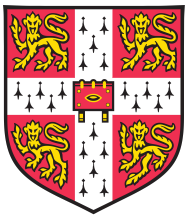
Association of cardiovascular and musculoskeletal biomarkers with clinical outcomes in chronic obstructive pulmonary disease.



This dissertation is submitted

by

Jilles Martijn Fermont

to

The Department of Medicine

for the degree of

Doctor of Philosophy (PhD)

in the subject of

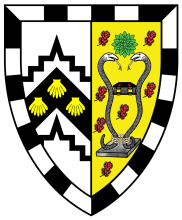
Medicine

University of Cambridge

Gonville & Caius College

Cambridge, United Kingdom

February 2019



Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

In accordance with the Degree Committee for the Faculty of Clinical Medicine, this dissertation does not exceed the prescribed limit of 60,000 words.

Summary

Association of cardiovascular and musculoskeletal biomarkers with clinical outcomes in chronic obstructive pulmonary disease.
Jilles M. Fermont

Background Chronic obstructive pulmonary disease (COPD) is one of the leading causes of death in the world. In addition to premature mortality, the consequent socio-economic burden is high, causing reduced quality of life, loss of productivity, and hospital admission. Diagnosis of COPD relies on lung function tests, which are inadequate and often leave the condition undiagnosed and thus untreated. There is a growing interest in the extra-pulmonary manifestations of COPD and assessing the predictive value of cardiovascular abnormalities, musculoskeletal weakness and plasma biomarkers for acute exacerbation of COPD, hospital admission and mortality, as there is currently no individual biomarker able to reliably identify or predict these common clinical outcomes. The aim of this research was to identify and evaluate the predictive value of existing and novel biomarkers for COPD, and determine if and how these biomarkers can predict the longer-term clinical outcomes using electronic health record data.

Methods Electronic databases were systematically searched and identified 61 studies, which were synthesised, including meta-analyses to estimate pooled hazard ratios of the associations between selected biomarkers and common clinical outcomes. Data derived from the Evaluating the Role of Inflammation in Chronic Airways disease

(ERICA) study were linked to electronic health record data (i.e. hospital admissions) and survival data. Predictive models for mortality and cardiovascular related hospital admission were developed using stratified multivariable Cox regression, and assessed by C-indices with 10-fold cross-validation. Negative binomial regression was used to model the event rate of acute exacerbation of COPD and determine the risk of hospitalisation due to acute exacerbation of COPD, and the associated length of stay. Data from the UK Biobank were used to explore cause-specific deaths in COPD. Sex-specific all-cause and cause-specific mortality rates were age-standardised using the 2013 European Standard Population. Hazard ratios were estimated using Cox proportional hazards regression, adjusted for age and sex.

Findings Systematic review indicated that shorter six-minute walk distance, elevated heart rate, fibrinogen, C-reactive protein, and white cell count were associated with a higher risk of mortality. Shorter six-minute walk distance and elevated fibrinogen and C-reactive protein were associated with COPD exacerbation, and shorter six-minute walk distance and higher heart rate, C-reactive protein and interleukin-6 were associated with hospitalisation. Data from the ERICA cohort indicated no significant difference between the discriminative ability of a BODE Index with six-minute walk and BODE Index with short physical performance battery when predicting mortality. For most musculoskeletal measures, poorer performance was associated with higher rate or longer duration of hospitalised acute exacerbation of COPD. Measures of arterial stiffness and carotid intima-media thickness were not associated with cardiovascular events. Measures of exercise capacity were significantly associated with cardiovascular disease and improved the discriminative ability when added to Framingham risk factors. Data from the UK Biobank indicated COPD was associated with a higher risk of all-cause mortality, and cardiovascular death. In both men and women, COPD had an associated threefold higher risk of early mortality, including a fourfold higher risk of cardiovascular-related death in women, and threefold higher

risk of cardiovascular-related death in men.

Conclusions Epidemiological evidence indicates that musculoskeletal measures have the potential to replace the six-minute walk in the BODE Index for predicting mortality in COPD. In addition, physical capacity should be considered as an important treatable trait in reducing risk of hospitalisations for acute exacerbation of COPD. Data from the ERICA cohort does not support the use of objective measures of arterial stiffness and carotid intima-media thickness in addition to Framingham risk factors for predicting cardiovascular events within COPD. Findings in the UK Biobank indicated that COPD is associated with a higher risk of cardiovascular death but cancer and respiratory disease to be the leading causes.

Preface

The aim of this dissertation was to explore the relationships between existing and novel biomarkers, questionnaire data, and electronic health record data and determine if and how these biomarkers can predict common clinical outcomes (i.e. acute exacerbation of chronic obstructive pulmonary disease (COPD), hospitalisation, mortality) within a COPD population. To conduct the analyses I used clinical data from the Evaluation of the Role of Inflammation in Chronic Airways disease (ERICA) cohort – a unique and well-defined dataset containing numerous biomarkers including musculoskeletal and cardiovascular markers, and demographic data of individuals diagnosed with COPD – and linked these to mortality data and electronic health record data obtained from the UK Office for National Statistics and National Health Services (NHS) Digital (England), NHS Wales and NHS Scotland, respectively. To explore how some of the findings in the ERICA cohort present in a different cohort, I have analysed data from the UK Biobank.

"One of the challenges for biomedicine in the digital age is how to move from data to knowledge and from knowledge to action to enhance the lives of patients in real-world contexts."

Gibbons

Acknowledgements

This report presents independent research funded by GlaxoSmithKline (GSK). The views expressed are those of the author and not necessarily those of GSK or the University of Cambridge. My role and of others in each chapter in this thesis are described below. Although the work described is of my own, it would not have been possible without the input of others. I thank my supervisors Professor Ian Wilkinson at the Department of Medicine, division of Experimental Medicine and Immunotherapeutics, Dr Angela Wood at the Department of Public Health and Primary Care, Cardiovascular Epidemiology Unit, and Dr Hana Muellerova at GSK, Respiratory Epidemiology Unit for their support and guidance throughout my PhD. I thank Professor Michael Polkey and Dr Divya Mohan for their clinical input, Dr Emma Day and Dr Mellone Marchong, and Ms Alicia Gore for their support with the electronic health record data applications, and Mr Thomas Bolton for his support in the UK Biobank application. I would like to thank the ERICA consortium and all study participants. I am grateful to the University of Cambridge and GSK for funding my PhD studies in Cambridge with a 3-year PhD Studentship (RG79358). In addition, I thank the American Thoracic Society, the European Respiratory Society (ERS), the Centre for Health Economics at the University of York, and Gonville & Caius College for their Scholarships, allowing me to learn and disseminate my research nationally and internationally. In addition, I am grateful to be the recipient of an ERS Young Scientist 2018 award. Finally, I would like to thank Dr Christine Lu for her mentorship during my fellowship at Harvard University. On a personal note, I would like to thank my parents and twin brother for their love and support,

and my close friends in Cambridge, in particular Paulina Rowicka and Alessandro Morelli for their friendship and making Cambridge such a wonderful experience.

Chapter 1 I drafted the text. Professor Ian Wilkinson, Dr Angela Wood and Dr Hana Muellerova provided helpful feedback.

Chapter 2 I wrote the study protocol for registration with *PROSPERO* (CRD 42016052075). I and Dr Angela Wood designed the study. I, Dr Magnus Jensen, Dr Renata Ferrari, Dr Valeria Pires Di Lorenzo, Dr Jacob Marot, Dr Philipp Schuetz, Dr Henrik Watz, Dr Benjamin Waschki, and Dr Hana Muellerova contributed to the data collection. I and Dr Katya Masconi extracted the data. I conducted the analysis, and produced the results figures and tables. Dr Angela Wood and Dr Katya Masconi provided advice on the statistical analysis. I wrote the initial draft of the manuscript. Dr Katya Masconi, Professor Michael Polkey, Professor Ian Wilkinson, and Dr Angela Wood contributed to the writing of the manuscript. All co-authors within the ERICA consortium read and commented on the manuscript.

Chapter 3 I created the database specification dictionary with clinical input from Dr Divya Mohan (**Appendix D**). I conducted the data management and data cleaning. I prepared and submitted applications with the National Health Service (NHS) Digital, NHS Wales, and NHS Scotland to obtain hospital episode statistics, and the Office for National Statistics to obtain mortality data for data linkage purposes. Supporting these applications I built the website http://ericacopd.org. I linked hospital episode statistics and mortality data with the ERICA data. I decided on and conducted the statistical analyses, produced the relevant tables and figures, and drafted the text. Dr Angela Wood and Professor Ian Wilkinson provided helpful feedback.

Chapter 4 I, together with Dr Angela Wood, Professor Ian Wilkinson, Dr Hana Muellerova, and Professor Michael Polkey designed the study. The source data set came from a study conceived and directed by Dr Ruth Tal-Singer, Professor Ian

Wilkinson and Professor Michael Polkey, and with Dr Charlotte Bolton, Dr William Macnee, Dr John Cockroft, Dr Carmel McEniery, and Dr Jonathan Fuld as co-investigators who organised and carried out the clinical study. I, Dr Divya Mohan and Dr Marie Fisk obtained the data. I conducted the analysis and produced the results, figures and tables. Dr Angela Wood provided advice on the statistical analysis. I used *STATA* program *precalib* developed by Dr Stephen Kaptoge to create calibration plots. I wrote the initial draft of the manuscript. Professor Michael Polkey, Dr Angela Wood, Professor Ian Wilkinson, and Dr Divya Mohan contributed to the writing of the manuscript. All co-authors critiqued and commented on the manuscript.

Chapter 5 I prepared the application and obtained access to the UK Biobank (35826). I decided on and conducted the statistical analyses, produced the relevant tables and figures, and drafted the text. Dr Angela Wood provided advice on the statistical analysis. Dr Angela Wood and Professor Ian Wilkinson provided helpful feedback.

Chapter 6 I, together with Dr Hana Muellerova, Professor Ian Wilkinson, Dr Angela Wood, and Professor Michael Polkey designed the study. The source data set came from a study conceived and directed by Dr Ruth Tal-Singer, Professor Ian Wilkinson and Professor Michael Polkey, and with Dr Charlotte Bolton, Dr William Macnee, Dr John Cockroft, Dr Carmel McEniery, and Dr Jonathan Fuld as co-investigators who organised and carried out the clinical study. I, together with Dr Divya Mohan and Dr Marie Fisk obtained the data. I conducted the analysis and produced the results, figures and tables. Dr Angela Wood provided advice on the statistical analysis. I wrote the initial draft of the manuscript. Dr Hana Muellerova, Dr Angela Wood, Professor Michael Polkey, Professor Ian Wilkinson, and Dr Divya Mohan contributed to the writing of the manuscript. All co-authors critiqued and commented on the manuscript.

Chapter 7 I, together with Professor Ian Wilkinson, Dr Angela Wood, Dr Hana Muellerova, and Professor Michael Polkey designed the study. The source data set came from a study conceived and directed by Dr Ruth Tal-Singer, Professor Ian Wilkinson and Professor Michael Polkey, and with Dr Charlotte Bolton, Dr William Macnee, Dr John Cockroft, Dr Carmel McEniery, and Dr Jonathan Fuld as co-investigators who organised and carried out the clinical study. I, together with Dr Divya Mohan and Dr Marie Fisk obtained the data. I conducted the analysis and produced the results, figures and tables. Dr Angela Wood provided advice on the statistical analysis. I wrote the initial draft of the manuscript. Professor Ian Wilkinson, Professor Michael Polkey, Dr Angela Wood, and Dr Divya Mohan contributed to the writing of the manuscript. All co-authors critiqued and commented on the manuscript.

Chapter 8 I drafted the text. Professor Ian Wilkinson and Dr Angela Wood and Dr Hana Muellerova provided helpful feedback.

Abbreviations

4MGS	Four-metre gait speed
6MW	Six-minute walk
AECOPD	Acute exacerbation of COPD
AIx	Augmentation index
BMI	Body mass index
BODE	Body mass index, airflow obstruction, dyspnoea and exercise capacity
CAT	COPD assessment test
CIMT	Carotid intima-media thickness
COPD	Chronic obstructive pulmonary disease
CRP	C-reactive protein
CV	Cardiovascular
EHR	Electronic health record
ERICA	Evaluation of the role of inflammation in chronic airways disease
FEV_1	Forced expiratory volume in one second
FVC	Forced vital capacity
GFR	Glomerular filtration rate
GOLD	Global initiative for chronic obstructive lung disease
HDL	High-density lipoprotein
HR	Hazard ratio
ICD-10	10th revision of the international statistical classification of diseases and related health problems
IL	Interleukins
IQR	Inter-quartile range
IRR	Incidence risk ratio
MICE	Multiple imputation by chained equations
MOOSE	Meta-analysis of observational studies in epidemiology
MRC	Medical Research Council
NHS	National Health Services
ONS	Office for National Statistics
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PWV	Pulse wave velocity
QMVC	Quadriceps maximum voluntary contraction
SGRQ-C	St.George respiratory questionnaire for COPD
SMD	Standardised mean difference
SNIP	Sniff nasal inspiratory pressure
SPPB	Short physical performance battery
STROBE	Strengthening the reporting of observational studies in epidemiology
TNF-α	Tumour necrosis factor-alpha
TRIPOD	Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis
WCC	White cell count

Contents

List of Figures

List of Tables

Chronic obstructive pulmonary disease is a lung disease characterized by chronic obstruction of lung airflow that interferes with normal breathing and is not fully reversible

World Health Organisation

1

Introduction

Chapter summary

This chapter describes the definition, underlying biology and disease manifestation of chronic obstructive pulmonary disease (COPD). Incidence and prevalence of COPD are reported in addition to how COPD is diagnosed, and common clinical outcomes including COPD exacerbation, hospital admission and mortality. There is a great incentive to identify high risk individuals in an early stage of disease with a focus on extra-pulmonary manifestations of COPD (i.e. systemic disease). Biomarkers currently used in diagnosing and staging COPD are described, in addition to novel biomarkers that can potentially capture systemic disease that traditional lung function measures fail to capture. A specific focus is placed on cardiovascular and musculoskeletal measures. Several clinical studies are exploring the potential use of measuring extra-pulmonary manifestations of COPD, including the Evaluating the Role of Inflammation in Chronic Airways disease (ERICA) study and are briefly introduced. This chapter concludes with an outline of the thesis.

1.1 Background

1.1.1 Definition

The Global Initiative for Chronic Obstructive Lung Disease (GOLD) – an independent body working with health care professionals aimed at increasing the awareness and promoting evidence-based practice for lung disease – characterises chronic obstructive pulmonary disease (COPD) by a "persistent air flow limitation that is usually progressive" and by the association with "an enhanced chronic inflammatory response in the airways and the lung to noxious particles or gases".[107] It is characterised by acute exacerbations, also known as flare ups, and comorbidities with both contributing to disease severity. Previously differentiation was made between chronic bronchitis and lung emphysema, with bronchitis being characterised by productive sputum (i.e. phlegm) and emphysema by air trapping in the lungs due to breakage of the alveoli (i.e. air sacs). As most individuals have a combination of the two conditions, often just the term COPD is used (**Figure 1.1**, page 2). Chronic bronchitis, however, is still seen as a separate disorder and may exist in individuals even with normal spirometry.

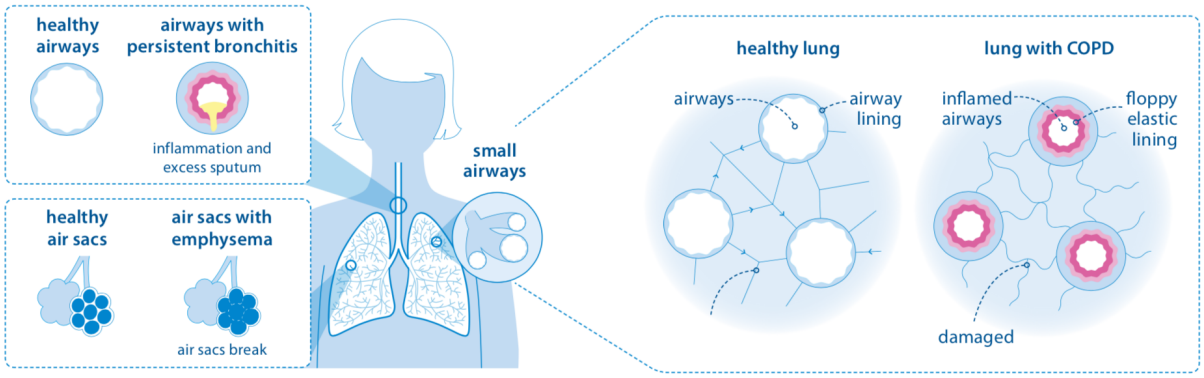


Figure 1.1: Chronic inflammation of the small airways. In addition to narrowing of the airways, lung tissue is damaged due to inflammation of the airway lining removing the elastic recoil. Image taken from the British Lung Foundation. "COPD: Living with chronic obstructive pulmonary disease".

1.1.2 Biology of COPD and manifestations

Chronic inflammation is present throughout the airways, parenchyma – functional lung tissue –, and in the pulmonary vasculature. Pathological changes include structural changes in the airway epithelial, smooth muscle and connective tissue with destruction of capillaries and the development of pulmonary hypertension and abnormal enlargement of the heart.[165] In addition to the lungs, COPD is also believed to lead to systemic problems such as impaired systemic muscle function and reduced exercise intolerance resulting from an increase in inflammatory markers such as macrophages, eosinophil count and neutrophils in various parts of the lung.[43] Cardiovascular disease is believed to be a common comorbidity in COPD. Even in mild cases of COPD, individuals may experience a reduced maximum heart rate and oxygen uptake.[34] Cardiovascular disease is believed to be a common comorbidity in COPD. Evidence suggest the two diseases go hand-in-hand[193,253] and can be explained by a so-called spill-over effect of inflammatory response (**Figure 1.2**, page 4).[272] As a result of chronic inflammation in the pulmonary vasculature, the walls gradually thicken starting with the intima, followed by an increase in smooth muscle, proteoglycans and collagen, and the infiltration of inflammatory cells into the vessel walls that in turn could lead to arterial stiffness.[229] The evidence of a spill-over effect, however, is inconclusive with some studies being unable to measure a relationship between productive cough and blood.[3] The statistical power of these studies is low, however, due to small sample sizes (n $<$50). Also, even in the absence of COPD, smoking itself may lead to systemic inflammation[241] and has been found to continue in former smokers.[126] In addition, ageing itself is accompanied with low-grade inflammation.[69] An alternative view is to see systemic inflammation as a multi-organ inflammation.[17] In the Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE) cohort, an association between common diseases such as heart disease and diabetes, and systemic inflammation was found.[180] Considering low-grade inflammation continues after smoking cessation, and ageing being an unmodifiable factor there is growing interest in targeting systematic inflammation therapeutically (i.e. secondary prevention).[294] Some evidence suggests that alongside smoking cessation, frequent exercise and

administering anti-inflammatory medication such as statins in those with systemic disease to lower the risk of COPD (i.e. primary prevention).[293]

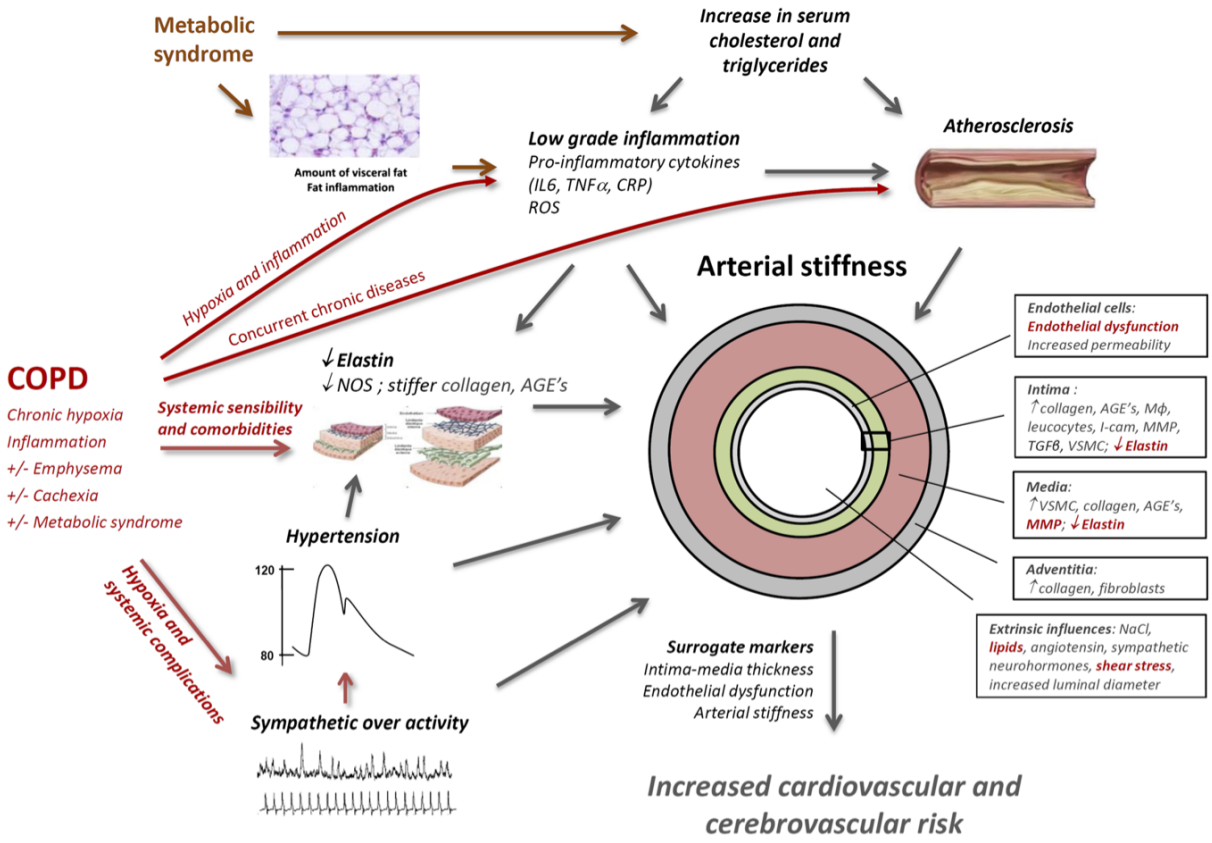


Figure 1.2: Inflammatory spill-over effect. "Mechanisms by which arterial stiffness is increased in chronic obstructive pulmonary disease".[272] With permission.

It takes several years for COPD to develop, with the airways gradually becoming narrower making it harder to breathe, causing chronic air trapping within the lungs. The usual onset of COPD is after the age of forty and most individuals will have a smoking history of at least ten pack years. Classical disease manifestations include chronic coughing, sputum production, dyspnoea (i.e. breathlessness), and an overall decline in lung function measured through spirometry (**Figure 1.3**, page 6). Systemic manifestations on the other hand include pulmonary hypertension, impaired systemic muscle function, unintended weight loss. Due to the various systemic

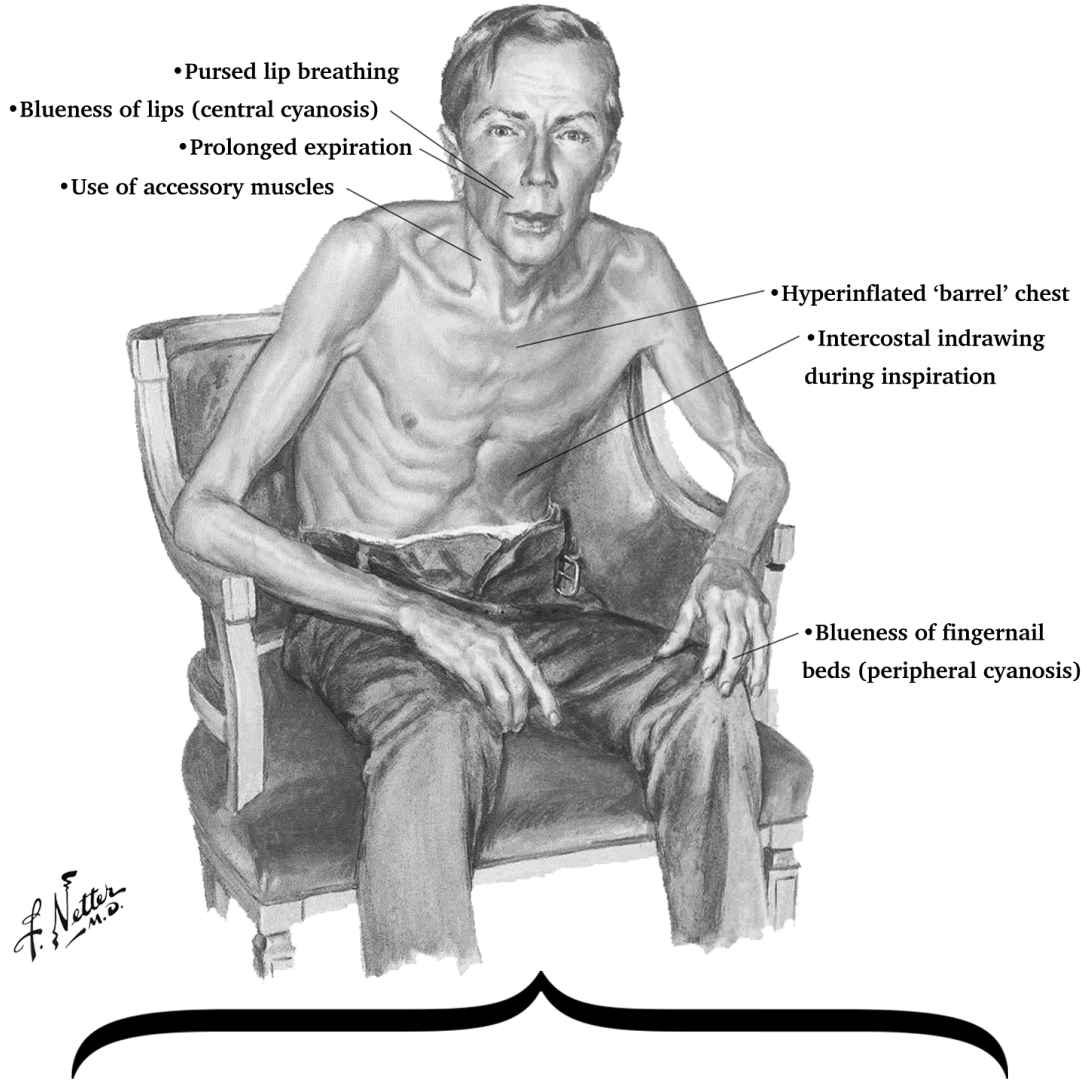
manifestations, it is suggested that COPD should be considered a systemic disease and research should focus on the metabolic and musculoskeletal manifestations.[218]

1.1.3 Incidence and prevalence of COPD

Globally, COPD is highly prevalent and positioned within the top ten of diseases with highest disease burden, measured by disability-adjusted life years, and increases with advancing age.[106] In particular in the developing world, the global burden of disease is expected to increase, primarily due to ageing populations and an increasing number of smokers.[239] In the United Kingdom (UK), approximately 1.2 million individuals are diagnosed with COPD with more than 100.000 newly diagnosed each year according to the British Lung Foundation. Diagnosis of COPD under the age of forty is uncommon but nearly 5% of people >40 have diagnosed COPD. Although COPD is dominated by men, within the UK there has been an increasing number of women with physician-based diagnosis of COPD. This particularly has been the case for socio-economically deprived areas and the northern-east part of the UK, affected by a threefold higher risk of COPD compared to more affluent areas. In terms of overall incidence of COPD in the UK, however, there has not been a significant change in recent years.[239] Incidence rates for men were 2.1/1000 people and 1.8/1000 for women in 2012.

1.1.4 Clinical outcomes

Common clinical outcomes in COPD include premature mortality, acute exacerbation of COPD, and hospital admission. Globally, the number of deaths related to COPD is highest in Bulgaria followed by North Korea and Greece with the UK positioned twelfth.[134] These high number of deaths reflect past smoking patterns, levels of air pollution, occupational exposures, and poorly ventilated indoor cooking fires. In the UK, premature mortality from COPD is almost twice as high compared the European average.[290] Hence, the Department of Health introduced a strategic agenda aimed at preventing, identifying, and treating COPD in earlier stages of disease hoping to improve life expectancy and quality of life.[205] The increasing number of governmental



Pulmonary manifestations

- Dyspnoea
- Wheezing
- Chronic cough
- Progressive exercise intolerance
- Sputum production
- Alteration in mental status
- Chest tightness
- Excess mucus in lungs
- Frequent respiratory infections

Systemic manifestations

- Decreased fat-free mass
- Impaired systemic muscle function
- Osteoporosis
- Anemia
- Depression
- Pulmonary hypertension
- Cor pulmonale
- Left-sided heart failure
- Unintended weight loss
- Lack of energy

Figure 1.3: Signs and symptoms of chronic obstructive pulmonary disease. Netter image modified.

publications, initiatives, and allocation of research funding emphasises the importance of addressing COPD and its consequences. In 2017, pollution has been the central topic of the Chief Medical Officer′s annual report, including recommendations targeted at improving the quality of air we breathe.[68] In 2018, to address the impact of common clinical outcomes, the National Health Services (NHS) England published the COPD RightCare Pathway report with a focus on early identification and accurate diagnosis to improve long-term care and management.[196] Other initiatives include funding allocation to the NHS Trust supporting the development of a self-management system for those affected, a National COPD Audit Programme aimed at improving provided care, and the NHS Business Services Authority Respiratory Dashboard that focuses on COPD exacerbations and inform best practice for drug prescription.

1.1.5 Diagnosing COPD

Diagnosing COPD relies heavily on spirometry results and is the most widely used marker for diagnosis and grading disease severity. An individual is required to exhale air completely from their lungs as quickly as possible after inhalation, expressed as the forced expiratory volume in one second (FEV_1). The volume of air exhaled after maximum inhalation indicates the forced vital capacity (FVC). Both measures are generally reduced with a lower score indicating worse lung function. Airflow limitation that is not fully reversible is confirmed by the presence of a post-bronchodilator FEV_1 <80% of predicted value in combination with an FEV_1/FVC <70% (**Figure 1.4**, page 8). These values have also been used by GOLD to categorise patients according to disease severity.[107] Initial categorisation was based on stages, according to lung function performance only, but have been replaced with groups based on COPD assessment test (CAT) or Medical Research Council (MRC) dyspnoea score and exacerbation history leading to hospital admission or not, in addition to airflow limitation. More recently, a system to classify COPD patients, primarily to address under-diagnosis, and assessment and management of COPD, was suggested by Agusti *et al.* where all smokers′ regardless of cough, dyspnoea or sputum should undergo spirometry. Agusti suggested differentiating between 'Simple COPD′ – referring to

those who are younger than 65 years of age with only mild/moderate airflow limitations and few symptoms and would be categorised as GOLD A, and 'Complex COPD′ – referring to all other patients who are more symptomatic and should be referred to specialist care, categorised as GOLD B, C and D.[4]

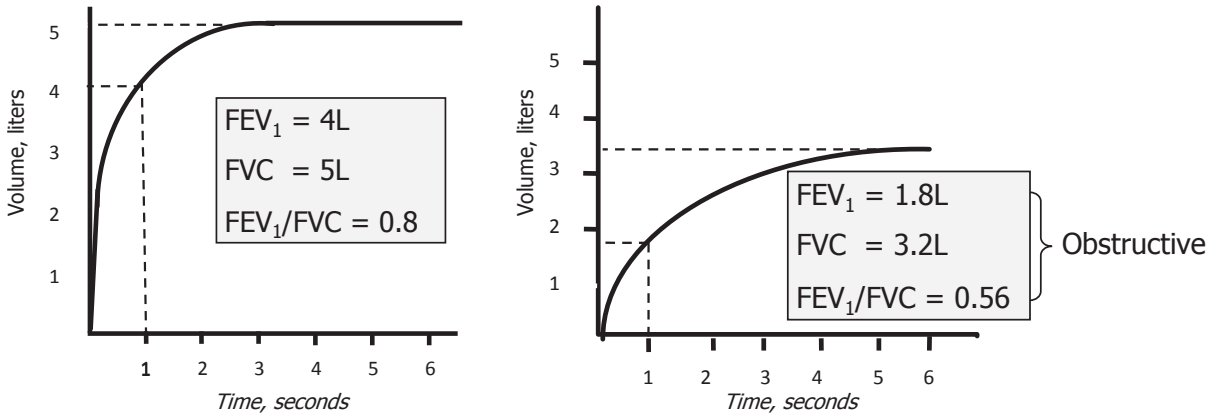


Figure 1.4: Spirometry results indicating lung obstruction. Image taken from the Global Initiative for Chronic Obstructive Lung Disease 2016 report.[107]

1.1.6 Established risk factors

Smoking is in most cases what causes COPD related death, followed by air pollution and particulates in household and occupation.[106] One of the most important and cost-effective interventions considered that could prevent COPD and improve lung function and survival, is smoking cessation.[29,261] However, especially amongst those with lower socio-economic status, there has been no significant reduction.[244] An increasing importance is placed on minimising risk factors and disease progression through prevention other than smoking cessation, as COPD is still largely underdiagnosed and therefore undertreated due to the current diagnosis that primarily relies on lung function measurements failing to capture the heterogeneity of the disease.[48,59,75,122,226]

1.1.7 Predicting clinical outcomes

It was believed that over time patients would worsen with increasing airflow limitation. However, clinicians have come to realise the disease is much more heterogeneous than initially thought and existing measures such as FEV_1 might not be appropriate as they fail to capture systemic disease. Replacing GOLD stages with groups based on symptoms and history of exacerbation in addition to airflow limitation, however, has not led to a significant difference being measured in terms of their ability to predict hospital admission and mortality.[141] Identifying individuals at high risk for common clinical outcomes has remained difficult. There is a growing interest in systemic manifestations. Assessing the predictive value of cardiovascular (CV) abnormalities, skeletal muscle weakness and plasma biomarkers for clinical outcomes are recognised to be of increasing clinical importance. In particular manifestations that can easily be measured in clinical practice and that support early stage detection. For example, inflammatory markers such as C-reactive protein (CRP) and fibrinogen have been assessed for their association between COPD and systemic inflammation and higher levels are often found in those diagnosed with COPD compared to those without. Other biomarkers such as reduced walking distance or lower limb muscle strength (i.e. muscle weakness) may well be predictors of clinical outcomes. Exercise testing such as six-minute walk (6MW) distance and cardiopulmonary exercise testing have already been recommended in the European Respiratory Society[211] and American Thoracic Society[11] guidelines.

1.1.8 Prognostic model development

Assessing the predictive value of biomarkers can be evaluated through the development of a multivariable prediction model. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines inform about the development, validation and updating of prediction models.[52] When combining prognostic factors that have been identified as important predictors for a particular outcome such as mortality or AECOPD into a multidimensional index, they indicate disease severity or levels of risk and can be used in

clinical practice by healthcare providers supporting in their decision making (**Figure 1.5**, page 11). Alternatively, they can help stratify patients for clinical trials and interventions (e.g. drug treatment, smoking cessation, or physical rehabilitation) aimed at modifying disease outcomes (e.g. AECOPD, mortality or hospital admission). The performance (i.e. discriminative ability) of such models is commonly expressed as the area under the receiver operating characteristic curve,[177] Somers′ D-statistic,[202] or Harrell′s C-statistic.[10] A model with a discriminative ability of 0.5 equates to random chance, whereas a performance of 1.0 is perfect prediction. As a rule of thumb, a C-statistic of 0.5-0.7 is considered weak, 0.7-0.8 good, and >0.8 very good.[129] However, there are no clear guidelines on the minimum level of model accuracy required. A persistent problem with risk prediction in lung disease, however, is the poor discriminative ability of predictive models. In addition, some models are based on too small datasets, lack statistical validation, fail to address missing data, are badly calibrated, or over-fit the data.[27,155] Simultaneously, typically little attention is given to the clinical practicalities such as cost, complexity, patient burden and time required, limiting the impact (i.e. improving patient outcomes) and widespread adoption of prediction models in clinical practice. Using electronic health record (EHR) data has been suggested as a potential solution to improve the uptake of prediction models.[108] Improvement of the presentation of models and the inclusion of provider and patient preferences are alternative examples that have been suggested to improve uptake.[144]

1.1.9 Multidimensional scoring systems

With the increased attention for extra-pulmonary manifestations in COPD, numerous scoring systems have been suggested and introduced aiming to improve prediction in clinical practice. For example, Celli and colleagues developed and validated a weighted multidimensional grading system based on body mass index, the degree of airflow obstruction, dyspnoea, and exercise capacity measured by the 6MW test (BODE). The BODE Index is better than FEV_1 to assess disease severity and predicting risk of death,[40] and found to be superior in predicting hospital

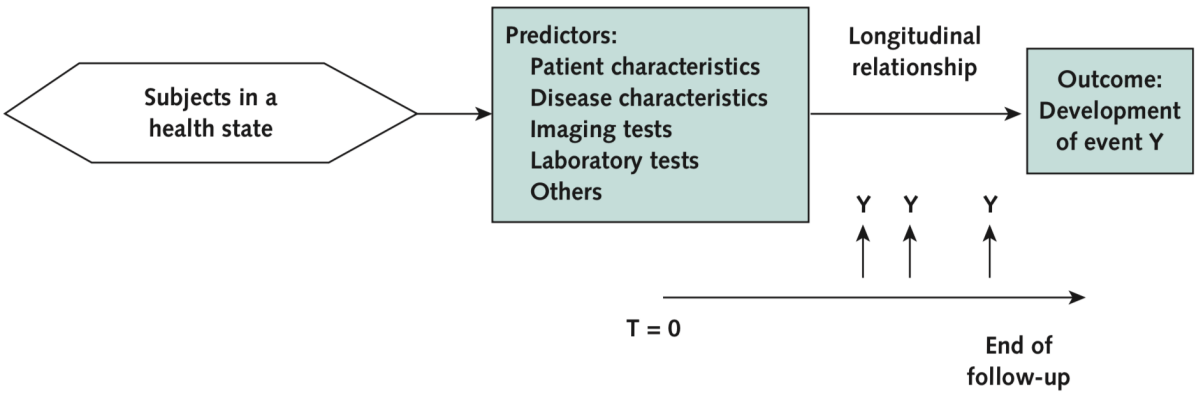


Figure 1.5: Prognostic multivariable modelling study. *Abbreviations*: T, moment of prediction. Y, time of the event. Image taken from the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement.[52]

admission when compared to the GOLD classifications.[208] Both GOLD and the BODE Index, however, appear to perform poorly at predicting mortality.[45] In addition, Puhan *et al.* concluded that the original BODE Index failed to accurately predict mortality and introduced an updated index (U-BODE) by re-assigning the number of points given to the variables included, in addition to developing a simplified ADO index based on age, dyspnoea and obstruction.[225]

Fisk *et al.* (2016, unpublished) compared three of these classification systems: GOLD stage, GOLD group, and BODE whilst assessing the relationship between COPD severity and the two most clinically important extra-pulmonary manifestations: CV disease and muscle weakness. The authors found that the cut-off points used in GOLD groups are incorrect and suggest the use of BODE Index quartiles in the evaluation of vascular and musculoskeletal phenotypes, in particular arterial stiffness and quadriceps weakness.

Considering COPD is irreversible, and treatments are still lacking, the focus of managing COPD is shifting towards prevention. Evaluating these systemic manifestations in COPD are considered to be a key unmet need, with some clinicians arguing to move away from disease labels such as COPD and focus on so-called ′treatable traits′ such as improving quadriceps muscle strength, addressing disease complexity and overlap of symptoms amongst diseases.[4]

1.1.10 Novel biomarkers

1.1.10.1 Inflammatory markers

Biomarkers of interest that are believed to be of predictive value include multiple inflammatory markers and CV and musculoskeletal measures. See **Appendix C** for detailed descriptions. For example, fibrinogen, white cell count (WCC), interleukins 6 (IL-6) and 8 (IL-8), tumour necrosis factor-alpha (TNF-α), and CRP are measures of inflammation and may be related to muscle or CV problems in COPD patients. Both CRP and fibrinogen are acute phase reactants and considered key regulators of inflammation. Levels increase with the presence of acute systematic inflammation. In 2015, the Food and Drug Administration approved fibrinogen as a prognostic marker for mortality and COPD exacerbations.[179] White cell count (i.e. leukocytes) are essential part of the immune system with elevated levels indicating inflammation. Persistent systemic inflammation is linked to poor clinical outcomes.[6] Interleukins 6 and 8, with IL-8 being a leukocyte chemotactic cytokine – low molecular weight proteins that stimulate recruitment of leukocytes – produced by various cells, are also inflammatory markers and play a key role in immune responses and production is activated by inflammatory stimuli.

1.1.10.2 Cardiovascular markers

Cardiovascular markers of interest include pulse wave velocity (PWV), augmentation index (AIx), and carotid intima-media thickness (CIMT). For example, elevated PWV has been reported in patients with COPD but the predictive value is not yet known. Similarly, these markers are predictors in other populations and have the potential of clinically predicting CV disease in COPD. In addition, despite the lack of widespread assessment, these markers are already known to a majority of respiratory and CV physicians. The PWV indicates the velocity of circulating blood flow and is a measure of vessel stiffness. Augmentation index reflects the pressure from the ascending aortic (i.e. central wave) and influences central blood pressure. Both can be measured using pulse wave analysis, a simple and reliable method.[286] Higher values reflect increasing pres-

sure on the arterial system. It is suggested that those with systemic inflammation and vascular dysfunction may be more likely to suffer from CV disease and mortality.[120,181] Findings in the Anglo-Cardiff Collaborative Trial (ACCT), a large cohort of healthy normotensive individuals, suggest AIx to be an appropriate measure for those aged <50, and PWV to be an appropriate measure for individuals aged >50.[173] Systematic analysis of PWV in predicting CV disease suggested model improvement to identify high-risk populations.[19] Carotid intima-media thickness is a non-invasive measure of atherosclerotic burden but also reflects arterial remodelling, and is used as a proxy measure for CV disease. Albeit a weak predictor in fully adjusted models, evidence suggest increased CIMT to be linked with a higher risk of future cerebrovascular and CV events.[25,26] In the ARIC (Atherosclerosis Risk In Communities) Study[195] with nearly 16.000 individuals recruited from the general population, and The Multi-Ethnic Study of Atherosclerosis (MESA) study with nearly 7000 adults with absence of clinically diagnosed CV disease,[95] CIMT was found to improve the predictive ability for coronary heart disease. Despite the suggestion of CIMT being an important risk factor for determining CV outcomes such as stroke and myocardial infarction, conclusive evidence for its usefulness in COPD is lacking.[262]

1.1.10.3 Musculoskeletal markers

In recent years there has also been an increasing interest in examining the predictive value of functional activities of the musculoskeletal system. Exercise limitation (i.e. impaired exercise tolerance) including musculoskeletal weakness is common within a COPD population, especially during and after acute exacerbation of COPD (AECOPD).[221] Level of impairment is thought to correspond with lung function but is prevalent in all stages of COPD.[77] Determinants of exercise capacity include 6MW distance and quadriceps strength.[109] The 6MW test is a well-known test to assess exercise intolerance and can evaluate an individual′s functional exercise capacity. It is primarily used in chronic respiratory disease and heart failure. The goal of the 6MW is to walk as far as possible in six minutes. The test has been evaluated for its ability to predict mortality, hospitalisation and exacerbation in a variety of diseases including chronic heart

failure, peripheral arterial occlusive disease,[187] and COPD allowing stratification of patients for clinical trials.[38] Alternative measures include resting heart rate (i.e. measure of cardiac efficiency) and musculoskeletal measures short physical performance battery (SPPB), quadriceps maximum voluntary contraction (QMVC), and sniff nasal inspiratory pressure (SNIP). The SPPB is a battery of tests (i.e. four-metre gait speed, balance, chair stand) used to evaluate the physical performance of the lower extremities.[116] Quadriceps muscle weakness is a surrogate marker of functional activity (i.e. quadriceps muscle strength) and found in 30-40% of COPD patients, regardless of breathlessness or the level of airflow limitation.[237,292] Quadriceps muscle weakness is an indicator of declining muscle function and can be estimated through, for example, predicted quadriceps strength – a regression equation incorporating age, gender, height and fat-free mass.[237] Sniff nasal inspiratory pressure is a surrogate marker of respiratory muscle function (i.e. inspiratory muscle strength) and can be used to identify respiratory muscle weakness. On the contrary to lung function, musculoskeletal functioning can be trained and improved with exercise training. Also, efforts are made in developing new medications to target musculoskeletal dysfunction. Not only does preventing or limiting declining exercise tolerance and physical training improve mortality, it also improves quality of life.[171]

Current evidence on the association between these selected CV and musculoskeletal biomarkers and the occurrence of clinical outcomes within a COPD population, however, is limited.

1.1.11 Clinical studies

Multiple cohort studies aimed at improving outcomes in COPD of which several are addressing the need for novel biomarkers in predicting common clinical outcomes exist. Examples include US-based "Exercise in Health and Chronic Obstructive Pulmonary Disease" study that aimed to assess oxidative stress in relation to peripheral muscle dysfunction (NCT02300064), the French "Role of Systemic Inflammation in Increase of Cardio-vascular Risk in Chronic Obstructive Pulmonary Disease (BPCO)" study (NCT02888886) that assessed the relationship between inflammatory markers and CV mortality, and the Belgium-based "Systemic Consequences and

Comorbidities in Mild/Moderate Chronic Obstructive Pulmonary Disease (COPD), Time for Action!" study (NCT01314807) that evaluates skeletal muscle dysfunction and physical inactivity in COPD. UK-based studies, which are part of a partnership between multiple medical centres with an interest in COPD, and Innovate UK and GlaxoSmithKline include the ECLIPSE,[270] Assessment of Risk in Chronic Airways Disease Evaluation (ARCADE),[101] and Evaluating the Role of Inflammation in Chronic Airways disease (ERICA) studies.[184] All three form part of a consortium that has the overarching aim to fill the biomarker gap and support stratified medicine. The ERICA study, however, is rather unique and differs from all other cohort studies through capturing multiple musculoskeletal and CV measures in addition to lung performance and questionnaire data. In addition, the cohort allows linkage with EHR data and survival data.

1.2 Thesis outline

Spirometry has limitations, COPD is more complex and heterogeneous than initially thought, and effective treatments are lacking. In particular due to the absence of treatment and the frequency of adverse outcomes, early diagnosis and identifying high risk individuals in an early stage of disease is crucial. The availability of EHR data has sparked the hope of being able to capture the heterogeneity of COPD, and assess the relationship between new potential risk factors and common clinical outcomes. Key scientific questions include which phenotypic traits (i.e. CV and musculoskeletal) influence treatment and clinical outcomes? Do EHR data enable to capture the various disease manifestations of COPD? Can COPD management be improved to better identify high risk individuals in an early stage of disease, allowing timely intervention, and treating the right patients?

The overall aim of my doctoral research is to identify and evaluate the relationships between existing and novel biomarkers, and questionnaire data, and EHR data and determine if and how these biomarkers can predict common clinical outcomes (i.e. AECOPD, hospitalisation, mortality) within a COPD population. **Chapter 2** describes systematically synthesised published

evidence on the associations between selected CV and musculoskeletal biomarkers that are not yet widely used in clinical practice but may potentially better capture systemic problems in COPD than conventional measures, and the occurrence of clinical outcomes including exacerbations, hospitalisation, and mortality within a COPD population. **Chapter 3** describes the ERICA cohort, baseline values of key variables captured, and differences between recruitment sites and sex, and examined the presence of missing data and relationships between variables of interest. **Chapter 4** evaluates the association between the measures of skeletal muscle function and all-cause mortality in stable COPD patients, and with the assumption that a relationship would be found to investigate whether a BODE Index in which the 6MW component was replaced by alternative musculoskeletal measures retained predictive ability when predicting death. **Chapter 5** describes cause of death in a COPD defined population identified in the UK Biobank, and compares survival risk with a non-COPD defined population. Comparisons between findings in the UK Biobank and the ERICA cohort are made. **Chapter 6** evaluates the relationship between musculoskeletal measures and the risk of hospital admissions due to AECOPD, and determined a relationship between musculoskeletal measures and length of hospital stay for initial AECOPD using routinely collected hospital electronic health records and clinical data from the ERICA cohort. **Chapter 7** describes incidence of fatal- and non-fatal CV disease within COPD, and evaluated the association of classical Framingham risk factors with subsequent fatal and non-fatal CV events in stable COPD patients using clinical data from the ERICA cohort and EHR data, and determined the association of measures of arterial stiffness and incident CV disease and their added value above and beyond Framingham risk factors. In addition, it describes the association of alternative measures including musculoskeletal function, thought to better capture systemic problems, and CV disease, and their added value above and beyond Framingham risk factors. **Chapter 8** summarises and discusses the evidence available before this dissertation, the added value of the work presented in this doctoral research, and discusses the implications of all the available evidence. In addition, analysis for future research are proposed. **Appendix A** provides a list of research items I authored during the PhD. **Ap-**

pendix B includes a data dictionary I created providing descriptions and ranges of values of the variables (i.e. baseline and follow-up questionnaire data) captured in the ERICA study. **Appendix C** includes the ERICA study protocol. **Appendix D** includes the data completion form I created to collect data for the meta-analysis. **Appendix E** includes a table with cause of death descriptions.

2

Biomarkers and clinical outcomes in COPD – a systematic review & meta-analysis

Chapter summary

Background Conventional measures to evaluate chronic obstructive pulmonary disease (COPD) may fail to capture systemic problems particularly musculoskeletal weakness and cardiovascular disease. Identifying these manifestations and assessing their association with clinical outcomes (i.e. mortality, exacerbation, and COPD hospital admission) is of increasing clinical importance.

Objective To assess associations between six-minute walk distance, heart rate, fibrinogen, C-reactive protein, white cell count, interleukins 6 and 8, tumour necrosis factor-alpha, quadriceps maximum voluntary contraction, sniff nasal inspiratory pressure, short physical performance battery, pulse wave velocity, carotid intima-media thickness and augmentation index, and clinical outcomes in patients with stable COPD.

Methods We systematically searched electronic databases (August 2018) and identified 61

studies, which were synthesised, including Meta-Analyses and Systematic Reviews of Observational Studies (MOOSE) guidelines and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

Findings Shorter six-minute walk distance and elevated heart rate, fibrinogen, C-reactive protein and white cell count were associated with higher risk of mortality. Pooled hazard ratios were 0.80 (95% CI 0.73 to 0.89) per 50 m longer six-minute walk distance, 1.10 (95% CI 1.02 to 1.18) per 10 bpm higher heart rate, 3.13 (95% CI 2.14 to 4.57) per twofold increase in fibrinogen, 1.17 (95% CI 1.06 to 1.28) per twofold increase in C-reactive protein and 2.07 (95% CI 1.29 to 3.31) per twofold increase in white cell count. Shorter six-minute walk distance and elevated fibrinogen and C-reactive protein were associated with exacerbation, and shorter six-minute walk distance, higher heart rate, C-reactive protein and interleukin-6 were associated with hospitalisation. Few studies examined associations with musculoskeletal measures.

Conclusion Findings suggest six-minute walk distance, heart rate, C-reactive protein, fibrinogen, and white cell count are associated with clinical outcomes in stable COPD patients. Use of musculoskeletal measures to assess outcomes in COPD patients requires further investigation.

2.1 Background

Chronic obstructive pulmonary disease (COPD) is one of the leading causes of death worldwide, with prevalence of 5.6% (3.2 million) in 2015 projected to increase to 7.8% by 2030.[172] The consequent socio-economic burden of COPD is high, causing reduced quality of life, loss of productivity, increased hospital admissions and premature mortality.[206,289] One important and cost-effective intervention is smoking cessation.[29,261] However, increasing importance is placed on improving risk factors and slowing down disease progression by addressing non-pulmonary aspects of the condition.[48,59,75,122,226]

Spirometry is the most widely used marker of disease severity and progression. No longer is it believed that all patients will worsen over time with increasing airflow limitation. Clinicians have

now identified that COPD is more heterogeneous than initially thought and existing measures such as forced expiratory volume in one second (FEV_1) may fail to capture systemic disease[5] and have divergent trajectories.[154]

Chronic obstructive pulmonary disease also leads to systemic problems, such as skeletal muscle weakness and cardiovascular (CV) disease, the latter accounting for a third of deaths in COPD.[174] While multiple studies have shown that quadriceps involvement in COPD is associated with worse outcomes,[171,215,259] it has also been postulated that these features result from an increase in inflammatory markers like C-reactive protein (CRP) and fibrinogen,[43] with a spill-over effect of inflammatory response proposed as the underling mechanism.[272] Thus capturing systemic manifestations such as exercise intolerance, CV abnormalities, skeletal muscle weakness, and plasma biomarkers are recognised to be of increasing clinical importance.[17]

We aim to systematically synthesise the published evidence on the associations between selected CV and musculoskeletal biomarkers that are not yet widely used in clinical practice but may potentially better capture systemic problems in COPD than conventional measures, and the occurrence of clinical outcomes including exacerbations, hospitalisation, and mortality within a COPD population. Individual studies and a limited number of reviews[47,151,158] have assessed the association between selected biomarkers and clinical outcomes, however, to our knowledge no published study has systematically synthesised this evidence.

2.2 Methods

2.2.1 Search strategy

The review protocol was registered with PROSPERO (CRD 42016052075). The systematic review includes electronic searches in the Ovid versions of MEDLINE, Embase, Cochrane Library, CINAHL, and Web of Science. Search terms related to pulmonary disease were combined with terms related to CV and musculoskeletal measure, clinical outcome, and study design (**Table 2.1**, page 23). A meta-analysis was carried out following meta-analysis of observational studies

in epidemiology (MOOSE) guidelines and reported in accordance with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines to identify prospective studies assessing the relationship between CV or musculoskeletal measures and the occurrence of clinical outcomes in COPD.[185,254]

2.2.2 Biomarkers and outcomes

Biomarkers that may capture systemic problems in COPD and are not yet widely used in clinical practice were included: six-minute walk (6MW) distance, resting heart rate, quadriceps maximum voluntary contraction (QMVC), sniff nasal inspiratory pressure (SNIP), short physical performance battery (SPPB), pulse wave velocity (PWV), carotid intima-media thickness (CIMT), and augmentation index (AIx). Data relevant to inflammation were fibrinogen, CRP, white cell count (WCC), interleukin-6, (IL-6) -8 (IL-8), and tumour necrosis factor-alpha (TNF-α). Clinical outcomes of interest included: mortality, exacerbation, and hospitalisation. Mortality was defined as all-cause mortality. Exacerbation was defined as patients who either had a change in medication, which required increase or initiation of steroids or antibiotics, or were admitted to hospital due to COPD. Hospitalisation, a subset of COPD exacerbation by definition, was limited to only exacerbations that resulted in admissions related to COPD.

Two reviewers independently completed the selection and review of articles. Full-text papers and reviews found in the initial search were cross-referenced. Studies that satisfied the full-text paper selection criteria included: (i) primary research; (ii) had a sample size $\geq$ 50 with COPD; (iii) assessing a relevant biomarker; (iv) full-text paper in English; (v) a general population (e.g. not a single gender); (vi) did not include unstable COPD patients (e.g. currently in acute exacerbation, currently hospitalised or recruited on discharge); and (vii) were prospective studies with a follow-up period $\geq$ 6 months.

Table 2.1: Search strategies

Terms related to pulmonary disease	Terms related to selected CV and musculoskeletal biomarker	Terms related to clinical adverse outcomes	Terms related to study design
1. exp Pulmonary Disease, Chronic Obstructive/	9. biological marker.tw.	36. exp Cardiovascular Diseases/	52. exp Cohort Studies/
2. chronic obstructive pulmonary disease.tw.	10. systemic inflammation.tw.	37. cardiovascular disease.tw.	53. cohort stud$.tw.
3. COPD.tw.	11. exp Leukocytes/	38. exp Hospitalization/	54. exp Prospective Studies/
4. pulmonary emphysema.tw.	12. exp Interleukin-6/	39. hospitali$.tw.	55. prospective.tw.
5. chronic bronchitis.tw.	13. exp Interleukin-8/	40. patient admission.tw.	56. longitudinal stud$.tw.
6. exp Forced Expiratory Volume/	14. exp Fibrinogen/	41. exp Death/	57. exp Case-Control Studies/
7. exp Vital Capacity/	15. exp Tumor Necrosis Factor-alpha/	42. death.tw.	58. case-control stud$.tw.
8. or/1-7	16. exp C-Reactive Protein/	43. exp Mortality/	59. exp Randomized Controlled Trials as Topic/
	17. exp Carotid Intima-Media Thickness/	44. mortality.tw.	60. rct.tw.
	18. CIMT.tw.	45. outcome.tw.	61. or/52-60
	19. exp Pulse Wave Analysis/	46. exp Prognosis/	62. 8 and 35 and 51 and 61
	20. pulse wave velocity.tw.	47. prognos$.tw.	
	21. PWV.tw.	48. exp Survival Analysis/	
	22. augmentation index.tw.	49. survival.tw.	
	23. AIx.tw.	50. exacerbation.tw.	
	24. exp Heart Rate/	51. or/36-50	
	25. 6 minute walk$.tw.		
	26. 6mwt.tw.		
	27. 6mwd.tw.		
	28. exp Quadriceps Muscle/		
	29. quadriceps max$ voluntary contraction.tw.		
	30. qmvc.tw.		
	31. sniff nasal inspiratory pressure.tw.		
	32. snip.tw.		
	33. short physical performance battery.tw.		
	34. sppb.tw.		
	35. or/9-34		

*Terms related to pulmonary disease were combined with terms related to cardiovascular and musculoskeletal biomarker, clinical outcome and study design. Search strategy was used with Medline (Ovid) and modified as necessary for use with the other databases. For medical subject headings (MeSH) terms, all subheadings selected. *Abbreviations*: exp, exploded MeSH term. tw, text word.

2.2.3 Data extraction and quality assessment

Where possible, adjusted (i.e. age, sex, body mass index (BMI), and smoking status) and unadjusted hazard ratios (HR) for mortality were collected, as well as model performance measures (e.g. C-statistic). Sample sizes, mean values, and standard deviations of the biomarkers for individuals with and without the event (i.e. mortality, exacerbation, or hospitalisation) were extracted from published studies to estimate standardised mean differences (SMD). Where data were not published, the corresponding authors were contacted and asked to provide data by completing a data collection form (**Appendix B**). Three reminders were sent over a period of four months. For studies reporting the same cohort, data from the study with the most completed outcome data, largest sample size, or with the longest follow-up were used. The quality of each study was based on QUADAS-2 quality assessment criteria.[285] Scoring was based on the follow-up period, sample size, reporting of adjustment factors, method of defining COPD, age of study participants, and study type (**Table 2.2**, page 27). Scores range from 0-15, where fifteen is considered of highest quality.

2.2.4 Statistical analysis

To synthesize and analyse quantitative data, while accounting for heterogeneity by incorporating between study variability of effect sizes, results from the studies were assessed with random effects meta-analysis. Data were graphically displayed using forest plots. Where necessary and possible, HRs were converted to the selected unit effect measure. Hazard ratios for log-transformed biomarkers represent a twofold increase in the biomarker. To address uncertainty, we excluded studies with a quality score in the bottom QUADAS-2 score tertile (1^{st}, 15-12/ 2^{nd}, 11-9/ 3^{rd}, 8-0). Funnel plots – scatterplots of observed outcome against the standard error – were generated to assess potential publication bias. Asymmetry in the plots may indicate publication bias. Galbraith plots – radial plots of the ratio of observed outcome to standard error against reciprocal of standard error – were generated to assess heterogeneity in effect sizes.[124] Results from a fixed-effects meta-analysis were compared against those from a random-effects meta-

analysis. Finally, meta-regression was conducted where possible to analyse the impact of length of follow-up, year of publication, and the mean age of the cohort. Trend analysis was performed using ANOVA.

2.3 Findings

The systematic review yielded 2852 unique references from five electronic databases. After screening the abstracts, 61 articles met the selection criteria (**Figure 2.1**, page 26 and **Table 2.3**, page 29). The age of participants of the included studies ranged from 40-80 years of age, with an approximate median age of 65 years. The sample sizes ranged from 53-20192 subjects, with a median size of 237. The follow-up period ranged between six and 423 months, with an approximate median time of 36 months. The evaluation of COPD longitudinally to identify predictive surrogate end-points (ECLIPSE) and BMI, airflow obstruction, dyspnoea and exercise (BODE) cohorts were the most studied cohorts.

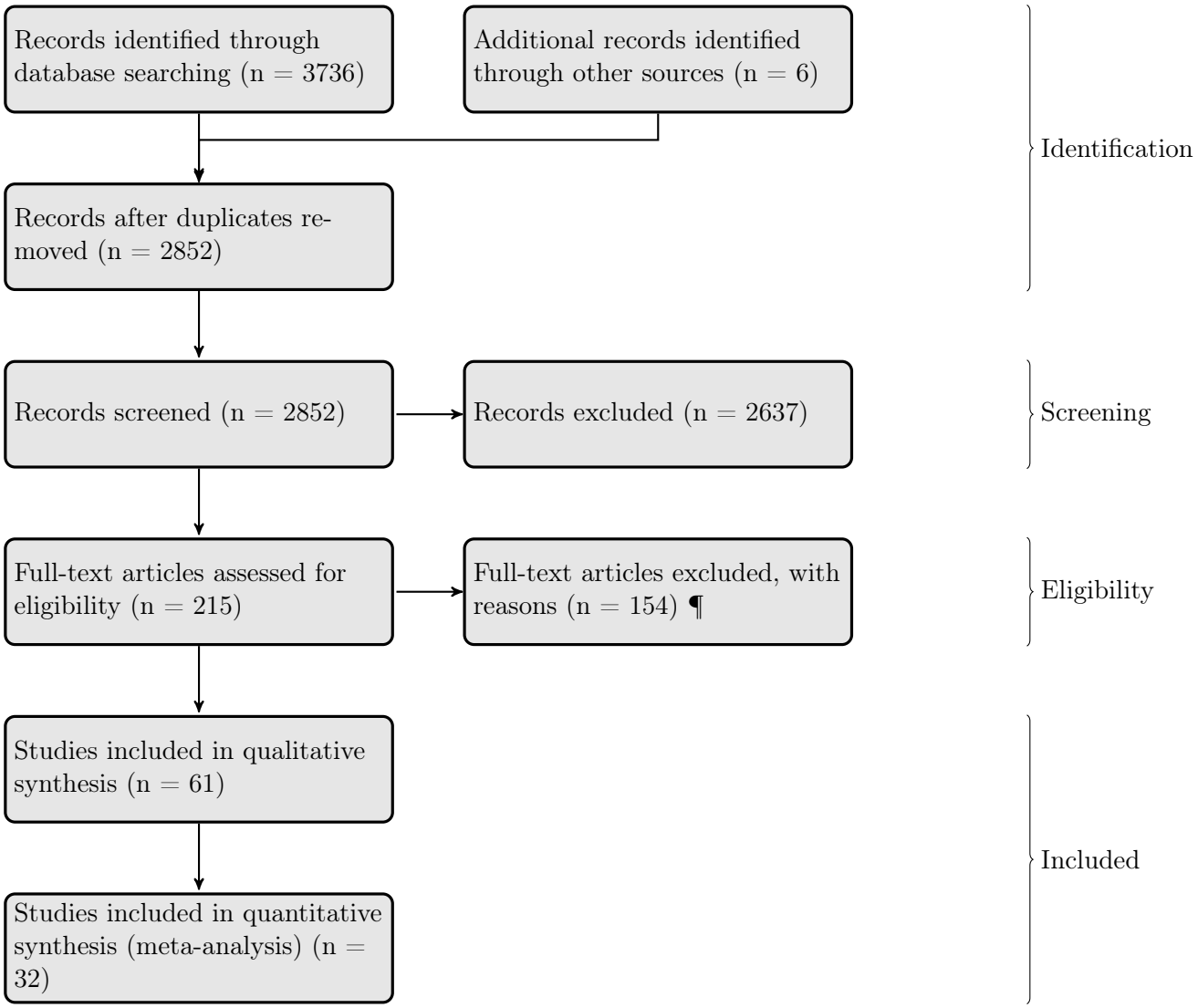


Figure 2.1: Flow diagram of studies included in the review, based on the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols.[185] Records identified: MEDLINE, n = 1175; Embase, n = 1597; Cochrane, n = 56; CINAHL, n = 143; Web of Science, n = 765. *Abbreviations*: COPD, chronic obstructive pulmonary disease.
¶Reasons for exclusion: only conference abstracts available (n = 41); cross-sectional i.e. there was no follow-up (n = 29); sample size of <50 with COPD (n = 18); follow-up period of <6 months (n = 21); other reasons e.g. single gender assessment (n = 17); assessed unstable COPD patients i.e. acute exacerbation or hospitalised (n = 13); not primary research e.g. review papers (n = 7); no assessment of relationship between COPD and relevant outcome (n = 6); full text unavailable in the English language (n = 2)

Table 2.2: QUADAS-2 scoring.

Study ID	Follow-up period	Sample size	Reporting of sample characteristics	Method of defining COPD	Age study participants	Study type	Quality score	Tertile
Agu002	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2] <65	[1] Cohort or case-control	15	3/3
Agu003	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	15	3/3
Dah059	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2] <65	[1] Cohort or case-control	15	3/3
Dah061	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2] <65	[1] Cohort or case-control	15	3/3
Jen108	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	15	3/3
Man128	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	15	3/3
Man129	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	15	3/3
Par148	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	15	3/3
Val187	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	15	3/3
Hur203	4]>36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	15	3/3
Cas038	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	14	3/3
Cot051	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	14	3/3
Tho180	[4]>36 months	[4]>500	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	14	3/3
Cel039	[2] 12-36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	13	3/3
Mil135	[2] 12-36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	13	3/3
Mul141	[2] 12-36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	13	3/3
Spr174	[2] 12-36 months	[4]>500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	13	3/3
Dah060	[4]>36 months	[4]>500	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	12	3/3
Blu029	[4]>36 months	[2] 250-500	[2] All four	[2] Spirometry	[1] 65-70	[0] RCT	11	2/3
Hus104	[2] 12-36 months	[2] 250-500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	11	2/3
Sin172	[2] 12-36 months	[2] 250-500	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	11	2/3
Swa178	[4]>36 months	[0]<250	[2] All four	[2] Spirometry	[2] <65	[1] Cohort or case-control	11	2/3
War193	[4]>36 months	[2] 250-500	[2] All four	[2] Spirometry	0]>70	[1] Cohort or case-control	11	2/3
Was195	[4]>36 months	[0]<250	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	11	2/3
Can037	[2] 12-36 months	[2] 250-500	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	10	2/3
Cot053	[4]>36 months	[2] 250-500	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	10	2/3
Cot054	[4]>36 months	[2] 250-500	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	10	2/3
Gro097	[2] 12-36 months	[2] 250-500	[2] All four	[2] Spirometry	[2]<65	[0] RCT	10	2/3
Gro099	[4]>36 months	[2] 250-500	[2] All four	[2] Spirometry	0]>70	[0] RCT	10	2/3
Liu123	[2] 12-36 months	[2] 250-500	[2] All four	[2] Spirometry	[2]<65	[0] RCT	10	2/3
Liu124	[4]>36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	10	2/3
Pin155	[2] 12-36 months	[2] 250-500	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	10	2/3
Dur073	[4]>36 months	[2] 250-500	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[0] RCT	9	2/3
Fag078	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	9	2/3
Hop103	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[2]<65	[1] Cohort or case-control	9	2/3
Moo138	4]>36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[2]<65	[1] Cohort or case-control	9	2/3
Ozg144	4]>36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[2]<65	[1] Cohort or case-control	9	2/3
Cot052	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	8	1/3
de 065	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	8	1/3

Study ID	Follow-up period	Sample size	Reporting of sample characteristics	Method of defining COPD	Age study participants	Study type	Quality score	Tertile
Don069	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	8	1/3
Dre070	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	8	1/3
Gak090	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	8	1/3
Jen107	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	8	1/3
Lac119	4]>36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	8	1/3
Wed196	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	8	1/3
Bu,032	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	0]>70	[1] Cohort or case-control	7	1/3
Bud033	[2] 12-36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[2]<65	[1] Cohort or case-control	7	1/3
Den066	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	0]>70	[1] Cohort or case-control	7	1/3
Fer083	[2] 12-36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[2]<65	[1] Cohort or case-control	7	1/3
Mon137	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	0] RCT	7	1/3
Moy139	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	0]>70	[1] Cohort or case-control	7	1/3
Moy140	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	0]>70	[1] Cohort or case-control	7	1/3
Pow159	[2] 12-36 months	[0]<250	[2] All four	[2] Spirometry	[1] 65-70	0] RCT	7	1/3
Ant017	[2] 12-36 months	[0]<250	0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	6	1/3
Baf020	[2] 12-36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	6	1/3
Daj062	[2] 12-36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	6	1/3
Fer084	[2] 12-36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	6	1/3
Guo100	[2] 12-36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[2]<65	[0] RCT	6	1/3
See167	[2] 12-36 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	6	1/3
Wan191	[2] 12-36 months	[0]<250	2] All four	[2] Spirometry	0]>70	[0] RCT	6	1/3
Mar131	[0] <12 months	[0]<250	[0] Age + gender + BMI or smoking status	[2] Spirometry	[1] 65-70	[1] Cohort or case-control	4	1/3

Abbreviations: BMI, body mass index. RCT, randomised controlled trial.

Table 2.3: Search results systematic review (number of studies = 61).

Study	Country	Sample size (cohort name)	Average age or range	COPD definition	Disease severity	Follow-up (months)	QUADAS-2 score (out of 15)	Source	Markers(s)	Outcome(s)
Agusti *et al.*, 2012 [6]	Spain	1755 (ECLIPSE)	64	Spirometry	GOLD II-IV	36	15	Publication	6MWD, fibrinogen, CRP, WCC, IL-6, IL-8, TNF-α	Mortality, exacerbation
Agusti *et al.*, 2013 [5]	Spain	2101 (ECLIPSE)	64	Spirometry	GOLD II-IV	36	15	Publication	6MWD	Mortality, hospitalisation
Antonelli-Incalzi *et al.*, 2016 [13]	Italy	134	69	Spirometry	Severe COPD	>24	6	Questionnaire	6MWD, heart rate, fibrinogen, WCC	Mortality
Bafadhel *et al.*, 2011 [15]	UK	115	69	Physician-diagnosis and spirometry	GOLD I-IV	12	6	Publication	WCC, CRP	Exacerbation
Blumenthal *et al.*, 2016 [23]	USA	326 (INSPIRE-II)	66	Spirometry	GOLD A-D	60	11	Publication	6MWD, CRP	Mortality, hospitalisation
Bu *et al.*, 2001 [30]	China	56 (MLCC)	71 (median)	Spirometry	GOLD II-IV	24	7	Publication	6MWD	Exacerbation
Budweiser *et al.*, 2007 [31]	Germany	98	65	Sx, Spirometry	GOLD IV	>24	7	Publication	WCC	Mortality
Cano *et al.*, 2004 [33]	France	309	72	Spirometry	-	>8	10	Dataset	6MWD, CRP	Mortality, hospitalisation
Casanova *et al.*, 2008 [36]	USA, Spain	576	68	Hx, Spirometry	-	>36	14	Publication	6MWD	Mortality
Celli *et al.*, 2012 [42]	Spain	1843 (ECLIPSE)	64	Spirometry	GOLD II-IV	36	13	Questionnaire	6MWD, heart rate, fibrinogen, WCC, IL-6, IL-8, TNF-α	Mortality
Cote *et al.*, 2007 [56]	USA, Spain, Venezuela	205 (BODE)	67	Spirometry	GOLD I-IV	24	8	Publication	6MWD	Hospitalisation
Cote *et al.*, 2007 [57]	USA	365 (BODE)	68	Sx, Spirometry	GOLD I-IV	67	10	Publication	6MWD	Mortality
Cote *et al.*, 2008 [55]	USA, Spain, Venezuela	1379 (BODE)	66	Spirometry	GOLD I-IV	>25	14	Publication	6MWD	Mortality
Cote *et al.*, 2008 [58]	USA, Spain	444 (BODE)	66	Spirometry	GOLD I-IV	>37	10	Publication	6MWD, heart rate	Mortality
Dahl *et al.*, 2001 [63]	Denmark	8955 (CCHS)	58-62	Spirometry	-	60	15	Publication	Fibrinogen	Hospitalisation
Dahl *et al.*, 2007 [64]	Denmark	1302 (CCHS)	68	Spirometry	-	96	12	Publication	CRP	Mortality, hospitalisation
Dahl *et al.*, 2011 [65]	Denmark	7974 (CCHS, CGPS)	49, 60	Spirometry, ICD-8(491-92), ICD-10(J41,J44)	-	33	15	Publication	CRP	Hospitalisation
Dajczman *et al.*, 2015 [66]	Canada	237	69	Spirometry	-	12	6	Publication	6MWD	Mortality
de Torres *et al.*, 2008 [72]	USA, Spain	218 (BODE)	65	Spirometry	GOLD I-IV	36	8	Publication	6MWD, CRP	Mortality
Deng *et al.*, 2014 [74]	China	116	71	Hx, Sx, Spirometry	-	32	7	Publication	CRP	Mortality
Donaldson *et al.*, 2005 [78]	UK	148	69 (median)	Spirometry	-	>86	8	Publication	Fibrinogen, IL-6, IL-8	Exacerbation

Study	Country	Sample size (cohort name)	Average age or range	COPD definition	Disease severity	Follow-up (months)	QUADAS-2 score (out of 15)	Source	Markers(s)	Outcome(s)
Dreyse *et al.*, 2015 [80]	Chile	100	69	Spirometry	GOLD stage I-IV	24	8	Questionnaire	6MWD, heart rate, fibrinogen, CRP, IL-6	Mortality, exacerbation, hospitalisation
Durheim *et al.*, 2015 [82]	USA	326 (INSPIRE-II)	67	Spirometry	GOLD A-D	60	9	Publication	6MWD	Mortality, hospitalisation
Faganello *et al.*, 2010 [89]	Brasil	120	65	Spirometry	GOLD I-IV	12	9	Questionnaire	6MWD, CRP, WCC, IL-6, IL-8, TNF-α	Exacerbation, hospitalisation
Ferrari *et al.*, 2011 [93]	Brasil	95	67	Spirometry	GOLD I-IV	36	7	Questionnaire	6MWD	Exacerbation
Ferrari *et al.*, 2013 [92]	Brasil	53	64	Spirometry	GOLD I-IV	36	6	Questionnaire	6MWD, CRP, WCC, IL-6, TNF-α	Mortality, exacerbation, hospitalisation
Gaki *et al.*, 2011 [99]	Greece	117	65	Spirometry	-	24	8	Publication	6MWD, CRP, fibrinogen	Hospitalisation
Groenewegen *et al.*, 2008 [111]	Netherlands	277 (COSMIC)	63	Spirometry	-	12	10	Publication	CRP, fibrinogen, TNF-α	Hospitalisation
Grolimund *et al.*, 2015 [112]	Switzerland	469 (ProHOSP)	74	Sx, Spirometry	GOLD I-IV	>60	10	Questionnaire	CRP, WCC	Mortality
Guo *et al.*, 2014 [115]	China	69	40-75	Spirometry	GOLD II-III	12	6	Publication	IL-6, IL-8, TNF-α	Exacerbation
Hopkinson *et al.*, 2007 [128]	UK	64	62	Spirometry	GOLD I-IV	12	9	Dataset	QMVC	Exacerbation
Hurst *et al.*, 2010 [132]	Spain	2138 (ECLIPSE)	63	Spirometry	GOLD II-IV	36	13	Questionnaire	6MWD, heart rate, fibrinogen, CRP, WCC, IL-6, IL-8, TNF-α	Exacerbation
Husebo *et al.*, 2014 [133]	Norway	403 (Bergen COPD)	44-76 years	Spirometry	GOLD II-IV	36	11	Questionnaire	6MWD, heart rate, CRP, WCC, IL-6, TNF-α	Mortality, exacerbation, hospitalisation
Jennings *et al.*, 2009 [137]	USA	194	67	Spirometry	-	12	8	Questionnaire	6MWD	Exacerbation, hospitalisation
Jensen *et al.*, 2013 [139]	Denmark	2645 (CCHS)	59	Spirometry	GOLD I-IV	>423	15	Questionnaire	Heart rate, CRP, fibrinogen	Mortality, hospitalisation
Lacasse *et al.*, 2005 [152]	Canada	147	65	Sx, Hx, Spirometry	-	>21	8	Publication	Heart rate	Mortality
Liu *et al.*, 2011 [159]	China	114	70	Spirometry	GOLD I-IV	>24	10	Publication	6MWD, fibrinogen, CRP	Mortality
Man *et al.*, 2008 [167]	Canada	4787 (LHS)	54	Spirometry	-	84	15	Publication	CRP	Mortality
Mannino *et al.*, 2012 [169]	USA	8507 (NHANES III)	40-80+	Spirometry	GOLD I-IV	216	15	Publication	Fibrinogen	Mortality
Marino *et al.*, 2014 [170]	Brasil	63	71 (median)	Spirometry	GOLD II-III	6	4	Questionnaire	6MWD	Exacerbation
Miller *et al.*, 2013 [180]	Spain	2164 (ECLIPSE)	63	Spirometry	GOLD II-IV	36	13	Publication	6MWD, IL-6, IL-8	Mortality, exacerbation
Monninkhof *et al.*, 2003 [186]	Netherlands	248 (COPE)	65	Spirometry	Almost exclusively GOLD II	12	7	Questionnaire	6MWD	Exacerbation, hospitalisation
Moore *et al.*, 2010 [188]	UK	110 (COPE)	63	Spirometry	GOLD I-IV	6	9	Publication	SNIP	Mortality
Moy *et al.*, 2013 [190]	USA	169	71	Hx, Spirometry	GOLD I-IV	16	7	Publication	6MWD	Exacerbation, hospitalisation
Moy *et al.*, 2014 [189]	USA	167	71	Hx, Spirometry	GOLD I-IV	16	7	Publication	6MWD, CRP, IL-6	Exacerbation, hospitalisation

Study	Country	Sample size (cohort name)	Average age or range	COPD definition	Disease severity	Follow-up (months)	QUADAS-2 score (out of 15)	Source	Markers(s)	Outcome(s)
Mullerova *et al.*, 2015 [192]	Spain	2138 (ECLIPSE)	63	Spirometry	GOLD II-IV	36	13	Questionnaire	6MWD heart rate, CRP, fibrinogen, WCC, IL-6, TNF-α	Hospitalisation
Ozgur *et al.*, 2012 [210]	Turkey	73	59	Spirometry	GOLD II-IV	48	9	Publication	6MWD	Hospitalisation
Parker *et al.*, 2014 [212]	USA	1339 (NHANES III)	40-80+	Spirometry	GOLD I-IV	115	15	Publication	CRP, fibrinogen, WCC	Mortality
Pinto-Plata *et al.*, 2012 [219]	USA	253	65	Hx, Spirometry	GOLD I-IV	36	10	Publication	6MWD, IL-6, IL-8, TNF-α	Mortality
Powrie *et al.*, 2007 [223]	UK	142	66	Spirometry	-	12	7	Publication	CRP, IL-6, IL-8	Exacerbation
Seemungal *et al.*, 2001 [235]	UK	137 (East London COPD)	68	Spirometry	-	18	6	Publication	Fibrinogen, IL-6	Exacerbations
Singh *et al.*, 2010 [242]	Spain	488 (ECLIPSE)	64	Spirometry	GOLD II-IV	12	11	Publication	CRP, IL-6, IL-8	Exacerbation
Spruit *et al.*, 2012 [248]	Denmark, UK, USA, Canada	2110 (ECLIPSE)	63	Spirometry	GOLD II-IV	36	13	Publication	6MWD	Mortality, hospitalisations
Swallow *et al.*, 2007 [259]	UK	162	64	Spirometry	GOLD I-IV	48	11	Dataset	QMVC	Mortality
Thomsen *et al.*, 2013 [260]	Denmark	8020 (CHCHS, CGPS)	67	Spirometry	GOLD I-IV	60	14	Publication	CRP, fibrinogen, WCC	Exacerbations
Valvi *et al.*, 2012 [265]	USA	20192 (ARIC, CHS)	45-64	Spirometry	GOLD I-IV	>116	15	Publication	Fibrinogen	Mortality
Wang *et al.*, 2014 [278]	China	136	72	Spirometry	GOLD I-III	12	6	Publication	6MWD	Exacerbation
Wang *et al.*, 2014 [277]	China	331 (ChiCTR-TRC)	63	Spirometry	GOLD I-IV	12	10	Publication	6MWD, IL-6, IL-8, TNF-α	Hospitalisation
Warnier *et al.*, 2014 [280]	Netherlands	405	73	Spirometry, ICPC (R91, R95)	GOLD I-IV	84	11	Publication	Heart rate	Mortality
Waschki *et al.*, 2011 [281]	Germany	170	65	Spirometry	GOLD I-IV	48	11	Questionnaire	6MWD, heart rate, fibrinogen, CRP, WCC, IL-6	Mortality
Wedzicha *et al.*, 2000 [282]	UK	93	67	Spirometry	-	12	8	Publication	Fibrinogen, IL-6	Exacerbation

Abbreviations: COPD, Chronic Obstructive Pulmonary Disease; Sx, signs and symptoms; Hx, medical history; ICD-9, International Disease Classification of Diseases ninth revision; 6MWD, Six-Minute Walk Distance; WCC, White Cell Count; CRP, C-Reactive Protein; IL, Interleukins; TNF-α, Tumour Necrosis Factor-alpha; QMVC, Quadriceps Maximum Voluntary Contraction; SNIP, Sniff Nasal Inspiratory Pressure. Data are displayed as mean and standard deviations. Disease severity is either given as GOLD stage (I-IV) or group (A-D), as described by GOLD. [107]

2.3.1 Included biomarkers

The most frequently reported biomarkers in the studies were: 6MW distance (56%), CRP (39%), fibrinogen (28%), IL-6 (25%), IL-8 (16%), WCC (16%), TNF-α (11%), and resting heart rate (8%), with few assessing CIMT, PWV, and AIx. With the exception of the 6MW distance, very few musculoskeletal biomarkers (i.e. QMVC, SNIP, and SPPB) were reported for their association with clinical outcomes within COPD. The majority of studies (n = 34) included mortality as an outcome measure, followed by exacerbation (n = 25) and hospitalisation (n = 15). Of these, eleven studies investigated two outcomes, and only one investigated all three outcomes.

