorbidities shown in this thesis, future work needs to investigate their association with HR-QoL. To that end, three linked studies have been conducted or planned to investigate the association between non-CVD comorbidities and HR-QoL:

(i) Quality of life study 1: HF and osteoarthritis publication on HR-QoL. In the first study the impact of non-CVD comorbidity on cardiovascular symptom specific physical limitations was investigated(29) (see E-Appendix E18). General practice patients with hypertension, ischaemic heart disease or HF with or without comorbid osteoarthritis (OA) were surveyed on their physical limitations related to their chest pain and shortness of breath. This survey data was linked to their clinical data and the comorbid CVD groups with OA were compared to those without, using linear regression. This study found that OA added to the CVD symptom related physical limitations and there was some evidence of interaction between the OA and HF observed for chest pain limitations. This study indicates the potential importance of non-CVD comorbidities for HR-QoL which includes symptoms as a component. The possible interaction provides further evidence that non-CVD comorbidities require including in HF specific management.

(ii) Quality of life study 2: 2C-HF prospective hospital investigation. In the second study, a new cohort of 180 HF patients were recruited from specialist HF outpatients' clinics and surveyed at baseline and at 1 month and 3 months follow-up (see E-Appendix E19 for ethics committee protocol). These surveys included generic and HF specific health measures that were linked to their clinical HF and comorbidity data extracted at baseline. This data included the comorbidity severity and change measures used in the CPRD studies and will be used to test these measures for their association with HR-QoL and change in HR-QoL over three months. The opportunity with this cohort analysis will be the inclusion of HF severity measures into the prognostic model.

(iii) Quality of life study 3: Swedish HF registry study. A third study is planned (see E-Appendix E20 for protocol) using the Swedish HF register (S-HFR). Whilst there are local HF registers in the UK there is no national register or routinely collected measures of quality of life in administration data. In the Swedish healthcare system, the national quality registers are obliged to incorporate patient reported outcomes for

certification at a high level. This is a unique database which has prospectively collected patient-reported outcome measures and there is data on around 50,000 HF patients from the hospital and outpatient HF population in Sweden. The data includes information on different comorbidities and patient-centred questionnaires on physical health, symptoms and quality-of-life which have been repeated over time. The data for this study has already been obtained and will be used to explore the influence of comorbidities on quality of life with a specific focus on the factors that moderate or mediate these associations. This analysis will provide the basis for the development of nurse interventions to improve outcomes in this group.

14.5 Conclusions

Heart failure is a serious chronic disease where, in the general practice population, a third of patients are admitted into hospital within a month of diagnosis and a third die within a year. Non-CVD comorbidities are common in this population and are associated with higher risk of being admitted into hospital or dying earlier. As with HF, chronic disease comorbidities progress in severity from new onset to end stage disease. This thesis has shown that prognosis, which is the prediction of outcomes within a disease population, also changes as the HF and comorbidities progress and change overtime. The dynamic nature of comorbidities is significantly and independently associated with increasing risk of poor outcomes and provides the potential for using routinely collected and clinically meaningful information to identify patients whose risk is changing.

This information is important for both public health and the clinical management of individual patients. Common non-CVD comorbidities experienced individually or in combination provide key targets for primary and secondary prevention. At a public policy level this information needs to be included in specialist disease management guidelines to improve outcomes in the HF population. For individual patient management this information is important so that interventions that optimise HF and comorbidity treatments can be targeted at the highest risk groups or so that treatments can be tailored to individuals that have the most potential benefit. More work needs to be done to elucidate the underlying mechanisms that explain the increased risk associated with non-CVD comorbidities so that interventions can be developed that target the index disease, the comorbidities, the disease combination or the patients' management of their combined diseases.

References

(1) Hennekens C, Buring J. *Epidemiology in Medicine*. 1st ed. Boston/Toronto: Little Brown and Company; 1987.

(2) Rothman K. Causes. *American Journal of Epidemiology*. 1976;(104):587-592.

(3) Hippocrates. On airs, waters, and places. *Medical Classics.* 1938;(3):19.

(4) Newsom SW. Pioneers in infection control: John Snow, Henry Whitehead, the Broad Street pump, and the beginnings of geographical epidemiology. *Journal of Hospital Infection*. 2006;64(3):210-216.

(5) World Health Organization. *Noncommunicable Diseases Country Profiles 2014*. Switzerland: World Health Organization; 2014.

(6) Porta M, Last JM. *Dictionary of Epidemiology*. 5th ed. Oxford: Oxford University Press; 2008.

(7) Brachman P. Epidemiology. In: Baron S, editor. *Medical Microbiology*. 4th ed. Galveston: Galveston (TX): University of Texas Medical Branch; 1996. Chapter 9.

(8) Killewo J, Heggenhougen K, Quah S editors. *Epidemiology and Demography in Public Health*. London: Elsevier; 2010.

(9) Rothman K. *Epidemiology, An Introduction*. 2nd ed. New York: Oxford University Press; 2012.

(10) Rothman KJ. *Induction and latent periods*. *American Journal of Epidemiology*. 1981;114(2):253-259.

(11) de Mutsert R, Jager KJ, Zoccali C, Dekker FW. *The effect of joint exposures: examining the presence of interaction*. Kidney International 2009 Apr;75(7):677-681.

(12) Haynes B, Sackett D, Guyatt G, Tugwell P. *Clinical Epidemiology: How to Do Clinical Practice Research*. 3rd ed.: LWW; 2011.

(13) Ben-Shlomo Y, Brookes S, Hickman M. *Epidemiology, Evidence-Based Medicine and Public Health*. 6th ed. Oxford: Wiley-Blackwell; 2013.

(14) Mathers CD, Loncar D. Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLoS Medicine.* 2006;3(11):e442.

(15) Kundi M. Causality and the interpretation of epidemiologic evidence. *Ciência and Saúde Coletiva.* 2007;12(2):419-428.

(16) World Health Organization. *Evaluation and use of eeidemiological evidence for environmental health risk assessment*. WHO; 2000.

(17) Whitehead D. Is there a place for epidemiology in nursing? *Nursing Standard*. 2000;14(42):35-39.

(18) Hall A. Qualitative research and its role in nursing knowledge. *Nursing Times*. 2006;102(20):32-35.

(19) Mulhall A. The case for a more epidemiologically informed nursing profession. *Nursing Times Research.* 2000;5(1):65-73.

(20) Boorse C. A rebuttal on health. In: Humber J, Almeder R, editors. *What is disease?* Totowa, NJ: Humana Press.; 1997. p.7-8.

(21) Scully JL. *What is a disease? EMBO Reports*. 2004;5(7):650-653.

(22) Ereshefsky M. Defining 'health' and 'disease'. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*. 2009;40(3):221-227.

(23) World Health Organization. 1*0 facts on noncommunicable diseases*. 2013; Available at: http://www.who.int/features/factfiles/noncommunicable_diseases/en/. Accessed November 11th, 2014.

(24) Stein RK, Perrin E, Pless IB, Gortmaker S, Perrin J, Walker D, et al. Severity of illness: concepts and measurements. *The Lancet*. 1987;330(8574):1506-1509.

(25) Huntley AL, Johnson R, Purdy S, Valderas JM, Salisbury C. Measures of Multimorbidity and Morbidity Burden for Use in Primary Care and Community Settings: A Systematic Review and Guide. *The Annals of Family Medicine*. 2012;10(2):134-141.

(26) Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Dieases*. 1987;40(5):373-383.

(27) Kadam UT, Schellevis FG, van der Windt DA, de Vet HC, Bouter LM, Croft PR. Morbidity severity classifying routine consultations from English and Dutch general practice indicated physical health status. *Journal of Clinical Epidemiology*. 2008;61(4):386-393.

(28) Prior JA, Jordan KP, Kadam UT. Associations between cardiovascular disease severity, osteoarthritis co-morbidity and physical health: a population-based study. *Rheumatology*. 2014;53(10):1794-1802.

(29) Rushton CA, Kadam UT. Impact of non-cardiovascular disease comorbidity on cardiovascular disease symptom severity: A population-based study. *International Journal Cardiology*. 2014;175(1):154-161.

(30) Murray SA, Kendall M, Boyd K, Sheikh A. Illness trajectories and palliative care. *The BMJ*. 2005;330(7498):1007-1011.

(31) Lynn J, Adamson D. *Living Well at the End of Life. Adapting Health Care to Serious Chronic Illness in Old Age*. Washington: Rand Health; 2003.

(32) Summerton N. Making a diagnosis in primary care: symptoms and context. *The British Journal of General Practice.* 2004;54(505):570-571.

(33) Engel GL. The need for a new medical model: a challenge for biomedicine. *Science*. 1977;196(4286):129-136.

(34) World Health Organisation. *Basic documents*. 39th ed. 1992. WHO.

(35) Saracci R. The World Health Organisation needs to reconsider its definition of health. *The BMJ*. 1997;314(7091):1409-1410.

(36) Bircher J. Towards a dynamic definition of health and disease. *Medicine, Health Care and Philosophy*. 2005;8(3):335-341.

(37) Royal College of Nursing. *Defining Nursing*. London: RCN; 2003.

(38) Santana MJ, Feeny D, Institute of Health Economics. *The Importance of Measuring Health-related Quality of Life*. Alberta,Canada: Institute of Health Economics; 2008.

(39) Organization of Economic Cooperation and Development (OECD). *Health Reform: Meeting the Challenge of Ageing and Multiple Morbidities.* OECD Publishing; 2011. http://dx.doi.org/10.1787/9789264122314-en

(40) Van der Akker M, Buntinx F, Knottnerus J. "Comorbidity or multimorbidity: what's in a name? A review of the literature". *European Journal of General Practice.* 1996;2(2):65-70.

(41) Feinstein A. "The pre-therapeutic classification of co-morbidity in chronic disease". *Journal of Chronic Disease.* 1970;23:455-468.

(42) de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity: a critical review of available methods. *Journal of Clinical Epidemiology.* 2003;56(3):221-229.

(43) Fried L, Ferrucci L, Darer J, Williamson J, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2004;59(3):255-263.

(44) Coulter A, Roberts S, Dixon A. *Delivering Better Services for People with Longterm Conditions*. London: The King's Fund; 2013.

(45) Smith SM, Soubhi H, Fortin M, Hudon C, O'Dowd T. Managing patients with multimorbidity: systematic review of interventions in primary care and community settings. *The BMJ*. 2012;345 doi:http://dx.doi.org/10.1136/bmj.e5205

(46) Bayliss EA, Steiner JF, Fernald DH, Crane LA, Main DS. Descriptions of barriers to self-care by persons with comorbid chronic diseases. *The Annals of Family Medicine.* 2003;1(1):15-21.

(47) Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *Journal of General Internal Medicine*. 2007;22 Suppl 3:391-395.

(48) Riegel B, Jaarsma T, Stromberg A. A middle-range theory of self-care of chronic illness. *Advances in Nursing Science.* 2012;35(3):194-204.

(49) Briesacher BA, Gurwitz JH, Soumerai SB. Patients at-risk for cost-related medication nonadherence: a review of the literature. *Journal of General Internal Medicine*. 2007;22(6):864-871.

(50) Barnes S, Gott M, Payne S, Parker C, Seamark D, Gariballa S, et al. Prevalence of Symptoms in a Community-Based Sample of Heart Failure Patients. *Journal of Pain and Symptom Management*. 2006;32(3):208-216.

(51) Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet*; 380(9836):37-43.

(52) Gott M, Barnes S, Parker C, Payne S, Seamark D, Gariballa S, et al. Predictors of the quality of life of older people with heart failure recruited from primary care. *Age Ageing*. 2006;35(2):172-177.

(53) Department of Health. *Improving chronic disease management*. London: Department of Health;2004.

(54) Robert Wood Johnson Foundation and the Johns Hopkins Bloomberg School of Public Health. *Chronic care: making the case for ongoing care*. New Jersey: Robert Wood Johnson Foundation; 2010.

(55) Tinetti ME, Bogardus ST, Agostini JV. Potential Pitfalls of Disease-Specific Guidelines for Patients with Multiple Conditions. *New England Journal of Medicine.* 2004;351(27):2870-2874.

(56) Boyd CM, Darer J, Boult C, Fried LP, Boult L,Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: Implications for pay for performance. *Journal of the American Medical Association*. 2005;294(6):716-724.

(57) Akner G. Analysis of multimorbidity in individual elderly nursing home residents. Development of a multimorbidity matrix. *Archives of Gerontology and Geriatrics*. 2009;49(3):413-419.

(58) Department of Health. *Long term conditions compendium of information*. London: Department of Health; 2012.

(59) Chamberlain AM, St Sauver J,L., Gerber Y, Manemann SM, Boyd CM, Dunlay SM, et al. Multimorbidity in Heart Failure: A Community Perspective. *American Journal of Medicine*. 2014;128(1):38-45.

(60) Kaplan MH, Feinstein AR. The importance of classifying initial co-morbidity in evaluating the outcome of diabetes mellitus. *Journal of Chronic Disease.* 1974;27(7-8):387-404.

(61) Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *Journal of Clinical Epidemiology*. 1992;45(6):613-619.

(62) Elixhauser A, Steiner C, Harris D, Coffey R. Comorbidity Measures for use with Administrative Data. *Medical Care* 1998;36(1):8-27.

(63) Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *The BMJ*. 2013;346:e5595.

(64) Fletcher R, Fletcher S, Fletcher G. *Clinical Epidemiology, The Essentials*. 5th ed. Philadelphia: Wolters Kluwer, Lippincott Williams and Wilkins; 2014.

(65) Sox H, Goodman S. The Methods of Comparative Effectiveness Research. *Annual Review of Public Health.* 2012;33:425-445.

(66) Hemingway H. Prognosis research: Why is Dr. Lydgate still waiting? *Journal of Clinical Epidemiology.* 2006;59(12):1229-1238.