2.3.2 Data synthesis

All 61 studies were included in the qualitative review, with 32 studies included in the quantitative data synthesis (**Figures 2.2 2.6 2.10**, pages 34-42; **Figures 2.3 2.7 2.11**, pages 35-43) and the sensitivity analyses (**Figures 2.4 2.5 2.8 2.9 2.12 2.13**, pages 36-45). Twenty (69%) studies reported data on mortality, nine (28%) reported data on COPD exacerbations, and eight (25%) reported data on COPD hospitalisation. Data from Faganello *et al.*[89] except for IL-8 were excluded as the same cohort but with a longer follow-up period was examined by Ferrari *et al.*[92] Data from Spruit *et al.*[248] were also excluded as the ECLIPSE cohort was examined in a more recent publication by Mullerova *et al.*[192] Additionally, data from Agusti *et al.* (ECLIPSE cohort)[7] and Durheim *et al.* (INSPIRE-II cohort)[82] were not included as more data were made available through Celli *et al.*[42] and Blumenthal *et al.*,[23] respectively. Results of the 6MWD reported by de Torres *et al.*[72] (BODE cohort, n = 218) were not included because these were covered by Cote *et al.* using a larger study sample (n = 365) and longer follow-up time.[57]

2.3.3 Association between cardiovascular and musculoskeletal measures, and clinical outcomes

2.3.3.1 Six-minute walk distance

Multiple studies, including ECLIPSE (n = 2138), BODE (n = 1379) and INSPIRE (investigating new standards for prophylaxis in reducing exacerbations)-II (n = 326), reported that COPD patients with a shorter 6MW distance at baseline have a higher number of clinical events over a follow-up period of at least six months. A 6MW distance of less than 350 metres was associated with higher risk of early mortality, according to Cote and colleagues,[55] while only Dajczman *et al.* found a significant difference in mortality with a cut-off point of 6MW distance ≤150 metres.[66] The 6MW-based model, authored by Cote *et al.* had a C-statistic of 0.75, similar to Waschki *et al.* (C-statistic = 0.77)[281] and higher than Casanova *et al.* (C-statistic = 0.70),[36] and Spruit *et al.* (C-statistic = 0.67) for a 6MW distance threshold of 334 metres.[248] The remaining studies, with relatively small sample sizes, indicated no statistical significant difference in 6MW distance between those with and without exacerbation.[56,80,93] Meta-analysis indicated that longer walking distances at baseline were associated with early mortality (HR 0.80 per 50 metres increase, 95% CI 0.73 to 0.89, p <0.01, I^2 = 99.4%, COPD exacerbation (SMD -0.27, 95% CI -0.41 to -0.13, p <0.01, I^2 = 53.0% and hospitalisation (SMD -0.48, 95% CI -0.66 to -0.30, p <0.01, I^2 = 61.3%). Galbraith plots showed several studies outside the 95% confidence intervals, indicating Ozgur *et al.*,[210] Marino *et al.*,[169] Monninkhof *et al.*,[186] and Dreyse *et al.*[80] to be the least consistent with the overall results, potentially causing bias (**Figure 2.14**, page 46). Removal of these studies did not alter findings. After removing studies with a quality score in the bottom tertile (≤8), SMDs for exacerbation (SMD -0.27 to -0.15) and hospitalisation (SMD -0.48 to -0.35) had a substantial change, resulting from the removal of studies with small sample sizes and short follow-up times. Meta-regression indicated no differences in HRs for studies with longer follow-up time or those more recently published, but suggests higher HRs for studies with older participants (p = 0.027; **Figure 2.15**, page 47).

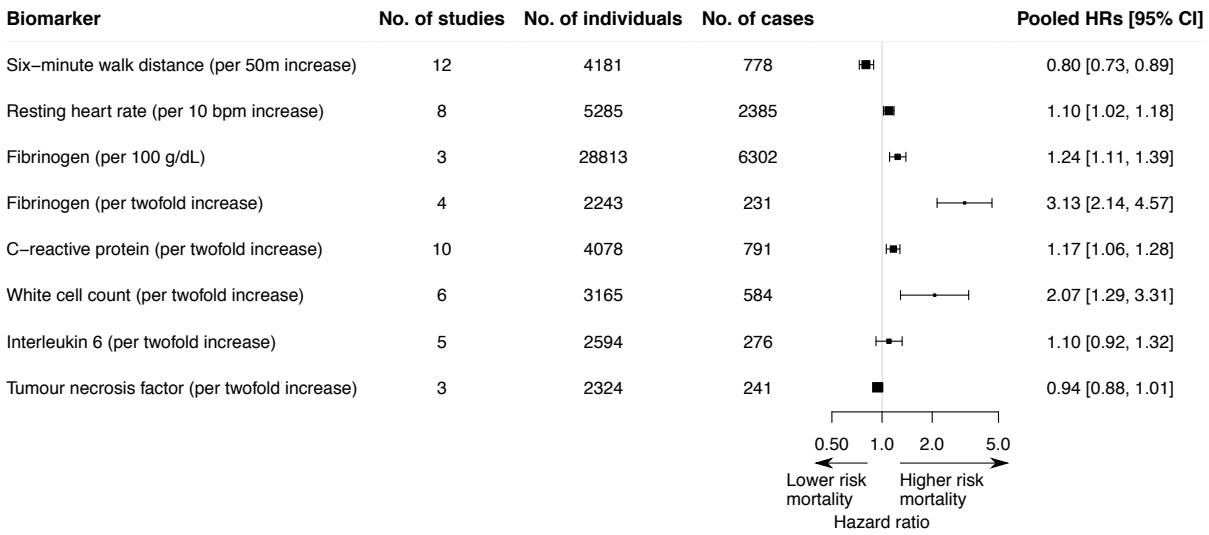


Figure 2.2: Pooled hazard ratios for the risk of mortality with 95% confidence intervals, by biomarker. Studies included: Ferrari *et al.*, 2013,[92] Celli *et al.*, 2012,[42] Blumenthal *et al.*, 2016,[23] de Torres *et al.*, 2008,[72] Cote *et al.*, 2007,[57] Dajczman *et al.*, 2015,[66] Waschki *et al.*, 2011,[281] Dreyse *et al.*, 2015,[80] Ozgur *et al.*, 2012,[210] Mannino *et al.*, 2012,[169] Jensen *et al.*, 2013,[139] Valvi *et al.*, 2012,[265] Liu *et al.*, 2011,[159] Grolimund *et al.*, 2015,[112] Budweiser *et al.*, 2007,[31] Husebo *et al.*, 2014,[133] Antonelli-Incalzi *et al.*, 2006,[13] Cano *et al.*, 2004,[33] Lacasse *et al.*, 2005,[152] and Warnier *et al.*, 2014.[280] See **Figure 2.3**, page 35 for full study details. Bars, 95% confidence intervals.

2.3.3.2 Resting heart rate

Jensen *et al.* estimated that having a resting heart below 65 beats per minute (bpm) compared to above 85 bpm (C-statistic = 0.59), was associated with increased survival of approximately ten years in Global initiative for chronic Obstructive Lung Disease (GOLD)[107] stage I, ∼7 years for GOLD stage II, and ∼6 years in GOLD stages III-IV.[139] Meta-analysis indicated that higher resting heart rates at baseline were associated with early mortality (HR 1.10 per 10 bpm, 95% CI 1.02 to 1.18, p = 0.01, I^2 = 99.4%), exacerbation (SMD 0.09 bpm, 95% CI 0.00 to 0.17, p = 0.05, I^2 = 0.0%), and hospitalisation (SMD bpm 0.21, 95% CI 0.15 to 0.28, p <0.01, I^2 = 10.0%). After removing studies with a quality score in the bottom tertile, HRs for mortality increased (1.10 to 1.15), and SMD (0.09 to 0.08) lost significance for exacerbation.

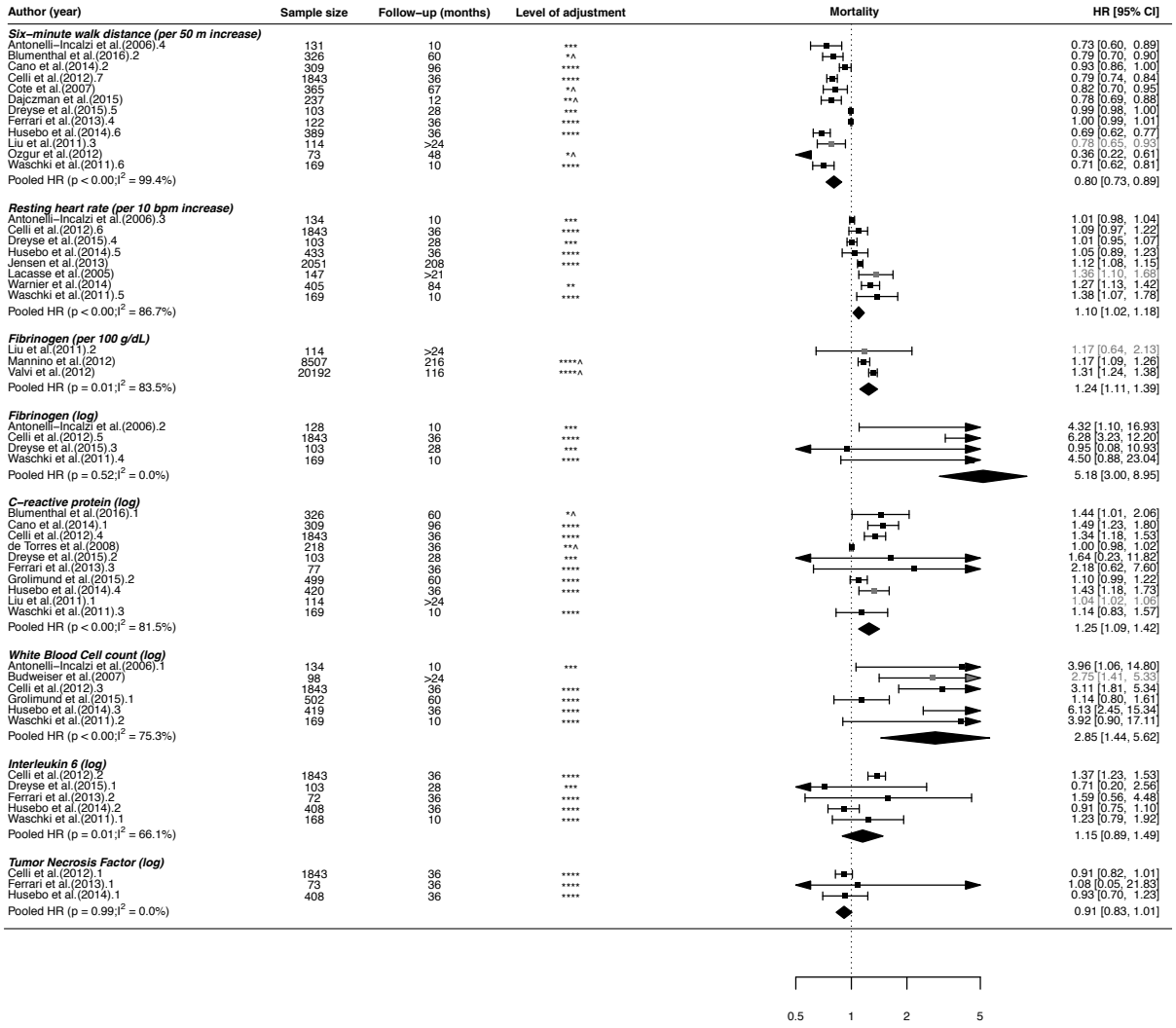


Figure 2.3: Adjusted hazard ratios for the risk of mortality with 95% confidence intervals, by biomarker. Values in grey are unadjusted and not included in the pooling of results. ****Adjusted for age, sex, body mass index, and smoking status, unless otherwise indicated. ∧Adjusted for additional confounders; Blumenthal *et al.* adjusted for age, Charlson index, duration of COPD, GOLD, and coping skills training; Cote *et al.* adjusted for body mass index and Charlson score; Dajzman *et al.* adjusted for age, sex, and FEV_1; Ozgur *et al.* adjusted for body mass index, IC/TLC, FEV_1, dyspnoea index, PaO2, and PaCO2; Mannino *et al.* adjusted for age, sex, body mass index, smoking status, ethnicity, diabetes, cardiovascular disease, education level, and poverty income ratio; Valvi *et al.* adjusted for age, sex, body mass index, smoking status, ethnicity, education level, diabetes mellitus, cardiovascular disease, and GOLD stage; de Torres *et al.* adjusted for age, sex, pack-year history, cardiovascular risk factors or disease, and treatment with inhaled corticosteroids. Bars, 95% confidence intervals.

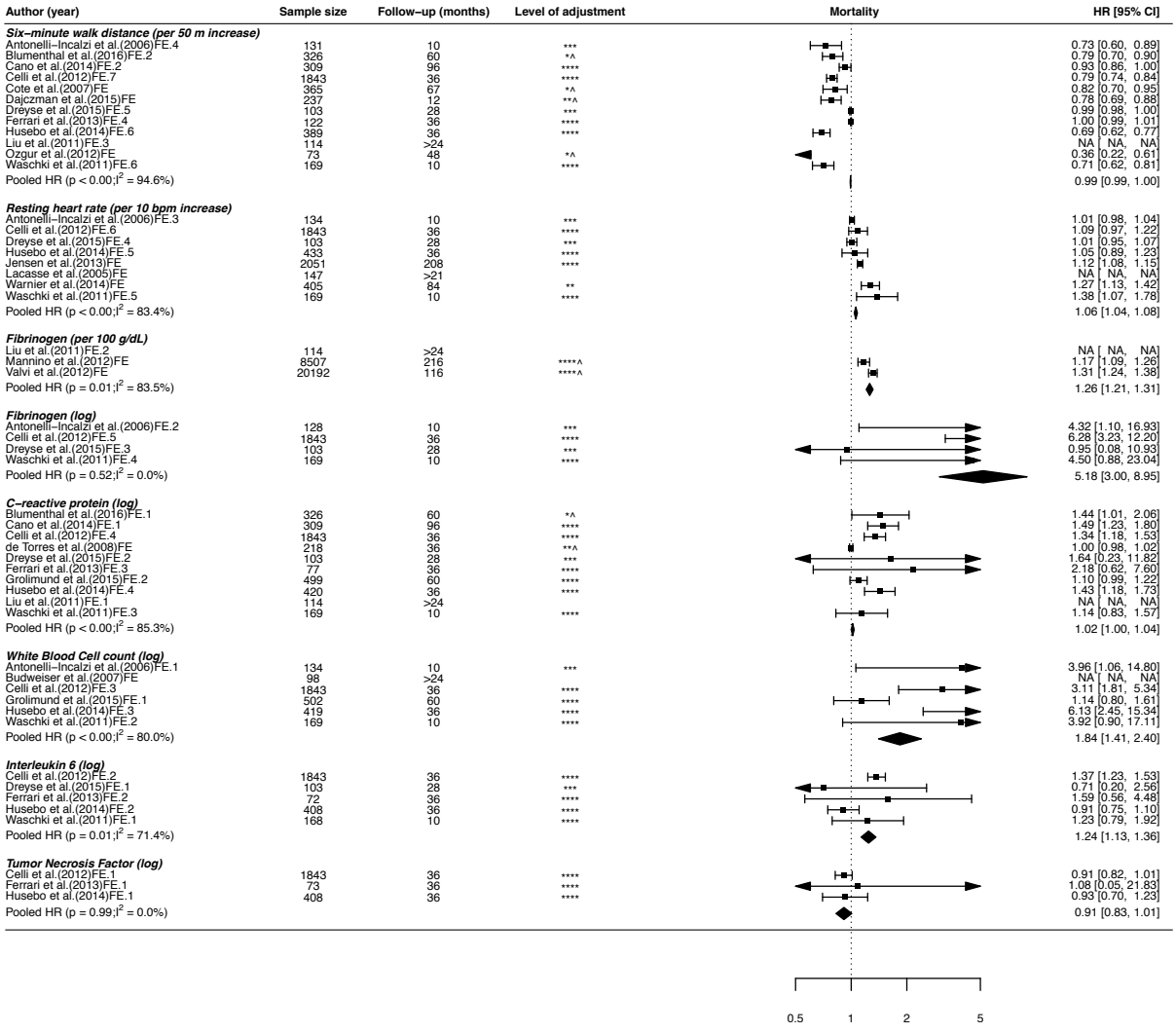


Figure 2.4: Hazard ratios for mortality using fixed-effect modelling, sensitivity analysis.

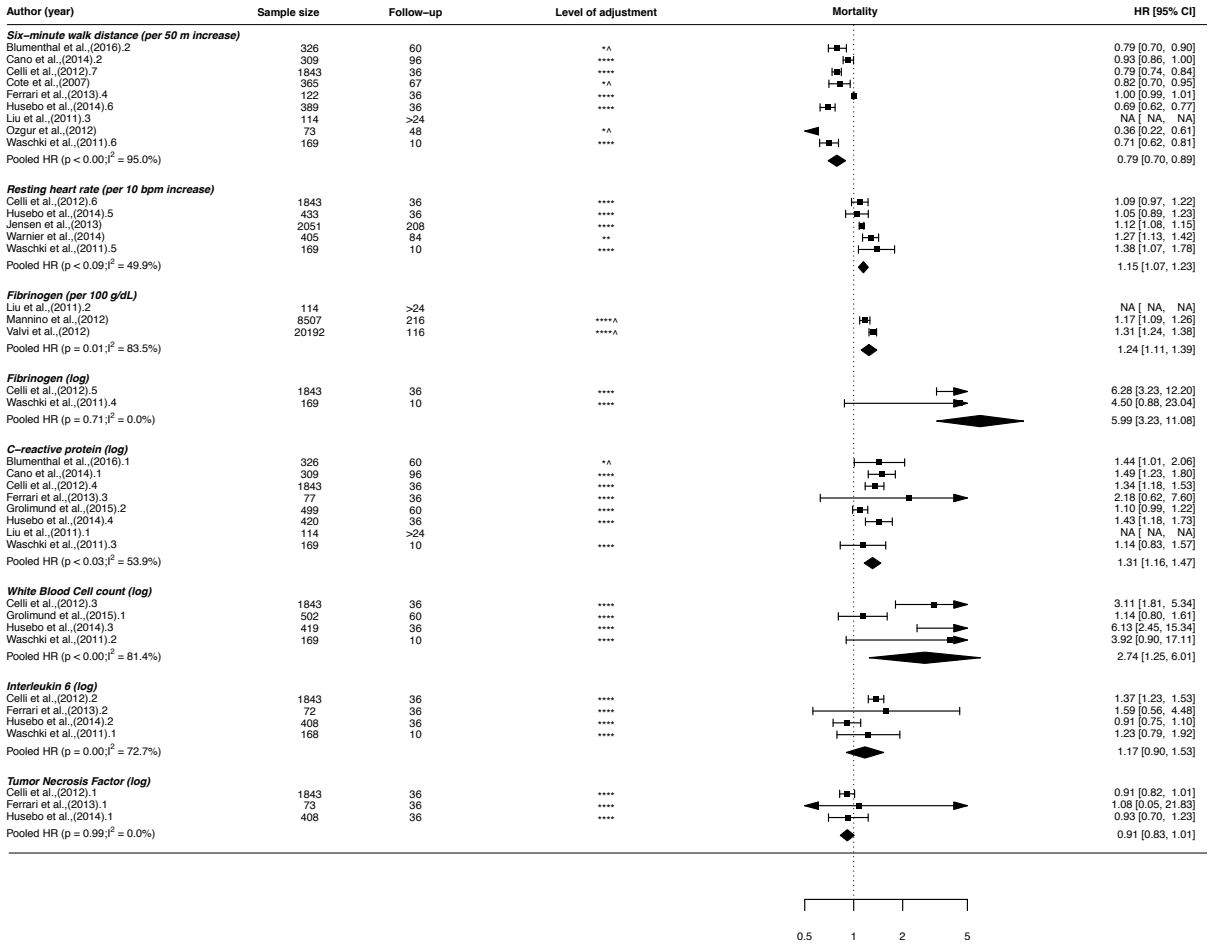


Figure 2.5: Hazard ratios for mortality, sensitivity analysis. Lower quartile QUADAS-2 scores removed.

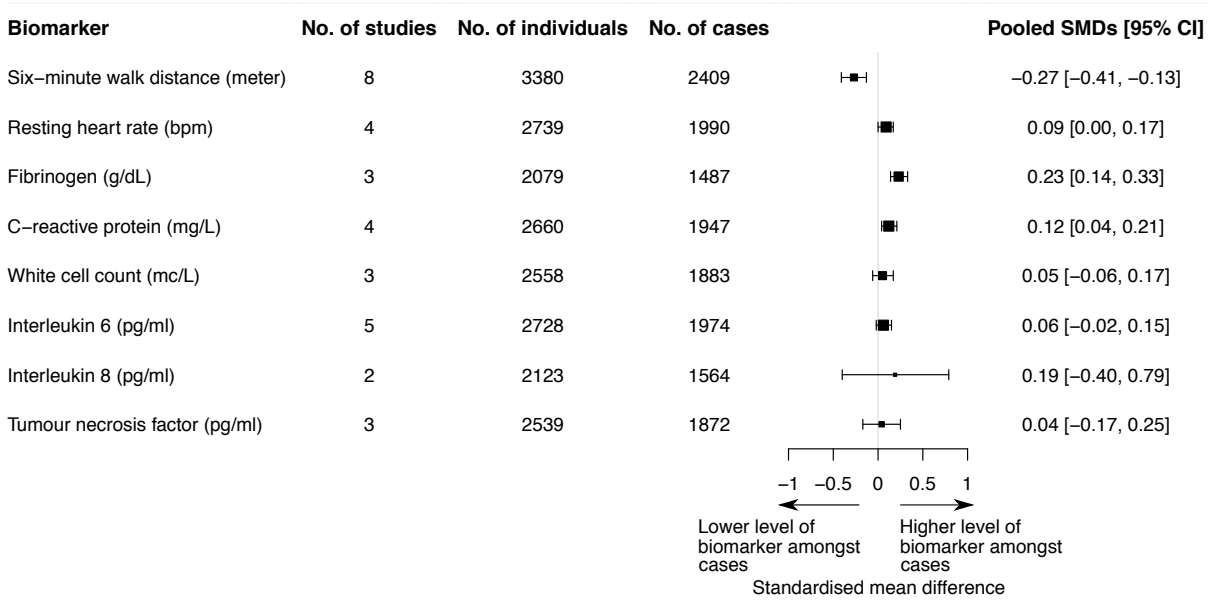


Figure 2.6: Pooled standardised mean differences with 95% confidence intervals for the risk of exacerbation, by biomarker. Studies included: Faganello *et al.*, 2010,[89] Cote *et al.*, 2007,[57] Dreyse *et al.*, 2015,[80] Ferrari *et al.*, 2011,[93] Monninkhof *et al.*, 2003,[186] Hurst *et al.*, 2010,[132] Husebo *et al.*, 2014,[133] Wedzicha *et al.*, 2000,[282] Jennings *et al.*, 2009,[137] and Marino *et al.*, 2014.[170] See **Figure 2.7**, page 39 for full study details. Bars, 95% confidence intervals.

2.3.3.3 Fibrinogen

Within the Copenhagen City Heart Study and Copenhagen General Population Study (n = 8020), Thomsen *et al.* reported a higher risk of exacerbation with elevated fibrinogen levels, however, only in combination with elevated levels of CRP and WCC at baseline (C-statistic = 0.73).[260] Celli *et al.* reported a similar C-statistic of 0.70 when including fibrinogen together with WCC, CRP, and other inflammatory markers to their predictive model.[42] Meta-analysis indicated that for mortality, there was a positive association with fibrinogen (HR 3.13 per twofold increase, 95% CI 2.14 to 4.57, p <0.01, I^2 = 0.0%, and HR 1.24 per 100 g/dL, 95% CI 1.11 to 1.39, p <0.01, I^2 = 83.5%).[159,169,265] Higher levels of fibrinogen were also associated with exacerbation (SMD 0.23 g/dL, 95% CI 0.14 to 0.33, p <0.01, I^2 = 0.0%).

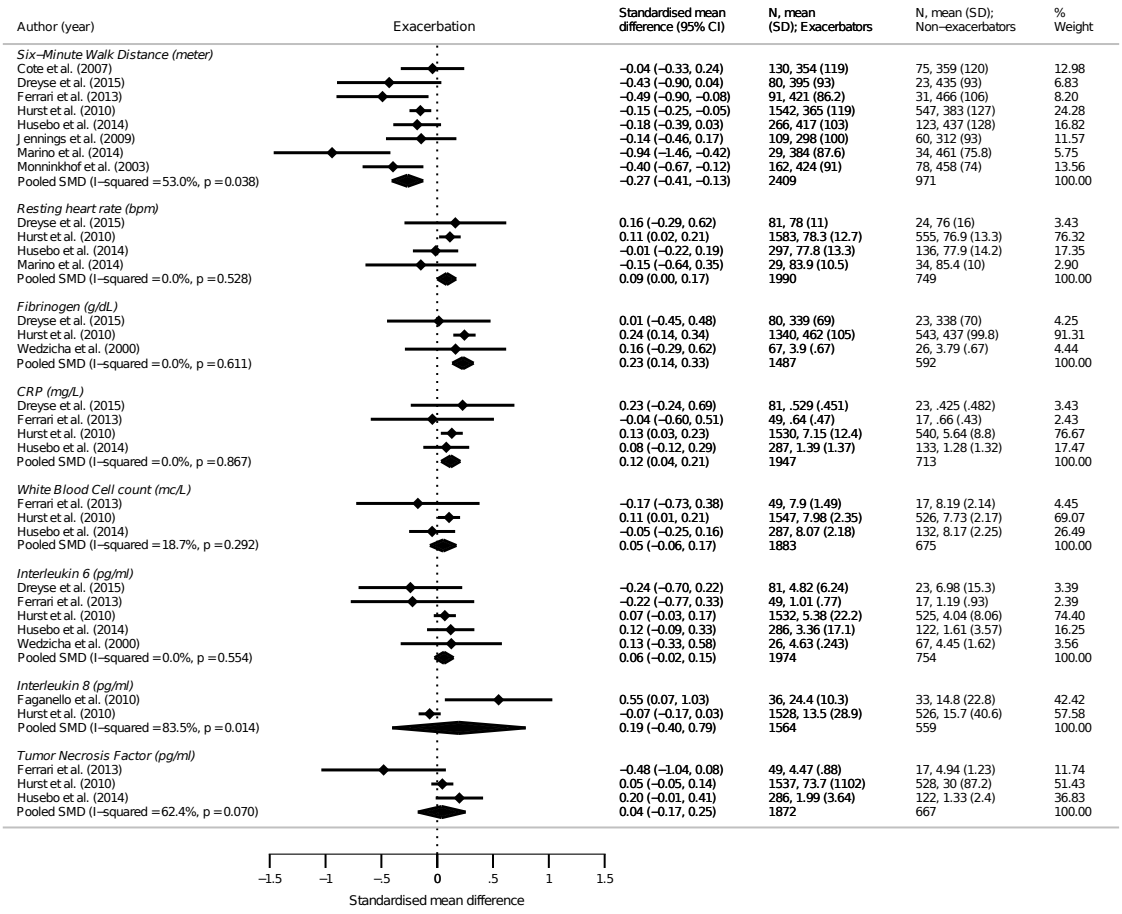


Figure 2.7: Standardised mean differences with 95% confidence intervals for the risk of exacerbation, by biomarker. Bars, 95% confidence intervals.

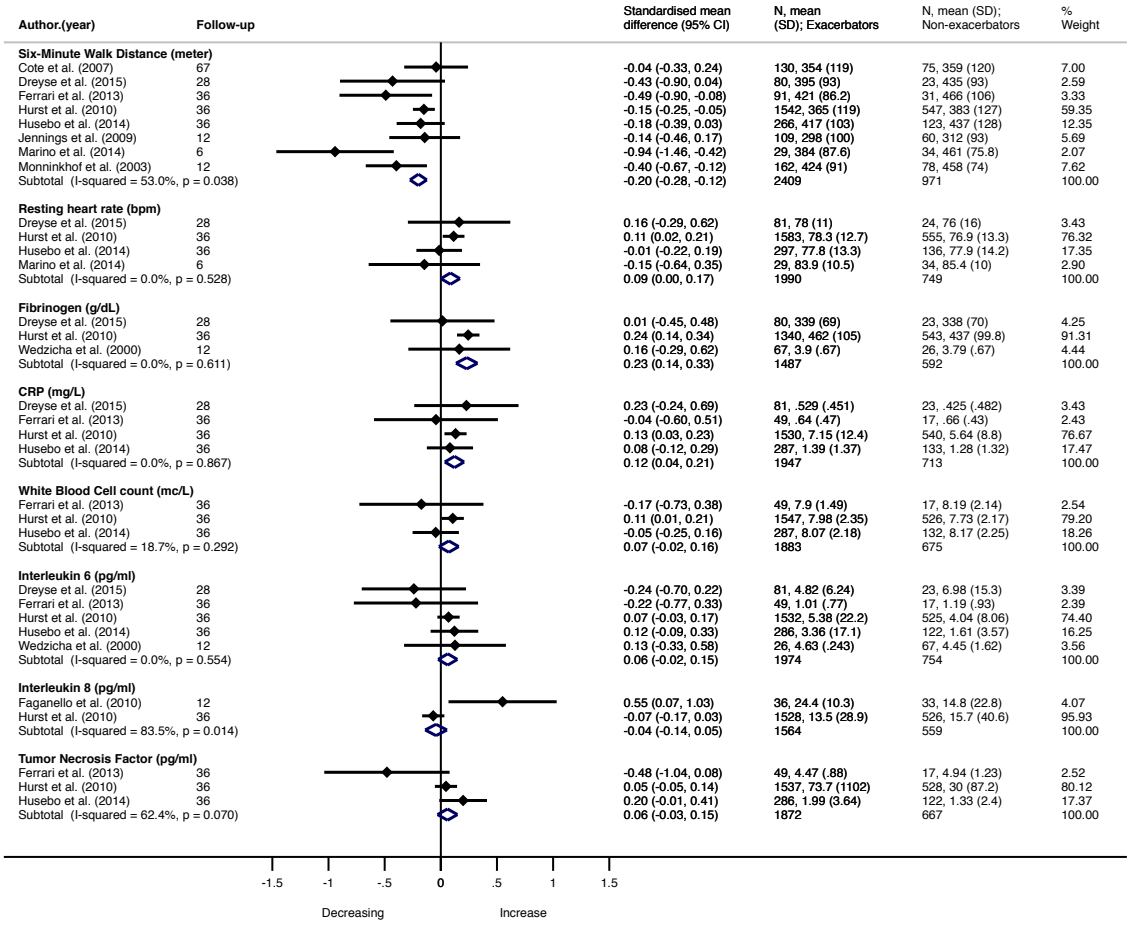


Figure 2.8: Standardised mean differences for exacerbation using fixed-effect modelling, sensitivity analysis.

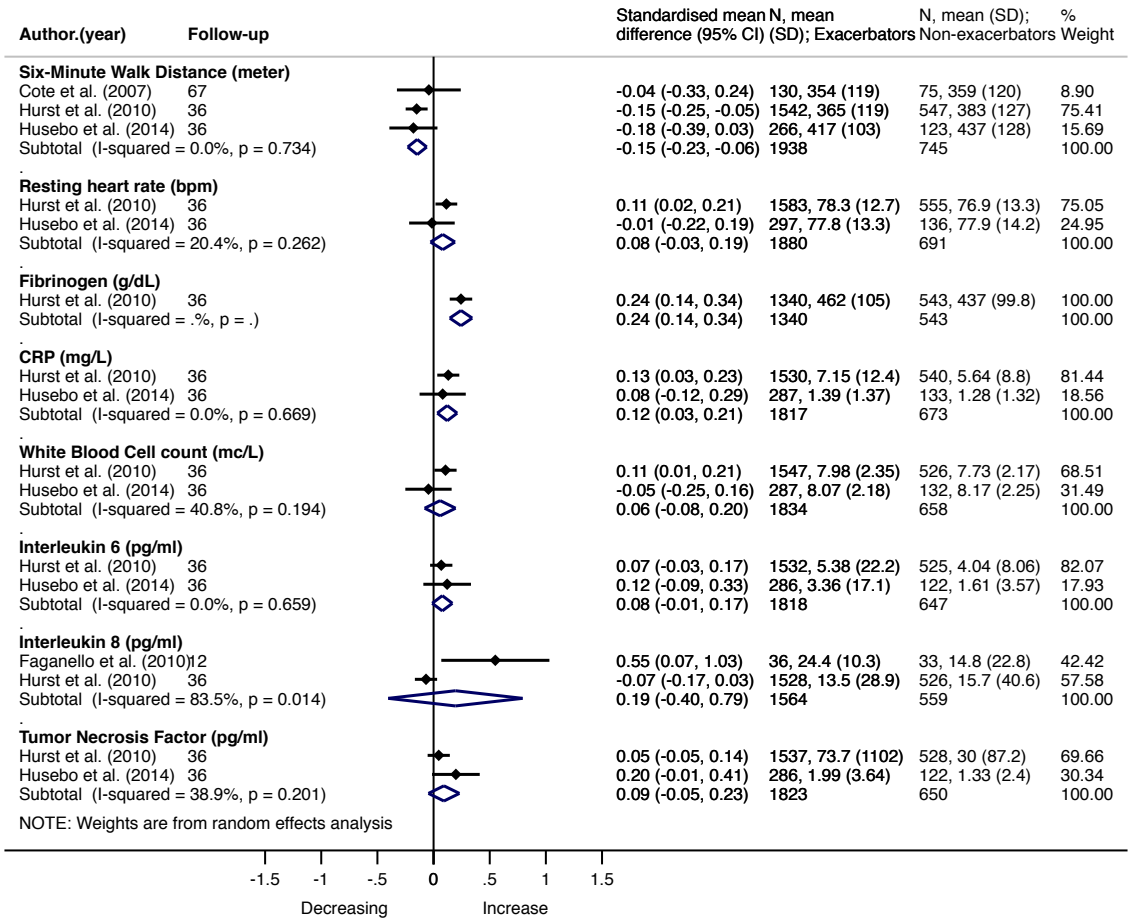


Figure 2.9: Standardised mean differences for exacerbation, sensitivity analysis. Lower quartile QUADAS-2 scores removed.

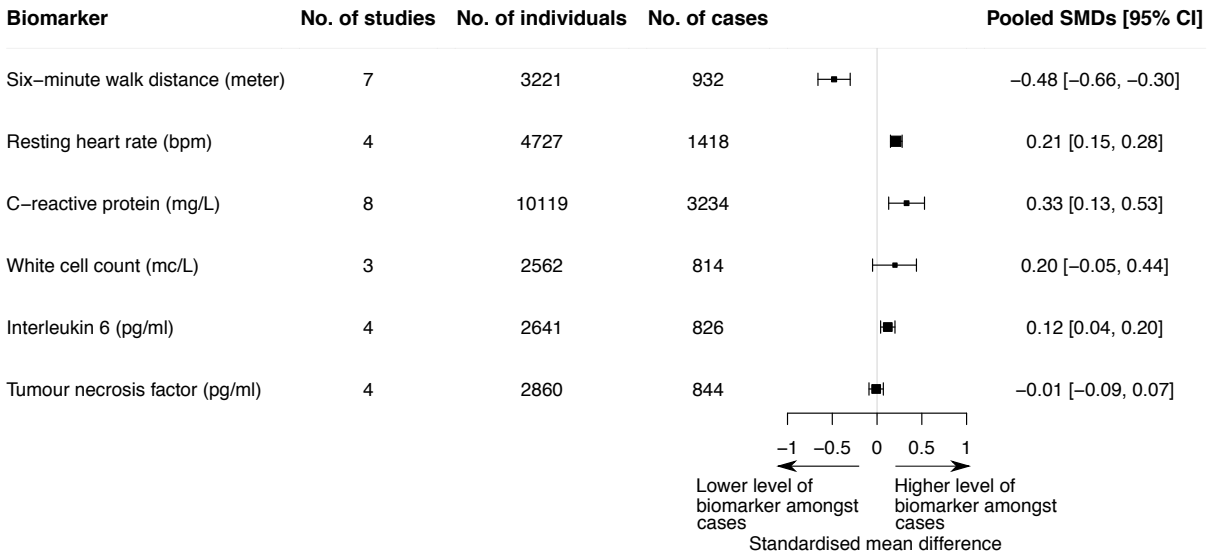


Figure 2.10: Pooled standardised mean differences with 95% confidence intervals for the risk of hospitalisation, by biomarker. Studies included: Ferrari et al., 2013,[92] Mullerova et al., 2015,[192] Dreyse et al., 2015,[80] Monninkhof et al., 2003,[186] Jensen et al., 2013,[139] Groenewegen et al., 2008,[111] Husebo et al., 2014,[133] Cano et al., 2004,[33] Jennings et al., 2009,[137] and Dahl et al., 2011.[65] See **Figure 2.11**, page 43 for full study details. Bars, 95% confidence intervals.

2.3.3.4 C-reactive protein

Moy *et al.* suggested that combining CRP with step count is a good predictor of acute exacerbations (C-statistic = 0.59) and hospital admission (C-statistic = 0.69).[189] However, de Torres *et al.* (BODE cohort, n = 218), reported no statistically significant associations between baseline CRP levels and mortality,[72] along with Grolimund *et al.* (ProHOSP, n = 469),[112] Ferrari *et al.*,[92] and Waschki *et al.*[281] There was also no difference in CRP levels at baseline for COPD exacerbation in the COSMIC study.[111] Meta-analysis indicated that individuals with higher levels of CRP measured at baseline had a higher risk of early mortality (HR 1.17 per twofold increase, 95% CI 1.06 to 1.28, p <0.01, I^2 = 81.5%). Higher levels of CRP were also associated with COPD exacerbations (SMD 0.12 mg/L, 95% CI 0.04 to 0.21, p <0.01, I^2 = 0.0%), and hospitalisation (SMD 0.33 mg/L, 95% CI 0.13 to 0.53, p <0.01, I^2 = 92.8%). After removing studies with a quality score in the bottom tertile, HRs for mortality increased (1.25 to 1.31), and decreased for hospitalisation (0.20 to 0.13). Meta-regression indicated no statistical signifi-

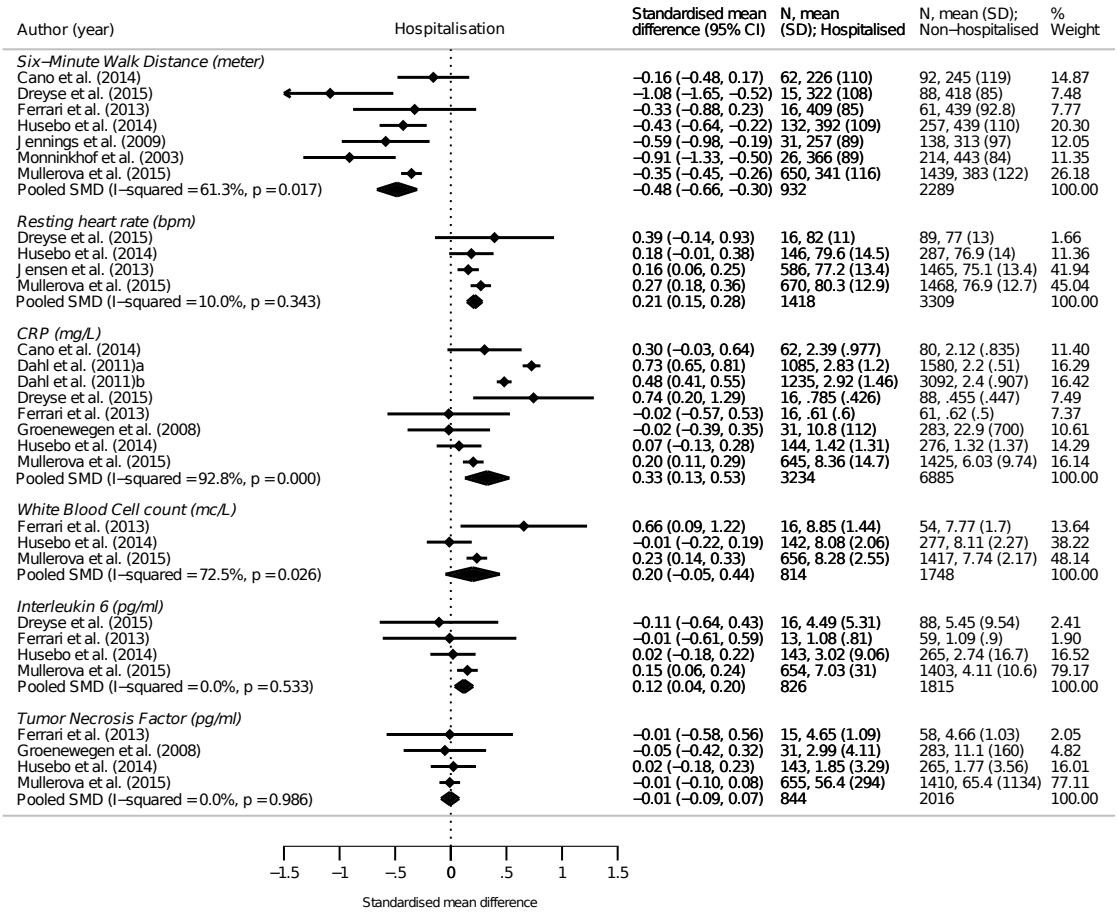


Figure 2.11: Standardised mean differences with 95% confidence intervals for the risk of hospitalisation, by biomarker. Bars, 95% confidence intervals.

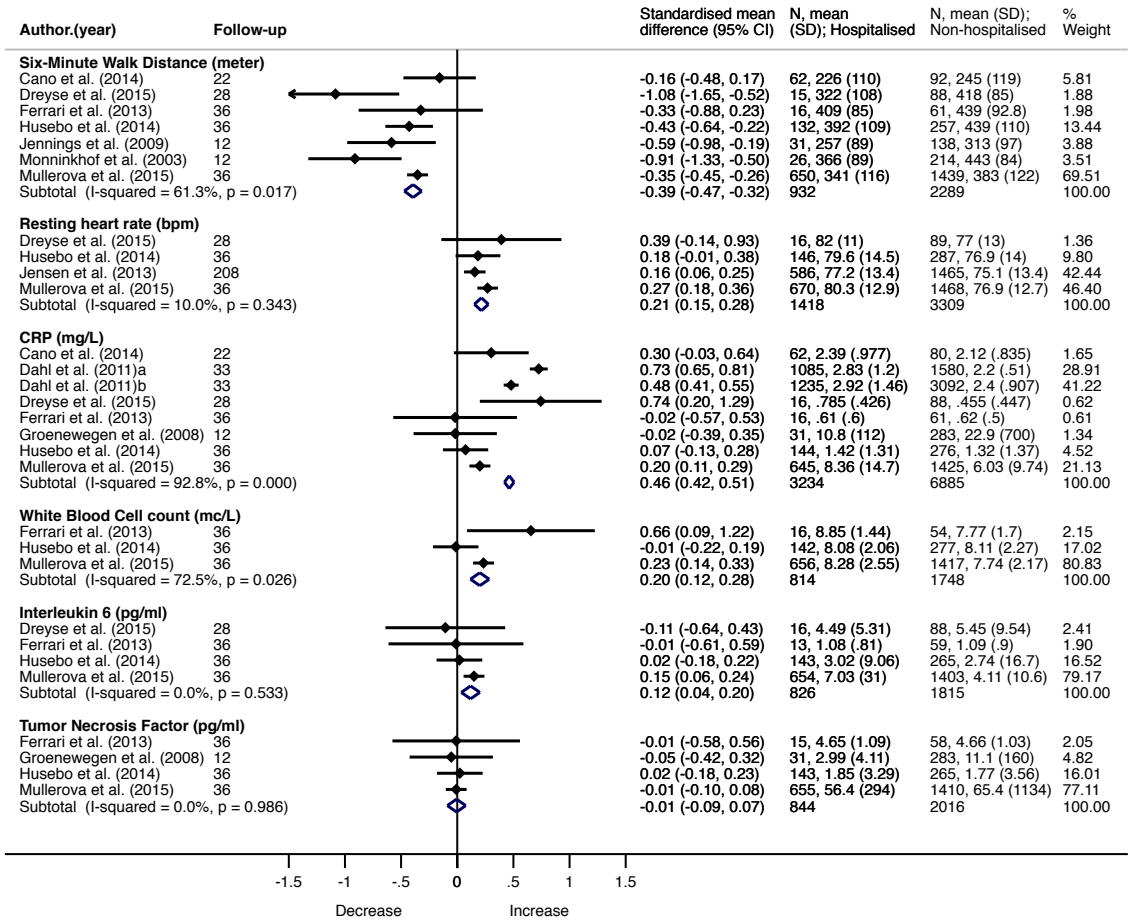


Figure 2.12: Standardised mean differences for hospitalisation using fixed-effect modelling, sensitivity analysis.

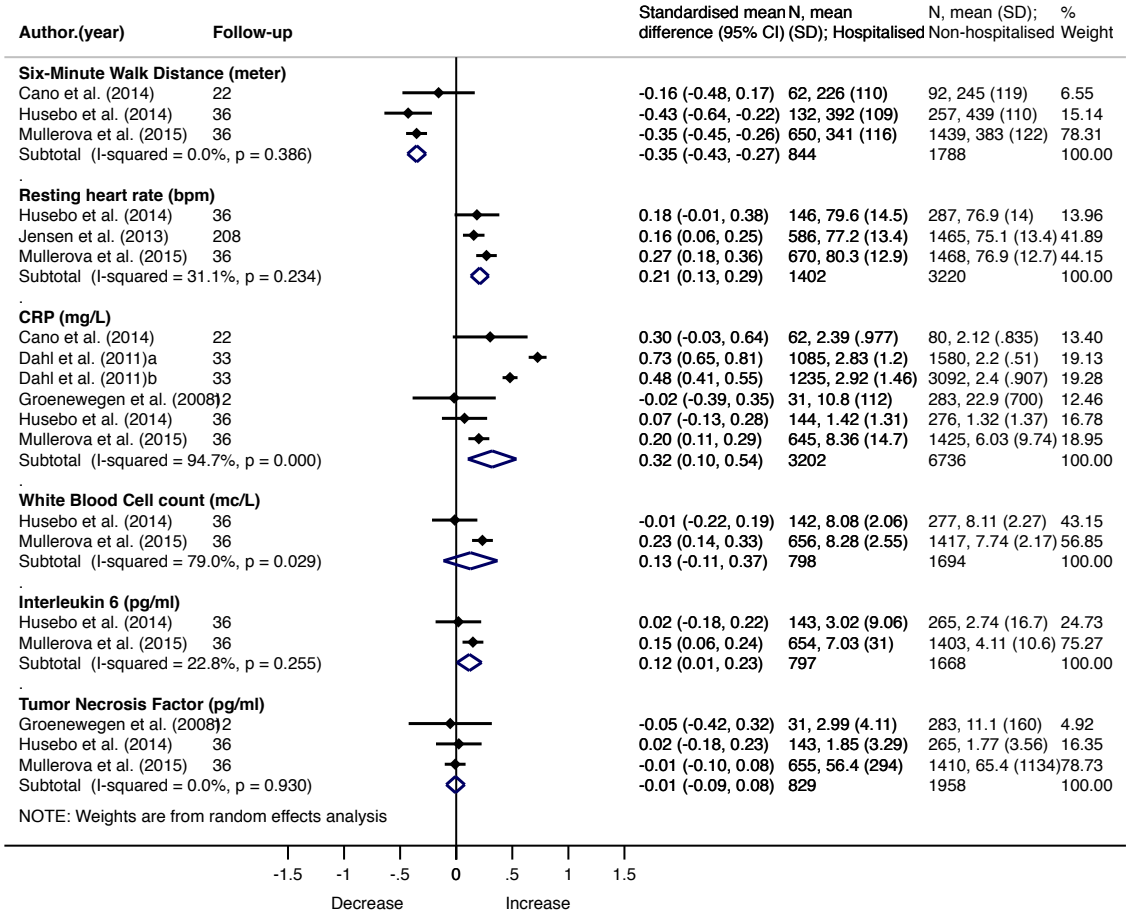


Figure 2.13: Standardised mean differences for hospitalisation, sensitivity analysis. Lower quartile QUADAS-2 scores removed.

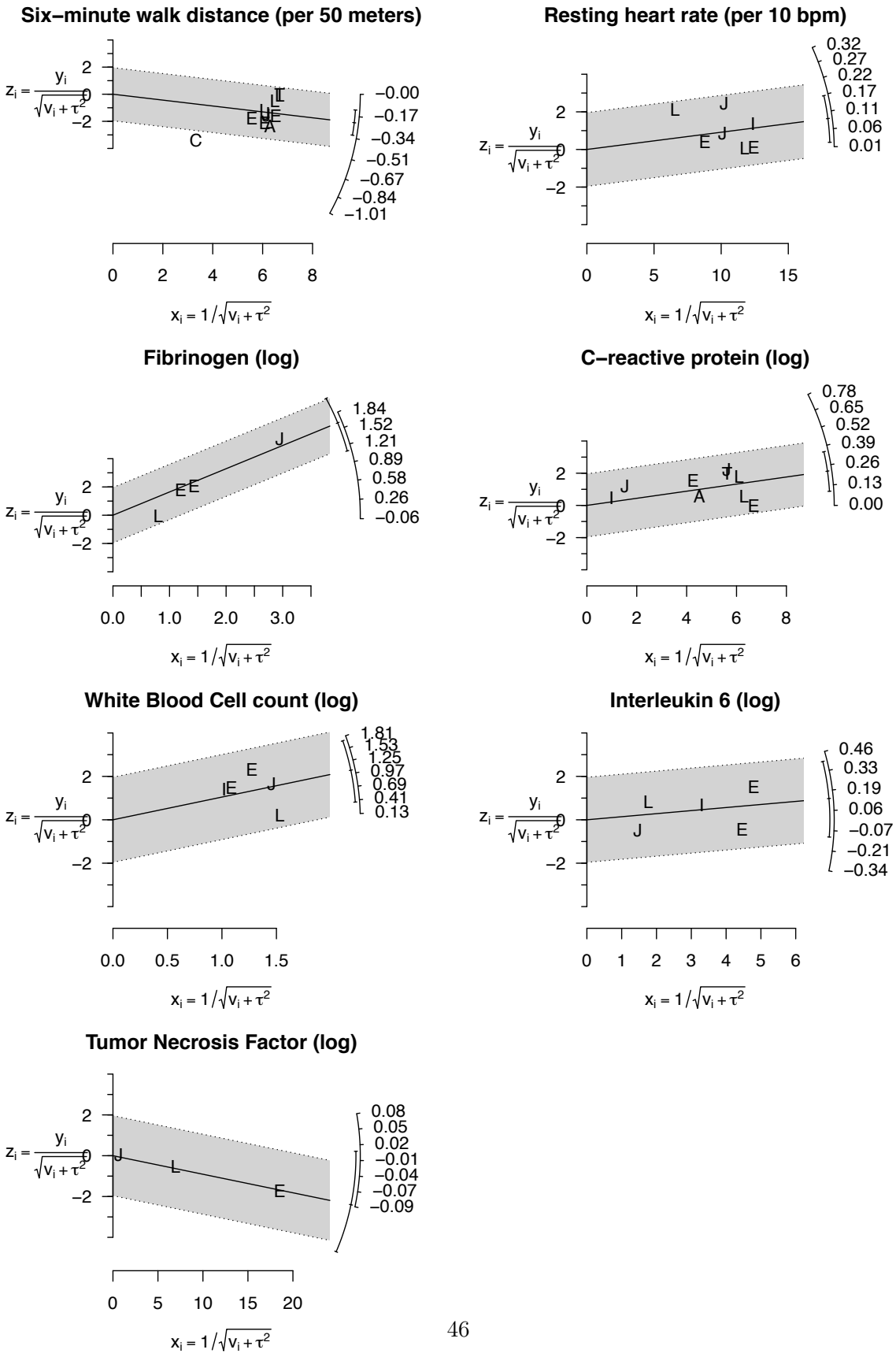


Figure 2.14: Galbraith plots for mortality.

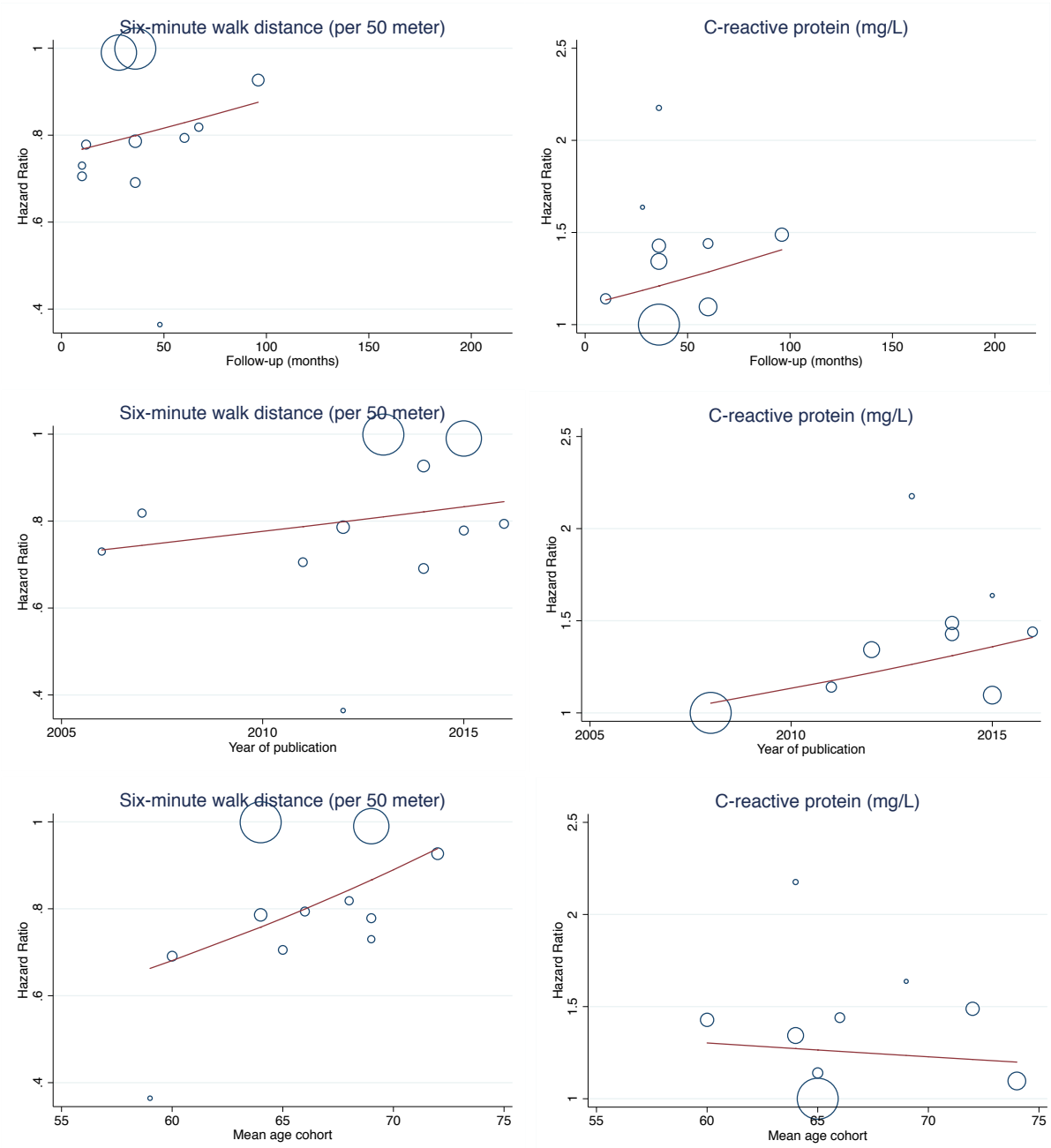


Figure 2.15: Meta-regression six-minute walk distance and C-reactive protein, hazard ratio mortality. Bubble plot with fitted meta-regression line.

cant difference for studies with longer follow-up time, with older participants and more recently published (**Figure 2.15**, page 47).

2.3.3.5 White cell count

Only a few studies compared baseline measures with clinical outcomes over time (≥6 months). Several studies reported COPD patients with higher WCC levels at baseline at higher risk of clinical outcomes.[31,42,132,192,260] However, Husebo *et al.* did not find higher baseline measures to be associated with a higher number of exacerbations during three years of follow-up.[133] Additionally, Grolimund *et al.* (ProHOSP, n = 469) did not find a statistically significant difference between WCC levels and mortality.[112] Meta-analysis indicated an association between higher levels of WCC at baseline and a higher risk of earlier death (HR 2.07 per twofold increase, 95% CI 1.29 to 3.31, p <0.01, I^2 = 75.3%). However, WCC levels were not associated with exacerbation (SMD 0.05, 95% CI -0.06 to 0.17, p = 0.38, I^2 = 18.7%) or hospitalisation (SMD 0.20, 95% CI -0.05 to 0.44, p = 0.12, I^2 = 72.5%). After removing studies with a quality score in the bottom tertile, HRs for mortality increased for fibrinogen (5.18 to 5.99; **Figure 2.5**, page 37).

2.3.3.6 Interleukin 6

Hurst *et al.* (ECLIPSE, n = 2138) did not find higher baseline measures to be associated with a higher number of exacerbations.[132] Additionally, Waschki *et al.*[281] and Wedzicha *et al.*[282] did not find higher IL-6 baseline levels to be associated with a higher risk of mortality. Meta-analysis indicated no association between IL-6 and earlier mortality (HR 1.10 per twofold increase, 95% CI 0.92 to 1.32, p = 0.28, I^2 = 66.1%). Neither was there an association with exacerbation (SMD 0.06, 95% CI -0.02 to 0.15, p = 0.16, I^2 = 0.0%). Increased levels were, however, associated with hospitalisation (SMD 0.12, 95% CI 0.04 to 0.20, p = 0.01, I^2 = 0.0%).

2.3.3.7 Interleukin 8

Interleukin 8 levels and its relation with clinical outcomes in COPD is not well reported. Within the ECLIPSE study (n = 2138), Hurst *et al.*, found that IL-8 levels at baseline was not a statistically significant predictor for exacerbations after one year of follow-up.[132] However, Celli *et al.*, who also used data from ECLIPSE (n = 1843), did find increased levels at baseline to be associated with a higher risk of mortality after three years of follow-up.[42] Meta-analysis indicated no association between IL-8 and exacerbation (SMD 0.19, 95% CI -0.40 to 0.79, p = 0.52, I^2 = 83.5%).

2.3.3.8 Tumour necrosis factor-alpha

Celli *et al.* (ECLIPSE, n = 1843) did not find a statistically significant difference between those who died after three years of follow-up and those still alive.[42] Hurst *et al.* (ECLIPSE, n = 2138) reported similar findings for exacerbations after one year of follow-up.[132] Additionally, Groenewegen *et al.* reported no statistically significant difference between the baseline TNF-α measure and clinical outcomes in the COSMIC cohort (n = 277), after one year of follow-up.[111] Meta-analysis indicated no associations between elevated levels of TNF-α and the risk of earlier death (HR 0.94 per twofold increase, 95% CI 0.88 to 1.01, p = 0.07, I^2 = 0.0%), nor for exacerbation (SMD 0.04, 95% CI -0.17 to 0.25, p = 0.71, I^2 = 62.4%), or hospitalisation (SMD -0.01, 95% CI -0.09 to 0.07, p = 0.88, I^2 = 0.0%).

2.3.3.9 Quadriceps maximum voluntary contraction

In recent years there has been an increasing interest in examining the predictive value of functional activities of the musculoskeletal system. The quadriceps muscle is of particular interest, being assessed using QMVC as a surrogate marker. However, only two studies assessing the same cohort of patients have assessed QMVC in relation to clinical outcomes, where quadriceps muscle function of 184 COPD patients using QMVC was found to be a good predictor of mortality after four years of follow-up (HR 0.88, 95% CI 0.77 to 1.00) with higher levels reducing

risk.[188,259]

2.3.3.10 Sniff nasal inspiratory pressure

Moore *et al.* reported a statistically significant association between baseline SNIP and mortality (HR 0.73 per 10 cmH_2O, 95% CI 0.63 to 0.84, C = 0.68) and suggest that, compared to pulmonary plethysmographs, a test commonly performed to measure functional residual capacity, SNIP is recommended because of its low cost and efficiency.[188]

No studies included in our systematic review reported associations with SPPB, PWV, CIMT, and AIx.

2.3.4 Publication bias

Publication bias was present in most biomarkers for all outcome measures, indicated through asymmetrical funnel plots (**Figure 2.16**, page 51). Larger studies appear in the top of the graphs with outliers near the bottom. Bias seemed to primarily occur due to the poor quality of small studies, which deviated most from the other studies. As indicated, the smaller studies have the tendency to show larger differences between those with the event compared to those without. Removal of studies that fell outside of the funnel plot did not alter findings.

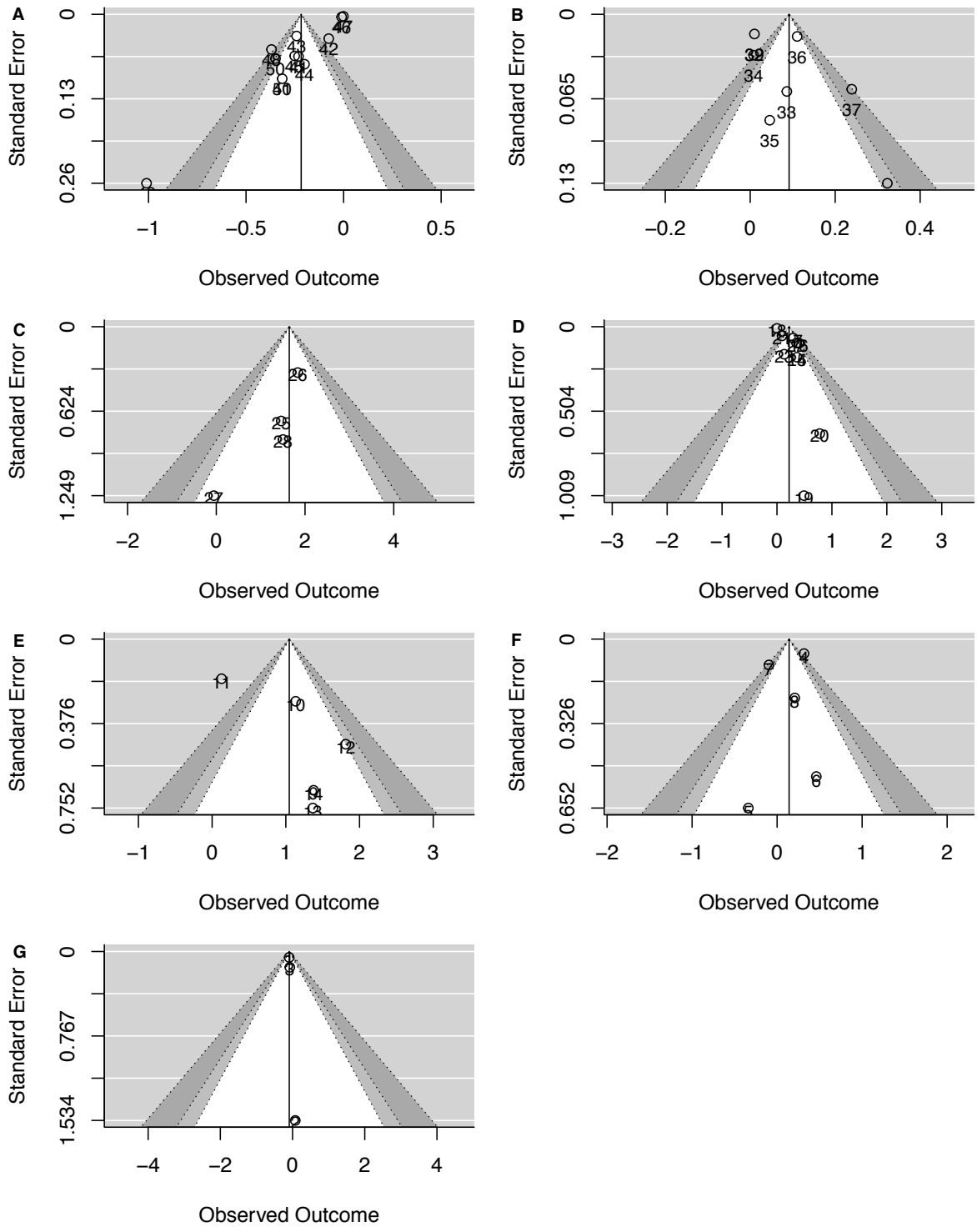


Figure 2.16: Funnel plots for mortality. A = six-minute walk distance. B = resting heart rate. C = fibrinogen. D = C-reactive protein. E = white cell count. F = interleukin-6. G = tumour necrosis factor-alpha

2.4 Discussion

This study systematically summarises and examines the association between multiple outcomes and biomarkers that may potentially better capture systemic problems in COPD patients and are not yet widely utilised in clinical practice. Our main findings indicate that stable COPD patients had higher risks of premature death if they had a shorter walking distance, and higher resting heart rate, fibrinogen, CRP and WCC at baseline, when followed-up over a period of at least six months. Only a shorter walking distance, and higher fibrinogen and CRP levels indicated a higher risk of COPD exacerbation. The risk for COPD-related hospital admission was higher with a shorter walking distance, and higher resting heart rate and CRP and IL-6 levels.

No studies evaluating SPPB, CIMT, PWV, and AIx were included in our systematic review. However, a small number of publications have assessed these in relation to clinical outcomes in COPD (which did not meet our selection criteria). Based on a meta-analysis of seventeen studies, mainly in the general population, a SPPB score <10 (range 0-12) was found to be predictive of all-cause mortality.[216] The gait speed, one of SPPBs components, was also found to predict hospital readmission in elderly COPD patients.[148] The non-invasive CIMT and its role in clinical outcomes in COPD patients has not been largely investigated. However, it has been shown that patients with COPD, in particular smokers, are at higher risk of an elevated CIMT due to atherosclerotic plaque formation and developing arterial stiffness as a result of hypoxaemia.[49,94,135] Other studies found associations with PWV[213] and AIx.[181] We have identified gaps in the literature that need to be examined in order to address these research questions, and while the recommendation for clinical utility differs slightly, the evidence across the studies suggest that the use of musculoskeletal measures to assess outcomes in COPD patients are worth further investigation.

This review has some potential limitations. By focussing on stable (i.e. non-hospitalised) patients, our results may not be generalisable to unstable COPD patients. Additionally, study

heterogeneity exists due to differences between studies in definitions of stable COPD, the duration of stability prior to study enrolment, patient selection criteria, length of follow-up, and outcome definitions. We aimed to address this by using random effects modelling, using SMDs (which are robust to varying lengths of follow-up[114]), and sensitivity analysis. We did not have access to individual patient data, which would allow us to model time-to-event data, adjust for a common set of confounders and estimate the discriminative ability of the biomarkers. Where possible, HRs are presented with adjustment for age, sex, BMI, and smoking status. Ideally studies should be adjusted for disease severity (i.e. FEV_1) and comorbidities like hypertension and diabetes.[180] And even so, continuous variables such as FEV_1 might not have been linear in the included studies and should have been transformed appropriately. Fitting a linear model to nonlinear data may result in biased estimates. Other potential sources of bias include e.g. history of sleep apnoea, number of previous hospital admissions prior study enrolment, and years of COPD. Also, most studies included had short follow-up times of about six months. Selecting a longer follow-up period, for example of a year, as a cut-off point would result in too few studies being included. Simultaneously, while over a too long time period or time, the predictive value of a biomarker would diminish.

Future investigation should focus on evaluating and validating the predictive ability of COPD biomarkers, preferably in large studies with longer follow-up time. Emphasis should be placed on ensuring biomarkers are generalisable (i.e. more diversity in ethnicity and comorbidities) and practical for clinical use. Tests such as the 6MW distance are not well adopted for clinical practice as they require time and space.[23] Future research could focus on the validation of fast and simple tests such as the SPPB or its components. These are easier and faster to conduct, require less space, and patients are less likely to require oxygen. Newly developed risk models could help monitor clinically diagnosed COPD patients in an early stage of disease to identify patients at high risk for mortality, exacerbation, and hospitalisation. Some work is already underway, with the SPIROMICS (Subpopulations and Intermediate Outcome Measures in COPD) study group developing a debility score aiming to identify COPD patients with debility, i.e. extreme breath-

lessness, decreased exercise capacity, and poor health status.[54] Additionally, the Evaluation of the Role of Inflammation in Chronic Airways disease (ERICA) study cohort could help provide answers to these questions aiming to fill the biomarker gap.

2.5 Conclusions

These findings suggest that 6MW distance, resting heart rate, fibrinogen, CRP, WCC, and IL-6 are associated with clinical outcomes in COPD. The review process elicited very few studies that examined the association between musculoskeletal measures (e.g. SPPB and QMVC) and COPD. While the recommendation for clinical utility differs slightly, the evidence across the studies suggest these are worth further investigation.

3

Evaluating the Role of Inflammation in Chronic Airways disease (ERICA) study

Chapter summary This chapter describes the Evaluating the Role of Inflammation in Chronic Airways disease (ERICA) study, a multi-centre non-interventional epidemiological observational study of individuals clinically diagnosed with chronic obstructive pulmonary disease (COPD). Baseline values of key variables are compared by recruitment site and sex using standard non-parametric statistics. Missing data are described. Partial correlations are used to examine the relationships between variables. Hospital episode statistics and causes of death are briefly described. In total, 714 individuals were included in the analysis. Most individuals had GOLD (Global initiative for chronic Obstructive Lung Disease) stage II. There were noticeable differences in most baseline variables such as age, resting heart rate and walking distance between recruitment sites and sexes. Individuals at Cardiff had worse scores for most baseline variables compared to other sites, for example, they scored lower on the musculoskeletal mea-

sures, had worse arterial stiffness, and higher COPD impact scores. For most variables (e.g. age, forced expiratory volume in one second, and walking distance) there were significant differences between sexes but not lung function, body mass index, and inflammatory markers. Most primary hospital admissions related to pulmonary- and cardiac disease. A majority of deaths were attributed to pulmonary disease. Compared to other COPD cohorts, the ERICA study has a relatively small sample size but is unique in terms of data density, including measures of cardiovascular and musculoskeletal function in addition to lung function, medical- and family history, and biochemical measures.

3.1 Background

The Evaluating the Role of Inflammation in Chronic Airways disease (ERICA) study is a multicentre observational, non-interventional, epidemiological cohort study, with a sample size of 734 individuals diagnosed with chronic obstructive pulmonary disease (COPD), established to identify cardiovascular (CV) and musculoskeletal biomarkers that could be targeted to improve the outcomes of COPD patients. It is a unique dataset that contains numerous biomarkers and demographic data measured longitudinally on patients diagnosed with COPD. The data collection is tied to capture events and changes related to respiratory, musculoskeletal and CV function including changes in therapy allowing to investigate the prevalence and significance of CV and musculoskeletal manifestations of COPD. Five UK centres with an interest in COPD undertook this study: Cambridge (n = 90), Edinburgh (n = 102), Cardiff (n = 374), Nottingham (n = 107) and London (n = 61). The ERICA study is part of a consortium based on a partnership between academia and industry. The consortium includes additional cohort studies ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points)[270] and ARCADE (Assessment of Risk in Chronic Airways Disease Evaluation).[101]

Key variables captured in the ERICA study relate to musculoskeletal and CV function. Musculoskeletal measures of particular interest are the short physical performance battery (SPPB)

and its components four-metre gait speed (4MGS), balance and chair stand. Others include quadriceps maximum voluntary contraction (QMVC), six-minute walk (6MW) distance, and sniff nasal inspiratory pressure (SNIP). Key CV function measures include arterial pulse wave velocity (PWV), carotid-intima media thickness (CIMT), and augmentation index (AIx). These measures are not yet widely used in clinical practice but may potentially better capture systemic problems in COPD than conventional measures, and the occurrence of clinical outcomes (i.e. COPD exacerbations, hospital admission, mortality) within a COPD population.

The aim of this chapter was to describe the ERICA cohort, baseline values of key variables captured and differences between recruitment sites and sex, and to examine any missing data and relationships between variables of interest.

3.2 Methods

3.2.1 Study details of the ERICA study

The patient population includes adults aged >40 years with a clinical diagnosis of COPD, post-bronchodilator spirometry forced expiratory volume in one second (FEV_1/ forced vital capacity (FVC) ratio <0.7 and FEV_1 ≤80% of predicted normal, current or ex-smoker with a smoking history of at least ten pack years, and who were clinically stable for more than four weeks from any exacerbation requiring treatment with oral steroids or antibiotics or hospitalisation were eligible.[184] Patients with the inability to provide written informed consent, a known diagnosis of α1-anti-trypsin deficiency, known neurological or skeletal muscle disease, pregnancy, and ongoing participation in a trial of an experimental drug were excluded from the ERICA study. Individuals were prospectively recruited either from existing databases, through clinicians at outpatient COPD clinics, or by advertisements at one of the study sites. In Cardiff, individuals were selected from the ARCADE study – a CV screening cohort – who consented for the ERICA study. At the other sites patients were recruited opportunistically from COPD clinics.

Clinical measures, blood samples, medical history, and questionnaire data were collected

starting December 2011 and individuals were followed-up every six months for up to 30 months via postal or telephone questionnaire, or until death. Standard operating procedures were developed to standardise measurements between the different study sites. For blood biomarkers, up to 50 ml of blood was drawn and analysed at the local National Health Services biochemical and haematological laboratory. Full details of the study protocol, including standardisation procedures, have been provided elsewhere,[184] and are available on ericacopd.org (created by J.M. Fermont). The study was registered with the UK Clinical Trials Gateway. See **Appendix C** for the ERICA study protocol, and **Appendix D** for a detailed data dictionary (created by J.M. Fermont).

Fibrinogen and neutrophil levels, amongst other parameters such as high-sensitivity C-reactive protein (CRP) and lipid profiles, were measured for all ERICA study participants via withdrawing up to 50 ml venous blood. Arterial stiffness was measured via aortic PWV, AIx, and central blood pressure using the non-invasive SphygmoCor system.[286] During each cardiac cycle mean arterial pressure indicated average blood pressure. The highest pressure on the cardiac system was indicated by the systolic blood pressure. The CIMT was measured on both sides of the neck through imaging using ultrasound (triggered R-wave) with a linear probe 7-12 MHz and three electrocardiogram electrodes. Patients lied down and rested five minutes before scanning but were seated for examination. Quadriceps maximum voluntary contraction was measured using the technique described by Edwards *et al.* where the best effort of six contractions was recorded.[85] The 6MW distance was recorded as the distance walked by the patient as quickly as possible for six minutes.[14] Body mass index (BMI) was estimated through bodyweight and height, and categorised according to the World Health Organization.[291] Exacerbation history, one year preceding the study, was defined as self-reported antibiotics and/or steroids use in the last twelve months. Disease severity was defined according to Global initiative for chronic Obstructive Lung Disease (GOLD) classification.[107] The SPPB has a range of 0-12 points and its three subtests score 0-4. Total SPPB score is the sum of points of each component. Functional limitation was defined by a SPPB cut-off score of <10.[20,216] In addition, the SPPB was

categorised in groups <3 vs. 3-11 vs. 12.[214] The COPD assessment test (CAT) was categorised in groups <10 (low), 10-20 (medium, 21-30 (high) and >30 (very high).[105] Total St. George's Respiratory Questionnaire for COPD (SGRQ-C) score, consisting of a symptoms, activity, and impact component, was estimated using the item-weighted algorithm provided by Jones.[142] Both questionnaires reflect on the impact of COPD on activities of daily living with higher scores indicating increasing limitations. Self-reported comorbidities at baseline were elicited with, for example, the questions "Have you ever required antibiotics for your chest?", and "If yes, how many courses of steroids have you required in the last 12 months?".

3.2.2 Outcome measures

Clinical data from the ERICA study were linked with mortality data obtained from the UK Office for National Statistics (ONS) and hospital episode statistics (HES) obtained from the National Health Services (NHS) Digital, NHS Scotland and NHS Wales. Applications for data linkage were prepared by J.M. Fermont and submitted March 2016. Data were received November 2017. The following information was used for data linkage purposes only: NHS number, date of birth, postcode, forename and surname. Causes of death were categorised by CV and respiratory physicians according to descriptions provided on death certificates (**Appendix E**).

3.2.3 Missing values in baseline characteristics

Some individuals had missing values for various baseline characteristics (**Table 3.1**, page 65). These were examined to assess the level and type of missing data, the missing data patterns and other basic descriptive statistics. Extreme values were examined individually. Based on thresholds determined by a clinician part of the ERICA consortium, unusual values were removed: glucose (n = 20), glycated haemoglobin (HbA1c; n = 9), glomerular filtration rate (GFR; n = 27), fibrinogen (n = 4), white cell count (WCC; n = 1), haemoglobin (n = 2), neutrophils (n = 1), total cholesterol (n = 1), low-density lipoprotein (LDL; n = 18), and CIMT (n = 1).

3.2.4 Statistical analysis

Demographics were described using number and percentage for categorical variables such as Medical Research Council (MRC) dyspnoea score and GOLD stage,[107] and the median with interquartile range (IQR) for continuous variables such as age and BMI. Categorical data were analysed using chi-square tests, unless e.g. the expected cell frequency condition fails, in which case the Fisher/s exact test was used. Continuous and ordinal data were analysed using the Wilcoxon-Mann-Whitney test and Kruskal-Wallis test. Associations between the clinical measures were quantified using partial Spearman's rank correlations adjusted for age, sex, FEV_1, and recruitment site. Values <0.30 were considered weak, 0.30-0.50 as moderate, and >0.50 as strong.[51] The maximum number of independent variables to be included in the models was determined by the number of events. According to Van Belle's statistical rules of thumb approximately ten events per variable are required to obtain reliable regression coefficient estimates.[266]

All tests were two-sided and of statistical significance at an alpha level of 0.05. Analyses were performed using *STATA* version 13.0 (College Station, Texas) and *R* (R Foundation). Patient demographics and group comparisons are displayed in figures and tabular form.

3.2.5 Patients/ consent and permission to publish

Ethics approval and written informed patient consent was obtained in writing from all study participants and permits the processing and publishing of all data included in this dissertation. Each patient has been allocated a unique study number. Ethics approval was granted by the National Research Ethics Service Committee East of England - Cambridge South and registered under reference 11/EE/0357. The ERICA study was funded by the Technology Strategy Board and the Medical Research Council.

3.3 Findings

3.3.1 Missing data

In total, there was <10% missing values with most for CIMT, followed by PWV, 6MW distance, QMVC and SNIP. The majority of missing values for CIMT were present in London, for PWV in Nottingham, and for 6MW in Cambridge. Most of the missing values related to a single variable, for example, 46 (6%) had missing values for CIMT only. These were likely to be missing due to the difficulty of obtaining high quality images. Eleven individuals (2%) had missing values for both CIMT and PWV. Only four individuals had missing values for CIMT, PWV, 6MW and QMVC (**Figures 3.1 3.2**, pages 61-62).

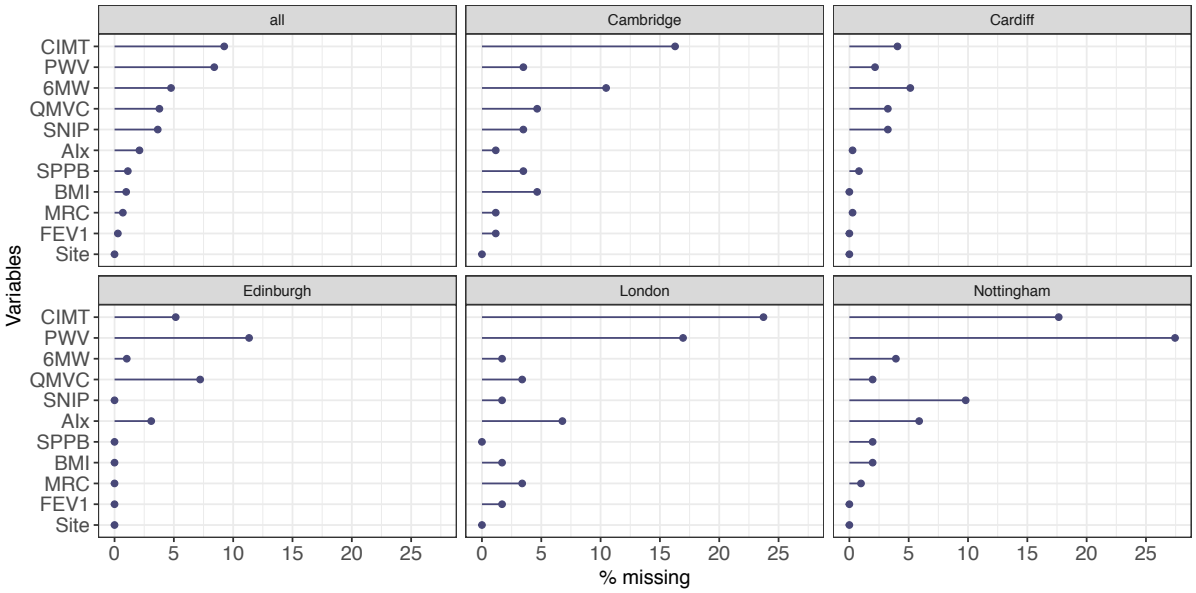


Figure 3.1: Missing values displayed by total and recruitment site. CIMT = carotid intima-media thickness. *Abbreviations*: PWV, pulse wave velocity. 6MW, six-minute walk. QMVC, quadriceps maximum voluntary contraction. SNIP, sniff nasal inspiratory pressure. AIx, augmentation index. SPPB, short physical performance battery. BMI, body mass index. MRC, Medical Research Council dyspnoea. FEV_1, forced expiratory volume in one second.

It is unlikely these values are missing completely at random. For example, plots suggest that missing values in walking distances are not missing completely at random: walking distance seemed to be missing for those with (A) lower QMVC, (B) higher BMI, (C) worse MRC dyspnoea

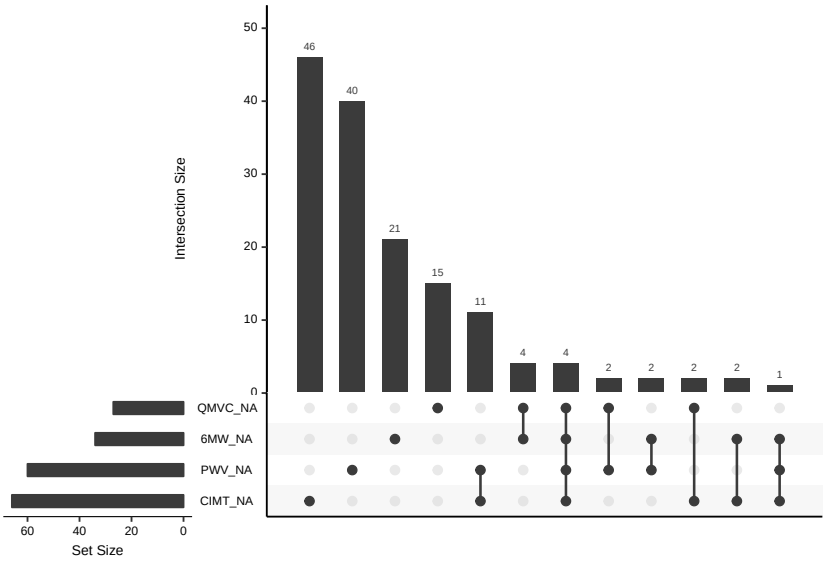


Figure 3.2: Percentage and pattern of missing values in key baseline characteristics. Abbreviations: QMVC, quadriceps maximum voluntary contraction. *Abbreviations*: 6MW, six-minute walk. PWV, pulse wave velocity. CIMT, carotid-intima media thickness.

scores, and (D) worse FEV_1 (**Figure 3.3**, page 63).