(67) Croft PR, Dunn KM, Raspe H. Course and prognosis of back pain in primary care: The epidemiological perspective. *Pain*. 2006 5;122(1–2):1-3.

(68) Krumholz HM. Outcomes Research: Myths and Realities. *Circulation: Cardiovascular Quality and Outcomes.* 2009;2(1):1-3.

(69) CancerGuide: Statistics. *The median isn't the message*. 2002; Available at: http://cancerguide.org/median_not_msg.html. Accessed July 31st, 2014.

(70) Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Medicine*. 2013;10(2):e1001380.

(71) Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *Plos Medicine.* 2013;10(2):e1001381.

(72) Jacobs DR,Jr, Kroenke C, Crow R, Deshpande M, Gu DF, Gatewood L, et al. PREDICT: A simple risk score for clinical severity and long-term prognosis after hospitalization for acute myocardial infarction or unstable angina: the Minnesota heart survey. *Circulation*. 1999;100(6):599-607.

(73) Abraham WT, Adams KF, Fonarow GC, Costanzo MR, Berkowitz RL, LeJemtel TH, et al. In hospital mortality in patients with acute decompensated heart failure requiring intravenous vasoactive medications: an analysis from the Acute Decompensated Heart Failure National Registry (ADHERE). *Journal of the American College of Cardiology*. 2005;46(1):57-64.

(74) Wyatt JC, Altman DG. Commentary: Prognostic models: clinically useful or quickly forgotten? *The BMJ*. 1995;311(7019):1539-1541.

(75) Heathfield HA, Wyatt J. Philosophies for the design and development of clinical decision-support systems. *Methods of Information in Medicine* 1993 Feb;32(1):1-8; discussion 9-17.

(76) Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW, et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *The BMJ.* 2013;346:e5793.

(77) Jaarsma T, Beattie J, Ryder M, Rutten F, McDonagh T, Mohacsi P, et al. Palliative care in heart failure:a position statement from the palliative care workshop of theHeart Failure Association of the European Society of Cardiology. *European Journal of Heart Failure*. 2009;11(1):433-443.

(78) Beaglehole R, Bonita R, Kjellstrom T. *Basic Epidemiology*. Geneva: World Health Organization; 1993.

(79) Masson S, Latini R, Anand IS, Barlera S, Angelici L, Vago T, et al. Prognostic value of changes in N-terminal pro-brain natriuretic peptide in Val-HeFT (Valsartan Heart Failure Trial). *Journal of the American College of Cardiology.* 2008;52(12):997-1003.

(80) O'Connor CM, Whellan DJ, Wojdyla D, Leifer E, Clare RM, Ellis SJ, et al. Factors related to morbidity and mortality in patients with chronic heart failure with systolic dysfunction the HF-ACTION predictive risk score model. *Circulation: Heart Failure*. 2012;5(1):63-71.

(81) Cowie MR, Sarkar S, Koehler J, Whellan DJ, Crossley GH, Tang WH, et al. Development and validation of an integrated diagnostic algorithm derived from parameters monitored in implantable devices for identifying patients at risk for heart failure hospitalization in an ambulatory setting. *European Heart Journal.* 2013;34(31):2472-2480.

(82) Jowett N, Thompson D. *Comprehensive Coronary Care*. 4th Edition ed. London: Bailliere Tindall Elsvier; 2007.

(83) Mehta PA, Dubrey SW, McIntyre HF, Walker DM, Hardman SM, Sutton GC, et al. Improving survival in the 6 months after diagnosis of heart failure in the past decade: population-based data from the UK. *Heart* 2009;95(22):1851-1856.

(84) Davis R, Davies M, Gregory L editors. *ABC of Heart Failure*. 2nd ed. Oxford: Blackwell Publishing; 2006.

(85) Ashley E, Niebauer J. *Heart Failure. Cardiology Explained*. London: Remedica; 2004. Chapter 7.

(86) Mehta PA, Dubrey SW. High output heart failure. *QJM: An International Journal of Medicine*. 2009;102(4):235-241.

(87) De Keulenaer GW, Brutsaert DL. Systolic and Diastolic Heart Failure Are Overlapping Phenotypes Within the Heart Failure Spectrum. *Circulation*. 2011;123(18):1996-2005.

(88) Vinereanu D, Lim PO, Frenneaux MP, Fraser AG. Reduced myocardial velocities of left ventricular long-axis contraction identify both systolic and diastolic heart failure-a comparison with brain natriuretic peptide. European Journal of Heart Failure. 2005;7(4):512-519.

(89) Chatterjee K. Pathophysiology of systolic and diastolic heart failure. *Medical Clinics of North America.* 2012;96(5):891-899.

(90) Owan TE, Hodge DO, Herges RM, Jacobsen SJ, Roger VL, Redfield MM. Trends in prevalence and outcome of heart failure with preserved ejection fraction. *New England Journal of Medicine*. 2006;355(3):251-259.

(91) Lee DS, Gona P, Vasan RS, Larson MG, Benjamin EJ, Wang TJ, et al. Relation of Disease Pathogenesis and Risk Factors to Heart Failure With Preserved or Reduced Ejection Fraction: Insights From the Framingham Heart Study of the National Heart, Lung, and Blood Institute. *Circulation*. 2009;119(24):3070-3077.

(92) Komamura K. Similarities and Differences between the Pathogenesis and Pathophysiology of Diastolic and Systolic Heart Failure. *Cardiology Research and Practice*. 2013;2013(2013): http://dx.doi.org/10.1155/2013/824135

(93) McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Bohm M, Dickstein K, et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *European Journal of Heart Failure*. 2012;14(8):803-869.

(94) National Clinical Guideline Centre (UK). *Chronic heart failure: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care*. NICE guidelines [CG108]. London: National Institute for Health and Care Excellence; 2010.

(95) Bursi F, Weston SA, Redfield MM, Jacobsen SJ, Pakhomov S, Nkomo VT, et al. Systolic and diastolic heart failure in the community. *Journal of the American Medical Association*. 2006;296(18):2209-2216.

(96) Frans H Rutten, Karel G M Moons, Maarten-Jan M Cramer, Diederick E Grobbee, Nicolaas P A Zuithoff, an-Willem J Lammers, et al. Recognising heart failure in elderly patients with stable chronic obstructive pulmonary disease in primary care: cross sectional diagnostic study. *The BMJ*. 2005;331(7529):1379.

(97) Mant J, Doust J, Roalfe A, Barton P, Cowie MR, Glasziou P, et al. Systematic review and individual patient data meta-analysis of diagnosis of heart failure, with modelling of implications of different diagnostic strategies in primary care. *Health Technology Assessment*. 2009;13(32):1-207, iii.

(98) Johansson S, Wallander M, Ruigómez A, Rodríguez LAG. Incidence of newly diagnosed heart failure in UK general practice. *European Journal of Heart Failure*. 2001;3(2):225-231.

(99) Remes J, Miettinen H, Reunanen A, Pyorala K. Validity of clinical diagnosis of heart failure in primary health care. *European Heart Journal*. 1991;12(3):315-321.

(100) Davies M, Hobbs F, Davis R, Kenkre J, Roalfe AK, Hare R, et al. Prevalence of left-ventricular systolic dysfunction and heart failure in the Echocardiographic Heart of England Screening study: a population based study. *The Lancet*. 2001;358(9280):439-444.

(101) Cowie MR, Wood DA, Coats AJ, Thompson SG, Poole-Wilson PA, Suresh V, et al. Incidence and aetiology of heart failure; a population-based study. *European Heart Journa*l. 1999;20(6):421-428.

(102) National Institute of Clinical Excellence. *Chronic heart failure: national clinical guideline on diagnosis and management in primary and secondary care*. London: NICE guidelines [CG5]. 2003.

(103) NHS Employers. 2015/16 *General Medical Services (GMS) contract Quality and Outcomes Framework (QOF) Guidance for GMS contract 2015/16*. London: NHS Employers; 2015.

(104) World Health Organisation. *The global burden of disease: 2004 update*. WHO;2008.

(105) British Heart Foundation Statistics website (online). *Incidence of heart failure, by sex and age, England, Scotland, Wales and Northern Ireland 2011 (Table 2.7)*. Available at: http://www.bhf.org.uk/publications/view-publication.aspx?ps=1002097, 2013.

(106) Ho KK, Anderson KM, Kannel WB, Grossman W, Levy D. Survival after the onset of congestive heart failure in Framingham Heart Study subjects. *Circulation*. 1993;88(1):107-115.

(107) Majeed A, Williams J, de Lusignan S, Chan T. Management of heart failure in primary care after implementation of the National Service Framework for Coronary Heart Disease: a cross-sectional study. *Public Health*. 2005;119(2):105-111.

(108) British Heart Foundation Statistics Website (online). *Prevalence of heart failure, by sex and age, England, Scotland, Wales and Northern Ireland 2011 (Table 2.18)*. http://www.bhf.org.uk/publications/view-publication.aspx?ps=1002097. Available at: http://www.bhf.org.uk/publications/view-publication.aspx?ps=1002097, 2013.

(109) Hobbs FD, Korewicki J, Cleland JG, Eastaugh J, Freemantle N, IMPROVEMENT Investigators. The diagnosis of heart failure in European primary care: The IMPROVEMENT Programme survey of perception and practice. *European Journal of Heart Failure*. 2005;7(5):768-779.

(110) Mosterd A, Hoes AW. Clinical epidemiology of heart failure. *Heart*. 2007;93(9):1137-1146.

(111) Wong CY, Chaudhry SI, Desai MM, Krumholz HM. Trends in comorbidity, disability, and polypharmacy in heart failure. *American Journal of Medicine*. 2011;124(2):136-143.

(112) Hobbs FD, Roalfe AK, Davis RC, Davies MK, Hare R, Midlands Research Practices Consortium (MidReC). Prognosis of all-cause heart failure and borderline left ventricular systolic dysfunction: 5 year mortality follow-up of the Echocardiographic Heart of England Screening Study (ECHOES). *European Heart Journal.* 2007;28(9):1128-1134.

(113) Westlake C, Dracup K, Creaser J, Livingston N, Heywood JT, Huiskes BL, et al. Correlates of health-related quality of life in patients with heart failure. *Heart Lung*. 2002;31(2):85-93.

(114) Cowie MR, Komajda M, Murray-Thomas T, Underwood J, Ticho B, POSH Investigators. Prevalence and impact of worsening renal function in patients hospitalized with decompensated heart failure: results of the prospective outcomes study in heart failure (POSH). *European Heart Journal*. 2006;27(10):1216-1222.

(115) Komajda M, Hanon O, Hochadel M, Lopez-Sendon JL, Follath F, Ponikowski P, et al. Contemporary management of octogenarians hospitalized for heart failure in Europe: Euro Heart Failure Survey II. *European Heart Journal*. 2009;30(4):478-486.

(116) Azevedo A, Bettencourt P, Alvelos M, Martins E, Abreu-Lima C, Hense H-, et al. Health-related quality of life and stages of heart failure. *International Journal of Cardiology*. 2008;129(2):238-244.

(117) Murphy NF, Simpson CR, McAlister FA, Stewart S, MacIntyre K, Kirkpatrick M, et al. National survey of the prevalence, incidence, primary care burden, and treatment of heart failure in Scotland. *Heart*. 2004;90(10):1129-1136.

(118) Hospital Episode Statistics (online). 'Primary diagnosis: 3 character, 2011–12'. HES online website. Accessed . Available at: http://www.hesonline.nhs.uk/Ease/ContentServer?siteID=1937&categoryID=203, 2013.

(119) Petersen S, Rayner M, Wolstenholme J. *Coronary heart disease statistics: heart failure supplement*. London: British Heart Foundation;2002.

(120) Stewart S, Ekman I, Ekman T, Oden A, Rosengren A. Population impact of heart failure and the most common forms of cancer: A study of 1 162 309 hospital cases in Sweden (1988 to 2004). *Circulation: Cardiovascular Quality and Outcomes.* 2010;3(6):573-580.

(121) Buck HG, Akbar JA, Zhang SJ, Bettger JAP. Measuring Comorbidity in Cardiovascular Research: A Systematic Review. *Nursing Research and Practice.* 2013:11.

(122) Krum H, Gilbert RE. Demographics and concomitant disorders in heart failure. *The Lancet*. 2003;362(9378):147-158.

(123) Lee DS, Austin PC, Stukel TA, Alter DA, Chong A, Parker JD, et al. "Dose-dependent" Impact of Recurrent Cardiac Events on Mortality in Patients with Heart Failure. *American Journal of Medicine*. 2009;122(2):162-169.e1.

(124) Gheorghiade M, Flaherty JD, Fonarow GC, Desai RV, Lee R, McGiffin D, et al. Coronary artery disease, coronary revascularization, and outcomes in chronic advanced systolic heart failure. International Journal of Cardiology. 2011;151(1):69-75.

(125) van der Wel MC, Jansen RWMM, Bakx JC, Bor HHJ, OldeRikkert MGM, van Weel C. Non-cardiovascular co-morbidity in elderly patients with heart failure outnumbers cardiovascular co-morbidity. *European Journal of Heart Failure.* 2007;9(6-7):709-715.

(126) Dahlstrom U. Frequent non-cardiac comorbidities in patients with chronic heart failure. *European Journal of Heart Failure.* 2005;7(3):309-316.

(127) Lang CC, Mancini DM. Non-cardiac comorbidities in chronic heart failure. *Heart.* 2007;93(6):665-671.

(128) Gambassi G, Forman DE, Lapane KL, Mor V, Sgadari A, Lipsitz LA, et al. Management of heart failure among very old persons living in long-term care: has the voice of trials spread? The SAGE Study Group. American Heart Journal. 2000;139(1 Pt 1):85-93.

(129) Braunstein JB, Anderson GF, Gerstenblith G, Weller W, Niefeld M, Herbert R, et al. Noncardiac comorbidity increases preventable hospitalizations and mortality among Medicare beneficiaries with chronic heart failure. *Journal of the American College of Cardiology*. 2003;42(7):1226-1233.