3.3.2 Descriptive statistics

Of the 734 individuals entered into the study, 729 met study inclusion criteria of whom 714 were able to be linked with NHS and ONS for hospital admission and survival status (**Figure 3.4**, page 64). Most individuals were recruited in Cardiff (n = 370). In total, 434 (61%) were male with a median baseline age of 67 years (range 43-89 years), median (IQR) BMI of 26.6 kg/m^2 (23.3 - 31.1) with two-thirds above normal bodyweight (**Table 3.1**, page 65, and **Figures 3.5 3.6**, pages 67-67). Cambridge and Nottingham had a significantly higher number of males compared to other sites, with Cardiff the least ($p < 0.001$). Overall, the median age was lowest in London ($p = 0.027$) but with the largest difference in median age by sex ($p = 0.019$). Body mass index was highest in London (23 kg/m^2) and lowest in Cardiff (28 kg/m^2; $p < 0.001$), whereas Nottingham had the highest difference in BMI between sexes ($p = 0.018$).

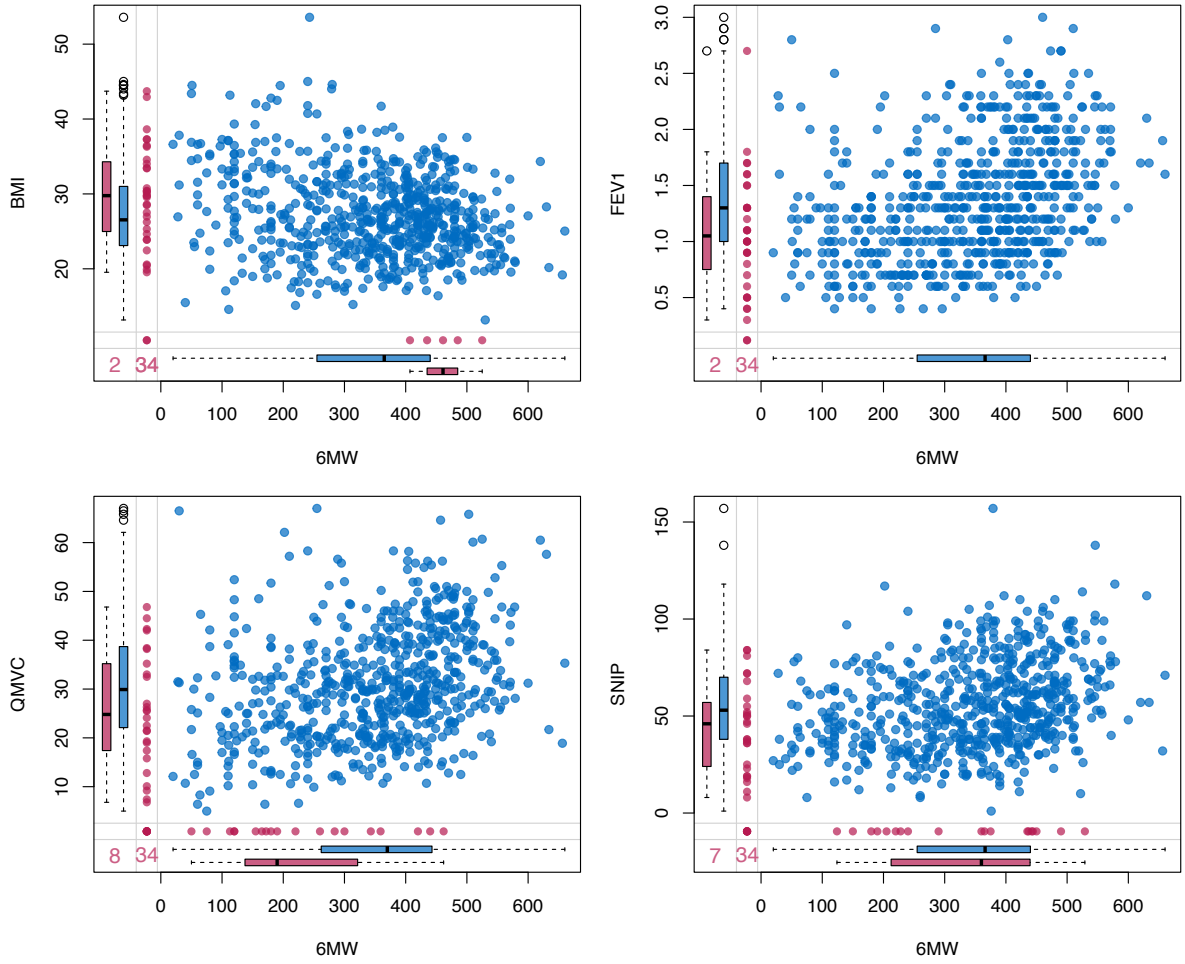


Figure 3.3: Missing data patterns. Margin plots in blue indicate the distribution of observed data given the other variable is observed. Red box plots indicate data distribution of the observed data given the other variable is missing. *Abbreviations*: BMI, body mass index. 6MW, six-minute walk. FEV_1, forced expiratory volume in one second. QMVC, quadriceps maximum voluntary contraction. SNIP, sniff nasal inspiratory pressure.

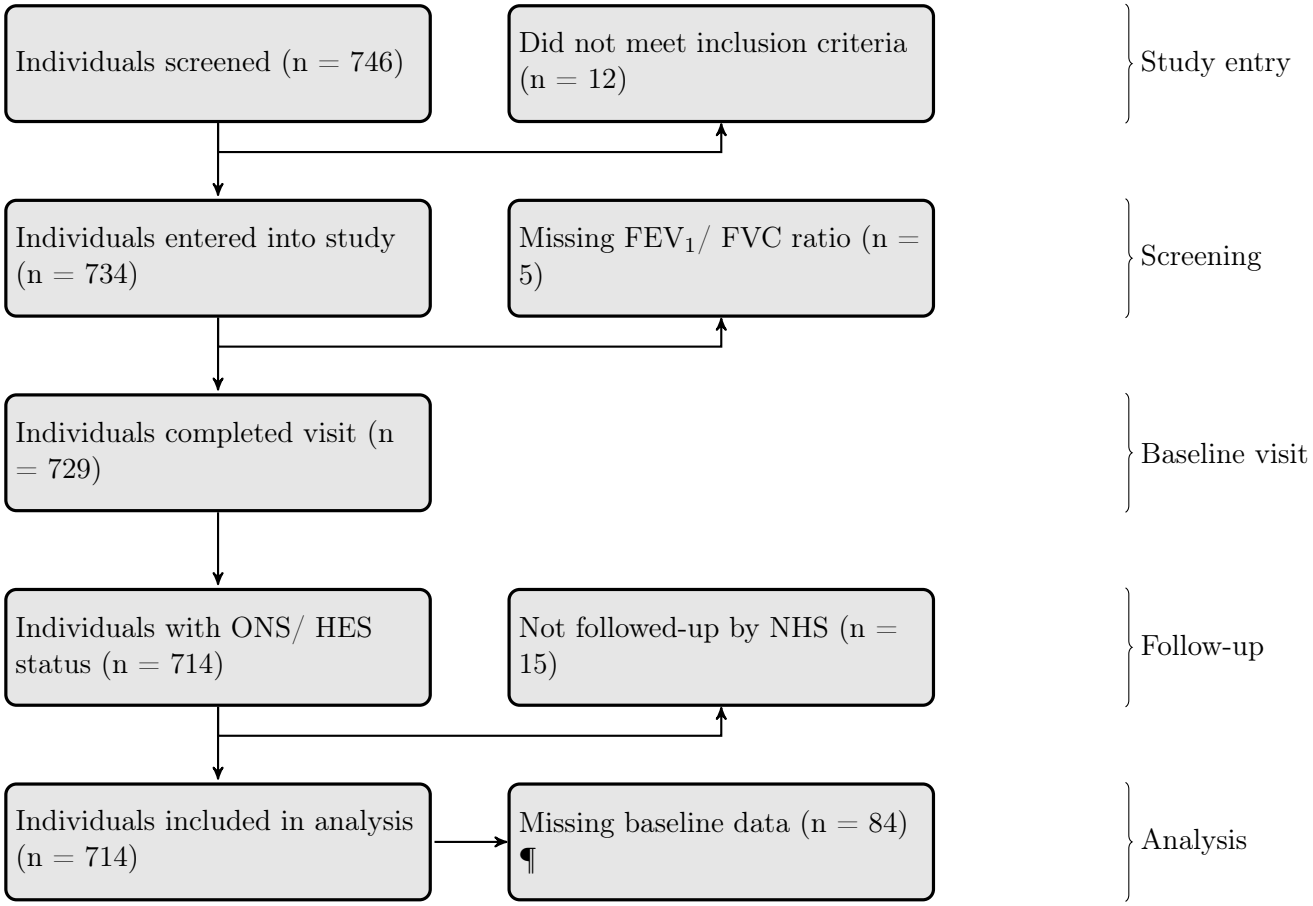


Figure 3.4: Participant enrolment flow diagram. *Abbreviations*: FEV_1 = forced expiratory volume one second. FVC = forced vital capacity. ONS, Office for National Statistics. HES, hospital episode statistics. NHS = National Health Services.
¶Missing baseline data: body mass index (n = 7), smoking status (n = 4), Medical Research Council dyspnoea score (n = 5), six-minute walk distance (n = 34), short physical performance battery (n = 8), quadriceps maximum voluntary contraction (n = 27), sniff nasal inspiratory pressure (n = 26).

Table 3.1: Baseline variables, by recruitment site.

Characteristic	Total	N (%)	Sex		Recruitment site				
			Female (n = 280)	Male (n = 434)	Cambridge (n = 86)	Edinburgh (n = 97)	Cardiff (n = 370)	Nottingham (n = 102)	London (n = 59)
Description									
Age (years)	67 (62-73)	714 (100)	66 (62-72)	68 (63-74)	69 (65-74)	69 (64-75)	67 (62-73)	68 (62-72)	65 (60-72)
Sex - male, n (%)	434 (61)	714 (100)	-	-	67 (78)	57 (59)	197 (53)	73 (72)	40 (68)
BMI (kg/m^2)	27 (23-31)	707 (99)	27 (23-31)	27 (24-31)	26 (23-30)	26 (23-29)	28 (24-32)	27 (22-31)	23 (20-27)
Lung function									
FEV_1 (litre)	1.3 (0.9-1.7)	712 (100)	1.0 (0.8-1.3)	1.5 (1.1-2.0)	1.3 (1.0-1.8)	1.3 (1.0-1.7)	1.2 (0.9-1.6)	1.5 (1.1-1.8)	1.0 (0.7-1.5)
FEV_1 % predicted	53 (40-65)	712 (100)	55 (43-65)	52 (39-65)	53 (35-61)	53 (41-66)	54 (42-66)	54 (44-63)	40 (27-60)
Smoking status - current, n (%)	218 (31)	710 (99)	93 (33)	125 (29)	14 (16)	31 (32)	137 (37)	29 (29)	7 (12)
MRC dyspnoea score, ≥2, n (%)	646 (91)	709 (99)	261 (94)	385 (89)	75 (88)	92 (95)	345 (94)	81 (80)	53 (93)
GOLD - stage IV, n (%)	67 (9)	713 (100)	14 (5)	53 (12)	12 (14)	3 (3)	28 (8)	5 (5)	19 (32)
Exacerbation history (1 year), ≥1	473 (67)	709 (99)	212 (77)	261 (60)	60 (71)	64 (66)	242 (65)	68 (67)	39 (68)
Productive cough - yes, n (%)	327 (46)	708 (99)	128 (46)	199 (46)	30 (35)	37 (38)	177 (48)	56 (55)	27 (47)
Biochemical measures									
Glucose (mmol/L)	4.9 (4.5-5.4)	687 (96)	4.9 (4.5-5.4)	4.9 (4.5-5.3)	4.7 (4.4-5.2)	4.9 (4.6-5.3)	4.9 (4.6-5.4)	5.1 (4.7-5.5)	4.7 (4.3-5.1)
Fibrinogen (g/dL)	3.4 (2.9-3.9)	700 (98)	3.4 (3.0-3.9)	3.3 (2.9-4.0)	3.1 (2.8-3.4)	3.5 (3.0-3.8)	3.4 (2.9-4.0)	3.5 (3.1-4.2)	3.5 (3.0-3.9)
CRP (mg/L)	3.4 (1.6-7.5)	696 (97)	3.3 (1.5-7.7)	3.4 (1.7-7.2)	3.0 (1.5-6.3)	4.1 (2.0-9.1)	3.3 (1.6-7.7)	3.3 (1.4-7.8)	4.0 (1.0-7.0)
WCC (mcL)	7.1 (6.0-8.6)	704 (99)	7.1 (5.8-8.5)	7.2 (6.2-8.6)	7.0 (6.0-8.2)	7.4 (6.2-8.9)	7.1 (6.0-8.5)	7.5 (6.1-9.3)	7.1 (6.3-9.0)
GFR (mL/min/1.73 m^2)	87.2 (76.5-101.0)	679 (95)	86.0 (77.2-96.1)	89.3 (75.4-104.4)	77.6 (63.6-90.2)	90.2 (77.8-100.3)	91.7 (80.7-103.7)	78.6 (70.1-88.2)	98.6 (81.4-113.5)
Neutrophils (mm^3)	4.5 (3.6-5.6)	701 (98)	4.3 (3.4-5.6)	4.5 (3.7-5.6)	4.5 (3.8-5.5)	4.7 (3.8-5.8)	4.3 (3.5-5.4)	4.8 (3.6-5.8)	4.4 (3.4-5.9)
Haemoglobin (g/dL)	14.3 (13.4 - 15.3)	703 (98)	13.8 (12.9-14.5)	14.7 (13.8-15.6)	14.4 (13.5-15.0)	14.7 (13.8-15.5)	14.1 (13.05-15.2)	14.8 (13.8-15.9)	14.3 (13.5-15.0)
HbA1c (mmol/mol)	41 (38-45)	692 (97)	41 (38-46)	41 (38-44)	42 (39-45)	40 (37-43)	42 (40-46)	39 (37-41)	41 (39-43)
HDL (mmol/L)	1.4 (1.2-1.7)	706 (99)	1.5 (1.3-1.8)	1.3 (1.1-1.6)	1.5 (1.2-1.9)	1.4 (1.1-1.6)	1.3 (1.1-1.6)	1.6 (1.3-2.1)	1.5 (1.3-1.8)
Total cholesterol (mmol/L)	5.0 (4.3-5.8)	705 (99)	5.4 (4.7-6.1)	4.8 (4.2-5.6)	5.1 (4.2-5.9)	5.0 (4.4-5.7)	5.0 (4.3-5.9)	5.05 (4.4-5.9)	4.9 (4.3-5.6)

Characteristic	Total	N (%)	Sex		Recruitment site				
			Female (n = 280)	Male (n = 434)	Cambridge (n = 86)	Edinburgh (n = 97)	Cardiff (n = 370)	Nottingham (n = 102)	London (n = 59)
Cardiovascular									
status									
Heart rate (bpm)	74 (66-82)	702 (98)	76 (68-83)	73 (65-82)	74 (67-83)	72 (64-79)	73 (66-81)	78 (67-86)	79 (68-87)
SBP (mmHg)	142 (131-154)	706 (99)	140 (129-154)	143 (132-155)	143 (131-154)	139 (125-148)	144 (133-157)	141 (131-157)	136 (128-145)
Arterial stiffness									
CIMT (mm)	0.81 (0.71-0.96)	648 (91)	0.79 (0.69-0.94)	0.82 (0.72-0.97)	0.76 (0.69-0.86)	0.82 (0.72-0.96)	0.83 (0.74-0.96)	0.82 (0.66-0.96)	0.74 (0.62-0.91)
PWV (m/sec)	9.8 (8.4-11.8)	654 (92)	9.5 (8.2-11.2)	10.1 (8.5-12.1)	10.1 (8.5-11.9)	9.8 (8.9-12.3)	9.9 (8.5-11.8)	8.9 (6.9-10.5)	10.1 (8.2-12.0)
AIx (%)	28 (20-34)	699 (98)	31 (25-37)	25 (18-32)	27 (19-30)	29 (22-36)	28 (21-35)	27 (20-32)	24 (17-33)
Musculoskeletal									
measures									
6MW distance (metre)	366 (255-440)	680 (95)	340 (240-420)	383 (275-457)	420 (300-498)	399 (298-480)	346 (220-420)	363 (255-436)	423 (302-480)
SPPB score (0-12)	10 (8-11)	706 (99)	9 (7-11)	11 (9-12)	11 (10-12)	11 (9-11)	9 (7-11)	11 (9-12)	11 (10-12)
- 4MGS (0-4)	4 (3-4)	709 (99)	4 (3-4)	4 (4-4)	4 (4-4)	4 (4-4)	4 (3-4)	4 (4-4)	4 (4-4)
- Balance (0-4)	4 (4-4)	711 (100)	4 (3-4)	4 (4-4)	4 (4-4)	4 (4-4)	4 (3-4)	4 (4-4)	4 (4-4)
- Chair stand (0-4)	3 (1-4)	707 (99)	2 (1-4)	3 (1-4)	3 (2-4)	3 (2-4)	2 (1-3)	3 (1-4)	3 (2-4)
QMVC peak (kg)	30 (22-39)	687 (96)	22 (18-28)	35 (28-43)	31 (26-40)	30 (24-38)	28 (20-37)	33 (25-42)	31 (23-36)
SNIP (cm H_2O)	53 (38-70)	688 (96)	46 (33-61)	58 (42-77)	65 (45-81)	70 (59-83)	44 (32-58)	66 (51-81)	51 (39-67)
Questionnaires									
SGRQ-C (0-100)	51 ((34-66)	692 (97)	53 (37-70)	50 (32-64)	42 (29-58)	48 (32-60)	55 (37-71)	50 (32-63)	50 (36-63)
CAT (0-40)	20 (13-26)	702 (98)	21 (15-27)	19 (13-25)	16 (12-22)	19 (13-25)	21 (15-27)	20 (14-25)	18 (12-24)
Diabetes - yes	82 (12)	709 (99)	30 (11)	52 (12)	14 (16)	5 (5)	51 (14)	10 (10)	2 (4)
CVD drugs - yes	402 (56)	714 (100)	155 (55)	247 (57)	41 (48)	59 (61)	233 (63)	47 (46)	22 (37)

Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of 714 patients were included. Diabetes and the use of CVD drugs were self-reported. *Abbreviations*: BMI, body mass index. FEV_1, forced expiratory volume in one second. MRC, Medical Research Council. GOLD, global initiative for obstructive lung disease. CRP, C-reactive protein. WCC, white cell count. GFR, glomerular filtration rate. HbA1c, glycated haemoglobin. HDL, high-density lipoprotein. SBP, systolic blood pressure. CIMT, carotid intima-media thickness. PWV, pulse wave velocity. AIx, augmentation index. 6MW, six-minute walk. SPPB, short physical performance measure. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction. SNIP, sniff nasal inspiratory pressure. SGRQ-C, St. George respiratory questionnaire for COPD. CAT, COPD assessment test. CVD, cardiovascular disease.

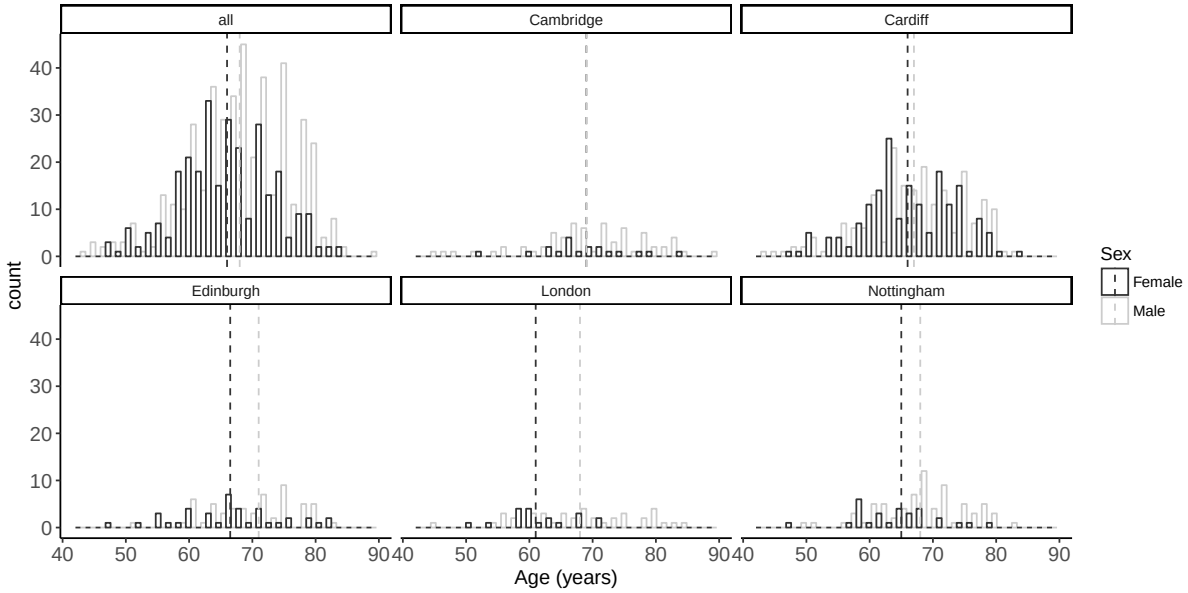


Figure 3.5: Histograms displaying the distribution of age, by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

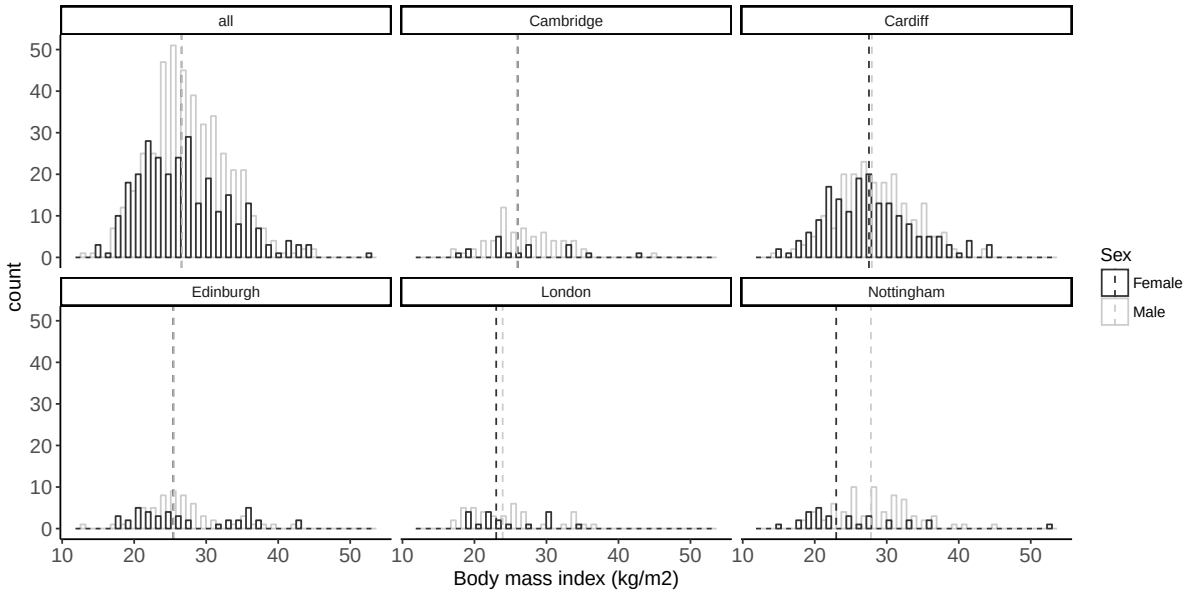


Figure 3.6: Histograms displaying the distribution of body mass index, by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

3.3.3 Lung function measures

All individuals were ex-smokers with at least ten pack years of smoking by default with 218 (31%) current smokers, and 406 (57%) were identified as GOLD stage II equating to a median (IQR) FEV_1 of 1.3 (0.9-1.7) litre. About 46% (n = 327) had self-reported productive cough – mucus or phlegm – on most mornings. The median number of self-reported exacerbations of COPD one year before study enrolment was 1 (0-3; **Table 3.1**, page 65, and **Figures 3.7 3.8**, pages 68-69). Individuals from London had a significantly lower FEV_1 % predicted compared to the other sites ($p < 0.001$). Edinburgh had a significant higher number of women with an exacerbation history than men ($p = 0.004$).

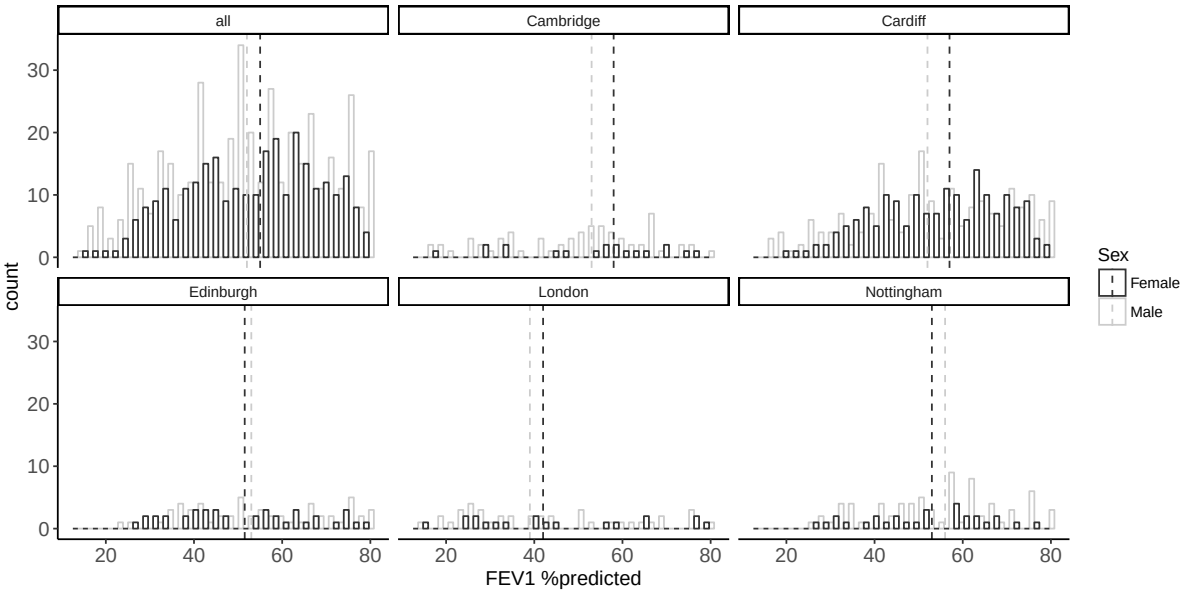


Figure 3.7: Histograms displaying the distribution of forced expiratory volume in one-second (FEV_1) percentage predicted, by sex and recruitment site. Dashed lines indicate median values by sex.

3.3.4 Biochemical measures

The median (IQR) levels of circulating inflammatory markers were 3.4 (2.9 - 3.9) g/dL for fibrinogen and 3.4 (1.6 - 7.5) mg/L for C-reactive protein (CRP). Median white cell count (WCC; i.e. leukocytes) was 7.1 (6.0 - 8.6) mcL. Neutrophils were 1.4 (1.2 - 1.7) mmol/L. There

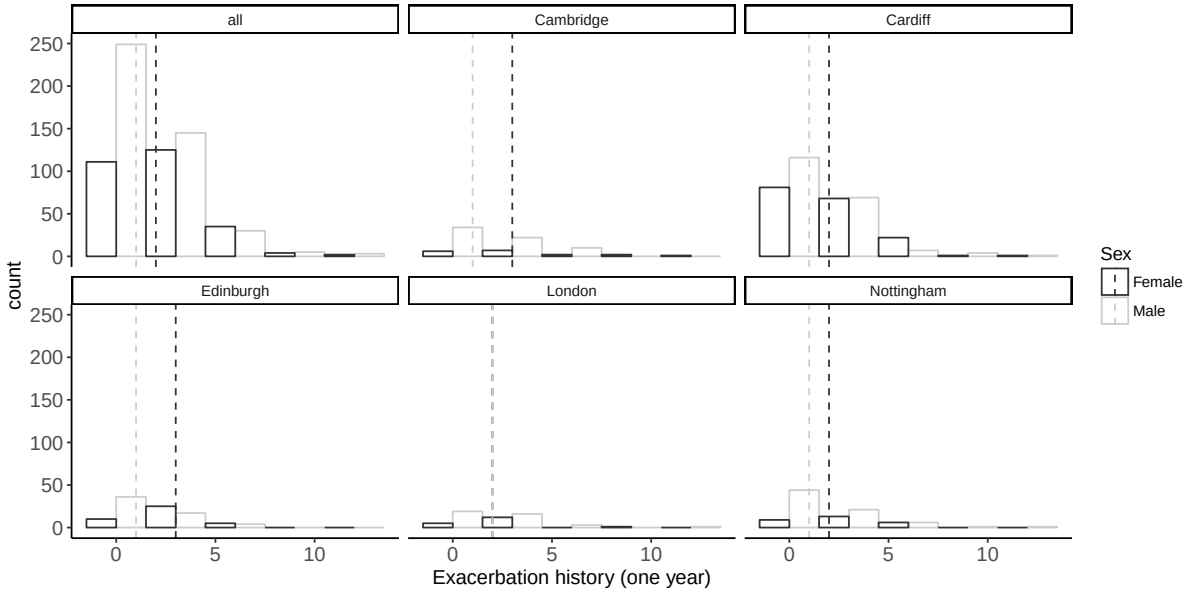


Figure 3.8: Histograms displaying the distribution of exacerbation history, one year before study enrolment, by sex and recruitment site. Dashed lines indicate median values by sex.

was a significant difference in fibrinogen levels between sites with the lowest in Cambridge (3.1; p = 0.001) but not for other inflammatory markers CRP (p = 0.470), WCC (p = 0.305), and neutrophils (p = 0.136; **Table 3.1**, page 65, and **Figures 3.9 3.10 3.11 3.12**, pages 70-71).

Median haemoglobin levels were 14.3 (13.4 - 15.3) g/L, 41 (38 - 45) mmol/mol for HbA1c and 4.9 (4.5 - 5.4) mmol/L for glucose. Haemoglobin levels were higher for men than for women (p <0.001). Median GFR values were 87 (76 - 101) mL/min/1.73 m^2. Glomerular filtration rates decreased with advancing age (p <0.001; **Table 3.1**, page 65, and **Figures 3.13 3.14 3.15**, pages 72-73).

Individuals from Cardiff had the lowest high-density lipoprotein (HDL) levels (1.3; p <0.001). Median HDL levels were 1.4 (1.2 - 1.7) mmol/L, with higher levels indicating a lower risk of heart disease. Median total cholesterol levels were 5.0 (4.3 - 5.8) mmol/L (**Table 3.1**, page 65, and **Figures 3.16 3.17**, pages 73-74).

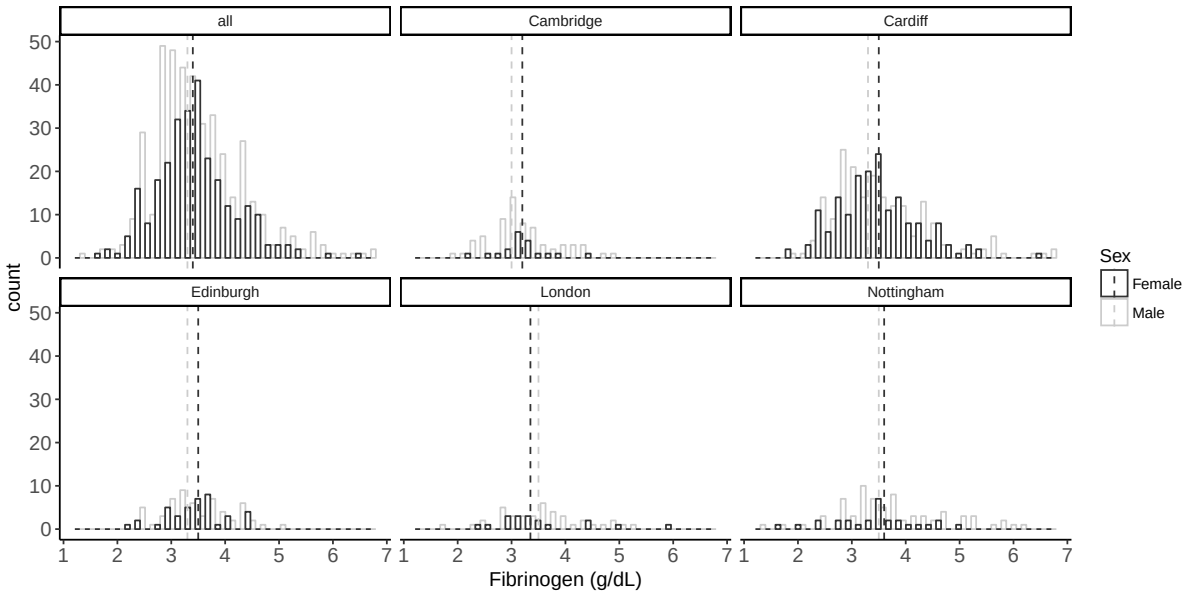


Figure 3.9: Histograms displaying the distribution of fibrinogen by sex and recruitment site. Dashed lines indicate median values by sex.

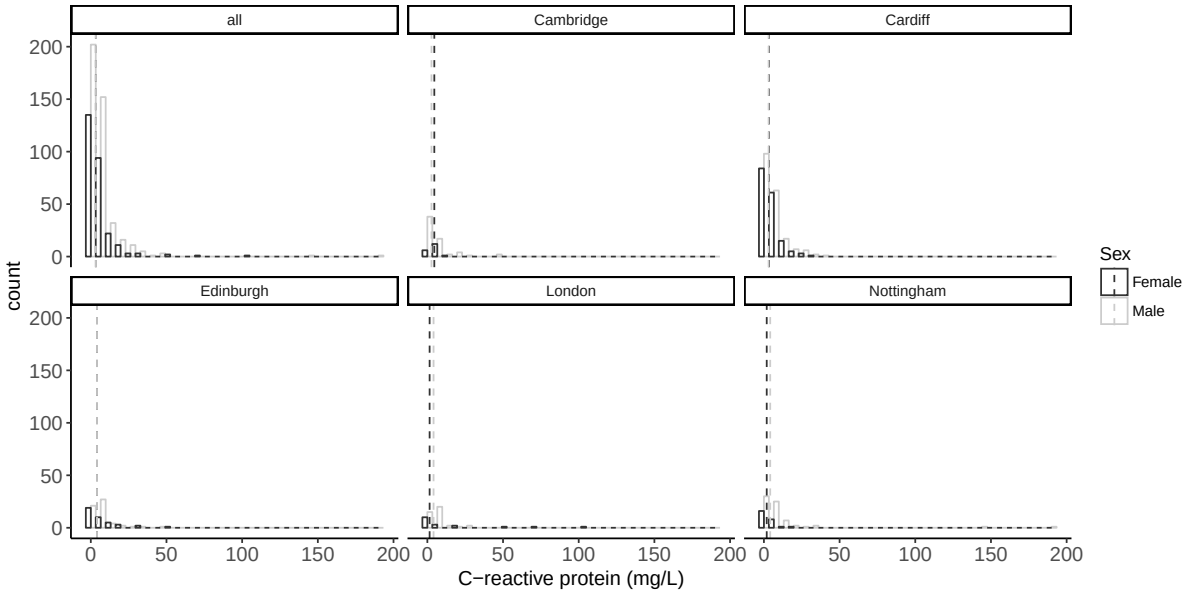


Figure 3.10: Histograms displaying the distribution of C-reactive protein by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

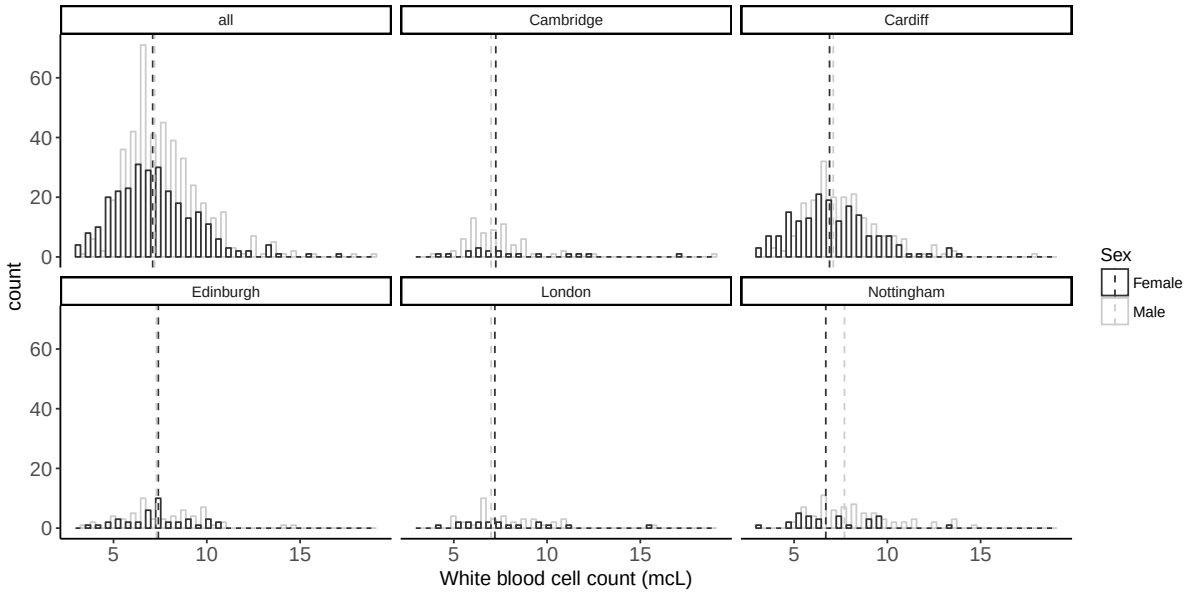


Figure 3.11: Histograms displaying the distribution of white cell count by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

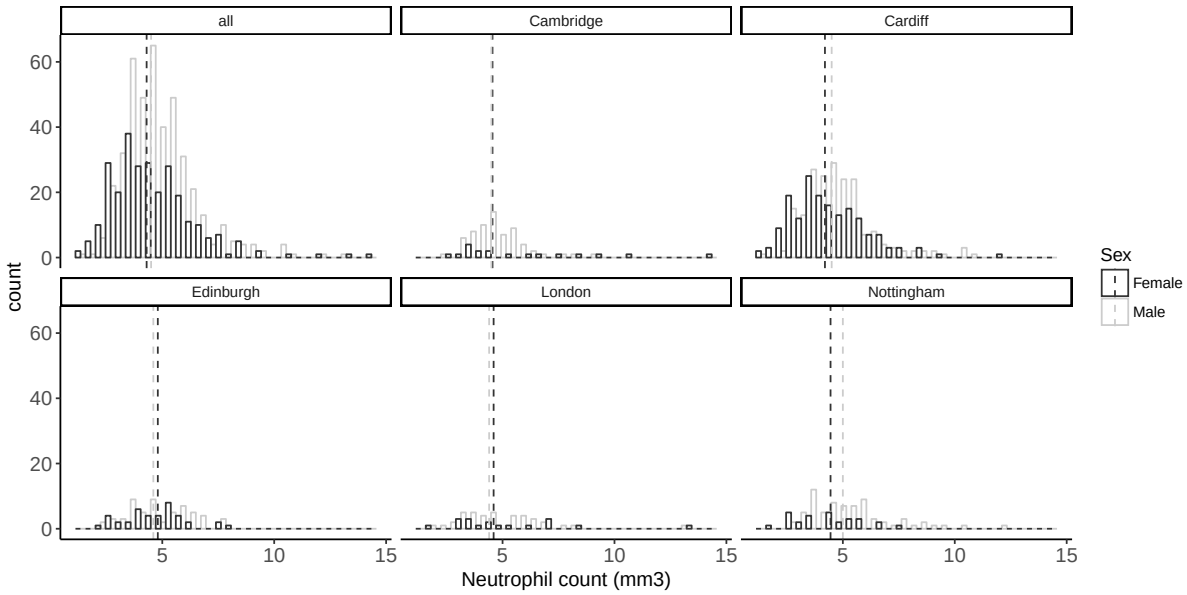


Figure 3.12: Histograms displaying the distribution of neutrophil count by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

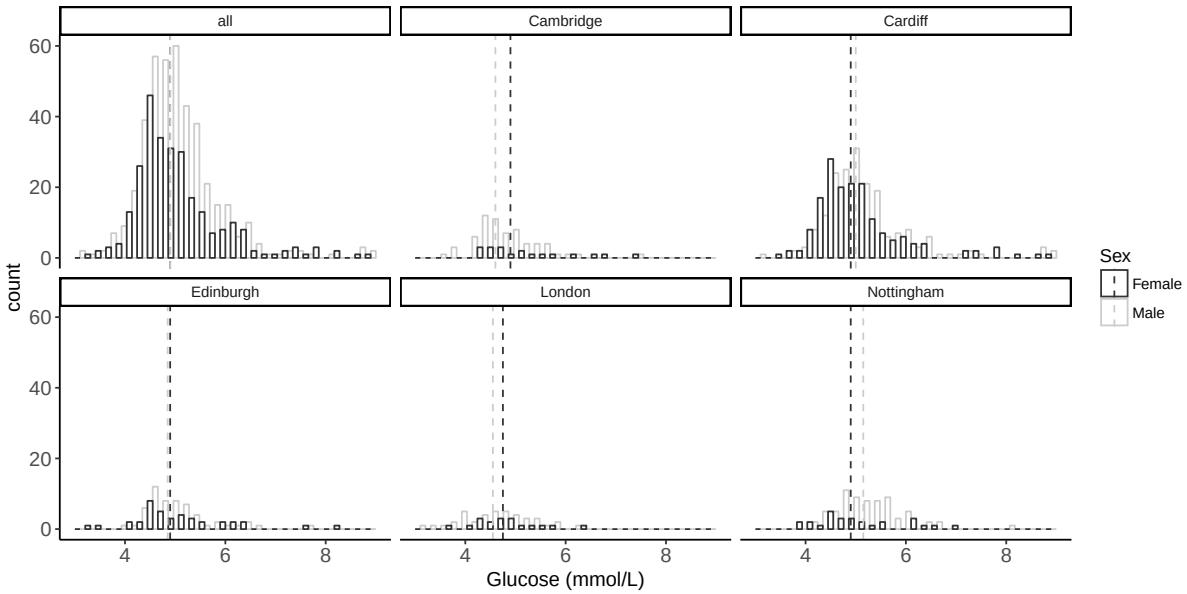


Figure 3.13: Histograms displaying the distribution of glucose level by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

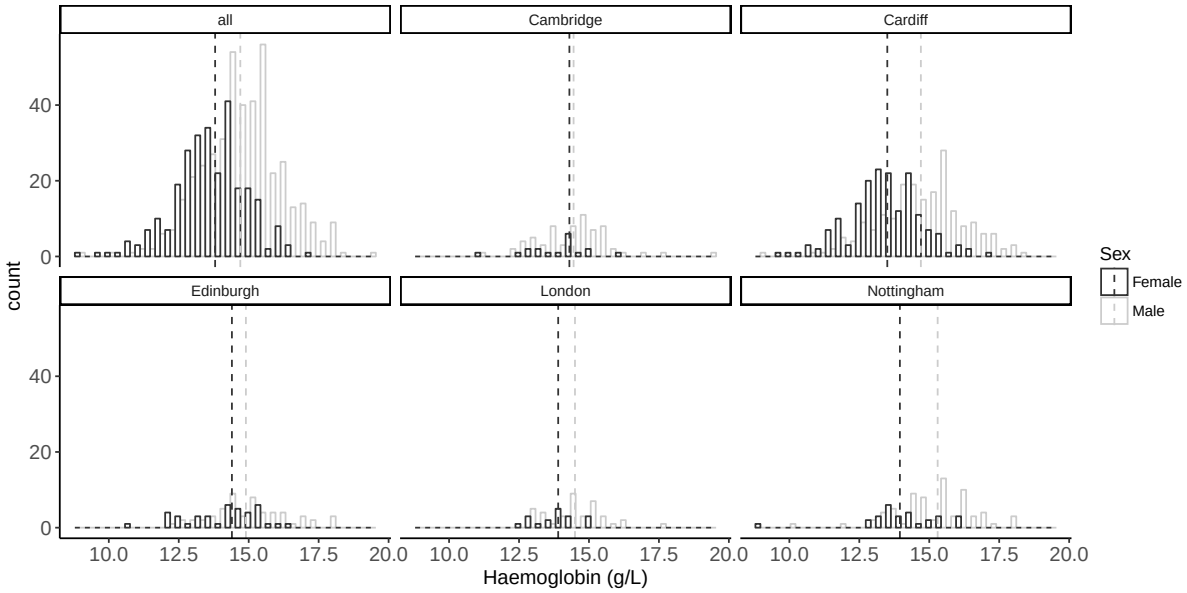


Figure 3.14: Histograms displaying the distribution of haemoglobin by sex and recruitment site. Dashed lines indicate median values by sex.

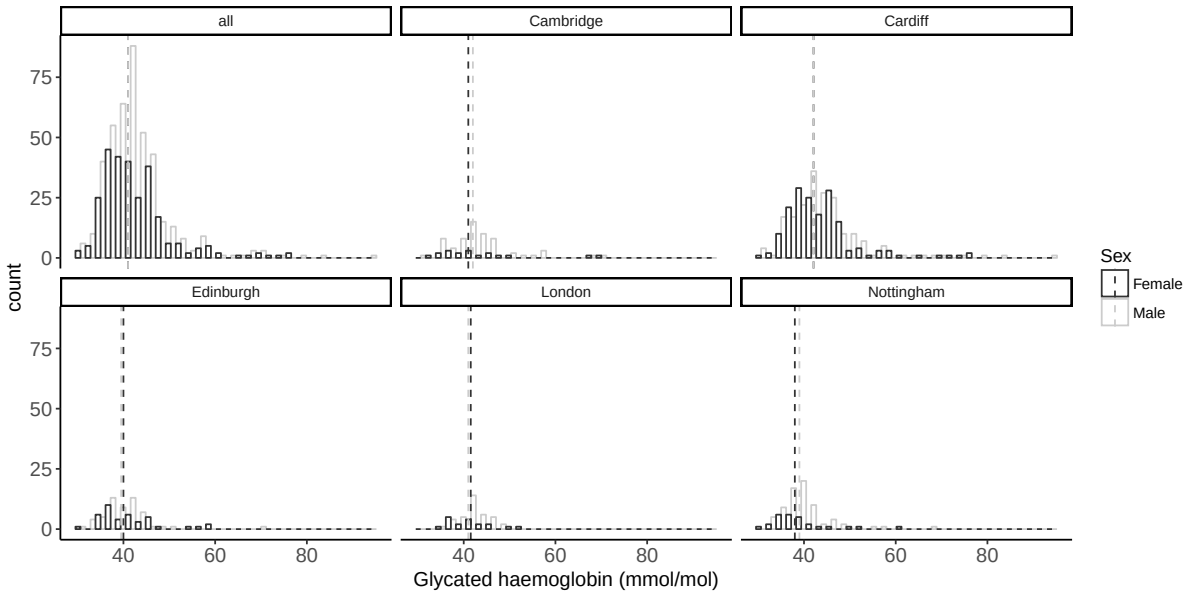


Figure 3.15: Histograms displaying the distribution of glycated haemoglobin by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

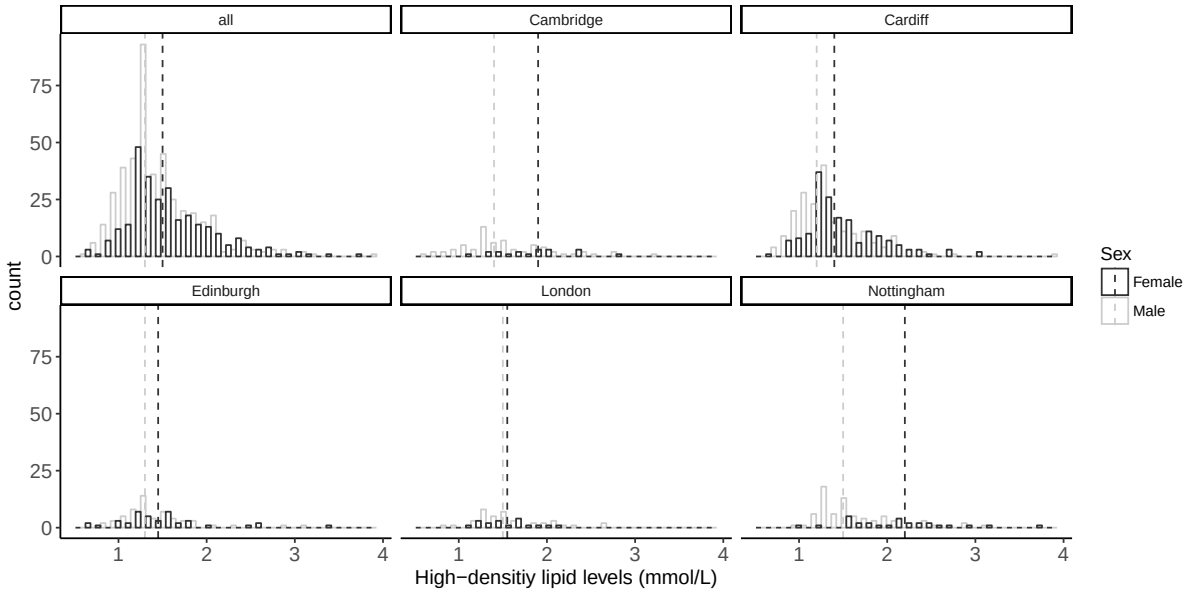


Figure 3.16: Histograms displaying the distribution of high-density lipoprotein levels by sex and recruitment site. Dashed lines indicate median values by sex.

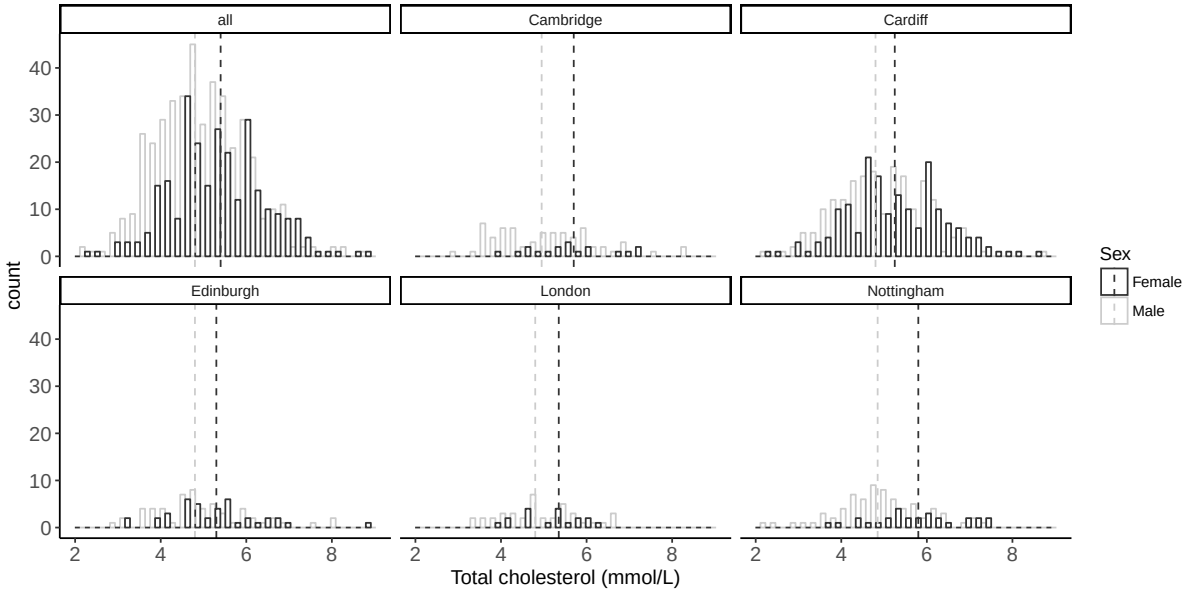


Figure 3.17: Histograms displaying the distribution of total cholesterol levels by sex and recruitment site. Dashed lines indicate median values by sex.

3.3.5 Musculoskeletal measures

The median (IQR) 6MW distance was 366 (255 - 440) metres, with shortest distances in Nottingham and longest in London (p <0.001). Men completed longer walking distances than women (p <0.001). More than 40% (n = 292) had functional limitation with a median total SPPB score of 10 (8 - 11), and SPPB components 4MGS 4 (3 - 4), balance 4 (4 - 4), and chair stand 3 (1 - 4) points. Men scored two median points higher than women (p <0.001; **Table 3.1**, page 65, and **Figures 3.18 3.19**, pages 75-75).

Median values of peak quadriceps contraction were 30 (22 - 39) kg. Scores were the lowest for those in Cardiff (28 kg) and highest in Nottingham (33 kg; p <0.001). Males had a median additional 13 kg of quadriceps strength compared to females (p <0.001). The sniff nasal inspiratory pressure had a median score of 53 (38 - 70) cm H_2O. Women, and generally those in Cardiff had significantly lower median SNIP scores compared to men and other sites respectively (p <0.001, p <0.001; **Table 3.1**, page 65, and **Figures 3.20 3.21**, pages 76-76).

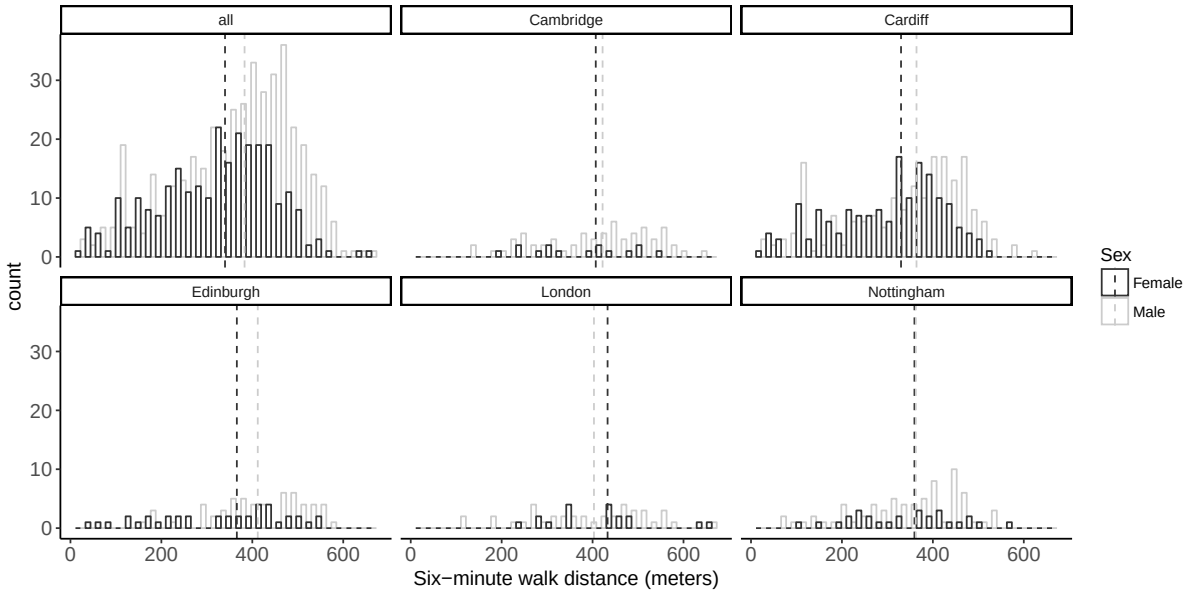


Figure 3.18: Histograms displaying the distribution of six-minute walk distance by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

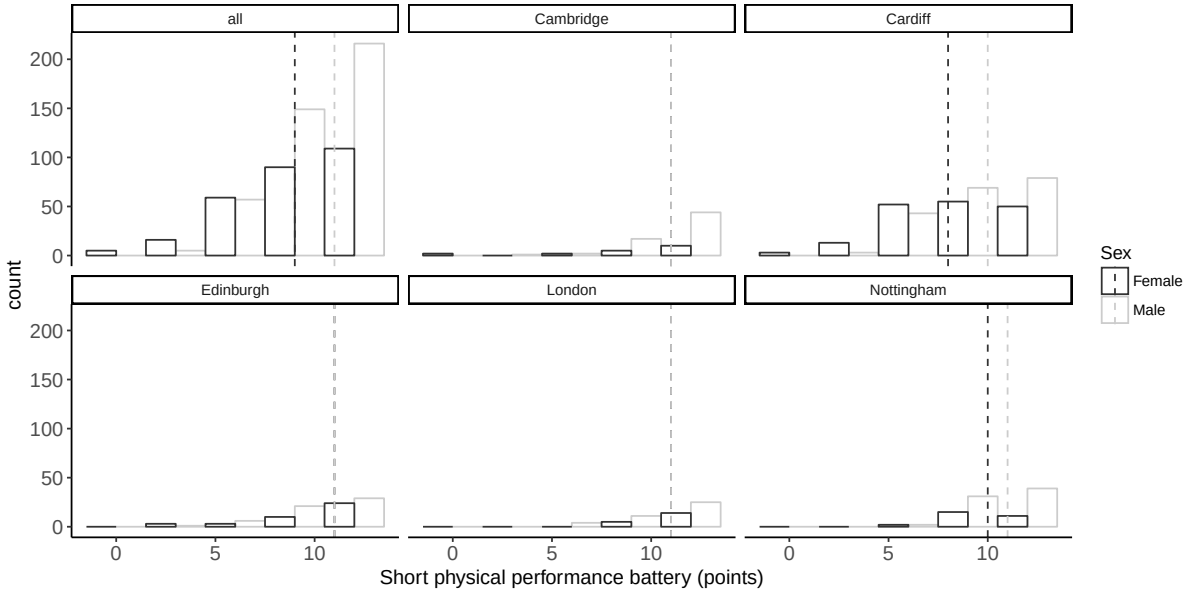


Figure 3.19: Histograms displaying the distribution of the short physical performance battery by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

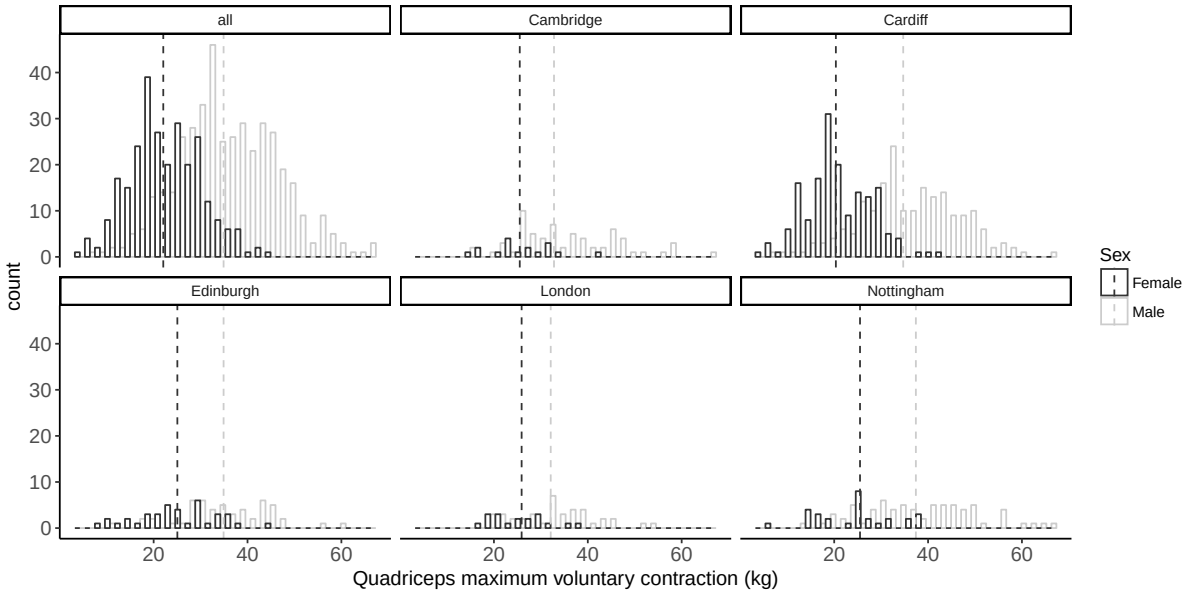


Figure 3.20: Histograms displaying the distribution of quadriceps maximum voluntary contraction by sex and recruitment site. Dashed lines indicate median values by sex.

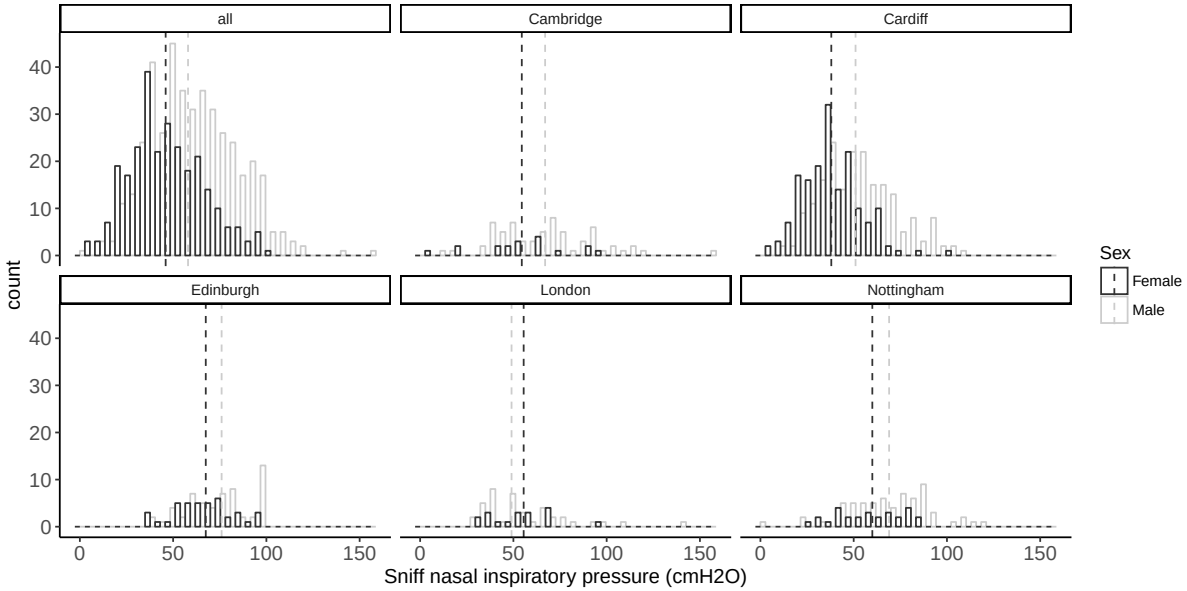


Figure 3.21: Histograms displaying the distribution of sniff nasal inspiratory pressure by sex and recruitment site. Dashed lines indicate median values by sex.

3.3.6 Cardiovascular measures

The median (IQR) resting heart rate was 74 (66-82) beats per minute, with about a third (n = 217) of the cohort a resting heart rate too high. Edinburgh (72 bpm) had the lowest resting heart rate with London (79 bpm) the highest (p = 0.002). Women had slightly higher resting heart rates than men (p = 0.006). Median value of the mean arterial pressure was 103 (95 - 111) mmHg. Median systolic blood pressure was 142 (131 - 154) mmHg, with 92% who had a systolic blood pressure >120 mmHg (**Table 3.1**, page 65, and **Figures 3.22 3.23 3.24**, pages 77-78).

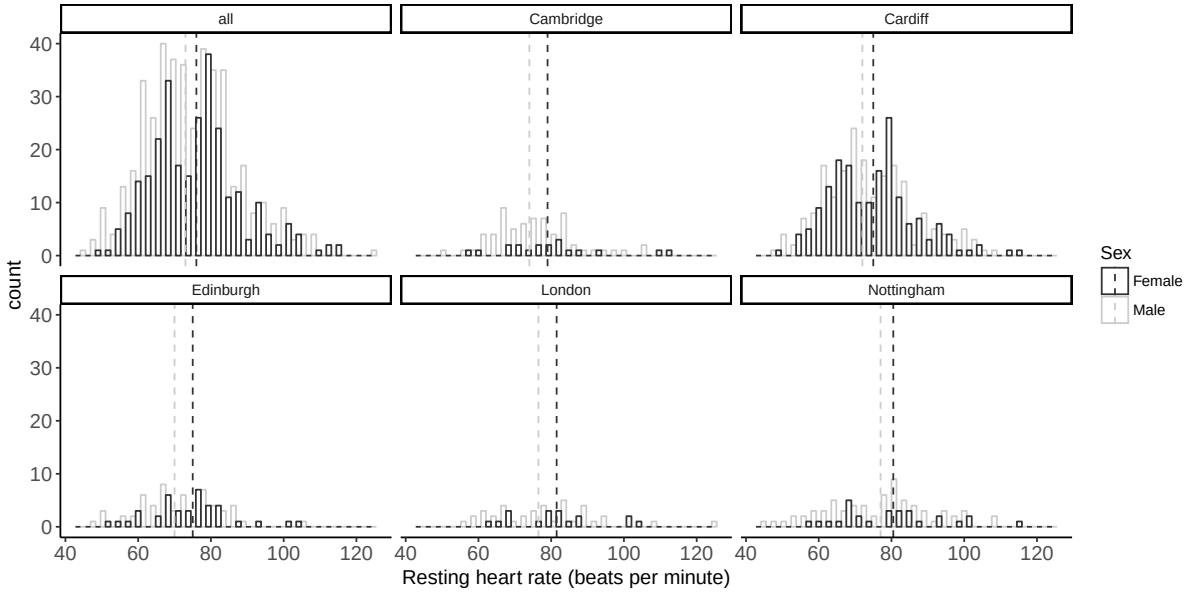


Figure 3.22: Histograms displaying the distribution resting heart rate by sex and recruitment site. Dashed lines indicate median values by sex.

Measures of arterial stiffness include CIMT, PWV, and AIx. Median CIMT was 0.81 (0.71 - 0.96) mm, with Cardiff (0.83) reporting the highest thickness followed by Edinburgh (0.82) and Nottingham (0.82; p = 0.002). Augmentation index had median values of 28% (20 - 34%) with Edinburgh (29%) and woman (31%) reporting highest values, p = 0.002 and p <0.001 respectively. Median PWV was 9.8 (8.4 - 11.8) m/sec. with nearly half of the cohort (n = 310) had an abnormally raised PWV. Cambridge (10.1), London (10.1) but also males (10.1) reported

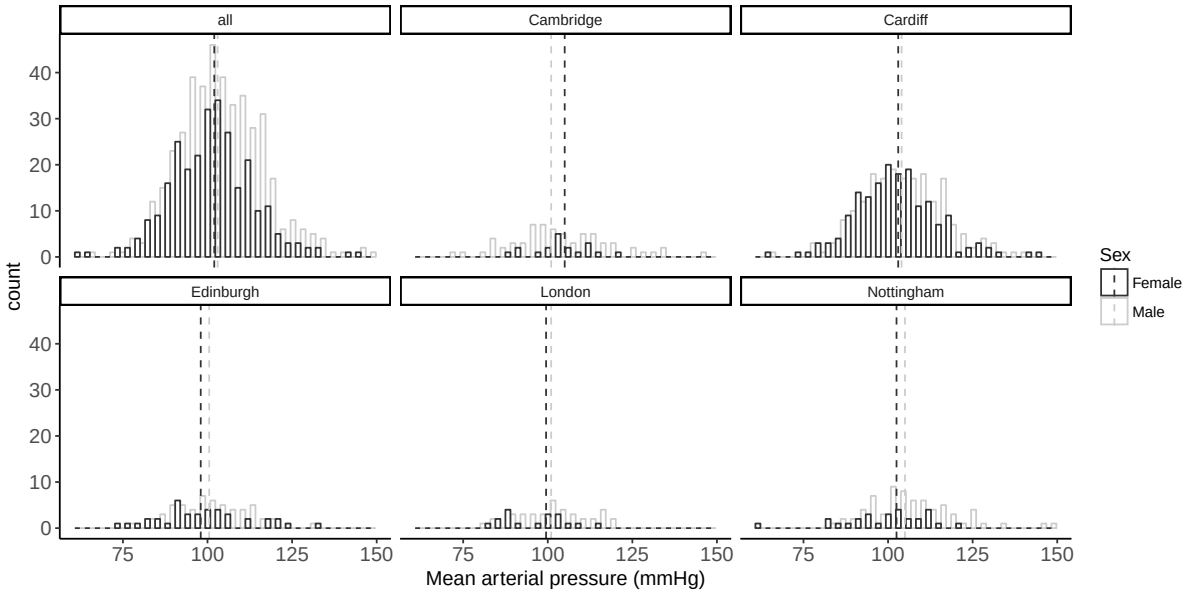


Figure 3.23: Histograms displaying the distribution of mean arterial pressure by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

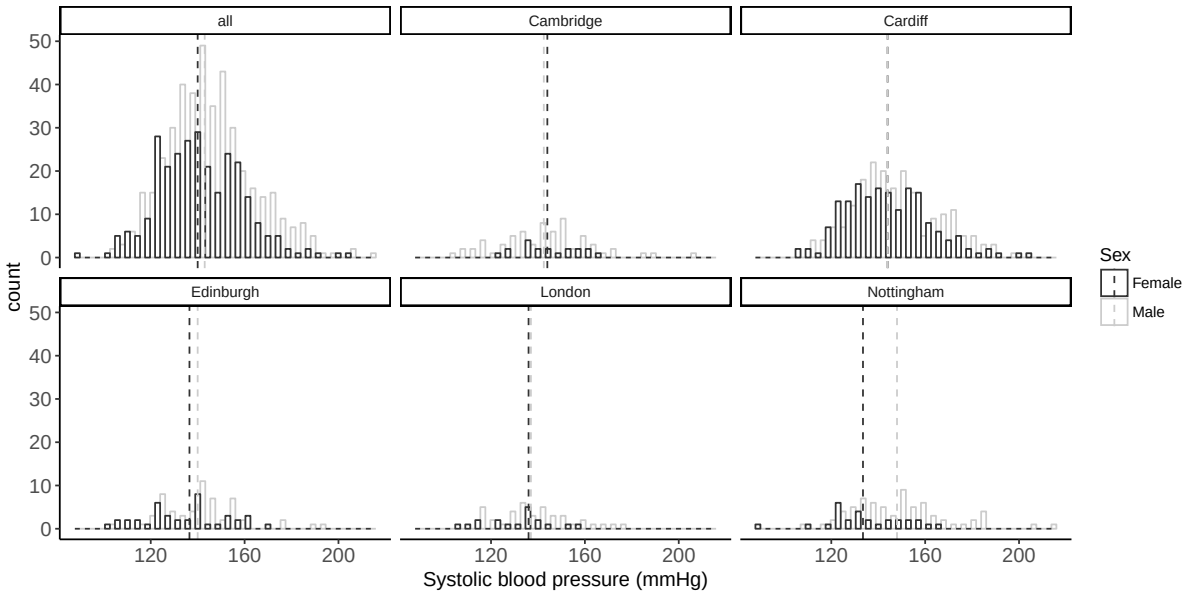


Figure 3.24: Histograms displaying the distribution of systolic blood pressure by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

highest median values (p = 0.002, p = 0.011 respectively; **Table 3.1**, page 65, and **Figures 3.25 3.26 3.27**, pages 79-80).

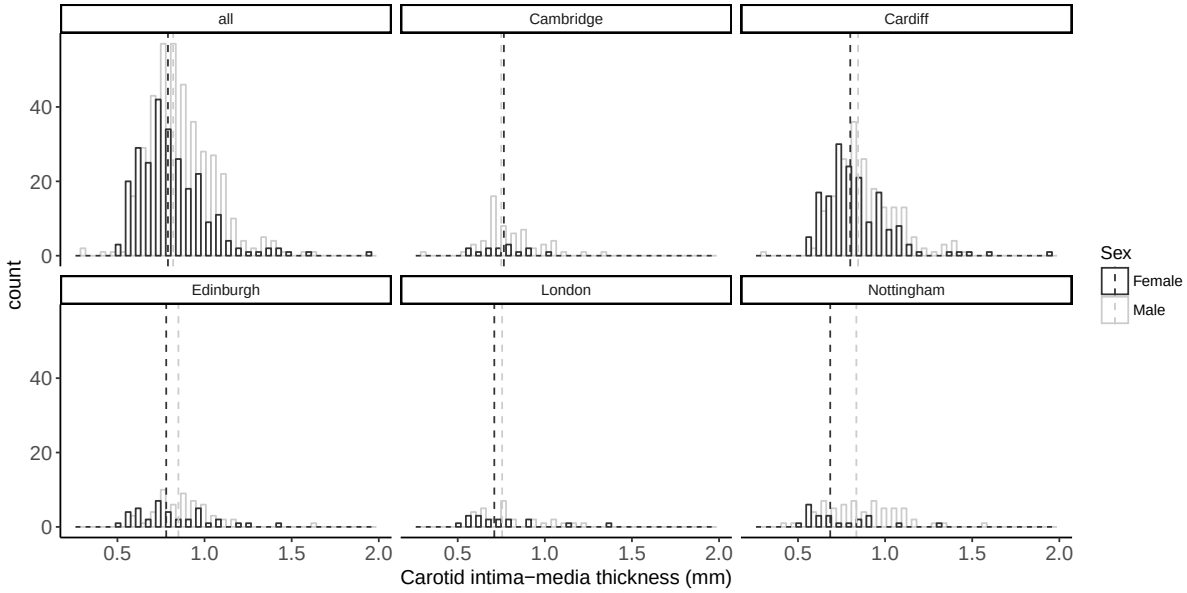


Figure 3.25: Histograms displaying the distribution of carotid intima-media thickness by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

3.3.7 Questionnaires

A proportion over 45% (n = 323) had high COPD Assessment Test (CAT) scores with a median (IQR) of 20 (13-26) points for the cohort, and median SGRQ-C points of 51 (34-66) with Cambridge (SGRQ-C, 42 points; CAT 16 points) scoring the lowest scores for both measures (p <0.001 and p <0.001, respectively). Most individuals (91%) were symptomatic with 40% (n = 281) feeling at least "short of breath when hurrying on the level or walking up a slight hill", and 64 (9%) feeling "too breathless to leave the house or I am breathless when dressing", measured by the MRC dyspnoea scale indicating perceived impact of breathlessness on mobility (i.e. physical activity). Eighty-two (12%) individuals self-reported diabetes, and 402 (56%) self-reported taking medications for treating CV disease. Cardiff (63%) and Edinburgh (61%)

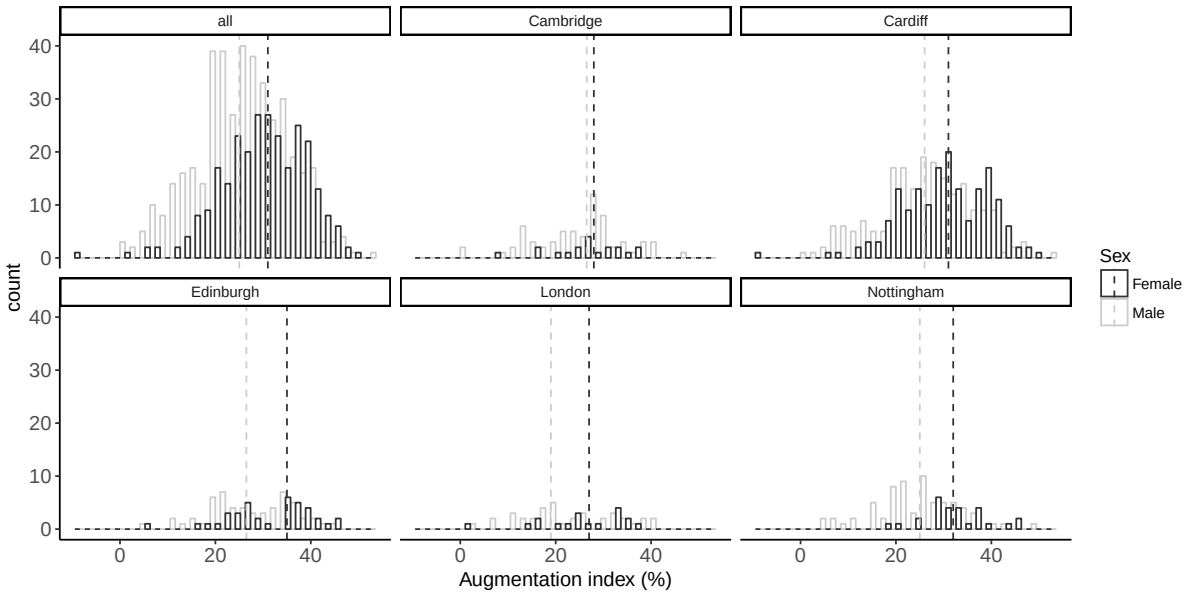


Figure 3.26: Histograms displaying the distribution of augmentation index by sex and recruitment site. Dashed lines indicate median values by sex.

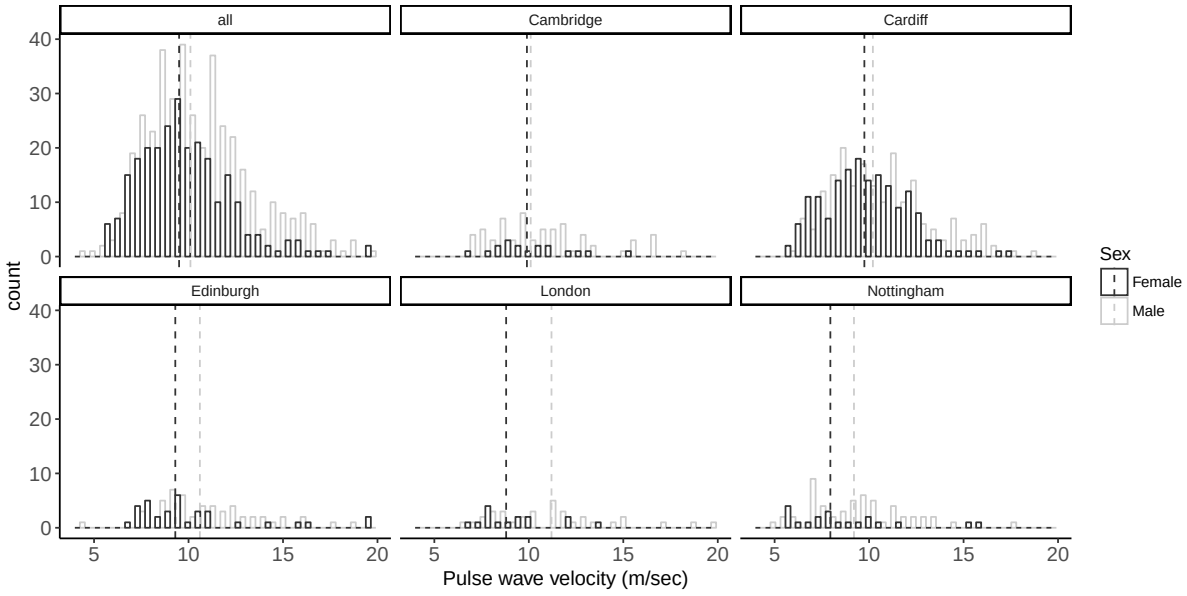


Figure 3.27: Histograms displaying the distribution of pulse wave velocity by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

had a significant higher number of individuals taking such medications with London the least (37%; p <0.001; **Table 3.1**, page 65, and **Figures 3.28 3.29**, pages 81-82).

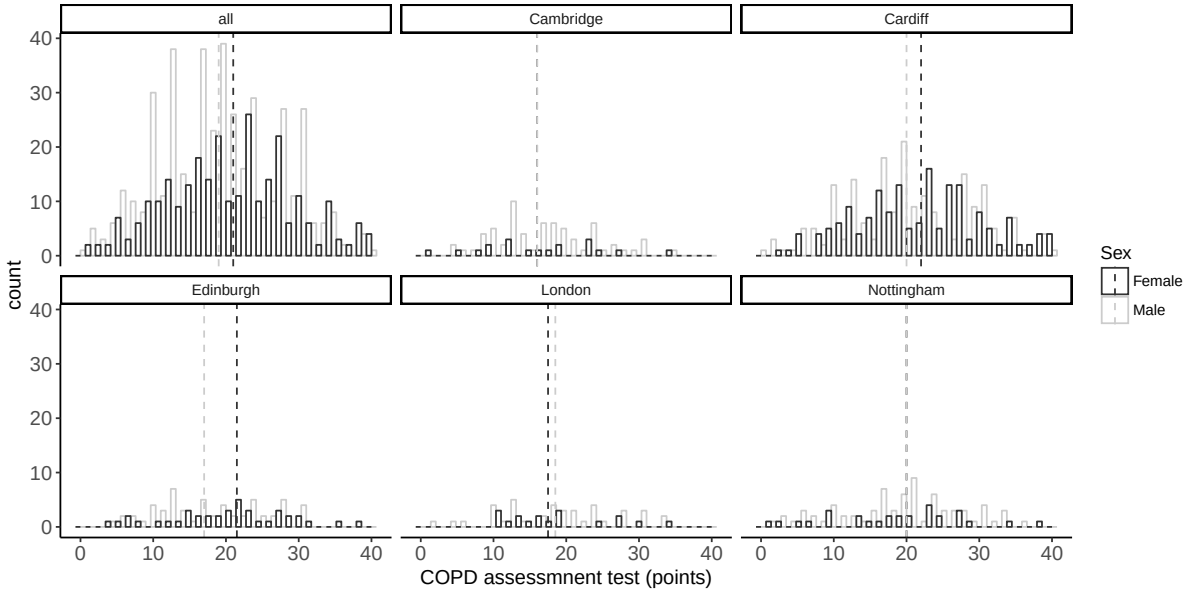


Figure 3.28: Histograms displaying the distribution of COPD assessment test scores by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

3.3.8 Correlations

There was a strong positive correlation between COPD impact measures CAT, MRC, and SGRQ-C; measures of physical functioning SPPB and its components (i.e. 4MGS, balance, and chair stand), and 6MW; between sex and QMVC; and between inflammatory markers fibrinogen and CRP. There was a strong negative correlation between MRC, SGRQ-C, CAT and 6MW. The 6MW distance and MRC dyspnoea score correlated strongly with most variables including CAT, SGRQ-C, and SPPB (**Figure 3.30**, page 83).