(130) Saczynski JS, Go AS, Magid DJ, Smith DH, McManus DD, Allen L, et al. Patterns of comorbidity in older adults with heart failure: The cardiovascular research network PRESERVE study. *Journal of the American Geriatrics Society.* 2013;61(1):26-33.

(131) van Deursen VM, Urso R, Laroche C, Damman K, Dahlström U, Tavazzi L, et al. Co-morbidities in patients with heart failure: an analysis of the European Heart Failure Pilot Survey. *European Journal of Heart Failure*. 2014;16(1):103-111.

(132) Ahluwalia SC, Gross CP, Chaudhry SI, Ning YM, Leo-Summers L, Van Ness P, et al. Impact of comorbidity on mortality among older persons with advanced heart failure. *Journal of General Internal Medicine*. 2012;27(5):513-519.

(133) Ather S, Chan W, Bozkurt B, Aguilar D, Ramasubbu K, Zachariah AA, et al. Impact of noncardiac comorbidities on morbidity and mortality in a predominantly male population with heart failure and preserved versus reduced ejection fraction. *Journal of the American College of Cardiology*. 2012;59(11):998-1005.

(134) Lee CS, Bidwell JT, Denfeld QE, Creber RM, Gelow JM, Mudd JO, et al. Comorbid illness profiles predict greater costs: An analysis of 417,477 heart failure admissions. *Journal of Cardiac Failure*. Conference Publication: (var.pagings). 2012;18(8 SUPPL. 1):S76.

(135) Sturm H, Haaijer-Ruskamp F, Veeger N, Balje-Volkers C., Swedberg K, van Gilst W. The relevance of comorbidities for heart failure treatment in primary care: A European survey. *European Journal of Heart Failure*. 2006;8:31-37.

(136) Blinderman CD, Homel P, Billings JA, Portenoy RK, Tennstedt SL. Symptom distress and quality of life in patients with advanced congestive heart failure. *Journal of Pain and Symptom Management*. 2008;35(6):594-603.

(137) Franzen K, Saveman BI, Blomqvist K. Predictors for health related quality of life in persons 65 years or older with chronic heart failure. *European Journal of Cardiovascular Nursing*. 2007;6(2):112-120.

(138) Oudejans I, Mosterd A, Zuithoff NP, Hoes AW. Comorbidity Drives Mortality in Newly Diagnosed Heart Failure: A Study Among Geriatric Outpatients. *Journal of Cardiac Failure*. 2012;18(1):47-52.

(139) Mogensen UM, Ersboll M, Andersen M, Andersson C, Hassager C, Torp-Pedersen C, et al. Clinical characteristics and major comorbidities in heart failure patients more than 85 years of age compared with younger age groups. *European Journal of Heart Failure*. 2011;13(11):1216-1223.

(140) Mentz RJ, Kelly JP, von Lueder TG, Voors AA, Lam CSP, Cowie MR, et al. Noncardiac Comorbidities in Heart Failure With Reduced Versus Preserved Ejection Fraction. *Journal of the American College of Cardiology*. 2014;64(21):2281-2293.

(141) Henkel DM, Redfield MM, Weston SA, Gerber Y, Roger VL. Death in heart failure: a community perspective. *Circulation: Heart Failure*. 2008 Jul;1(2):91-97.

(142) Marechaux S, Six-Carpentier MM, Bouabdallaoui N, Montaigne D, Bauchart JJ, Mouquet F, et al. Prognostic importance of comorbidities in heart failure with preserved left ventricular ejection fraction. *Heart Vessels.* 2011;26(3):313-320.

(143) Babayan ZV, McNamara RL, Nagajothi N, Kasper EK, Armenian HK, Powe NR, et al. Predictors of cause-specific hospital readmission in patients with heart failure. *Clinical Cardiology.* 2003;26(9):411-418.

(144) Chun S, Tu JV, Wijeysundera HC, Austin PC, Wang X, Levy D, et al. Lifetime analysis of hospitalizations and survival of patients newly admitted with heart failure. *Circulation: Heart Failure*. 2012;5(4):414-421.

(145) Smith GL, Lichtman JH, Bracken MB, Shlipak MG, Phillips CO, DiCapua P, et al. Renal Impairment and Outcomes in Heart Failure. Systematic Review and Meta-Analysis. *Journal of the American College of Cardiology*. 2006;47(10):1987-1996.

(146) Damman K, Valente MAE, Voors AA, O'Connor CM, van Veldhuisen DJ, Hillege HL. Renal impairment, worsening renal function, and outcome in patients with heart failure: an updated meta-analysis. *European Heart Journal.* 2014;35(7):455-469.

(147) Arnaudis B, Lairez O, Escamilla R, Fouilloux A, Fournier P, Monteil B, et al. Impact of chronic obstructive pulmonary disease severity on symptoms and prognosis in patients with systolic heart failure. *Clinical Research in Cardiology.* 2012;101(9):717-726.

(148) Kwon BJ, Kim DB, Jang SW, Yoo KD, Moon KW, Shim BJ, et al. Prognosis of heart failure patients with reduced and preserved ejection fraction and coexistent chronic obstructive pulmonary disease. *European Journal of Heart Failure*. 2010;12(12):1339-1344.

(149) Gustafsson I, Hildebrandt P, Seibaek M, Melchior T, Torp-Pedersen C, Kober L, et al. Long-term prognosis of diabetic patients with myocardial infarction: relation to antidiabetic treatment regimen. The TRACE Study Group. *European Heart Journal.* 2000;21(23):1937-1943.

(150) Savage MP, Krolewski AS, Kenien GG, Lebeis MP, Christlieb AR, Lewis SM. Acute myocardial infarction in diabetes mellitus and significance of congestive heart failure as a prognostic factor. *American Journal of Cardiology*. 1988;62:665-669.

(151) Belziti CA, Bagnati R, Ledesma P, Vulcano N, Fernandez S. Worsening renal function in patients admitted with acute decompensated heart failure: incidence, risk factors and prognostic implications. Revista Española de Cardiología. 2010;63(3):294-302.

(152) Blair JEA, Pang PS, Schrier RW, Metra M, Traver B, Cook T, et al. Changes in renal function during hospitalization and soon after discharge in patients admitted for worsening heart failure in the placebo group of the EVEREST trial. *European Heart Journal.* 2011;32(20):2563-2572.

(153) Forman DE, Butler J, Wang Y, Abraham WT, O'Connor CM, Gottlieb SS, et al. Incidence, predictors at admission, and impact of worsening renal function among patients hospitalized with heart failure. *Journal of the American College of Cardiology*. 2004;43(1):61-67.

(154) Gottlieb SS, Abraham W, Butler J, Forman DE, Loh E, Massie BM, et al. The prognostic importance of different definitions of worsening renal function in congestive heart failure. *Journal Cardiac Failure*. 2002;8(3):136-141.

(155) Kociol RD, Greiner MA, Hammill BG, Phatak H, Fonarow GC, Curtis LH, et al. Long-term outcomes of medicare beneficiaries with worsening renal function during hospitalization for heart failure. *American Journal of Cardiology.* 2010;105(12):1786-1793.

(156) Damman K, Jaarsma T, Voors AA, Navis G, Hillege HL, Van Veldhuisen DJ. Both in- and out-hospital worsening of renal function predict outcome in patients with heart failure: Results from the Coordinating Study Evaluating Outcome of Advising and Counseling in Heart Failure (COACH). *European Journal of Heart Failure.* 2009;11(9):847-854.

(157) Damman K, Navis G, Voors AA, Asselbergs FW, Smilde TD, Cleland JG, et al. Worsening renal function and prognosis in heart failure: systematic review and meta-analysis. *Journal of Cardiac Failure*. 2007;13(8):599-608.

(158) Pokhrel N, Maharjan N, Dhakal B, Arora RR. Cardiorenal syndrome: A literature review. *Experimental & Clinical Cardiology*. 2008;13(4):165-170.

(159) Brown AM, Cleland JG. Influence of concomitant disease on patterns of hospitalization in patients with heart failure discharged from Scottish hospitals in 1995. *European Heart Journal.* 1998;19(7):1063-1069.

(160) Thanassoulis G, Brophy JM, Richard H, Pilote L. Gout, allopurinol use, and heart failure outcomes. *Archives Internal Medicine*. 2010;170(15):1358-1364.

(161) Philbin EF, DiSalvo TG. Prediction of hospital readmission for heart failure: development of a simple risk score based on administrative data. *Journal of the American College of Cardiology*. 1999;33(6):1560-1566.

(162) Lee CS, Denfeld QE, Bidwell JT, Creber RM, Gelow JM, Mudd JO, et al. Comorbid illness profiles predict greater length of stay for women: An analysis of 93,242 heart failure admissions. *Circulation*. 2012;126 (21 suppl. 1.)

(163) Heiat A, Gross CP, Krumholz HM. Representation of the elderly, women, and minorities in heart failure clinical trials. Archives of Internal Medicine. 2002;162(15):1682-1688.

(164) Yamokoski LM, Hasselblad V, Moser DK, Binanay C, Conway GA, Glotzer JM, et al. Prediction of Rehospitalization and Death in Severe Heart Failure by Physicians and Nurses of the ESCAPE Trial. *Journal of Cardiac Failure*. 2007;13(1):8-13.

(165) Allen LA, Yager JE, Funk MJ, Levy WC, Tulsky JA, Bowers MT, et al. Discordance between patient-predicted and model-predicted life expectancy among ambulatory patients with heart failure. Journal of the American Medical Association. 2008;299(21):2533-2542.

(166) Chattoo S, Atkin KM. Extending specialist palliative care to people with heart failure: semantic, historical and practical limitations to policy guidelines. *Social Science and Medicine.* 2009;69(2):147-153.

(167) Harding R, Selman L, Beynon T, Hodson F, Coady E, Read C, et al. Meeting the communication and information needs of chronic heart failure patients. *Journal of Pain and Symptom Management*. 2008;36(2):149-156.

(168) Selman L, Harding R, Beynon T, Hodson F, Coady E, Hazeldine C, et al. Improving end-of-life care for patients with chronic heart failure: "Let's hope it'll get better, when I know in my heart of hearts it won't". *Heart* 2007;93(8):963-967.

(169) Ketchum ES, Levy WC. Multivariate risk scores and patient outcomes in advanced heart failure. *Congestive Heart Failure*. 2011;17(5):205-212.

(170) Vazquez R, Bayes-Genis A, Cygankiewicz I, Pascual-Figal D, Grigorian-Shamagian L, Pavon R, et al. The MUSIC Risk score: a simple method for predicting mortality in ambulatory patients with chronic heart failure. *European Heart Journal*. 2009;30(9):1088-1096.

(171) Kearney MT, Nolan J, Lee AJ, Brooksby PW, Prescott R, Shah AM, et al. A prognostic index to predict long-term mortality in patients with mild to moderate chronic heart failure stabilised on angiotensin converting enzyme inhibitors. *European Journal of Heart Failure.* 2003;5(4):489-497.

(172) Wedel H, McMurray JJ, Lindberg M, Wikstrand J, Cleland JG, Cornel JH, et al. Predictors of fatal and non-fatal outcomes in the Controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA): incremental value of apolipoprotein A-1, high-sensitivity C-reactive peptide and N-terminal pro B-type natriuretic peptide. *European Journal of Heart Failure.* 2009;11(3):281-291.

(173) Alla F, Briancon S, Juilliere Y, Mertes PM, Villemot JP, Zannad F. Differential clinical prognostic classifications in dilated and ischemic advanced heart failure: the EPICAL study. *American Heart Journal*. 2000;139(5):895-904.

(174) Fonarow GC, Adams J, K.F, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: Classification and regression tree analysis. *Journal of the American Medical Association*. 2005;293(5):572-580.

(175) Lee DS, Austin PC, Rouleau JL, Liu PP, Naimark D, Tu JV. Predicting Mortality among Patients Hospitalized for Heart Failure: Derivation and Validation of a Clinical Model. *Journal of the American Medical Association*. 2003;290(19):2581-2587.

(176) Felker GM, Leimberger JD, Califf RM, Cuffe MS, Massie BM, Adams J, et al. Risk stratification after hospitalization for decompensated heart failure. *Journal of Cardiac Failure.* 2004;10(6):460-466.

(177) Levy WC, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, et al. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation*. 2006;113(11):1424-1433.

(178) Manzano L, Babalis D, Roughton M, Shibata M, Anker SD, Ghio S, et al. Predictors of clinical outcomes in elderly patients with heart failure. *European Journal of Heart Failure*. 2011;13(5):528-536.

(179) Pocock SJ, Wang D, Pfeffer MA, Yusuf S, McMurray JJ, Swedberg KB, et al. Predictors of mortality and morbidity in patients with chronic heart failure. *European Heart Journal.* 2006;27(1):65-75.

(180) Villacorta H, Mesquita ET. Prognostic factors in patients with congestive heart failure. *Arquivos Brasileiros de Cardiologia*. 1999;72(3):343-362.

(181) Lainscak M, Hodoscek LM, Dungen HD, Rauchhaus M, Doehner W, Anker SD, et al. The burden of chronic obstructive pulmonary disease in patients hospitalized with heart failure. *The Wiener klinische Wochenschrift.* 2009;121(9-10):309-313.

(182) Ross JS, Mulvey GK, Stauffer B,et al. Statistical models and patient predictors of readmission for heart failure: A systematic review. *Archives of Internal Medicine*. 2008;168(13):1371-1386.

(183) Betihavas V, Davidson PM, Newton PJ, Frost SA, Macdonald PS, Stewart S. What are the factors in risk prediction models for rehospitalisation for adults with chronic heart failure? *Australian Critical Care*. 2012;25(1):31-40.