After adjustment, correlations between FEV_1 and age, and BMI, exacerbation history, and FEV_1 were weak: reducing lung capacity with advancing age, and higher BMI, and fewer exacerbations with higher FEV_1 (**Figure 3.31**, page 84).

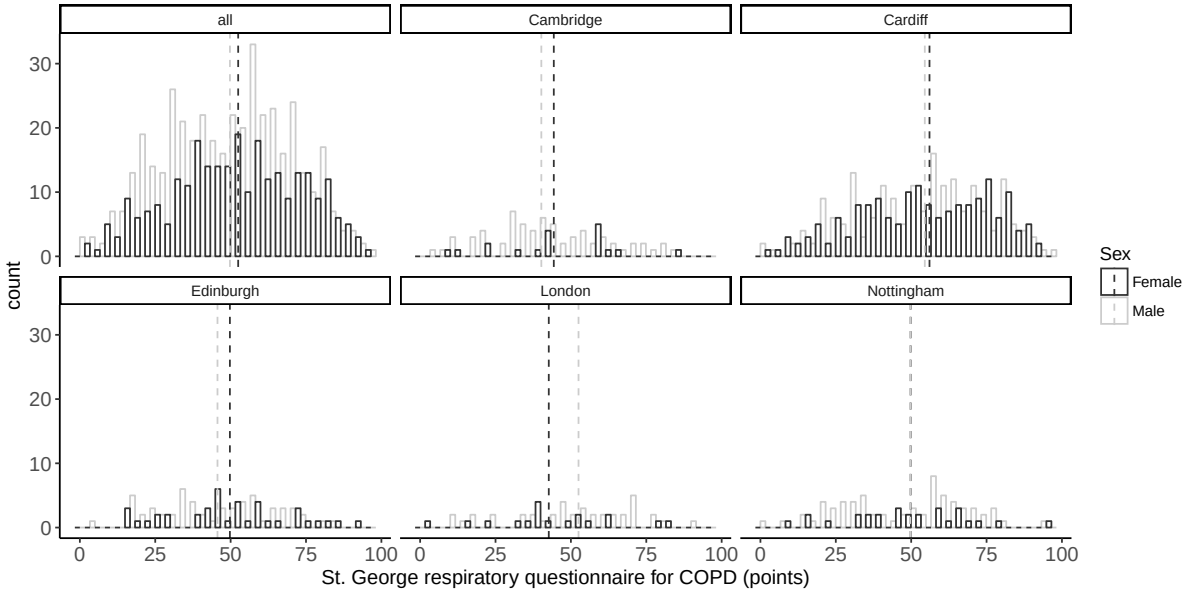


Figure 3.29: Histograms displaying the distribution of St George's respiratory questionnaire for COPD, by sex and recruitment site. Dashed lines indicate median values by sex. Where only one median line is visible, medians for both sexes are similar.

There was little evidence of correlation between inflammatory markers such as fibrinogen and WCC, and FEV_1. This was similar for other biochemical measures such as neutrophils and HDL cholesterol (**Figure 3.32**, page 85).

Data indicated a moderate positive association between 6MW distance, QMVC, and FEV_1 indicating longer walk distances and increasing quadriceps muscle strength with increasing lung capacity. Correlations between resting heart rate and other musculoskeletal measures SPPB or its components (4MGS r = 0.19, balance r = -0.01, and chair stand r = 0.11) and SNIP were weak (**Figure 3.33**, page 86).

There was a strong positive correlation between SPPB (4MGS r = 0.57, balance r = 0.26, and chair stand r = 0.57) and 6MW distance, indicating increasing physical functioning, except for balance, with longer walking distance. Correlations between resting heart rate, QMVC, SNIP, and FEV_1 were weak (**Figure 3.34**, page 87).

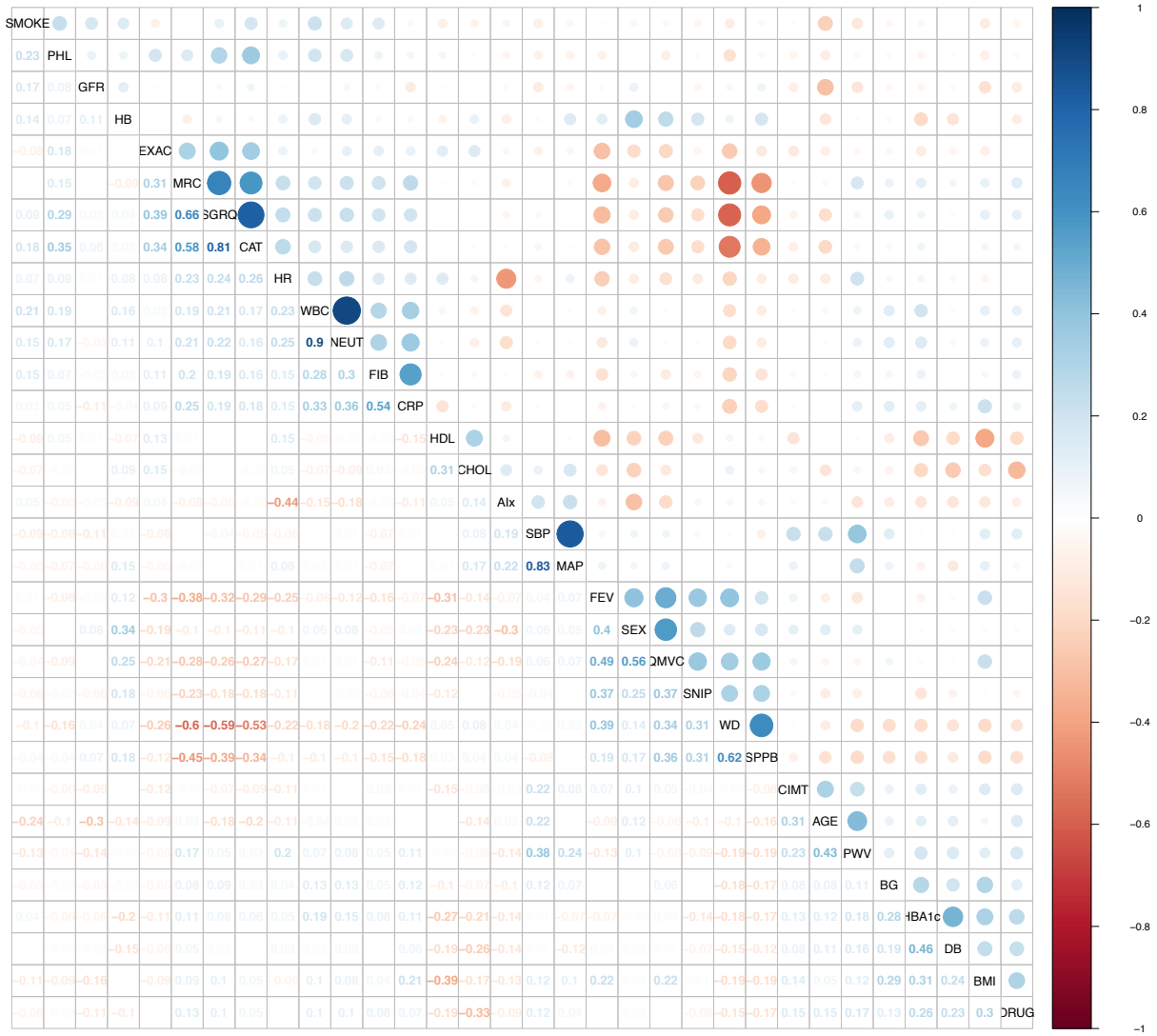


Figure 3.30: Correlation matrix of baseline variables. *Abbreviations*: GFR, Glomerular filtration rate. SMOKE, smoking status. PHL, phlegm. EXAC, exacerbations. MRC, dyspnoea scale. SGRQ, St George's respiratory questionnaire for COPD. CAT, COPD assessment test. HR, heart rate. WCC, white cell count. NEUT, neutrophils. FIB, fibrinogen. CRP, C-reactive protein. HDL, high-density lipoprotein. CHOL, total cholesterol. AIx, augmentation index. SBP, systolic blood pressure. MAP, mean arterial pressure. CIMT, carotid intima-media thickness. PWV, pulse wave velocity. BG, glucose. HBA1c, glycated haemoglobin. DB, diabetes. BMI, body mass index. DRUG = use of cardiovascular drugs. FEV_1, forced expiratory volume in one-second. HB, haemoglobin. QMVC, quadriceps maximum voluntary contraction. SNIP, sniff nasal inspiratory pressure. SPPB, short physical performance battery. WD, six-minute walk distance. Correlation coefficients with a values <0.30 were considered weak, 0.30 - 0.50 as moderate, and >0.50 as strong.[51]

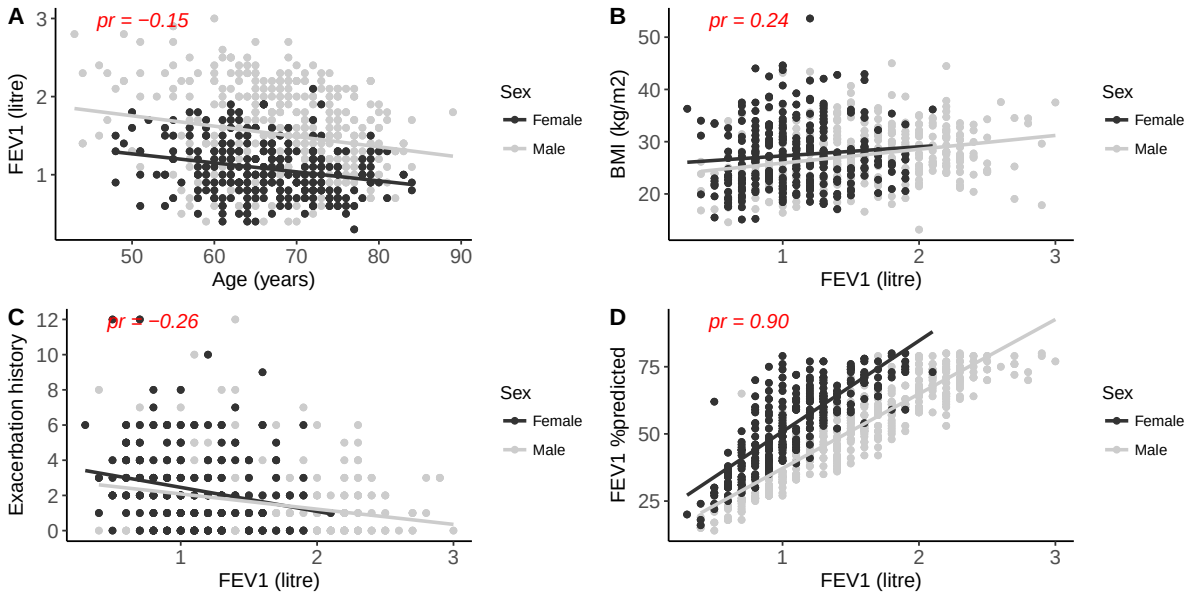


Figure 3.31: Scatter plots displaying the distribution of (A) age, (B) body mass index (BMI), (C) exacerbation history, and (D) forced expiratory volume in one second (FEV_1) percentage predicted, over FEV_1, by sex. Partial correlations (pr) are displayed in the top left corners.

Correlations between COPD symptom questionnaires and FEV_1 were moderate, with higher scores (i.e. increasingly symptomatic) associated with worse lung function (**Figure 3.35**, page 87).

There was an absence of correlation between arterial stiffness measures CIMT, PWV, AIx, and FEV_1 (**Figure 3.36**, page 88).

Increasing CIMT and PWV, but not AIx, were moderately associated with advancing age. Pulse wave velocity correlated moderately with higher systolic blood pressure. Higher AIx was strongly correlated with increasing resting heart rate (**Figure 3.37**, page 89).

3.3.9 Time to event outcomes

Survival data and electronic hospital records were available for 714 individuals, as fifteen individuals were not followed by the NHS. During 75 months of follow-up, a total of 149 deaths (21%) occurred. There was a higher proportion of deaths in Cardiff (49%) compared to the other sites (**Figure 3.38**, page 90). A majority of deaths were due to pulmonary causes (55%),

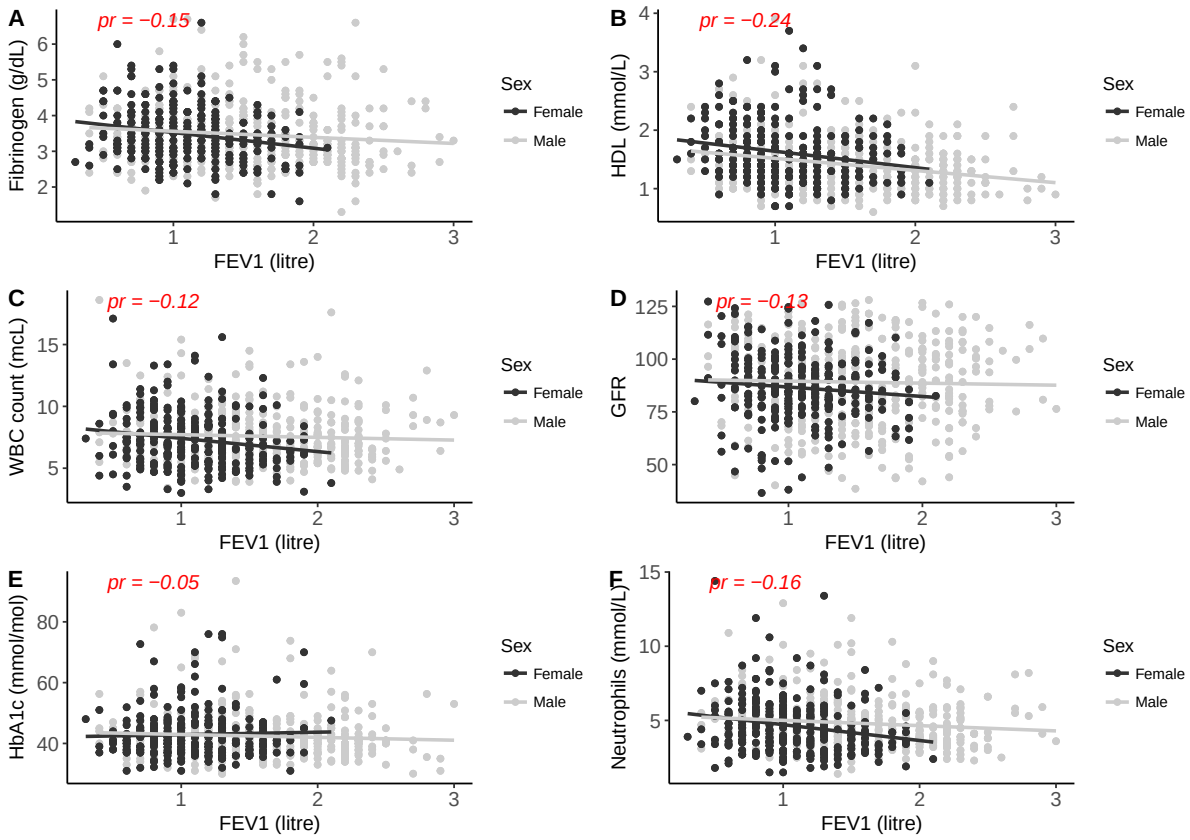


Figure 3.32: Scatter plots displaying the distribution of (A) fibrinogen, (B) high-density lipoprotein (HDL) cholesterol, (C) white cell count (WCC), (D) glomerular filtration rate (GFR), (E) glycated haemoglobin (HbA1c), and (F) neutrophils over forced expiratory volume in one second (FEV_1), by sex. Partial correlations (pr) are displayed in the top left corners.

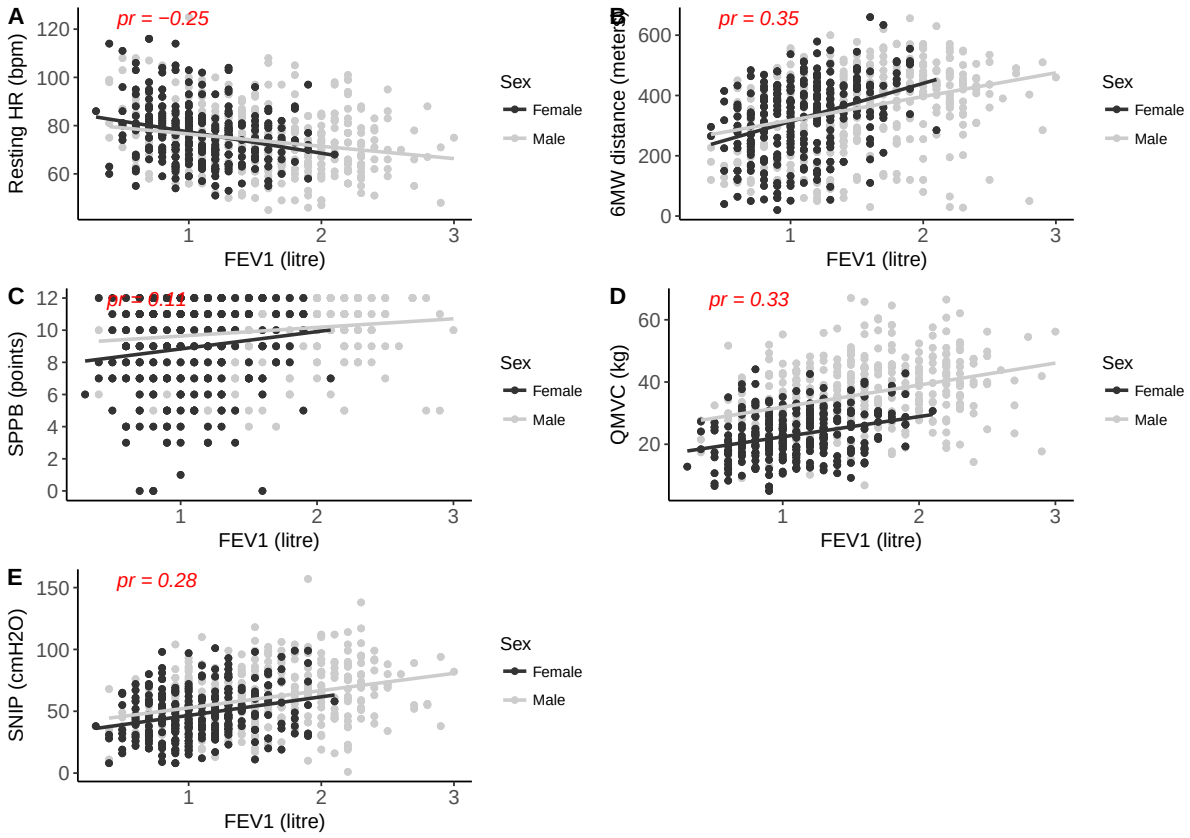


Figure 3.33: Scatter plots displaying the distribution of (A) resting heart rate (HR), (B) six-minute walk (6MW) distance, (C) short physical performance battery (SPPB), (D) quadriceps maximum voluntary contraction (QMVC), and (E) sniff nasal inspiratory pressure (SNIP) over forced expiratory volume in one second (FEV_1), by sex. Partial correlations (pr) are displayed in the top left corners.

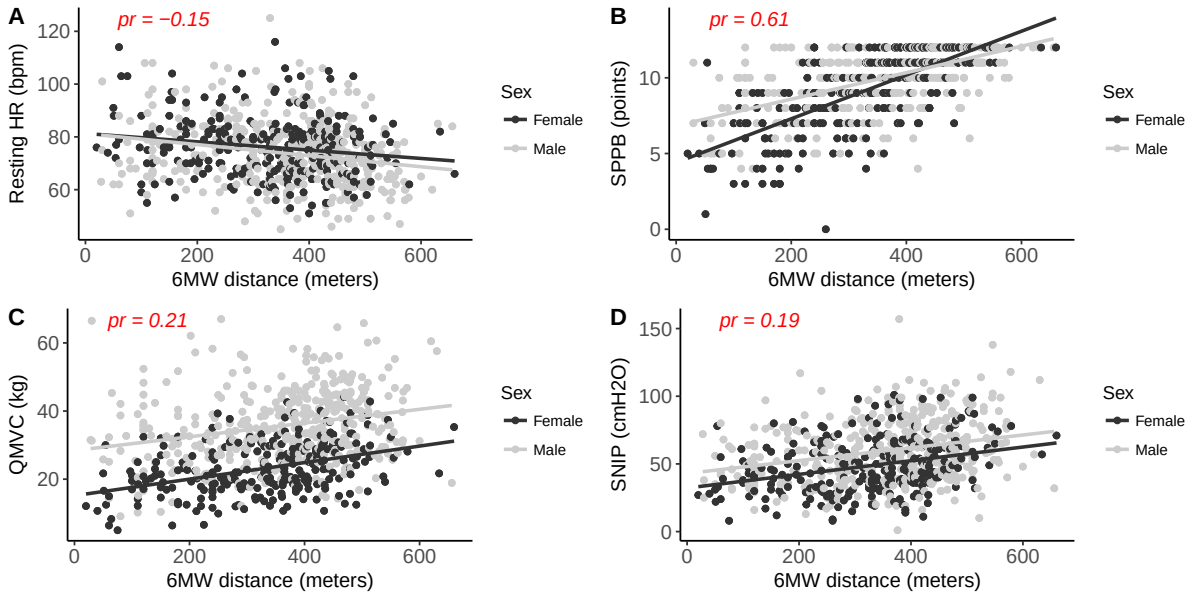


Figure 3.34: Scatter plots displaying the distribution of (A) resting heart rate (HR), (B) short physical performance battery (SPPB), (C) quadriceps maximum voluntary contraction (QMVC), and (D) sniff nasal inspiratory pressure over six-minute walk distance, by sex. Partial correlations (pr) are displayed in the top left corners.

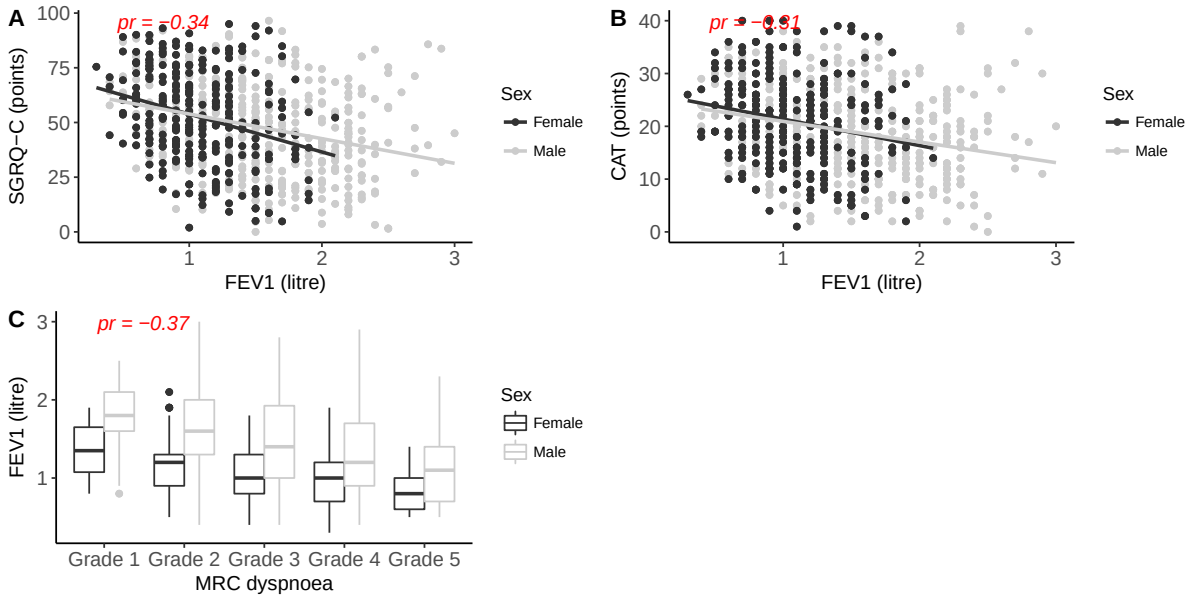


Figure 3.35: Scatter plots displaying the distribution of (A) St. George respiratory questionnaire for COPD (SGRQ-C), (B) COPD assessment test (CAT), and (C) Medical Research Council (MRC) dyspnoea score over forced expiratory volume in one-second (FEV_1), by sex. Partial correlations (pr) are displayed in the top left corners.

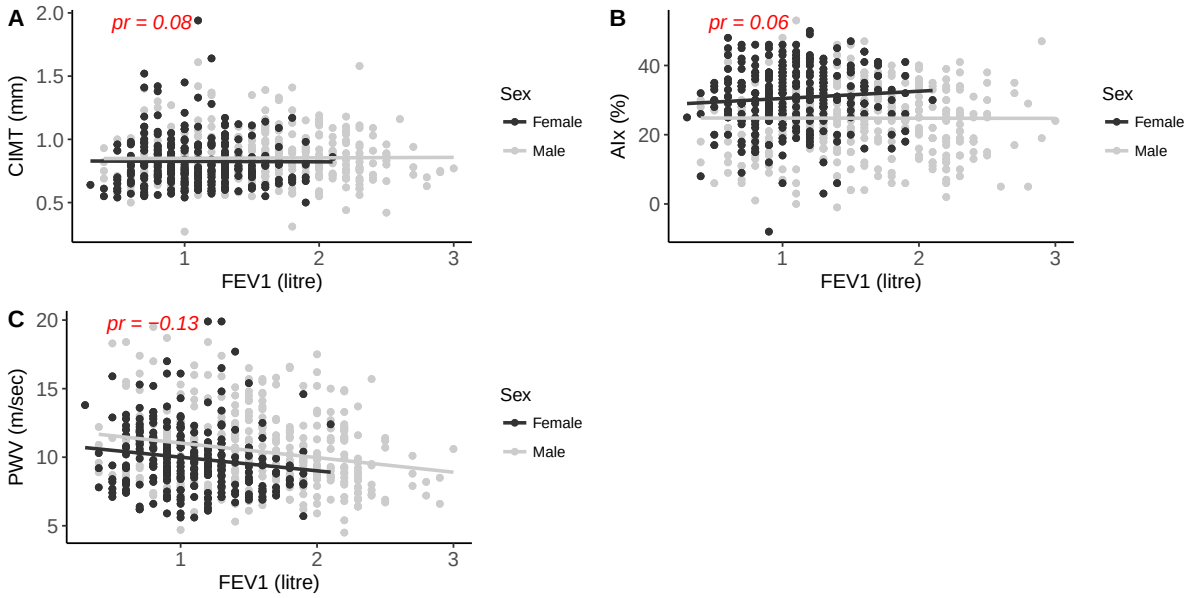


Figure 3.36: Scatter plots displaying the distribution of (A) carotid intima-media thickness (CMIT), (B) augmentation index (AIx), and (C) pulse wave velocity (PWV) over forced expiratory volume in one-second (FEV_1), by sex. Partial correlations (pr) are displayed in the top left corners.

followed by cancer (24%) with 12% a cardiac cause of death. Pulmonary-related cause of death increased with advancing disease severity (i.e. GOLD; **Figure 3.39**, page 90). Most deaths occurred amongst those aged 65-74 years (**Figure 3.40**, page 91).

Most hospital ICD-10 (10th revision of the international statistical classification of diseases and related health problems) diagnoses were recorded at the secondary position during hospital admission (**Figure 3.41**, page 91). Diseases of the respiratory and circulatory system were amongst the most common diagnosis, with "chronic lower respiratory diseases" (J40-J47) and "ischaemic heart diseases" (I20-I25) as one of the most frequently reported primary diagnosis (**Figure 3.42**, page 92).

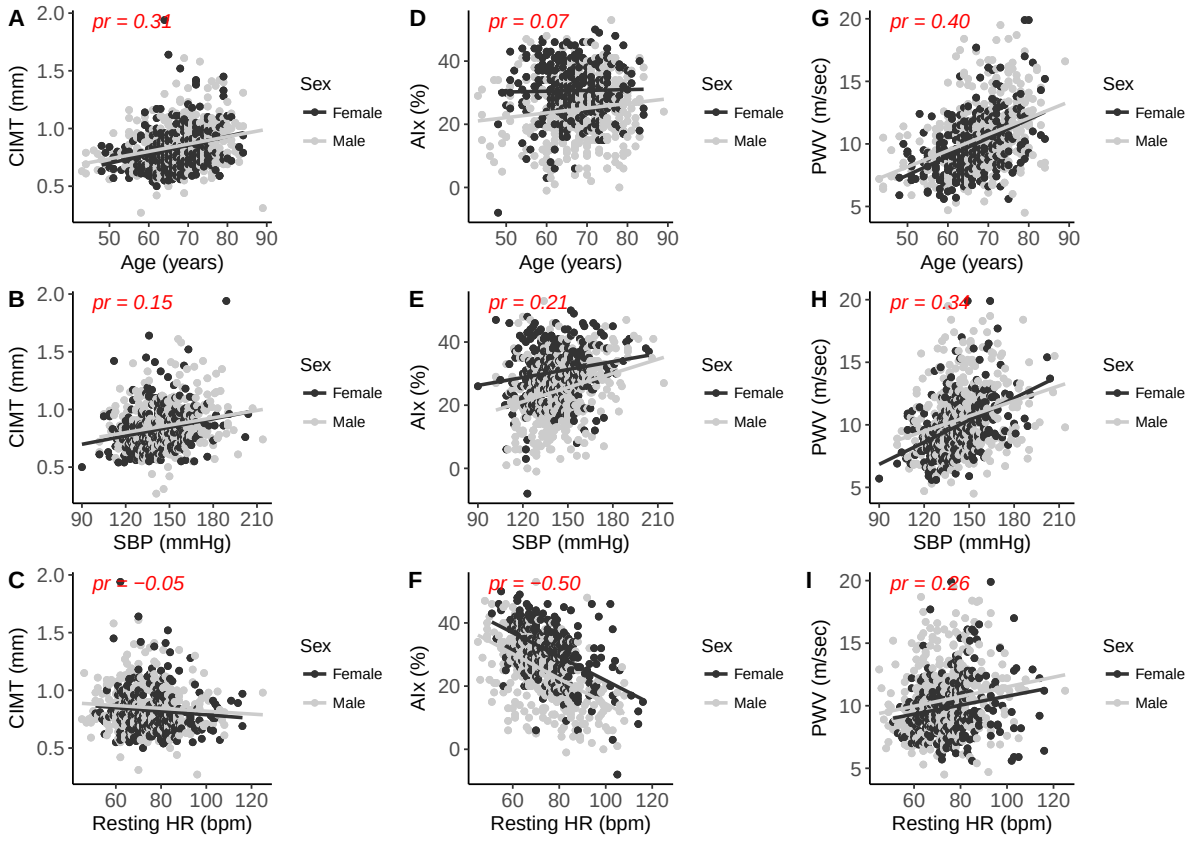


Figure 3.37: Scatter plots displaying the distribution of age, systolic blood pressure, and resting heart rate (HR) under (A-C) carotid intima-media thickness (CIMT), (D-F) augmentation index (AIx), and (G-I) pulse wave velocity (PWV), by sex. Partial correlations (pr) are displayed in the top left corners.

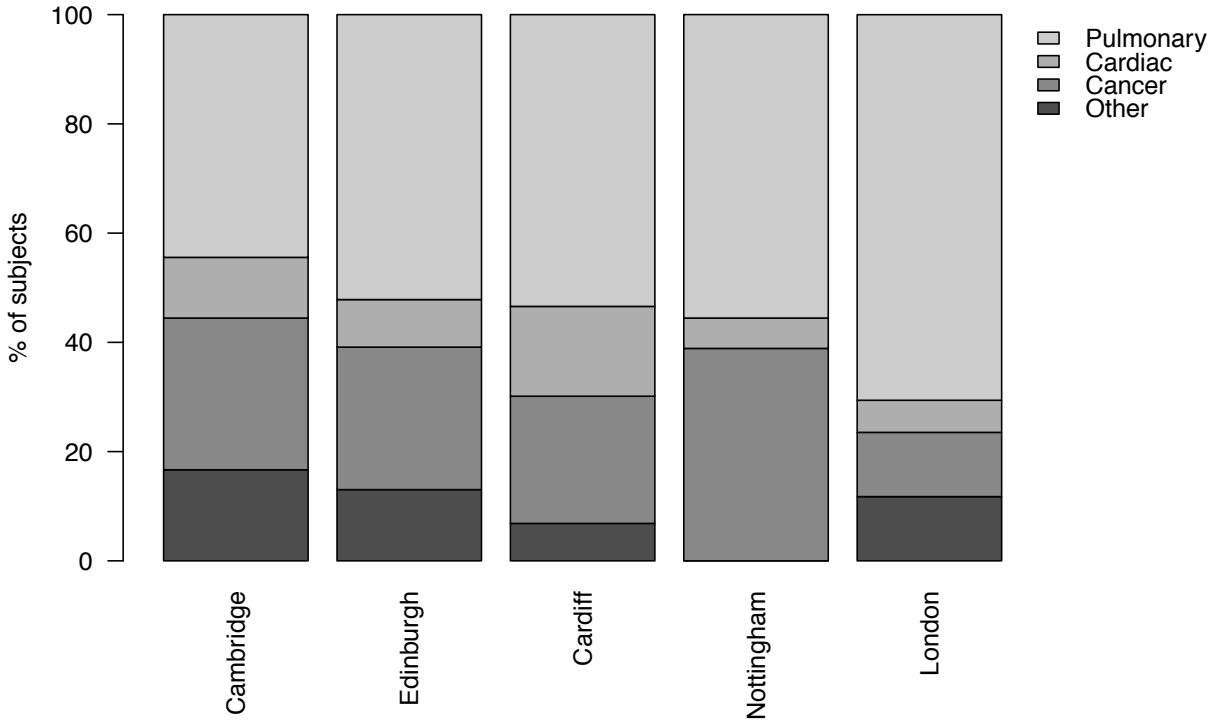


Figure 3.38: Cause of death (n = 149) during study period, by recruitment site. Mortality data were obtained from the Office for National Statistics. Deaths were categorised by cardiac and respiratory physicians.

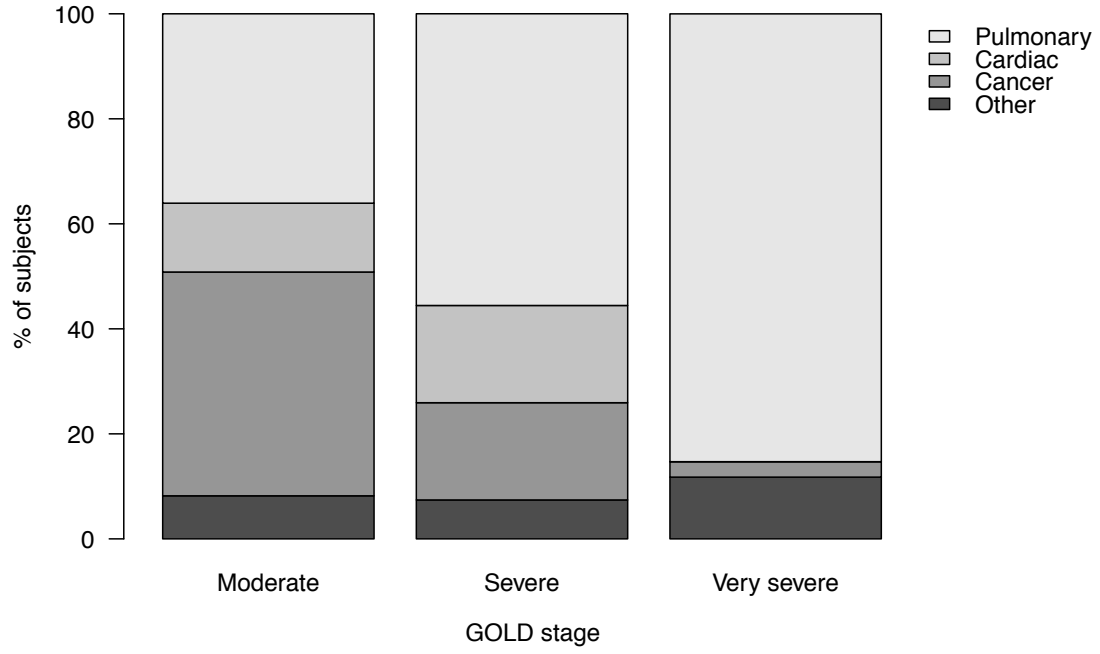


Figure 3.39: Deaths in the ERICA cohort, by cause and GOLD stage.

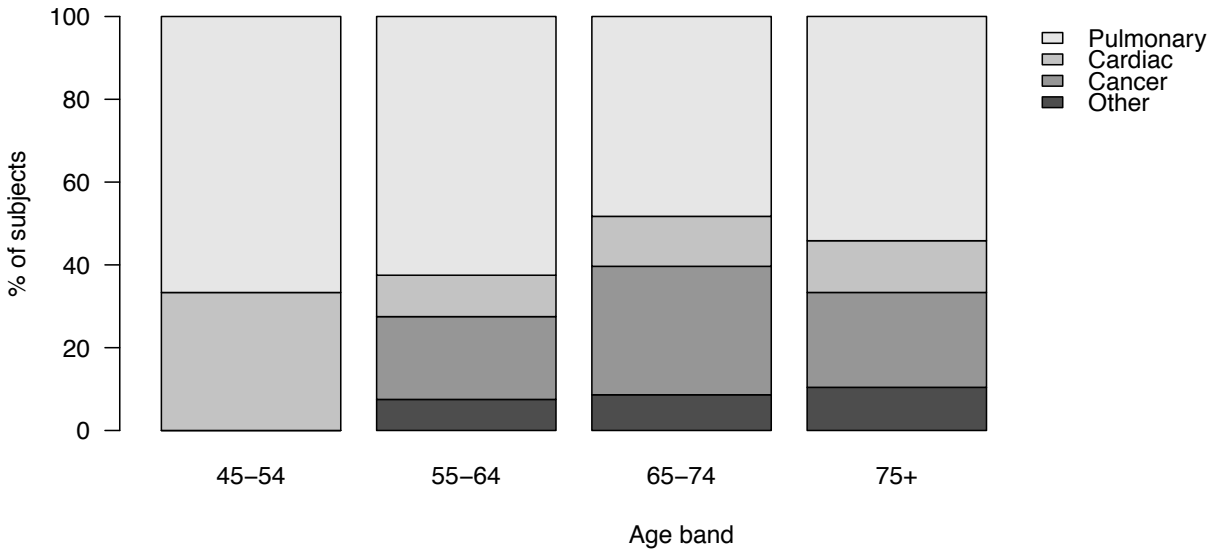


Figure 3.40: Deaths in the ERICA cohort, by age band.

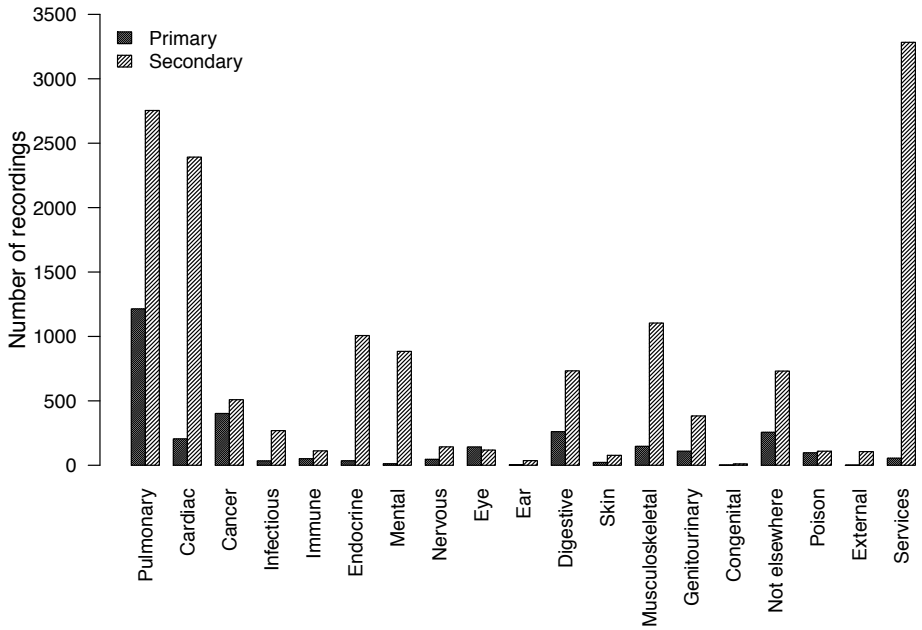


Figure 3.41: Hospital admissions extracted from electronic health record data using ICD-10 chapter coding, by primary and secondary position.

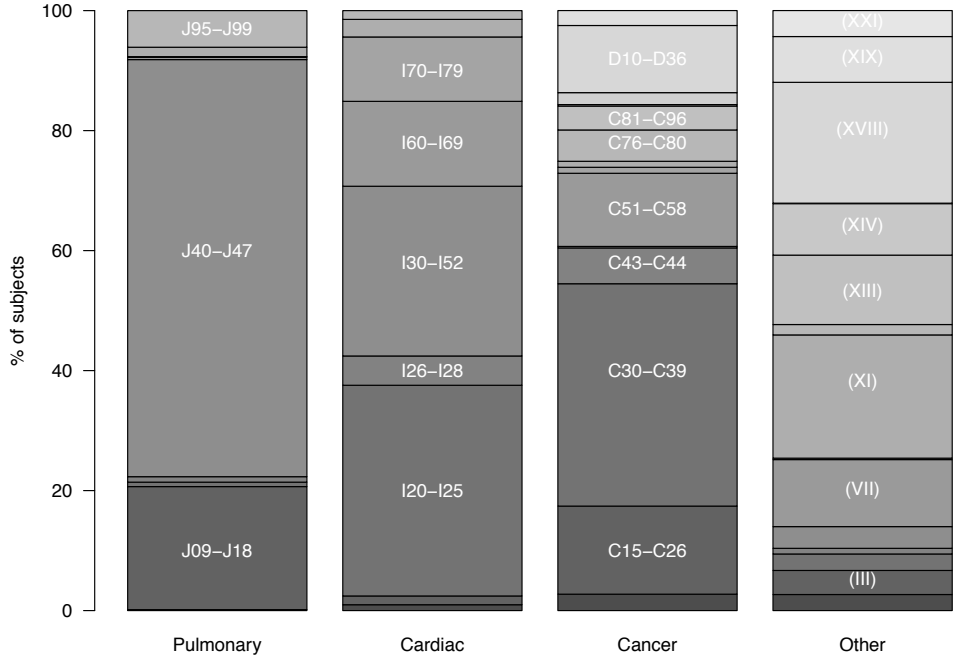


Figure 3.42: Primary hospital admissions extracted from electronic health record data, by ICD-10 chapter coding. J95-J99, Other diseases of the respiratory system. J40-J47, Chronic lower respiratory diseases. J09-J18, Influenza and pneumonia. I70-I79, Diseases of arteries, arterioles and capillaries. I60-I69, Cerebrovascular diseases. I30-I52, Other forms of heart disease. I26-I28, Pulmonary heart disease and diseases of pulmonary circulation. I20-I25, Ischaemic heart diseases. D10-D36, Benign neoplasms. C81-C96, Malignant neoplasms, stated or presumed to be primary, of lymphoid, haematopoietic and related tissue. C76-C80, Malignant neoplasms of ill-defined, secondary and unspecified sites. C51-C58, Malignant neoplasms of female genital organs. C43-C44, Melanoma and other malignant neoplasms of skin. C30-C39, Malignant neoplasms of respiratory and intrathoracic organs. C15-C26, Malignant neoplasms of digestive organs. XXI, Factors influencing health status and contact with health services. XIX, Injury, poisoning and certain other consequences of external causes. XVIII, Congenital malformations, deformations and chromosomal abnormalities. XIV, Diseases of the genitourinary system. XIII, Diseases of the musculoskeletal system and connective tissue. XI, Diseases of the digestive system. VII, Diseases of the eye and adnexa. III, Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism.

3.4 Discussion

This chapter describes the ERICA cohort including selection criteria, definitions used, outcomes measures captured, presence of missing values, and differences in key variables between sexes and recruitment sites, and the association between baseline characteristics. In addition, clinical outcomes are briefly described. Most individuals had GOLD stage II, indicating moderate lung disease. Noticeable were the differences between sexes. The cohort consisted predominantly of men, and women were about two years younger. Compared to men, women had higher number of previous exacerbations, cholesterol levels, resting heart rate and AIx, and were more symptomatic, measured by CAT, SGRQ-C, and MRC. Women had lower levels of neutrophils and haemoglobin, shorter walk distance, lower SPPB, QMVC, and SNIP scores, and CIMT and PWV scores. There were no differences in FEV_1%, BMI, and inflammatory markers between sexes. There was a strong correlation between 6MW distance and chair stand component of the SPPB, which requires quadriceps muscle strength, yet correlation between 6MW and QMVC was weak. Overall, there was a moderate correlation between measures of function and symptoms, and FEV_1. Correlations between spirometry and blood tests were weak. The SPPB scores were skewed negative with a longer left tail, in particular for men. Especially for the balance component, many individuals scored the highest possible score, indicating that the balance component may not be sensitive enough to capture performance differences in this population.

The study was originally designed and powered on the basis of a tertile analysis of variables PWV and QMVC, based on an estimated sample size of 800 individuals with COPD.[184] However, not all individuals could be included due to study ineligibility. In addition, the NHS and ONS did not follow (i.e. flag) all individuals for survival status and hospital admission, further reducing the sample size for analysis and therefore also statistical power. The overall sample size was relatively small, limiting e.g. subgroup analysis.

Patient questionnaires included self-reported questions such as history of exacerbations and use of steroids and CV drugs. Recall bias is common when collecting self-reported data with

increasing bias as time passes. For example, Frei *et al.* reported that patients were unable to accurately recall the number of COPD exacerbations.[96] The authors assessed >400 patients over a 6-month period and found inaccuracies in reporting. Quint *et al.* reported on the other hand that using daily diary cards can reliably recall the number of exacerbations in the first year.[227] No diary cards were used in the ERICA study.

In terms of generalisability, individuals were recruited throughout the UK. For each centre, baseline measures of ten volunteers were compared with those of the other sites.[184] Inter- and intra-user reliability was assessed through intra-class correlation coefficients. All sites required individual site training as well as centralised training in addition to standardisation visits. Only when sites were considered competent, sites could start with the recruitment. In addition to recruitment pamphlets, which might have led to self-selection bias, potential study participants were identified by local principle investigators and clinical care teams. Also, individuals with systemic inflammatory diseases like rheumatoid arthritis and inflammatory bowel disease were excluded and may therefore limit generalisability.

Assessments were anticipated to take place over two visits but had to be completed within three months. Some individuals had already undergone assessments in previous studies within the consortium such as either ARCADE, the ECLIPSE Extension study, PROACTIVE, MRC WP4 Consortium or Skeletal Muscle dysfunction study but these measures were only used if captured within the last three months of recruitment. More importantly, post-bronchodilator spirometry, one of the selection criteria, had to be performed before or on the first visit. Spirometry is the gold standard to diagnose COPD.

Despite the training and standardisation there was variation between recruitment sites for most variables. Populations at the different sites were not comparable on several measures. The centre in London had slightly healthier individuals, whereas individuals from Cardiff had more severe disease. For example, individuals from Cardiff scored lower on the musculoskeletal measures, had worse arterial stiffness and higher COPD impact scores compared to the other sites. Although, according to the ONS life expectancy is the worst in Scotland and Wales.[207]

There was, however, some evidence that data such as the 6MW distance were not missing completely at random. A possible cause could be that measurements such as QMVC and 6MW were either novel to most sites or required verbal encouragement and instructions. Along with PWV and CIMT, these measures were considered most likely to result variability between sites due to differences in equipment and measurements (i.e. expertise). Also, missing values for the 6MW could have resulted from the contraindications for conducting a 6MW test. Myocardial infarction during the previous month and unstable angina were absolute contraindications. A resting heart rate >120 bpm, a systolic blood pressure >180 mmHg, or a diastolic blood pressure >100 mmHg were relative contraindications.

The mean age in the ERICA cohort (67 years) was well above the usual age of COPD diagnosis and comparable to other cohorts including the Investigational Study of Psychological Intervention in Recipients of Lung Transplant (INSPIRE)-II cohort (66 years) and Body mass index, airflow Obstruction, Dyspnoea and Exercise (BODE) cohort (67 years) with the exception of the ECLIPSE study, one of the largest COPD cohorts. ERICA participants were slightly older compared to ECLIPSE participants (64 years). In terms of disease severity, ERICA participants were comparable to the other cohorts.

In contrast to many other cohort studies,[151,158] all individuals were clinically stable. The clinical pathway differs between stable and unstable COPD. The focus of disease management for individuals with stable COPD is on health education, disease coping strategy, treatment regulation, and prevention. Disease management for individuals with unstable COPD, however, is primarily focussed on medical intervention including non-invasive positive pressure ventilation and antibiotic treatment. When assessing the predictive performance of novel biomarkers, aimed at identifying high risk individuals in an early stage of disease allowing for preventative intervention such as exercise training, it is crucial to assess biomarkers that are clinically stable.

Most deaths were related to pulmonary disease followed by cancer, of which a large proportion had hospital visits (primary visits) related to chronic lower respiratory disease and lung cancer, as expected within a COPD population.

The ERICA study was designed as a multi-centre observational, non-interventional, epidemiological cohort study with prospective data collection. Strengths of observational cohort studies, in particular prospective ones, include the ability of evaluating the association between disease or baseline measures and multiple outcomes (e.g. acute exacerbation of COPD, hospital admission, and mortality) and allows for the calculation of disease rates. One of the disadvantages of prospective cohort studies is the associated cost. For example, follow-up data were collected every six months via postal or telephone questionnaire but for cost reasons there were no repeated measurements of CV and musculoskeletal measures taken, limiting the assessment of the predictive value of these biomarkers at different time points. Attrition (i.e. systematic loss to follow-up) is a potential bias in observational studies. However, the use of electronic health records prevented this, as the NHS collects every hospital visit, and the ONS reports any deaths. Linked electronic health record data therefore is one of the major strengths of this dissertation. It not only avoids recall bias, which is often the case with self-reported data, but also prevents lost to follow-up and therefore improves the reliability of findings.

In conclusion, the ERICA cohort is unique in terms of the density of the data captured in a COPD population. Other and larger cohort studies, such as ECLIPSE[270] and NHANES (National Health and Nutritional Examination Survey)-III,[168] do exist but none have captured musculoskeletal and CV function measures in addition to lung performance, biochemical markers, medical history, et cetera. ERICA is the first prospective observational cohort study that has enabled the examination of the relationship between inflammatory markers, musculoskeletal and CV measures within a COPD population, and evaluate their association with common clinical outcomes.

4

Can simpler measures substitute for the six-minute walk component of the BODE Index in predicting death in COPD in the ERICA cohort?

Chapter summary

Background The BODE (Body mass index, airflow Obstruction, Dyspnoea and Exercise) Index predicts mortality in chronic obstructive pulmonary disease but includes a six-minute walk component; the test would be simpler if the six-minute walk was replaced.

Objective We investigated whether a modified BODE Index in which six-minute walk was replaced by alternative measures of skeletal muscle performance, the short physical performance battery or components (i.e. four-metre gait speed, balance, and chair stand), quadriceps and nasal inspiratory muscle strength, retained predictive ability.

Methods We analysed 630 individuals with stable chronic obstructive pulmonary disease at

baseline from the Evaluation of the Role of Inflammation in Chronic Airways disease (ERICA) cohort study and in whom UK Office for National Statistics verified mortality data was available. Variables tested at baseline included short physical performance battery, quadriceps maximum voluntary contraction, and sniff nasal inspiratory pressure. Predictive models were developed using stratified multivariable Cox regression, and assessed by C-indices and calibration plots with 10-fold cross-validation and replication.

Findings During median two years of follow up, 60 (10%) individuals died. There was no significant difference between the discriminative ability of BODE based on six-minute walk (C-index 0.709, 95% CI, 0.680 to 0.737), BODE based on short physical performance battery (C-index 0.683, 95% CI, 0.647 to 0.712), BODE based on four-metre gait speed (C-index 0.676, 95% CI, 0.643 to 0.700), BODE based on balance (C-index 0.686, 95% CI, 0.651 to 0.713), and BODE based on nasal inspiratory muscle strength (C-index 0.676 (95% CI, 0.637 to 0.703) when predicting mortality. Quadriceps muscle voluntary contraction was not able to substitute the six-minute walk in the BODE Index without loss in discriminative ability.

Conclusion Short physical performance battery, its four-metre gait speed and balance components, and sniff nasal inspiratory pressure have the potential to replace the six-minute walk in the BODE Index without significant loss of predictive ability.

4.1 Background

4.1.1 Introduction

Chronic obstructive pulmonary disease (COPD) was responsible for 3.2 million deaths globally in 2015.[106] Accurate assessment of prognosis enables clinicians to focus their resources on their most vulnerable patients, to decide on safety for interventional procedures and may also be useful for stratification of entry into clinical trials. The best known prognostic index is the BODE Index, which generates a composite score from the Body mass index, airflow Obstruction, Dyspnoea, and Exercise capacity, with the latter measured by the six-minute walk (6MW) test

(**Table 4.1**, page 100),[40] which even on its own has predictive ability for mortality.[248] However, both BODE and the 6MW have received limited adoption in clinical practice, perhaps, in part, because the 6MW requires a minimum thirty-metre corridor, and with the necessity for a training walk with rest in between,[127] which may in practice take over 30 minutes. It is unsurprising therefore that NICE (National Institute for Health and Care Excellence) UK 2018 guidelines for COPD included a recommendation not to use the BODE Index for prognosis in COPD, as some components are time-consuming and not routinely available in primary care; of the four components of BODE, the 6MW test is the one that is most unavailable in primary care settings.[200]

Thus, a test that has sufficient predictive ability to replace the 6MW, and which is more clinically practical might facilitate uptake of prognostic scoring. In the original BODE study the authors considered individual measures which were known to have prognostic value and using the same approach, having reviewed recent literature, we identified other tests reflective of physical function. These were the short physical performance battery (SPPB), which is prognostic for mortality in older individuals in the general population,[157,274] and its components (i.e. four-metre gait speed (4MGS), balance, and chair stand); quadriceps strength[259] measured as maximum voluntary contraction force (QMVC) and maximal sniff nasal inspiratory pressure (SNIP),[188] as potential alternative measures of musculoskeletal function.

The aim of our analysis was to first evaluate the association between the measures of skeletal muscle function and all-cause mortality in stable COPD patients, and with the assumption that a relationship would be found to investigate whether a BODE Index in which the 6MW component ($BODE_{6MW}$) was replaced by alternative musculoskeletal measures retained predictive ability. Finally, we aimed to assess whether the prediction of all-cause mortality using BODE Index can be improved by the addition (as opposed to substitution) of these measures to the standard BODE.

Table 4.1: Multidimensional risk factors, published prediction models for the prediction of mortality in COPD. Placed in order of publication date.

	BODE[40]	HADO[87]	mBODE[58]	CPI[28]	DOMI BOX[150]	ADO[225]	U-BODE[225]	BODEx[246]	eBODE[246]	PILE[175]	ECLIPSE[42]	BODE-A[252]
Age				✓		✓					✓	
Sex				✓								
BMI	✓		✓	✓	✓		✓	✓	✓		✓	✓
Questionnaires (MRC, ATS, Fletcher, CRQ)	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
FEV_1 %	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
6MW distance	✓				✓		✓		✓		✓	✓
Exacerbation				✓	✓			✓	✓			
Exercise max. O_2 consumption			✓									
O_2-use			✓									
CVD				✓								
Blood Oxygen (PaO_2)				✓								
Health status		✓										
Activity		✓										
QMVC										✓		
Inflammatory markers (e.g. IL-6 and fibrinogen)										✓	✓	✓
Total sample size	625	611	444	8802	68	232	232	185	185	268	1843	549
Observed deaths (total No.)	162	94	206	166	22	79	79	71	71	83	168	26
Follow-up (months)	28	36	22	6-36	36	30	30	36	36	73	36	12
C-statistic	0.74	0.68	0.72	0.71	N/A	0.63	0.61	0.74	0.77	0.71	0.73	0.75

Abbreviations: BMI, body mass index. MRC, Medical Research Council. CRQ, Chronic Respiratory Questionnaire. ATS, American Thoracic Society. FEV_1, forced expiratory volume in one second. 6MW, six-minute walk. SGRQ, St.George Respiratory Questionnaire. CVD, cardiovascular disease. PaO_2, partial pressure of oxygen. QMVC, quadriceps maximum voluntary contraction. IL-6, interleukin 6.

4.2 Methods

4.2.1 Study design and participants

The ERICA study is a multi-centre observational, non-interventional, epidemiological cohort study, with a sample size of 729 stable global initiative for chronic obstructive lung disease (GOLD) stage II-IV[107] COPD patients.[184] A baseline assessment was undertaken between December 2011 and January 2014, with planned interval mortality obtained from the UK Office for National Statistics (ONS) last updated for this study in November 2017. Analyses were limited to three years of follow-up (August 2016).

4.2.2 Point assignment for components of BODE Index

Points for BODE Index were assigned and classified in quartiles as described by Celli *et al.*, with higher scores indicating a higher risk of mortality.[40] The SPPB has a range of 0-12 points and comprises three subtests scored 0-4. To preserve a four-category system, we combined 1 and 2 points. The SPPB itself was divided into 10-12, 7-9, 4-6 and <4 based on the cut-off score of <10 to define functional limitation,[20,216] and the distribution of the data. In the absence of a generally accepted categorisation system, SNIP and QMVC were divided into quartiles (**Table 4.2**, page 102).

4.2.3 Statistical analysis

Hazard ratios (HR) were estimated using multivariable Cox regression, stratified by recruitment centre, and adjusted for age and sex. Further analyses included body mass index (BMI), forced expiratory volume in one second (FEV_1), smoking status, and Medical Research Council (MRC) dyspnoea. Multiple predictive models were developed according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) and guidelines for clinical prediction models.[52,232,251] The pre-selected prediction models were: (i) BMI, (ii) BMI + MRC dyspnoea, (iii) BMI + MRC dyspnoea + FEV_1%, (iv) BMI + MRC dyspnoea

Table 4.2: Assignment of points.

Variable	**0 points**	**1 point**	**2 points**	**3 points**
BODE				
BMI (kg/m^2)	>21	≤21		
FEV_1 (% predicted)	≥65	50-64	36-49	≤35
Dyspnoea (MRC scale)	0-1	2	3	4
Six-minute walk distance (m)	≥350	250-349	150-249	≤149
Alternative musculoskeletal measures				
SPPB (points)	10-12	7-9	4-6	<4
Four-metre gait speed (points)	4	3	1-2	0
Balance (points)	4	3	1-2	0
Chair stand (points)	4	3	1-2	0
SNIP (cm H_2O)	≥71	54-70	39-53	≤38

Abbreviations: BMI, body-mass index. FEV_1, forced expiratory volume in one second. MRC, Medical Research Council. SPPB, short physical performance battery. SNIP, sniff nasal inspiratory pressure.

+ FEV_1% + 6MW, (v) BMI + MRC dyspnoea + FEV_1% + SPPB, and (vi) BMI + MRC dyspnoea + FEV_1% + SNIP. Linearity of continuous predictors was assessed visually. We tested for violation of the proportional hazards assumption by including time interactions and visually examining Arjas plots. Discrimination (i.e. Harrell′s C-statistic[10,202] and calibration (i.e. Hosmer-Lemeshow test[9] and calibration plots) were assessed using 10-fold cross validation with 200 replications.[245] Effect of missing data was assessed in sensitivity analyses using multivariable imputation by chained equations (MICE). Predictive mean matching was used for continuous variables, ordered logistic regression (as continuous) for ordinal variables, multinomial logistic regression for categorical variables, and logistic regression for binary variables. Derived variables such as SPPB (a composite score of 4MGS, balance, and chair stand) were estimated post MICE using passive imputation. To minimise potential overfitting caused from using the same imputed dataset for the training and test data for 10-fold cross-validation, we created 10 x 2 imputed datasets and used ten for model derivation and the other ten for model validation. Thus, within each step of the cross-validation, the training and test datasets were from two different imputations. We performed cross-validation separately for the ten pairs of imputed datasets before combining estimates of interest using Rubin′s rules. Observational data is reported according to the Strengthening The Reporting of OBservational Studies in Epidemiology

(STROBE) statement.[275]

4.3 Findings

4.3.1 Descriptive statistics

In total, 714 individuals were followed by the ONS for survival status, of which 630 had complete baseline data and were included in the primary analysis (**Figure 3.4**, page 64). Of the cohort, 386 (61%) were male, 192 (30%) were current smokers, 358 (57%) were identified as GOLD stage II, and the median baseline age was 67 years (range 43-84 years; **Tables 3.1 and 4.3**, pages 65 and 104). In total, 245 (39%) had defined functional limitation (SPPB score <10) with a median (IQR) SPPB score of 10 (8-12), a median 6MW distance of 370 (268-440) metres, a median QMVC of 30 (22-39) kg, and a median maximal SNIP of 53 (38-70) cm H_2O.

4.3.2 Factors associated with all-cause mortality

The three years survival probability was 90% (88-93% CI) with an event rate of 3.3 (95% CI 2.6 to 4.3) per 100 person-years. Event rates differed by recruitment site: 2.4 (1.0 to 5.8) for Cambridge, 3.6 (1.9 to 7.0) for Edinburgh, 2.8 (1.9 to 4.1) for Cardiff, 3.3 (1.7 to 6.6) for Nottingham, and 7.1 (3.9 to 12.8) for London. In total 60 patients (10%) died after study enrolment, with a median follow-up time of two years. Age-adjusted multivariable analysis identified multiple markers associated with mortality including BMI (HR 0.91 per 1 point increase, 95% CI, 0.86 to 0.97, p = 0.002), 6MW distance (HR 0.85 per 30-metre increase, 95% CI, 0.78 to 0.92, p <0.001), SPPB (HR 0.81 per 1 point increase, 95% CI, 0.72 to 0.92, p = 0.002), 4MGS (HR 0.67 per 1 point increase, 95% CI, 0.49 to 0.93, p = 0.015), balance (HR 0.63 per 1 point increase, 95% CI, 0.48 to 0.82, p = 0.001), and SNIP (HR 0.81 per 10 cm H_2O increase, 95% CI, 0.69 to 0.95, p = 0.010; (**Figures 4.1 4.2** and **Table 4.4**, pages 105-107). Chair stand and QMVC were not associated with all-cause mortality, after further adjustment.

Table 4.3: Baseline characteristics by functional limitation (n = 630).

Characteristic	Total (%)	N (%)	SPPB, ≤9 points	SPPB ≥10 points	P value
Description					
Age (yrs.), median (IQR)	67 (62-73)	714 (100)	70 (63-75)	66 (62-71)	<0.001
Male	386 (61)	714 (100)	129 (53)	257 (67)	<0.001
Body mass index (kg/m^2), median (IQR)	27 (23-31)	707 (99)	28 (24-32)	26 (23-29)	<0.001
Lung function					
FEV_1 %predicted, median (IQR)	53 (40-65)	712 (100)	52 (39-63)	54 (41-66)	0.265
Current smoker	192 (30)	709 (99)	73 (30)	119 (31)	0.767
MRC dyspnoea score		709 (99)			
1	54 (9)	-	7 (3)	47 (12)	<0.001
≥2	576 (91)	-	238 (97)	338 (88)	
GOLD		713 (100)			
Stage II	358 (57)	-	131 (53)	227 (59)	0.357
Stage III	216 (34)	-	89 (36)	127 (33)	
Stage IV	56 (9)	-	25 (10)	31 (8)	
Musculoskeletal measures					
6MW distance (metre), median (IQR)	370 (268-440)	680 (95)	265 (174-344)	420 (360-470)	<0.001
SPPB (0-12), median (IQR)	10 (8-12)	706 (99)			
No functional limitation, ≥10	385 (61)	-			
Functional limitation <10	245 (39)	-			
4MGS score (0-4), median (IQR)	4 (3-4)	709 (99)	3 (3-4)	4 (4-4)	<0.001
Balance points (0-4), median (IQR)	4 (4-4)	711 (100)	4 (3-4)	4 (4-4)	<0.001
Chair stand score (0-4), median (IQR)	3 (1-4)	707 (99)	1 (1-1)	3 (3-4)	<0.001
QMVC peak (kg), median (IQR)	30 (22-39)	687 (96)	25 (19-33)	32 (26-41)	<0.001
SNIP (cm H_2O), median (IQR)	53 (38-70)	688 (96)	44 (32-61)	59 (44-74)	<0.001

Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of 630 individuals are included. P-values estimated using Wilcoxon-Mann-Whitney test for continuous data, and Chi-square test for categorical data. *Abbreviations*: MRC, Medical Research Council. GOLD, global initiative for obstructive lung disease. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction. SNIP, sniff nasal inspiratory pressure.

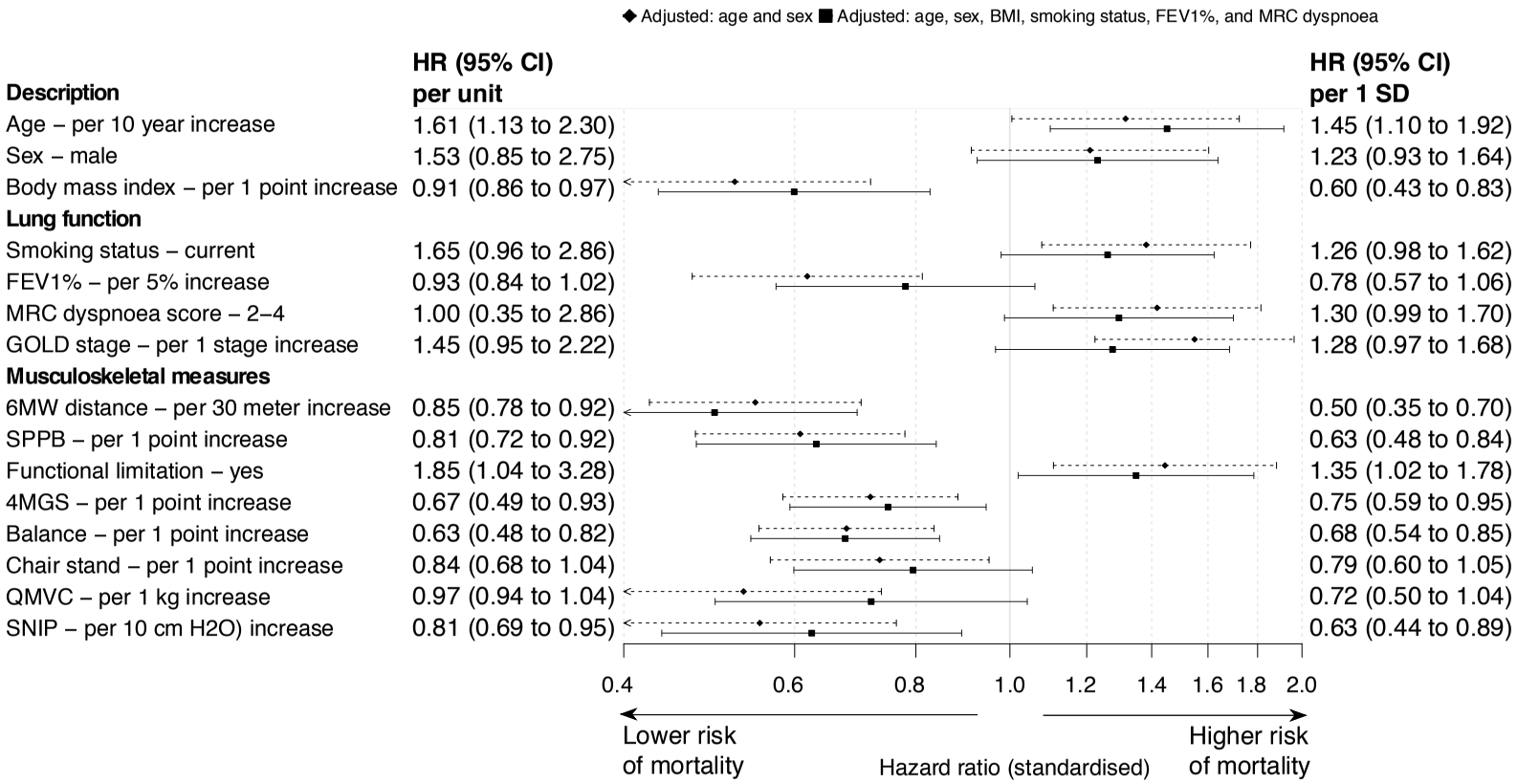


Figure 4.1: Forest plot displaying standardized hazard ratios (HR) for three years of follow-up, by category. The top dashed lines indicate HR for adjustment of age and sex. The second full lines indicate HR further adjustment of BMI, smoking status, $FEV_1\%$, and dyspnoea scores. *Abbreviations*: SD, standard deviation. CI, confidence intervals. BMI, body mass index. $FEV_1\%$, predicted forced expiratory volume one second. MRC, Medical Research Council. GOLD, global initiative for chronic obstructive lung disease. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction. SNIP, sniff nasal inspiratory pressure.

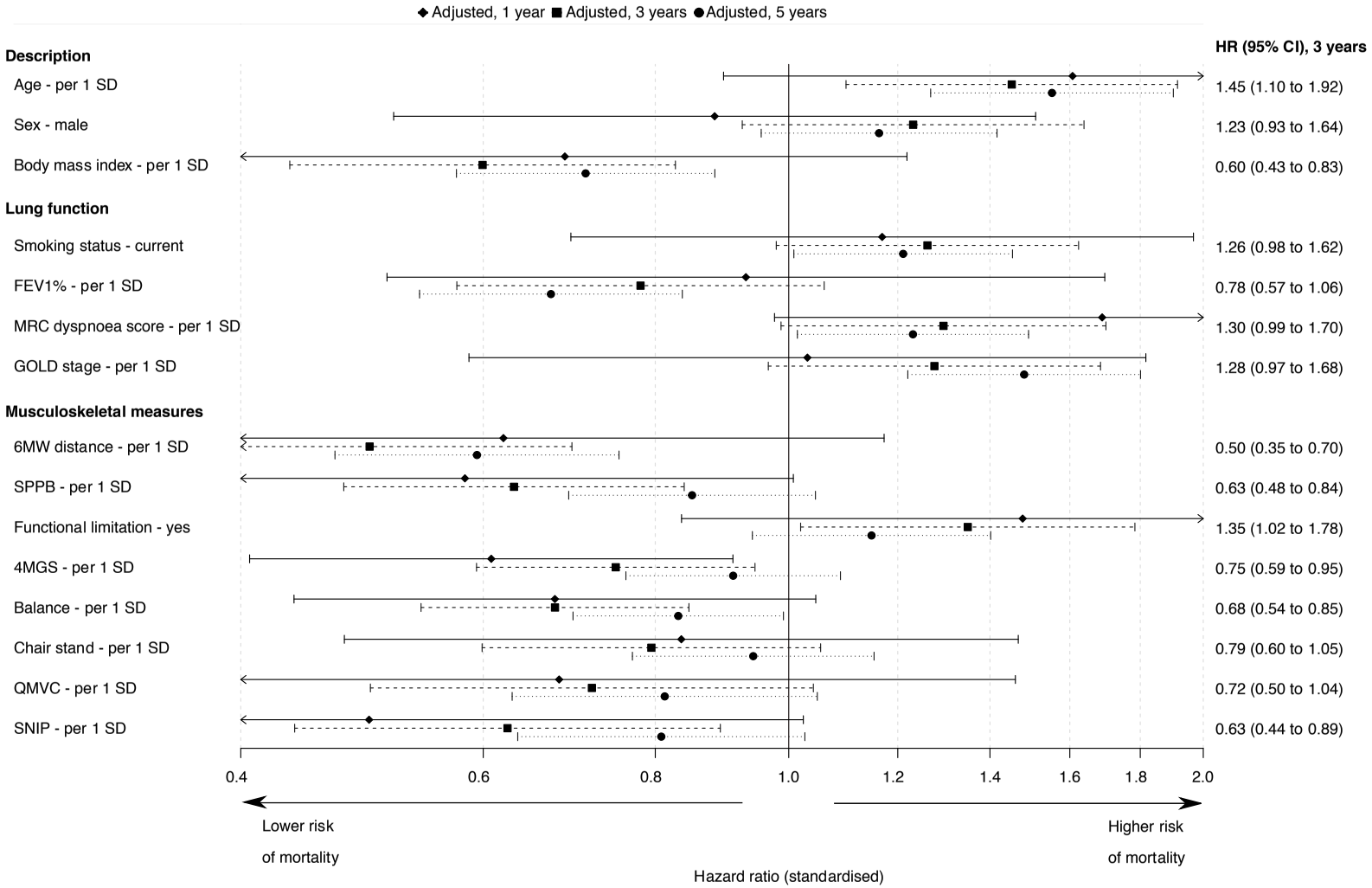


Figure 4.2: Forest plot displaying standardized adjusted hazard ratios, by years of follow-up. Hazard ratios were estimated using Cox regression. All analyses were adjusted for age, sex, body mass index, smoking status, $FEV_1\%$, and MRC dyspnoea score. Hazard ratios displayed are after 3 years of follow-up. Number of deaths was 15 after 1 year, 60 after 3 years, and 121 after 5 years of follow-up. *Abbreviations*: SD, standard deviation. HR, hazard ratio. CI, confidence intervals. BMI, body mass index. $FEV_1\%$, predicted forced expiratory volume one second. MRC, Medical Research Council. GOLD, global initiative for chronic obstructive lung disease. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction. SNIP, sniff nasal inspiratory pressure.

Table 4.4: Adjusted multivariable associations, with the occurrence of death, by years of follow-up (n = 630).

Baseline Characteristics	3 years (n = 60 deaths)				5 years (n = 121 deaths)			
	Hazard ratio (95% CI) [a]	P value [c]	Hazard ratio (95% CI) [b]	P value [c]	Hazard ratio (95% CI) [a]	P value [c]	Hazard ratio (95% CI) [b]	P value [c]
Description								
Age - per 10 year increase	1.42 (1.01 to 2.01)	0.046	1.61 (1.13 to 2.30)	0.008	1.54 (1.20 to 1.97)	0.001	1.76 (1.36 to 2.28)	<0.001
Sex - male	1.48 (0.83 to 2.63)	0.186	1.53 (0.85 to 2.75)	0.154	1.35 (0.91 to 2.02)	0.138	1.36 (0.91 to 2.04)	0.133
Body mass index - per 1 point increase	0.89 (0.84 to 0.94)	<0.001	0.91 (0.86 to 0.97)	0.002	0.92 (0.89 to 0.96)	<0.001	0.94 (0.91 to 0.98)	0.002
Lung function								
Smoking status - current	2.02 (1.18 to 3.46)	0.011	1.65 (0.96 to 2.86)	0.072	1.77 (1.20 to 2.61)	0.004	1.51 (1.02 to 2.25)	0.04
FEV_1 - per 5% increase %predicted	0.86 (0.79 to 0.94)	0.001	0.93 (0.84 to 1.02)	0.114	0.84 (0.79 to 0.89)	<0.001	0.88 (0.82 to 0.95)	<0.001
MRC dyspnoea score - 2-4	1.44 (0.52 to 4.01)	0.487	1.00 (0.35 to 2.86)	0.999	1.93 (0.84 to 4.42)	0.121	1.33 (0.57 to 3.09)	0.507
GOLD stage - per increase to next stage	1.95 (1.36 to 2.80)	<0.001	1.45 (0.95 to 2.22)	0.086	1.83 (1.36 to 2.46)	<0.001	1.83 (1.36 to 2.46)	<0.001
Musculoskeletal measures								
Six-minute walk distance - per 30 metre increase	0.87 (0.82 to 0.92)	<0.001	0.85 (0.78 to 0.92)	<0.001	0.89 (0.85 to 0.92)	<0.001	0.88 (0.84 to 0.94)	<0.001
SPPB score (0-12) - per 1 point increase	0.80 (0.71 to 0.89)	<0.001	0.81 (0.72 to 0.92)	0.002	0.93 (0.85 to 1.02)	0.125	0.93 (0.85 to 1.02)	0.125
Functional limitation (SPPB) - yes	2.13 (1.24 to 3.66)	0.006	1.85 (1.04 to 3.28)	0.036	1.33 (0.88 to 2.00)	0.173	1.33 (0.88 to 2.00)	0.173
4MGS score (0-4) - per point increase	0.63 (0.48 to 0.84)	0.002	0.67 (0.49 to 0.93)	0.015	0.88 (0.69 to 1.13)	0.31	0.88 (0.69 to 1.13)	0.31
Balance score (0-4) - per increase of 1 point	0.63 (0.49 to 0.81)	<0.001	0.63 (0.48 to 0.82)	0.001	0.80 (0.65 to 0.99)	0.04	0.80 (0.65 to 0.99)	0.04
Chair stand score (0-4) - per point increase	0.79 (0.65 to 0.96)	0.02	0.84 (0.68 to 1.04)	0.112	0.96 (0.82 to 1.11)	0.565	0.96 (0.82 to 1.11)	0.565
QMVC peak - per 1 kg increase	0.95 (0.92 to 0.97)	<0.001	0.97 (0.94 to 1.00)	0.082	0.96 (0.94 to 0.98)	0.111	0.98 (0.96 to 1.00)	0.111
SNIP - per 10 cm H_2O increase	0.77 (0.67 to 0.89)	<0.001	0.81 (0.69 to 0.95)	0.01	0.91 (0.82 to 1.01)	0.082	0.91 (0.82 to 1.01)	0.082

Hazard ratios were estimated using Cox regression. All analyses were stratified by recruitment centre. Data after one-year follow-up are not included due to too few events (n = 15). [a] Adjusted for age and sex [b] Adjusted for age, sex, body mass index, smoking status, FEV_1%, and MRC dyspnoea score [c] P values based on Cox regression. *Abbreviations*: CI, confidence intervals. FEV_1, forced expiratory volume in one second. MRC, Medical Research Council. GOLD, global initiative for obstructive lung disease. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction. SNIP, sniff nasal inspiratory pressure.

4.3.3 Predictive models

Predictive modelling indicated slightly higher HR for SPPB and its components compared to $BODE_{6MW}$ (**Figure 4.3**, page 109). The C-statistic was the highest for $BODE_{6MW}$ (C = 0.709, 95% CI, 0.680 to 0.737) but there was no significant difference in discriminative ability compared to $BODE_{SPPB}$ (C = 0.683, 95% CI, 0.647 to 0.712; **Figure 4.3 and Table 4.5**, pages 109 and 112). Neither was there a significant difference in risk discrimination when compared with the $BODE_{4MGS}$ (C = 0.676, 95% CI, 0.643 to 0.700), $BODE_{BALANCE}$ (C = 0.686, 95% CI, 0.651 to 0.713), and the $BODE_{SNIP}$ (C = 0.676, 95% CI, 0.637 to 0.703). When comparing $BODE_{SPPB}$ with its components, there were no significant differences in risk discrimination between indices. Calibration tests and plots of the hazard models indicate good model fit and calibration for 3-year prediction of mortality (**Figure 4.4**, page 111).

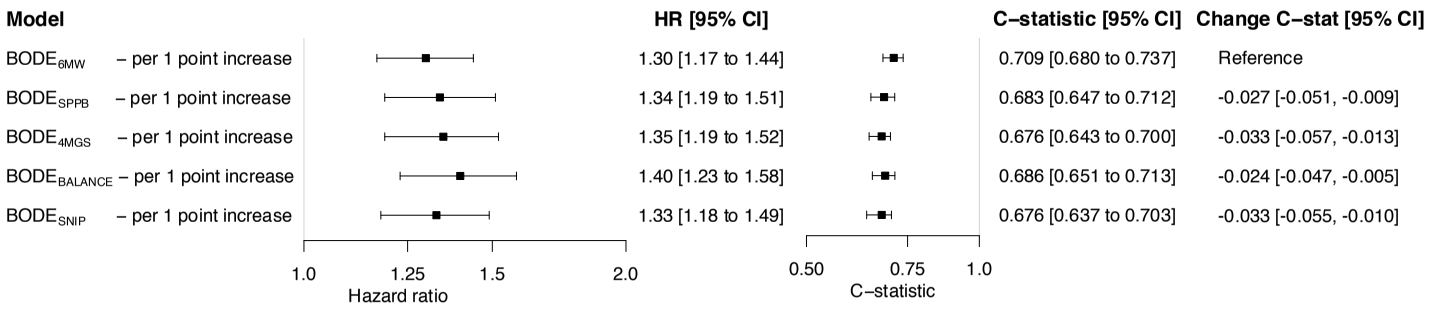


Figure 4.3: Hazard ratios and C-indices with change scores for various BODE models. All models were stratified by recruitment centre. *Abbreviations*: HR, hazard ratio. CI, confidence interval. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. SNIP, sniff nasal inspiratory pressure.

4.3.4 Practical example of using the BODE Index

Estimating an individuals′ BODE score with these indices is similar to the original BODE Index (**Table 4.2**, page 102), with scores ranging between 0 and 10 points (**Tables 4.6 4.7** and **Figure 4.5**, pages 113 and 114). For example, when using the $BODE_{SNIP}$, an individual with a BMI of $\leq$21 (1 point), FEV_1 of 36-49% predicted (2 points), dyspnoea score of four (3 points), and SNIP score of 54-70 cm H_2O (2 points), has a total score of eight points out of ten. The $BODE_{SNIP}$ Index quartile 1 was defined by a score of 0-2, quartile 2 by a score of 3-4, quartile 3 by a score of 5-6, and quartile 4 by a score of 7-10. A $BODE_{SNIP}$ Index score of eight would then indicate a 23% predicted risk of early mortality over the next three years (**Figure 4.4**, page 111). In other words, out of 100 individuals with the same $BODE_{SNIP}$ score, 23 are likely to die within the next three years.

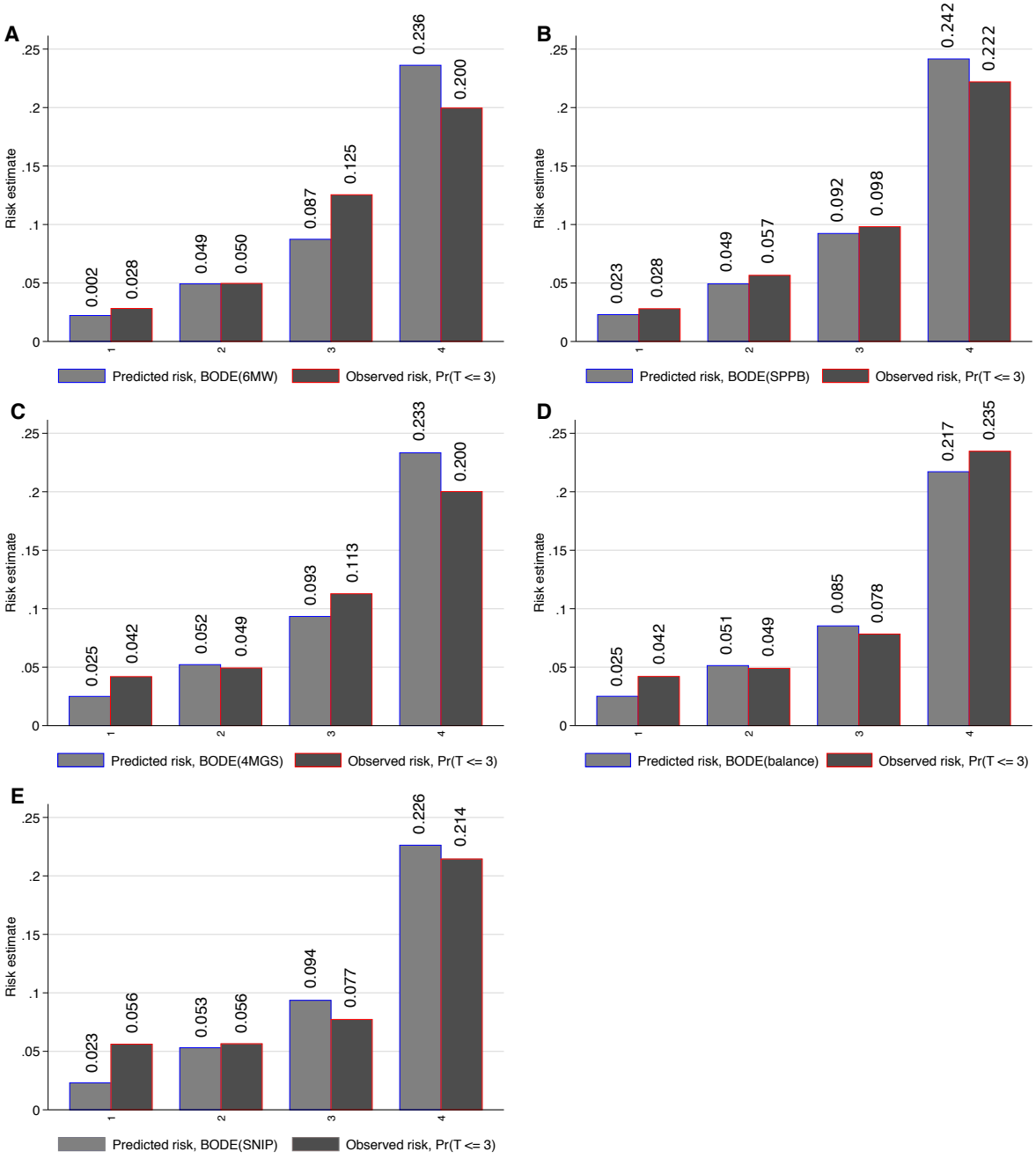


Figure 4.4: Predicted and observed mortality by risk quartiles with associated 95% confidence intervals (n = 630). X-axis indicates four risk groups with predicted mortality ranges. Y-axis indicates 3-year risk for mortality as a probability. Labels indicate predicted and observed risks. (A) BODE_{6MW}, (B) $\text{BODE}_{\text{SPPB}}$, (C) $\text{BODE}_{\text{4MGS}}$, (D) $\text{BODE}_{\text{BALANCE}}$, and (E) $\text{BODE}_{\text{SNIP}}$.

Table 4.5: Cox proportional hazards regression analyses for all-cause mortality, using continuous data.

Variable	Model 1: BMI	Model 2: BMI, MRC	Model 3: BMI, MRC, FEV_1%	Model 4: $BODE_{6MW}$	Model 5: $BODE_{SPPB}$	Model 6: $BODE_{4MGS}$	Model 7: $BODE_{BALANCE}$	Model 8: $BODE_{SNIP}$
Variable	**Hazard Ratio (95% CI)**							
BMI - per 1 point increase	0.90 (0.85 to 0.95)	0.91 (0.86 to 0.95)	0.91 (0.87 to 0.96)	0.89 (0.85 to 0.94)	0.90 (0.86 to 0.95)	0.91 (0.86 to 0.95)	0.91 (0.87 to 0.96)	0.92 (0.87 to 0.97)
MRC dyspnoea score - per 1 point increase		1.38 (1.10 to 1.72)	1.27 (0.99 to 1.63)	0.92 (0.68 to 1.23)	1.05 (0.80 to 1.38)	1.15 (0.89 to 1.49)	1.15 (0.90 to 1.49)	1.21 (0.94 to 1.55)
FEV_1 - per 5% increase %predicted			0.93 (0.85 to 1.02)	0.98 (0.89 to 1.07)	0.93 (0.84 to 1.02)	0.94 (0.85 to 1.03)	0.92 (0.83 to 1.01)	0.95 (0.86 to 1.04)
6MW - per 30 metre increase				0.84 (0.78 to 0.91)				
SPPB (component) - per increase of 1 point					0.80 (0.71 to 0.90)	0.61 (0.45 to 0.84)	0.63 (0.49 to 0.82)	
SNIP - per 10 cm H_2O increase								0.81 (0.69 to 0.94)
C-index	0.634 (0.600 to 0.658)	0.650 (0.620 to 0.676)	0.646 (0.607 to 0.672)	0.709 (0.680 to 0.737)	0.682 (0.647 to 0.712)	0.676 (0.642 to 0.700)	0.685 (0.651 to 0.712)	0.676 (0.637 to 0.703)
Goodness of fit, chi2(3)	10.63	2.57	1.95	2.64	1.57	0.89	4.28	2.67
P >chi2	0.014	0.464	0.583	0.451	0.665	0.827	0.233	0.445
Change in C-statistic	-0.075 (-0.106 to -0.048)	-0.060 (-0.082 to -0.037)	-0.064 (-0.083 to -0.041)	Reference	-0.027 (-0.052 to -0.009)	-0.033 (-0.057 to -0.013)	-0.024 (-0.047 to -0.005)	-0.033 (-0.055 to -0.010)

All models were stratified by recruitment centre. Goodness of fit estimates was based on quartiles of risk. *Abbreviations*: CI, confidence intervals. BMI, body mass index. MRC, Medical Research Council. FEV_1%, predicted forced expiratory volume one second. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. SNIP, sniff nasal inspiratory pressure.

Table 4.6: Risk indices using point system, by survival status.

Risk indices	Median (IQR)	Survivors	Non-survivors	P value ¶
$BODE_{6MW}$ (0-10)	3 (1-5)	3 (1-5)	5 (2-7)	<0.001
$BODE_{SPPB}$ (0-10)	3 (1-4)	2 (1-4)	4 (2-6)	<0.001
$BODE_{4MGS}$ (0-10)	3 (1-4)	2 (1-4)	4 (2-6)	<0.001
$BODE_{BALANCE}$ (0-10)	2 (1-4)	2 (1-4)	4 (2-6)	<0.001
$BODE_{SNIP}$ (0-10)	4 (2-6)	4 (2-5)	5 (4-7)	<0.001

¶Wilcoxon-Mann-Whitney test. *Abbreviations*: 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. SNIP, sniff nasal inspiratory pressure.

Table 4.7: Risk indices using point system, by cause of death.

Risk indices	Pulmonary	Cardiac	Cancer	Other	P value ¶
$BODE_{6MW}$ (0-10)	6 (4-7)	2 (1-7)	2 (1-4)	3 (1-7)	0.004
$BODE_{SPPB}$ (0-10)	5 (3-6)	2 (1-5)	2 (1-4)	4 (2-6)	0.004
$BODE_{4MGS}$ (0-10)	6 (3-6)	2 (1-5)	2 (2-3)	3 (2-6)	<0.001
$BODE_{BALANCE}$ (0-10)	5 (3-6)	4 (1-6)	3 (1-3)	4 (2-5)	0.007
$BODE_{SNIP}$ (0-10)	6 (4-8)	4 (2-6)	3 (3-4)	5 (3-7)	0.005

¶Kruskal-Wallis equality-of-populations rank test. *Abbreviations*: 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. SNIP, sniff nasal inspiratory pressure.

Compared to the composite scoring, use of continuous data did not improve discriminative ability for any of the BODE indices significantly (**Table 4.8**, page 116). When assessing BODEs′ individual scoring components (i.e. BMI, FEV_1%, dyspnoea, 6MW), most of BODEs′ predictive ability was attributed to the 6MW component (C = 0.648, 95% CI, 0.609 to 0.673; **Figure 4.6**, page 115). When replacing the 6MW component with the SPPB, or its components 4MGS or balance, the C-index changed from 0.671 (95% CI 0.641 to 0.693) to 0.667 (95% CI 0.627 to 0.694), 0.670 (95% CI 0.634 to 0.694), and 0.682 (95% CI 0.646 to 0.702) respectively. When replacing the 6MW component with SNIP, the C-index changed from 0.671 to 0.672 (95% CI 0.629 to 0.695).

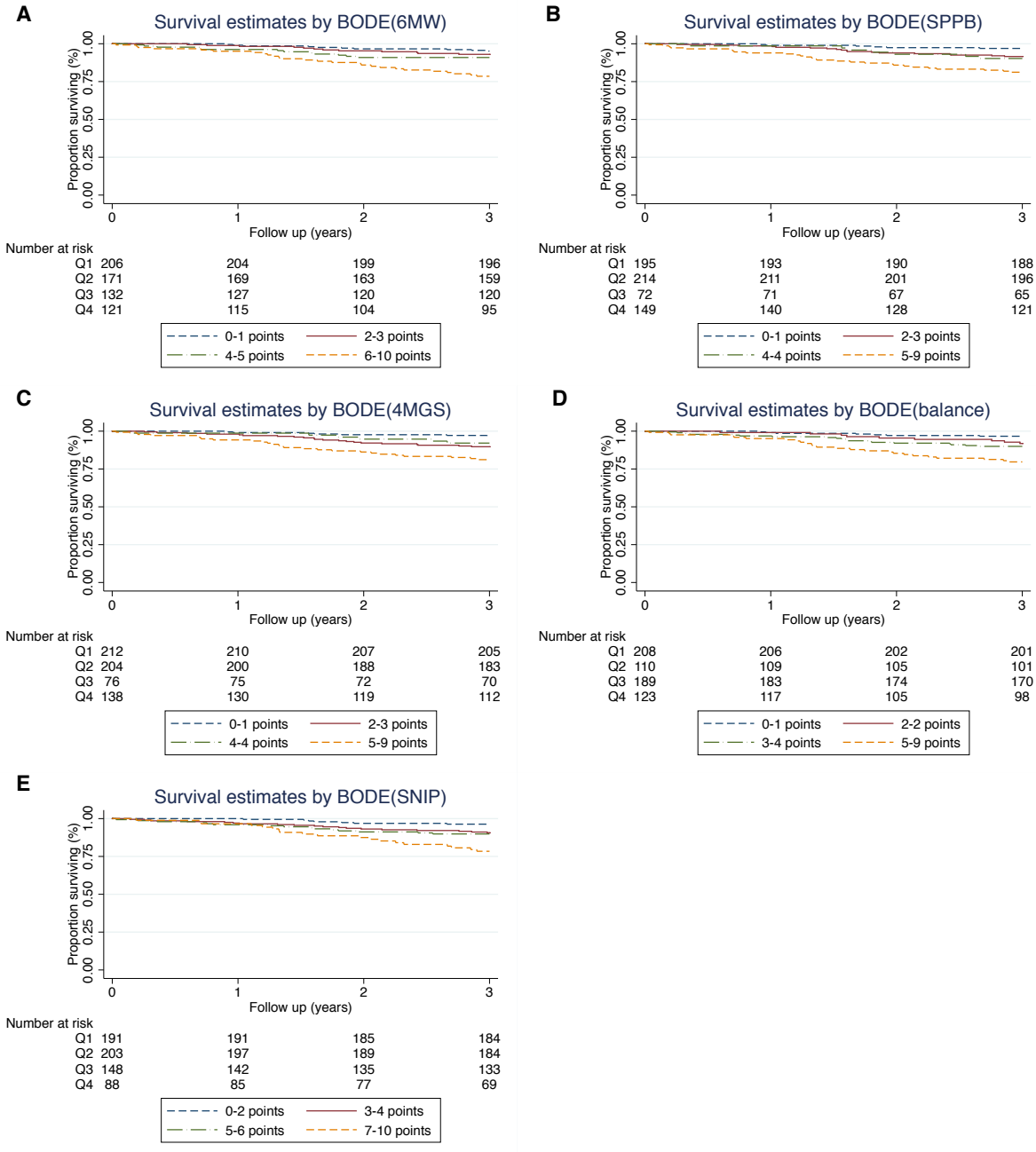


Figure 4.5: Survival risk indices, by quartiles with number at risk at different time points: (A) BODE_{6MW}, (B) $\text{BODE}_{\text{SPPB}}$, (C) $\text{BODE}_{\text{4MGS}}$, (D) $\text{BODE}_{\text{BALANCE}}$, and (E) $\text{BODE}_{\text{SNIP}}$. Mortality data obtained from the UK Office of National Statistics. *Abbreviations*: 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. SNIP, sniff nasal inspiratory pressure.

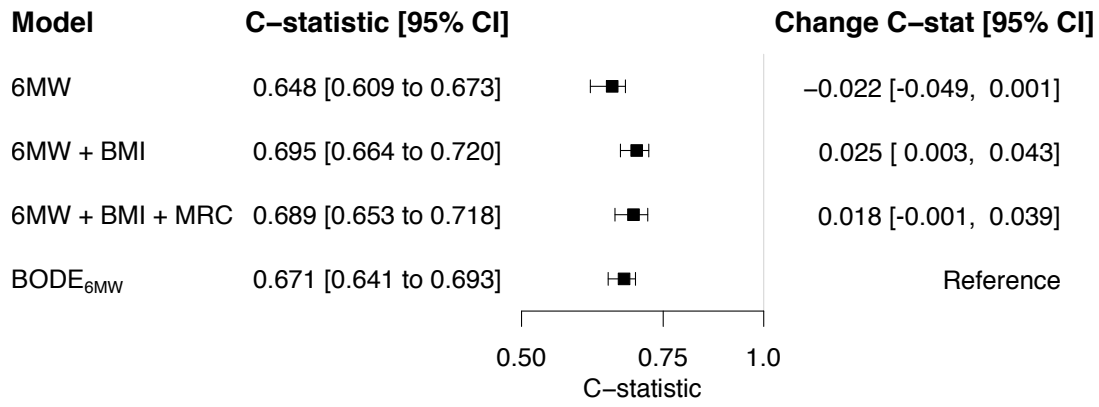


Figure 4.6: Change scores C-index. Individual components BODE. *Abbreviations*: CI, confidence interval. 6MW, six-minute walk. BMI, body mass index. MCR, Medical Research Council.

Table 4.8: Cox proportional hazards regression analyses for all-cause mortality during follow-up, using composite scores for BODE indices.

Variable	Model 1: BMI	Model 2: BMI, MRC	Model 3: BMI, MRC, FEV_1%	Model 4: $BODE_{6MW}$	Model 5: $BODE_{SPPB}$	Model 6: $BODE_{4MGS}$	Model 7: $BODE_{BALANCE}$	Model 8: $BODE_{SNIP}$
Variable				**Hazard Ratio (95% CI)**				
BMI - per 1 point increase	0.90 (0.85 to 0.95)	0.91 (0.86 to 0.95)	0.91 (0.87 to 0.96)	~	~	~	~	~
MRC dyspnoea score	~	1.38 (1.10 to 1.72)	1.27 (0.99 to 1.63)	~	~	~	~	~
FEV_1 - per 5% increase %predicted	~	~	0.93 (0.85 to 1.02)	~	~	~	~	~
BODE	~	~	~	1.30 (1.17 to 1.44)	1.34 (1.19 to 1.51)	1.35 (1.19 to 1.52)	1.40 (1.23 to 1.58)	1.33 (1.18 to 1.49)
C-index	0.608 (0.559 to 0.641)	0.649 (0.608 to 0.679)	0.649 (0.609 to 0.685)	0.671 (0.641 to 0.693)	0.667 (0.627 to 0.694)	0.670 (0.634 to 0.694)	0.682 (0.646 to 0.702)	0.672 (0.629 to 0.695)
Goodness of fit, chi2(3)	0.22	5.17	7.91	0.75	3.88	3.08	5.43	3.5
P >chi2	0.894	0.16	0.048	0.862	0.274	0.379	0.143	0.321
Change in C-statistic	-0.063 (-0.101 to -0.033)	-0.021 (-0.044 to -0.001)	-0.021 (-0.050 to 0.010)	Reference	-0.003 (-0.018 to 0.012)	-0.001 (-0.013 to 0.014)	0.011 (-0.003 to 0.027)	0.001 (-0.015 to 0.016)

All models were stratified by recruitment centre. *Abbreviations*: CI, confidence intervals. BMI, body mass index. MRC, Medical Research Council. FEV_1%, predicted forced expiratory volume one second. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. SNIP, sniff nasal inspiratory pressure.

The $BODE_{6MW}$ had a significantly higher C-index (0.709, 95% CI 0.680 to 0.737) compared to the ADO Index[225] (age, dyspnoea and obstruction; FEV_1; C = 0.649, 95% CI 0.604 to 0.678). Adding age or any musculoskeletal measures to the $BODE_{6MW}$ did not significantly improve the predictive ability of BODE (**Figure 4.7**, page 117). Measuring just SPPB resulted in a C-statistic of 0.617, 95% CI 0.580 to 0.645).

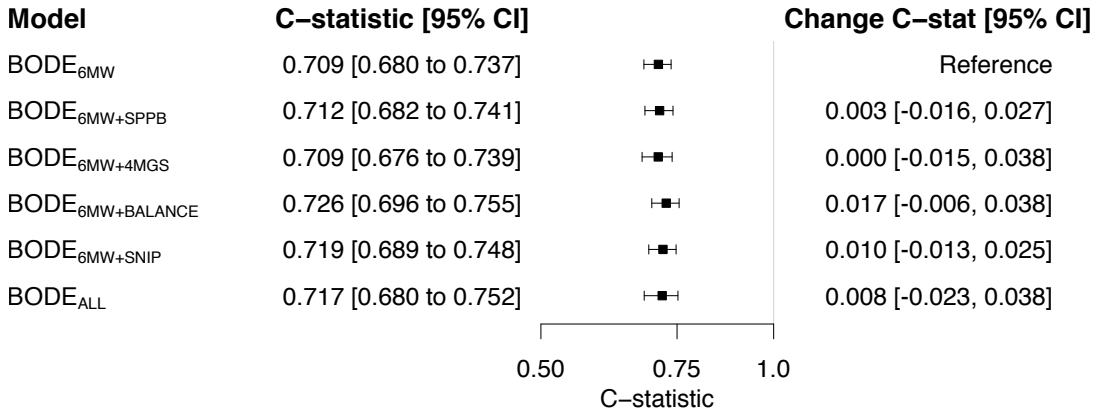


Figure 4.7: C-indices with change scores, alternative BODE models. All models were stratified by recruitment centre. $BODE_{ALL}$ includes body mass index, MRC dyspnoea score, FEV_1, 6MW, short physical performance battery components four-metre gait speed and balance, and sniff nasal inspiratory pressure.

4.3.5 Sensitivity analysis

All 714 individuals (n = 71 deaths after three years of follow-up) were included in sensitivity analyses using multiple imputation of missing baseline values (**Figures 4.8 4.9 4.10**, pages 118-120). Hazard ratios decreased for all models except for the model based on SNIP, which increased only slightly. Cross-validated C-indices decreased but were unchanged between the different models (**Table 4.9**, page 121).

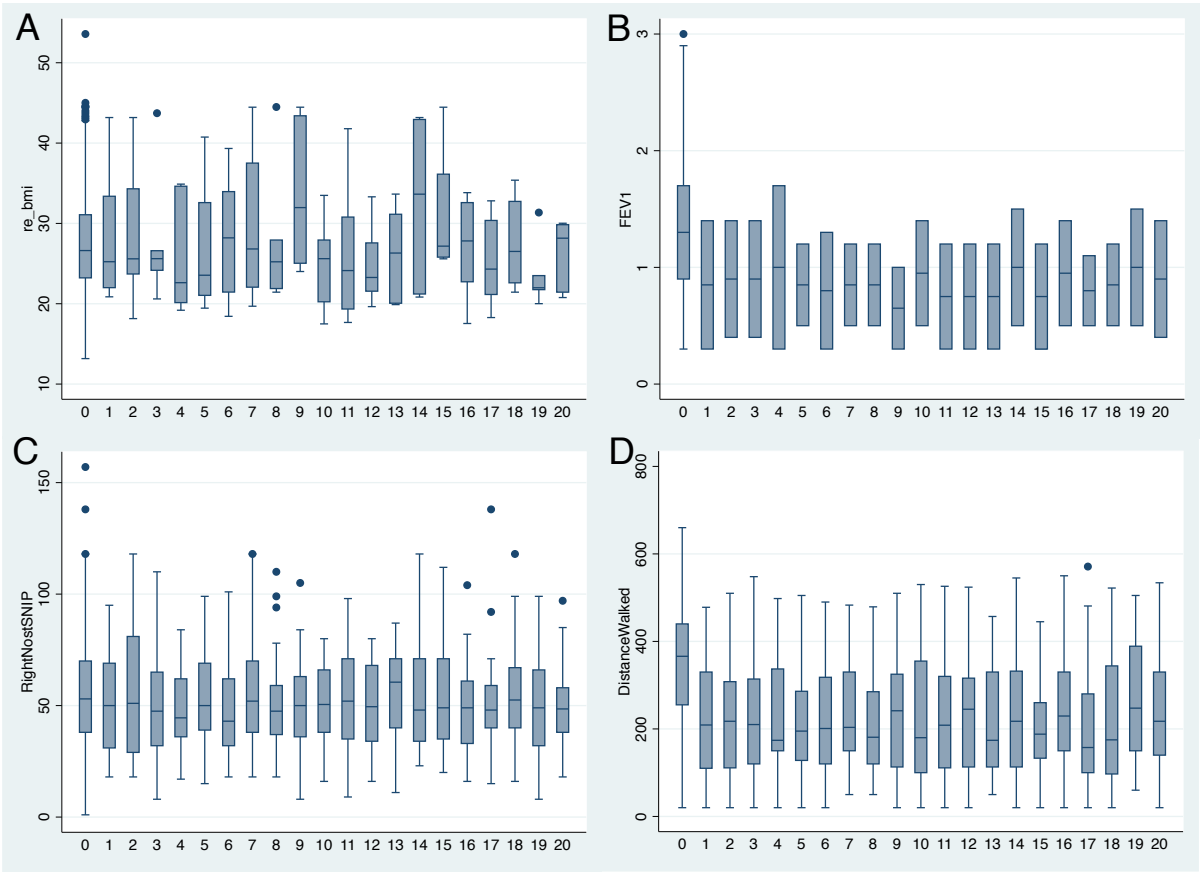


Figure 4.8: Boxplots, after multiple imputation. (A) body mass index. (B) forced expiratory volume in one second. (C) sniff nasal inspiratory pressure. (D) six-minute walk distance. Boxplots in the first column presents data from the complete-case dataset. Boxplots in columns 1-20 present imputed data.

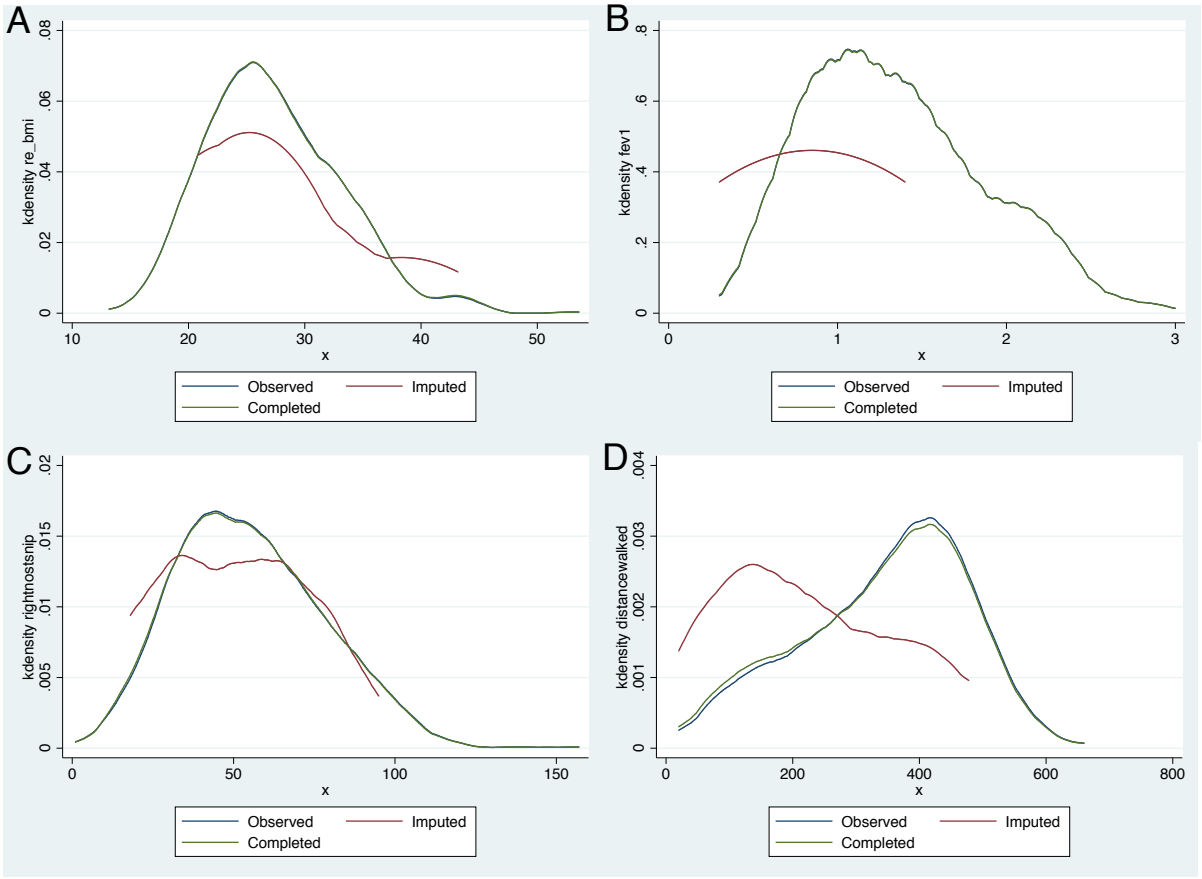


Figure 4.9: Kernel density plots, after multiple imputation. (A) body mass index. (B) forced expiratory volume in one second. (C) sniff nasal inspiratory pressure. (D) six-minute walk distance.

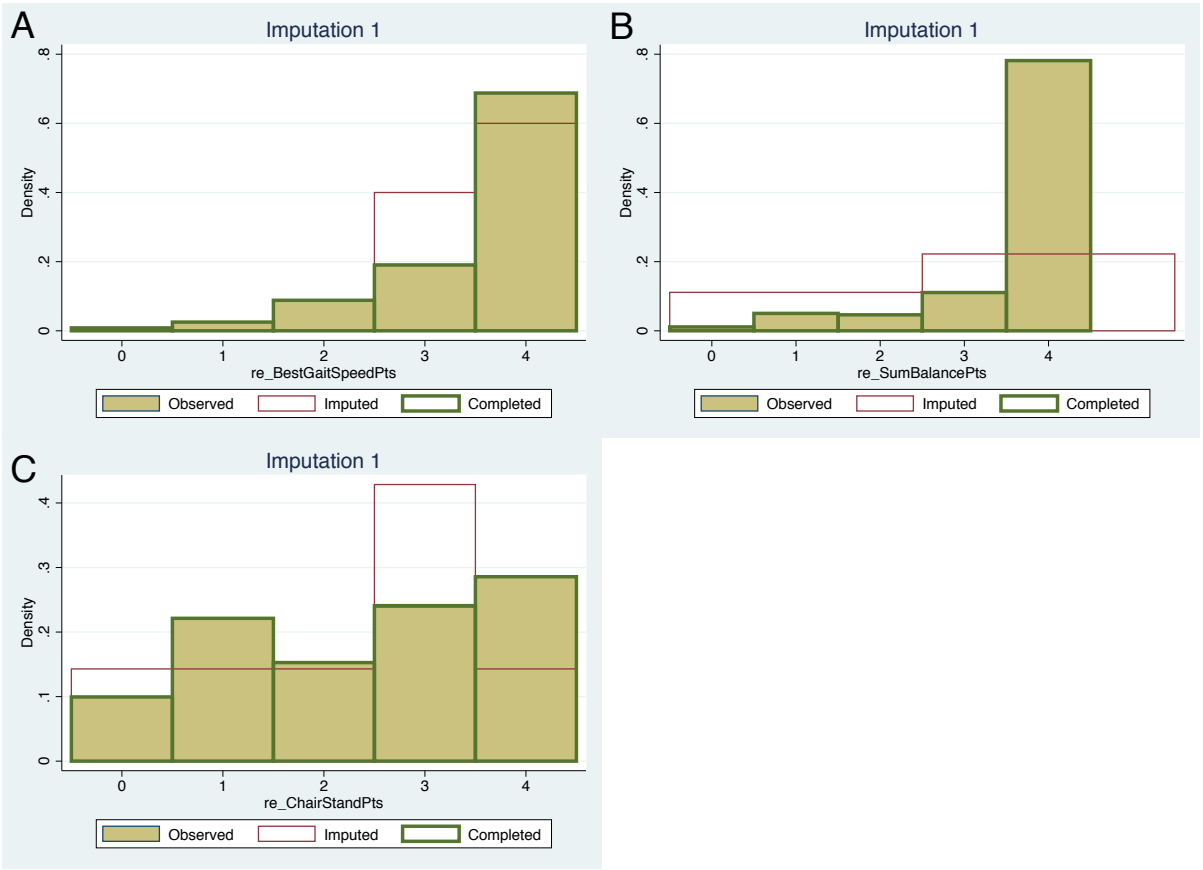


Figure 4.10: Diagnostic histograms, after multiple imputation. (A) four-metre gait speed. (B) balance. (C) chair stance. All three variables are components of the short physical performance battery (SPPB), and were together with the sniff nasal inspiratory pressure score used to estimate the SPPB total score.

Table 4.9: Cox proportional hazards regression analyses for all-cause mortality, using continuous data after multiple imputation (n = 714).

Model	Model 1: BMI	Model 2: BMI, MRC	Model 3: BMI, MRC, FEV_1%	Model 4: $BODE_{6MW}$	Model 5: $BODE_{SPPB}$	Model 6: $BODE_{4MGS}$	Model 7: $BODE_{BALANCE}$	Model 8: $BODE_{SNIP}$
Variable	**Hazard Ratio (95% CI)**							
BMI - per 1 point increase	0.91 (0.87 to 0.96)	0.92 (0.88 to 0.96)	0.93 (0.88 to 0.97)	~	~	~	~	~
MRC dyspnoea score	~	1.42 (1.16 to 1.74)	1.31 (1.04 to 1.63)	~	~	~	~	~
FEV_1 - per 5% increase %predicted	~	~	0.99 (0.97 to 1.00)	~	~	~	~	~
BODE	~	~	~	1.29 (1.17 to 1.42)	1.33 (1.19 to 1.48)	1.33 (1.20 to 1.49)	1.37 (1.22 to 1.53)	1.33 (1.20 to 1.48)
C-index	0.546 (0.529 to 0.560)	0.617 (0.606 to 0.629)	0.662 (0.650 to 0.673)	0.690 (0.679 to 0.699)	0.661 (0.650 to 0.671)	0.660 (0.650 to 0.672)	0.673 (0.658to 0.685)	0.670 (0.658 to 0.682)

All models were stratified by recruitment centre. *Abbreviations*: CI, confidence intervals. BMI, body mass index. MRC, Medical Research Council. FEV_1%, predicted forced expiratory volume one second. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. SNIP, sniff nasal inspiratory pressure.

4.4 Discussion

The main finding of the present study was that SPPB or its 4MGS and balance components, or SNIP may substitute 6MW in BODE for the prediction of all-cause mortality in stable COPD patients (GOLD stage II-IV). The study confirms prior observations that each test on its own is associated with prognosis in simple age and gender adjusted analysis, but QMVC and chair stand, however, performed less well.

4.4.1 Critique of the method

This study has limitations. Firstly, there is no independent validation cohort with a fully comparable dataset. We addressed this issue using a cross-validation technique approach and estimated C-indices through random partitioning of the dataset. Secondly, baseline data differed amongst the recruitment centres but was addressed through stratification by centre. Thirdly, there were missing data with evidence that some was not at random (**Figure 3.3**, page 63). Analysing complete-case data may have introduced bias, and although HRs and C-statistics of the models shifted following MICE, the main conclusions were unchanged. Subgroup analysis, for example, assessing the predictive ability for different age groups was not possible due to the limited number of events. Results should be interpreted with caution, however. Overall, the study sample size was rather small limiting statistical power, making it difficult to provide robust conclusions but instead our findings encourage further analysis in larger cohorts.

In many diseases and scoring systems, for example the ADO Index,[225] age is a strong predictor of death, and unsurprisingly we found the risk of death to rise with age. However, our data showed that ADO compared to BODE_{6MW} provides less discriminative ability, or even adding age to the models did not provide any significant difference in discriminative ability.

The ERICA cohort consisted primarily of individuals GOLD staged II-III, making generalisations to those with very mild COPD, or those with very advanced disease difficult. Additional deaths occurred beyond the three years of follow-up included in the primary analysis. However,

three years of follow-up was chosen because insufficient deaths occur over a short time frame while over a long time period, the predictive ability of BODE diminishes, both because ageing is a strong predictive variable and because measured variables are so distant from the point of death. Consistent with this, some very large COPD trials such as the TOwards a Revolution in COPD Health (TORCH)[174] and Study to Understand Mortality and Morbidity in COPD (SUMMIT)[256] have used 3-year follow-up.

The BODE Index is a point-based system using cut-off points, but Puhan *et al.* have already pointed out the poor calibration of the original BODE Index resulting in an Updated BODE Index.[225] Such a score would have detracted from utility of the BODE at the time of its conception but would be readily available as a phone or web based app now. However, our analysis failed to demonstrate the superiority of continuous rather than categorical data. While this may reflect lack of statistical power, or due to the fact that the SPPB and its components are categorical, our data do not suggest any advantage for a model based on continuous variables. On the other hand, even when outcomes lack statistical significance – a statement about the likelihood of findings being due to chance – this differs from the clinical significance. Clinical significance considers rather the practical value or relevance of a specific biomarker, or using continuous variables for example, which does not necessarily consider statistical significance.

Some of our data did not have defined quartiles of normality and therefore these categories were created from the dataset. This arbitrary point assignment may not be optimal and undermine validity. We do not know how well these quartiles would map to alternative or non-trial populations and to what extent the categorisation we created or the matching to prognostic impact would be maintained. Therefore, the validity of this scoring should be tested in an alternative and independent cohort.

4.4.2 Significance of the findings

By 2030, chronic respiratory diseases, cardiovascular diseases and cancer are each predicted to be responsible for a third of deaths globally.[172] Existing COPD risk indices have so far failed

to achieve widespread clinical adoption. In some cases this may be due to insufficient clinical validation data, but in the case of BODE we speculate that impracticalities relating to the 6MW component, especially such test duration, and the requirement of space and equipment, may play a role.

The SPPB, however, is a simple test to measure lower limb function that requires only a chair, clock and a four-metre flat surface taking less than five minutes. In early 2018, the European Medicines Agency approved the SPPB as a measure of frailty for disease associated with musculoskeletal decline.[88] In fact, our data suggest that even substitution of a single test that is quick (e.g. 4MGS or balance) does not result in any significant loss in predictive ability compared with $BODE_{6MW}$.

Another quick and simple measurement that may be performed at the bedside is the SNIP. Measuring SNIP, however, requires a respiratory pressure metre costing approximate £1000, which many primary care units do not have. Maximal SNIP reflects diaphragm function and is therefore directly impacted by hyperinflation[222,238] and also is susceptible to generalised cachectic influences. In a separate cohort we found that SNIP was predictive of survival and indeed to some extent slightly outperformed direct measurement of hyperinflation.[188] Hyperinflation has previously also been reported as predictor of poor prognosis.[35]

Several studies have tried to improve the BODE Index by adding additional markers (**Table 4.1**, page 100).[267] Within the Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE) study, for example, inflammatory markers such as fibrinogen and C-reactive protein were added. However, only interleukin-6 improved the models′ predictive performance,[42] and other inflammatory markers did not improve performance significantly. Moreover, adding exacerbation history does not show to substantially improve the prognostic capacity of BODE for mortality.[246]

With regard to our final hypothesis, we were not able to demonstrate any significant improvement in the predictive ability of the $BODE_{6MW}$ by adding alternative musculoskeletal measures, specifically SPPB. We believe this is because they are likely to capture the same phenotypic

information as 6MW. Strong correlations between 6MW and SPPB (and its components) have been described previously[149] and were to be expected as they both depend on lower limb function. The QMVC was moderately correlated with 6MW but performed less well. Relationships between QMVC and SPPB were also expected since SPPB was designed as a test of lower limb function. We suspect this explains why the SPPB can be easily substituted for 6MW but also why it conferred no additional value when added to 6MW.

The superiority of the balance component is of interest and may reflect the impact of co-morbidities in COPD beyond those traditionally captured by chest physicians, or indeed by the current protocol. In particular impaired balance may reflect multiple pathologies beyond musculoskeletal weakness. Sensory input is one such that would be impaired in patients with poorly controlled diabetes or alcohol related neuropathy. Visual impairment might also impair balance, as would cognitive issues. Although, only 128 individuals (20%) had a score below the maximum four points. This may indicate that these individuals, especially those with 0-1 points were extremely ill and possibly frail.

4.5 Conclusions

We conclude that the SPPB, and its 4MGS and balance components have potential to replace the 6MW component in the BODE Index for our cohort without significant loss of predictive ability in all-cause mortality, thus potentially enhancing the uptake of such risk measures in clinical practice. Large prospective validation of such simpler measures to replace the 6MW component for use in prognostic tools is warranted.

5

Causes of death in COPD using the UK Biobank

Chapter summary

Background There is uncertainty about the incidence of fatal cardiovascular disease amongst the chronic obstructive pulmonary disease (COPD) population. Early studies have suggested that cardiovascular death occurs in one tenth of COPD patients, whilst more recent studies suggest this is closer to one third. In the Evaluation of the Role of Inflammation in Chronic Airways disease (ERICA) cohort there were only few cardiac deaths, much lower than expected. Further investigations are required to confirm/ replicate these findings. The primary aim of this study was to determine cause of death in the UK Biobank cohort by estimating age-standardised all-cause and cause-specific mortality rates by sex in COPD, and compare these rates with those found in the ERICA cohort. Secondly, we aimed to estimate the age-standardised non-fatal cardiovascular disease incidence rates by sex in COPD, and compare these with those found in the ERICA cohort.

Methods We analysed survival and causes of death of individuals with at least two spirom-

etry measurements and complete information for sex, height and smoking status aged 40 years and older identified in the UK Biobank data (n = 150,542) recruited between 2006 and 2018. We compared outcomes between individuals with defined COPD and non-COPD. Chronic obstructive pulmonary disease was defined based on spirometry results and smoking history. Causes of death and non-fatal cardiovascular related hospital admission were obtained from the UK Office for National Statistics and the National Health Services respectively, and categorised using ICD-10 coding. Sex-specific mortality rates were age-standardised using the 2013 European Standard Population. Hazard ratios were estimated using age- and sex-adjusted Cox regression.

Findings Cumulative survival was 98% (98-98% CI) at 5 years, 97% (97-97% CI) at 7 years, and 96% (96-97% CI) at 9 years. The highest cause-specific mortality rates were cancer-related. In both men and women, COPD had an associated threefold higher risk of early mortality. In women, COPD had an associated fourfold higher risk of cardiac-related death and ninefold higher risk of pulmonary death. In men, COPD was associated with a threefold higher risk of cardiac-related death and sevenfold higher risk of pulmonary death.

Interpretation Findings in the UK Biobank indicate cancer to be the leading cause of death in COPD with a lower cardiac-death rate than expected. Our findings may reflect the downward trend in fatal cardiovascular disease incidence/ improved cardiovascular survival in COPD, and an increased prevalence of death from other causes especially cancer and respiratory disease (i.e. trumped by the speed of progression of these diseases). Chronic obstructive pulmonary disease is, however, associated with higher all-cause mortality, including cardiac-specific death. Differences in fatal and non-fatal cardiovascular disease event rates, and the associated risk by sex suggests tailored COPD management and treatment to be important.

5.1 Background

There is controversy regarding the incidence of fatal and non-fatal cardiovascular (CV) disease amongst the chronic obstructive pulmonary disease (COPD) population. Systematic analysis of

CV comorbidity in patients with COPD reported individuals with COPD to have at least a twice as high risk of non-fatal CV disease compared to those without.[46] However, high levels of heterogeneity were found between studies. Cardiovascular co-morbidity is thought to significantly contribute to both in-hospital and post-discharge mortality in COPD.[8,53,98] Yet, a systematic review on the risk of myocardial infarction (MI) and mortality in COPD suggests there is evidence indicating COPD to be associated with a higher risk of MI and mortality after MI but not in-hospital mortality.[231] In addition, the recent Korean Health and Nutritional Examination Survey (KNHANES, n = 24,429) found COPD to be associated with all-cause mortality but not with increased CV mortality.[247]

Early studies have suggested that CV death occurs in only one tenth of COPD patients,[32] whilst more recent studies suggest this closer to a third. A review on all-cause mortality in COPD published in 2007 found that 35% of deaths have a respiratory cause, 27% CV, 21% cancer, 10% other and 8% are unknown.[174]

The reliability of classifying causes of death from COPD, however, has been questioned with potential under reporting of respiratory disease being the underlying cause of death.[81,118,138] Competing risks of multiple diseases associated with mortality, especially for ageing populations who are more likely to have multiple morbidities, may serve as an explanation. Likewise, minor diseases such as pneumonia, which are common in COPD patients and particularly in the elderly, have found to be often misclassified as an underlying cause of death.[178]

Furthermore, previous studies did not age-standardise death rates or separate by sex. For most causes of death, rates differ by age and require death rates to be adjusted to a standard age distribution. Failing to age-standardise mortality rates may result in variation of overall death rates when comparing two or more populations. Mortality rates may also differ by sex. For example, in the TOwards a Revolution in COPD Health (TORCH) study, causes of death were similar between sexes but with higher mortality rates for men.[39] In the Body mass index, airflow Obstruction, Dyspnoea and Exercise (BODE) study, also higher all-cause mortality rates and death related to pulmonary disease were found in men.[71] In the Evaluation of the Role of

Inflammation in Chronic Airways disease (ERICA) cohort about a third of individuals had non-fatal CV disease with only very few CV related deaths. Out of 714 individuals with a total of 149 deaths during the study period, only eighteen (12%) of individuals died of cardiac causes, much lower than expected. Further investigations, however, are required to confirm/replicate these findings in both UK and non-UK populations.

One of the largest and most detailed UK-based population cohort studies with data made available to any qualified health researcher is the UK Biobank. The UK Biobank contains a large number of measurements, including lung performance measures obtained using spirometry. It has followed-up individuals for survival status, including cause-specific death, and captured any hospital admission including CV related ones since study enrolment.

The primary aim of this study was to determine cause of death within the UK Biobank, estimate age-standardised all-cause mortality and cause-specific rates by sex and COPD status, and compare these rates with those found in the ERICA cohort. Secondly, we aimed to estimate the age-standardised non-fatal CV disease incidence rates by sex and COPD status, and compare these with those found in the ERICA cohort study.

5.2 Methods

5.2.1 Study design and participants

Data from the UK Biobank (application P35826) were linked to mortality data obtained from the UK Office for National Statistics, and electronic health record (EHR) data (i.e. hospital admission data) obtained from the National Health Services. Data were collected between 2006 and 2010 with continuous follow-up of survival status. The UK Biobank is a large long-term prospective national and international biobank study in the United Kingdom with approximately 500.000 individuals aged between 40-69 years. Individuals recruited in the UK Biobank were expected to develop common diseases over time including lung and CV disease: 5000

cases of COPD and 10.000 cases of myocardial infarction and coronary death eight years after recruitment. Data were captured at 22 assessment sites across the UK and included demographics, medical history, measures of pulmonary function, and others. Full details are provided elsewhere.[255] Analysis were limited to those of white European ancestry, at least two forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC) measures, and complete information for spirometry method, sex, standing height, and smoking status, and aged above forty years.

5.2.2 Definition of COPD

Individuals with COPD at baseline were identified using the following selection criteria: (i) COPD defined as post-bronchodilator FEV_1 of 80% or less of their predicted value, a baseline FEV_1/FVC ratio of $<$0.7 measured by spirometry, (ii) a smoking history of at least ten pack-years, and (iii) Global initiative for Obstructive Lung Disease (GOLD) stage $\geq$II.[107]

5.2.3 Statistical analysis

Based on baseline disease, individuals were categorised according to defined COPD or non-COPD (reference group). The primary outcome measure was cause-specific incidence of death within the UK Biobank by COPD status. Causes of death were categorised according to the international classification of diseases and related health problems 10^{th} revision coding (ICD-10; **Table 5.1**, page 133). The secondary outcome measure was hospitalised non-fatal CV disease derived through EHR data. These data were first cleaned for episode status and events were extracted from both primary and secondary positions of ICD-10 coding (**Table 5.2**, page 135). Time to event (i.e. time to death or time to first hospitalised non-fatal CV disease) was defined as time to death or admission from the baseline visit date to date of death or admission, or censored at November 2018. Time to hospitalised non-fatal CV disease was defined as time to first admission from the baseline visit to date of admission, or censored at November 2018. Sex specific all-cause and cause-specific mortality rates were age-standardised using the 2013

European Standard Population.[198] Hazard ratios (HRs) were estimated using Cox proportional hazards regression, adjusted for age and sex. Demographics were described using number and percentage for categorical variables, and the medians (inter-quartile ranges (IQR)) for continuous variables.

Table 5.1: Definitions of underlying (primary) cause of death: ICD10, available in UK Biobank.

End point	ICD-10 codes
Diseases of the circulatory system (IX)	F01, G45, I05-I15, I20-I28, I30-I52, I60-I89, Q20-Q28, R96
Vascular dementia	F01
Transient cerebral ischaemic attacks and related syndromes	G45
Chronic rheumatic heart diseases	I05-I09
Hypertensive diseases	I10-I15
Ischaemic heart diseases	I20-I25
Pulmonary heart disease and diseases of pulmonary circulation	I26-I28
Other forms of heart disease	I30-I52
Cerebrovascular diseases	I60-I69
Diseases of arteries, arterioles and capillaries	I70-I79
Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified	I80-I89
Congenital malformations of the circulatory system	Q20-Q28
Other sudden death, cause unknown	R96
Diseases of the respiratory system (X)	J09-J18, J20-J22, J30-J47, J60-J70, J80-J86, J90-J99
Influenza and pneumonia	J09-J18
Other acute lower respiratory infections	J20-J22
Other diseases of upper respiratory tract	J30-J39
Chronic lower respiratory diseases	J40-J47
Lung diseases due to external agents	J60-J70
Other respiratory diseases principally affecting the interstitium	J80-J84
Suppurative and necrotic conditions of lower respiratory tract	J85-J86
Other diseases of pleura	J90-J94
Other diseases of the respiratory system	J95-J99
Neoplasms (II)	C00-C26, C30-C41, C43-C58, C60-C97, D10-D48
Malignant neoplasms of lip, oral cavity and pharynx	C00-C14
Malignant neoplasms of digestive organs	C15-C26
Malignant neoplasms of respiratory and intrathoracic organs	C30-C39
Malignant neoplasms of bone and articular cartilage	C40-C41
Melanoma and other malignant neoplasms of skin	C43-C44
Malignant neoplasms of mesothelial and soft tissue	C45-C49
Malignant neoplasm of breast	C50
Malignant neoplasms of female genital organs	C51-C58
Malignant neoplasms of male genital organs	C60-C63
Malignant neoplasms of urinary tract	C64-C68

End point	ICD-10 codes
Malignant neoplasms of eye, brain and other parts of central nervous system	C69-C72
Malignant neoplasms of thyroid and other endocrine glands	C73-C75
Malignant neoplasms of ill-defined, secondary and unspecified sites	C76-C80
Malignant neoplasms, stated or presumed to be primary, of lymphoid, haematopoietic and related tissue	C81-C96
Malignant neoplasms of independent (primary) multiple sites	C97
Benign neoplasms	D10-D36
Neoplasms of uncertain or unknown behaviour	D37-D48
Other	
Certain infectious and parasitic diseases (I)	A00-A09, A15-A28, A30-A49, A80-A89, B00-B09, B15-B64, B90-B94, B99
Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (III)	D55-D77, D80-D89
Endocrine, nutritional and metabolic diseases (IV)	E00-E07, E10-E16, E20-E35, E65-E68, E70-E90
Mental and behavioural disorders (V)	F00, F02-F19, F30-F39, F80-F89
Diseases of the nervous system (VI)	G00-G14, G20-G26, G30-G32, G35-G37, G40-G41, G47, G50-G64, G70-G73, G80-G83, G90-G99
Diseases of the digestive system (XI)	K20-K31, K35-K38, K40-K46, K50-K52, K55-K67, K70-K77, K80-K87, K90-K93
Diseases of the skin and subcutaneous tissue (XII)	L00-L08, L50-L54, L80-L99
Diseases of the musculoskeletal system and connective tissue (XIII)	M00-M03, M05-M25, M40-M43, M45-M49, M60-M63, M70-M90
Diseases of the genitourinary system (XIV)	N00-N08, N10-N23, N25-N51, N80-N98
Congenital malformations, deformations and chromosomal abnormalities (XVII)	Q00-Q07, Q38-Q45, Q60-Q89
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (XVIII)	R00-R09, R47-R69, R95, R97-R99
External causes of morbidity and mortality (XX)	V01-V29, V40-V49, V80-V97, W00-W99, X00-X19, X30-X49, X58-X99, Y10-Y34, Y40-Y69, Y83-Y89
Codes for special purposes (XXII)	U00-U49

Abbreviations: ICD-10, international classification of diseases 10th edition.

Table 5.2: Definitions of diagnoses for non-fatal cardiovascular disease – main ICD10

End point	ICD-10 codes
Diseases of the arteries	I70.2, I72, I73.9-I79, E10.5, E11-E14
Peripheral arterial disease	I70.2, I73.9, E10.5, E11-E14
Diseases of arteries, arterioles and capillaries	I72, I74-I79
Coronary heart disease	I20.0-I20.1, I20.8-I21-I25
Angina	I20.1, I20.8-I20.9
Unstable angina	I20.0, I24
Coronary heart disease not otherwise specified	I25
Acute MI, and certain current complications following acute MI	I21, I23
Subsequent myocardial infarction	I22
All stroke	I60, I61, I.62, I63, I64, I65-I69, F01, G46.3-G46.7, G458, G459
Subarachnoid haemorrhage	I60
Intra-cerebral haemorrhage	I61
Cerebral infarction	I63
Stroke, not specified as haemorrhage or infarction	I64
Stroke syndromes	G46.3-G46.7
Transient ischaemic attack	G458, G459
Other stroke	I62, I65-I69, F01
Heart failure	I11.0, I13.0, I13.2, I50
Heart failure	I50
Hypertensive heart disease with (congestive) heart failure	I11.0
Hypertensive heart and renal disease with (congestive) heart failure	I13.0
Hypertensive heart and renal disease with both (congestive) heart failure and renal disease	I13.2

Atrial fibrillation and flutter (I48) and hypertensive diseases (I10-I15) were considered risk factors and therefore not included. *Abbreviations*: ICD-10, international classification of diseases 10th edition. MI, myocardial infarction.

5.3 Findings

5.3.1 Descriptive statistics

From a total of 502 595 individuals in the UK Biobank, 472 866 were of white European ancestry but only 150 542 individuals had at least two FEV_1 and FVC measures, complete phenotype data (i.e. height and sex) and were aged >40 years and therefore included in the primary analysis (**Figure 5.1**, page 137). Those without complete phenotype data (i.e. excluded from analysis) were one year older (p <0.001) and more likely to be male (p <0.001) compared to those included. Most missing data were present for FEV_1 (68%) and FVC (25%) measures. Of those included, 9926 (7%) individuals were identified with COPD. The selected cohort in UK Biobank (n = 150 542) had slightly more women (58%) than men, the median age 57 years (range 41-72), 45% were overweight, median FEV_1 (IQR) was 2.7 (2.2-3.2) litre, and a third was an ever smoker at baseline with 10% GOLD staged II or above (**Table 5.3**, page 136). Those with COPD were older (p <0.001), had higher body mass index (BMI; p <0.001), had lower FEV_1 (p <0.001) and were more likely to be male (p <0.001) compared to those without COPD.

Table 5.3: Baseline characteristics (n = 150 542).

Characteristic	Selected cohort	Non-COPD	COPD
No. (%) of participants	150 542	140 616	9926
Age at survey, median (IQR), years	57 (50-63)	57 (50-63)	61 (56-65)
Male sex, No. (%)	62 844 (42)	57 416 (41)	5428 (55)
BMI, median (IQR), kg/m^2	27 (24-30)	27 (24-30)	28 (25-31)
FEV_1, median (IQR), litre	2.7 (2.2-3.2)	2.7 (2.3-3.3)	1.9 (1.5-2.3)
Ever smoker, No. (%)	46 279 (31)	36 353 (26)	9926 (100)
GOLD stage, ≥II	23 118 (9)	13 192 (9)	9926 (100)

Values are given as the medians and interquartile ranges (IQR), or No. of cases (%). Baseline data of 150 542 patients included. *Abbreviations*: BMI, body mass index. FEV_1, forced expiratory volume in one second. GOLD, global initiative for obstructive lung disease.

5.3.2 Event and mortality rates

Cumulative survival was 98% (98-98% CI) at 5 years, 97% (97-97% CI) at 7 years, and 96% (96-97% CI) at 9 years. Crude all-cause mortality and cause-specific mortality rates gradually

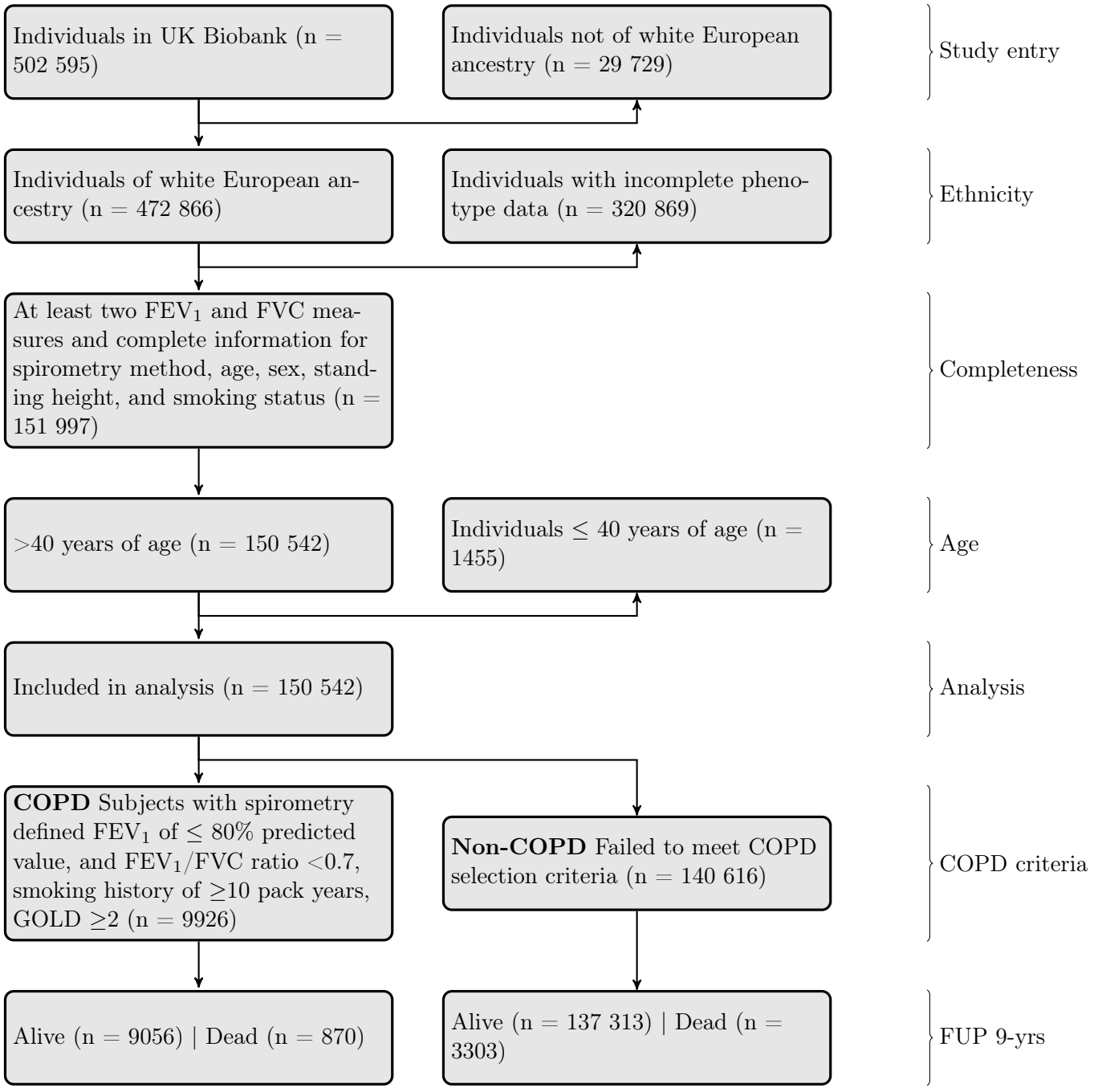


Figure 5.1: Flow diagram of sample selection strategy, UK Biobank. Total number of deaths (n = 4173). *Abbreviations*: FUP, follow-up period. FEV_1, forced expiratory volume one second. FVC, forced vital capacity. GOLD, global initiative for obstructive lung disease.

increased throughout the study period for both sexes (**Figure 5.2**, page 139). During 7-year median follow-up a total of 4173 (3%) individuals died: 870 (9%) with COPD and 3303 (2%) non-COPD. Age- and sex-adjusted survival rates indicated those with COPD to be at higher risk of mortality (**Figure 5.3**, page 140). The leading primary (underlying) cause of death was cancer followed by diseases of the circulatory system, respiratory system, and other (**Figure 5.4**, page 141). Individuals with defined COPD had a higher proportion of deaths related to cardiac and respiratory diseases, and a lower proportion of deaths related to cancer compared to those without COPD, regardless of sex.

Age- and sex-adjusted all-cause mortality rates after 9-year follow-up were higher in men (0.53 per 100 person-years (95% CI 0.48 to 0.58)) than in women (0.28 (95% CI 0.25 to 0.31); **Figure 5.5**, page 142). Death due to cancer had the highest rate followed by cardiac disease. Chronic obstructive pulmonary disease was associated with higher all-cause and cause-specific mortality, in particular men. All-cause mortality for individuals with COPD (1.23 [95% CI 0.97-1.50] per 100 person-years for men, and 0.77 [95% CI 0.53-1.01] for women) were higher compared to those without COPD (0.45 [95% CI 0.40-0.50] for men, and 0.25 [95% CI 0.22-0.28] for women). Cardiac death rates for COPD were 0.29 (95% CI 0.17-0.41) per 100 person-years for men, and 0.11 (95% CI 0.04-0.18) for women compared to 0.10 (95% CI 0.08-0.12) and 0.03 (95% CI 0.02-0.04) for men and women without COPD, respectively.

Age- and sex-adjusted HRs for all-cause mortality were 2.70 (95% CI 2.43 to 2.93) and 3.06 (95% CI 2.69-3.47) for men and women with COPD, respectively (**Figure 5.5**, page 142). For all-cause and cause-specific mortality, HRs were higher for women. Hazard ratios for cardiac death for women with COPD were 3.55 (95% 2.57-4.89), and 2.81 (95% 2.32-3.41) for men with COPD.

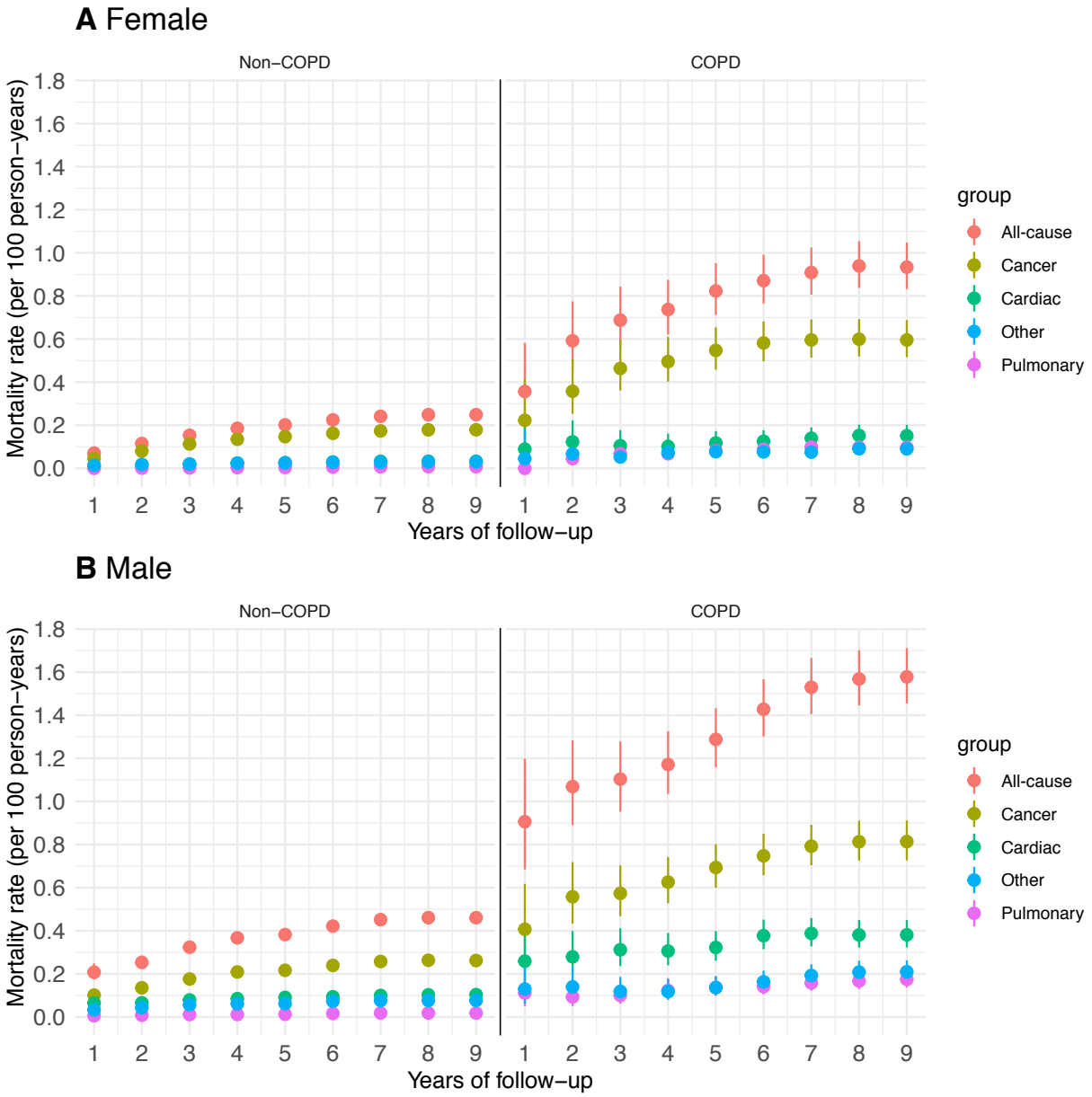


Figure 5.2: Crude annual all-cause and cause-specific mortality rates with 95% CI, by sex and COPD status. Figure (A) indicates event rates by COPD status for women. Figure (B) indicates event rates by COPD status for men.

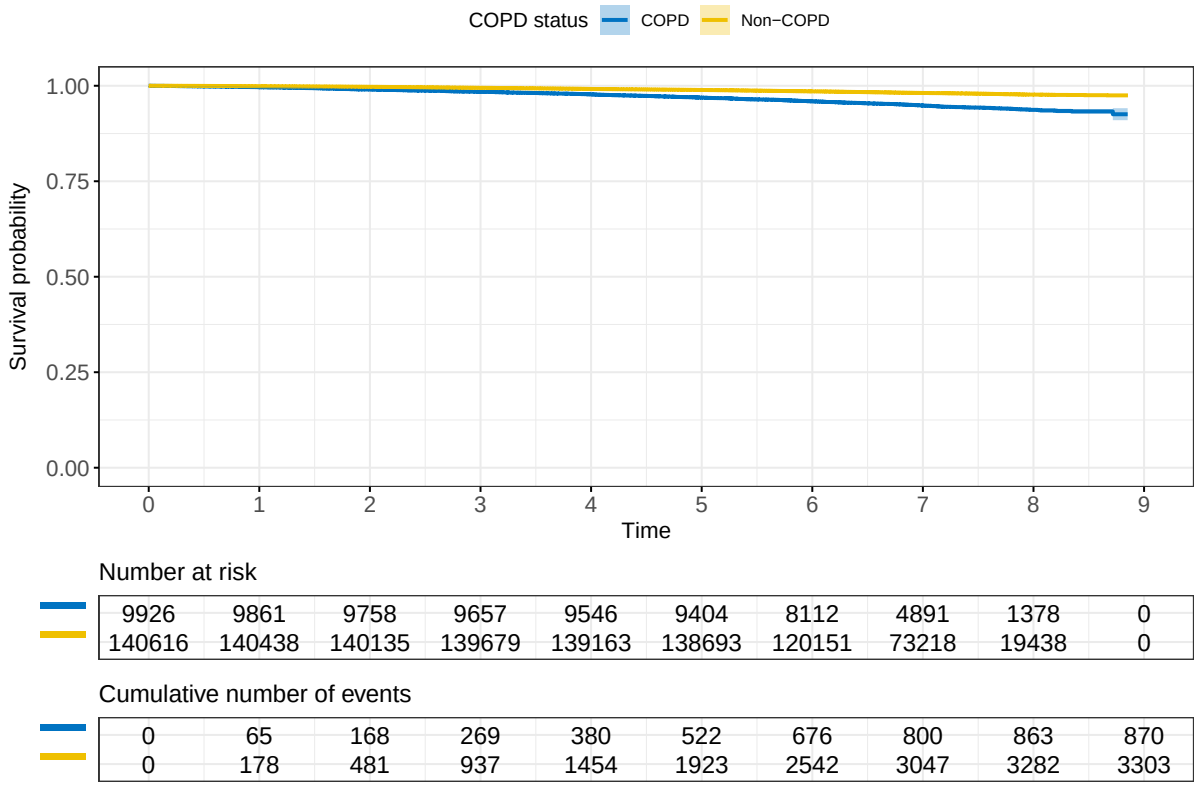


Figure 5.3: Age- and sex-adjusted survival, by COPD status. Time in years. Risk table indicates the number of individuals at risk of death during the study period, and cumulative number of events at specific time points.

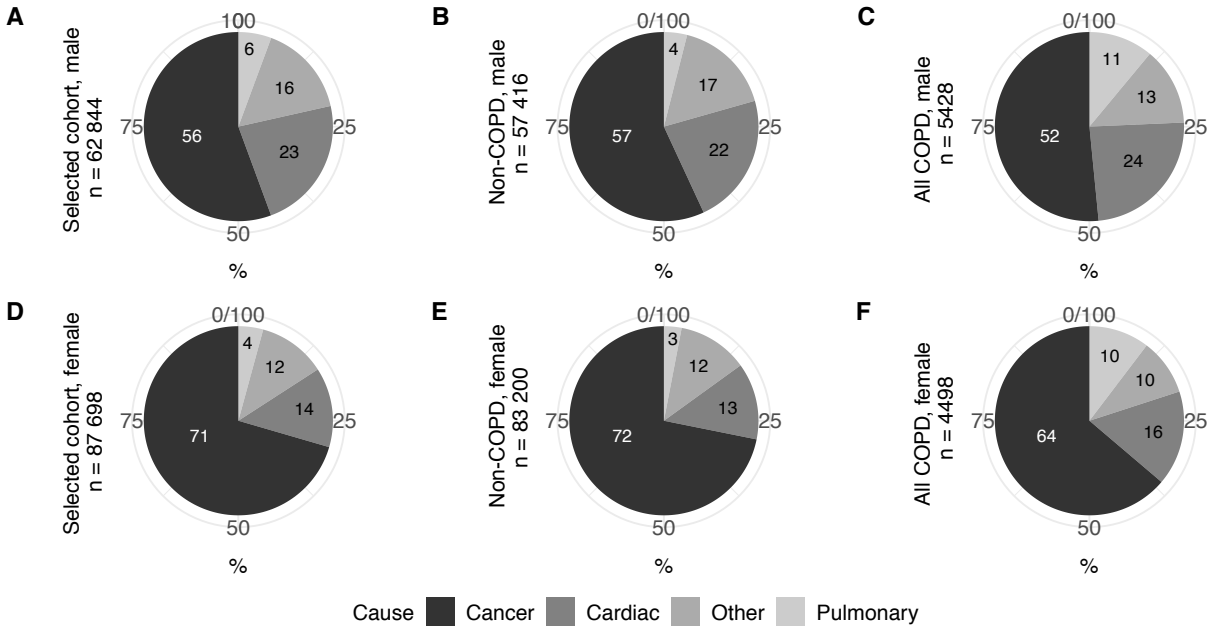


Figure 5.4: Cause-specific death by sex of (A, D) selected cohort: males (n = 2427 deaths) and females (n = 1746), (B, E) non-COPD: males (n = 1847 deaths) and females (n = 1456), and (C, F) COPD: males (n = 580 deaths) and females (n = 290).

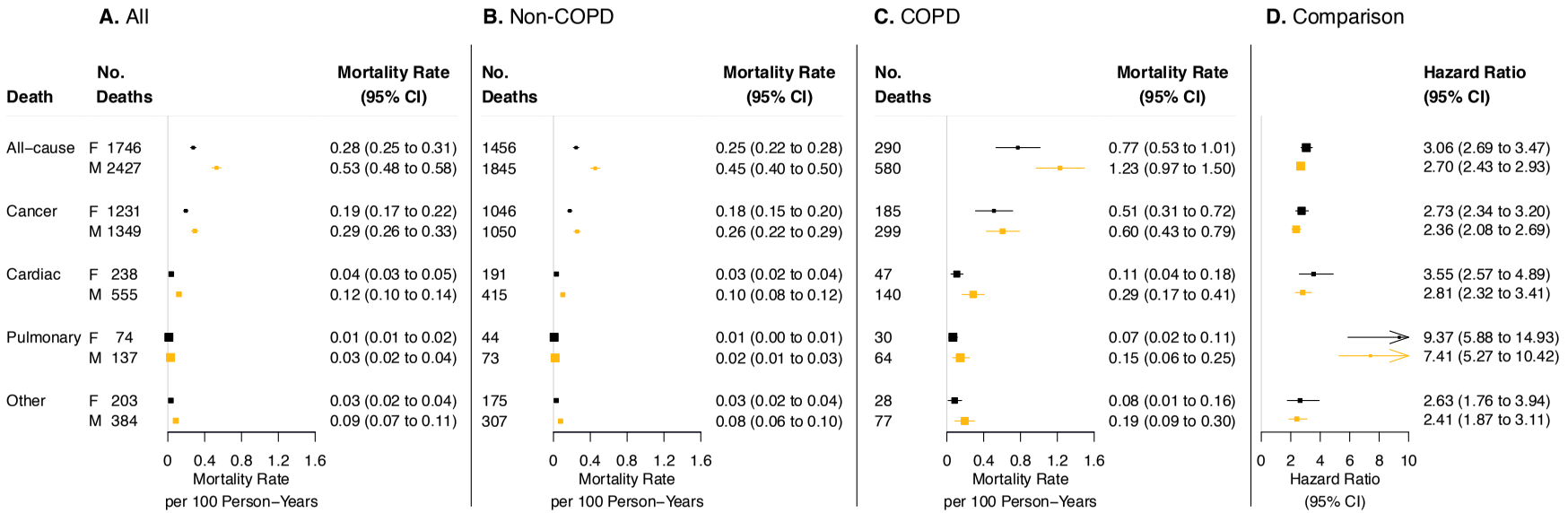


Figure 5.5: Age-standardised mortality rates and hazard ratios with 95% CI after nine years of follow-up, by sex. The selected cohort (A) included 150 542 participants over 1053 741 person-years: 87698 women over 615 956 person-years, and 62844 men over 437 785 person-years. Figures (A-C) present all-cause and cause-specific mortality rates, by sex. Figure (D) presents all-cause and cause-specific hazard ratios, by sex. Top rows indicate results for female. Second rows indicate results for male (coloured yellow).

5.3.3 Non-fatal cardiovascular disease

During the study period there were in total 13 800 (9%) individuals admitted to hospital for non-fatal CV disease. Age-adjusted overall non-fatal CV event rate was higher for men than women (2.10 [95% CI 2.01-2.20] vs. 0.87 [95% CI 0.82-0.92] per 100 person-years; **Figure 5.6**, page 144). The group with defined COPD had a higher non-fatal CV event rate than the non-COPD group, especially men (3.22 [95% CI 2.82-3.63] vs. 1.69 [95% CI 1.60-1.77] per 100 person-years). Event rates for women with COPD compared to those without were 1.90 [95% CI 1.54-2.26] vs. 0.81 [95% CI 0.75-0.86] per 100 person-years, respectively.

Age- and sex-adjusted HRs for non-fatal CV disease rates were 2.28 (95% CI 2.10-2.46) for women with COPD, and 1.88 (95% CI 1.78-1.99) for men with defined COPD at baseline (**Figure 5.6**, page 144).

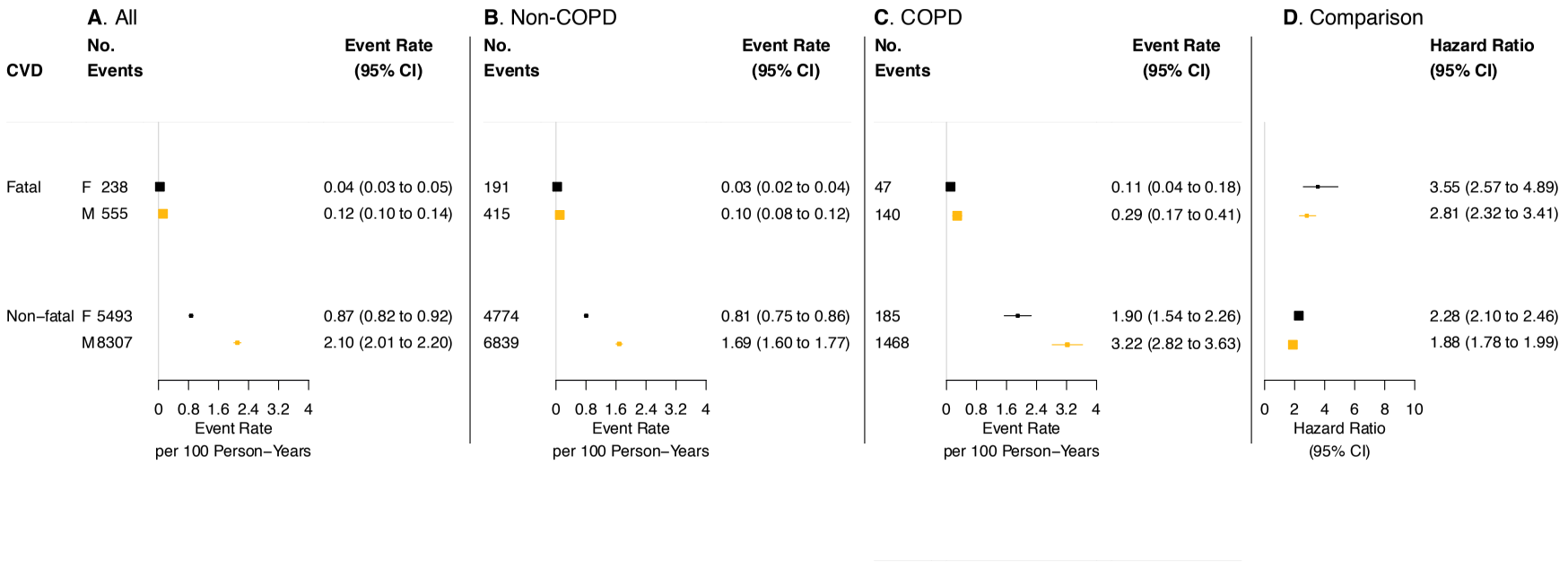


Figure 5.6: Age-standardised fatal- and non-fatal event rates and hazard ratios with 95% CI after nine years of follow-up, by sex. Figures A-C present fatal- and non-fatal cardiac-specific rates, by sex. Figure (D) presents cardiac-specific hazard ratios, by sex. Top rows indicate results for women. Second rows indicate results for men (coloured yellow).

5.3.4 Comparison with the ERICA cohort

In comparison with the ERICA cohort, UK Biobank participants were a median six years younger, and had a median 0.6 litre (1.9 vs. 1.3 litre and 16% predicted (69% vs. 53%) higher FEV_1 (**Table 5.4**, page 146). In the UK Biobank there were fewer men (55% vs. 61%) and more current smokers (43% vs. 31%). Body mass index, biochemical measures (i.e. white cell count, neutrophils, and haemoglobin), resting heart rate and systolic blood pressure values were similar between the two cohorts.

Age-standardised all-cause mortality rates were higher for both sexes in the ERICA cohort (**Figure 5.7**, page 147). Cardiac death rates were higher for men, and slightly higher for both sexes in the ERICA cohort. Namely, cardiac death rates in the UK Biobank cohort were 0.11 (95% CI 0.04-0.18) per 100 person-years compared to 0.17 (95% CI 0.00-0.41) in the ERICA cohort, for women. Cardiac death rates in the UK Biobank cohort were 0.29 (95% CI 0.17-0.41) per 100 person-years compared to 0.31 (95% CI 0.01-0.67) in the ERICA cohort, for men. In the UK Biobank mortality rates were highest for cancer followed by cardiac and other causes of death. The lowest rate of death was due to pulmonary disease, whereas the ERICA cohort had the highest ratio of death due to pulmonary disease.

When comparing non-fatal CV disease event rates between the two cohorts, event rates for men were rather similar between the two cohorts. Rates in the UK Biobank cohort were 3.22 (95% CI 2.82-3.63) per 100 person-years, compared to 3.21 (95% CI 2.17-4.32) in the ERICA cohort (**Figure 5.8**, page 147). Event rates for women in the UK Biobank cohort were lower than those for women in the ERICA cohort.

Table 5.4: Comparison of baseline characteristics between the UK Biobank and ERICA cohorts.

Characteristic	Total, UK Biobank (COPD)	N (%)	Total, ERICA (COPD)	N (%)
Description				
Age, median (IQR), years	61 (56-65)	9926 (100)	67 (62-73)	714 (100)
Male sex, No. (%)	5428 (55)	9926 (100)	434 (61)	714 (100)
BMI, median (IQR), kg/m^2	28 (25-31)	9926 (100)	27 (23-31)	707 (99)
Lung function				
FEV_1, median (IQR), litre	1.9 (1.5-2.3)	9926 (100)	1.3 (0.9-1.7)	712 (100)
FEV_1, median (IQR), % predicted	69 (60-75)	9926 (100)	53 (40-65)	712 (100)
Smoking status - current, n (%)	4287 (43)	9926 (100)	218 (31)	710 (99)
Ever smoker, n (%)	9926 (100)	9926 (100)	714 (100)	714 (100)
GOLD stage, ≥II	9926 (100)	9926 (100)	713 (100	713 (100)
Shortness of breath walking on level ground (MRC IV)	1031 (10)	3321 (33)	146 (21)	709 (99)
Biochemical measures				
WCC (mcL)	7.8 (6.6-9.2)	9502 (96)	7.1 (6.0-8.6)	704 (99)
Neutrophils (mm^3)	4.8 (3.9-5.9)	9486 (96)	4.5 (3.6-5.6)	701 (98)
Haemoglobin (g/dL)	14.4 (13.6-15.3)	9502 (96)	14.3 (13.4-15.3)	703 (98)
Cardiovascular status				
Heart rate (bpm)	72 (64-80)	9317 (100)	74 (66-82)	702 (98)
SBP (mmHg)	142 (129-156)	9302 (94)	142 (131-154)	706 (99)
PWV (m/sec)	10.2 (8.1-12.1)	3283 (33)	9.8 (8.4-11.8)	654 (92)
AIx (%)	22 (17-30)	172 (2)	28 (20-34)	699 (98)

Values are given as the medians and interquartile ranges (IQR), or No. of cases (%). *Abbreviations*: BMI, body mass index. FEV_1, forced expiratory volume in one second. GOLD, global initiative for obstructive lung disease. MRC, Medical Research Council. WCC, white cell count. SBP, systolic blood pressure. MAP, mean arterial pressure. CIMT, carotid intima-media thickness. PWV, pulse wave velocity. AIx, augmentation index.