(184) Alba AC, Agoritsas T, Jankowski M, Courvoisier D, Walter SD, Guyatt GH, et al. Risk Prediction Models for Mortality in Ambulatory Patients With Heart Failure: A Systematic Review. *Circulation: Heart Failure*. 2013;6(5):881-889.

(185) Aaronson KD, Schwartz JS, Chen TM, Wong KL, Goin JE, Mancini DM. Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation. *Circulation.* 1997;95(12):2660-2667.

(186) Ouwerkerk W, Voors AA, Zwinderman AH. Factors Influencing the Predictive Power of Models for Predicting Mortality and/or Heart Failure Hospitalization in Patients With Heart Failure. *Journal of the American College of Cardiology: Heart Failure*. 2014;2(5):429-436.

(187) Rahimi K, Bennett D, Conrad N, Williams TM, Basu J, Dwight J, et al. Risk Prediction in Patients With Heart FailureA Systematic Review and Analysis. *Journal of the American College of Cardiology: Heart Failure*. 2014;2(5):440-446.

(188) Pocock SJ, Ariti CA, McMurray JJ, Maggioni A, Kober L, Squire IB, et al. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *European Heart Journal*. 2013;34(19):1404-1413.

(189) Senni M, Parrella P, De Maria R, Cottini C, Bohm M, Ponikowski P, et al. Predicting heart failure outcome from cardiac and comorbid conditions: The 3C-HF score. *International Journal of Cardiology*. 2013;163(2):206-211.

(190) Farkas J, Nabb S, Zaletel-Kragelj L, Cleland JG, Lainscak M. Self-rated health and mortality in patients with chronic heart failure. *European Journal of Heart Failure*. 2009;11(5):518-524.

(191) Solomon SD, Anavekar N, Skali H, McMurray JJ, Swedberg K, Yusuf S, et al. Influence of ejection fraction on cardiovascular outcomes in a broad spectrum of heart failure patients. *Circulation*. 2005;112(24):3738-3744.

(192) Komajda M, Carson PE, Hetzel S, McKelvie R, McMurray J, Ptaszynska A, et al. Factors associated with outcome in heart failure with preserved ejection fraction: findings from the Irbesartan in Heart Failure with Preserved Ejection Fraction Study (I-PRESERVE). *Circulation: Heart Failure*. 2011;4(1):27-35.

(193) Subramanian D, Subramanian V, Deswal A, Mann D. New Predictive Models of Heart Failure Mortality Using Time-Series Measurements and Ensemble Models. *Circulation.Heart Failure*. 2011;4(4):456-462.

(194) Giolo SR, Krieger JE, Mansur AJ, Pereira AC. Survival analysis of patients with heart failure: Implications of time-varying regression effects in modeling mortality. *PLoS ONE. 2012;7(6):* DOI:10.1371/journal.pone.0037392

(195) Sy RW, Chawantanpipat C, Richmond DR, Kritharides L. Development and validation of a time-dependent risk model for predicting mortality in infective endocarditis. *European Heart Journal*. 2011;32(16):2016-2026.

(196) van Staa TP, Leufkens HG, Abenhaim L, Zhang B, Cooper C. Oral corticosteroids and fracture risk: relationship to daily and cumulative doses. *Rheumatology (Oxford).* 2000;39(12):1383-1389.

(197) General Practitioners Committee. The new GMS contract explained. London: BMA; 2004.

(198) Zlotnick C, Tam TW, Soman LA. Life course outcomes on mental and physical health: the impact of foster care on adulthood. *American Journal of Public Health*. 2012;102(3):534-540.

(199) Woodward LJ, Fergusson DM. Life course outcomes of young people with anxiety disorders in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2001;40(9):1086-1093.

(200) Doll R. Proof of causality: deduction from epidemiological observation. *Perspectives in Biology and Medicine*. 2002;45(4):499-515.

(201) Copi I, Cohen C, Flage D. *Essentials of Logic*. 2nd ed. New Jersey: Pearson Education; 2007.

(202) Popper K. *The logic of scientific discovery*. London: Hutchinson; 1997.

(203) Hill A. *The environment and disease: Association or causation? Proceedinigs of the Royal Society of Medicine*.1965;58:295-300.

(204) Coggion D, Rose G, Barker D. *Epidemiology for the Uninitiated*. 5th ed. London: BMJ Publishing Group; 2003.

(205) Wolbers M, Koller MT, Stel VS, Schaer B, Jager KJ, LeffondrÃ K, et al. Competing risks analyses: objectives and approaches. *European Heart Journal*. 2014;35(42):2936-2941.

(206) Fine J, Gray R. A proportional hazards model for the subdistribution of a competing risk. *Journal American Statistical Association*. 1999;94::496-509.

(207) McNeil D. *Epidemiological Research Methods*. New York: John Wiley ans Sons; 1996.

(208) Somerville M, Kumaran K, Anderson R. *Public Health and Epidemiology at a Glance*. Oxford: Wiley-Blackwell; 2012.

(209) Gordis L. *Epidemiology*. 4th ed. Philadelphia: Saunders, Elsevier; 2009.

(210) Sibbald B, Roland M. Understanding controlled trials. Why are randomised controlled trials important? The BMJ. 1998;316(7126):201-201.

(211) Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. The BMJ. 2010;340:b5087

(212) Ho AM-, Dion PW, Ng CSH, Karmakar MK. Understanding immortal time bias in observational cohort studies. *Anaesthesia*. 2013;68(2):126-130.

(213) Schlesselman J. *Case Control Studies; design, conduct, analysis*. New York: Oxford University Press; 1982.

(214) Vandenbroucke JP, Pearce N. Case–control studies: basic concepts. *International Journal of Epidemiology*. 2012;41(5):1480-1489.

(215) Pearce N. What does the odds ratio estimate in a case-control study? *International Journal of Epidemiology*. 1993;22(6):1189-1192.

(216) Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. *Journal of Clinical Epidemiology*. 1999;52(12):1165-1172.

(217) Sato T. Risk ratio estimation in case-cohort studies. Environmental Health Perspectives. 1994;102:53-56.

(218) Langholz B, Richardson D. Are nested case-control studies biased? *Epidemiology*. 2009;20(3):321-329.

(219) Breslow, NE. Statistics in epidemiology: the case-control study. Journal American Statistics Association. 1996;91(433):14-28.

(220) Essebag V, Platt R, Abrahamowicz M, Pilote L. Comparison of nested case-control and survival analysis methodologies for analysis of time-dependent exposure. *BMC Medical Research Methodology.* 2005;5(1):1-6.

(221) Houwelingen H, Putter H. *Dynamic Prediction in Clinical Survival Analysis*. Boca Raton, FL: CRC Press; 2012.

(222) Hancock G, Mueller R editors. *The Reviewer's Guide to Quantitative Methods in the Social Sciences*. Abingdon, Oxfordshire: Routledge; 2010.

(223) Berry SD, Ngo L, Samelson EJ, Kiel DP. Competing risk of death: an important consideration in studies of older adults. *Journal of the American Geriatrics Society*. 2010;58(4):783-787.

(224) Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Statistics in Medicine*. 1999;18(6):695-706.

(225) Steyerberg E. *Clinical prediction models: A Practical Approach to Development, Validation, and Updating.* Springer; 2008.

(226) Kleinbaum D, Klein M. *Logistic regression(statistics for biology and health)*. 3rd ed. New York: Springer-Verlag New York Inc.; 2010.

(227) Gould W. "Interpreting logistic regression in all its forms". *Stata Technical Bulletin*. 2000;9(53):19-29.

(228) Park HA. An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean Academy of Nursing*. 2013;43(2):154-164.

(229) Petrie A, Sabin C. *Medical Statistics at a Glance*. 3rd ed. 2009: Wiley-Blackwell.

(230) Bowers D. *Medical Statistics from Scratch*. West Sussex: John Wiley and Sons; 2002.

(231) Bloom MS, Schisterman EF, Hediger ML. The use and misuse of matching in case-control studies: the example of PCOS. *Fertility and Sterility*. 2007;88(3):707-710.

(232) Yarnell J. *Epidemiology and prevention: a systems based approach*. 1st ed. United States: Radcliffe Publishing; 2010.

(233) David M, Ware R, Donald M, Alati R. Assessing generalisability through the use of disease registers: findings from a diabetes cohort study. *The BMJ Open*. 2011;1(1):e000078-2011-000078.

(234) Newton J, Garner S. *Disease Registers in England: A report commissioned by the Department of Health Policy Research Programme in support of the White Paper entitled Saving Lives: Our Healthier Nation*. Oxford: Institute of Health Sciences 2002.

(235) Dugmore K, Furness P, Leventhal B, Moy C. Beyond the 2011 Census in the United Kingdom With an international perspective. *International Journal of Market Research*. 2011;53(5):619-650.

(236) Health and Social Care Information Centre. *Health Survey for England 2013; volume 2, methods and documentation*. 2014; Available at: http://www.hscic.gov.uk/catalogue/PUB16076/HSE2013-Methods-and-docs.pdf. Accessed 4th February, 2015.

(237) Williams T, van Staa T, Puri S, Eaton S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Therapeutic Advances in Drug Safety*. 2012;3(2):89-99.

(238) Tate AR, Beloff N, Al-Radwan B, Wickson J, Puri S, Williams T, et al. Exploiting the potential of large databases of electronic health records for research using rapid search algorithms and an intuitive query interface. *The Journal of the American Medical Informatics*. 2014;21(2):292-298.

(239) Rushton CA, Satchithananda DK, Jones PW, Kadam UT. Non-cardiovascular comorbidity, severity and prognosis in non-selected heart failure populations: A systematic review and meta-analysis. *International Journal of Cardiology*. 2015;196:98-106.

(240) Krumholz HM, Wang Y, Mattera JA, Wang Y, Han LF, Ingber MJ, et al. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure. *Circulation*. 2006;113(13):1693-1701.

(241) Harry Hemingway, Richard D Riley, Douglas G Altman. Ten steps towards improving prognosis research. *The BMJ*. 2009;339.

(242) Hayden J, Cote P, Steenstra I, Bombardier C. Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. *Journal of Clinical Epidemiology*. 2008;61:552-560.e

(243) Douglas G Altman. Systematic reviews of evaluations of prognostic variables. *The BMJ*. 2001;323(7306):224-228.

(244) Stroup DF, Berlin JA, Morton SC,et al. Meta-analysis of observational studies in epidemiology: A proposal for reporting. Journal of the American Medical Association 2000;283(15):2008-2012.

(245) Costantino G, Montano N, Casazza G. When should we change our clinical practice based on the results of a clinical study? Searching for evidence: PICOS and PubMed. *Internal and Emergency Medicine*. 2015;10(4):525-527.

(246) Burch M. Heart failure in the young. *Heart*. 2002;88(2):198-202.

(247) Stephen Gillam, Nicholas Steel. The Quality and Outcomes Framework—where next? *The BMJ*. 2013;346.

(248) Sachdeva A, Horwich TB, Fonarow GC. Comparison of usefulness of each of five predictors of mortality and urgent transplantation in patients with advanced heart failure. *American Journal of Cardiology*. 2010;106(6):830-835.

(249) Bruch C, Sindermann J, Breithardt G, Gradaus R. Prevalence and prognostic impact of comorbidities in heart failure patients with implantable cardioverter-defibrillator. *Europace*. 2007;9(8):681-686.

(250) Saxon LA, Bristow MR, Boehmer J, Krueger S, Kass DA, Marco TD, et al. Predictors of sudden cardiac death and appropriate shock in the Comparison of Medical Therapy, Pacing, and Defibrillation in Heart Failure (COMPANION) Trial. *Circulation*. 2006;114(25):2766-2772.

(251) McAlister FA, Ezekowitz J, Tarantini L, Squire I, Komajda M, Bayes-Genis A, et al. Renal dysfunction in patients with heart failure with preserved versus reduced ejection fraction: impact of the new Chronic Kidney Disease-Epidemiology Collaboration Group formula. *Circulation: Heart Failure*. 2012;5(3):309-314.

(252) Desai AS. Heart Failure With Preserved Ejection FractionTime for a New Approach? *Journal of the American College of Cardiology*. 2013;62(4):272-274.

(253) World Health Organization. *Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus*, WHO/NMH/CHP/CPM/11.1. 2011; Accessed 13th March, 2014.

(254) The National Collaborating Centre for Chronic Conditions. *Chronic Kidney Disease: National clinical guideline for early identification and management in adults in primary and secondary care*. London: Royal College Physicians; 2008.

(255) De Isla LP, Zamorano J, Hernandez N, Contreras L, Rodrigo JL, Almeria C, et al. Prognostic factors and predictors of in-hospital mortality of patients with heart failure with preserved left ventricular ejection fraction. *Journal of Cardiovascular Medcine.* 2008;9(10):1011-1015.

(256) Greenberg G, Cohen E, Garty M, Iakobishvili Z, Sandach A, Behar S, et al. Outcomes of acute heart failure associated with acute coronary syndrome versus other causes. *Acute Cardiac Care.* 2011;13(2):87-92.

(257) Otero-Ravina F, Grigorian-Shamagian L, Fransi-Galiana L, Nazara-Otero C, Fernandez-Villaverde JM, del Alamo-Alonso A, et al. Morbidity and mortality among heart failure patients in Galicia, N.W. Spain: the GALICAP Study. *International Journal of Cardiology*. 2009;136(1):56-63.

(258) Martinez-Selles M, Martinez E, Cortes M, Prieto R, Gallego L, Fernandez-Aviles F. Determinants of long-term survival in patients hospitalized for heart failure. *Journal of Cardiovascular Medicine*. 2010;11(3):164-169.

(259) Senni M, Santilli G, Parrella P, De Maria R, Alari G, Berzuini C, et al. A Novel Prognostic Index to Determine the Impact of Cardiac Conditions and Co-Morbidities on One-Year Outcome in Patients With Heart Failure. *American Journal of Cardiology*. 2006;98(8):1076-1082.

(260) Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technology Assessment*. 2003;7(41):1-90.

(261) Lefebvre C, Manheimer E, Glanville J. Searching for Studies. In: Higgins J, Green S, editors. *Cochrane Handbook for Systemativ Reviews of Interventions*. 1st ed.: John Wiley and Sons; 2008. p. 95-150.

(262) David Woods, Kate Trewheellar. Medline and Embase complement each other in literature searches. *The BMJ*. 1998;316(7138):1166.

(263) Sampson M, Barrowman NJ, Moher D, Klassen TP, Pham B, Platt R, et al. Should meta-analysts search Embase in addition to Medline? *Journal of Clinical Epidemiology*. 2003 10;56(10):943-955.

(264) Damarell R, Tieman J, Slade R, Davidson P.
Development of a heart failure filter for Medline: an objective approach using evidence-based clinical practice guidelines as an alternative to hand searching. 2011; Available at: http://www.biomedcentral.com/1471-2288/11/12. Accessed 5th January, 2013.

(265) Geersing G, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons K.
Search Filters for Finding Prognostic and Diagnostic Prediction Studies in Medline to Enhance Systematic Reviews.
. PLoS ONE 2012;7(2):e32844.doi:10.1371/journal.pone.0032844.

(266) Wilczynski N, Haynes R.
Optimal Search Strategies for Detecting Clinically Sound Prognostic Studies in EMBASE: An Analytic Survey. *Journal of the American Medical Informatics Association* 2005;12(4):481-485.

(267) Mulrow C, Cook D editors. *Systematic Reviews, Synthesis of Best Evidence for Health Care Decisions.*. Philadelphia: American College of Physicians; 1998.

(268) Higgins J, Deeks J. Selecting Studies and Collecting Data. In: Higgins J, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. Available from www.cochrane-handbook.org.: The Cochrane Collaboration; 2011. p. Chapter 7.

(269) Atkins D, Best D, Briss P, Eccles M, Falck-Ytter Y, Flottorp S. Grading quality of evidence and strength of recommendations. *The BMJ*. 2004;328(7454):1490.

(270) Hayden J, van der Windt D, Cartwright J, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. *Annals of Internal Medicine*. 2013;158(4):280-286.

(271) Perel P. *Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients.* 2008; Available at: http://www.bmj.com/content/336/7641/425. Accessed January, 2013.

(272) Egger M, Smith G, Altman D editors. *Systematic Reviews in Health Care, Meta-analysis in Context*. 2nd ed. London: BMJ Publishing Group; 2001.

(273) Khan K, Kunz R, Kleijnen J, Antes G. *Systematic Reviews to Support Evidence-based Medicine: how to review and apply findings of healthcare research.* London: Theroyal Society of Medicine Press; 2003.

(274) Deeks J, Higgins J, Altman D. Analysing data and undertaking meta-analyses. In: Higgins J, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]* Available from www.cochrane-handbook.org.: The Cochrane Collaboration; 2001: Chapter 9.

(275) Khan K, Kunz R, Kleijnen J, Antes G. *Systematic Reviews, To Support Evidence Based Medicine*. 2nd ed. London: Hodder Arnold; 2011.

(276) DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clinical Trials*. 1986;7(3):177-188.

(277) Kranzler J. *Statistics for the Terrified*. 3rd ed. Upper Saddle River, NJ: Prentice Hall; 2003.

(278) Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods*. 2006;11(2):193-206.

(279) da Costa BR, Juni P. Systematic reviews and meta-analyses of randomized trials: principles and pitfalls. *European Heart Journal*. 2014;35(47):3336-3345.

(280) Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI, Health Outcomes, Policy, and Economics (HOPE) Collaborative Group. Understanding heterogeneity in meta-analysis: the role of meta-regression. *International Journal of Clinical Practice*. 2009;63(10):1426-1434.

(281) Gough D, Oliver S, Thomas J. *An Introduction to Systematic Reviews*. Los Angeles: Sage; 2012.

(282) Galbraith RF. The radial plot: Graphical assessment of spread in ages. *International Journal of Radiation Applications and Instrumentation.Part D.Nuclear Tracks and Radiation Measurements*. 1990;17(3):207-214.

(283) Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology*. 2001;54(10):1046-1055.

(284) Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *The BMJ*. 1997;315(7109):629-634.

(285) Burger AJ, Tsao L, Aronson D. Prognostic impact of diabetes mellitus in patients with acute decompensated heart failure. *American Journal of Cardiology*. 2005;95(9):1117-1119.

(286) de Boer RA, Doehner W, van der Horst ICC, Anker SD, Babalis D, Roughton M, et al. Influence of Diabetes Mellitus and Hyperglycemia on Prognosis in Patients >=70 Years Old With Heart Failure and Effects of Nebivolol (Data from the Study of Effects of Nebivolol Intervention on Outcomes and Rehospitalization in Seniors With Heart Failure [SENIORS]). *American Journal of Cardiology*. 2010;106(1):78-86.e1.

(287) Gustafsson I, Brendorp B, Seibaek M, Burchardt H, Hildebrandt P, Kober L, et al. Influence of diabetes and diabetes-gender interaction on the risk of death in patients hospitalized with congestive heart failure. *Journal of the American College of Cardiology*. 2004;43(5):771-777.

(288) Issa VS, Amaral AF, Cruz FD, Ayub-Ferreira SM, Guimaraes GV, Chizzola PR, et al. Glycemia and prognosis of patients with chronic heart failure--subanalysis of the Long-term Prospective Randomized Controlled Study Using Repetitive Education at Six-Month Intervals and Monitoring for Adherence in Heart Failure Outpatients (REMADHE) trial. *American Heart Journal*. 2010;159(1):90-97.

(289) Berry C, Brett M, Stevenson K, McMurray JJ, Norrie J. Nature and prognostic importance of abnormal glucose tolerance and diabetes in acute heart failure. *Heart*. 2008;94(3):296-304.

(290) Gerstein HC, Swedberg K, Carlsson J, McMurray JJ, Michelson EL, Olofsson B, et al. The hemoglobin A1c level as a progressive risk factor for cardiovascular death, hospitalization for heart failure, or death in patients with chronic heart failure: an analysis of the Candesartan in Heart failure: Assessment of Reduction in Mortality and Morbidity (CHARM) program. *Archives of Internal Medicine*. 2008;168(15):1699-1704.

(291) MacDonald MR, Jhund PS, Petrie MC, Lewsey JD, Hawkins NM, Bhagra S, et al. Discordant short- and long-term outcomes associated with diabetes in patients with heart failure: importance of age and sex: a population study of 5.1 million people in Scotland. *Circulation: Heart failure*. 2008;1(4):234-241.

(292) Ahmed A, Aban IB, Vaccarino V, Lloyd-Jones D, Goff D, J., Zhao J, et al. A propensity-matched study of the effect of diabetes on the natural history of heart failure: variations by sex and age. *Heart*. 2007; 93(12):1584-1590.

(293) Flores-Le Roux JA, Comin J, Pedro-Botet J, Benaiges D, Puig-de Dou J, Chillaron JJ, et al. Seven-year mortality in heart failure patients with undiagnosed diabetes: An observational study. *Cardiovascular Diabetology*. 2011;10.

(294) From AM, Leibson CL, Bursi F, Redfield MM, Weston SA, Jacobsen SJ, et al. Diabetes in heart failure: prevalence and impact on outcome in the population. *American Journal of Medicine.* 2006;119(7):591-599.

(295) Greenberg BH, Abraham WT, Albert NM, Chiswell K, Clare R, Stough WG, et al. Influence of diabetes on characteristics and outcomes in patients hospitalized with heart failure: a report from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). *American Heart Journal.* 2007;154(2):277.e1-277.e8.

(296) Iversen KK, Kjaergaard J, Akkan D, Kober L, Torp-Pedersen C, Hassager C, et al. The prognostic importance of lung function in patients admitted with heart failure. *European Journal of Heart Failure.* 2010;12(7):685-691.

(297) De Blois J, Simard S, Atar D, Agewall S, Norwegian Heart Failure Registry. COPD predicts mortality in HF: the Norwegian Heart Failure Registry. *Journal of Cardiac Failure*. 2010;16(3):225-229.

(298) Macchia A, Monte S, Romero M, D'Ettorre A, Tognoni G. The prognostic influence of chronic obstructive pulmonary disease in patients hospitalised for chronic heart failure. *European Journal of Heart Failure.* 2007;9(9):942-948.

(299) Rusinaru D, Saaidi I, Godard S, Mahjoub H, Battle C, Tribouilloy C. Impact of Chronic Obstructive Pulmonary Disease on Long-Term Outcome of Patients Hospitalized for Heart Failure. *American Journal of Cardiology*. 2008;101(3):353-358.

(300) Breidthardt T, Socrates T, Noveanu M, Klima T, Heinisch C, Reichlin T, et al. Effect and clinical prediction of worsening renal function in acute decompensated heart failure. *American Journal of Cardiology.* 2011;107(5):730-735.

(301) Campbell RC, Sui X, Filippatos G, Love TE, Wahle C, Sanders PW, et al. Association of chronic kidney disease with outcomes in chronic heart failure: A propensity-matched study. *Nephrology Dialysis Transplantation*. 2009;24(1):186-193.

(302) Gotsman I, Zwas D, Planer D, Admon D, Lotan C, Keren A. The significance of serum urea and renal function in patients with heart failure. *Medicine.* 2010;89(4):197-203.

(303) Hamaguchi S, Tsuchihashi-Makaya M, Kinugawa S, Yokota T, Ide T, Takeshita A, et al. Chronic kidney disease as an independent risk for long-term adverse outcomes in patients hospitalized with heart failure in Japan. Report from the Japanese Cardiac Registry of Heart Failure in Cardiology (JCARE-CARD). *Circulation Journal*. 2009;73(8):1442-1447.

(304) Hillege HL, Nitsch D, Pfeffer MA, Swedberg K, McMurray JJ, Yusuf S, et al. Renal function as a predictor of outcome in a broad spectrum of patients with heart failure. *Circulation.* 2006;113(5):671-678.

(305) Kimura H, Hiramitsu S, Miyagishima K, Mori K, Yoda R, Kato S, et al. Cardio-renal interaction: Impact of renal function and anemia on the outcome of chronic heart failure. *Heart Vessels.* 2010;25(4):306-312.

(306) Ismailov RM, Goldberg RJ, Lessard D, Spencer FA. Decompensated heart failure in the setting of kidney dysfunction: a community-wide perspective. *Nephron Clinical Practice*. 2007;107(4):c147-55.

(307) Olandoski M, De Lima RR, Da Silva MMF, Pecoits-Filho R, Barboza AO, Erbano BO, et al. Interaction of anemia and decrease in renal function on survival of patients with heart failure. *International Journal of Cardiology*. 2012;154(3):338-340.

(308) Petretta M, Scopacasa F, Fontanella L, Carlomagno A, Baldissara M, de Simone A, et al. Prognostic value of reduced kidney function and anemia in patients with chronic heart failure. *Journal of Cardiovascular Medicine*. (Hagerstown) 2007(11):909-916.

(309) Takagi A, Iwama Y, Yamada A, Aihara K, Daida H. Estimated glomerular filtration rate is an independent predictor for mortality of patients with acute heart failure. *Journal of Cardiology*. 2010;55(3):317-321.

(310) Aronson D, Burger AJ. The Relationship Between Transient and Persistent Worsening Renal Function and Mortality in Patients With Acute Decompensated Heart Failure. *Journal of Cardiac Failure*. 2010;16(7):541-547.

(311) Maeder MT, Rickli H, Pfisterer ME, Muzzarelli S, Ammann P, Fehr T, et al. Incidence, clinical predictors, and prognostic impact of worsening renal function in elderly patients with chronic heart failure on intensive medical therapy. *American Heart Journal*. 2012;163(3):407-14, 414.e1.

(312) Waldum B, Westheim AS, Sandvik L, Flønaes B, Grundtvig M, Gullestad L, et al. Renal function in outpatients with chronic heart failure. *Journal of Cardiac Failure*. 2010;16(5):374-380.

(313) Go AS, Yang J, Ackerson LM, Lepper K, Robbins S, Massie BM, et al. Hemoglobin level, chronic kidney disease, and the risks of death and hospitalization in adults with chronic heart failure - The anemia in chronic heart failure: Outcomes and Resource Utilization (ANCHOR) Study. *Circulation*. 2006;113(23):2713-2723.

(314) Testani JM, McCauley BD, Chen J, Coca SG, Cappola TP, Kimmel SE. Clinical characteristics and outcomes of patients with improvement in renal function during the treatment of decompensated heart failure. *Journal of Cardiac Failure*. 2011;17(12):993-1000.

(315) Smith GL, Vaccarino V, Kosiborod M, Lichtman JH, Cheng S, Watnick SG, et al. Worsening renal function: what is a clinically meaningful change in creatinine during hospitalization with heart failure? *Journal of Cardiac Failure*. 2003;9(1):13-25.

(316) Davis III JM, Roger VL, Crowson CS, Kremers HM, Therneau TM, Gabriel SE. The presentation and outcome of heart failure in patients with rheumatoid arthritis differs from that in the general population. *Arthritis and Rheumatology.* 2008;58(9):2603-2611.

(317) Chaudhry SI, Wang Y, Gill TM, Krumholz HM. Geriatric conditions and subsequent mortality in older patients with heart failure. *Journal of the American College of Cardiology*. 2010;55(4):309-316.

(318) Chaudhry SI, McAvay G, Chen S, Whitson H, Newman AB, Krumholz HM, et al. Risk factors for hospital admission among older persons with newly diagnosed heart failure: Findings from the cardiovascular health study. *Journal of the American College of Cardiology*. 2013;61(6):635-642.