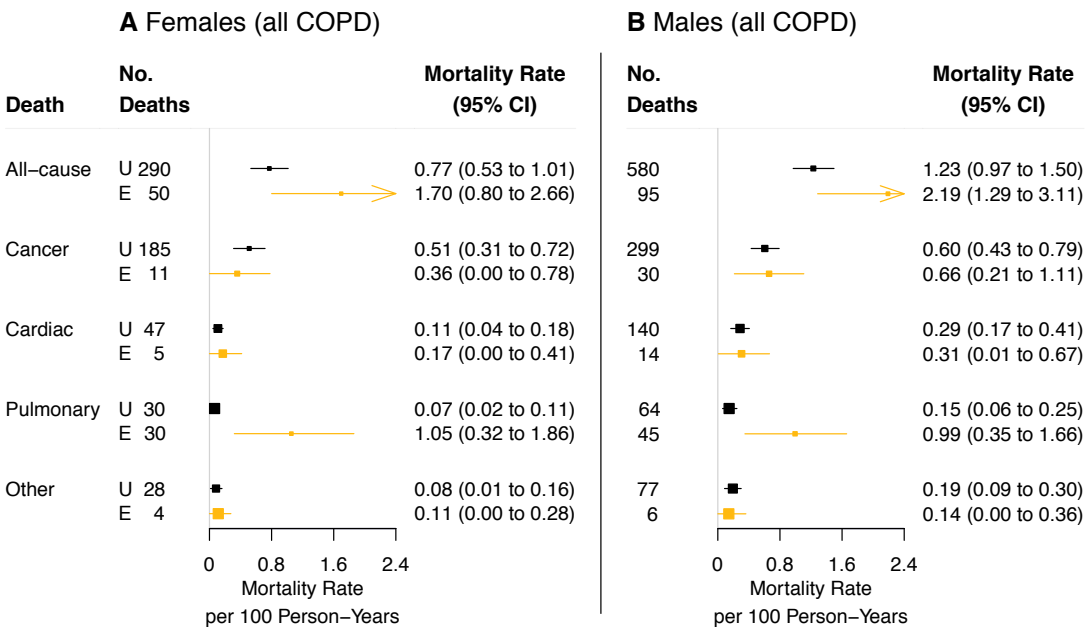


Figure 5.7: Age-standardised mortality rates and hazard ratios with 95% CI after nine years of follow-up, by sex and cohort study. Figures (A) presents all-cause and cause-specific mortality rates for women, by cohort. Figure (B) presents all-cause and cause-specific mortality ratios, by sex and cohort. Top rows indicate results from the selected UK Biobank cohort (U). Second rows indicate results from the ERICA cohort (E; coloured yellow).

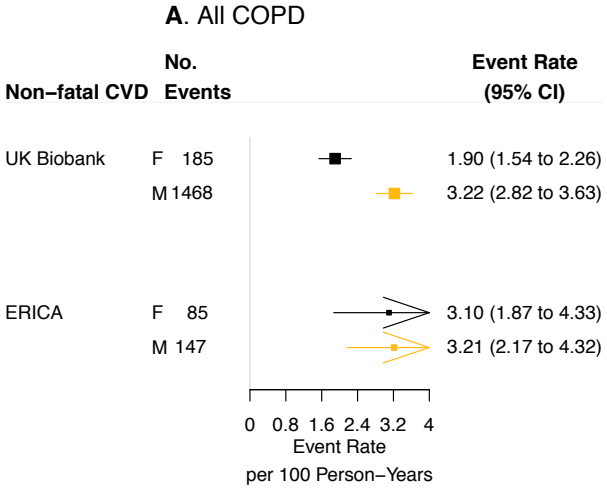


Figure 5.8: Age-standardised non-fatal cardiovascular event rates and hazard ratios with 95% CI after nine years of follow-up, by sex. Top rows indicate results for women (F). Second rows indicate results for men (M; coloured yellow).

5.4 Discussion

This study evaluated age-standardised all-cause and cause-specific mortality, and non-fatal CV disease rates by COPD status and sex, and the associated risk of COPD with these outcomes. The main findings of this analysis were that despite the higher event rates for men, women with COPD were at higher risk of both fatal- and non-fatal CV disease. In addition, COPD in men was associated with a 2.7-fold higher risk of early mortality, including a 2.8-fold higher risk of cardiac-related death and 7.4-fold higher risk of pulmonary death. Whereas in women, COPD was associated with a threefold higher risk of early mortality including an 3.6-fold higher associated risk of cardiac-related death and 9.4-fold higher risk of pulmonary death. In addition, COPD in men was associated with a 1.9-fold higher risk of non-fatal CV disease, and 2.3-fold higher risk in women.

Mortality rates were higher in ERICA than in the UK Biobank, although cardiac-related death rates were similar. Non-fatal CV disease rates for women were higher in the ERICA cohort, yet for men these were similar. Pulmonary-related death was the primary cause of death in ERICA, indicating a population with more severe pulmonary disease.

Differences between findings in the UK Biobank and the ERICA cohort may be explained by the variation in populations, despite limiting analyses to individuals with COPD specific spirometry results aged forty years of age and over. The population in ERICA was specifically recruited because of their lung disease, and may have had more severe lung disease (i.e. worse FEV_1 and more shortness of breath) than individuals recruited in UK Biobank. Besides the larger proportion of males (61%) recruited in ERICA, there may have been differences in geographical distribution between cohorts. Most participants in the ERICA study were recruited in Cardiff followed by Nottingham, Edinburgh, Cambridge and London. Individuals from London were slightly healthier compared to other sites and individuals from Cardiff (52%) had more severe lung disease. Participants in the UK Biobank were primarily recruited in England/Wales

with about 10% from Scotland. Overall, the UK Biobank cohort was younger, healthier and included more women compared to ERICA, and therefore the two cohorts may not have been that well comparable. Simultaneously in terms of the generalisability of the UK Biobank cohort, compared to nationally representative data sources, individuals enrolled in the UK Biobank were more likely to be female and older with higher socio-economic status than non-participants.[97] Hence, in terms of generalisability, it may be difficult to generalise findings to other populations. Findings need to be replicated in other and independent COPD cohorts to confirm findings.

For both cohorts, cause of death data were taken from death certificates provided by the UK ONS. In addition to ONS, UK Biobank is also linked to the Cancer Registry, whereas ERICA is not. However, this should have no impact on causes of death that were included in these analyses as both cohorts use the same ICD-10 coding provided by ONS. However, the reliability of classifying causes of death has been questioned and post-mortem analysis may show different underlying causes of death. A systematic problem, beyond the control of those analysing death data is that generally physicians completing death certificates usually know the patient′s background, which may potentially lead to overestimation of individuals who died of pulmonary disease and miss cardiac death. The appointment of a clinical endpoint committee, similar to the TORCH study,[174] as an alternative method could have been more reliable but was not feasible.

In terms of comparing baseline CV status between the two cohorts, there were too few observations of objectives measures of arterial stiffness (i.e. augmentation index and pulse wave velocity), and carotid intima-media thickness to compare these. There were also too few observations to make a comparison in the experienced shortness of breath, being a symptomatic marker of disease impact. Considering other lung function measures were worse for ERICA participants, it is likely that individuals in ERICA would have experienced higher levels of breathlessness than those in UK Biobank.

This study has several potential limitations. Analyses were not adjusted for comorbidities,

therefore not addressing any competing risks of other diseases associated with early mortality than COPD. Apart from lung function measures and standard socio-demographics such as age and sex, both cohorts captured different baseline measures making the inclusion of additional covariates in the analyses difficult. In the UK Biobank, of the full white European ancestry cohort (n = 472 866), about 1200 cases had a technical failure when completing spirometry, another ∼1250 cases had unknown reasons for not completing this test, and almost another 1000 individuals could not have their lung function measured. These individuals may have systematically differed from the population included in the analyses and may have had worse lung function. Analyses were neither adjusted for drug treatment or smoking status at baseline, potentially inducing bias. For example, the intake of drugs related to treating CV disease may be associated with either increasing or decreasing the risk of CV disease.[79,228,233,243,295] At baseline 56% of individuals in the ERICA study were taking CV drugs indicating their CV risk was already being addressed before study enrolment and may explain the lower incidence of CV related disease. In addition, disease management such as maintenance medications has changed over time,[22] and there are treatment differences between those with and without COPD when admitted to hospital for CV related events.[249]

Despite the literature indicating a third of deaths in COPD to be related to cardiac disease, age-standardised mortality rates do not support the high incidence of cardiac-related death in COPD. Leading causes of death in COPD were either cancer or pulmonary disease related, and did not differ that much from the non-COPD population, which had also cancer as the leading cause of death. Also, data from the UK Biobank evaluating cause-specific death in the general population found 25% of men died of causes related to CV disease with 53% of cancer and 6% of respiratory disease.[103] In women only 12% died of CV disease with death attributed to cancer (69%) as the main cause. About 4% died of respiratory disease. Besides reporting standardised event rates by sex, which are more useful and reliable when analysing these outcomes instead of proportions, and especially when comparing findings between studies, over the past few decades

cardiac-related death in the UK general population has halved for both sexes.[224] The reduction in CV related death is primarily the result of prevention and improved heart disease treatment and management.[16,76] Also, where CV disease hospitalisation increased, the number of cardiac deaths declined by nearly 70% between 1980-2013 with similar declines in coronary heart disease and stroke.[21] Our data may indicate a similar reduction in CV related death in COPD. Finally, our results also add to the evidence of the importance of providing appropriate health care interventions with consideration of sex differences, including increasing the awareness of COPD in women and tailoring treatment strategies for prevention and treatment.[136]

5.5 Conclusions

Findings in the UK Biobank indicate cancer to be the leading cause of death in COPD with a lower cardiac-death rate than expected. Our findings may reflect the downward trend in fatal CV disease incidence/ improved CV survival in COPD, and an increased prevalence of death from other causes especially cancer and respiratory disease (i.e. trumped by the speed of progression of these diseases). Chronic obstructive pulmonary disease is, however, associated with higher all-cause mortality, including cardiac-specific death. Differences in fatal and non-fatal CV disease event rates, and the associated risk by sex emphasises the importance of tailored COPD treatment and management.

6

Exercise capacity traits and their association with COPD exacerbations requiring hospital admission: UK ERICA cohort linked with national hospital statistics

Chapter summary

Background Acute exacerbations of chronic obstructive pulmonary disease (COPD) frequently result in hospitalisation. Few reliable predictors exist, the strongest being exacerbation history. Improving exercise capacity is known to reduce rate of hospitalisations. Our aim was to assess the associations between musculoskeletal measures and risk and duration of acute exacerbation of COPD requiring hospital admission.

Methods Clinical data from the Evaluation of the Role of Inflammation in Chronic Airways

disease (ERICA) cohort were linked with hospital episode statistics capturing acute exacerbation of COPD-related admissions. Negative binomial regression was used to evaluate associations of musculoskeletal measures: six-minute walk distance, short physical performance battery and its components, and quadriceps muscle strength with hospitalised acute exacerbation of COPD and hospital length of stay for acute exacerbation of COPD.

Findings Of 714 individuals with COPD, 291 individuals experienced 762 hospitalised acute exacerbation of COPD during five-year follow up. Poorer performance of musculoskeletal measures was associated with rate or longer duration of acute exacerbation of COPD. Six-minute walk distance (incidence risk ratio (IRR) 1.67 per 30m decrements, 95% CI 1.42-1.97), lower short physical performance battery score (IRR 1.08 per 1 point decrease, 95% CI 1.01-1.14) and weaker quadriceps maximum voluntary contraction (IRR 1.02 per 1 kg decrease, 95% CI 1.00-1.03) were associated with rates of hospitalised acute exacerbation of COPD. Similar associations were observed for acute exacerbations of COPD-related hospital length of stay.

Interpretation Musculoskeletal measures were significantly associated with rate and duration of hospitalised acute exacerbation of COPD. Physical capacity should be considered an important treatable trait in reducing risk of hospitalised acute exacerbation of COPD, its assessment incorporated in risk indices evaluating future exacerbation risk, and its improvement should form a part of routine care for COPD.

6.1 Background

Acute exacerbations of COPD (AECOPD) are acute episodic flare-ups that often lead into hospital admission and are associated with high mortality and morbidity. According to 2016-17 statistics of the National Health Services (NHS) Digital, more than 128,000 individuals with a specific code for COPD exacerbation (ICD-10 J44.0, J44.1, J44.8, J44.9) in the United Kingdom were admitted to hospital, of which 97% were emergency admissions with a median hospital length of stay of three days.[197]

Overall, there is a paucity of validated and reliable measures predicting risk of AECOPD. Prior exacerbation history is currently the strongest factor reliably predicting future risk of AECOPD including hospital admissions.[131,234] Blood biomarkers such as fibrinogen[179] and white cell count,[260] have been associated with a higher risk of AECOPD, but are not routinely used in clinical practice. It has been proposed to use predictive models to improve discriminative ability and identify high-risk individuals for AECOPD in an early stage. In 2017, a systematic review of published prediction models for AECOPD concluded that out of thirty prediction models none but two were validated; all but one failed to meet practical applicability.[113]

Exercise capacity is a promising marker of early deconditioning and is strongly associated with a higher risk of mortality.[104,281] There is a possibility that such exercise traits may help identifying those people who have already started deconditioning at an early stage of disease and may experience accelerated disease progression. In particular, AECOPD also contributes to a decline in exercise capacity.[125] The relationship of shorter 6MW distance with a higher risk of mortality and exacerbation risk has been well-established.[220,248] Further, there is a small body of evidence indicating that other musculoskeletal measures used in assessment of exercise capacity such as the short physical performance battery (SPPB) and the quadriceps maximum voluntary contraction (QMVC) are also associated with a higher risk of mortality[216,259] and all-cause hospital readmission[148] in COPD. In particular, the four-metre gait speed (4MGS) test, a component of SPPB, was found to predict hospital readmission after AECOPD.[148] With the exception of the 6MW distance,[248] to our knowledge, there is no evidence of studies that have examined the association between exercise capacity measures such as SPPB, QMVC, and AECOPD requiring hospital admission.

There are scientific and practical considerations for expanding our knowledge of musculoskeletal measures and their role in predicting COPD outcomes. Each of these measures tests different properties of exercise, including cardiovascular capacity, lower limb function or musculoskeletal weakness more specifically, coordination and balance. These different exercise capacity components may reflect the impact of co-morbidity in COPD beyond those traditionally captured.

Practical applicability in routine practice is another important consideration of expanding the repository of exercise tests. Compared to the 6MW test, SPPB and its components are faster and easier to complete in clinical practice, requiring a four-metre flat surface, chair and stopwatch only. In addition, to facilitate generalisability of musculoskeletal measures in primary care, it is important to evaluate their associations with COPD outcomes in different COPD populations.

Our primary aim was to evaluate the relationship between exercise capacity assessed with musculoskeletal measures and risk of hospital admissions due to AECOPD. Further, we aimed to determine a relationship between musculoskeletal measures and length of hospital stay for initial AECOPD. To address these questions, we used a novel study design approach of combining routinely collected hospital electronic health record data with a prospective COPD disease cohort recruited in the Evaluation of the Role of Inflammation in Chronic Airways disease (ERICA) cohort.

6.2 Methods

6.2.1 Study design and participants

Observational data is reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement.[275] Data were used from the ERICA cohort, a multi-centre observational, non-interventional, epidemiological study with a sample size of 729 stable COPD patients (Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade II-IV).[107] Full study design and participant details are available in the published ERICA cohort protocol.[184] Patient level cohort data were linked to hospital admission data obtained from the NHS admitted patient care dataset, Hospital Episodes Statistics (HES) in England, Scotland and Wales, which captures all hospital admissions for AECOPD, since cohort baseline visit until November 2017. Analyses were limited to five years of follow-up. Baseline data were collected between December 2011 and January 2014. Demographic, clinical and patient reported measures and biomarkers including musculoskeletal measures were collected at baseline. Prior

exacerbation history was defined as self-reported antibiotics and/or steroids use in the previous year (twelve months before baseline). Body mass index (BMI) was categorised according to the World Health Organization.[291] Disease severity was defined by GOLD stage and estimated as described by the GOLD.[107] Productive cough (i.e. phlegm) was defined using questionnaire data and considered a surrogate marker of inflammation. Where there was missing data for productive cough, data from the phlegm question of the St. George Respiratory Questionnaire for COPD (SGRQ-C), and the COPD Assessment Test (CAT) were used. Productive cough was dichotomised (never vs. other).

6.2.2 Study outcomes

The primary outcome measure was hospital admission for AECOPD. These data were first cleaned for episode status and inpatient (i.e. hospitalised) AECOPD episodes were identified using validated criteria (**Table 6.1**, page 158).[230] Acute exacerbations of COPD were extracted from both primary and secondary positions of international classification of diseases and related health problems 10^{th} revision coding (ICD-10). Only so-called definite and possible hospitalised AECOPD were considered for this analysis (**Figure 6.1**, page 159). Priority was given to definite AECOPD. Only episodes during the study follow-up were evaluated. Admission and discharge dates were used to determine hospital length of stay (i.e. number of days) for initial AECOPD.

6.2.3 Potential predictor variables

All significant variables reported by Hurst *et al.*[132] and musculoskeletal markers captured in the ERICA cohort were considered. A full list of predictor variables is shown in **Table 6.2**, page 159 including demographics, lung function measurements, blood markers, questionnaire data, and exercise capacity traits. Measures of particular interest were SPPB and its components (i.e. 4MGS, balance, chair stand), QMVC, and 6MW distance. Exacerbation history was dichotomised (0 vs. $\geq$1).

Table 6.1: ICD-10 codes to ascertain acute exacerbation in COPD in the hospital episode statistics.

End point	ICD-10 codes	Disease/ Category	Use to ascertain AECOPD usage
J22	Lower respiratory tract infection	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J41	Simple and mucopurulent chronic bronchitis	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J41.0	Simple chronic bronchitis	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J41.1	Mucopurulent chronic bronchitis	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J41.8	Mixed simple and mucopurulent chronic bronchitis	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J42	Unspecified chronic bronchitis	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J43	Emphysema	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J43.0	MacLeod's syndrome	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J43.1	Panlobular emphysema	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J43.2	Centrilobular emphysema	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J43.8	Other emphysema	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J43.9	Emphysema, unspecified	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J44	Other chronic obstructive pulmonary disease	Possible	Use if COPD diagnosed in primary care data in First position of any finished consultant episode
J44.0	Chronic obstructive pulmonary disease with acute lower respiratory infection	Definite	Any position of any finished consultant episode as per validation study
J44.1	Chronic obstructive pulmonary disease with acute exacerbation, unspecified	Definite	Any position of any finished consultant episode as per validation study
J44.8	Other specified chronic obstructive pulmonary disease	Possible	Ditto
J44.9	Chronic obstructive pulmonary disease, unspecified	Possible	First position of any finished consultant episode as per validation study
J45	Asthma	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J45.0	Predominantly allergic asthma	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J45.1	Nonallergic asthma	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J45.8	Mixed asthma	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J45.9	Asthma, unspecified	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J45	Asthma	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J46	Status asthmaticus	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J47.0	Bronchiectasis with acute lower respiratory infection	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J47.1	Bronchiectasis with (acute) exacerbation	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J47.9	Bronchiectasis, uncomplicated	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J96.0	Acute respiratory failure	Potential	If increased sensitivity required, use if COPD diagnosed in primary care data in First position of any finished consultant episode
J96.2	Acute and chronic respiratory failure	Potential	Use to ascertain AECOPD usage

Abbreviations: ICD-10, international disease classification tenth edition. AECOPD, acute exacerbation of chronic obstructive pulmonary disease.

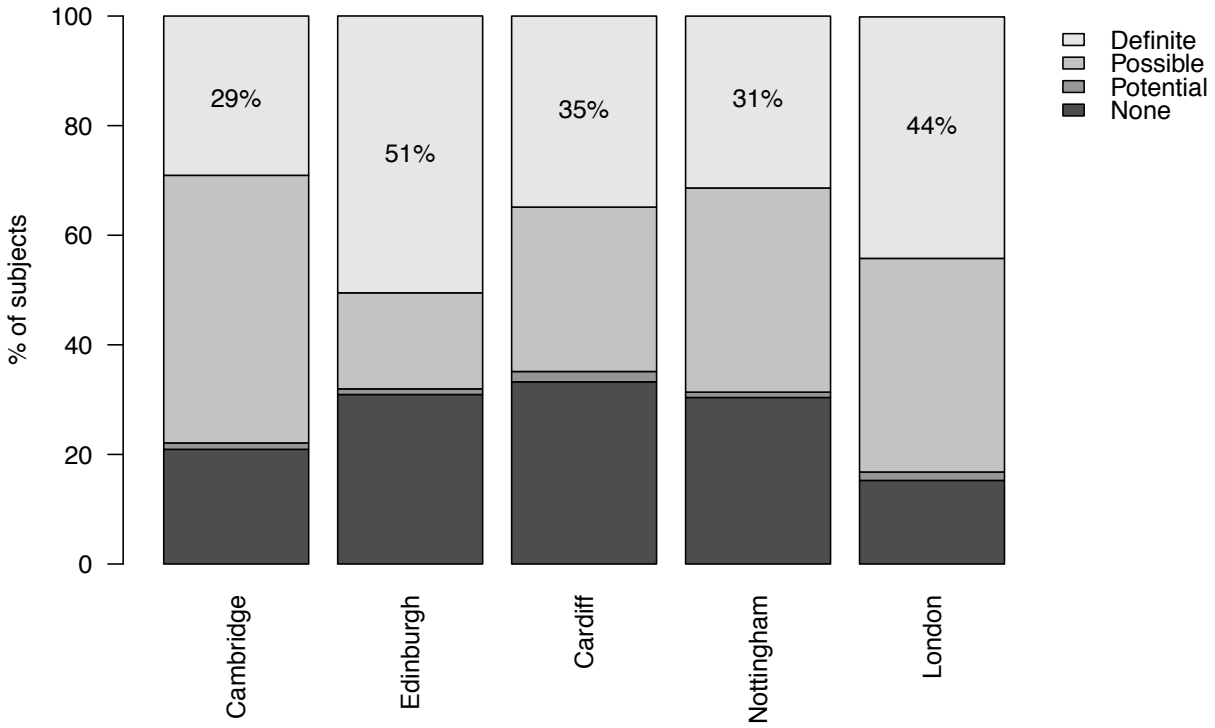


Figure 6.1: Type of acute exacerbation of COPD, by recruitment centre. Hospital admission data obtained from the National Health Service (NHS) Digital, NHS Wales, and NHS Scotland.

Table 6.2: Covariates considered

Description	**Lung function**	**Biochemical measures**	**Cardiovascular status**	**Questionnaires**	**Musculoskeletal measures**
Age	FEV_1	Glucose	Resting heart rate	SGRQ-C	6MW distance
Sex	Smoking status	Fibrinogen		CAT	SPPB
BMI	Exacerbation history	CRP			4MGS
	Phlegm	GFR			Balance
		Neutrophils			Chair stand
		Haemoglobin			QMVC
		Total cholesterol			

White cell count and Medical Research Council dyspnoea score were omitted due to collinearity with musculoskeletal measures. *Abbreviations*: BMI, body mass index. FEV_1, forced expiratory volume in one second. CRP, C-reactive protein. GFR, glomerular filtration rate. SGRQ-C, St. George respiratory questionnaire for COPD. CAT, COPD assessment test. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction.

6.2.4 Statistical analysis

Missing values were present. Only complete cases were considered. Negative binomial regression was used to examine the association between musculoskeletal measures and (i) the rate of AECOPD within the study period, and (ii) length of hospital stay (per day). Analyses were adjusted for exposure times (time between baseline visit date and earliest of death, or end of study period). Regression estimates are presented as incidence-rate ratios (IRR). Markers transformed on the natural log scale were exponentiated by a factor of 0.736 to represent a two-fold increase in risk.

Relationships between baseline variables were quantified using Spearman′s pair-wise correlations; values <0.30 were considered weak, 0.30-0.50 as moderate, and >0.50 as strong (**Figure 3.30**, page 83).[51] All analyses were stratified by recruitment site, and adjusted for age and sex. Further analyses were adjusted for BMI, smoking status, and covariates found to be of significance in the main multivariate model by Hurst *et al.*,[132] namely exacerbation history (previous year), forced expiratory volume in one second (FEV_1) measured in litres, and productive cough. Covariates were tested for collinearity resulting in the omission of Medical Research Council (MRC) dyspnoea score and white cell count (WCC). Predictors for the final analyses were derived sequentially, firstly estimating the association of each individual variable fully adjusted, following stepwise regression including the significant variables only, whilst considering collinearity and clinical utility. Only predictors with a significance level above α 0.1 for backward selection and α 0.05 for forward stepwise selection were considered. For each stepwise regression model, likelihood ratio tests were conducted to determine if independent variables should remain in the model or not, and the maximum number of variables considered in each model were based on the least number of events.[266]

As sex and exacerbation history can act as effect modifiers, in sensitivity analyses, we explored analysis stratified by these factors and tested for interactions.

6.3 Findings

6.3.1 Missing data

Missing values were present and described in **Figures 6.2 6.3**, pages 161-162. Those with missing values for 6MW distance (n = 31) had a higher rate of AECOPD-related hospital admission (p = 0.047).

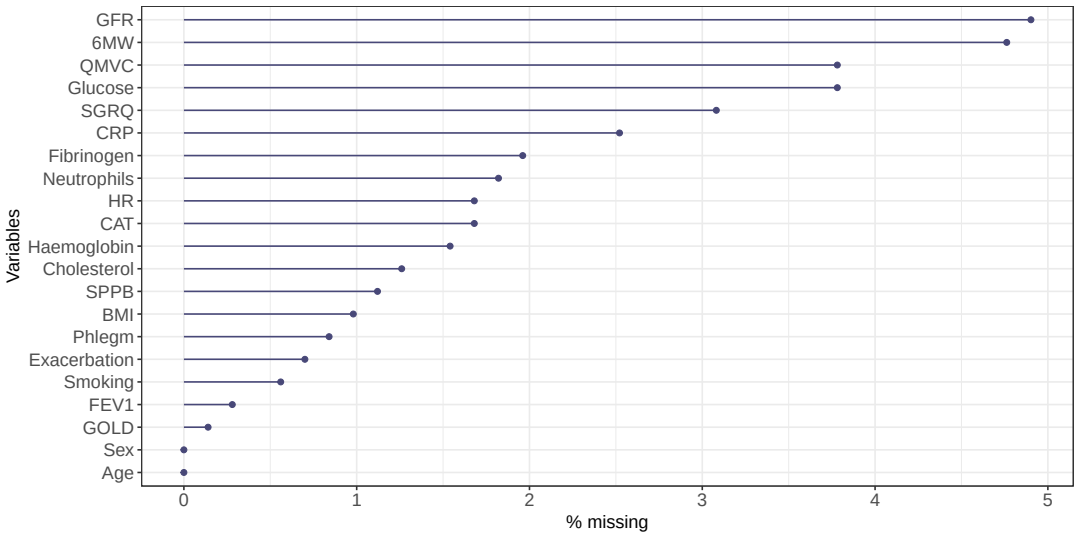


Figure 6.2: Percentage of missing values. *Abbreviations*: GFR, glomerular filtration rate. 6MW, six-minute walk. QMVC, quadriceps maximum voluntary contraction. CRP, C-reactive protein. HR, heart rate. CAT, COPD assessment test. SPPB, short physical performance battery. BMI, body mass index. FEV_1, forced expiratory volume in one second. GOLD, global initiative for obstructive lung disease.

6.3.2 Descriptive statistics

The mean number of acute exacerbations for COPD were 1.3 vs. 0.6 with a variance of 6.2 vs. 2.6 for those with an exacerbation history compared to those without, indicating over-dispersed count data. In total, 714 individuals with stable COPD were included in the analysis, of whom 291 (41%) experienced at least one hospital admission for AECOPD during the study follow-up; 159 (22%) had multiple events (**Figure 3.4**, page 64). The resulting event rate for hospitalised AECOPD was 11 events (95% CI 10-13) per 100 person-years (**Figure 6.4**, page 163). Overall,

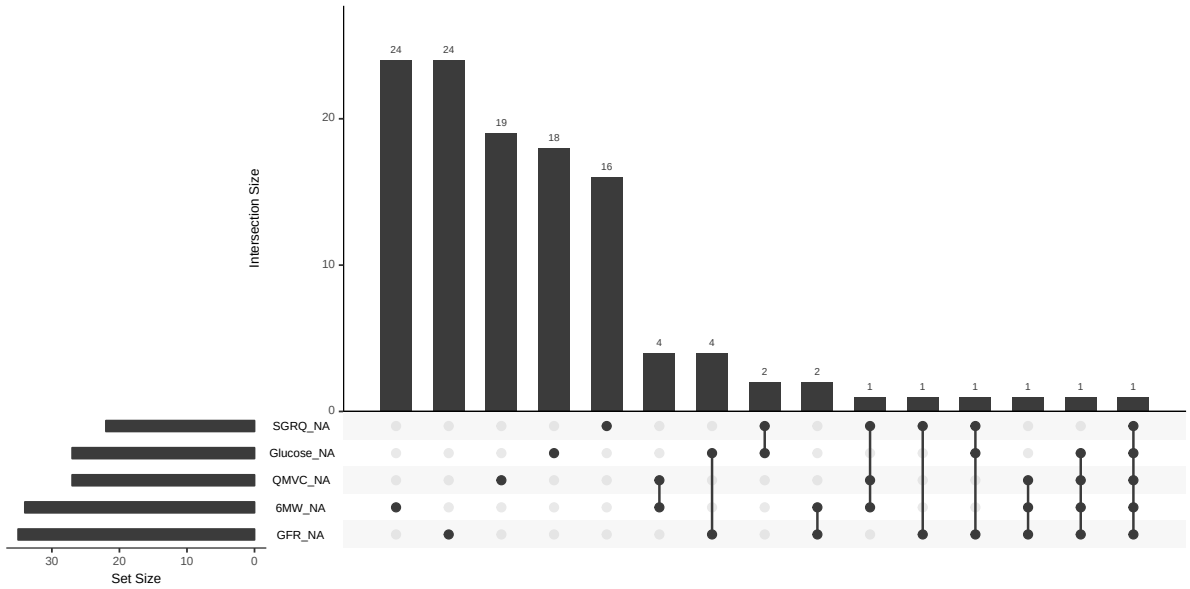


Figure 6.3: Missing data patterns. *Abbreviations*: SGRQ, St. George respiratory questionnaire for COPD. QMVC, quadriceps maximum voluntary contraction. 6MW, six-minute walk. GFR, glomerular filtration rate.

127 (18%) individuals died and, of these, the majority 103 (81%) had died following hospitalised AECOPD. At baseline, the mean age of the cohort was 67 ± 8 years with 61% males. A third of the cohort was overweight, another third obese. Exacerbations during the year prior to baseline were reported by 67% individuals with a corresponding mean of 2 (interquartile range (IQR; 1-4) events per person-year. Mean FEV_1 was 1.3 ± 0.5 litre with third current smokers. About half of the cohort (51%) experienced breathlessness on exercise (MRC grade ≥3) and 46% had productive cough on most mornings (**Table 6.3**, page 165). Median length of hospital stay for initial AECOPD-related admission was 3 (IQR 1-7) days.

For those readmitted, the median time to hospital readmission was 179 (54-421) days, of whom 65 individuals (41%) were readmitted to hospital within 90 days after initial admission and had a median length of stay of 3 (2-7) days. Those with an exacerbation history at baseline were younger ($p = 0.027$), male ($p < 0.001$), had lower forced expiratory volume in one second ($p < 0.001$), worse dyspnoea scores ($p = 0.002$), and higher inflammatory levels of fibrinogen ($p = 0.001$) and C-reactive protein ($p = 0.019$) compared to those without. Shorter walking

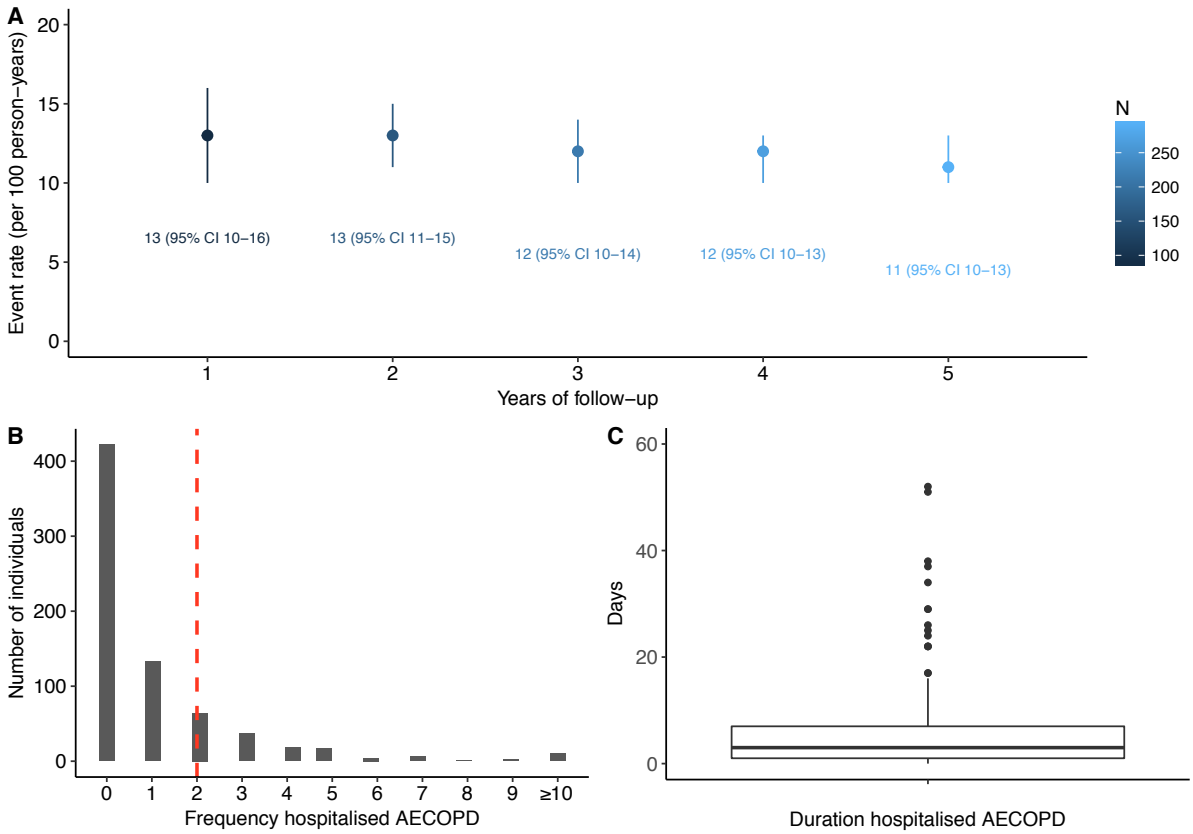


Figure 6.4: Yearly event rates, frequency and duration of AECOPD-related hospital admission. Figures display (A) mean event rates with 95% confidence intervals per 100 person-years during study period, (B) AECOPD frequency, and (C) AECOPD duration. Depth of blue indicates the cumulative number of individuals with first AECOPD during the study period: 1 year (n = 86), 2 years (n = 160), 3 years (n = 213), 4 years (n = 266) and 5 years of follow-up (n = 291). Red dashed line indicates the median number of hospital admissions for AECOPD amongst those experienced an AECOPD. *Abbreviations*: N, indicates the number of participants. AECOPD, acute exacerbation of chronic obstructive pulmonary disease.

distance (p <0.001), lower SPPB scores (p = 0.003), or the components 4MGS (p <0.001) and chair stand (p = 0.003) but not balance (p = 0.630), and QMVC (p <0.001) were also reported for those with an exacerbation history at baseline.

6.3.3 Factors associated with rate of AECOPD-related hospital admission

Musculoskeletal measures of 6MW distance, SPPB and its 4MGS, and chair stand components, and QMVC were associated with the risk of AECOPD-related hospital admission (**Figure 6.5** and **Table 6.4**, pages 167-166). Balance was not associated with the outcome. Six-minute walk distance (IRR 1.13 per 30 metre decrease, 95% CI 1.08 to 1.17, p <0.001), FEV_1 (IRR 0.84 per 100 ml increase, 95% CI 0.81 to 0.86, p <0.001) or disease severity measured by GOLD (IRR 2.51 per increase to next stage, 95% CI 2.04 to 3.10, p <0.001), and males (IRR 2.41, 95% CI 1.77 to 3.29, p <0.001) had the highest associated IRRs. Stepwise regression, including variables fully adjusted and significantly associated with AECOPD-related hospital admission rate only, retained the following predictors: males (IRR 2.14, 95% CI 1.55 to 2.96, p <0.001), FEV_1, (IRR 0.88 per 100 ml increase, 95% CI 0.85 to 0.91, p <0.001), exacerbation history ≥1 (IRR 1.96, 95% CI 1.39 to 2.76, p <0.001), CAT (IRR 1.03 per 1 point increase, 95% CI 1.01 to 1.05, p = 0.010), resting heart rate (IRR 1.01 per 1 bpm increase, 95% CI 1.00 to 1.03, p = 0.025), and 6MW distance (IRR 1.08 per 30 metre decrease, 95% CI 1.04 to 1.12, p <0.001; **Table 6.5**, page 167).

6.3.4 Factors associated with AECOPD-related hospital stay

Including data from individuals admitted to hospital only (n = 291), multivariable analysis identified multiple markers to be associated with AECOPD-related hospital stay (**Figure 6.6** and **Table 6.8**, pages 169-170). All musculoskeletal measures, except for QMVC were associated with longer AECOPD-related hospital stay (**Table 6.4** and **Figure 6.5**, pages 166-167). Age (IRR 1.83 per 10 year increase, 95% CI 1.48 to 2.26, p <0.001), 6MW (IRR 1.14 per 30 metre decrease, 95% CI 1.08 to 1.20, p <0.001), and SPPB (IRR 1.18 per 1 point decrease, 95% 1.10

Table 6.3: Baseline characteristics.

Characteristic	Total	Without AECOPD	With AECOPD
Description		**Median (IQR) or n (%)**	
Age (years)	67 (62-73)	68 (62-74)	67 (63-72)
Sex, n (%)			
Male	434 (61)	262 (62)	172 (59)
Female	280 (39)	160 (38)	120 (41)
Body mass index (kg/m^2)	27 (23-31)	27 (24-31)	26 (23-31)
Musculoskeletal measures			
6MW distance (metre)	366 (255-440)	398 (298-462)	326 (210-404)
SPPB (0-12)	10 (8-11)	10 (8-12)	10 (8-11)
No functional limitation, $\geq$10, n (%)	414 (58)	254 (61)	160 (55)
Functional limitation $<$10, n (%)	292 (41)	163 (39)	129 (45)
- 4MGS score (0-4)	4 (3-4)	4 (3-4)	4 (3-4)
- Balance points (0-4)	4 (4-4)	4 (4-4)	4 (4-4)
- Chair stand score (0-4)	3 (1-4)	3 (1-4)	2 (1-3)
QMVC peak (kg), median (IQR)	30 (22-39)	31 (23-40)	28 (20-35)
Lung function			
FEV_1	1.3 (0.9-1.7)	1.5 (1.1-1.9)	1.1 (0.8-1.4)
Smoking status, n (%)			
Current	218 (31)	131 (31)	87 (30)
Former	492 (69)	291 (69)	201 (70)
GOLD, n (%)			
Grade II	406 (57)	291 (69)	115 (40)
Grade III	240 (34)	112 (27)	128 (44)
Grade IV	68 (10)	19 (5)	48 (16)
Exacerbation history, 1 year ($\geq$1)	473 (66)	247 (59)	226 (79)
Phlegm, n (%)			
Never	46 (7)	237 (57)	144 (50)
Other	662 (94)	181 (43)	146 (50)
Biochemical measures			
log Glucose (mmol/L)	1.59 (1.50-1.69)	4.9 (4.5-5.3)	4.9 (4.5-5.4)
log Fibrinogen (g/dL)	1.22 (1.06-1.36)	1.19 (1.03-1.36)	1.25 (1.10-1.39)
log C-reactive protein (mg/L)	1.21 (0.47-2.00)	1.10 (0.48-1.85)	1.39 (0.43-2.19)
GFR (mL/min/1.73 m^2)	87 (76-101)	87 (77-100)	88 (76-102)
Neutrophils (mm^3)	4.5 (3.6-5.6)	4.3 (3.5-5.5)	4.8 (3.7-5.7)
Haemoglobin (g/dL)	14.3 (13.4-15.3)	14.3 (13.4-15.2)	14.4 (13.4-15.4)
Total cholesterol (mmol/L)	5.0 (4.3-5.8)	5.0 (4.2-5.7)	5.0 (4.3-5.9)
Cardiovascular status			
Heart rate (bpm)	74 (66-82)	72 (65-81)	77 (67-84)
Questionnaires			
SGRQ-C (0-100)	51 (34-66)	43 (29-61)	57 (45-71)
CAT (0-40)	20 (13-26)	18 (12-24)	22 (17-28)

Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of study participants are included. *Abbreviations*: MRC, Medical Research Council. GOLD, global initiative for obstructive lung disease. WCC, white cell count. SGRQ-C, St George's respiratory questionnaire for COPD. CAT, COPD assessment test. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction.

Table 6.4: Adjusted multivariable associations with frequency of AECOPD-related hospital admission.

Baseline Characteristics	5 year (n = 714, of whom 291 had AECOPD)			
	Incidence risk ratio (95% CI). Adjusted for age and sex [a]	P value [c]	Incidence risk ratio (95% CI). Multivariable adjusted [b]	P value [c]
Description				
Age - per 10 year increase	0.97 (0.81 to 1.17)	0.773	0.88 (0.74 to 1.04)	0.126
Sex - male	1.01 (0.74 to 1.39)	0.932	2.41 (1.77 to 3.29)	<0.001
Body mass index - per 1 point increase	0.96 (0.94 to 0.98)	0.002	1.00 (0.98 to 1.02)	0.947
Lung function				
FEV_1 - per 100 ml increase	0.83 (0.80 to 0.85)	<0.001	0.84 (0.81 to 0.86)	<0.001
Smoking status - current	1.25 (0.89 to 1.76)	0.192	1.15 (0.84 to 1.57)	0.382
GOLD stage - per increase to next stage	2.71 (2.21 to 3.33)	<0.001	2.51 (2.04 to 3.10)	<0.001
Exacerbation history (1 year), ≥1	2.52 (1.79 to 3.53)	<0.001	1.94 (1.40 to 2.67)	<0.001
Productive cough - yes	1.90 (1.00 to 3.61)	0.049	1.04 (0.79 to 1.38)	0.768
Biochemical measures				
Glucose - per 1 log unit increase	1.43 (0.53 to 3.87)	0.477	1.77 (0.69 to 4.53)	0.231
Fibrinogen - per 1 log unit increase	3.43 (1.71 to 6.88)	0.001	1.95 (1.03 to 3.68)	0.04
CRP - per 1 log unit increase	1.18 (1.03 to 1.35)	0.018	1.10 (0.98 to 1.25)	0.116
GFR - per 1 unit increase	1.01 (1.00 to 1.02)	0.093	1.00 (0.99 to 1.01)	0.621
Neutrophils - per 1 unit increase	1.22 (1.12 to 1.33)	<0.001	1.14 (1.05 to 1.24)	0.001
Haemoglobin - per 1 unit increase	0.98 (0.89 to 1.09)	0.722	0.96 (0.88 to 1.06)	0.429
Total cholesterol - per 1 unit increase	1.00 (0.87 to 1.15)	0.955	0.93 (0.82 to 1.06)	0.269
Cardiovascular status				
Heart rate - per 1 bpm increase	1.04 (1.02 to 1.05)	<0.001	1.02 (1.01 to 1.03)	<0.001
Questionnaire data				
SGRQ-C - per 4 point increase	1.13 (1.10 to 1.17)	<0.001	1.07 (1.03 to 1.10)	<0.001
CAT - per 1 point increase	1.09 (1.07 to 1.11)	<0.001	1.05 (1.03 to 1.07)	<0.001
Musculoskeletal measures				
Six-minute walk distance - per 30 metre decrease	1.19 (1.15 to 1.24)	<0.001	1.13 (1.08 to 1.17)	<0.001
SPPB score (0-12) - per 1 point decrease	1.14 (1.07 to 1.22)	<0.001	1.08 (1.01 to 1.14)	0.019
Functional limitation (SPPB) - yes	1.68 (1.21 to 2.33)	0.002	1.22 (0.91 to 1.64)	0.179
4MGS score (0-4) - per 1 point decrease	1.46 (1.20 to 1.76)	<0.001	1.19 (1.00 to 1.41)	0.048
Balance score (0-4) - per 1 point decrease	1.11 (0.93 to 1.33)	0.246	1.07 (0.91 to 1.25)	0.434
Chair stand score (0-4) - per 1 point decrease	1.25 (1.11 to 1.40)	<0.001	1.14 (1.02 to 1.26)	0.016
QMVC peak - per 1 kg decrease	1.05 (1.03 to 1.07)	<0.001	1.02 (1.00 to 1.03)	0.039

Incidence rate ratios were estimated based on negative binomial regression. All analyses were adjusted for recruitment site. [a] Adjusted for age and sex [b] Further adjusted for body mass index, smoking status, forced expiratory volume in one second, productive cough, and exacerbation history. [c] P values based on negative binomial regression.

Variables MRC dyspnoea score and white cell count were omitted due to collinearity. *Abbreviations*: CI, confidence intervals. FEV_1, forced expiratory volume in one second. GOLD, global initiative for obstructive lung disease. GFR, glomerular filtration rate. SGRQ-C, St. George respiratory questionnaire for COPD. CAT, COPD assessment test. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction.

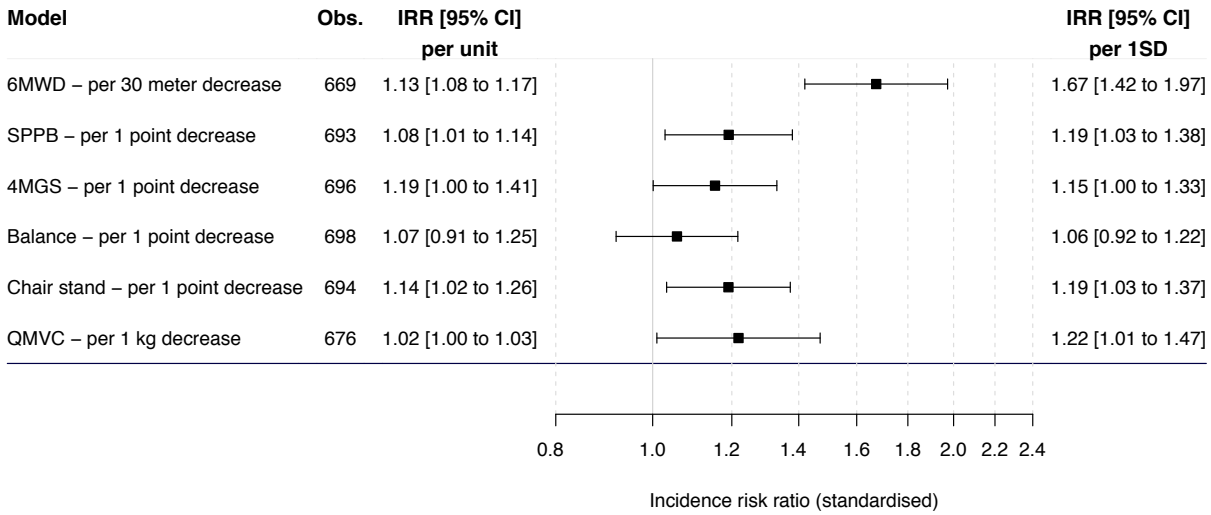


Figure 6.5: Associations of baseline musculoskeletal measures and rate of hospitalised acute exacerbation of chronic obstructive pulmonary disease in the ERICA cohort. Risk indicated as incidence risk ratios (IRR). Estimates derived using negative binomial regression. Analyses adjusted for recruitment site. Age, sex, body mass index, smoking status, forces expiratory volume in one second, productive cough, and exacerbation history were included as covariates. *Abbreviations*: Obs, number of observations included in analysis. IRR, incidence risk ratios. CI, confidence intervals. SD, standard deviation. 6MWD, six-minute walk distance. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction.

Table 6.5: Factors associated with rate of AECOPD-related hospital admission in the stepwise multivariable model.

	Stepwise selection (n = 610)	
Factor	**IRR (95% CI)**	**P value**
Sex - male	2.14 (1.55 to 2.96))	<0.001
FEV_1 - per 100 ml increase	0.88 (0.85 to 0.91)	<0.001
Exacerbation history, ≥1	1.96 (1.39 to 2.76)	<0.001
CAT - per 1 point increase	1.03 (1.01 to 1.05)	0.010
Resting heart rate – per 1 bpm increase	1.01 (1.00 to 1.03)	0.025
6MW distance – per 30 metre decrease	1.08 (1.04 to 1.12)	<0.001

Adjusted for recruitment site. *Abbreviations*: IRR, incidence risk ratios. CI, confidence intervals. FEV_1, forced expiratory volume in one second. CAT, COPD assessment test. 6MW, six-minute walk.

Table 6.6: Factors associated with rate of AECOPD-related hospital admission, by exacerbation history.

Factor	Exacerbation history (n = 439) IRR (95% CI)	P value	No exacerbation history (n = 222) IRR (95% CI)	P value	P-value ¶
Sex – male	2.03 (1.43 to 2.87)	<0.001	3.80 (1.75 to 8.26)	0.001	0.029
FEV_1, – per 100 ml increase	0.87 (0.83 to 0.90)	<0.001	0.89 (0.83 to 0.95)	0.001	0.896
Exacerbation history, ≥1	N/A	N/A	N/A	N/A	N/A
CAT – per 1 point increase	1.02 (1.00 to 1.05)	0.029	1.04 (1.00 to 1.09)	0.051	0.865
Resting heart rate – per 1 bpm increase	1.01 (1.00 to 1.03)	0.07	1.01 (0.99 to 1.04)	0.363	0.760
6MW distance – per 30 metre decrease	1.06 (1.02 to1.11)	0.005	1.14 (1.05 to 1.23)	0.002	0.174

¶P values of interaction with exacerbation history. Adjusted for recruitment site. *Abbreviations*: IRR, incidence risk ratios. CI, confidence intervals. FEV_1, forced expiratory volume in one second. CAT, COPD assessment test. 6MW, six-minute walk.

Table 6.7: Factors associated with rate of AECOPD-related hospital admission, by sex.

Factor	Female (n = 257) IRR (95% CI)	P value	Male (n = 404) IRR (95% CI)	P value	P-value ¶
Sex – male	N/A	N/A	N/A	N/A	N/A
FEV_1, – per 100 ml increase	0.86 (0.80 to 0.92)	<0.001	0.88 (0.84 to 0.91)	<0.001	0.607
Exacerbation history, ≥1	3.48 (1.70 to 7.13)	0.001	1.38 (0.95 to 1.99)	0.087	0.030
CAT – per 1 point increase	1.02 (0.99 to 1.06)	0.125	1.03 (1.00 to 1.05)	0.034	0.905
Resting heart rate – per 1 bpm increase	1.03 (1.01 to 1.05)	0.014	1.00 (0.99 to 1.02)	0.567	0.193
6MW distance – per 30 metre decrease	1.06 (1.00 to 1.13)	0.057	1.09 (1.04 to 1.14)	<0.001	0.6743

¶P values of interaction with sex. Adjusted for recruitment site. *Abbreviations*: IRR, incidence risk ratios. CI, confidence intervals. FEV_1, forced expiratory volume in one second. CAT, COPD assessment test. 6MW, six-minute walk.

to 1.27, p <0.001) were the strongest associated variables. Stepwise regression, including variables fully adjusted and significantly associated with hospital length of stay only, retained the following predictors: age (IRR 1.53 per 10 year increase, 95% CI 1.18 to 1.98, p = 0.001), BMI (IRR 0.93 per 1 point increase, 95% CI 0.90 to 0.96, p <0.001), glucose (IRR 2.89 per twofold increase, 95% CI 1.18 to 7.05, p = 0.020), and SPPB (IRR 1.19 per 1 point decrease, 95% CI 1.10 to 1.30, p <0.001; **Table 6.9**, page 171).

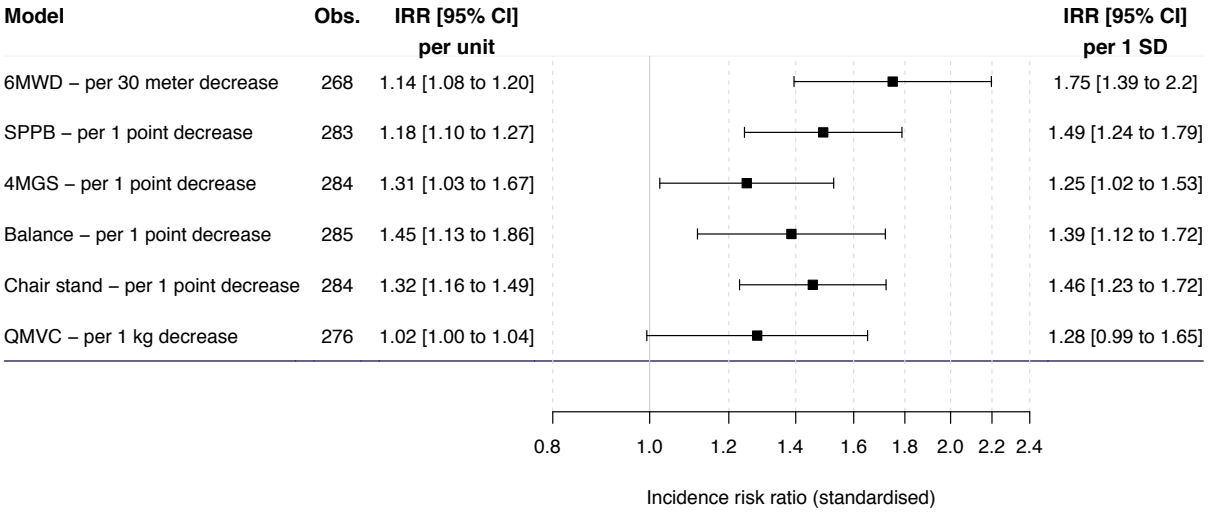


Figure 6.6: Associations of baseline musculoskeletal measures and hospital length of stay after admission for acute exacerbation of chronic obstructive pulmonary disease in the ERICA cohort. Risk indicated as incidence risk ratios (IRR). Estimates derived using negative binomial regression. Analyses adjusted for recruitment site. Age, sex, body mass index, smoking status, forces expiratory volume in one second, productive cough, and exacerbation history were included as covariates. *Abbreviations*: Obs, number of observations included in analysis. IRR, incidence risk ratios. CI, confidence intervals. SD, standard deviation. 6MWD, six-minute walk distance. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction.

6.3.5 Sensitivity analysis for rate of AECOPD-related hospital admission

Overall, IRRs were higher for men and 6MW distance for those with no exacerbation history (**Tables 6.6 6.10**, pages 168 and 172). Incidence risk ratios of exacerbation history were higher for women when stratifying by sex (**Tables 6.7 6.11**, pages 168 and 173). When testing for interactions, both exacerbation history and sex were significant.

Table 6.8: Adjusted multivariable associations with AECOPD length of stay.

Baseline characteristics	5 year (n = 291 individuals with AECOPD)			
	Incidence risk ratio (95% CI). Adjusted for age and sex [a]	P value [c]	Incidence risk ratio (95% CI). Multivariable adjusted [b]	P value [c]
Description				
Age - per 10 year increase	1.78 (1.45 to 2.20)	<0.001	1.83 (1.48 to 2.26)	<0.001
Sex - male	0.84 (0.58 to 1.21)	0.354	0.84 (0.56 to 1.26)	0.399
Body mass index - per 1 point increase	0.96 (0.94 to 0.99)	0.011	0.96 (0.93 to 0.99)	0.009
Lung function				
FEV_1 - per 100 ml increase	0.96 (0.91 to 1.00)	0.063	0.97 (0.93 to 1.02)	0.269
Smoking status - current	1.39 (0.93 to 2.09)	0.11	1.2 (0.78 to 1.87)	0.409
GOLD stage - per increase to next stage	1.14 (0.86 to 1.50)	0.374	1.15 (0.87 to 1.53)	0.335
Exacerbation history (1 year), ≥1	0.63 (0.41 to 0.97)	0.035	0.62 (0.39 to 0.97)	0.037
Productive cough - yes	0.75 (0.34 to 1.66)	0.483	1.12 (0.77 to 1.62)	0.559
Biochemical measures				
Glucose - per 1 log unit increase	7.89 (2.67 to 23.33)	<0.001	8.78 (2.81 to 27.49)	<0.001
Fibrinogen - per 1 log unit increase	2.50 (1.11 to 5.61)	0.027	3.14 (1.37 to 7.18)	0.007
CRP - per 1 log unit increase	1.07 (0.92 to 1.24)	0.407	1.14 (0.97 to 1.35)	0.107
GFR - per 1 unit increase	0.99 (0.98 to 1.00)	0.05	0.98 (0.97 to 1.00)	0.014
Neutrophils - per 1 unit increase	1.07 (0.97 to 1.18)	0.164	1.04 (0.93 to 1.16)	0.525
Haemoglobin - per 1 unit increase	0.94 (0.83 to 1.05)	0.273	0.91 (0.80 to 1.03)	0.134
Total cholesterol - per 1 unit increase	0.93 (0.81 to 1.08)	0.358	0.93 (0.79 to 1.09)	0.349
Cardiovascular status				
Heart rate - per 1 bpm increase	1.00 (0.99 to 1.02)	0.478	1.00 (0.98 to 1.01)	0.665
Questionnaire data				
SGRQ-C - per 4 point increase	1.00 (0.96 to 1.03)	0.857	1.02 (0.97 to 1.06)	0.449
CAT - per 1 point increase	0.99 (0.97 to 1.01)	0.504	1.00 (0.98 to 1.03)	0.892
Musculoskeletal measures				
Six-minute walk distance - per 30 metre decrease	1.11 (1.05 to 1.16)	<0.001	1.14 (1.08 to 1.20)	<0.001
SPPB score (0-12) - per 1 point decrease	1.15 (1.06 to 1.24)	<0.001	1.18 (1.10 to 1.27)	<0.001
Functional limitation (SPPB) - yes	1.84 (1.27 to 2.68)	0.001	2.01 (1.37 to 2.94)	<0.001
4MGS score (0-4) - per 1 point decrease	1.29 (1.01 to 1.65)	0.045	1.31 (1.03 to 1.67)	0.029
Balance score (0-4) - per 1 point decrease	1.44 (1.12 to 1.84)	0.004	1.45 (1.13 to 1.86)	0.003
Chair stand score (0-4) - per 1 point decrease	1.24 (1.09 to 1.40)	0.001	1.32 (1.16 to 1.49)	<0.001
QMVC peak - per 1 kg decrease	1.03 (1.01 to 1.05)	0.002	1.02 (1.00 to 1.04)	0.056

Incidence rate ratios were estimated based on negative binomial regression. All analyses were adjusted for recruitment site. [a] Adjusted for age and sex [b] Further adjusted for body mass index, smoking status, forced expiratory volume in one second, productive cough, and exacerbation history. [c] P values based on negative binomial regression.
Variables MRC dyspnoea score and white cell count were omitted due to collinearity. *Abbreviations*: CI, confidence intervals. FEV_1, forced expiratory volume in one second. GOLD, global initiative for obstructive lung disease. GFR, glomerular filtration rate. SGRQ-C, St. George respiratory questionnaire for COPD. CAT, COPD assessment test. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction.

Table 6.9: Factors associated with AECOPD-related hospital admission length of stay in the stepwise multivariable model.

	Stepwise selection (n = 233)	
Factor	**IRR (95% CI)**	**P value**
Age - per 10 year increase	1.53 (1.18 to 1.98)	0.001
BMI - per 1 point increase	0.93 (0.90 to 0.96)	<0.001
Glucose - per twofold increase	2.89 (1.18 to 7.05)	0.020
SPPB - per 1 point decrease	1.19 (1.10 to 1.30)	<0.001

Adjusted for recruitment site. *Abbreviations*: BMI, body mass index. SPPB, short physical performance battery. IRR, incidence risk ratio. CI, confidence interval.

Table 6.10: Adjusted multivariate associations with AECOPD frequency, by exacerbation history.

Baseline characteristics	5 year (n = 714, of whom 291 had AECOPD)			
	Exacerbation history (n = 473)		No exacerbation history (n = 236)	
	Incidence risk ratio (95% CI). Multivariable adjusted [a]	P value [b]	Incidence risk ratio (95% CI). Multivariable adjusted [a]	P value [b]
Description				
Age - per 10 year increase	0.90 (0.74 to 1.09)	0.283	0.69 (0.47 to 1.02)	0.063
Sex - male	2.05 (1.46 to 2.89)	<0.001	5.39 (2.55 to 11.41)	<0.001
Body mass index - per 1 point increase	0.98 (0.96 to 1.01)	0.128	1.08 (1.02 to 1.15)	0.01
Lung function				
FEV_1 - per 100 ml increase	0.85 (0.82 to 0.88)	<0.001	0.80 (0.75 to 0.86)	<0.001
Smoking status - current	1.07 (0.75 to 1.52)	0.723	1.10 (0.58 to 2.11)	0.762
GOLD stage - per increase to next stage	2.17 (1.73 to 2.74)	<0.001	3.74 (2.36 to 5.93)	<0.001
Exacerbation history (1 year), ≥1				
Productive cough - yes	1.14 (0.84 to 1.56)	0.408	1.23 (0.63 to 2.42)	0.539
Biochemical measures				
Glucose - per 1 log unit increase	1.71 (0.57 to 5.10)	0.337	2.50 (0.36 to 17.33)	0.353
Fibrinogen - per 1 log unit increase	1.98 (0.99 to 3.97)	0.055	1.96 (0.44 to 8.60)	0.375
CRP - per 1 log unit increase	1.18 (1.03 to 1.35)	0.018	0.90 (0.66 to 1.22)	0.5
GFR - per 1 unit increase	1.00 (0.99 to 1.01)	0.458	0.99 (0.97 to 1.01)	0.264
Neutrophils - per 1 unit increase	1.15 (1.06 to 1.26)	0.002	1.07 (0.88 to 1.30)	0.518
Haemoglobin - per 1 unit increase	0.95 (0.86 to 1.05)	0.302	1.12 (0.88 to 1.41)	0.358
Total cholesterol - per 1 unit increase	0.95 (0.83 to 1.10)	0.497	1.01 (0.72 to 1.40)	0.969
Cardiovascular status				
Heart rate - per 1 bpm increase	1.02 (1.01 to 1.03)	0.003	1.02 (0.99 to 1.05)	0.122
Questionnaire data				
SGRQ-C - per 4 point increase	1.07 (1.03 to 1.11)	0.001	1.07 (1.00 to 1.15)	0.037
CAT - per 1 point increase	1.04 (1.02 to 1.06)	<0.001	1.07 (1.02 to 1.11)	0.004
Musculoskeletal measures				
Six-minute walk distance - per 30 metre decrease	1.11 (1.06 to 1.16)	<0.001	1.16 (1.06 to 1.26)	0.001
SPPB score (0-12) - per 1 point decrease	1.11 (1.04 to 1.19)	0.003	0.94 (0.80 to 1.10)	0.435
Functional limitation (SPPB) - yes	1.40 (1.01 to 1.95)	0.046	0.81 (0.43 to 1.54)	0.524
4MGS score (0-4) - per 1 point decrease	1.28 (1.07 to 1.54)	0.008	0.82 (0.48 to 1.42)	0.476
Balance score (0-4) - per 1 point decrease	1.13 (0.95 to 1.35)	0.175	0.71 (0.45 to 1.12)	0.14
Chair stand score (0-4) - per 1 point decrease	1.18 (1.05 to 1.33)	0.006	1.02 (0.82 to 1.28)	0.827

Incidence rate ratios were estimated based on negative binomial regression. All analyses were stratified by recruitment centre and exacerbation history. [a] Adjusted for age, sex, body mass index, smoking status, forced expiratory volume in one second, and productive cough. [b] P values based on negative binomial regression.

Variables MRC dyspnoea score and white cell count were omitted due to collinearity. *Abbreviations*: CI, confidence intervals. FEV_1, forced expiratory volume in one second. GOLD, global initiative for obstructive lung disease. GFR, glomerular filtration rate. SGRQ-C, St. George respiratory questionnaire for COPD. CAT, COPD assessment test. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction.

Table 6.11: Adjusted multivariate associations with AECOPD frequency, by sex.

Baseline characteristics	5 year (n = 714, of whom 291 had AECOPD)			
	Male (n = 434)		Female (n = 280)	
	Incidence risk ratio (95% CI) [a]	P value [b]	Incidence risk ratio (95% CI) [a]	P value [b]
Description				
Age - per 10 year increase	0.89 (0.72 to 1.11)	0.313	0.77 (0.58 to 1.02)	0.073
Sex - male				
Body mass index - per 1 point increase	1.02 (0.99 to 1.06)	0.237	0.98 (0.95 to 1.01)	0.242
Lung function				
FEV_1 - per 100 ml increase	0.84 (0.81 to 0.87)	<0.001	0.81 (0.76 to 0.87)	<0.001
Smoking status - current	0.87 (0.57 to 1.32)	0.502	1.51 (0.96 to 2.36)	0.073
GOLD stage - per increase to next stage	2.87 (2.21 to 3.73)	<0.001	2.01 (1.41 to 2.87)	<0.001
Exacerbation history (1 year), ≥1	1.50 (1.03 to 2.18)	0.033	3.45 (1.78 to 6.67)	<0.001
Productive cough - yes	1.05 (0.72 to 1.51)	0.81	1.19 (0.78 to 1.82)	0.409
Biochemical measures				
Glucose - per 1 log unit increase	0.66 (0.19 to 2.31)	0.516	6.67 (1.68 to 26.56)	0.007
Fibrinogen - per 1 log unit increase	1.42 (0.61 to 3.3)	0.412	2.96 (1.08 to 8.14)	0.035
CRP - per 1 log unit increase	1.08 (0.92 to 1.26)	0.373	1.18 (0.97 to 1.45)	0.104
GFR - per 1 unit increase	1.00 (0.99 to 1.01)	0.69	1.00 (0.99 to 1.02)	0.947
Neutrophils - per 1 unit increase	1.01 (0.90 to 1.13)	0.899	1.30 (1.16 to 1.46)	<0.001
Haemoglobin - per 1 unit increase	0.96 (0.86 to 1.08)	0.525	0.93 (0.78 to 1.10)	0.383
Total cholesterol - per 1 unit increase	0.93 (0.79 to 1.10)	0.403	0.93 (0.76 to 1.13)	0.469
Cardiovascular status				
Heart rate - per 1 bpm increase	1.01 (1.00 to 1.03)	0.082	1.03 (1.01 to 1.05)	0.001
Questionnaire data				
SGRQ-C - per 4 point increase	1.06 (1.02 to 1.11)	0.004	2.40 (1.54 to 3.74)	<0.001
CAT - per 1 point increase	1.05 (1.02 to 1.08)	<0.001	1.06 (1.01 to 1.11)	0.023
Musculoskeletal measures				
Six-minute walk distance - per 30 metre decrease	1.12 (1.07 to 1.18)	<0.001	1.04 (1.01 to 1.08)	0.005
SPPB score (0-12) - per 1 point decrease	1.03 (0.94 to 1.12)	0.546	1.15 (1.05 to1.26)	0.002
Functional limitation (SPPB) - yes	1.03 (0.71 to 1.50)	0.876	1.70 (1.05 to 2.73)	0.029
4MGS score (0-4) - per 1 point decrease	1.09 (0.84 to 1.40)	0.527	1.40 (1.10 to 1.78)	0.007
Balance score (0-4) - per 1 point decrease	0.95 (0.71 to 1.27)	0.732	1.11 (0.91 to 1.35)	0.303
Chair stand score (0-4) - per 1 point decrease	1.06 (0.93 to 1.21)	0.378	1.35 (1.13 to 1.61)	0.001

Incidence rate ratios were estimated based on negative binomial regression. Analyses were stratified by recruitment centre and sex. [a] Adjusted for age, body mass index, smoking status, forced expiratory volume in one second, productive cough, and exacerbation history. [b] P values based on negative binomial regression.

¶Variables MRC dyspnoea score and white cell count were omitted due to colinearity. *Abbreviations*: CI, confidence intervals. FEV_1, forced expiratory volume in one second. GOLD, global initiative for obstructive lung disease. GFR, glomerular filtration rate. SGRQ-C, St. George respiratory questionnaire for COPD. CAT, COPD assessment test. 6MW, six-minute walk. SPPB, short physical performance battery. 4MGS, four-metre gait speed. QMVC, quadriceps maximum voluntary contraction.

6.4 Discussion

This is the first study demonstrating the potential use of SPPB in predicting AECOPD-related hospital admission. In addition, 6MW distance, known to predict mortality in individuals with COPD, was found to be associated with AECOPD-related hospital admission, after adjusting for common and known predictive covariates. Our findings show that multiple measures including multiple exercise capacity traits were associated with both AECOPD-related hospital admission rate and duration. The strongest associated measure for admission, however, was lung function measured as FEV_1, or GOLD stage as defined by FEV_1% predicted. For AECOPD-related hospital duration, both SPPB and 6MW distance, following age, were the strongest associated measures. Sensitivity analysis indicated that 6MW distance might potentially be more useful in predicting admission rate in males and those with no history of exacerbation. Exacerbation history is known to reliably predict future risk of AECOPD including hospital admissions. In our cohort, women had a significantly higher number of previous exacerbations at baseline compared to men. In addition, exacerbation history was stronger associated with AECOPD-related admission for women than men and therefore may be more useful for future prediction of AECOPD-related hospital admission in women.

Over a study period of five years, over 40% had at least one admission with nearly one fifth of the cohort experiencing multiple visits. In addition, many individuals had readmissions within six months after initial admission, with an equal amount of time spent in hospital as the initial admission. Of those readmitted, about 20% died within the first year after initial admission. Unfortunately, the study was not designed to evaluate the associations between baseline variables and hospital readmission but considering the high number of readmissions and deaths following AECOPD-related admission, monitoring individuals at set intervals to identify those at high-risk could prove useful allowing timely intervention and preventing or minimising the number of readmission and premature mortality. Recently EDGE, for example, a mobile self-managing COPD platform has been introduced to monitor symptoms including heart rate to recognise and

start treatment of exacerbations early.[90]

The 6MW test is a reflection of cardiovascular status, in addition to lower limb function (i.e. musculoskeletal function), but perhaps also a good proxy measure of overall health. The 6MW test, however, has received limited adoption in clinical practice. The 4MGS – a test of lower limb function – being faster and more practical than the 6MW performed slightly less well than the 6MW, but might have more potential in clinical practice when considering the clinical practicalities.

This study has several limitations. Hospital episode statistics were obtained from the NHS Digital (England), NHS Scotland and NHS Wales. Apart from admission and discharge dates, we did not have spell data (i.e. total continuous stay and use of a hospital bed) available for individuals registered with the NHS Scotland and NHS Wales. The study period covered the time from study enrolment until the end of study, or death. Some individuals, however, may have been admitted to hospital for AECOPD shortly before study enrolment and these events will have not been electronically captured but potentially through self-reported exacerbation history. Although, self-reported data is known to suffer from recall bias.[96] Due to the limited number of events we were unable to stratify by GOLD stage (i.e. indicator of disease severity), exacerbation history, and sex to assess the association between baseline measures and AECOPD-related hospital stay (i.e. duration) for those admitted to hospital. Also, we explored for non-linearity of variables considered but had not enough power to identify any difference. For SPPB scores, a majority of individuals scored towards the highest possible score. Despite this, a high proportion of individuals with COPD had physical limitations but it may indicate that SPPB is not sensitive enough to discriminate sufficiently between those with and without the event. There were differences for most measures between recruitment sites. For example, individuals from London were slightly healthier compared to other sites, and individuals from Cardiff had more severe lung disease; however we caution that the departments at each of the five participating hospitals had variations in practice making analysis in difference in prognosis between sites of doubtful value. We addressed this by adjusting for recruitment site in our analyses. Even

though the ERICA study included participants from centres throughout the UK, the cohort consisted primarily of individuals GOLD staged II-III. This limits generalising results to those with mild or advanced COPD. Future studies, using larger cohorts and/or different geographical populations, should replicate our findings. Missing data was present, reducing the overall sample size and statistical power limiting to make robust conclusions. In order to optimise the analysis, we included as many observations as possible and reported the number of observations included in each analysis. Analyses were adjusted for productive cough, believed to be an indicator of inflammation. A large proportion had productive cough on most mornings but there was no significant association with the outcomes in our cohort. According to Hurst and colleagues, WCC and MRC dyspnoea score were found to be significantly associated with AECOPD.[132] We excluded these variables from the analysis due to collinearity, allowing us to evaluate the association of novel exercise capacity traits with the outcomes of interest.

This study has several strengths. Firstly, individuals were clinically stable upon recruitment. Secondly, event rates were stable throughout the study period, which is not only encouraging but also rates were comparable to those in large cohort studies including Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE).[270] Hospital admissions were identified using validated criteria, and only definite and possible episodes were included in the analysis. In contrast to self-reported hospital admission, which may suffer from under-reporting,[203,236] AECOPD episodes were captured using electronic health record data. Individuals had different observation periods. The use of study inclusion and admission (i.e. event) dates allowed to adjust for exposure time and therefore used the correct probability distributions.

Considering that both the SPPB and 6MW distance were associated with admission rate and duration, one potential scenario would include to use both tests to determine the physical capacity of an individual whilst assessing their risk of hospital admission rate within five years, and their hospital length of stay for initial AECOPD-related admission. It has been shown that exercise capacity can be improved through pulmonary rehabilitation, and is known to positively impact hospital utilisation. Exercise capacity traits such as the SPPB and 6MW may be good

candidates for training purposes, in addition to assessing an individual′s exercise capacity. Of these two, the SPPB is clinically more practical than 6MW, as there is often too little time available in clinic to perform a 6MW test and often patients are in need of additional oxygen. In addition, those who are very ill might have difficulty completing the 6MW test due to the intensity and physical impact. Our results show, however, no causality but merely an association between musculoskeletal strength and AECOPD admission and length of stay.

Future studies, using larger cohorts, should assess the predictive value of exercise capacity traits including SPPB and its component 4MGS, and the 6MW test, and demonstrate if physical training such as strengthening of the quadriceps improves pulmonary outcomes. In addition, evaluating these measures repeatedly at different time intervals would allow estimation of the association between AECOPD-related hospital admission rate, duration, and readmission at different time points.

6.5 Conclusions

There is potential for the use of SPPB in predicting AECOPD-related hospital admission and length of hospital stay within a COPD population with moderate to severe disease. The SPPB is a measure requiring low commitment that could be measured routinely.

7

Do arterial stiffness or carotid intima-media thickness improve on a Framingham approach when predicting cardiovascular disease in COPD?

Chapter summary

Background Individuals with chronic obstructive pulmonary disease (COPD) have increased risk of fatal and non-fatal cardiovascular disease. Cardiovascular risk is traditionally predicted using 'Framingham' risk factors (age, sex, smoking, high-density lipoprotein, total cholesterol, systolic blood pressure, diabetes, and the prescription of drugs to treat cardiovascular disease). However, newer measures, specifically arterial stiffness and carotid intimal thickness are thought to better capture systemic disease, and therefore may better identify high-risk individuals. Here we aimed to evaluate these measures against Framingham' risk factors for the prediction of cardiovascular events in a UK COPD population.

Methods Clinical data from the Evaluation of the Role of Inflammation in Chronic Airways disease (ERICA) cohort were linked with UK National Health Services electronic health record data, with cardiovascular events identified using ICD-10 coding. Non-fatal cardiovascular events were obtained from the UK Office for National Statistics, and adjudicated by cardiovascular and respiratory physicians. Associations were estimated using stratified multivariable Cox regression, and assessed by C-indices with 10-fold cross-validation and replication.

Findings Out of 714 individuals, 237 (33%) had at least one cardiovascular event during median follow-up of 4.5 years. Of the Framingham risk factors, age (hazard ratio (HR) 1.40 per 10-year increase, 95% CI 1.16 to 1.70), systolic blood pressure (HR 0.92 per 10 mmHg increase, 95% CI 0.84 to 0.99), self-reported diabetes (HR 3.07, 95% CI 2.21 to 4.27), and self-report use of drugs to treat cardiovascular disease (HR 2.10, 95% CI 1.52 to 2.90) were significantly associated with mortality. Measures of arterial stiffness and carotid intima-media thickness were not associated with cardiovascular events. Measures of exercise capacity four-metre gait speed (HR 1.07 per one sec. increase; C = 0.717) and six-minute walk distance (HR 0.91 per 30 metre increase; C = 0.728) were significantly associated with cardiovascular disease and improved the discriminative ability when added to Framingham risk factors.

Interpretation Our data does not support the use of objective measures of arterial stiffness and carotid intima-media thickness in addition to Framingham risk factors for predicting cardiovascular events within COPD. Similarly, blood pressure measurement and smoking status do neither add to the predictive ability of Framingham within the ERICA cohort. Age, systolic blood pressure, diabetes and cardiovascular drugs, and exercise capacity measures four-metre gait speed and six-minute walk distance are predictive. Moreover, despite the presence of cardiovascular disease, cardiac death is not common in patients with COPD. This may reflect a downward trend in fatal cardiovascular disease incidence/ improved cardiovascular survival, or an increased prevalence of death from respiratory causes.

7.1 Background

Chronic obstructive pulmonary disease (COPD) leads to a higher risk of cardiovascular (CV) disease, increasing the risk of non-fatal CV disease 2.5 times and a third dying of cardiac causes.[46,174] Reduced lung function is associated with systemic inflammation.[276] The elevated CV risk in COPD may be explained by the increased inflammatory burden and consequential effects leading to aortic stiffness and increased arteriosclerosis load. Current CV disease risk prediction algorithms focus mostly on the classical Framingham risk factors: age, sex, smoking, high-density lipoprotein (HDL), total cholesterol, systolic blood pressure (SBP, diabetes, and the use of drugs to treat CV disease.[62] The Framingham risk score predicts an individual′s 10-year risk of coronary heart disease, peripheral artery disease, and heart failure and improves the prediction of CV events and premature death.[156]

The predictive ability of the Framingham risk factors, the inclusion of alternative markers, and the comparison of the discriminative ability of Framingham with alternative measures has been evaluated in various population groups including elderly[70] and a multitude of comorbidities including diabetes,[143] metabolic syndrome,[279] and chronic kidney disease.[283] How Framingham risk factors and several other biomarkers for CV disease perform in a COPD population, however, is unclear.

Elevated pulse wave velocity (PWV), a measure of aortic stiffness, has been reported in patients with COPD,[94] but its predictive value is not known. Likewise, other proxy measures of arterial stiffness augmentation index (AIx) and carotid intima-media thickness (CIMT) are also predictors in some populations and have the potential of clinically predicting CV disease. Both arterial stiffness and CIMT are predictors of CV disease in the general population.[119,162,258] Adding these to the Framingham Risk Score was found to provide, albeit minor, improvement in the predictive ability.[73] Using these alternative measures, in particular measures of arterial stiffness, is thought to identify high-risk patients in an early stage of disease, which in turn could lead to opportunities to slow disease progression and to support decision makers in their

judgements regarding treatment planning and resource allocation. Evidence across several small studies, taken together, suggests that measures of arterial stiffness are worth further investigation but conclusive evidence is lacking.[91]

Thus, the aims of our study were firstly to determine incidence of fatal- and non-fatal CV disease, and evaluate the association of classical Framingham risk factors and with subsequent fatal and non-fatal CV events in stable Global initiative for chronic Obstructive Lung Disease (GOLD) stage II-IV[107] COPD patients using clinical data from the Evaluation of the Role of Inflammation in Chronic Airways disease (ERICA) cohort and UK electronic health record data. Secondly, we aimed to determine the association of measures of arterial stiffness and CIMT, and incident CV disease, and assess their added value above and beyond Framingham risk factors. Lastly, we used the opportunity to determine the association of alternative measures including musculoskeletal function, thought to better capture systemic problems, and CV disease, and their added value above and beyond Framingham risk factors.

7.2 Methods

7.2.1 Study design and participants

The ERICA study is a multi-centre observational, non-interventional, epidemiological cohort study, with 729 stable GOLD stage II-IV[107] COPD patients, established to identify important CV and musculoskeletal biomarkers that could be targeted to improve the outcomes of COPD patients. Full details of the protocol have been provided elsewhere.[184] Data captured included demographics, pulmonary function measures, biochemical markers, measures of arterial stiffness (i.e. PWV, AIx, and CIMT), and measures of musculoskeletal function (i.e. four-metre gait speed (4MGS) and six-minute walk (6MW) test).

7.2.2 Clinical measures

After four hours of fasting, with no bronchodilators for six hours, and ten minutes of supine rest CV measures were taken using an SphygmoCor system. Pulse wave velocity (i.e. velocity of blood pressure pulse) was measured between the femoral and carotid arteries, and the average of two measurements was taken, as described by Wilkinson *et al.*,[286] Carotid intima-media thickness (i.e. extend of arteriosclerotic process) of the common carotid arteries was measured using B-mode ultrasound at a distance of 1 cm from the carotid bulb with a 7–12 MHz linear probe to estimate the extent of atherosclerosis for each individual.[61] For each artery 3 x 10 sec. loops were recorded. The thickest artery of the two was included in the analysis. Augmentation index was derived from the ascending aortic pressure waveform.

Diabetes status and use of drugs for CV disease treatment were self-reported and captured at baseline. Cardiovascular disease-related treatment included drugs such as simvastatin, warfarin, eplerenone, bendroflumethiazide, digoxin, and ramipril. Disease severity was defined according to GOLD classification.[107] Points for the BODE Index, which generates a composite score from the Body mass index (BMI), airflow Obstruction, Dyspnoea, and Exercise capacity, with the latter measured by the 6MW test, were assigned as described by Celli *et al.*[40]

7.2.3 Cardiovascular events

Clinical data were linked to electronic healthcare records (i.e. admitted patient care data) obtained from the UK National Health Service (NHS) Digital, NHS Scotland, and NHS Wales. Non-fatal CV events were extracted from both primary and secondary international statistical classification of diseases and related health problems 10th revision coding (ICD-10) positions, and included diseases of the arteries, all stroke, and heart failure (**Table 7.1**, page 184). Atrial fibrillation and flutter, and hypertensive diseases were excluded. Fatal events were obtained from the UK Office for National Statistics (ONS) and cardiac deaths were adjudicated by CV and respiratory physicians.

Table 7.1: Definitions of diagnoses by ICD-10 coding.

End point	**ICD-10 codes**
All cardiovascular disease	E10.5, E11-E14, F01, G46.3-G46.7, G458, G459 I11.0, I13.0, I13.2, I20.0-I20.1, I20.8-I21-I25, I50, I60, I61, I.62, I63, I64, I65-I69, I70.2, I71.3-I71.9, I72, I73.9-I79, R96 + cardiac death
Diseases of the arteries	I70.2, I72, I73.9-I79, E10.5, E11-E14
Peripheral arterial disease	I70.2, I73.9, E10.5, E11-E14
Diseases of arteries, arterioles and capillaries	I72, I74-I79
Coronary heart disease	I20.0-I20.1, I20.8-I21-I25
Angina	I20.1, I20.8-I20.9
Unstable angina	I20.0, I24
Coronary heart disease not otherwise specified	I25
Acute myocardial infarction (MI), and certain current complications following acute MI	I21, I23
Subsequent myocardial infarction	I22
All stroke	I60, I61, I.62, I63, I64, I65-I69, F01, G46.3-G46.7, G458, G459
Subarachnoid haemorrhage	I60
Intra-cerebral haemorrhage	I61
Cerebral infarction	I63
Stroke, not specified as haemorrhage or infarction	I64
Stroke syndromes	G46.3-G46.7
Transient ischaemic attack	G458, G459
Other stroke	I62, I65-I69, F01
Heart failure	I11.0, I13.0, I13.2, I50
Heart failure	I50
Hypertensive heart disease with (congestive) heart failure	I11.0
Hypertensive heart and renal disease with (congestive) heart failure	I13.0
Hypertensive heart and renal disease with both (congestive) heart failure and renal disease	I13.2
Cardiac death	Adjudicated
Other vascular deaths	R96, I71.3-I71.9
Sudden death, cause unknown	R96
Abdominal aortic aneurysm	I71.3-I71.9

Atrial fibrillation and flutter (I48) and hypertensive diseases (I10-I15) were considered risk factors and therefore not included.

7.2.4 Outcomes and predictors of interest

The primary outcome measure was defined as the new occurrence (first event) of fatal or non-fatal CV disease. Time to event was calculated from the difference between the baseline visit date and either the date of death or first CV event up to November 2017, when follow-up discontinued. The association between Framingham risk factors (i.e. age, sex, smoking, HDL, total cholesterol, systolic blood pressure, diabetes, and the use of drugs to treat CV disease) and CV disease were evaluated. In addition, measures of arterial stiffness, CIMT and alternative measures thought to better capture systemic disease, and CV disease and their added value above and beyond Framingham risk factors were evaluated.

7.2.5 Statistical analysis

Hazard ratios (HRs) were estimated using Cox regression, stratified by recruitment centre, and adjusted for age and sex. In addition, PWV was adjusted for mean arterial pressure (MAP) and heart rate, and AIx was adjusted for heart rate and height. Further analyses included Framingham risk factors as covariates. Discrimination (i.e. Harrell′s C-statistic)[10,202] was assessed using 10-fold cross validation with 200 replications.[245] Hazard ratios for log-transformed biomarkers represent a twofold increase in the biomarker. Associations between the clinical measures were quantified using Spearman's rank correlations, with values <0.30 considered as weak, 0.30-0.50 as moderate, and >0.50 as strong.[51]

There were missing values. Data were assessed for the level and type of missing data, and completion patterns (**Figures 3.1 3.2**, pages 61 and 62). There were about 10% missing values for variables CIMT (n = 66) and PWV (n = 60), with $<5\%$ missing values for other variables. Missing values were addressed using multiple imputations using chained equations (MICE). The time-to-event outcome was included using the non-parametric Nelson-Aalen estimator. Predictive mean matching was used for continuous variables, ordered logistic regression (as continuous) for ordinal variables, multinomial logistic regression for categorical variables, and logistic regression for binary variables. Derived variables such as the BODE Index (a composite score of BMI,

forced expiratory volume in one second (FEV_1), Medical Research Council dyspnoea score, and 6MW distance) and GOLD stage were estimated post MICE using passive imputation. Observational data is reported according to the Strengthening The Reporting of OBservational Studies in Epidemiology (STROBE) statement.[275]

7.3 Findings

7.3.1 Descriptive statistics

Of the 729 individuals included in the study, 714 were linkable with hospital admission and survival records, and included in the analysis (**Figure 3.4**, page 64). The mean age was 67 years old (range 43-89 years) and 434 (61%) individuals were male. A third of the cohort smoked, and 402 individuals (56%) were taking CV drugs at baseline. Median (interquartile range (IQR)) SBP was 142 mmHg (131-154), PWV 9.8 m/sec (8.4-11.8), CIMT 0.81 (0.71-0.96), and AIx 28% (20-34; **Tables 7.1 7.2 7.3**, pages 184-187). Baseline characteristics have been reported in **Figures 3.25 3.26 3.27**, pages 79-80.

7.3.2 Association of Framingham risk factors with CV events, and their predictive value

In total, six individuals had a fatal CV event and 231 individuals (33%) experienced a non-fatal CV event during median follow up for 4.5 years. The CV incidence rate was 8.8 (95% CI 7.7 to 10.0) per 100 person-years. Of the Framingham risk factors, only age, SBP, and self-reported diabetes and use of drugs to treat CV disease were significantly associated with CV events (**Figure 7.1** and **Table 7.2**, pages 189 and 187). Systolic blood pressure was negatively associated with CV disease; primarily for those with mild COPD (HR 0.88, 95% CI 0.78 to 0.99, p = 0.030; aged 70 years and above (HR 0.87, 95% CI 0.78 to 0.97, p = 0.012; taking CV drugs (HR 0.86, 95% CI 0.78 to 0.95, p = 0.003), and males (HR 0.87, 95% CI 0.79 to 0.96, p = 0.007).

Table 7.2: Hazard ratios for cardiovascular disease with measured baseline levels of risk factors.

	Median (IQR) or n (%)	HR (95% CI) [a]	P value	HR (95% CI) [b]	P value
Framingham risk factors					
Age - per 10 year increase	67 (62-73)	1.50 (1.26 to 1.78)	<0.001	1.40 (1.16 to 1.70)	<0.001
Sex - males	434 (61)	1.16 (1.02 to 1.06)	0.282	1.17 (0.88 to 1.56)	0.283
Smoking - current	218 (31)	0.93 (0.69 to 1.25)	0.641	0.93 (0.69 to 1.26)	0.654
HDL - per 1 mmol/L increase	1.4 (1.2-1.7)	0.61 (0.44 to 0.85)	0.003	0.94 (0.67 to 1.31)	0.721
Total cholesterol - per 1 mmol/L increase	5.0 (4.3-5.8)	0.71 (0.63 to 0.81)	<0.001	0.94 (0.82 to 1.08)	0.392
SBP - per 10 mmHg increase	142 (131-154)	0.92 (0.85 to 0.99)	0.023	0.92 (0.84 to 0.99)	0.031
Diabetes - yes	82 (12)	4.18 (3.11 to 5.63)	<0.001	3.07 (2.21 to 4.27)	<0.001
CV drug treatment - yes	402 (56)	2.61 (1.94 to 3.52)	<0.001	2.10 (1.52 to 2.90)	<0.001

Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of 714 patients are included. All models are stratified by recruitment site.
[a] Adjusted for age and sex
[b] Adjusted for Framingham risk factors: age, sex, smoking, high-density lipoprotein, total cholesterol, systolic blood pressure, diabetes, cardiovascular drug treatment. *Abbreviations*: CI, confidence interval. HDL, high-density lipoprotein. SBP, systolic blood pressure. CV, cardiovascular.

Table 7.3: Hazard ratios for cardiovascular disease with measured baseline levels of risk factors.

	Median (IQR) or n (%)	HR (95% CI) [a]	P value	HR (95% CI) [b]	P value
Measures of arterial stiffness					
PWV - per 1 m/sec increase	9.8 (8.4-11.8)	1.04 (0.98 to 1.10)	0.171	0.99 (0.93 to 1.06)	0.843
CIMT - per 1 mm increase	0.81 (0.71-0.96)	1.27 (0.62 to 2.60)	0.512	1.22 (0.58 to 2.54)	0.602
AIx - per 5% increase	28 (20-34)	0.85 (0.78 to 0.92)	<0.001	0.93 (0.85 to 1.02)	0.129

Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of 714 patients are included. All models are stratified by recruitment site.
[a] Adjusted for age and sex
[b] Adjusted for Framingham risk factors: age, sex, smoking, high-density lipoprotein, total cholesterol, systolic blood pressure, diabetes, cardiovascular drug treatment.
PWV further adjusted for mean arterial pressure and resting heart rate. AIx further adjusted for resting heart rate and height. *Abbreviations*: CI, confidence interval. PWV, pulse wave velocity. CIMT, carotid intima-media thickness. AIx, augmentation index.

The discriminative ability of all Framingham risk factors combined had a C-statistic of 0.701, 95% CI 0.695 to 0.706). Self-reported use of CV drugs (C = 0.638, 95% CI 0.630 to 0.647) and diabetes (C = 0.616, 95% CI 0.607 to 0.622) followed by age (C = 0.594, 95% CI 0.588 to 0.602) contributed most to the discriminative ability.

Risk factor	Median (IQR) or n (%)	HR (95% CI)[a]	HR (95% CI)[b]	C–statistic
Framingham				0.701 [0.695 to 0.706]
Age – per 10 year increase	67 (62–73)	1.50 (1.26 to 1.78)	1.40 (1.16 to 1.70)	0.594 [0.588 to 0.602]
Sex – male	434 (61)	1.16 (1.02 to 1.06)	1.17 (0.88 to 1.56)	0.551 [0.542 to 0.558]
Smoking – current	218 (31)	0.93 (0.69 to 1.25)	0.93 (0.69 to 1.26)	0.537 [0.528 to 0.544]
HDL – per 1 mmol/L increase	1.4 (1.2–1.7)	0.61 (0.44 to 0.85)	0.94 (0.67 to 1.31)	0.566 [0.559 to 0.573]
Cholesterol – per 1 mmol/L increase	5.0 (4.3–5.8)	0.71 (0.63 to 0.81)	0.94 (0.82 to 1.08)	0.616 [0.609 to 0.620]
SBP – per 10 mmHg increase	142 (131–154)	0.92 (0.85 to 0.99)	0.92 (0.84 to 0.99)	0.539 [0.531 to 0.544]
Diabetes – yes	82 (12)	4.18 (3.11 to 5.63)	3.07 (2.21 to 4.27)	0.616 [0.607 to 0.622]
CV drugs – yes	402 (56)	2.61 (1.94 to 3.52)	2.10 (1.52 to 2.90)	0.638 [0.630 to 0.647]

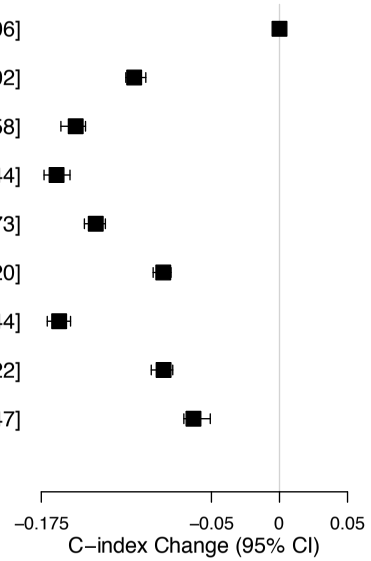


Figure 7.1: Framingham risk factors at baseline, their hazard ratios and discriminative ability for cardiovascular disease. Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of 714 patients are included. All models are stratified by recruitment site.
[a] Adjusted for age and sex
[b] Adjusted for Framingham risk factors: age, sex, smoking, high-density lipoprotein, total cholesterol, systolic blood pressure, diabetes, cardiovascular drug treatment. *Abbreviations*: CI, confidence interval. HDL, high-density lipoprotein. SBP, systolic blood pressure. CV, cardiovascular. There were <5% missing values for descriptive variables such as BMI and smoking status. Missing values were addressed using multiple imputations using chained equations.

7.3.3 Association of arterial stiffness measures with CV events, and their predictive value

Except for the AIx (HR 0.85 per 5% increase, 95% CI 0.78 to 0.92, p <0.001), none of the arterial stiffness measures were significantly associated with CV disease, after adjustment for age and sex (**Figure 7.2** and **Table 7.3**, pages 191 and 187). After further adjustment for Framingham risk factors neither was AIx. Arterial stiffness did not statistical significantly change discriminative ability of Framingham.

Risk factor	Median (IQR) or n (%)	HR (95% CI)[a]	HR (95% CI)[b]	C–statistic
Framingham				0.701 [0.695 to 0.706]
+ PWV – per 1 m/sec increase	9.8 (8.4–11.8)	1.04 (0.98 to 1.10)	0.99 (0.93 to 1.06)	0.700 [0.694 to 0.705]
+ CIMT – per 1 mm increase	0.81 (0.71–0.96)	1.27 (0.62 to 2.60)	1.22 (0.58 to 2.54)	0.700 [0.693 to 0.705]
+ AIx – per 5% increase	28 (20–34)	0.85 (0.78 to 0.92)	0.93 (0.85 to 1.02)	0.698 [0.692 to 0.703]

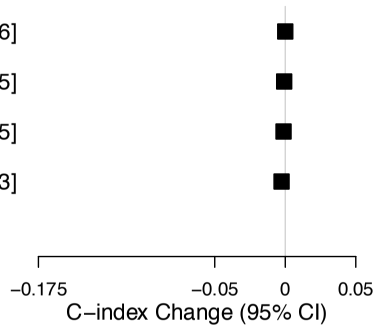


Figure 7.2: Arterial stiffness at baseline, their hazard ratios and discriminative ability for cardiovascular disease. Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of 714 patients are included. All models are stratified by recruitment site.

[a] Adjusted for age and sex

[b] Adjusted for Framingham risk factors: age, sex, smoking, high-density lipoprotein, total cholesterol, systolic blood pressure, diabetes, cardiovascular drug treatment. PWV further adjusted for mean arterial pressure and resting heart rate. AIx further adjusted for resting heart rate and height. *Abbreviations*: CI, confidence interval. PWV, pulse wave velocity. CIMT, carotid intima-media thickness. AIx, augmentation index. There were about 10% missing values for variables CIMT (n = 66) and PWV (n = 60). Missing values were addressed using multiple imputations using chained equations.

7.3.4 Association of alternative measures with CV events, and their predictive value

Multivariable analysis identified several alternative measures that were associated with CV events (**Figure 7.3** and **Table 7.4**, pages 193 and 194). Inflammatory markers C-reactive protein (CRP; HR 1.11 per twofold increase, 95% CI 1.02 to 1.21, p = 0.013) and fibrinogen (HR 1.59 per twofold increase, 95% CI 1.07 to 2.39, p = 0.023), blood glucose (HR 1.94 per twofold increase, 95% CI 1.07 to 3.52, p = 0.030), BMI (HR 1.04 per 1 kg/m^2 increase, 95% CI 1.01 to 1.06, p = 0.002), and FEV_1 as GOLD stage (HR 1.27 per 1 stage increase, 95% CI 1.04 to 1.56, p = 0.021) were associated with a higher risk of CV events. Musculoskeletal measures 4MGS (HR 0.72 per 1 second increase, 95% CI 0.71 to 0.72, p = 0.009) and 6MW distance (HR 0.73 per 30 metre increase, 95% CI 0.72 to 0.73, p <0.001) were associated with a lower risk of CV events. Predictive modelling indicated statistical significant improvement in risk discrimination when adding 4MGS (C = 0.717, 95% CI 0.712 to 0.722) or 6MW distance (C = 0.728, 95% CI 0.723 to 0.733) to the Framingham risk factors. Adding BMI, 4MGS, 6MW, and BODE all together to the Framingham risk factors resulted in a C-index of 0.731 (95% CI 0.727 to 0.737).

Risk factor	Median (IQR) or n (%)	HR (95% CI)[a]	HR (95% CI)[b]	C–statistic
Framingham				0.701 [0.695 to 0.706]
+ CRP – per twofold increase	1.21 (0.47 to 2.01)	1.13 (1.05 to 1.23)	1.11 (1.02 to 1.21)	0.703 [0.698 to 0.708]
+ Fibrinogen – per twofold increase	1.22 (1.06 to 1.36)	1.71 (1.15 to 2.56)	1.59 (1.07 to 3.52)	0.700 [0.694 to 0.705]
+ Glucose – per twofold increase	1.59 (1.46 to 4.59)	2.40 (1.33 to 4.33)	1.94 (1.07 to 3.52)	0.701 [0.693 to 0.706]
+ BMI – per 1 kg/m2 increase	27 (23–31)	1.07 (1.04 to 1.09)	1.04 (1.01 to 1.06)	0.712 [0.707 to 0.718]
+ GOLD – per 1 stage increase	1 (1–2)	1.19 (0.98 to 1.45)	1.27 (1.04 to 1.56)	0.701 [0.695 to 0.706]
+ 4MGS – per 1 second increase	4.2 (3.5–5.2)	1.10 (1.05 to 1.16)	1.07 (1.02 to 1.13)	0.717 [0.712 to 0.722]
+ 6MW distance – per 30 meter increase	366 (255 to 440)	0.91 (0.88 to 0.94)	0.91 (0.89 to 0.95)	0.728 [0.723 to 0.733]
+ BODE – per 1 point increase	3 (1–5)	1.12 (1.07 to 1.18)	1.13 (1.07 to 1.19)	0.714 [0.709 to 0.720]
+ BMI, 4MGS, 6MW, BODE – per 1 unit increase				0.731 [0.727 to 0.737]

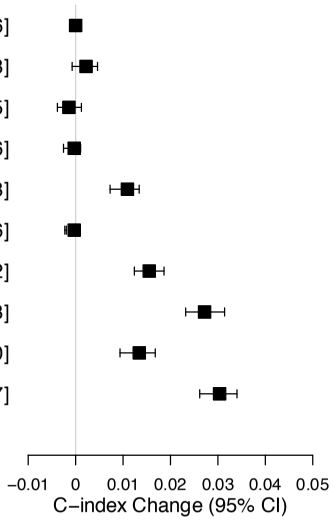


Figure 7.3: Alternative measures at baseline, their hazard ratios and discriminative ability for cardiovascular disease. Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of 714 patients are included. All models are stratified by recruitment site.

[a] Adjusted for age and sex

[b] Adjusted for Framingham risk factors: age, sex, smoking, high-density lipoprotein, total cholesterol, systolic blood pressure, diabetes, cardiovascular drug treatment. *Abbreviations*: CI, confidence interval. CRP, C-reactive protein. BMI, body mass index. GOLD, global initiative for chronic obstructive lung disease. 4MGS, four-metre gait speed. 6MW, six-minute walk. BODE, body mass index, obstruction, dyspnoea, exercise. There were <5% missing values for biochemical markers including fibrinogen and cholesterol. Missing values were addressed using multiple imputations using chained equations.

Table 7.4: Hazard ratios for cardiovascular disease with measured baseline levels of risk factors.

	Median (IQR) or n (%)	HR (95% CI) [a]	P value	HR (95% CI) [b]	P value
Other risk factors					
CRP - per twofold increase	1.21 (0.47 to 2.01)	1.13 (1.05 to 1.23)	0.002	1.11 (1.02 to 1.21)	0.013
Fibrinogen - per twofold increase	1.22 (1.06 to 1.36)	1.71 (1.15 to 2.56)	0.009	1.59 (1.07 to 2.39)	0.023
Glucose - per twofold increase	1.59 (1.50 to 1.69)	2.40 (1.33 to 4.33)	0.004	1.94 (1.07 to 3.52)	0.03
BMI - per 1 kg/m^2 increase	27 (23-31)	1.07 (1.04 to 1.09)	<0.001	1.04 (1.01 to 1.06)	0.002
GOLD - per 1 stage increase	1 (1-2)	1.19 (0.98 to 1.45)	0.083	1.27 (1.04 to 1.56)	0.021
4MGS - per 1 second increase	4.2 (3.5-5.2)	1.10 (1.05 to 1.16)	<0.001	1.07 (1.02 to 1.13)	0.009
6MW distance - per 30 metre increase	366 (255 to 440)	0.91 (0.88 to 0.94)	<0.001	0.91 (0.89 to 0.95)	<0.001
BODE - per 1 point increase	3 (1-5)	1.12 (1.07 to 1.18)	<0.001	1.13 (1.07 to 1.19)	<0.001

Values are given as the median and interquartile range (IQR), or No. of cases (%). Baseline data of 714 patients are included. All models are stratified by recruitment site.
[a] Adjusted for age and sex
[b] Adjusted for Framingham risk factors: age, sex, smoking, high-density lipoprotein, total cholesterol, systolic blood pressure, diabetes, cardiovascular drug treatment. *Abbreviations*: CI, confidence interval. CRP, C-reactive protein. BMI, body mass index. GOLD, global initiative for chronic obstructive lung disease. 4MGS, four-metre gait speed. 6MW, six-minute walk. BODE, body mass index, obstruction, dyspnoea, exercise.

7.4 Discussion

Four of the classical Framingham risk factors were shown to have prognostic power for the prediction of CV events, but contrary to our hypothesis predictive power was not improved by any of the measured biomarkers of CV function. The inverse relationship with SBP may be due to confounding and can possibly be explained by advancing age related to frailty.[204] By examining other data from the ERICA dataset we show that several other measures, specifically CRP, fibrinogen, BMI, GOLD stage, 4MGS, and 6MW, were significantly associated with CV events above and beyond Framingham risk factors. The C-statistic of Framingham risk factors with BMI, 4MGS, 6MW and BODE Index combined, is similar to adding the 6MW only. This indicates 6MW is the main component driving the improvement of discriminative ability.

Tests such as the 4MGS and 6MW distance are proxy measures of overall mobility and

physical functioning including CV fitness.[161] Exercise capacity and CV fitness are known to be associated with fatal and non-fatal CV disease.[201] Simultaneously, exercise-based cardiac rehabilitation reduces risk of heart disease.[12] GOLD stage (i.e. FEV_1) is a reflection of airflow limitation and also known to be associated with CV disease.[191] For every 10% reduction in lung performance, fatal- and non-fatal CV disease were reported to increase by 28% and 20%, respectively.[240] Fibrinogen is useful in identifying high risk individuals for COPD exacerbation and early mortality.[83] And although analysis of cross-sectional data from the ERICA study indicated no relationship between fibrinogen and CV manifestations in COPD,[183] analysis of the prospective data suggests fibrinogen to be associated with fatal- and non-fatal CV disease in COPD. That the associations between previously mentioned biomarkers and CV events remain significance after adjustment for Framingham risk factors indicate their potential value for identifying high risk individuals within a COPD population.

Age- and sex-adjusted Cox regression indicated a negative association between AIx and CV events. This association disappeared, however, after including other Framingham risk factors. There was also a negative association with smoking but this was not significant. About a third of the ERICA population were current smokers, yet all individuals had at least ten pack-years of smoking. Individuals with COPD and such a smoking history might already be at a higher risk for most diseases including CV disease, regardless of their smoking status, and therefore may explain why smoking status did not add any value in predicting CV events.

7.4.1 Strength and limitations

It could be argued that we did not have enough patients to detect any difference in arterial stiffness and CIMT measurements between high- and low-risk individuals. It should also be noted that while a systematic review including seventeen studies concluded that arterial stiffness was a strong predictor of CV disease,[273] this was primarily the case for individuals at higher risk and more severe disease including renal failure, whom were not included in the ERICA study. Definitions of CV disease and arterial stiffness measurement differed amongst included studies

and most failed to adjust for height and MAP. Also, we excluded hypertension, atrial fibrillation and flutter as outcomes, since these are rather risk factors and not necessarily a CV event, and may therefore limit comparability with the studies included in the systemic review. Differences in baseline CV risk may also explain the absence of association between arterial stiffness and CV events.[273] With the few number of events we were unable to examine this. The Framingham Risk Score has not yet been calibrated for the COPD population. The recalibration, an index of accuracy, allows to adapt the risk score to the COPD population, addressing potential over- or underestimation, and therefore may return different risk estimates.

This study has potential limitations. There was no validated algorithm available to identify CV events in electronic health record data. Previously CV disease endpoints have been validated using the Clinical Practice Research Datalink (CPRD) classification algorithm[121] combined with extensive clinical input.[18] We did, however, not have access to CPRD data and defined CV disease based on classifications used by the Emerging Risk Factors Collaboration.[86] Efforts are being made to reach a consensus in the use of clinical CV endpoints.[123] We extracted CV events from both primary and secondary ICD-10 positions. Selecting the primary position only would indicate the underlying cause of diagnosis specifically but most CV events were recorded in secondary positions, indicating the primary admission might be related to something else than for cardiac reasons with CV disease seen as comorbidity. We did not have an independent validation cohort, which is commonly required when assessing the predictive ability of biomarkers. To address this, we used cross validation techniques (i.e. random partitioning of the dataset). In addition, we used replication in the cross validation to account for the relatively small number of observations aiming to prevent potential overfitting of the models.[245] Baseline data varied amongst recruitment centres for most variables. We addressed this by stratifying by recruitment centre. Generalisability is limited to those with mild to moderate disease, as the majority of the cohort was GOLD staged II-III. Overall, the study had a relatively small sample size, which might have contributed to several Framingham risk factors not to be associated with CV events in COPD. We used multiple imputation techniques nonetheless to improve the statistical power

and precision. Missing data is common in clinical studies and epidemiological research. Ignoring missing data and analysing complete data only may introduce bias and provide misleading results.[250] Multiple imputation replaces missing values with estimates based on the observed data. The correct and appropriate use of multiple imputation techniques is expected to improve the validity of clinical findings.

7.4.2 Significance of the findings

Cardiovascular disease is predicted to be accountable for a third of deaths globally.[172] However, in the ERICA cohort only very few cardiac deaths were reported with most deaths related to pulmonary disease. This might be the result of biased death certificates, potentially leading to misclassification and underreporting of cardiac death.[138] Since the nineties, death rates for circulatory disease, primarily due to ischaemic heart disease, and cancer have both declined largely.[209] Regardless, despite the recommendations of measuring arterial stiffness in clinical practice[164] we found no evidence in the ERICA cohort that would support screening for arterial stiffness in COPD. In addition, most missing values were present for measures of arterial stiffness, likely because producing high quality data is difficult and time consuming. Instead, our data supports the inclusion of musculoskeletal measures in predicting CV events in COPD. Both the 6MW test and 4MGS statistically improved the discriminative ability, with the 4MGS having more potential when considering clinical practicalities. It is faster and easier to complete than the 6MW test and requires only a stopwatch and a short flat walking surface. Moreover, despite the presence of CV disease, cardiac death is not common in patients with COPD. This may reflect a downward trend in fatal CV disease incidence/ improved CV survival, or an increased prevalence of death from respiratory causes.

7.5 Conclusions

We show that age, diabetes and taking drugs to treat CV disease are predictors of CV disease in a COPD population but measures of arterial stiffness and CIMT do not provide any additional value to predicting CV disease. In addition, alternative measures thought to better capture systemic problems, in particular the 6MW distance and the 4MGS test were significantly associated with CV events, and may improve the predictive ability above and beyond Framingham risk factors.

8

General Discussion

Summary of the key findings

The overall aim of this doctoral research was to identify and evaluate the relationships between existing and novel biomarkers, and questionnaire data and electronic health record data, and determine if and how these biomarkers can predict common clinical outcomes (i.e. acute exacerbation of chronic obstructive pulmonary disease (COPD), hospitalisation, and mortality) within a COPD population.

Some of the key scientific questions in COPD at present relate to the evaluation of the usefulness of novel biomarkers such as cardiovascular (CV) and musculoskeletal measures in predicting common clinical outcomes. The six-minute walk (6MW) distance is by far the most studied biomarker in COPD. Before this research, evidence about the usefulness of alternative measures four-metre gait speed (4MGS), short physical performance battery (SPPB), quadriceps maximum voluntary contraction (QMVC), and sniff nasal inspiratory pressure (SNIP) in predicting clinical outcomes in COPD was limited. Similarly, there was little known about the associations

of objective CV measures augmentation index (AIx), pulse wave velocity (PWV) and carotid intima-media thickness (CIMT), and clinical outcomes in COPD. This research aimed to fill this gap and contribute to the unmet need of evaluating extra-pulmonary manifestations in COPD.

Findings of this research indicate that regardless of the clinical outcome, the 6MW test is the superior test compared to any of the alternative measures in terms of discriminative ability. In the systematic review and meta-analysis (**Chapter 2**), alongside inflammatory marker C-reactive protein, 6MW distance was also the only marker associated with mortality, exacerbation and hospitalisation. Nevertheless, analysis in **Chapter 3** indicated that most of the missing data following CV measures were related to the 6MW test. Though, there are alternative measures that could potentially replace the 6MW test in predicting clinical outcomes in COPD. Analysis in **Chapter 4** showed that the SPPB and its 4MGS and balance components, and the SNIP have potential to replace the 6MW component in the BODE (Body mass index, airflow Obstruction, Dyspnoea and Exercise) Index when predicting mortality without significant loss in discriminative ability. In addition to predicting mortality, analysis in **Chapter 6** showed the SPPB to also have potential in predicting acute exacerbations of COPD-related hospital admission and length of stay. The QMVC, however, the exception of musculoskeletal measures, was found not to perform that well in predicting clinical outcomes in COPD. Despite the superior performance of the 6MW in predicting outcomes in COPD, it is encouraging that these alternative measures have the potential of replacing the 6MW in predicting clinical outcomes in COPD.

Moreover, analysis of the Evaluating the Role of Inflammation in Chronic Airways disease (ERICA) and UK Biobank cohorts in **Chapter 5** showed that despite the believe cardiac death is common in COPD, the primary cause of death is related to cancer and pulmonary disease. This may also explain findings in **Chapter 7**, where CV measures of interest were not predictive for clinical outcomes. Then again, alternative measures 4MGS and especially the 6MW test improved the predictive ability of a Framingham approach when predicting CV disease in COPD.

8.1 Strengths and limitations

The ERICA cohort is a well-phenotyped cohort. Despite ERICA being an observational cohort study, making it difficult to prove causality, the study was prospective by design which in turn reduces the likelihood of reversed causality and allowed to calculate risk estimates. Missing data is common in clinical observational studies as a result of e.g. failed recordings or measurement errors, study participants skipping visits, or individuals being lost to follow-up. Despite clear study protocols, there were missing data, in particular for CV measures and 6MW with evidence that some data were not missing at random. With sensitivity analysis we aimed to address this, comparing complete-case with imputed data but the presence of missing data was likely related to the difficulty of generating high quality data of these measures, and the physical intensity required to perform the 6MW test. Although there were missing data for baseline variables and follow-up questionnaire data, the greatest strength of this research is the linkage with electronic health record (EHR) data that indicate any hospital admission or death report for each study participant for the full study duration. Only fifteen individuals (2%) of all those recruited could not be included in the analysis because they were not followed-up. A comparison of self-reported CV events with EHR data showed almost perfect agreement (>80%) after the first year of follow-up. In addition, the use of EHR data allowed evaluation of the association between multiple biomarkers with multiple outcomes. Electronic health record data are better in capturing the heterogeneity of COPD than cohort studies relying on questionnaire data only. On the other hand, a large retrospective observational study evaluating the use of EHR data in predicting outcomes emphasised potential biases in using EHR data affecting the ability of predicting clinical outcomes.[2]

The ERICA study was conducted in multiple centres throughout the UK, increasing the diversity of study participants and facilitating the generalisation of findings. Baseline data differed though amongst the recruitment centres. The centre in London, for example, had slightly healthier individuals, whereas individuals from Cardiff had more severe disease. This

was addressed by stratification of recruitment centre. Also, most individuals in the ERICA cohort were diagnosed with mild to moderate disease, making it difficult to generalise findings to those in early stage or the very advanced disease. There were no socio-economic details captured in the ERICA cohort, therefore analyses were not adjusted for this. Socio-economic status may affect the incidence rates of AECOPD and mortality.

The overall sample size of the ERICA study was relatively small, reducing statistical power and limiting the making of robust conclusions. The ERICA study was originally designed and powered on the basis of a tertile analysis of variables PWV and QMVC, based on an estimated sample size of 800 individuals with COPD.[184] However, this sample size is not sufficient when developing or evaluating multivariable prognostic models and may have resulted in failing to capture the significance of associations of specific biomarkers. In order to produce robust findings that are measurable and comparable, the sample size should preferably have been estimated based on the D or C-statistic.[140] In developing the modified BODE Indices based on alternative musculoskeletal measures we used multiple imputation to maximise the sample size, and used cross validation with replication to prevent model overfitting and to avoid relying on the availability of another independent dataset, but in essence findings rather generate new hypothesis and do require replication in larger and non-UK cohorts.

I have attempted to obtain access to alternative datasets including the UB Biobank, CPRD and US-based Million Veterans Program (MVP). Financial and time restrictions prevented obtaining access to at least CPRD and MVP data. To some extent we have been able to validate some of the findings in the UK Biobank. Using data from the UK Biobank we have tried to replicate findings related to CV disease incidence in COPD, in particular cardiac death. Both the ERICA and UK Biobank cohorts are likely a better representation on the UK than the TORCH trial, which was a highly selective population. We could not validate findings related to musculoskeletal weakness, since these biomarkers were not captured by the UK Biobank. For only a limited number of individuals CIMT and PWV were recorded. To identify CV disease, events were captured using ICD-10 coding based on the Emerging Risk Factors Collaboration.

However, there was absence of a validated algorithm. Using a different set of codes would likely have resulted in a different number of individuals with defined CV disease. Despite these weaknesses, this research has many strengths in addition to having unique features, and serves as a great foundation for further analysis.

8.2 Public health implications

The hypothesis was that several new predictors would be useful in predicting clinical outcomes in COPD. Of all measures evaluated, age and the 6MW test were most predictive of the outcomes but findings in the ERICA cohort indicate several significant associations between musculoskeletal measures and common clinical outcomes in COPD. These novel biomarkers may have potential for inclusion in risk prediction in primary care settings, such as the BODE Index, aimed at identifying high-risk individuals in an earlier stage of disease when timely intervention is still possible. The BODE Index was introduced in 2004 but it has failed to be widely adopted clinically, likely due to space and time constraints relating to the 6MW. These alternative musculoskeletal measures require only low commitment allowing for routine measurement and thus may improve the uptake of risk prediction indices in clinical practice. For example, the chair stand component of the SPPB in particular could be useful as a standalone test in time-limited settings such as primary care. More importantly, no other study provides data suggesting that in patients with stable COPD, SPPB or the chair stand are associated with hospitalised AECOPD incidence as well as related length of stay, and this information further adds support for SPPB being used as a drug development tool and endpoint for clinical trials addressing AECOPD, especially since the European Medicines Agency (EMA) favours the SPPB as the measure of choice in the assessment of frailty.

Findings also indicate that the assessment of physical capacity and its improvement should form a part of routine care for COPD in order to, for example, reduce acute exacerbation of COPD (AECOPD) risk. We demonstrated that simple and amenable to routine care exercise

capacity test, SPPB or its components like chair stand have similar associations with a higher risk of AECOPD requiring hospital admission as the 6MW test. These specific musculoskeletal markers are modifiable traits, in particular chair stand and balance are modifiable measures, making them ideal for training and testing purposes, and could be incorporated in physical rehabilitation programs, in addition for usage in predicting clinical outcomes. Not only is effective treatment lacking, prevention of these clinical outcomes is much more efficient and has a lower cost associated for healthcare systems. The SPPB, and in particular its chair stand component, are usable in both primary and secondary care, for evaluating risk of mortality, and hospital admission and stay in COPD. The SPPB can aid in decision making and prioritising healthcare resources. In addition, most notably pulmonary rehabilitation can increase physical capacity.[263] Although the available data are mixed, some reports suggest that novel strategies can reduce hospital admission rates by early application of telemedicine techniques. However, these interventions have costs and thus in terms of prioritising patients who will derive most benefit it is suggested that SPPB is a useful too.

These exercise capacity traits can also easily be combined with wearables and other electronic devices. Technological advancement has allowed for capturing a wealth of information at increased accuracy through, for example, monitoring physical activity. Evidence shows that even simple pedometers can improve physical activity,[176] and smartphone-based physical activity is well received by both patients and providers.[160]. In 2013, EDGE a mobile self-managing COPD platform was introduced aimed at monitoring symptoms including oxygen levels and heart rate to recognise and start treatment of exacerbations early.[90] Medopad, another recent development, allows in addition to recording symptoms and disease related questions, individuals to perform a standardised 6MW test. These and other wearables and devices could in turn be linked to EHR data and primary care data to feed data allowing to monitor disease progression, and warn patients and health care providers timely when intervention is needed.[84]

To facilitate the adoption of newly developed risk models, including the BODE Index based on the SPPB, interactive platforms such as Shiny – an *R* package for building interactive web

applications – could prove useful to clinicians and promote uptake of the Index. When communicating the clinical utility of risk models, decision curve analysis would an appropriate method, which would in turn also facilitate the adoption of risk prediction models in clinical practice.[271]

If fatal- and non-fatal major CV events in COPD are less common, reflected through a downward trend in CV disease incidence/ improved CV survival and an increased prevalence of death from other causes especially respiratory and cancer, it is important to update clinical guidelines. Considering the limited resources available and increasing healthcare expenditure it is important to avoid unnecessary diagnostic testing, for example, measuring arterial stiffness in COPD when they are not predictive of clinical outcomes.

8.3 Future research

Future investigations could focus on evaluating and validating the predictive ability of biomarkers in larger studies with longer follow-up times. Emphasis should be placed on ensuring biomarkers are generalisable (i.e. more diversity in ethnicity and comorbidities) and practical for clinical use. Most studies on COPD are conducted in the developed world, particularly in Europe where clinical guidelines are already in place with good diagnostic facilities. Many studies have too small sample sizes and/or too short follow-up periods, are cross-sectional in design resulting in over-estimation of effect sizes, or lack generalisability to a 'real world population' limiting generalisability or even preventing the estimation of the predictive value of a risk factor.[47,102] Especially with the increasing interest of including genetic data in risk prediction models, large sample sizes are required. Future investigations could focus on the external validation of existing risk models or perhaps combined models, ideally tailored to the individual with potentially adding novel predictors such as genetic variants aimed at maximising patient benefit. The use of genetic data may improve prediction accuracy, and potentially identify novel genetic causes that e.g. play a role in the development of lung disease and lifestyle behaviour.[1] Recommended techniques include Mendelian randomisation (MR) analysis – a method using genetic variants

to confirm causality between risk factor and outcome. Alternative techniques that may further improve prediction accuracy, and are particularly well suited for genetic data, include machine learning (ML) techniques such as random forest and neural networks.[284] The basic principle of ML relates to data inference; using estimates from past samples to predict new data using statistical, probabilistic and optimisation tools.[182] Moreover, ML allows predicting risks and outcomes for alternative populations based on minimal datasets requiring population socio-demographic characteristics only.[163] For example, these models can potentially be applied to understudied populations where the number of deaths due to COPD is the highest such as in India and Bangladesh. Studies like the Bangladesh Risk of Acute Vascular Events study (BRAVE) – a 16,000-person case-control study of CV disease – and BangladEsh Longitudinal Investigation of Emerging Vascular Events (BELIEVE) – a 100,000-person prospective cohort study in Bangladesh – could provide opportunities to study this and capture lung function measurements in addition to CV related ones. These datasets contain phenotypic and genotypic data including lung function test results and clinical outcomes.

More specifically, I propose future research directions related to assessing frailty and multi-organ tissue loss in COPD, assessing the clinical impact of risk prediction, estimating years of life lost (YLL) due to COPD, including genetic data in risk prediction, and I highlight the importance of considering the cost implications of risk prediction for the healthcare system.

8.3.0.1 Frailty and multi-organ tissue loss

Other areas of particular importance in COPD relate to frailty and predicting related events such falls and fractures using hospital admission data. In 2012, Gale *et al.* linked increased physical impairment and frailty in COPD patients.[100] Maddocks *et al.* assessed the prevalence of frailty in COPD and its effect on pulmonary rehabilitation completion.[166] Frailty was found in 25% of patients and resulted in non-completion of the program. Frailty is common in COPD, and there is increasing interest in assessing this within the COPD population.[24] Moreover, Lahouse and colleagues assessed frailty in COPD and the risk of mortality, and found increased

frailty with severe airflow limitation, shortness of breath, and frequent exacerbations.[153] As a result of osteoporosis – one of the clinical features of frailty – fractures are common in COPD patients and is characterised by decreased skeletal resistance.[37,110] The use of steroids in COPD is thought to contribute to these fractures but this has not been studied well. In addition, systemic inflammation is believed to prolong fracture healing time and increase complication rates.[50]

Frailty, also known as age-related physical disability (ICD-10 R54), can be defined using modified Fried criteria.[145] Although there is overlap, the difference between sarcopenia and frailty is that sarcopenia refers to muscle mass atrophy related to ageing, and frailty relates to a geriatric syndrome linked with a higher risk of falls, fractures and hospitalisation.[44] According to the European working group on sarcopenia in older people, sarcopenia (i.e. ICD-10 M62.84) is defined as low muscle mass and weakness measured by the fat free mass index (FFMI) and 4MGS with cut-off points of FFMI <8.5 kg/m^2 for men and <5.75 kg/m^2 for woman, and <0.8 m/s respectively.[60] Modified Fried criteria include (i) self-reported unintentional weight loss, (ii) muscle weakness derived through predicted quadriceps strength estimated using Seymour′s equation,[237] (iii) exhaustion measured by the COPD assessment test (CAT) [item 8 ≥ 3] or St. George′s respiratory questionnaire for COPD (SGRQ-C) [Q10.f], (iv) self-reported slowness of walking (<3 mph), slowness while walking measured by the SGRQ-C [Q12.c] or a score of <4 in 4MGS, and (v) low levels of activity measured by CAT [Q5] or SGRQ-C [Q13 or Q14]. A frailty score of 0 would be considered not frail, 1-2 pre-frail, and >2 as being frail. Preliminary findings in the ERICA and evaluation of COPD longitudinally to identify predictive surrogate end-points (ECLIPSE) cohorts show disease progression to be associated with increasing tissue loss, both pulmonary and extra-pulmonary.[41] Future research could examine the association of musculoskeletal markers and clinical outcomes in those with a so-called multi-organ loss of tissue COPD phenotype.

8.3.0.2 Net reclassification index

No significant loss in discriminative ability does not necessarily mean there is no meaningful difference. The C-index is a well-known model performance measure but merely an indicator of discriminative ability. Despite its popularity it has been argued that this metric has limitations such as limited clinical relevance, and when changes are marginal its interpretation becomes difficult. This was indeed the case when comparing the C-indices of the different models, based on the 6MW, SPPB or other measures. The net reclassification improvement (NRI) is a quantitative method – focussing on clinical outcomes rather than model performance – to estimate the number of individuals that are correctly or incorrectly reclassified. For example, when comparing the BODE based on 6MW with BODE based on SPPB the C-index may have no significant difference but the NRI may actually indicate the number of individuals that are being reclassified into a lower or higher BODE Index quartile. Despite NRI is popular and may indicate the potential public health implications, simultaneously it has been suggested to provide misleading results with high NRI statistics actually being a result of poorly fitted models.[217] In addition, clear risk thresholds are required that do currently not exist for the BODE Index, and with >3 risk categories, which is the case with the BODE Index, NRI may therefore potentially not be suitable.[146]

Alternatively, having a risk index with good discriminative ability and/or high NRI that is unlikely to be widely adopted in clinical practice, it might be worthwhile to make a trade-off between model performance and clinical practicality. Using indirect methods such as discrete choice experiments to elicit preferences of various stakeholders (e.g. clinicians and statisticians) may provide useful insights in how much discriminative ability stakeholders are willing to trade for improved clinical uptake (i.e. benefit-risk assessment).[269]

8.3.0.3 Years of life lost

Beyond standardised mortality rates, estimating YLL quantifies reduced life expectancy due to disease. It is a measure of premature mortality and considers the number of years an individual

would have lived without the disease. Years of life lost takes the age of death into account, providing bigger weights at younger age. Years of life lost can be estimated by sex and based on the number of deaths at different age categories and multiplied by the remaining years of life expected. It enables healthcare planners to set priorities in addressing disease interventions at a population level. Proposed analysis includes estimating sex-specific potential YLL due to premature all-cause and cause-specific mortality caused by COPD in both the ERICA and UK Biobank cohorts.

8.3.0.4 Genetic risk prediction

Conventional diagnostics rely primarily on spirometry. However, more recently there is increasing interest in stratifying individuals based on their genetic make-up and determine their risk for future events. Single gene testing provides information of a single gene function, whereas whole-genome sequencing (WGS) provides information about the full genetic makeup of the individual organism at a single point in time. It is believed that WGS will benefit patients through identifying individuals at high risk for common clinical outcomes such as early mortality and CV events including heart attacks and stroke. Genetic risk scores, where genomic data is integrated with conventional risk factors are increasingly being developed, as evidence indicates the inclusion of genetic data in risk prediction to outperform models based on traditional predictors only.[147]

Findings in the ERICA cohort indicated only very few cardiac deaths. Evidence indicates the possible involvement of a genetic component in COPD. Those affected could potentially benefit from individual risk profiling based on genetic or genomic sequencing – collectively referred to as next generation sequencing (NGS).[194] A possible causal factor could be a genetic determinant such as MMP-12 – a gene known to play a role in lung damage and associated with COPD onset.[130] Mendelian randomisation analysis suggest that increased levels of MMP-12 may protect patients from coronary heart disease but increases COPD susceptibility.[257] MMP-12 and other genetic activity can be measured using the appropriate gene arrays. Future research could include

to develop and validate a prediction model using, for example, ML techniques including genetic and other risk factors identified through MR analysis, in order to stratify COPD patients and identify those who are at high risk for clinical outcomes in an early stage of disease. Proposed analysis include to determine which stratifiers could identify those who are likely to have clinical outcomes, determine the association between selected biomarkers and their relationship with clinical outcomes, and if genetic markers such as MMP-12 protects patients from coronary heart disease but increase COPD susceptibility. Other analysis could include developing, validating and comparing the predictive ability of genetic and non-genetic risk models, assess the predictive ability of these models comparing population data of different cohorts, ideally from different countries, assess the feasibility of applying the risk models to understudied populations using ML techniques, and estimate the effectiveness of the different models.

8.3.0.5 Healthcare system

The number of people who could be saved, for example, if AECOPD could be diagnosed earlier depends highly on the ability of a healthcare system to identify high-risk individuals in an early stage of disease. Population-based screening programs allow screening at a single point in time or sequentially – depending on set risk thresholds an individual might require additional testing (e.g. genetic sequencing following spirometry) or may support improving lifestyle behaviour.[287] The impact of risk prediction, however, is often limited. Risk models frequently fail widespread clinical adoption due to their impracticalities, methodological shortcomings or associated costs.[67,113,267] In addition, providing individuals with personalised risk information has so far not shown to improve lifestyle or screening adherence.[264] With healthcare becoming increasingly complex and challenging, there is a need for a multidisciplinary approach. Epidemiology and health economics each consider different areas of evidence and can enhance the analysis, providing healthcare policy makers with more robust and comprehensive data. Whilst healthcare expenditure continues to rise, high emphasis is placed on the economic evaluation of healthcare innovation. Despite clear guidelines on evaluating clinical and cost effectiveness

exist,[199] a majority of published studies lack any health economic assessment.[268] For NGS this is partly attributed to the lack of clinical trials.[288] There is no current evidence on the additional health benefits produced by the extra information from NGS, or whether the analysis plus clinical action based on additional findings provides value for money to the healthcare system. Policy makers are unclear about the clinical utility of NGS and are concerned it could be an expensive addition to existing diagnostics.[117] In order to estimate the full impact of risk prediction, the clinical applicability and cost implications of genomic risk prediction needs to be evaluated.

Future research could assess the impact of genetic and genomic data, and the use of EHR data in predicting clinical outcomes whilst considering the clinical practicalities and cost implications. The objectives could be to develop a decision analytical framework for the economic evaluation (i.e. cost-effectiveness and budget impact analysis), calculate the costs of using conventional and genetic-based risk prediction, and estimate the incremental cost and effects of a genetic-based model to determine the clinical utility and health economic impact of risk prediction in COPD, and provide health policy recommendations for healthcare improvement.

Appendix A: research items authored during the PhD

Publications

- Schwarze *et al.* (2019). "The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom." In: *Genetics in Medicine*. DOI: 10.1038/s41436-019-0618-7

- Fermont *et al.* (2019). "Biomarkers and clinical outcomes in COPD - a systematic review and meta-analysis." In: *Thorax*. DOI: 10.1136/thoraxjnl-2018-211855.

- Fermont *et al.* (2017). "The EQ-5D-5L is a valid approach to measure health related quality of life in patients undergoing bariatric surgery." In: *PLOS ONE* 12.12, pp. 1-13. DOI: 10.1371/journal.pone.0189190.

- Fermont *et al.* (2017). "Clinical applicability and cost of a 46-gene panel for genomic analysis of solid tumours: Retrospective validation and prospective audit in the UK National Health Service." In: *PLOS Medicine* 14.2, pp. 1-26. DOI: 1key71/journal.pmed.1002230.

Conference proceedings

- Fermont *et al.* (2018). "OA2141 The value of short physical performance battery as an alternative component of the BODE Index in predicting death in COPD in the ERICA cohort." In Proceedings of the European Respiratory Society International Congress: vol. 52. Suppl 62. European Respiratory Journal. DOI: 10.1183/13993003.congress-2018.OA214.

- Fermont *et al.* (2017). "P78 Cardiovascular and musculoskeletal phenotypes and the clinical outcomes in COPD: a systematic review and meta-analysis." In Proceedings of

the British Thoracic Society Winter Meeting: vol. 72. Suppl 3. Thorax, A125-A126. DOI: 10.1136/thoraxjnl-2017-210983.220.

- Fermont *et al.* (2017). "P96 Death related to cardiovascular disease in chronic obstructive pulmonary disease." In Proceedings of the British Thoracic Society Winter Meeting: vol. 72. Suppl 3. Thorax, A135-A135. DOI: 10.1136/thoraxjnl-2017-210983.238.

Protocols

- Fermont *et al.* (2017). "Preferences of patients and clinicians for genomic diagnostic technologies in healthcare: a systematic review protocol." In: *PROSPERO*. CRD: 42017062294.

- Fermont *et al.* (2016). "Cardiovascular and musculoskeletal phenotypes and the clinical outcomes in chronic obstructive pulmonary disease: a systematic review and meta-analysis protocol." In: *PROSPERO*. CRD: 42016052075.

Submitted

- Fermont *et al.* (2019). "Can simpler measures substitute for the 6-minute walk component of the BODE Index?"

- Fermont *et al.* (2019). "Risk assessment for hospital admissions in patients with COPD; multi-centre prospective study in the UK."

In preparation

- Fermont *et al.* (2019). "Predicting fatal and non-fatal cardiovascular disease in COPD; do arterial stiffness or carotid intima-media thickness improve on a Framingham approach?"

- Fermont *et al.* (2019). "Causes of death in COPD using the UK Biobank Resource."

- Fermont *et al.* (2019). "The cost-effectiveness and budget impact of tumour profiling in lung cancer."

- Fermont *et al.* (2019). "Providers′ hidden cost of preauthorization: a cost model."
- Fermont *et al.* (2019). "Preferences of patients and clinicians for genomic diagnostic technologies in healthcare: A systematic review."

Appendix B: data completion form meta-analysis

TITLE:

STUDY ID:

META-ANALYSIS / META-REGRESSION

DATA COMPLETION FORM

GUIDANCE

- This questionnaire is designed to collect data on selected cardiovascular and musculoskeletal markers in clinically stable patients diagnosed with COPD measured at baseline for the outcome measures mortality, exacerbation and hospitalisation.
- For each outcome measure we are collecting the sample size, mean and standard deviation (SD), by those with and those without the event, for the variables 6MWD (six-minute walk distance), WBC (white blood cell count), log-CRP (C-reactive protein), IL-6 (interleukin-6), IL-8 (interleukin-8), fibrinogen, TNF-alpha, leukocytes, resting heart rate, QMVC (quadriceps maximal voluntary contraction) and SNIP (sniff nasal inspiratory pressure).
- Those with the event are defined as those with any number of exacerbations or hospitalisation. Those without the event are defined as those with no exacerbation or hospitalisation during the study period.
- In addition, for mortality we are collecting the unadjusted and adjusted (i.e. age, sex, BMI and smoking status) hazard ratio with 95% confidence interval (95% CI) for each variable. We will ask for which confounders you have been able to adjust for.
- If data were obtained from your publication it will already be completed but please check the values and correct if necessary. If you have not collected a particular outcome or variable please leave it blank.
- Please provide estimates in the indicated units. Also, note the preferred magnitude of effect.
- After completing please return to: jmf88@medschl.cam.ac.uk

STUDY DETAILS

DESIGN

Indicate the study type

☐Cohort

☐Case-control

☐Other, (e.g. nested case-control or RCT): *please describe study design*

LENGTH OF FOLLOW-UP

Indicate the minimum length of follow-up in months

Minimum: 36 *months*

OUTCOME: MORTALITY

Marker		Survivors			Non-survivors			Units used for effect measure	Unadjusted Hazard Ratio		Adjusted Hazard Ratio	
	Units	Sample size	Mean	SD	Sample size	Mean	SD		Estimate	95% CI	Estimate	95% CI
6MWD	meter	n	mean	SD	n	mean	SD	per 50 meter	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
WBC	mc/L	n	mean	SD	n	mean	SD	mc/L	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
log-CRP	mg/L	n	mean	SD	n	mean	SD	mg/L	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
IL-6	pg/ml	n	mean	SD	n	mean	SD	pg/ml	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
IL-8	pg/ml	n	mean	SD	n	mean	SD	pg/ml	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
Fibrinogen	g/dL	n	mean	SD	n	mean	SD	g/dL	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
TNF-alpha	pg/ml	n	mean	SD	n	mean	SD	pg/ml	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
Leukocytes	x 10^9	n	mean	SD	n	mean	SD	x 10^9	Unadjust	Lower –	Adjusted	Lower –

	cells/L							cells/L	ed HR	upper	HR	upper
Resting heart rate	bpm	n	mean	SD	n	mean	SD	per 10 bpm	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
QMVC	kg	n	mean	SD	n	mean	SD	per 5 kg	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper
SNIP	cmH_2O	n	mean	SD	n	mean	SD	per 10 cmH_2O	Unadjusted HR	Lower – upper	Adjusted HR	Lower – upper

ADJUSTMENT VARIABLES

Select the variables you have adjusted for in the estimation of adjusted hazard ratios

☐Age
☐Sex
☐BMI
☐Smoking status

OUTCOME: EXACERBATION							
Marker		Exacerbators			Non-exacerbators		
	Units	Sample size	Mean	SD	Sample size	Mean	SD
6MWD	meter	n	mean	SD	n	mean	SD
WBC	mc/L	n	mean	SD	n	mean	SD
log-CRP	mg/L	n	mean	SD	n	mean	SD
IL-6	pg/ml	n	mean	SD	n	mean	SD
IL-8	pg/ml	n	mean	SD	n	mean	SD
Fibrinogen	g/dL	n	mean	SD	n	mean	SD
TNF-alpha	pg/ml	n	mean	SD	n	mean	SD
Leukocytes	x 10^9 cells/L	n	mean	SD	n	mean	SD
Resting heart rate	bpm	n	mean	SD	n	mean	SD
QMVC	kg	n	mean	SD	n	mean	SD
SNIP	cmH_2O	n	mean	SD	n	mean	SD

OUTCOME: HOSPITALISATION							
Marker		Hospitalisation			Non-hospitalisation		
	Units	Sample size	Mean	SD	Sample size	Mean	SD
6MWD	meter	n	mean	SD	n	mean	SD
WBC	mc/L	n	mean	SD	n	mean	SD
log-CRP	mg/L	n	mean	SD	n	mean	SD
IL-6	pg/ml	n	mean	SD	n	mean	SD
IL-8	pg/ml	n	mean	SD	n	mean	SD
Fibrinogen	g/dL	n	mean	SD	n	mean	SD
TNF-alpha	pg/ml	n	mean	SD	n	mean	SD
Leukocytes	x 10^9 cells/L	n	mean	SD	n	mean	SD
Resting heart rate	bpm	n	mean	SD	n	mean	SD
QMVC	kg	n	mean	SD	n	mean	SD
SNIP	cmH_2O	n	mean	SD	n	mean	SD

This is the end of the data completion form. Thank you for your input! Please return the form to: jmf88@medschl.cam.ac.uk

Appendix C: ERICA protocol

COPD JOURNAL OF CHRONIC OBSTRUCTIVE PULMONARY DISEASE

informa healthcare

COPD, 11:552–559, 2014
ISSN: 1541-2555 print / 1541-2563 online

DOI: 10.3109/15412555.2014.898031

ORIGINAL RESEARCH

Evaluating the Role of Inflammation in Chronic Airways Disease: The ERICA Study

Divya Mohan,[1] Nichola S. Gale,[2] Carmel M. McEniery,[3] Charlotte E. Bolton[4], John R. Cockcroft,[2] William MacNee,[5] Jonathan Fuld,[3] David A. Lomas,[6] Peter M.A. Calverley,[7] Dennis J. Shale,[2] Bruce E. Miller,[8] Ian B. Wilkinson,[3] Ruth Tal-Singer,[8] and Michael I. Polkey,[1] on behalf of the ERICA Consortium

1 NIHR Respiratory Biomedical Research Unit, Royal Brompton & Harefield NHS Foundation Trust and Imperial College, London, UK

2 Department of Cardiology, Wales Heart Research Institute, Cardiff University, Cardiff, UK

3 Clinical Pharmacology Unit, Addenbrookes Hospital, University of Cambridge, Cambridge, UK

4 Nottingham Respiratory Research Unit, University of Nottingham, Nottingham, UK

5 MRC Centre for Inflammation, University of Edinburgh, Edinburgh, UK

6 Faculty of Medical Sciences, University College London, London, UK

7 School of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK

8 Respiratory Therapy Area Unit, GlaxoSmithKline, King of Prussia, PA, USA

Abstract

Extrapulmonary manifestations are recognized to be of increasing clinical importance in Chronic Obstructive Pulmonary disease. To investigate cardiovascular and skeletal muscle manifestations of COPD, we developed a unique UK consortium funded by the Technology Strategy Board and Medical Research Council comprising industry in partnership with 5 academic centres. ERICA (Evaluating the Role of Inflammation in Chronic Airways disease) is a prospective, longitudinal, observational study investigating the prevalence and significance of cardiovascular and skeletal muscle manifestations of COPD in 800 subjects. Six monthly follow up will assess the predictive value of plasma fibrinogen, cardiovascular abnormalities and skeletal muscle weakness for death or hospitalization.

As ERICA is a multicentre study, to ensure data quality we sought to minimise systematic observer error due to variations in investigator skill, or adherence to operating procedures, by staff training followed by assessment of inter- and intra-observer reliability of the four key measurements used in the study: pulse wave velocity (PWV), carotid intima media thickness (CIMT), quadriceps maximal voluntary contraction force (QMVC) and 6-minute walk distance (6MWT). This report describes the objectives and methods of the ERICA trial, as well as the inter- and intra-observer reliability of these measurements.

Introduction

Although COPD is primarily a lung disease, it is now widely recognised that COPD is a heterogeneous condition with a range of extra-pulmonary manifestations including cachexia (1), peripheral muscle dysfunction (2, 3), cardiovascular disease (4, 5) and osteoporosis (5, 6) that have an effect on the severity of the condition.

Two of these extrapulmonary manifestations, namely cardiovascular and skeletal muscle dysfunction, represent a key unmet need in patients with COPD that require the development of new therapies. Cardiovascular disease is the second-leading cause of death in patients with COPD (7), and even subjects with mild spirometric abnormalities have an increased risk of admission or death from cardiovascular causes (8). Similarly skeletal muscle weakness (2, 3) and biopsy abnormalities (9) exist even in patients with mild airflow obstruction and are associated with an increased risk of death (10).

A combination of systemic and local factors such as physical inactivity, oxidative stress, cachexia, exposure to cigarette smoke and inflammation are thought

Keywords: inflammation, muscle, Cardiovascular, extrapulmonary, fibrinogen, arterial stiffness

Correspondence to: Dr Divya Mohan, Muscle Laboratory, Level 1 Fulham wing, Royal Brompton Hospital, London, SW3 6NP, United Kingdom, phone: 0207 351 8029, email: d.mohan@rbht.nhs.uk

to contribute towards the skeletal muscle dysfunction seen in COPD (11). Since pulmonary rehabilitation is a highly effective therapy in COPD that increases quadriceps strength (12), targeting this abnormality is likely to translate into patient benefit. Objectively measured physical activity relates to muscle mass (3), especially in mild disease, and in a survival analysis by Waschki and co-workers (13), the combination of physical activity measurement and assessment of vascular status predicted mortality better than either alone, suggesting that the cardiovascular and skeletal muscle phenotypes are not identical.

Persistent systemic inflammation has been linked with poorer outcomes in COPD and has been identified as a novel COPD phenotype (14). Recent data suggest that fibrinogen is a promising, stable biomarker of systemic inflammation, and that elevated fibrinogen levels relate to frequent exacerbations and mortality in COPD (15–17). Previous studies have suggested that almost a third of COPD patients suffer from 2 or more exacerbations per year, with a fifth of COPD patients requiring hospitalizations over the course of 1 year (18). For these reasons, the COPD Biomarkers Qualification Consortium (CBQC) has submitted fibrinogen for consideration for qualification as a drug development tool by the Food and Drug Administration (FDA). CBQC was established in 2010 with the aim of collating anonymised data from clinical and observational trials (Figure 1).

Trials such as ECLIPSE (19), GSK-supported investigator-sponsored eclipse extension study NTR3221, ARCADE (20), PROactive (Clinicaltrials.gov number NCT01388218) and MRC/ABPI WP4 (Clinicaltrials.gov number NCT01620645) will contribute data towards ERICA, thus allowing a sufficiently large dataset to conclusively establish the value of biomarkers or drug development tools (DDTs) as stratification tools (21). Nevertheless there remains a paucity of data to assess whether fibrinogen will also be a satisfactory biomarker for extrapulmonary manifestations of COPD.

Plasma fibrinogen independently predicts cardiovascular risk in the general, healthy population (22), however, the value of fibrinogen in the prediction of the cardiovascular and skeletal muscle manifestations of COPD, and in the interplay between these phenotypes, requires further evaluation. For this reason we conceived the ERICA (Evaluating the Role of Inflammation in Chronic Airways disease) study. The study has three specific aims. First, to determine how effectively plasma fibrinogen predicts the cardiovascular and/or skeletal muscle manifestations of COPD. Second, to determine how fibrinogen and other specific measures of cardiovascular and muscle function predict longer-term outcomes including death, disability and hospital admission, and third, to determine the extent to which subsets of COPD patients with cardiovascular or muscle manifestations overlap. For this purpose, a cardiovascular manifestation is defined as an abnormally raised aortic pulse wave velocity and a skeletal muscle manifestation is defined as quadriceps muscle weakness.

The current report describes the objectives and methods of the ERICA trial, and the standardisation procedures undertaken with the objective to improve inter- and intra-observer reliability of measurements used in the study.

Methods

Subjects

A maximum of 800 COPD patients are to be recruited over a period of 2 years. The study is powered on the basis of a tertile analysis of the two key cardiovascular and muscular biomarkers, systemic arterial stiffness as measured by aortic pulse wave velocity (PWV) and skeletal muscle function, measured as quadriceps maximal voluntary contraction (QMVC). Assuming an average PWV of 10 (SD 1.0) m/s and a minimal clinically relevant difference of 0.4 m/s, 230 patients per tertile will provide 90%

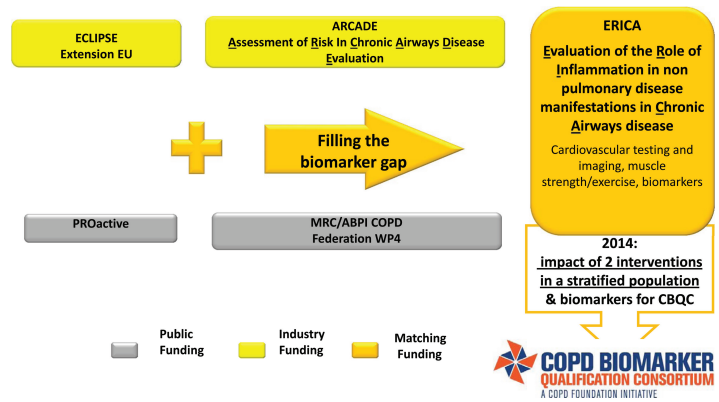


Figure 1. Clinical and Observational studies contributing data towards the COPD Biomarker Qualification Consortium (CBQC) and their sources of funding.

Table 1. Inclusion and exclusion criteria for study participants

Inclusion	Exclusion
Clinical diagnosis of COPD	Inability to provide written, informed consent
Baseline post-bronchodilator Forced expiratory volume in the first second (FEV_1) of 80% or less of their predicted value, a baseline FEV_1/forced vital capacity (FVC) ratio of < 0.7	A known diagnosis of α1-antitrypsin deficiency, known neurological or skeletal muscle disease
Age > 40 years	Pregnancy
A smoking history of at least 10 pack years	Ongoing participation in a trial of an experimental drug
Clinical stability for at least 4 weeks, without any hospitalisations or exacerbations requiring treatment at the time of study measurements	

power at $p < 0.01$ to detect this difference between the top and bottom quartiles. For QMVC, assuming an average QMVC of 32 (SD 8) kg, 220 patients per tertile will provide 90% power at a significance of $p < 0.01$ in order to detect the minimum clinical difference of 3 kg between the top and bottom tertiles. Allowing for a 10% dropout rate and incomplete datasets, approximately 800 patients were calculated to be required. Recruitment is on target to finish in autumn 2013. Table 1 describes the inclusion and exclusion criteria for subject participation in the study. All participants provided written, informed consent.

Study design

ERICA is an on-going longitudinal, observational, prospective study being conducted at 5 centres in the UK, which is presently funded for 2 years by the UK Technology Strategy Board/MRC. As the study was not a trial, the study is registered with the UK Clinical Research Network Study Portfolio with UKCRN ID 11101 (http://public.ukcrn.org.uk/Search/StudyDetail.aspx?StudyID=11101); the UKCRN is a publically searchable database.

Following baseline visits to perform study measurements, participants are followed up at 6 monthly intervals for 2 years with telephone or postal questionnaires to assess the frequency of COPD exacerbations. For this study, we have defined exacerbations as self-reported increase in COPD symptoms that required treatment with antibiotics and/or steroids and severe exacerbations as those that require hospital admission. The development of cardiac co-morbidity is assessed through new self-reported cardiac symptoms such as exertional chest pain or ankle swelling, physician diagnoses of angina, myocardial infarction, stroke or hypertension and the introduction of new concomitant cardiac medication since the last patient visit or questionnaire.

To assess the impact of COPD on the patient, we used Medical Research Council (MRC) dyspnoea scores, COPD Assessment Tool and the St George's Respiratory COPD Questionnaire, whilst physical activity is self-reported. There are no prohibited medications in the study. All subjects continued their routine prescribed medications throughout the study and the patient's physician may offer treatments (e.g. medication change, rehabilitation) in line with the patients' needs. These treatment changes are captured at the 6 monthly calls/questionnaires. The study is being conducted in accordance with the Declaration of Helsinki and good clinical practice guidelines.

The research was given a favourable opinion by the Cambridge South East Research Ethics Committee and the local research and development departments at each participating site. The trial steering committee comprises physicians and scientists from five UK academic centres, two academic physicians independent of the recruiting centres and representatives from GlaxoSmithKline.

Outcome measurements

Study assessments are performed at baseline over two visits. Standardised procedures are used in all centres as defined in the study procedure manual. Measurements that are of primary interest are plasma fibrinogen, aortic PWV, carotid intima media thickness (CIMT), 6-minute walk distance (6MWT) and QMVC. Methods for these 5 procedures are described here, but all study parameters are listed in Table 2. For all study procedures a Standard Operating Procedure (an SOP) was generated to which all partners adhered. Patients will be registered for long-term health outcomes through Hospital Episode Statistics (HES), a central UK database recording all admissions to National Health Service (NHS) hospitals, and the NHS Information Centre from the Office for National Statistics (previously the Medical Research Information Service), which can report on the status of study participants and provide follow up data for longitudinal studies within the UK.

Fibrinogen

For determination of plasma fibrinogen, whole blood is collected into a vacutainer tube (sodium citrate as the anticoagulant) by venepuncture from a peripheral vein after a 4-hour fast. Plasma is prepared by centrifugation at 2000 × *g* for 10 min. Plasma fibrinogen is measured in fresh plasma samples using an automated, modified Clauss method [HemosIL Fibrinogen-C XL, Instrumentation Laboratories(23)]. The assay method is a direct measurement of functional fibrinogen and is the method most commonly used in clinical laboratories. Daily testing on a fibrinogen calibrator was carried out at the Royal Brompton and Harefield NHS hospital laboratory and Addenbrookes hospital laboratory to assess inter-assay variability.

Table 2. Description of standardized assessments carried out on all study subjects as described in the study manual and existing studies that are carrying out these assessments for the COPD Biomarker Qualification Consortium

Assessment	Description	Existing studies contributing data from assessments towards ERICA
Post bronchodilator Spirometry: FEV_1 and FVC	Performed within one hour of administration of patient's own bronchodilators	ARCADE, ECLIPSE extension, PROactive, Longitudinal determination of skeletal muscle dysfunction in COPD, MRC WP4 consortium
Cardiovascular Assessments		
Blood Pressure	Average of final 2 out of 3 measurements, taken after 10 minutes of rest	ARCADE, ECLIPSE extension, Longitudinal determination of skeletal muscle dysfunction in COPD, MRC WP4 consortium
12-lead Electrocardiogram (ECG)	A standard 12 lead (10 electrode) ECG recorded with the patient in a supine position and at a recording speed of 25 mm/sec	ARCADE, PROactive, MRC WP4 consortium
Arterial Stiffness	Assessments of carotid-femoral pulse wave velocity (PWV) and aortic augmentation index (AIx) via Sphygmocor device after 4 hours of fasting and 6 hours without bronchodilators.	ARCADE, Longitudinal determination of skeletal muscle dysfunction in COPD
Carotid Intima Media Thickness (CIMT)	B-mode ultrasound measurement of bilateral common carotid intima media thickness to assess subclinical atherogenesis (25)	ARCADE
Skeletal muscle assessments		
Quadriceps Maximal Volitional Contraction (QMVC)	Best effort from 6 volitional, isometric quadriceps contractions of the right leg as described by Edwards et al. (27). Predicted QMVC was calculated according to Seymour et al. (2)	PROactive, Longitudinal determination of skeletal muscle dysfunction in COPD, MRC WP4 consortium
Sniff Nasal Inspiratory Pressure (SNIP)	A non-invasive measure of inspiratory muscle strength using a hand-held MicroRPM (respiratory pressure meter). The most negative of a minimum 3 efforts will be used for data analysis	PROactive
Exercise/ Physical performance assessments		
6-minute walk test (6MWT)	Supervised walking test as per American Thoracic Society guidelines (26), but without a practice test	ARCADE, ECLIPSE extension, PROactive, MRC WP4 consortium
Short Physical Performance Battery	A composite assessment of lower extremity function comprising standing balance, 4 meter gait speed and sitting-to-standing speed (35)	MRC WP4 consortium
Anthropometrics	Height, weight, body mass index (BMI).	ARCADE, ECLIPSE extension, PROactive, Longitudinal determination of skeletal muscle dysfunction in COPD, MRC WP4 consortium
Bioimpedance and Fat-free mass	Estimated using single-frequency (50 kHz) bioelectrical impedance analysis via TANITA BC 418 MA (Tanita Corporation, Tokyo, Japan).	ARCADE, ECLIPSE extension, PROactive, Longitudinal determination of skeletal muscle dysfunction in COPD, MRC WP4 consortium
Health Outcomes		
Health status and symptom assessment	COPD specific St George's COPD respiratory questionnaire (SGRQ-C) and COPD assessment tool (CAT)	ARCADE, ECLIPSE extension, PROactive, Longitudinal determination of skeletal muscle dysfunction in COPD, MRC WP4 consortium
Breathlessness, exacerbation frequency and treatment, physical activity and smoking habit	Medical Research Council (MRC) dyspnoea score and postal questionnaire at baseline and repeated at 6 monthly intervals for 2 years	
Biomarkers		
Blood samples	Taken after 4 hours of fasting. Samples stored for plasma, serum and DNA and analysed for plasma fibrinogen, urea and electrolytes, full blood count, high sensitivity C-reactive protein, glucose, Hba1c and lipid profile.	
Urine	Spot urine sample; taken after 4 hours of fasting	

Pulse wave velocity

Following 10 minutes of supine rest, brachial blood pressure was measured three times, and an average of the final two readings was used for analysis. Aortic PWV is measured via the SphygmoCor device (AtCor, West Ryde, Australia), between the carotid and femoral arteries, using a piezoelectric tonometer placed over the artery and ECG gating, as previously described in detail (24). The path length is calculated by subtracting the distance between the carotid pulse and supra-sternal notch, from the femoral artery supra-sternal notch distance. Measurements are made following 4 hours of fasting, and 6 hours without bronchodilator use.

Carotid intima media thickness

Carotid intima media thickness (25) was measured via B-mode ultrasound, using a 7–12 MHz linear probe. Measurements were taken after 10 minutes of supine rest. Both

the right and left common carotid arteries are scanned at a distance of 1 cm from the carotid bulb. Images are not ECG gated, and three 10 second loops are recorded for each carotid artery. Images are then transferred in DICOM format to be analysed via Vascular Tools 5 software (Medical Imaging Application LLC, Coralville, USA).

Six-minute walk test

Six-minute walk distance is measured in accordance with the guidelines of the American Thoracic Society (26) except that a practice walk was not performed due to time constraints. Although subjects could set their own walking pace, it was emphasized that they cover as many laps as possible over a standard 30 m, level track during the 6 minutes. Subjects were permitted to use their usual medications prior to the test, and were given standardised encouragement only at the end of each minute during the walking test. Where oxygen was required during the walking test, an additional researcher carried the oxygen cylinder for the patient, but behind the patient in order not to influence the patient's pace.

Quadriceps maximal volitional contraction

QMVC force was measured using the technique of Edwards et al. (27) and expressed as a percentage of predicted values using the equations developed by Seymour and co-workers (2). Patients were verbally encouraged to make a maximal contraction by pushing out (i.e. extension) against an inextensible strap placed above the ankle. The manoeuvre is repeated six times with a minimum 20-second interval between efforts. We used the highest value of contraction which could be sustained for 1 second for analysis.

Harmonising inter-site data collection

Prior to study recruitment, all centres participated in centralised training, individual site training and standardisation visits. Amongst study measurements, QMVC, 6MWT, PWV and CIMT were identified as most prone to systematic error due to variability in equipment and expertise across sites. Intra- observer reliability measurements were therefore carried out using 10 volunteer subjects at each site. Additionally, unlike the 6MWT, QMVC measurement was novel to most sites, and requires the observer to verbally encourage as well as correct patient technique, therefore inter-observer reliability measurements were carried out on 10 volunteers at each site. Intra-class correlation coefficients were used to measure inter- and intra-user reliability. Statistical analysis was carried out using IPB SPSS v 19.

Results

Inter-assay %CV values based on a fibrinogen calibrator tested daily are 6.7% for Royal Brompton Hospital and 9.4% for Cambridge University Hospital.

Intra- and inter-observer variability are shown in Table 3. The initial intra-class correlation coefficient (ICC) at one centre for QMVC was entered erroneously; a repeat set revealed an ICC of 0.60. As this was less than the agreed target of 0.85, recommendations were implemented and a final intra-observer repeatability ICC of 0.96 was observed.

Table 3. Intra-observer reliability for Pulse Wave Velocity (PWV), 6-minute walk test (6MWT) and quadriceps maximal volitional contraction (QMVC) and inter-observer reliability measurements for QMVC in 10 volunteers as measured by final intra-class correlation coefficients

	Intra-observer reliability (n = 10)				Inter-observer reliability (n = 10)
Site	PWV (m/s)	6MWT (m)	QMVC (kg)	CIMT (mm)	QMVC (kg)
1	0.99	0.98	0.97	0.98	0.99
2	1.00	0.92	0.90	0.99	0.88
3	1.00	0.98	0.98	0.97	0.98
4	0.99	0.99	0.96	0.89	0.92
5	0.96	0.99	0.96	0.94	0.98
Overall	1.00	0.99	0.95	0.98	0.98

Intra-class correlation coefficients for intra-observer reliability of Pulse Wave Velocity (PWV), 6 minute walk test (6MWT), Quadriceps Maximal Volitional Contraction (QMVC) and Carotid intima media thickness (CIMT), and inter-observer reliability for QMVC measurements, measured on 10 volunteers at each of the 5 participating sites.

Once a site demonstrated competency in all relevant research techniques, they were allowed to recruit study participants. Table 4 reports the baseline characteristics of the first 10 subjects recruited at each site, in order to provide an example of the likely eventual type of patients who will be recruited to the cohort.

Discussion

The main conclusions drawn from setting up ERICA are first that technically demanding measurements, like PWV, carotid media thickness and maximal voluntary contraction force can be made in patients with COPD. Secondly, with relatively little training measurements can be made with good repeatability and low inter-observer variability. Finally, participants recruited to the trial so far appear representative of a typical convenience cohort for COPD trials and there does not appear to be a great deal of variance between individual sites.

Harmonising inter-site data collection is essential in multi-centre studies for the production of valid, reliable results. Centralised training followed by individual site visits has enabled standardisation of techniques, and collection of inter- and intra-observer reliability allowed identification of problems prior to the commencement of study recruitment. We would therefore endorse the current practice that researcher competency in performing novel techniques should be formally evaluated in multi-site trials prior to subject recruitment to ensure good quality data.

Currently if planning a large trial of either an anabolic or cardiovascular therapeutic that addresses extrapulmonary disease manifestations in COPD, an investigator would be hampered by insufficient detail regarding subsets of patients who are most likely to

Table 4. Baseline characteristics of the first 50 study recruits at participating sites

	Centre 1 $n = 10$	Centre 2 $n = 10$	Centre 3 $n = 10$	Centre 4 $n = 10$	Centre 5 $n = 10$	Overall $n = 50$
Age (years)	68.6 (12.2)	65.9 (7.2)	68.1 (9.0)	67.1 (6.3)	71.8 (5.9)	68.3 (8.3)
Male: Female	10:0	4:6	5:5	6:4	9:1	34:16
Height (m)	173.3 (10.2)	165.8 (10.8)	165.5 (7.6)	170.8 (9.7)	172.2 (5.4)	169.5 (9.2)
Weight (kg)	85.2 (22.4)	72.8 (16.9)	66.4 (15.1)	80.0 (19.6)	78.1 (16.2)	76.5 (18.7)
Fat-free Mass (kg)	59.2 (11.9)	48.9 (12.2)	45.4 (9.8)	53.9 (14.1)	55.8 (6.9)	52.6 (11.9)
BMI (kg/m^2)	27.9 (5.4)	26.3 (4.8)	24.2 (5.3)	27.2 (5.5)	26.3 (5.1)	26.4 (5.1)
Smoking (pack-years)	52 (29)	32 (14)	39 (23)	66 (31)	53 (31)	48 (28)
Current smokers (%)	10	10	60	10	10	20
FEV_1 (% predicted)	47.4 (18.5)	57.8 (10.5)	59.7 (18.3)	60.1 (15.5)	44.3 (17.2)	53.9 (17.0)
FEV_1 (L)	1.34 (0.53)	1.43 (0.39)	1.43 (0.51)	1.61 (0.43)	1.26 (0.54)	1.41 (0.48)
FVC (L)	3.29 (1.09)	2.50 (0.58)	3.02 (0.80)	3.10 (0.87)	3.50 (0.63)	3.08 (0.85)
FEV_1/ FVC	0.42 (0.13)	0.57 (0.08)	0.47 (0.10)	0.53 (0.13)	0.34 (0.12)	0.47 (0.14)
6MW distance (m)	354 (113)	324 (100)	425 (112)	352 (97)	384 (118)	368 (109)
Fibrinogen (g/L)	3.5 (0.72)	3.89 (0.95)	3.44 (0.44)	3.24 (0.59)	3.5 (0.42)	3.51 (0.66)
Seated Systolic BP (mm Hg)	140 (15)	140 (18)	142 (21)	133 (14)	147 (18)	140 (17)
Seated Diastolic BP (mm Hg)	77 (6)	77 (13)	80 (12)	79 (5)	83 (10)	79 (12)
PWV (m/s)	10.5 (1.8)	9.3 (3.9)	11.7 (3.9)	7.5 (2.8)	11.8 (3.1)	10.2 (3.5)
QMVC (kg)	35.4 (12.6)	31.7 (15.0)	29.1 (7.7)	34.5 (8.6)	36.0 (9.4)	33.4 (10.8)
QMVC (% predicted)	71.5 (16.5)	73.9 (26.6)	74.8 (16.3)	77.5 (15.3)	79.3 (18.6)	75.4 (18.5)
SNIP (cm H_2O)	70.9 (12.1)	59.1 (22.6)	72.7 (18.7)	72.4 (16.6)	55.4 (19.7)	66.1 (19.0)
SPPB score (of 12)	9.8 (1.87)	9.9 (2.33)	10.4 (1.78)	10.7 (1.25)	10.6 (1.71)	10.3 (1.78)
MRC score (of 5)	2.8 (0.92)	2.8 (1.03)	2.1 (0.74)	2.1 (0.74)	2.9 (1.29)	2.54 (0.99)
SGRQ-C score	48.8 (16.4)	59.3 (20.2)	39.4 (13.6)	47.5 (20.8)	45.7 (21.7)	48.1 (19.1)
CAT score	17.8 (6.9)	19.5 (9.0)	17.7 (6.7)	17.5 (9.5)	16.5 (7.7)	17.8 (7.8)

Values expressed as mean, with standard deviation in brackets.

benefit. Although cardiovascular and skeletal muscle dysfunction are serious and common co-morbidities in COPD, they are not observed in all patients diagnosed with the disease. At this time it is also unclear whether there is an association between manifestations of skeletal muscle or cardiovascular dysfunction in COPD and the knowledge of their functional consequences is limited. It would be particularly attractive to have a blood biomarker which permitted selection of patients with these disease manifestations for clinical trials assessing efficacy as large all comer trials are likely to fail (28).

More data are available for cardiovascular disease than skeletal muscle weakness, but in both cases the evidence is that these extra-pulmonary manifestations of COPD are present only in a minority of patients. In a review of approximately 46,000 case records of patients managed in a Kaiser Permanante program several cardiovascular co-morbidities were identified in a minority of patients (at most 25%) though the group as a whole had a 2–3-fold increased risk of subsequent cardiovascular-related admission (8). In relation to the measures used in this study only very small data sets exist; in the study of Maclay et al. (29) roughly half the participants were above the threshold of 10 m/s considered to represent increased future cardiovascular risk. Skeletal muscle dysfunction, while common, is present in only a minority of patients with COPD whether judged by weakness (2) or muscle size (3). Interestingly the prevalence of skeletal muscle weakness, probably around 30% of patients, is not greatly influenced by disease severity judged by FEV_1 (3). We have previously discussed the difficulties of developing an anabolic agent for patients with COPD elsewhere (28).