(319) Fernandez-Berges D, Consuegra-Sanchez L, Felix-Redondo FJ, Robles NR, Galan Montejano M, Lozano-Mera L. Clinical characteristics and mortality of heart failure. INCAex study. *Revista Clínica Española*. 2013;213(1):16-24.

(320) Dunlay SM, Redfield MM, Weston SA, Therneau TM, Hall Long K, Shah ND, et al. Hospitalizations after heart failure diagnosis a community perspective. *Journal of the American College of Cardiology*. 2009;54(18):1695-1702.

(321) Aranda J, J.M, Johnson JW, Conti JB. Current trends in heart failure readmission rates: Analysis of medicare data. *Clinical Cardiology*. 2009;32(1):47-52.

(322) Mosterd A, Cost B, Hoes AW, de Bruijne MC, Deckers JW, Hofman A, et al. The prognosis of heart failure in the general population: The Rotterdam Study. *European Heart Journal.* 2001;22(15):1318-1327.

(323) Gotsman I, Zwas D, Planer D, Azaz-Livshits T, Admon D, Lotan C, et al. Clinical Outcome of Patients with Heart Failure and Preserved Left Ventricular Function. *American Journal of Medicine.* 2008;121(11):997-1001.

(324) Hamaguchi S, Kinugawa S, Goto D, Tsuchihashi-Makaya M, Yokota T, Yamada S, et al. Predictors of long-term adverse outcomes in elderly patients over 80 years hospitalized with heart failure. - A report from the Japanese Cardiac Registry of Heart Failure in Cardiology (JCARE-CARD). *Circulation Journal.* 2011;75(10):2403-2410.

(325) Harjola V-, Follath F, Nieminen MS, Brutsaert D, Dickstein K, Drexler H, et al. Characteristics, outcomes, and predictors of mortality at 3 months and 1 year in patients hospitalized for acute heart failure. *European Journal of Heart Failure*. 2010;12(3):239-248.

(326) MacIntyre K, Capewell S, Stewart S, Chalmers JWT, Boyd J, Finlayson A, et al. Evidence of improving prognosis in heart failure: Trends in case fatality in 66 547 patients hospitalized between 1986 and 1995. *Circulation.* 2000;102(10):1126-1131.

(327) Mahjoub H, Rusinaru D, Souliere V, Durier C, Peltier M, Tribouilloy C. Long-term survival in patients older than 80 years hospitalised for heart failure. A 5-year prospective study. *European Journal of Heart Failure*. 2008;10(1):78-84.

(328) Pons F, Lupon J, Urrutia A, Gonzalez B, Crespo E, Diez C, et al. Mortality and Cause of Death in Patients With Heart Failure: Findings at a Specialist Multidisciplinary Heart Failure Unit. *Revista Española de Cardiología*. 2010;63(3):303-314.

(329) Rusinaru D, Mahjoub H, Goissen T, Massy Z, Peltier M, Tribouilloy C. Clinical features and prognosis of heart failure in women. A 5-year prospective study. *International Journal of Cardiology*. 2009;133(3):327-335.

(330) Shiba N, Watanabe J, Shinozaki T, Koseki Y, Sakuma M, Kagaya Y, et al. Analysis of chronic heart failure registry in the Tohoku district - Third year follow-up. *Circulation Journal*. 2004;68(5):427-434.

(331) Gorelik O, Almoznino-Sarafian D, Shteinshnaider M, Alon I, Tzur I, Sokolsky I, et al. Clinical variables affecting survival in patients with decompensated diastolic versus systolic heart failure. *Clinical Research in Cardiology*. 2009;98(4):224-232.

(332) Tribouilloy C, Buiciuc O, Rusinaru D, Malaquin D, Levy F, Peltier M. Long-term outcome after a first episode of heart failure. A prospective 7-year study. *International Journal of Cardiology.* 2010;140(3):309-314.

(333) Barsheshet A, Shotan A, Cohen E, Garty M, Goldenberg I, Sandach A, et al. Predictors of long-term (4-year) mortality in elderly and young patients with acute heart failure. *European Journal of Heart Failure*. 2010;12(8):833-840.

(334) Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, et al. Factors identified as precipitating hospital admissions for heart failure and clinical outcomes: findings from OPTIMIZE-HF. *Archives of Internal Medicine*. 2008;168(8):847-854.

(335) Ahmed A, Aronow WS, Fleg JL. Predictors of mortality and hospitalization in women with heart failure in the Digitalis Investigation Group trial. *American Journal of Therapy*. 2006;13(4):325-331.

(336) Huynh BC, Rovner A, Rich MW. Long-term survival in elderly patients hospitalized for heart failure: 14-year follow-up from a prospective randomized trial. *Archives of Internal Medicine*. 2006;166(17):1892-1898.

(337) Krumholz HM, Chen YT, Wang Y, Vaccarino V, Radford MJ, Horwitz RI. Predictors of readmission among elderly survivors of admission with heart failure. *American Heart Journal*. 2000;139(1 Pt 1):72-77.

(338) Wang L, Porter B, Maynard C, Bryson C, Sun H, Lowy E, et al. Predicting risk of hospitalization or death among patients with heart failure in the veterans health administration. A*merican Journal of Cardiology.* 2012;110(9):1342-1349.

(339) O'Connor CM, Abraham WT, Albert NM, Clare R, Gattis Stough W, Gheorghiade M, et al. Predictors of mortality after discharge in patients hospitalized with heart failure: An analysis from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). *American Heart Journal.* 2008;156(4):662-673.

(340) Bouvy ML, Heerdink ER, Leufkens HG, Hoes AW. Predicting mortality in patients with heart failure: a pragmatic approach. *Heart*. 2003;89(6):605-609.

(341) Barlera S, Tavazzi L, Franzosi MG, Marchioli R, Raimondi E, Masson S, et al. Predictors of mortality in 6975 patients with chronic heart failure in the gruppo italiano per lo studio della streptochinasi nell'infarto miocardico-heart failure trial proposal for a nomogram. *Circulation*: *Heart Failure*. 2013;6(1):31-39.

(342) Mascarenhas J, Lourenco P, Lopes R, Azevedo A, Bettencourt P. Chronic obstructive pulmonary disease in heart failure. Prevalence, therapeutic and prognostic implications. *American Heart Journal.* 2008;155(3):521-525.

(343) Iversen KK, Kjaergaard J, Akkan D, Kober L, Torp-Pedersen C, Hassager C, et al. Chronic obstructive pulmonary disease in patients admitted with heart failure. *Journal of Internal Medicine*. 2008;264(4):361-369.

(344) Rodríguez-Artalejo F, Guallar-Castillón P, Pascual CR, Otero CM, Montes AO, García A, et al. Health-Related quality of life as a predictor of hospital readmission and death among patients with heart failure. *Archives of Internal Medicine*. 2005;165(11):1274-1279.

(345) O'Loughlin C, Murphy NF, Conlon C, O'Donovan A, Ledwidge M, McDonald K. Quality of life predicts outcome in a heart failure disease management program. *International Journal of Cardiology.* 2010;139(1):60-67.

(346) Heidenreich PA, Spertus JA, Jones PG, Weintraub WS, Rumsfeld JS, Rathore SS, et al. Health status identifies heart failure outpatients at risk for hospitalization or death. *Journal of the American College of Cardiology*. 2006;47(4):752-756.

(347) Carson P, Tam SW, Ghali JK, Archambault WT, Taylor A, Cohn JN, et al. Relationship of quality of life scores with baseline characteristics and outcomes in the African-American Heart Failure Trial. *Journal of Cardiac Failure.* 2009;15(10):835-842.

(348) Allen LA, Gheorghiade M, Reid KJ, Dunlay SM, Chan PS, Hauptman PJ, et al. Identifying patients hospitalized with heart failure at risk for unfavorable future quality of life. *Circulation: Cardiovascular Quality and Outcomes*. 2011;4(4):389-398.

(349) Chin MH, Zhang JX, Rathouz PJ. Transitions in Health Status in Older Patients with Heart Failure. Southern Medical Journal. 2003;96(11):1096-1106.

(350) Setoguchi S, Stevenson LW, Schneeweiss S. Repeated hospitalizations predict mortality in the community population with heart failure. *American Heart Journal.* 2007;154(2):260-266.

(351) Rutten FH, Grobbee DE, Hoes AW. Differences between general practitioners and cardiologists in diagnosis and management of heart failure: a survey in every-day practice. *European Journal of Heart Failure.* 2003;5(3):337-344.

(352) Nielsen OW, Hilden J, McDonagh T, Fischer Hansen J. Survival differences between heart failure in general practices and in hospitals. *Heart.* 2003;89(11):1298-1302.

(353) Domanski MJ, Krause-Steinrauf H, Massie BM, Deedwania P, Follmann D, Kovar D, et al. A comparative analysis of the results from 4 trials of beta-blocker therapy for heart failure: BEST, CIBIS-II, MERIT-HF, and COPERNICUS. *Journal of Cardiac Failure*. 2003;9(5):354-363.

(354) Kannel WB, McGee DL. Diabetes and cardiovascular risk factors: the Framingham study. *Circulation* 1979;59(1):8-13.

(355) Nichols GA, Gullion CM, Koro CE, Ephross SA, Brown JB. The incidence of congestive heart failure in type 2 diabetes: an update. *Diabetes Care*. 2004;27(8):1879-1884.

(356) Lim H, MacFadyen RJ,Lip GH. DIabetes mellitus, the renin-angiotensin-aldosterone system, and the heart. Archives of Internal Medicine. 2004;164(16):1737-1748.

(357) MILLER JA. Impact of Hyperglycemia on the Renin Angiotensin System in Early Human Type 1 Diabetes Mellitus. *Journal of the American Society of Nephrology*. 1999;10(8):1778-1785.

(358) Fava S, Azzopardi J, Muscat HA, Fenech FF. Factors that influence outcome in diabetic subjects with myocardial infarction. *Diabetes Care*. 1993;16(12):1615-1618.

(359) Gerstein HC, Rosenstock J. Insulin Therapy in People Who Have Dysglycemia and Type 2 Diabetes Mellitus: Can It Offer Both Cardiovascular Protection and Beta-Cell Preservation? *Endocrinology Metabolism Clinics of North America.* 2005;34(1):137-154.

(360) Grundy SM, Benjamin IJ, Burke GL, Chait A, Eckel RH, Howard BV, et al. Diabetes and cardiovascular disease: a statement for healthcare professionals from the American Heart Association. *Circulation*. 1999;100(10):1134-1146.

(361) Wingard DL, Barrett-Connor EL, Scheidt-Nave C, McPhillips JB. Prevalence of cardiovascular and renal complications in older adults with normal or impaired glucose tolerance or NIDDM. A population-based study. *Diabetes Care.* 1993;16(7):1022-1025.

(362) Natali A, Vichi S, Landi P, Severi S, L'Abbate A, Ferrannini E. Coronary atherosclerosis in Type II diabetes: angiographic findings and clinical outcome. *Diabetologia*. 2000;43(5):632-641.

(363) Wexler DJ, Grant RW, Meigs JB, Nathan DM, Cagliero E. Sex disparities in treatment of cardiac risk factors in patients with type 2 diabetes. *Diabetes Care*. 2005;28(3):514-520.

(364) GORDON T, KANNEL WB, HJORTLAND MC, McNAMARA PM. Menopause and Coronary Heart DiseaseThe Framingham Study. *Annals of Internal Medicine*. 1978 August 1;89(2):157-161.

(365) Huiart L, Ernst P, Suissa S. Cardiovascular morbidity and mortality in copd. *CHEST Journal.* 2005;128(4):2640-2646.

(366) Sidney S, Sorel, Michael A1Quesenberry, Jr.,Charles P., DeLuise C, Lanes S, Eisner MD. Copd and incident cardiovascular disease hospitalizations and mortality: Kaiser permanente medical care program. *CHEST Journal*. 2005;128(4):2068-2075.

(367) Salisbury AC, Reid KJ, Spertus JA. Impact of Chronic Obstructive Pulmonary Disease on Post-Myocardial Infarction Outcomes. *American Journal of Cardiology*. 2007;99(5):636-641.

(368) Sin DD, Logan AG, Fitzgerald FS, Liu PP, Bradley TD. Effects of continuous positive airway pressure on cardiovascular outcomes in heart failure patients with and without Cheyne-Stokes respiration. *Circulation* 2000;102(1):61-66.

(369) Sin DD, Man SFP. Skeletal muscle weakness, reduced exercise tolerance, and COPD: is systemic inflammation the missing link? *Thorax*. 2006 January;61(1):1-3.

(370) Agusti A, Soriano JB. COPD as a systemic disease. *COPD*. 2008;5(2):133-138.

(371) Polak JF, Holman BL, Wynne J, Colucci WS. Right ventricular ejection fraction: An indicator of increased mortality in patients with congestive heart failure associated with coronary artery disease. *Journal of the American College of Cardiology*. 1983;2(2):217-224.

(372) Kjoller E, Kober L, Iversen K, Torp-Pedersen C, Trace Study Group. Importance of chronic obstructive pulmonary disease for prognosis and diagnosis of congestive heart failure in patients with acute myocardial infarction. *European Journal of Heart Failure*. 2004;6(1):71-77.

(373) Hawkins NM, Petrie MC, Jhund PS, Chalmers GW, Dunn FG, McMurray JJV. Heart failure and chronic obstructive pulmonary disease: diagnostic pitfalls and epidemiology. *European Journal of Heart Failure.* 2009;11(2):130-139.

(374) Caroci AdS, Lareau SC. Descriptors of dyspnea by patients with chronic obstructive pulmonary disease versus congestive heart failure. *Heart & Lung: The Journal of Acute and Critical Care.* 2004;33(2):102-110.

(375) Gustafsson F, Torp-Pedersen C, Brendorp B, Seibaek M, Burchardt H, Kober L, et al. Long-term survival in patients hospitalized with congestive heart failure: relation to preserved and reduced left ventricular systolic function. *European Heart Journal*. 2003;24(9):863-870.

(376) Shlipak MG, Smith GL, Rathore SS, Massie BM, Krumholz HM. Renal function, digoxin therapy, and heart failure outcomes: evidence from the digoxin intervention group trial. *Journal of the American Society of Nephrology.* 2004;15(8):2195-2203.

(377) De Silva R, Nikitin NP, Witte KKA, Rigby AS, Goode K, Bhandari S, et al. Incidence of renal dysfunction over 6 months in patients with chronic heart failure due to left ventricular systolic dysfunction: Contributing factors and relationship to prognosis. *European Heart Journal.* 2006;27(5):569-581.

(378) Hebert K, Dias A, Delgado MC, Franco E, Tamariz L, Steen D, et al. Epidemiology and survival of the five stages of chronic kidney disease in a systolic heart failure population. *European Journal of Heart Failure*. 2010;12(8):861-865.

(379) Rusinaru D, Buiciuc O, Houpe D, Tribouilloy C. Renal function and long-term survival after hospital discharge in heart failure with preserved ejection fraction. *International Journal of Cardiology*. 2011;147(2):278-282.

(380) Hillege HL, Girbes ARJ, De Kam PJ, Boomsma F, De Zeeuw D, Charlesworth A, et al. Renal function, neurohormonal activation, and survival in patients with chronic heart failure. *Circulation.* 2000;102(2):203-210.

(381) Tonelli M, Bohm C, Pandeya S, Gill J, Levin A, Kiberd BA. Cardiac risk factors and the use of cardioprotective medications in patients with chronic renal insufficiency. *American Journal of Kidney Diseases.* 2001;37(3):484-489.

(382) Ljungman S, Laragh JH, Cody RJ. Role of the kidney in congestive heart failure. Relationship of cardiac index to kidney function. *Drugs*. 1990;39 Suppl 4:10-21; discussion 22-4.

(383) Sarnak MJ, Levey AS, Schoolwerth AC, Coresh J, Culleton B, Hamm LL, et al. Kidney disease as a risk factor for development of cardiovascular disease: a statement from the American Heart Association Councils on Kidney in Cardiovascular Disease, High Blood Pressure Research, Clinical Cardiology, and Epidemiology and Prevention. *Hypertension.* 2003;42(5):1050-1065.

(384) Shah B, Greaves K. The Cardiorenal Syndrome: A Review. *International Journal of Nephrology*. 2011;2011:920195. doi:10.406

(385) Butler J, Forman DE, Abraham WT, Gottlieb SS, Loh E, Massie BM, et al. Relationship between heart failure treatment and development of worsening renal function among hospitalized patients. *American Heart Journal.* 2004;147(2):331-338.

(386) Echemann M, Zannad F, Briancon S, Juilliere Y, Mertes PM, Virion JM, et al. Determinants of angiotensin-converting enzyme inhibitor prescription in severe heart failure with left ventricular systolic dysfunction: The EPICAL study. *American Heart Journal*. 2000;139(4):624-631.

(387) Ljungman S, Kjekshus J, Swedberg K. Renal function in severe congestive heart failure during treatment with enalapril (the Cooperative North Scandinavian Enalapril Survival Study [CONSENSUS] Trial). *American Journal of Cardiology*. 1992;70(4):479-487.

(388) Ezekowitz J, McAlister FA, Humphries KH, Norris CM, Tonelli M, Ghali WA, et al. The association among renal insufficiency, pharmacotherapy, and outcomes in 6,427 patients with heart failure and coronary artery disease. *Journal of the American College of Cardiology*. 2004;44(8):1587-1592.

(389) Read WL, Tierney RM, Page NC, Costas I, Govindan R, Spitznagel EL, et al. Differential prognostic impact of comorbidity. *Journal of Clinical Oncology*. 2004;22(15):3099-3103.

(390) MacDonald MR, Petrie MC, Varyani F, Ostergren J, Michelson EL, Young JB, et al. Impact of diabetes on outcomes in patients with low and preserved ejection fraction heart failure: an analysis of the Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity (CHARM) programme. *European Heart Journal.* 2008 Jun;29(11):1377-1385.

(391) Adams KF, Sueta CA, Gheorghiade M, O'Connor CM, Schwartz TA, Koch GG, et al. Gender Differences in Survival in Advanced Heart Failure: Insights From the FIRST Study. *Circulation.* 1999;99(14):1816-1821.

(392) Swan JW, Walton C, Godsland IF, Clark AL, Coats AJS, Oliver MF. Insulin resistance in chronic heart failure. *European Heart Journal.* 1994;15(11):1528-1532.

(393) Kistorp C, Galatius S, Gustafsson F, Faber J, Corell P, Hildebrandt P. Prevalence and characteristics of diabetic patients in a chronic heart failure population. *International Journal of Cardiology*. 2005;100(2):281-287.

(394) Sin DD, Wu L, Man SFP. The relationship between reduced lung function and cardiovascular mortality*: A population-based study and a systematic review of the literature. *CHEST Journal*. 2005;127(6):1952-1959.

(395) Olson TP, Beck KC, Johnson BD. Pulmonary Function Changes Associated With Cardiomegaly in Chronic Heart Failure. *Journal of Cardiac Failure*. 2007;13(2):100-107.

(396) Pison C, Malo JL, Rouleau JL, Chalaoui J, Ghezzo H, Malo J. Bronchial hyperresponsiveness to inhaled methacholine in subjects with chronic left heart failure at a time of exacerbation and after increasing diuretic therapy. *Chest*. 1989;96(2):230-235.

(397) Lizak MK, Zakliczyński M, Jarosz A, Zembala M. The Influence of Chronic Heart Failure on Pulmonary Function Tests in Patients Undergoing Orthotopic Heart Transplantation. *Transplantation Proceedings*. 2009 ;41(8):3194-3197.

(398) Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D, et al. A More Accurate Method To Estimate Glomerular Filtration Rate from Serum Creatinine: A New Prediction Equation. *Annals of Internal Medicine.* 1999;130(6):461-470.

(399) Smilde TD, van Veldhuisen D, Navis G, Voors AA, Hillege HL. Drawbacks and prognostic value of formulas estimating renal function in patients with chronic heart failure and systolic dysfunction. *Circulation*. 2006;114(15):1572-1580.

(400) Froissart M, Rossert J, Jacquot C, Paillard M, Houillier P. Predictive Performance of the Modification of Diet in Renal Disease and Cockcroft-Gault Equations for Estimating Renal Function. *Journal of the American Society of Nephrology*. 2005;16(3):763-773.

(401) Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Annals of Internal Medicine*. 2006;145(4):247-254.

(402) Sterne J, Egger M, Moher D. Addressing Reporting Biases. In: Higgins J, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. Available from www.cochrane-handbook.org.: The Cochrane Collaboration; 2011. p. Chapter 10.

(403) Moher D, Liberati A, Tetzlaff J, Altman D. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *The BMJ*. 2009;339:b2535.doi: 10.1136/bmj.b2535

(404) Chisholm J. *The Read clinical classification*. *The BMJ*. 1990;300(6732):1092.

(405) Hawkins NM, Scholes S, Bajekal M, Love H, O'Flaherty M, Raine R, et al. Community care in England: reducing socioeconomic inequalities in heart failure. *Circulation*. 2012;126(9):1050-1057.

(406) Maru S, Koch GG, Stender M, Clark D, Gibowski L, Petri H, et al. Antidiabetic drugs and heart failure risk in patients with type 2 diabetes in the U.K. primary care setting. *Diabetes Care*. 2005;28(1):20-26.

(407) Ury HK. Efficiency of case-control studies with multiple controls per case: continuous or dichotomous data. *Biometrics*. 1975;31(3):643-649.

(408) Maclure M, Mittleman MA. Should we use a case-crossover design? *Annual Review of Public Health*. 2000;21:193-221.

(409) Flegal KM, Brownie C, Haas J. The effects of exposure misclassification on estimates of relative risk. *American Journal of Epidemiology*. 1986;123(4):736-751.

(410) Global Strategy for the Diagnosis, Management and Prevention of COPD. *Global Initiative for Chronic Obstructive Lung Disease (GOLD).* 2014; Available at: http://www.goldcopd.org/. Accessed 4th April, 2014.

(411) Smith CJ, Gribbin J, Challen KB, Hubbard RB. The impact of the 2004 NICE guideline and 2003 General Medical Services contract on COPD in primary care in the UK. *QJM: An International Journal of Medicine*. 2008;101(2):145-153.

(412) Cooke CR, Joo MJ, Anderson SM, Lee TA, Udris EM, Johnson E, et al. The validity of using ICD-9 codes and pharmacy records to identify patients with chronic obstructive pulmonary disease. *BMC Health Services Research*. 2011 Feb 16;11:37-6963-11-37.

(413) Quint JK, Müllerova H, DiSantostefano RL, Forbes H, Eaton S, Hurst JR, et al. Validation of chronic obstructive pulmonary disease recording in the Clinical Practice Research Datalink (CPRD-GOLD). *The BMJ*. Open 2014;4(7).

(414) British Medical Association. *Royal Pharmaceutical Society of Great Britain. British National Formulary, No. 58*. London: BMJ Group and Pharmaceutical Press; 2009.

(415) Güder G, Rutten FH, Brenner S, Angermann CE, Berliner D, Ertl G, et al. The Impact of Heart Failure on the Classification of COPD Severity. Journal of Cardiac Failure. 2012 8;18(8):637-644.

(416) Sandford A, Chagani T, Weir T, Connett J, Anthonisen N, Parã P. Susceptibility Genes for Rapid Decline of Lung Function in the Lung Health Study. *American Journal of Respiratory and Critical Care Medicine*. 2001;163(2):469-473.

(417) Schmidt SL, Nambiar AM, Tayob N, Sundaram B, Han MK, Gross BH, et al. Pulmonary Function Measures Predict Mortality Differently in Idiopathic Pulmonary Fibrosis versus Combined Pulmonary Fibrosis and Emphysema. *The European Respiratory Journal*. 2010;38(1):176-183.

(418) National Clinical Guideline Centre (UK). *Type 1 diabetes: Diagnosis and management of type 1 diabetes in children, young people and adults, NICE guidelines [CG15]*. London: National Institute for Health and Care Excellence; 2004.

(419) National Clinical Guideline Centre (UK). *Type 2 diabetes: The management of type 2 diabetes, NICE guidelines [CG87].* London: National Institute for Health and Care Excellence; 2009.

(420) World Health Organization. *Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus Abbreviated Report of a WHO Consultation.* Switzerland: WHO; 2011.

(421) Department of Health. *National Service Framework for Diabetes: Standards.* London: Department of Health; 2001.

(422) McDonald HI, Thomas SL, Millett ER, Nitsch D. CKD and the Risk of Acute, Community-Acquired Infections Among Older People With Diabetes Mellitus: A Retrospective Cohort Study Using Electronic Health Records. *American Journal of Kidney Disease*. 2015;66(1):60-8.

(423) The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *New England Journal of Medicine*. 1993;329(14):977-986.

(424) Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. *The Lancet.* 1998;352(9131):837-853.

(425) Iribarren C, Karter AJ, Go AS, Ferrara A, Liu JY, Sidney S, et al. Glycemic control and heart failure among adult patients with diabetes. *Circulation*. 2001;103(22):2668-2673.

(426) Parry HM, Desmukh H, Levin D, Van Zuydam N, Elder DHJ, Morris AD, et al. Both High and Low HbA1c Predict Incident Heart Failure in Type 2 Diabetes Mellitus. *Circulation: Heart Failure*. 2015;8(2):236-42.

(427) Aguilar D, Bozkurt B, Ramasubbu K, Deswal A. Relationship of Hemoglobin A1C and Mortality in Heart Failure Patients with Diabetes. *Journal of the American College of Cardiology*. 2009;54(5):422-428.

(428) Eshaghian S, Horwich TB, Fonarow GC. An unexpected inverse relationship between HbA1c levels and mortality in patients with diabetes and advanced systolic heart failure. *American Heart Journa*l. 2006;151(1):91.e1-91.e6.

(429) Engoren M, Schwann TA, Arslanian-Engoren C, Maile M, Habib RH. U-Shape Association Between Hemoglobin A1c and Late Mortality in Patients With Heart Failure After Cardiac Surgery. *American Journal of Cardiology*. 2013;111(8):1209-1213.

(430) Currie CJ, Peters JR, Tynan A, Evans M, Heine RJ, Bracco OL, et al. Survival as a function of HbA(1c) in people with type 2 diabetes: a retrospective cohort study. The Lancet. 2010 Feb 6;375(9713):481-489.

(431) Nicholas J, Charlton J, Dregan A, Gulliford MC. Recent HbA1c Values and Mortality Risk in Type 2 Diabetes. Population-Based Case-Control Study. *PLoS ONE*. 2013;8(7):e68008.

(432) National Institute for Health and Care Excellence. *Chronic kidney disease: early identification and management of chronic kidney disease in adults in primary and secondary care: NICE guidelines [CG182].* London: National Institute for Health and Care Excellence; 2014.

(433) Damman K, Voors AA, Navis G, van Veldhuisen DJ, Hillege HL. The Cardiorenal Syndrome in Heart Failure. *Progress in Cardiovascular Diseases*. 2011;54(2):144-153.

(434) Kidney Disease Improving Global Outcomes. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney International Supplements*. 2013;3(1):1-150.

(435) Grams ME, Plantinga LC, Hedgeman E, Saran R, Myers GL, Williams DE, et al. Validation of CKD and Related Conditions in Existing Datasets: A Systematic Review. *American Journal of Kidney Diseases*. 2010;57(1):44-54.

(436) Liao L, Anstrom KJ, Gottdiener JS, Pappas PA, Whellan DJ, Kitzman DW, et al. Long-term costs and resource use in elderly participants with congestive heart failure in the Cardiovascular Health Study. *American Heart Journal*. 2007;153(2):245-252.

(437) Tonelli M, Klarenbach SW, Lloyd AM, James MT, Bello AK, Manns BJ, et al. Higher estimated glomerular filtration rates may be associated with increased risk of adverse outcomes, especially with concomitant proteinuria. *Kidney International.* 2011;80(12):1306-1314.

(438) Ergul Y, Nisli K, Avci B, Omeroglu RE. Dilated cardiomyopathy associated with dystrophic epidermolysis bullosa: Role of micronutrient deficiency?. *Turk Kardiyoloji Dernegi Arsivi.* 2011;39(4):328-331.

(439) Testani JM, McCauley BD, Chen J, Shumski M, Shannon RP. Worsening Renal Function Defined as an Absolute Increase in Serum Creatinine Is a Biased Metric for the Study of Cardio-Renal Interactions. *Cardiology*. 2010;116(3):206-212.

(440) National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease:Evaluation, classification and stratification. *American Journal of Kidney Diseases*. 2002;39:S1-S266.

(441) Khan NA, Ma I, Thompson CR, Humphries K, Salem DN, Sarnak MJ, et al. Kidney Function and Mortality among Patients with Left Ventricular Systolic Dysfunction. *Journal of the American Society of Nephrology*. 2006;17(1):244-253.

(442) Ather S, Bavishi C, McCauley MD, Dhaliwal A, Deswal A, Johnson S, et al. Worsening renal function is not associated with response to treatment in acute heart failure. *International Journal of Cardiology.* 2015;167(5):1912-1917.

(443) National Institute for Health and Care Excellence. *Chronic kidney disease: early identification and management of chronic kidney disease in adults in primary and secondary care; NICE guidelines [CG73]*. London: National Institute for Health and Care Excellence; 2009.

(444) National Institute for Health and Clinical Excellence. *Diabetes in Adults Quality Standard; NICE quality standards [QS6].* London: National Institute for Health and Care Excellence; 2013.

(445) Craig R, Mindell J editors. *Health Survey for England – 2011. Volume 1: Health, social care and lifestyles*. London: Health and Social Care Information Centre; 2012.

(446) Office for National Statistics. *2011 Census: Ethnic group, local authorities in the United Kingdom. 2013;* Available at: http://www.ons.gov.uk/ons/rel/census/2011-census/key-statistics-and-quick-statistics-for-local-authorities-in-the-united-kingdom---part-1/rft-ks201uk.xls.

(447) Lainscak M, Anker SD. Prognostic factors in chronic heart failure. A review of serum biomarkers, metabolic changes, symptoms, and scoring systems. *Herz.* 2009;34(2):141-147.

(448) Department for Communities and Local Government. *English Indices of Deprivation 2010: Guidance Doument [Online].* 2011; Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6871/1871208.pdf. Accessed June 8th, 2015.

(449) van der Heijden GJ, Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. *Journal of Clinical Epidemiology*. 2006;59(10):1102-1109.

(450) Miller R. *Beyond ANOVA: Basics of Applied Statistics.* London: Chapman & Hall; 1997.

(451) Newson R. *Confidence intervals and p-values for delivery to the end user.* 2003(3):245-269.

(452) Grace-Martin K. Multicollinearity in Linear Regression Models - Centering Variables to Reduce Multicollinearity.2008. Available at: http://EzineArticles.com/2077352. Accessed 4th November, 2015.

(453) Andersson T, Alfredsson L, Kallberg H, Zdravkovic S, Ahlbom A. Calculating measures of biological interaction. *European Journal of Epidemiology*. 2005;20(7):575-579.

(454) Hosmer DW, Lemeshow S. Confidence Interval Estimation of Interaction. *Epidemiology*. 1992;3(5):452-456.

(455) Walter S, Tiemeier H. Variable selection: current practice in epidemiological studies. European Journal of Epidemiology. 2009;24(12):733-736.

(456) Greenland S. Modeling and variable selection in epidemiologic analysis. *American Journal of Public Health.* 1989;79(3):340-349.

(457) Greenland S. Invited commentary: variable selection versus shrinkage in the control of multiple confounders. *American Journal of Epidemiology*. 2008;167(5):523-9; discussion 530-1.

(458) Cook NR. Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction. *Circulation.* 2007;115(7):928-935.

(459) Hardin J, Hilbe J editors. *Generalized Linear Models and Extensions*. 2nd Edition ed. Texas: Stata Press; 2007.

(460) McFadden D. “Conditional logit analysis of qualitative choice behavior.”. In: Zarembka P, editor. *Frontiers in Econometrics*. Cambridge: Academic Press; 1974. p. 105-142.

(461) National Institute for Health and Care Excellence. Chronic kidney disease: early identification and management of chronic kidney disease in adults in primary and secondary care; NICE guidelines [CG182]. London: National Institute for Health and Care Excellence; 2014.

(462) Aranda JM,Jr, McIntyre SE, Klodell CT,Jr, York KM, Dragstedt CA, Chaille PJ, et al. Initial heart rate and systolic blood pressure predict outcomes in chronic heart failure patients who are evaluated for cardiac transplant. *Clinical Cardiology*. 2007;30(6):282-287.

(463) Lloyd-Jones DM. Cardiovascular Risk Prediction: Basic Concepts, Current Status, and Future Directions. *Circulation*. 2010;121(15):1768-1777.

(464) Roger VL, Weston SA, Redfield MM, Hellermann-Homan JP, Killian J, Yawn BP, et al. Trends in heart failure incidence and survival in a community-based population. *Journal of the American Medical Association.* 2004;292(3):344-350.

(465) Goldberg RJ, Ciampa J, Lessard D, Meyer TE, Spencer FA. Long-term survival after heart failure: A contemporary population-based perspective. *Archives of Internal Medicine*. 2007;167(5):490-496.

(466) Goodlin SJ. Palliative Care in Congestive Heart Failure. *Journal of the American College of Cardiology* 2009;54(5):386-396.

(467) Gross NJ, Petty TL, Friedman M, Skorodin MS, Silvers GW, Donohue JF. Dose response to ipratropium as a nebulized solution in patients with chronic obstructive pulmonary disease. A three-center study. "The American Review of Respiratory Disease. 1989;139(5):1188-1191.

(468) Higgins BG, Powell RM, Cooper S, Tattersfield AE. Effect of salbutamol and ipratropium bromide on airway calibre and bronchial reactivity in asthma and chronic bronchitis. *European Respiratory Journal.* 1991;4(4):415-420.

(469) Vathenen AS, Britton JR, Ebden P, Cookson JB, Wharrad HJ, Tattersfield AE. High-dose inhaled albuterol in severe chronic airflow limitation. *The American Review of Respiratory Disease*.1988;138(4):850-855.

(470) Schols AM, Wesseling G, Kester AD, de Vries G, Mostert R, Slangen J, et al. Dose dependent increased mortality risk in COPD patients treated with oral glucocorticoids. *European Respiratory Journal.* 2001;17(3):337-342.

(471) Horita N, Miyazawa N, Morita S, Kojima R, Inoue M, Ishigatsubo Y, et al. Evidence suggesting that oral corticosteroids increase mortality in stable chronic obstructive pulmonary disease. *Respiratory Research*. 2014;15:37-9921-15-37.

(472) Walters JA, Walters EH, Wood-Baker R. Oral corticosteroids for stable chronic obstructive pulmonary disease. *Cochrane Database Systematic Reviews*. 2005;(3)(3):CD005374.

(473) Schmidt SAJ, Johansen MB, Olsen M, Xu X, Parker JM, Molfino NA, et al. The impact of exacerbation frequency on mortality following acute exacerbations of COPD: a registry-based cohort study. *The BMJ Open*. 2014;4(12).

(474) Damman K, Jaarsma T, Voors AA, Navis G, Hillege HL, van Veldhuisen DJ, et al. Both in- and out-hospital worsening of renal function predict outcome in patients with heart failure: results from the Coordinating Study Evaluating Outcome of Advising and Counseling in Heart Failure (COACH). European *Journal of Heart Failure.* 2009;11(9):847-854.

(475) Dib N, Dinsmore J, Lababidi Z, White B, Moravec S, Campbell A, et al. One-Year Follow-Up of Feasibility and Safety of the First U.S., Randomized, Controlled Study Using 3-Dimensional Guided Catheter-Based Delivery of Autologous Skeletal Myoblasts for Ischemic Cardiomyopathy (CAuSMIC Study). *Journal of the American College of Cardiology: Cardiovascular Interventions*. 2009;2(1):9-16.

(476) Pyram R, Kansara A, Banerji MA, Loney-Hutchinson L. Chronic kidney disease and diabetes. *Maturitas.* 2012;71(2):94-103.

(477) Ahlbom A, Alfredsson L. Interaction: A word with two meanings creates confusion. *European Journal of Epidemiology.* 2005;20(7):563-564.

(478) Iseki K, Iseki C, Ikemiya Y, Fukiyama K. Risk of developing end-stage renal disease in a cohort of mass screening. *Kidney International*. 1996;49(3):800-805.

(479) Halbesma N, Brantsma AH, Bakker SJ, Jansen DF, Stolk RP, De Zeeuw D, et al. Gender differences in predictors of the decline of renal function in the general population. Kidney International. 2008;74(4):505-512.

(480) Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. *Biostatistics.* 2005;6(2):227-239.

(481) Hosmer D, Lemesbow. Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics - Theory and Methods.* 1980;9(10):143-169.

(482) Harrell FJ. *Regression Modeling Strategies*. New York: Springer; 2001.

(483) Morrato EH, Elias M, Gericke CA. Using population-based routine data for evidence-based health policy decisions: lessons from three examples of setting and evaluating national health policy in Australia, the UK and the USA. *Journal of Public Health*. 2007;29(4):463-471.

(484) Jorm L. Routinely collected data as a strategic resource for research: priorities for methods and workforce. *Public Health Research and Practice*. 2015;25(4):e2541540.

(485) Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Medicine. 2015;12(10):e1001885.

(486) Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *Journal of Epidemiology and Community Health.* 2012;66(12):1182-1186.

(487) Fan J, Han F, Liu H. Challenges of Big Data Analysis. *National Science Review*. 2014;1(2):293-314.

(488) Kaplan RM, Chambers DA, Glasgow RE. Big data and large sample size: a cautionary note on the potential for bias. *Clinical and Translational Science*. 2014;7(4):342-346.

(489) Nuzzo R. Scientific method: statistical errors. *Nature*. 2014;506(7487):150-152.

(490) Greenwald AG, Gonzalez R, Harris RJ, Guthrie D. Effect sizes and p values: what should be reported and what should be replicated? Psychophysiology.1996;33(2):175-183.

(491) Overhage JM, Overhage LM. Sensible use of observational clinical data. *Statistical Methods in Medical Research*. 2013;22(1):7-13.

(492) Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *British Journal of Clinical Pharmacology*. 2010;69(1):4-14.

(493) Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *British Journal of General Practice*. 2010;60(572):e128-36.

(494) Grimes DA, Schulz KF. Compared to what? Finding controls for case-control studies. *The Lancet*. 2005;365(9468):1429-1433.

(495) Breslow NE. Statistics in epidemiology: the case-control study. *Journal of the American Statistics Association*. 1996;91(433):14-28.

(496) Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. American Journal of Epidemiology. 2007;166(6):646-655.

(497) Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. American Journal of Epidemiology. 1995;142(12):1255-1264. (498) Donders

(498) Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology.* 2006;59(10):1087-1091.

(499) Vach W. *Logistic regression with missing values in the covariates*. New York: Springer; 1994.

(500) Hougardy JM, Delanaye P, Le Moine A, Nortier J. Estimation of the glomerular filtration rate in 2014 by tests and equations: strengths and weaknesses. Revue Médicale de Bruxelles. 2014;35(4):250-257.

(501) Oakes D. Survival times: Aspects of partial likelihood (with discussion). *International Statistics Review*. 1981; 49:235-264.

(502) Borgan O, Langholz B. Non-parametric estimation of relative mortality from nested case-control studies. *Biometrics*. 1993;49:593-602.

(503) Langholz B, Goldstein L. Risk set sampling in epidemiologic cohort studies. *Statistical Science*. 1996;11:35-53.

(504) Messite J, Stellman SD. Accuracy of death certificate completion: the need for formalized physician training. JAMA. 1996;275(10):794-796.

(505) Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bulletin of the World Health Organization. 2005;83(3):171-177.

(506) Murray CJL, Kulkarni SC, Ezzati M. Understanding the Coronary Heart Disease Versus Total Cardiovascular Mortality Paradox: A Method to Enhance the Comparability of Cardiovascular Death Statistics in the United States. *Circulation*. 2006;113(17):2071-2081.

(507) Snyder ML, Love S, Sorlie PD, Rosamond WD, Antini C, Metcalf PA, et al. Redistribution of heart failure as the cause of death: the Atherosclerosis Risk in Communities Study. *Population Health Metrics.* 2014;12:10-10.

(508) Black WC, Haggstrom DA, Gilbert Welch H. All-Cause Mortality in Randomized Trials of Cancer Screening. Journal of the National Cancer Institute. 2002;94(3):167-173.

(509) Hulley SB, Walsh JM, Newman TB. Health policy on blood cholesterol. Time to change directions. *Circulation*. 1992;86(3):1026-1029.

(510) Pitt B. Candesartan reduced mortality and hospital admissions in chronic heart failure. *American College of Physicians Journal Club.* 2004;140(2):32-33.

(511) Packer M, Bristow MR, Cohn JN, Colucci WS, Fowler MB, Gilbert EM, et al. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. New England Journal Medicine. 199696;334(21):1349-1355.

(512) Rushton CA, Green J, Jaarsma T, Walsh P, Strömberg A, Kadam UT. The challenge of multimorbidity in nurse education: An international perspective. *Nurse Education Today*. 2015; 35(1): 288-292.