Fibrinogen is attractive as a biomarker, as it is a commonly available test, acceptable, relatively inexpensive and easy measure in clinical practice. Inflammatory markers such as TNF-alpha, IL-6, CRP and p-selectin have been shown to relate to disease severity in some studies, however individual variability for these markers is high (30). A panel of 34 inflammatory markers was recently assessed in the ECLIPSE study, with plasma fibrinogen emerging as the most repeatable biomarker in stable patients with COPD. Although other inflammatory markers such as C-reactive protein and interleukin-6 were also raised in COPD patients as compared to healthy controls, these biomarkers displayed wide variability in stable subjects with COPD over 3 months. Fibrinogen has additionally been shown to relate to

disease severity and is predictive of death both in COPD and other conditions (15,31).

However it is largely unknown to what extent fibrinogen has a predictive value to diagnose cardiovascular dysfunction and especially skeletal muscle dysfunction in COPD. Data from very large studies have demonstrated that COPD patients with a self-reported history of cardiovascular disease have higher fibrinogen levels (16), but the relationship between fibrinogen and other measures with predictive value for future cardiovascular disease, such as pulse wave velocity and carotid intima media thickness used in ERICA, is unknown. Eickhoff and colleague used a third detailed and predictive measure, flow mediated dilatation, and found no relationship with fibrinogen, but their cohort was limited to 60 patients (32). Several studies that have investigated fibrinogen and cardiovascular disease in COPD have often excluded those with severe cardiovascular disease (15), which runs the risk of biasing results, and also leaves a data gap for patients most at risk from cardiovascular death. Importantly therefore pre-existing cardiovascular disease is not an exclusion criterion for our study.

Although both quadriceps strength, measured as QMVC and fibrinogen were related to lower physical activity in COPD patients measured by accelerometry in a prior smaller report (33), its relationship to skeletal muscle strength has not been examined before in a large cohort of patients. Our study additionally measures several aspects of muscle function, specifically QMVC, Sniff Nasal Inspiratory Pressure (SNIP) (34), Short Physical Performance Battery (SPPB) (35) and 6-minute walk distance. The short physical performance battery may prove to be of particular interest from a regulatory perspective since it is widely used in academic gerontology. The 4-metre gait speed, which is a component of the SPPB predicts death in elderly people and is reproducible in COPD (36).

Conclusions

In conclusion, ERICA is the first large prospective study to examine the interplay between fibrinogen, skeletal muscle and cardiovascular manifestations of COPD, as well as their relation to exacerbations and mortality. At the conclusion of the study we will be able to determine whether cardiovascular and muscle dysfunction phenotypic "sets" commonly overlap and to what extent fibrinogen is a useful marker of these sets. Identification of the relationship between co-morbidities and potential predictive biomarkers of COPD will help the development of future therapies, and may be useful diagnostically.

Acknowledgments

Authors Mohan, Wilkinson, Singer, and Polkey contributed equally to this work.

With special thanks to Sridevi Nagarajan and, Mellone Marchong and Jessica Middlemiss from the Cambridge Clinical Trials Unit, Addenbrookes Hospital, Cambridge, CB2 2QQ, the research teams at the participating centres and Suresh Chahwala from the Technology Strategy Board, North Star House, North Star Avenue, Swindon, SN2 1UE.

This work was supported by the NIHR Respiratory Biomedical Research Unit at the Royal Brompton and Harefield NHS Foundation Trust and Imperial College who part fund the salaries of MIP and DM. IBW and CMM are funded by the British Heart Foundation and the Cambridge Biomedical Research Centre, and receive support for the East Anglian CLRN. All 5 study centres are supported by the NIHR Clinical Research Network.

Declaration of Interest Statement

The authors have no conflicts of interest. Michael Polkey, Divya Mohan, Nichola Gale, Carmel McEniery and Ian Wilkinson have had grants paid to their institute by the Technology Strategy Board. Charlotte Bolton was paid for an educational, non-promotional lecture in 2011 and her institute received consultancy fees from an advisory board in 2012. John Cockcroft has received consultancy fees from Merck, Forest Pharmaceuticals and GlaxoSmithKline (GSK), lecturing fees from Menarini and Sanofi, and his institute has received grants from the Technology Strategy Board and Medical Research Council. William MacNee has received royalties from Health Press Ltd, Hodder & Stoughton Ltd, People's Medical Publishing House, Imperial College Press Ltd, consultancy fees from Pfizer, GSK, Novartis, Almirall, Janssen and received payments from GSK, Novartis and Pfizer and their institute has received money from the Technology Strategy Board. David Lomas has received board membership, lecture, travel expenses and consultancy fees from GSK, consultancy fees from Talecris/Grifols and travel expenses from Boehringer Ingelheim (BI), and his institute has received grants from GlaxoSmithKline. Peter Calverley has received travel expenses from the Medical Research Council, board membership fees from Nycomed, BI, consultancy fees from Novartis, expert testimony fees from Forest, payment for lectures from GSK, Astrazeneca, Takeda and travel expenses from BI, and his institute has received lecture fees from Novartis and Pfizer. Bruce Miller is an employee of GSK, and is married to a GSK employee and shareholder. Ruth Tal-Singer is employed by GSK and owns shares in GSK. Michael Polkey has received consulting fees from Novartis and Philips, travel support from GSK Almirall and BI and his institute has received consulting fees from BI, Lilly, Regeneron and GSK, lecture fees from Chiesi as well as grants from Technology Strategy Board, AstraZeneca and GSK. The authors alone are responsible for the writing and the content of the paper.

References

1. Schols AM, Slangen J, Volovics L, Wouters EF. Weight loss is a reversible factor in the prognosis of chronic obstructive

pulmonary disease. Am J Respir Crit Care Med 1998; 157(6 Pt 1):1791–1797.
2. Seymour JM, Spruit MA, Hopkinson NS, Natanek SA, Man WD, Jackson A, et al. The prevalence of quadriceps weakness in COPD and the relationship with disease severity. Euro Respir J 2010; 36(1):81–88.
3. Shrikrishna D, Patel M, Tanner RJ, Seymour JM, Connolly BA, Puthucheary ZA, et al. Quadriceps wasting and physical inactivity in patients with COPD. ERJ 2012; 40(5):1115–1122.
4. McAllister DA, Maclay JD, Mills NL, Mair G, Miller J, Anderson D, et al. Arterial stiffness is independently associated with emphysema severity in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2007; 176(12):1208–1214.
5. Sabit R, Bolton CE, Edwards PH, Pettit RJ, Evans WD, McEniery CM, et al. Arterial stiffness and osteoporosis in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2007; 175(12):1259–1265.
6. Bolton CE, Ionescu AA, Shiels KM, Pettit RJ, Edwards PH, Stone MD, et al. Associated loss of fat-free mass and bone mineral density in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2004; 170(12):1286–93.
7. Mannino DM, Doherty DE, Sonia Buist A. Global Initiative on Obstructive Lung Disease (GOLD) classification of lung disease and mortality: findings from the Atherosclerosis Risk in Communities (ARIC) study. Respir Med 2006; 100(1):115–122.
8. Sidney S, Sorel M, Quesenberry CP, Jr., DeLuise C, Lanes S, Eisner MD. COPD and incident cardiovascular disease hospitalizations and mortality: Kaiser Permanente Medical Care Program. Chest 2005; 128(4):2068–2075.
9. Natanek SA, Gosker HR, Slot IG, Marsh GS, Hopkinson NS, Man WD, et al. Heterogeneity of quadriceps muscle phenotype in chronic obstructive pulmonary disease (COPD); implications for stratified medicine? Muscle Nerve 2013; 48(4):488–497.
10. Swallow EB, Reyes D, Hopkinson NS, Man WD, Porcher R, Cetti EJ, et al. Quadriceps strength predicts mortality in patients with moderate to severe chronic obstructive pulmonary disease. Thorax 2007; 62(2):115–120.
11. Gea J, Agusti A, Roca J. Pathophysiology of muscle dysfunction in COPD. J Appl Physiol 2013; 114(9):1222–1234.
12. Troosters T, Gosselink R, Decramer M. Short- and long-term effects of outpatient rehabilitation in patients with chronic obstructive pulmonary disease: a randomized trial. Am J Med 2000; 109(3):207–212.
13. Waschki B, Kirsten A, Holz O, Muller KC, Meyer T, Watz H, et al. Physical activity is the strongest predictor of all-cause mortality in patients with COPD: a prospective cohort study. Chest 2011; 140(2):331–342.
14. Agusti A, Edwards LD, Rennard SI, MacNee W, Tal-Singer R, Miller BE, et al. Persistent systemic inflammation is associated with poor clinical outcomes in COPD: a novel phenotype. PloS One 2012; 7(5):e37483.
15. Duvoix A, Dickens J, Haq I, Mannino D, Miller B, Tal-Singer R, et al. Blood fibrinogen as a biomarker of chronic obstructive pulmonary disease. Thorax 2013; 68:670–676.
16. Valvi D, Mannino DM, Mullerova H, Tal-Singer R. Fibrinogen, chronic obstructive pulmonary disease (COPD) and outcomes in two United States cohorts. Int J Chron Obstruct Pulmon Dis 2012; 7:173–182.
17. Thomsen M, Dahl M, Lange P, Vestbo J, Nordestgaard BG. Inflammatory biomarkers and comorbidities in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2012; 186(10):982–988.
18. Hurst JR, Vestbo J, Anzueto A, Locantore N, Mullerova H, Tal-Singer R, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. N Engl J Med 2010; 363(12):1128–1138.
19. Vestbo J, Anderson W, Coxson HO, Crim C, Dawber F, Edwards L, et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). European Respir J 2008; 31(4):869–873.
20. Gale NS, Munnery M, Albarrati A, Munnery I, Shale DJ, Cockcroft JR. Assessment of risk factors in chronic airways disease evaluation (ARCADE). Artery Res 2012; 6(4):158.
21. Casaburi R, Celli B, Crapo J, Criner G, Croxton T, Gaw A, et al. The COPD Biomarker Qualification Consortium (CBQC). J COPD. 2013; 10(3):367–377.
22. Kaptoge S, Di Angelantonio E, Pennells L, Wood AM, White IR, Gao P, et al. C-reactive protein, fibrinogen, and cardiovascular disease prediction. N Engl J Med 2012; 367(14):1310–1320.
23. Clauss A. [Rapid physiological coagulation method in determination of fibrinogen]. Acta Haematol 1957; 17(4):237–246.
24. Wilkinson IB, Fuchs SA, Jansen IM, Spratt JC, Murray GD, Cockcroft JR, et al. Reproducibility of pulse wave velocity and augmentation index measured by pulse wave analysis. J Hypertens 1998; 16(12 Pt 2):2079–2084.
25. Csanyi A, Egervari A. Simple clinical method of average intima-media thickness measurement in the common carotid artery. VASA Zeitschrift fur Gefasskrankheiten 1996;25(3):242–248.
26. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166(1):111–117.
27. Edwards RH, Young A, Hosking GP, Jones DA. Human skeletal muscle function: description of tests and normal values. Clin Sci Mol Med 1977; 52(3):283–290.
28. Steiner MC, Roubenoff R, Tal-Singer R, Polkey MI. Prospects for the development of effective pharmacotherapy targeted at the skeletal muscles in chronic obstructive pulmonary disease: a translational review. Thorax 2012; 67(12):1102–1109.
29. Maclay JD, McAllister DA, Mills NL, Paterson FP, Ludlam CA, Drost EM, et al. Vascular dysfunction in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2009; 180(6):513–520.
30. MacNee W. Systemic inflammatory biomarkers and co-morbidities of chronic obstructive pulmonary disease. Ann Med 2013; 45(3):291–300.
31. Dickens JA, Miller BE, Edwards LD, Silverman EK, Lomas DA, Tal-Singer R. COPD association and repeatability of blood biomarkers in the ECLIPSE cohort. Respir Res 2011; 12:146.
32. Eickhoff P, Valipour A, Kiss D, Schreder M, Cekici L, Geyer K, et al. Determinants of systemic vascular function in patients with stable chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2008; 178(12):1211–1218.
33. Waschki B, Spruit MA, Watz H, Albert PS, Shrikrishna D, Groenen M, et al. Physical activity monitoring in COPD: compliance and associations with clinical characteristics in a multicenter study. Respir Med 2012; 106(4):522–530.
34. Moore AJ, Soler RS, Cetti EJ, Amanda Sathyapala S, Hopkinson NS, Roughton M, et al. Sniff nasal inspiratory pressure versus IC/TLC ratio as predictors of mortality in COPD. Respir Med. 2010; 104(9):1319–1325.
35. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994; 49(2):M85–94.
36. Kon SS, Patel MS, Canavan JL, Clark AL, Jones SE, Nolan CM, et al. Reliability and validity of the four metre gait speed in COPD. ERJ 2013; 42(2):333–340.

Appendix D: data dictionary

ERICA DATABASE SPECIFICATION DOCUMENT
BASELINE AND FOLLOW UP QUESTIONNAIRES

TABLE: Baseline data
GENERIC

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Trial	trial	Pre-filled by application			
Site	site	Text: n01=Cambridge, n02=Edinburg, n03=Cardiff, n04=Nottingham, n05=London	Drop down: Cambridge, Edinburgh, Cardiff, Nottingham, London	n01=Cambridge, n02=Edinburg, n03=Cardiff, n04=Nottingham, n05=London	Only one option possible
Label	label	Pre-filled by application	Person ID, date of birth, status		
Subject ID	personid	Assigned by database			
Visit cycle	visitcycle	Assigned by database			
Form cycle	formcycle	Assigned by database			
Repeat number	repeatnumber	Assigned by database			
Site ID	siteid	Text: 1=N01, 2=N02, 3=N03, 4=N04, 5=N05	Drop down: Cambridge, Edinburgh, Cardiff, Nottingham, London	1=N01, 2=N02, 3=N03, 4=N04, 5=N05	Only one option possible
Visit date	visitdate	Date	dd/mm/yyyy		
Date of birth	edob	Date	dd/mm/yyyy		
Sub ID	subid	Assigned by database			
Macro ID	macroid	Assigned by database			
Gender	eligsex	Number: 1=Female, 2=Male	Drop down: Female, Male	1=Female, 2=Male	Only one option possible
Informed consent	consent1	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Informed consent	consall	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of informed consent	consentdate	Date	dd/mm/yyyy		Should be ≤ *visitdate*
Age	ageeligibility	Number	nn yrs		Must be ≥40
Ethnicity	ethinicity	Number: 1=Black/ African-Caribbean/ Sub-Saharan, 2=White, 3=Asian, 4=Other	Drop down: Black/ African-Caribbean/ Sub-Saharan, White, Asian, Other	1=Black/ African-Caribbean/ Sub-Saharan, 2=White, 3=Asian, 4=Other	Only one option possible
Sort ID	sortid	Number	nnn		Only one option possible

TABLE: Baseline data
ANTHROPOMETRY

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Height	height	Number	nnn cm		Must be in range 130-200
Fat mass	fatmass_anthro	Number	nn.n kg		Must be in range 0-100
Weight (Tanita)	weight	Number	nn.n kg		Must be in range 0-200
Fat free mass:	fatfreemas_antr	Number	nn.n kg		Must be in range 0-130 (~70% of *weight*)

Body fat (Tanita)	bodyfat	Number	nn.n %		Must be in range 0-100
Total body water	totalbodywater	Number	nn.n kg		Must be in range 0-130 (~70% of *weight*)
Impedance, whole	impedence	Number	nnn Ω		Must be in range 0-1500
Date of anthropometry	anthrdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*

TABLE: Baseline data
SPIROMETRY

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Spirometry undertaken at this visit?	spirothisvisit	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of spirometry	spirodate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Previous spirometry undertaken?	prevspirometry	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Previous FEV_1	prevfev1	Number	nn L		Must be <3.5
Previous FEV_1/FVC ratio	prevfevfvcratio	Number	n.n %		Must be <0.7
Date of previous spirometry	prevdatespir	Date	dd/mm/yyyy		Should be < *spirodate*
FVC	fvc	Number	n.nn L		Must be <4.5
FVC percentage	fvcpercent	Number	nn %predicted		Must be in range 5-160
FEV_1	fev1	Number	n.nn L		Must be in range 0-4
FEV_1 percentage	fev1percent	Number	nn %predicted		Must be ≤80%
FEV_1/FVC ratio	fev_fvcratio	Number	n.nn %		Must be <0.7
ECG undertaken?	ecgundertaken	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of ECG	ecgdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
ECG signed by a clinician?	ecgsignci	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

TABLE: Baseline data
SHORT PHYSICAL PERFORMANCE BATTERY (SPPB)

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
SPPB undertaken?	sppbdone	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of SPPB	sppbdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Balance side-by-side, time	sidebysidetime	Number	n.n sec.		Must be ≤10
Correction - Balance side-by-side, time	re_sidebysidetime	Number	n.n sec.		Must be ≤10
Balance side-by-side, points	balancesidepts	Number	n		Must be in range 0-1
Correction - Balance side-by-side, points	re_balancesidepts	Number	n		Must be in range 0-1
Balance side-by-side, not done	spsidbyside_nd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Correction - Balance side-by-side, not done	re_spsidbyside_nd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Balance semi-tandem, time	balasemitantime	Number	n.n sec.		Must be ≤10. Variable *sidebysidetime* must be ≥10

Correction - Balance semi-tandem, time	re_balasemitantime	Number	n.n sec.		Must be ≤10. Variable *sidebysidetime* must be ≥10
Balance semi-tandem, points	balasemitandpts	Number	n		Must be in range 0-1
Correction - Balance semi-tandem, points	re_balasemitandpts	Number	n		Must be in range 0-1
Balance semi-tandem, not done	spsemitand_nd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Correction - Balance semi-tandem, not done	re_spsemitand_nd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Balance tandem, time	baltandemtime	Number	n.n sec.		Must be ≤10. Variables *sidebysidetime* and *balasemitantime* must be ≥10
Correction - Balance tandem, time	re_baltandemtime	Number	n.n sec.		Must be ≤10. Variables *sidebysidetime* and *balasemitantime* must be ≥10
Balance tandem, points	baltandempts	Number	n		Must be in range 0-2. One additional point if 3-9.99 seconds. Zero points if <3 seconds.
Correction - Balance tandem, points	re_baltandempts	Number	n		Must be in range 0-2. One additional point if 3-9.99 seconds. Zero points if <3 seconds.
Balance tandem, not done	spbaltand_nd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Correction - Balance tandem, not done	re_spbaltand_nd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Sum points of balance	sumbalancepts	Number	n		Must be in range 0-4
Correction - Sum points of balance	re_sumbalancepts	Number	n		Must be in range 0-4
Effort 1 of 4m gait speed test	gaitspedef1time	Number	n.n sec.		Must be in range 0-60
Effort 2 of 4m gait speed test	gaitspedef2time	Number	n.n sec.		Must be in range 0-60
Best time of effort 1 and 2 of 4m gait speed test	re_gaitsped_besttime	Number	n.n sec.		Must be in range 0-60
Converted points 4m gait speed test	bestgaitspeed	Number	n		Must be in range 0-4. One point if > 8.70 sec.; two points if 6.21-8.70 sec.; three points if 4.82-6.20 sec.; four points if <4.82 sec.
Correction - Converted points 4m gait speed test	re_bestgaitspeedpts	Number	n		Must be in range 0-4. One point if > 8.70 sec.; two points if 6.21-8.70 sec.; three points if 4.82-6.20 sec.; four points if <4.82 sec.
Gait speed test, not done	sp_gaitspeednd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Correction - Gait speed test, not done	re_sp_gaitspeednd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Chair stand, time	chairstandtime	Number	n.n sec.		Must be in range 0-120

Correction - Chair stand, time	re_chairstandtime	Number	n.n sec.		Must be in range 0-120
Chair stand, points	chairstandpts	Number	n		Must be in range 0-4. One point if 16.70-60 sec.; two points if 13.70-16.69 sec.; three points if 11.20-13.69 sec.; four points if <11.20 sec.
Correction - Chair stand, points	re_chairstandpts	Number	n		Must be in range 0-4. One point if 16.70-60 sec.; two points if 13.70-16.69 sec.; three points if 11.20-13.69 sec.; four points if <11.20 sec.
Chair stand test, not done	sp_chairstandnd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Correction - Chair stand test, not done	re_sp_chairstandnd	Number: 1=?, 2=?	Drop down: ?, ?	1=?, 0=?	Only one option possible
Total sum of SPPB	totalsumsppb	Number	n		Must be in range 0-12. Summation of variables *sumbalancepts*, *bestgaitspeed* and *chairstandpts*
Correction - Total sum of SPPB	re_totalsumsppb	Number	n		Must be in range 0-12. Summation of variables *sumbalancepts*, *bestgaitspeed* and *chairstandpts*
All 3 SPPB components completed?	re_all_3_sppb_comps	Text: Yes, No	Drop down: Yes, No	Yes, No	Only one option possible

TABLE: Baseline data
6 MINUTE WALK TEST (6MWT)

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
6MWT undertaken?	walktestdone	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of 6MWT	walktestdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Distance walked	distancewalked	Number	nnn metres		Must be in range 0-999
Pre-walk O_2 saturations	prewalk02satur	Number	nn %		Must be in range 50-100
Post-walk O_2 saturations	postwalk02satur	Number	nn %		Must be in range 50-100
Pre-walk borg rating	prewalkborgrate	Number	nn		Must be in range 0-10
Post-walk borg rating	postwalkborgrat	Number	nn		Must be in range 0-10
Did the patient require O_2 supplementation?	o2supplerecd	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Did the patient require O_2 supplementation?	o2requiredlts	Number	n L		Must be in range 0-10

TABLE: Baseline data
VENEPUNCTURE/URINE SAMPLE

Home

CRF field name	short field name	Field type	Description	Format	Validation rules

(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Blood sample taken?	bloodsmpletaken	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of blood sample taken	bloodsmpledate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Sodium	sodium	Number	nnn mEq/L		Must be in range 115-146
Potassium	potassium	Number	n.n mEq/L		Must be in range 2.9-6.0
Creatinine	creatinine	Number	nn mg/dL		Must be in range 50-700
Glucose	glucose	Number	n.n mmol/L		Must be in range 3.0-9.0
Glycated haemoglobin (HBA1c)	hba1c	Number	nn mmol/mol		Must be in range 30-100
Correction - Glycated haemoglobin (HBA1c)	re_hba1c	Number	nn mmol/mol		Must be in range 30-100
Glomerular filtration rate	gfr	Number	nn.nn		Must be in range 0-130
Fibrinogen	fibrinogen	Number	n.n g/dL		Must be in range 1.0-7.0
High-sensitivity c-reactive protein (HSCRP)	hscrp	Number	n.nn mg/L		Must be in range 0-200
White blood cell count	wbc	Number	n.n mcL		Must be in range 3-20x10^9
Haemoglobin	haemoglobin	Number	nnn g/L		Must be in range 6-20
Correction - Haemoglobin	re_haemoglobin	Number	nnn g/L		Must be in range 6-20
Platelets	platelets	Number	nnn mcL		Must be in range 0-800x10^9
Neutrophils	neutrophils	Number	n.n mm3		Must be in range 1-15x10^9
Total cholesterol	totalcholestrol	Number	n.n mmol/L		Must be <10.0
LDL cholesterol	ldlcholestrol	Number	n.n mmol/L		Must be <5.0
HDL cholesterol	hdlcholestrol	Number	n.n mmol/L		Must be <5.0
Triglycerides	triglycerides	Number	n.n mg/dL		Must be <5.0
Urine sample collected?	urinesmpletaken	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of urine sample collected	urinesmpledate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Serum stored	serumstoredone	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

TABLE: Baseline data
BLOOD PRESSURE AND ARTERIAL STIFFNESS

Home

CRF field name	**short field name**	**Field type**	**Description**	**Format**	**Validation rules**
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Date of blood pressure and arterial stiffness	bpartstifdat	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Seated brachial blood pressure (systolic)	seatedsysbp	Number	nnn mmHg		Must be in range 50-250
Seated brachial blood pressure (diastolic)	seateddiabp1	Number	nn mmHg		Must be in range 40-150
Seated central blood pressure (systolic)	seatcentsysbp	Number	nnn mmHg		Must be in range 50-250
Seated central blood pressure (diastolic)	seatcentbpdia	Number	nn mmHg		Must be in range 40-150

Seated heart rate (SphygmoCor)	seathr	Number	nn bpm		Must be in range 40-200
Seated mean arterial pressure	seatedmap	Number	nn mmHg		Must be in range 40-200
Seated augmentation index	seatedaugindex	Number	nn %		Must be in range -10-60
Supine pulse wave velocity – notch-carotid (proximal)	notchcarotidp	Number	nn mm		Must be in range 30-200
Supine pulse wave velocity – notch-femoral (distal)	notchfemorald	Number	nnn mm		Must be in range 300-1000
Supine blood pressure (systolic)	supinesysbp	Number	nn mmHg		Must be in range 50-250
Supine blood pressure (diastolic)	supinediabp1	Number	nn mmHg		Must be in range 40-150
Pulse wave velocity	pwv1	Number	n.n m/sec		Must be in range 4-20
Supine heart rate (SphygmoCor)	suphr	Number	nn bpm		Must be in range 40-200

TABLE: Baseline data
CAROTID INTIMA-MEDIA THICKNESS (CIMT) Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
CIMT done	imtdone	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of CIMT	imtdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Right CIMT diameter	imt_diam_r	Number	n.nn mm		Must be in range 3.0-15
Left CIMT diameter	imt_diam_l	Number	n.nn mm		Must be in range 3.0-15
Right CIMT test	imt_cimt_r	Number	n.nn mm		Must be in range 0.0-2.0
Left CIMT test	imt_cimt_l	Number	n.nn mm		Must be in range 0.0-2.0

TABLE: Baseline data
SNIFF NASAL INSPIRATORY PRESSURE (SNIP) Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
SNIP undertaken?	snipdone	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of SNIP	snipdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
SNIP right nostril (highest value)	rightnostsnip	Number	nn cmH2O		Must be in range 0-200

TABLE: Baseline data
QUADRICEPS MAXIMAL VOLUNTARY CONTRACTION (QMVC) Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
QMVC undertaken?	qmvcdone	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Date of QMVC	qmvcdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Highest/best effort QMVC	besteffort_qmvc	Number	nn.n kg		Must be in range 0-100

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
GENERIC Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
ERICA questionnaire completion date	ericquescompdat	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Age derived	agederied	Number	nn yrs		Must be ≥40
Birth weight	birthweight	Number	n.n kg		Must be in range 0.2-7
Are you employed?	employ	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Have you had any jobs involving work with any of these?	job	Number: 0=No, 1=Coal, 2=Asbestos, 3= Chemicals, 4=Dust	Drop down: No, Coal, Asbestos, Chemicals, Dust	0=No, 1=Coal, 2=Asbestos, 3= Chemicals, 4=Dust	Multiple options possible
Have you had any jobs involving work with any of these?	jobyn	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Multiple options possible
If yes, please specify approximately how long for	joblong	Number	nn		Must be in range 0-999
If yes, please specify approximately how long for	jobunit	Number: 1=Days, 2= Weeks, 3=Months, 4=Years	Drop down: Days, Weeks, Months, Years	1=Days, 2= Weeks, 3=Months, 4=Years	Multiple options possible
Are you married or cohabiting?	marr	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If not married/ co-habiting, are you:	ntmarr	Number: 1=Single / never married, 2=Widowed, 3=Divorced, 4=Separated	Drop down: Single / never married, Widowed, Divorced, Separated	1=Single / never married, 2=Widowed, 3=Divorced, 4=Separated	Only one option possible
If widowed/ divorced or separated, what year?	marryr	Date	yyyy		Should be ≤ *visitdate*

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
LUNG HEALTH

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Do you get short of breath?	srtbreath	Number: 0=No, 1=All the time, 2=Worse at certain times of the day, 3=With exercise only, 4=Night only	Drop down: No, All the time, Worse at certain times of the day, With exercise only, Night only	0=No, 1=All the time, 2=Worse at certain times of the day, 3=With exercise only, 4=Night only	Only one option possible
Do you cough?	eqcough	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If you cough, do you produce phlegm (sputum)?	eqphlegm	Number: 0=Never, 1=Yes, most mornings, 2=At least three months per year, 3=Only with exacerbations, 4=Occasionally	Drop down: Never, Yes, most mornings, At least three months per year, Only with exacerbations, Occasionally	0=Never, 1=Yes, most mornings, 2=At least three months per year, 3=Only with exacerbations, 4=Occasionally	Only one option possible
When were you diagnosed with COPD?	copdage	Number	nn yrs		Must be in range 30-100
When were you diagnosed with COPD?	copddiayear	Date	yyyy		Should be ≤ *visitdate*
Can you climb a flight of stairs without stopping?	eqstairs	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

Which statement best describes your breathlessness?	descbreath	Number: 1=I only get breathless with strenuous exercise, 2=I get short of breath when hurrying on the level or walking up a slight hill, 3=I walk slower than people on the level/stop for breath when walking at own pace, 4=I stop for breath after walking about 100 yards or after a few minutes on the level, 5=I am too breathless to leave the house or I am breathless when dressing	Drop down: I only get breathless with strenuous exercise, I get short of breath when hurrying on the level or walking up a slight hill, I walk slower than people on the level/stop for breath when walking at own pace, I stop for breath after walking about 100 yards or after a few minutes on the level, I am too breathless to leave the house or I am breathless when dressing	1=I only get breathless with strenuous exercise, 2=I get short of breath when hurrying on the level or walking up a slight hill, 3=I walk slower than people on the level/stop for breath when walking at own pace, 4=I stop for breath after walking about 100 yards or after a few minutes on the level, 5=I am too breathless to leave the house or I am breathless when dressing	Only one option possible
Have you ever smoked cigarettes?	smoked	Number: 0=No, 1= Yes - I currently smoke, 2= Yes - but I have given up	Drop down: No, Yes - I currently smoke, Yes - but I have given up	0=No, 1= Yes - I currently smoke, 2= Yes - but I have given up	Only one option possible
If you smoke or have smoked, how many cigarettes did you smoke each day?	smokenum	Number	n per day		Must be in range 1-999
If you smoke or have smoked, how many years was this for?	smokeyrs	Number	n yrs		Must be in range 1-100
What age did you start smoking?	smokeage	Number	n yrs old		Must be in range 7-100
If you have given up, how many years ago?	smokegiven	Number	n yrs ago		Must be in range 0-100
Have you been a cigar smoker?	cigar	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, how many	cigarnum	Number	n per day		Must be in range 1-999
How many years have you smoked cigars?	cigaryrs	Number	n yrs		Must be in range 0-100
Have you ever regularly smoked social drugs, for e.g. cannabis?	drug	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Total pack years	totalpackyears	Number	nn yrs		Must be ≥10
Have you ever required steroids?	steroids	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Have you ever required antibiotics for your chest?	antibi	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, how many courses of steroids have you required in the last 12 months?	stcourse	Number	n coure(s)		Must be in range 0-999
If yes, how many courses of antibiotics have you required in the last 12 months?	abcourse	Number	n course(s)		Must be in range 0-999
When was your last course of steroids/ antibiotics?	lascousestds	Number	n wks ago		Must be in range 0-999

Do you have oxygen at home?	oxygen	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, how many hours a day, have you been advised to use it?	oxyhrs	Number	n hrs/day		Must be in range 0-24
Do you snore?	snore	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Do you have sleep apnoea (OSA)?	apnoea	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
How likely are you to doze off or fall asleep in the following situations in contrast to just feeling tired?	doze	Number: 1=Sitting and reading, 2=Watching TV, 3=Sitting inactive in a public place, 4=As passenger in a car for an hour without break, 5=Lying down to rest during the day when circumstances permit, 6=Sitting and talking to someone, 7=Sitting quietly after lunch without alcohol, 8=In a car, while stopped for a few minutes in traffic	Drop down: Sitting and reading, Watching TV, Sitting inactive in a public place, As passenger in a car for an hour without break, Lying down to rest during the day when circumstances permit, Sitting and talking to someone, Sitting quietly after lunch without alcohol, In a car, while stopped for a few minutes in traffic	1=Sitting and reading, 2=Watching TV, 3=Sitting inactive in a public place, 4=As passenger in a car for an hour without break, 5=Lying down to rest during the day when circumstances permit, 6=Sitting and talking to someone, 7=Sitting quietly after lunch without alcohol, 8=In a car, while stopped for a few minutes in traffic	
Chance of dozing	dozescale	Number: 0=Would never doze, 1=Slight chance of dozing, 2=Moderate chance of dozing, 3=High chance of dozing	Drop down: Would never doze, Slight chance of dozing, Moderate chance of dozing, High chance of dozing	0=Would never doze, 1=Slight chance of dozing, 2=Moderate chance of dozing, 3=High chance of dozing	Only one option possible
Have you ever done pulmonary rehabilitation?	pulrehab	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, did you complete the whole course?	pulrehabcomp	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
How long ago was the course?	pulrehabyrs	Number	n yrs		Must be in range 0-100
What limits your walking?	pulrehablim	Number: 0=Breathlessness, 1=Legs, 2=Nothing/other	Drop down: Breathlessness, Legs, Nothing/other	0=Breathlessness, 1=Legs, 2=Nothing/other	Only one option possible
Have you lost or gained any weight in the last 12 months?	eqweight	Number: 0=Stayed about same, 1=Lost weight, 2=Gained weight	Drop down: Stayed about same, Lost weight, Gained weight	0=Stayed about same, 1=Lost weight, 2=Gained weight	Only one option possible

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
ACTIVITIES Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
How often do you take part in sport or activities that are mildly energetic, moderately energetic or vigorous?	sport	Number: 1=Mildly energetic, 2=Moderately energetic, 3=Vigorous	Drop down: Mildly energetic, Moderately energetic, Vigorous	1=Mildly energetic, 2=Moderately energetic, 3=Vigorous	Multiple options possible

How often do you take part in sport or activities that are mildly energetic, moderately energetic or vigorous?	sportscale	Number: 0=Never/ hardly ever, 1=Three times a week, 2=Once or twice a week, 3=About once or three times a month	Drop down: Never/ hardly ever, Three times a week, Once or twice a week, About once or three times a month	0=Never/ hardly ever, 1=Three times a week, 2=Once or twice a week, 3=About once or three times a month	Multiple options possible
Please give the average number of hours per week that you spend in such activities	sportav	Number: 1=Mildly energetic, 2=Moderately energetic, 3=Vigorous	Drop down: Mildly energetic, Moderately energetic, Vigorous	1=Mildly energetic, 2=Moderately energetic, 3=Vigorous	Multiple options possible
Please give the average number of hours per week that you spend in such activities	sporthrs	Number	n hrs/week		Must be in range 0-140
In the past week, on average, for how long did you walk outside your home/workplace on each weekday?	walkweekdayeq1	Text	nn.nn.nn hrs/min		Must be in range 0-24
In the past week, on average, for how long did you walk outside your home/workplace on each weekend day?	walkweekendeq1	Text	nn.nn.nn hrs/min		Must be in range 0-24
Walking, not recorded	walknoeq1	Number: 1=Yes	Drop down: Yes	1=Yes	Only one option possible
In the past, on average, for how long did you cycle on each weekday?	cycleweekdayeq1	Text	nn.nn.nn hrs/min		Must be in range 0-24
In the past, on average, for how long did you cycle on each weekend day?	cycleweekendeq1	Text	nn.nn.nn hrs/min		Must be in range 0-24
Cycling, not recorded	cyclenoeq1	Number: 1=Yes	Drop down: Yes	1=Yes	Only one option possible
How would you describe your usual walking pace?	walkpace	Number: 1=Slow pace (<3mph), 2=Steady average pace, 3=Brisk pace, 4=Fast pace (>4mph)	Drop down: Slow pace (<3mph), Steady average pace, Brisk pace, Fast pace (>4mph)	1=Slow pace (<3mph), 2=Steady average pace, 3=Brisk pace, 4=Fast pace (>4mph)	Only one option possible

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
PAST MEDICAL HISTORY

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Have you ever been told by your doctor that you have high blood pressure?	highbp	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If high blood pressure, in what year?	highbpyr	Date	yyyy		Should be ≤ *visitdate*
If high blood pressure, are you on therapy for it?	bpther	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

If high blood pressure therapy, in what year?	bptheryr	Date	yyyy		Should be ≤ *visitdate*
Please specify the drug name.	bpdrug	Max characters 50	Free text		
Have you ever been told by your doctor that you have high cholesterol?	highchol	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If high cholesterol, in what year?	highcholyr	Date	yyyy		Should be ≤ *visitdate*
If high cholesterol, are you on therapy for it?	cholther	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If high cholesterol therapy, in what year?	choltheryr	Date	yyyy		Should be ≤ *visitdate*
Please specify the drug name.	choldrug	Max characters 50	Free text		
Have you ever been told by your doctor that you have peripheral vascular disease?	pvd	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If peripheral vascular disease, in what year?	pvdyr	Date	yyyy		Should be ≤ *visitdate*
If peripheral vascular disease, are you on therapy for it?	pvdther	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If peripheral vascular disease therapy, in what year?	pvdtheryr	Date	yyyy		Should be ≤ *visitdate*
Please specify the drug name.	pvddrug	Max characters 50	Free text		
Have you ever been told by your doctor that you have atrial fibrillation?	afib	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If atrial fibrillation, in what year?	afibyr	Date	yyyy		Should be ≤ *visitdate*
If atrial fibrillation, are you on therapy for it?	afibther	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If atrial fibrillation therapy, in what year?	afibtheryr	Date	yyyy		Should be ≤ *visitdate*
Please specify the drug name.	afibdrug	Max characters 50	Free text		
Have you ever been told by your doctor that you have diabetes?	diab	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If diabetes, in what year?	diabyr	Date	yyyy		Should be ≤ *visitdate*
If diabetes, which type?	diabtype	Number: 0=Don't know, 1=Type I, 2=Type II	Drop down: Don't know, Type I, Type II	0=Don't know, 1=Type I, 2=Type II	Only one option possible
Has a doctor told you that you have had angina?	angina_eq1	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If angina, in what year?	anginayear	Date	yyyy		Should be ≤ *visitdate*
Has a doctor told you that you have had a heart attack?	heartattack_eq1	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If heart attack, in what year?	heartattackyear	Date	yyyy		Should be ≤ *visitdate*

Has a doctor told you that you have had a stroke or transient attack?	stroke_eq1	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If stroke or TIA, in what year?	strokeyear	Date	yyyy		Should be ≤ *visitdate*
If stroke or TIA what type?	stroketype_eq1	Number: 1=Stroke, 2=TIA, 3=Other, specify	Drop down: Stroke, TIA, Other, specify	1=Stroke, 2=TIA, 3=Other, specify	Only one option possible
If other, please specify?	otherstroke	Max characters 50	Free text		
Any other heart trouble suspected or confirmed?	eq_hrtrouble	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If any other heart trouble, please specify	eq_hrtrobother	Max characters 50	Free text		

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
FAMILY HISTORY

Home

CRF field name	**short field name**	**Field type**	**Description**	**Format**	**Validation rules**
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Was your father ever diagnosed with high blood pressure?	fatherhighbp	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If father high blood pressure, younger than 60 when diagnosed?	fathyoung60bp	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your father ever diagnosed with angina?	fatheangina	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If father angina, younger than 60 when diagnosed?	fathyng60angna	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your father ever diagnosed with a heart attack?	fatherheartatk	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If father heart attack, younger than 60 when diagnosed?	dadyng60hrtatk	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your father ever diagnosed with a stroke?	fatherstroke	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If father stroke, younger than 60 when diagnosed?	fathyng60stroke	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your father ever diagnosed with peripheral vascular disease?	fatherpvd	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If father peripheral vascular disease, younger than 60 when diagnosed?	fathyng60pvd	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your father ever diagnosed with diabetes?	fatherdiabetes	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If father diabetes, younger than 60 when diagnosed?	fathyoung60diab	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

Was your father ever diagnosed with asthma?	fatherasthma	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If father asthma, younger than 60 when diagnosed?	fathyng60asthma	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your father ever diagnosed with COPD?	fathercopd	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If father COPD, younger than 60 when diagnosed?	fathyng60copd	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your mother ever diagnosed with high blood pressure?	motherhighbp	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If mother high blood pressure, younger than 60 when diagnosed?	mumyoung60bp	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your mother ever diagnosed with angina?	motherangina	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If mother angina, younger than 60 when diagnosed?	mumyng60angina	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your mother ever diagnosed with a heart attack?	mumheartattack	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If mother heart attack, younger than 60 when diagnosed?	mumyng60hrtatk	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your mother ever diagnosed with a stroke?	motherstroke	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If mother stroke, younger than 60 when diagnosed?	mumyng60stroke	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your mother ever diagnosed with peripheral vascular disease?	motherpvd	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If mother peripheral vascular disease, younger than 60 when diagnosed?	mumyng60pvd	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your mother ever diagnosed with diabetes?	motherdiabetes	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If mother diabetes, younger than 60 when diagnosed?	mumyng60diabe	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your mother ever diagnosed with asthma?	motherasthma	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If mother asthma, younger than 60 when diagnosed?	mumyng60asthma	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Was your mother ever diagnosed with COPD?	mothercopd	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If mother COPD, younger than 60 when diagnosed?	mumyng60copd	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

Do you have any sibling?	anysibling	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Were your brother(s)/sister(s) ever diagnosed with high blood pressure?	siblinghighbp	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If brother(s)/sister(s) high blood pressure, younger than 60 when diagnosed?	siblingyng60bp	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Were your brother(s)/sister(s) ever diagnosed with angina?	siblingangina	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If brother(s)/sister(s) angina, younger than 60 when diagnosed?	sibyng60angina	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Were your brother(s)/sister(s) ever diagnosed with a heart attack?	siblinghrtattk	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If brother(s)/sister(s) heart attack, younger than 60 when diagnosed?	sibyng60hrtatk	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Were your brother(s)/sister(s) ever diagnosed with a stroke?	siblingstroke	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If brother(s)/sister(s) stroke, younger than 60 when diagnosed?	sibyng60stroke	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Were your brother(s)/sister(s) ever diagnosed with peripheral vascular disease?	siblingpvd	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If brother(s)/sister(s) peripheral vascular disease, younger than 60 when diagnosed?	sibyng60pvd	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Were your brother(s)/sister(s) ever diagnosed with diabetes?	siblingdiabetes	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If brother(s)/sister(s) diabetes, younger than 60 when diagnosed?	sibyng60diabe	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Were your brother(s)/sister(s) ever diagnosed with asthma?	siblingasthma	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If brother(s)/sister(s) asthma, younger than 60 when diagnosed?	sibyng60asthma	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Were your brother(s)/sister(s) ever diagnosed with COPD?	siblingcopd	Number: 0=No, 1=Yes, 2=Don't know	Drop down: No, Yes, Don't know	0=No, 1=Yes, 2=Don't know	Only one option possible
If brother(s)/sister(s) COPD, younger than 60 when diagnosed?	sibyng60copd	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
MEDICATION Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Inhaler medications	inhoth	Max characters 50	Free text		
How many times per day?	inhaler	Number	n		Must be in range 0-20
Do you have a nebuliser?	nebyn	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If so, please let us know what drugs you regularly use in your nebuliser	neboth	Max characters 50	Free text		
How many times per day?	nebuliser	Number	n		Must be in range 0-20
Are there any other medications you have not mentioned to us in question 17 or above?	othdrugyn	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Name other medications	othermedname	Max characters 50	Free text		
How many times per day?	othermedtype	Number	n		Must be in range 0-20
Other medications	otherdrg1	Max characters 50	Free text		
Other medications	otherdrg2	Max characters 50	Free text		
Other medications	otherdrg3	Max characters 50	Free text		
Other medications	otherdrg4	Max characters 50	Free text		

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
OTHER INFORMATION Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
On average, how much alcohol do you consume in a week?	alcohol	Number	n units		Must be in range 0-999

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
COPD ASSESSMENT TEST (CAT) Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Date of CAT	catdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Q1 – cough	cough	Number: 0-5			Only one option possible
Q2 – phlegm	phlegm	Number: 0-5			Only one option possible
Q3 – chest	chest	Number: 0-5			Only one option possible
Q4 – stairs	stairs	Number: 0-5			Only one option possible
Q5 – activity	activity	Number: 0-5			Only one option possible
Q6 – confident	confident	Number: 0-5			Only one option possible
Q7 – sleep	sleep	Number: 0-5			Only one option possible

Q8 – energy	energy	Number: 0-5			Only one option possible
Total score of CAT	totscore	Number			Must be in range 0-40 Summation of variables cough, *phlegm, chest, stairs, activity, confident, sleep* and *energy*

TABLE: ERICA PARTICPIANT'S QUESTIONNAIRE
ST. GEORGE'S RESPIRATORY QUESTIONNAIRE FOR COPD PATIENTS (SGRQ-C) Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Date of SGRQ-C	sgdate	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Please select one box to show how you describe your current health	curhealth	Number: 0=Very good, 1=Good, 2=Fair, 3=Poor, 4=Very poor	Drop down: Very good, Good, Fair, Poor, Very poor	0=Very good, 1=Good, 2=Fair, 3=Poor, 4=Very poor	Only one option possible
Q1 – cough	sgcough	Number: 1=Most days a week, 2=Several days a week, 3=Only with chest infections, 4=Not at all	Drop down: Most days a week, Several days a week, Only with chest infections, Not at all	1=Most days a week, 2=Several days a week, 3=Only with chest infections, 4=Not at all	Only one option possible
Q1: cough – score	sgcoughscore	Number	0, 28.1, 46.3, 80.6		Only one option possible Derived from variable *sgcough*
Q2: phlegm	sgphlegm	Number: 1=Most days a week, 2=Several days a week, 3=Only with chest infections, 4=Not at all	Drop down: Most days a week, Several days a week, Only with chest infections, Not at all	1=Most days a week, 2=Several days a week, 3=Only with chest infections, 4=Not at all	Only one option possible
Q2: phlegm – score	sgphlegmscore	Number	0, 30.2, 47, 76.8		Only one option possible Derived from variable *sgphlegm*
Q3: shortness of breath	sgbreath	Number: 1=Most days a week, 2=Several days a week, 3=Not at all	Drop down: Most days a week, Several days a week, Not at all	1=Most days a week, 2=Several days a week, 3=Not at all	Only one option possible
Q3: shortness of breath – score	sgbreathscore	Number	0, 50.3, 87.2		Only one option possible Derived from variable *sgbreath*
Q4: attacks of wheezing	sgwheez	Number: 1=Most days a week, 2=Several days a week, 3=A few days a month, 4=Only with chest infections, 5=Not at all	Drop down: Most days a week, Several days a week, A few days a month, Only with chest infections, Not at all	1=Most days a week, 2=Several days a week, 3=A few days a month, 4=Only with chest infections, 5=Not at all	Only one option possible

Q4: attacks of wheezing – score	sgwheezscore	Number	0, 36.4, 45.6, 71, 86.2		Only one option possible Derived from variable *sgwheez*
Q5: How many attacks of chest trouble did you have during the last year?	sgattack	Number: 1=Three or more attacks, 2=One or two attacks, 3=None	Drop down: Three or more attacks, One or two attacks, None	1=Three or more attacks, 2=One or two attacks, 3=None	Only one option possible
Q5: How many attacks of chest trouble did you have during the last year? – score	sgattackscore	Number	0, 52.3, 80.1		Only one option possible Derived from variable *sgattack*
Q6: How often do you have good days (with little chest trouble)?	sggood	Number: 1=No good days, 2=A few good days, 3=Most days are good, 4=Every day is good	Drop down: No good days, A few good days, Most days are good, Every day is good	1=No good days, 2=A few good days, 3=Most days are good, 4=Every day is good	Only one option possible
Q6: How often do you have good days (with little chest trouble)? – score	sggoodscore	Number	0, 38.5, 76.7, 93.3		Only one option possible Derived from variable *sggood*
Q7: If you have a wheeze, is it worse in the morning?	sgmorning	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Q7: If you have a wheeze, is it worse in the morning? – score	sgmorningscore	Number	0, 62		Only one option possible Derived from variable *sgmorning*
Q8: How would you describe your chest condition?	sgchest	Number: 1=Causes me a lot of problems or is the most important problem I have, 2=Causes me a few problems, 3=Causes no problem	Drop down: Causes me a lot of problems or is the most important problem I have, Causes me a few problems, Causes no problem	1=Causes me a lot of problems or is the most important problem I have, 2=Causes me a few problems, 3=Causes no problem	Only one option possible
Q8: How would you describe your chest condition? – score	sgqchestconscor	Number	0, 34.6, 82.9		Only one option possible Derived from variable *sgchest*
Q9: Getting washed or dressed	sgwash	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q9: Getting washed or dressed – score	sgwashscore	Number	0, 82.8		Only one option possible Derived from variable *sgwash*
Q9: Walking around the home	sghome	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q9: Walking around the home – score	sghomescore	Number	0, 80.2		Only one option possible Derived from variable *sghome*
Q9: Walking outside on the level	sgwalklev	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q9: Walking outside on the level – score	sgwalklevscore	Number	0, 81.4		Only one option possible

					Derived from variable *sgwalklev*
Q9: Walking up a flight of stairs	sgwlkst	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q9: Walking up a flight of stairs – score	sgwlkstscore	Number	0, 76.1		Only one option possible Derived from variable *sgwlkst*
Q9: Walking up hills	sgwlkhill	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q9: Walking up hills – score	sgwlkhillscore	Number	0, 75.1		Only one option possible Derived from variable *sgwlkhill*
Q10: My cough hurts	sgchurts	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q10: My cough hurts – score	sgchurtsscore	Number	0, 81.1		Only one option possible Derived from variable *sgchurts*
Q10: My cough makes me tired	sgctired	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q10: My cough makes me tired – score	sgctiredscore	Number	0, 79.1		Only one option possible Derived from variable *sgctired*
Q10: I am breathless when I talk	sgbrtlk	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q10: I am breathless when I talk – score	sgbrtlkscore	Number	0, 84.5		Only one option possible Derived from variable *sgbrtlk*
Q10: I am breathless when I bend over	sgbrbend	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q10: I am breathless when I bend over – score	sgbrbendscore	Number	0, 76.8		Only one option possible Derived from variable *sgbrbend*
Q10: My cough or breathing disturbs my sleep	sgcsleep	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q10: My cough or breathing disturbs my sleep – score	sgcsleepscore	Number	0, 87.9		Only one option possible Derived from variable *sgcsleep*
Q10: I get exhausted easily	sgexhaus	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q10: I get exhausted easily – score	sgexhausscore	Number	0, 84		Only one option possible Derived from variable *sgexhaus*
Q11: My cough or breathing is embarrassing in public	sgcembarras	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q11: My cough or breathing is embarrassing in public – score	sgcembarrscore	Number	0, 74.1		Only one option possible Derived from variable *sgcembarras*
Q11: My chest trouble is a nuisance to my family, friends or neighbours	sgfamily	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible

Q11: My chest trouble is a nuisance to my family, friends or neighbours – score	sgfamilyscore	Number	0, 79.1		Only one option possible Derived from variable *sgfamily*
Q11: I get afraid or panic when I cannot get my breath	sgpanic	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q11: I get afraid or panic when I cannot get my breath – score	sgpanicscore	Number	0, 87.7		Only one option possible Derived from variable *sgpanic*
Q11: I feel that I am not in control of my chest problem	sgcontrol	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q11: I feel that I am not in control of my chest problem – score	sgcontrolscore	Number	0, 90.1		Only one option possible Derived from variable *sgcontrol*
Q11: I have become frail or an invalid because of my chest	sgfrail	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q11: I have become frail or an invalid because of my chest – score	sgfrailscore	Number	0, 89.9		Only one option possible Derived from variable *sgfrail*
Q11: Exercise is not safe for me	sgexsafe	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q11: Exercise is not safe for me – score	sgexsafescore	Number	0, 75.7		Only one option possible Derived from variable *sgexsafe*
Q11: Everything seems too much of an effort	sgeffort	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q11: Everything seems too much of an effort – score	sgeffortscore	Number	0, 84.5		Only one option possible Derived from variable *sgeffort*
Q12: I take a long time to get washed or dressed	sgwashtime	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q12: I take a long time to get washed or dressed – score	sgwashtimescore	Number	0, 74.2		Only one option possible Derived from variable *sgwashtime*
Q12: I cannot take a bath or shower, or I take a long time	sgbath	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q12: I cannot take a bath or shower, or I take a long time – score	sgbathscore	Number	0, 81		Only one option possible Derived from variable *sgbath*
Q12: I walk slower than other people, or I stop for rests	sgwlkslow	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q12: I walk slower than other people, or I stop for rests – score	sgwlkslowscore	Number	0, 71.7		Only one option possible Derived from variable *sgwlkslow*

Q12: Jobs such as housework take a long time, or I have to stop for rests	sgjobs	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q12: Jobs such as housework take a long time, or I have to stop for rests – score	sgjobsscore	Number	0, 70.6		Only one option possible Derived from variable *sgjobs*
Q12: If I walk up one flight of stairs, I have to go slowly or stop	sgstslow	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q12: If I walk up one flight of stairs, I have to go slowly or stop – score	sgstslowscore	Number	0, 71.6		Only one option possible Derived from variable *sgstslow*
Q12: If I hurry or walk fast, I have to stop or slow down	sghurry	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q12: If I hurry or walk fast, I have to stop or slow down – score	sghurryscore	Number	0, 72.3		Only one option possible Derived from variable *sghurry*
Q12: My breathing makes it difficult to do things such as walk up hills, etc.	sgdiffgolf	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q12: My breathing makes it difficult to do things such as walk up hills, etc. – score	sgdiffgolfscore	Number	0, 74.5		Only one option possible Derived from variable *sgdiffgolf*
Q12: My breathing makes it difficult to do things such as carry heavy loads, etc.	sgdiffswim	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q12: My breathing makes it difficult to do things such as carry heavy loads, etc. – score	sgdiffswimscore	Number	0, 71.4		Only one option possible Derived from variable *sgdiffswim*
Q13: I cannot play sports or games	sgsports	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q13: I cannot play sports or games – score	sgsportsscore	Number	0, 64.8		Only one option possible Derived from variable *sgsports*
Q13: I cannot go out for entertainment or recreation	sgenter	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q13: I cannot go out for entertainment or recreation – score	sgenterscore	Number	0, 79.8		Only one option possible Derived from variable *sgenter*
Q13: I cannot go out of the house to do the shopping	sgshop	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q13: I cannot go out of the house to do the shopping – score	sgshopscore	Number	0, 81		Only one option possible Derived from variable *sgshop*
Q13: I cannot do housework	sghousework	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible

Q13: I cannot do housework – score	sghseworkscore	Number	0, 79.1		Only one option possible Derived from variable *sghousework*
Q13: I cannot move far from my bed or chair	sgbed	Number: 1=True, 0=False	Drop down: True, False	1=True, 0=False	Only one option possible
Q13: I cannot move far from my bed or chair – score	sgbedscore	Number	0, 94		Only one option possible Derived from variable *sgbed*
Q14: How does your chest trouble affect you?	sgaffect	Number: 0=It does not stop me doing anything I would like to do, 1=It stops me doing one or two things I would like to do, 2=It stops me doing most of the things I would like to do, 3=It stops me doing everything I would like to do	Drop down: It does not stop me doing anything I would like to do, It stops me doing one or two things I would like to do, It stops me doing most of the things I would like to do, It stops me doing everything I would like to do	0=It does not stop me doing anything I would like to do, 1=It stops me doing one or two things I would like to do, 2=It stops me doing most of the things I would like to do, 3=It stops me doing everything I would like to do	Only one option possible
Q14: How does your chest trouble affect you? – score	sgaffectscore	Number	0, 42, 84.2, 96.7		Only one option possible Derived from variable *sgaffect*

TABLE: Follow Up Questionnaire – 6, 12, 18, 24 and 32 months
CARDIAC HEALTH

Home

CRF field name	**short field name**	**Field type**	**Description**	**Format**	**Validation rules**
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Follow up date	fupvisitdate fupvisitdate1 fupvisitdate2 fupvisitdate3 fupvisitdate4	Date 999=missing	dd/mm/yyyy		Should be ≥ *visitdate*
Since we last saw you, has a doctor told you that you have had angina?	fupangina fupangina1 fupangina2 fupangina3 fupangina4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, did you go to hospital?	anghosp anghosp1 anghosp2 anghosp3 anghosp4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If you did go to hospital, did you stay overnight?	angover angover1 angover2 angover3 angover4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Angina – date of admission	angaddate angaddate1 angaddate2 angaddate3 angaddate4	Date	dd/mm/yyyy		Should be ≥ *visitdate*
Since we last saw you, has a doctor told you that you have had a heart attack?	fuphattack fuphattack1 fuphattack2 fuphattack3 fuphattack4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

If yes, when	hattackwhn hattackwhn1 hattackwhn2 hattackwhn3 hattackwhn4	Date	dd/mm/yyyy		Should be ≥ *visitdate*
If yes, did you go to hospital?	hattackhosp hattackhosp1 hattackhosp2 hattackhosp3 hattackhosp4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If you did go to hospital, did you stay overnight?	hattackover hattackover1 hattackover2 hattackover3 hattackover4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Since we last saw you, have you been told by a doctor that you have high blood pressure?	fuphyperten fuphyperten1 fuphyperten2 fuphyperten3 fuphyperten4	Number: 0=No, 1=Yes new diagnosis, 2=Was already diagnosed	Drop down: No, Yes new diagnosis, Was already diagnosed	0=No, 1=Yes, new diagnosis, 2=Was already diagnosed	Only one option possible
If you have been newly diagnosed, are you on treatment, if so which	hypertendiag hypertendiag1 hypertendiag2 hypertendiag3 hypertendiag4	Max characters 50	Free text		
Since we last saw you, have you been told by a doctor that you have high cholesterol?	fuphyperchol fuphyperchol1 fuphyperchol2 fuphyperchol3 fuphyperchol4	Number: 0=No, 1=Yes new diagnosis, 2=Was already diagnosed	Drop down: No, Yes new diagnosis, Was already diagnosed	0=No, 1=Yes, new diagnosis, 2=Was already diagnosed	Only one option possible
If you have been newly diagnosed, are you on treatment, if so which	hypercholdiag hypercholdiag1 hypercholdiag2 hypercholdiag3 hypercholdiag4	Max characters 50	Free text		
Since we last saw you, have you been told by a doctor that you have had a stroke or transient attack?	fupstroke fupstroke1 fupstroke2 fupstroke3 fupstroke4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, which type?	stroketype stroketype1 stroketype2 stroketype3 stroketype4	Number: 1=Stroke, 2=Transient Ischemic Attack, 999=Other	Drop down: Stroke, Transient Ischemic Attack, Other	1=Stroke, 2=Transient Ischemic Attack, 999=Other	Only one option possible
If other, specify	stroketypeoth stroketypeoth1 stroketypeoth2 stroketypeoth3 stroketypeoth4	Max characters 50	Free text		
If yes, when?	strokewhen strokewhen1 strokewhen2 strokewhen3 strokewhen4	Date	dd/mm/yyyy		Should be ≥ *visitdate*
If yes, did you go to hospital?	strokehosp strokehosp1 strokehosp2 strokehosp3 strokehosp4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If you did go to hospital, did you stay	strokeover strokeover1	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

overnight?	strokeover2 strokeover3 strokeover4				
Since we last saw you, have you been told by a doctor that you have diabetes?	fupdiab fupdiab1 fupdiab2 fupdiab3 fupdiab4	Number: 0=No, 1=Yes new diagnosis, 2=Was already diagnosed	Drop down: No, Yes new diagnosis, Was already diagnosed	0=No, 1=Yes, new diagnosis, 2=Was already diagnosed	Only one option possible
If you have been newly diagnosed, are you on treatment, if so which	diabtreat diabtreat1 diabtreat2 diabtreat3 diabtreat4	Max characters 50	Free text		
Since we saw you last, have you had any pain or discomfort in your chest?	chestpain chestpain1 chestpain2 chestpain3 chestpain4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, do you get pain or discomfort when you walk uphill or hurry?	walkhurry walkhurry1 walkhurry2 walkhurry3 walkhurry4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, do you get it when you walk at an ordinary pace on the level?	walkord walkord1 walkord2 walkord3 walkord4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, when you get any pain or discomfort in your chest, what do you do?	pain pain1 pain2 pain3 pain4	Number: 0=Stop, 1=Slow down, 2=Continue at the same pace	Drop down: Stop, Slow down, Continue at the same pace	0=Stop, 1=Slow down, 2=Continue at the same pace	Only one option possible
If yes, does it go away when you stand still?	standstill standstill1 standstill2 standstill3 standstill4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, how soon?	howsoon howsoon1 howsoon2 howsoon3 howsoon4	Number: 0=In ten minutes or less, 1=More than ten minutes	Drop down: In ten minutes or less, More than ten minutes	0=In ten minutes or less, 1=More than ten minutes	Only one option possible
Where do you get this pain or discomfort?	painsite painsite1 painsite2 painsite3 painsite4	Max characters 50	Free text		
Since we saw you last, have you had a severe pain across the front of your chest lasting half an hour or more?	sevpain sevpain1 sevpain2 sevpain3 sevpain4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, did you talk to a doctor about it?	talkdoc talkdoc1 talkdoc2 talkdoc3 talkdoc4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
What did he/she say about it?	doccom doccom1 doccom2 doccom3	Max characters 50	Free text		

	doccom4				
How many of these attacks have you had?	attacknum attacknum1 attacknum2 attacknum3 attacknum4	Number	n		Must be in range 0-999
Have you at any time since we last saw you been awoken at night by an attack of breathlessness?	nightbreath nightbreath1 nightbreath2 nightbreath3 nightbreath4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Have you ever had noticeable swelling of your ankles for at least one week?	ankswell ankswell1 ankswell2 ankswell3 ankswell4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Do you get pain in either leg on walking?	legpain legpain1 legpain2 legpain3 legpain4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible

TABLE: Follow Up Questionnaire – 6, 12, 18, 24 and 32 months
RESPIRATORY HEALTH AND COPD

Home

CRF field name	**short field name**	**Field type**	**Description**	**Format**	**Validation rules**
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Since we last saw you, have you been on a pulmonary rehabilitation course?	fuppulrehab fuppulrehab1 fuppulrehab2 fuppulrehab3 fuppulrehab4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, did you complete it?	rehabcomp rehabcomp1 rehabcomp2 rehabcomp3 rehabcomp4	Number: 0=No, 1=Yes, 2=Ongoing	Drop down: No, Yes, Ongoing	0=No, 1=Yes, 2=Ongoing	Only one option possible
Since we last saw you, have you changed your smoking habit?	smokehab smokehab1 smokehab2 smokehab3 smokehab4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Stopped	fupsmkhabstop fupsmkhabstop1 fupsmkhabstop2 fupsmkhabstop3 fupsmkhabstop4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
When?	fupskhabstpwhen fupskhabstpwhen1 fupskhabstpwhen2 fupskhabstpwhen3 fupskhabstpwhen4	Number	n months ago		Must be in range 0-250
Started	fupsmkhabstart fupsmkhabstart1 fupsmkhabstart2 fupsmkhabstart3 fupsmkhabstart4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
How many / day?	fupshabstarnu fupshabstarnu1 fupshabstarnu2 fupshabstarnu3 fupshabstarnu4	Number	n per day		Must be in range 0-999

Tried to stop but restarted	fupskhbrestart fupskhbrestart1 fupskhbrestart2 fupskhbrestart3 fupskhbrestart4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Changed amount	fupskhbchngeamt fupskhbchngeamt1 fupskhbchngeamt2 fupskhbchngeamt3 fupskhbchngeamt4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
How many / day?	fupshbchgeamt fupshbchgeamt1 fupshbchgeamt2 fupshbchgeamt3 fupshbchgeamt4	Number	n per day		Must be in range 0-999
Other – e.g. cigars	fupsmkhabothers fupsmkhabothers1 fupsmkhabothers2 fupsmkhabothers3 fupsmkhabothers4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Which statement best describes your breathlessness? – MRC dyspnea score	fupdescbreath fupdescbreath1 fupdescbreath2 fupdescbreath3 fupdescbreath4	Number: 1=I only get breathless with strenuous exercise, 2=I get short of breath when hurrying on the level or walking up a slight hill, 3=I walk slower than people of the same age on the level because of breathlessness or have to stop for breath when walking at my own pace on the level, 4=I stop for breath after walking about 100 yards or after a few minutes on the level, 5=I am too breathless to leave the house or I am breathless when dressing	Drop down: I only get breathless with strenuous exercise, I get short of breath when hurrying on the level or walking up a slight hill, I walk slower than people of the same age on the level because of breathlessness or have to stop for breath when walking at my own pace on the level, I stop for breath after walking about 100 yards or after a few minutes on the level, I am too breathless to leave the house or I am breathless when dressing	1=I only get breathless with strenuous exercise, 2=I get short of breath when hurrying on the level or walking up a slight hill, 3=I walk slower than people of the same age on the level because of breathlessness or have to stop for breath when walking at my own pace on the level, 4=I stop for breath after walking about 100 yards or after a few minutes on the level, 5=I am too breathless to leave the house or I am breathless when dressing	Only one option possible
Which statement best describes your breathlessness? – MRC dyspnea score	breathyn breathyn1 breathyn2 breathyn3 breathyn4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Have you had any courses of steroids since we last saw you for your breathing?	sterbreath sterbreath1 sterbreath2 sterbreath3 sterbreath4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, how many courses	sterbreyn sterbreyn1 sterbreyn2 sterbreyn3 sterbreyn4	Number	n courses		Must be in range 0-999
Have you had any courses of antibiotics since we last saw you for your breathing?	antibreath antibreath1 antibreath2 antibreath3 antibreath4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, how many	antibreyn	Number	n courses		Must be in range 0-999

courses	antibreyn1 antibreyn2 antibreyn3 antibreyn4				

TABLE: Follow Up Questionnaire – 6, 12, 18, 24 and 32 months
GENERAL HEALTH

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
How many times have you consulted your GP since we last saw you?	gp gp1 gp2 gp3 gp4	Number	n		Must be in range 0-999
Compared to 6 months ago, how would you rate your health in general now?	fuphealth fuphealth1 fuphealth2 fuphealth3 fuphealth4	Number: 1=Much better now than six months ago, 2=Somewhat better than six months ago, 3=About the same as six months ago, 4=Somewhat worse than six months ago, 5=Much worse than	Drop down: Much better now than six months ago, Somewhat better than six months ago, About the same as six months ago, Somewhat worse than six months ago, Much worse than six months	1=Much better now than six months ago, 2=Somewhat better than six months ago, 3=About the same as six months ago, 4=Somewhat worse than six months ago, 5=Much worse than six months ago	Only one option possible
Since we last saw you, have you noticed that your weight has changed?	weightch weightch1 weightch2 weightch3 weightch4	Number: 0=No change, 1=Lost weight intentionally, 2=Lost weight unintentionally, 3=Gained weight	Drop down: No change, Lost weight intentionally, Lost weight unintentionally, Gained weight	0= No change, 1=Lost weight intentionally, 2=Lost weight unintentionally, 3=Gained weight	Only one option possible
On average, how much alcohol do you consume in a week?	alchunit alchunit1 alchunit2 alchunit3 alchunit4	Number	n units		Must be in range 0-999 1 unit = 1 small glass of wine (125mL) = ½ pint of beer/ lager/ cider = 25 ml pub measure of spirit

TABLE: Follow Up Questionnaire – 6, 12, 18, 24 and 32 months
HOSPITAL ADMISSION

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
For your breathing or COPD, have you been admitted to hospital (including A&E or admitted) since we last saw you?	copdhosp copdhosp1 copdhosp2 copdhosp3 copdhosp4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
If yes, please specify the number of times in the last 6 months	copdnum copdnum1 copdnum2 copdnum3 copdnum4	Number	n		Must be in range 1-999
If yes, did you need to go to intensive care?	copdinten copdinten1 copdinten2 copdinten3 copdinten4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option possible
Have you been	hosprea	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Only one option

admitted to hospital (including A&E, day case or admitted) for any other reason since we last saw you?	hosprea1 hosprea2 hosprea3 hosprea4				possible
If yes, please specify the number of times	hospnum hospnum1 hospnum2 hospnum3 hospnum4	Number	n		Must be in range 1-999
Please specify the reason for hospitalization(s)	cause cause1 cause2 cause3 cause4	Max characters 50	Free text		
Please specify the dates (month)	causemonth causemonth1 causemonth2 causemonth3 causemonth4	Date	mm		Should be ≤ *visitdate*
Please specify the dates (year)	causeyear causeyear1 causeyear2 causeyear3 causeyear4	Date	yyyy		Should be ≤ *visitdate*

TABLE: Follow Up Questionnaire – 6, 12, 18, 24 and 32 months
MEDICATION

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
Please list below your medication	meds meds1 meds2 meds3 meds4	Max characters 50	Free text		
Is this a new prescription since we last saw you?	newpres newpres1 newpres2 newpres3 newpres4	Number: 1=Yes, 0=No	Drop down: Yes, No	1=Yes, 0=No	Multiple responses possible

TABLE: Follow Up Questionnaire – 6, 12, 18, 24 and 32 months
ACTIVITIES

Home

CRF field name	short field name	Field type	Description	Format	Validation rules
(something meaningful relating to what's on CRF)	*(name used in database)*	*(text, numerical, drop down with options, date etc)*	*(Y/N, drop down menu with options, free text etc)*	*(for stats purposes – used for variable labelling in STATA)*	*(Details of ALL validation)*
How often do you take part in sports or activities that are mildly energetic, moderately energetic or vigorous?	sportdr sportdr1 sportdr2 sportdr3 sportdr4	Number: 1=Mildly energetic, 2=Moderately energetic, 3=Vigorous	Drop down: Mildly energetic, Moderately energetic, Vigorous	1=Mildly energetic, 2=Moderately energetic, 3=Vigorous	Multiple options possible
How often do you take part in sports or activities that are mildly energetic, moderately energetic or vigorous?	sportscaledr sportscaledr1 sportscaledr2 sportscaledr3 sportscaledr4	Number: 0=Never/ hardly ever, 1=Three times a week, 2=Once or twice a week, 3=About once or three times a month	Drop down: Never/ hardly ever, Three times a week, Once or twice a week, About once or three times a month	0=Never/ hardly ever, 1=Three times a week, 2=Once or twice a week, 3=About once or three times a month	Multiple options possible
On average, for how	weekwalkhr	Number	n hrs		Must be in range 0-24

long did you walk outside your home/workplace on each weekday?	weekwalkhr1 weekwalkhr2 weekwalkhr3 weekwalkhr4				If not walked, enter '00'
On average, for how long did you walk outside your home/workplace on each weekday?	weekwalkmin weekwalkmin1 weekwalkmin2 weekwalkmin3 weekwalkmin4	Number	n min		Must be in range 0-60 If not walked, enter '00'
On average, for how long did you walk outside your home/workplace on each weekend day?	wnkdwalkhr wnkdwalkhr1 wnkdwalkhr2 wnkdwalkhr3 wnkdwalkhr4	Number	n hrs		Must be in range 0-24 If not walked, enter '00'
On average, for how long did you walk outside your home/workplace on each weekend day?	wnkdwalkmin wnkdwalkmin1 wnkdwalkmin2 wnkdwalkmin3 wnkdwalkmin4	Number	n min		Must be in range 0-60 If not walked, enter '00'

Appendix E: causes of death

Table 1: Cause of death (n = 149)

#	ID	Description 1	Description 2	Description 3	Description	Category
#	ID	Description 1	Description 2	Description 3	Description 4	Category
1	1	Carcinomatosis	Metastatic renal carcinoma	Bladder cancer; ischaemic heart disease	Type II diabetes; COPD; AF	Cancer
2	12	Infective exacerbation of COPD		Angina; atrial fibrillation		Pulmonary
3	20	Haemorrhage Per Urethra	Colovesical fistula	Diverticular disease	COPD; chronic nephritic syndrome	Other
4	39	End stage COPD	Unknown			Pulmonary
5	76	Infective exacerbation of COPD		Polymyalgia rheumatica		Pulmonary
6	206	Bronchopneumonia	Unknown	Bronchiectasis		Pulmonary
7	230	Pulmonary haemorrhage		Lung Cancer		Cancer
8	279	Metastatic cancer	Metastatic cancer - unknown origin			Cancer
9	416	Carcinomatosis	Lung cancer	COPD		Cancer
10	450	Caecal volvulus with large bowel ischaemia	Aortic stenosis; COPD; type two diabetes	Diabetes mellitus; hypertension		Other
11	487	Aspiration pneumonia	Cerebrovascular accident	Essential hypertension	COPD	Cardiac
12	536	Aspiration pneumonia	Transient ischaemic attack; right femoral fracture			Pulmonary
13	576	Pneumonia	COPD	Cigarette smoking		Pulmonary
14	584	COPD				Pulmonary
15	607	Mixed drug Intoxication (morphine and tramadol)	COPD	Suicide		Other
16	645	Bronchopneumonia	COPD			Pulmonary
17	661	Myocardial infarction		Abdominal aortic aneurysm; bladder cancer; COPD		Cardiac
18	685	Pneumonia	Metastatic gastric adenocarcinoma	Type 2 diabetes		Cancer
19	841	Infective exacerbation of COPD	Infective exacerbation of COPD	Severe COPD	Severe COPD	Pulmonary
20	924	Cerebral lymphoma		Cerebral lymphoma		Cancer
21	956	Pulmonary embolus	Exact cause of death unknown, post mortem results showed bullous emphysema and moderate coronary artery disease	COPD; pneumothorax		Other
22	964	Respiratory failure	Severe COPD	Severe pulmonary hypertension		Pulmonary
23	1093	Hospital acquired pneumonia				Pulmonary
24	1103	Type 2 respiratory failure	Fractured neck of femur-right	Congestive cardiac failure	Type 2 respiratory failure	Pulmonary
25	1147	Pneumonia	COPD	Severe COPD	Ischaemic heart disease; hypertension	Pulmonary
26	1155	End stage COPD		End stage COPD		Pulmonary
27	1171	Multi-organ failure	Infective exacerbation COPD	Influenza A		Pulmonary
28	1225	Bronchopneumonia	Severe COPD	Bronchopneumonia	Severe COPD	Pulmonary
29	1250	Bronchial carcinoma		COPD		Cancer
30	1284	Myocardial infarction	Ischeamic heart disease			Cardiac
31	1346	Aspiration pneumonia	Chronic lymphocytic leukemia	COPD		Cancer
32	1359	Bronchopneumonia	COPD	Bronchopneumonia	COPD	Pulmonary
33	1368	Multiple myeloma	COPD	Follicular non-hodgkins lymphoma	Multiple meloma	Cancer

#	ID	Description 1	Description 2	Description 3	Description 4	Category
34	1386	Presumed COPD				Pulmonary
35	1420	Possible ruptured abdominal aortic aneurysm				Other
36	1436	Right bronchopneumonia		COPD		Cancer
37	1476	Septic shock	Amyloidosis	E859	COPD	Other
38	1516	Probable exacerbation of COPD		Atrial fibrillation		Pulmonary
39	1543	Myocardial infarction	Ischaemic heart disease			Cardiac
40	1610	Complications of metastatic melanoma		Coplications of metastatis melanoma		Cancer
41	1714	COPD	Diabetes	Left ventricular heart failure	Bilateral bronchopneumonia	Pulmonary
42	1768	Ischaemic heart disease		COPD		Cardiac
43	1826	Pneumonia	COPD	Sacral insufficiency fracture		Pulmonary
44	1860	Pneumonia	COPD	Cancer of the jaw (operated)		Pulmonary
45	1890	Congestive cardiac failure	Congestive heart failure	Ischaemic heart disease		Cardiac
46	1925	Sepsis	Bronchopneumonia			Pulmonary
47	1983	Ischaemic heart disease				Cardiac
48	2007	Oesophageal carcinoma				Cancer
49	2120	Urosepsis				Other
50	2144	Metastatic lung carcinoma				Cancer
51	2172	Metastatic adenocarcinoma of left lung	Lung cancer			Cancer
52	2198	Bronchopneumonia		Sarcomatoid carcinoma of lung		Cancer
53	2206	Cardiac arrest (in hospital)				Cardiac
54	2239	Ventilatory failure	Ventilatory failure	COPD	Chronic obstructive airway disease	Pulmonary
55	2263	Metastatic adenocarcinoma of lung				Cancer
56	2312	Coronary artery atheroma				Cardiac
57	2336	Community acquired pneumonia	COPD			Pulmonary
58	2352	Infective exacerbation of COPD	COPD			Pulmonary
59	2371	Acute myeloid leukaemia				Cancer
60	2379	Bilateral pneumonia	Chest	COPD		Pulmonary
61	2468	Hospital acquired pneumonia		COPD		Pulmonary
62	2534	Ischaemic heart disease				Cardiac
63	2554	Rectal carcinoma				Cancer
64	2573	COPD		Previous pulmonary embolism; carcinoma colon		Pulmonary
65	2668	COPD				Pulmonary
66	2684	Multi organ failure		Sepsis		Other
67	2700	Bronchopneumonia	COPD			Pulmonary
68	2708	Sepsis		Bile leak		Other
69	2737	Infective exacerbation of COPD				Pulmonary
70	2825	Infective exacerbation of COPD				Pulmonary
71	2909	Chest infection	COPD; metastatic bladder cancer	Ischemic heart disease		Pulmonary

#	ID	Description 1	Description 2	Description 3	Description 4	Category
72	2997	Respiratory sepsis	COPD	Large ovarian cystic mass		Pulmonary
73	3025	Type two respiratory failure	Severe COPD	Chronic cardiac failure		Pulmonary
74	3033	Bronchopneumonia	Lung adenocarcinoma			Cancer
75	3042	Recurrent pneumonia	COPD			Pulmonary
76	3070	Metastatic pancreatic cancer				Cancer
77	3137	Cardiomegaly	Chest/heart	Ischaemic heart disease		Cardiac
78	3245	Metastatic carcinoma colon		COPD		Cancer
79	3314	Sepsis	Pneumonia	COPD; sigmoid lesion (operated)		Pulmonary
80	3402	Acute pulmonary oedema		Ischaemic and valvular heart disease		Cardiac
81	3663	Infective exacerbation of COPD				Pulmonary
82	3719	Community acquired pneumonia		COPD		Pulmonary
83	3727	Pneumonia	COPD	Chronic congestive cardiac failure		Pulmonary
84	3768	Haemopericardium	Aortic dissection			Cardiac
85	3901	COPD				Pulmonary
86	3929	Coronary artery thrombosis				Cardiac
87	4000	Bronchopneumonia				Pulmonary
88	4056	Carcinomatosis		Metastatic lung cancer		Cancer
89	4064	Pancreatic cancer with metastases				Cancer
90	4149	Metastatic oesophageal cancer				Cancer
91	4321	Type II respiratory failure	COPD			Pulmonary
92	4329	Type II respiratory failure		COPD		Pulmonary
93	4338	End stage COPD				Pulmonary
94	4450	Metastatic lung cancer		Chronic obstructive airway disease	COPD; hypertension	Cancer
95	4474	Aspiration pneumonia	COPD	Atrial fibrillation		Pulmonary
96	4502	Bronchogenic carcinoma				Cancer
97	4592	Infective exacerbation of COPD		Bronchiectasis		Pulmonary
98	4643	End stage COPD				Pulmonary
99	4661	End stage COPD				Pulmonary
100	4684	COPD				Pulmonary
101	4744	Bowel perforation		ischaemic bowel		Other
102	4757	Ischaemic heart disease				Cardiac
103	4790	COPD				Pulmonary
104	4829	Acute exacerbation COPD				Pulmonary
105	4846	Intracerebral haemorrhage		High grade brain tumour		Cancer
106	4876	End stage dementia				Other
107	4910	Cerebrovascular accident		Atrial fibrillation		Cardiac
108	4937	Respiratory failure	COPD			Pulmonary
109	4962	Pneumonia		COPD; pulmonary embolus		Pulmonary
110	4983	Cardiomegaly	COPD			Pulmonary
111	5078	Metastatic cancer of stomach				Cancer
112	5103	Hospital acquired pneumonia	COPD	Severe aortic stenosis (operated)		Pulmonary
113	5120	End stage severe COPD				Pulmonary
114	5128	Pneumonia		COPD		Pulmonary

#	ID	Description 1	Description 2	Description 3	Description 4	Category
115	5152	Lung cancer	Cancer			Cancer
116	5167	Bone and cerebral metastasis		Mixed large cell and small cell carcinoma of lung		Cancer
117	5192	Respiratory failure	Unknown	Severe COPD	Unknown	Pulmonary
118	5212	Hepatocellular carcinoma				Cancer
119	5288	Heart attack		Pancreatitis		Cardiac
120	5306	Pancreatic cancer				Cancer
121	5340	Community acquired pneumonia				Pulmonary
122	5402	Metastatic adenocarcinoma of the lung	Renal Failure			Cancer
123	5460	End stage COPD				Pulmonary
124	5557	Multiorgan failure		Chest sepsis on a background of severe COPD		Pulmonary
125	5643	Acute kidney failure		COPD end stage		Pulmonary
126	5703	Bronchopneumonia				Pulmonary
127	5735	Hospital acquired pneumonia; pulmonary embolism	Severe COPD	Hypertension		Pulmonary
128	5761	Aspiration pneumonia	Oesophageal adenocarcinoma	Lung cancer		Cancer
129	5876	Metastatic breast carcinoma				Cancer
130	5892	Chest sepsis	Severe COPD on long term oxygen therapy			Pulmonary
131	5905	Respiratory failure		COPD		Pulmonary
132	6002	Infective exacerbation of COPD		Myelo dysplasia		Pulmonary
133	6135	Pancreatic cancer				Cancer
134	6145	COPD				Pulmonary
135	6156	Acute exacerbation of COPD	Infective exacerbation of COPD			Pulmonary
136	6168	Pneumonia	Haemorrhagic pancreatitis	Gallstones	Diabetes type II	Other
137	6200	COPD	Exacerbation of COPD			Pulmonary
138	6224	Hospital acquired pneumonia		End stage COPD		Pulmonary
139	6254	Bronchopneumonia		Obstructive airway disease		Pulmonary
140	6382	Unascertained		Open conclusion		Other
141	6398	Chest Infection		COPD		Pulmonary
142	6463	Bronchopneumonia	Infective exacerbation of COPD	COPD		Pulmonary
143	6503	COPD		Cerebrovascular disease		Pulmonary
144	6531	End stage COPD				Pulmonary
145	6555	Type two respiratory failure		Severe COPD		Pulmonary
146	6581	Respiratory failure secondary to severe emphysema	Infective exacerbation of COPD			Pulmonary
147	6589	Acute left ventricular failure		Ischaemic heart disease		Cardiac
148	6600	Metastatic lung cancer				Cancer
149	6615	COPD				Pulmonary

Abbreviations: COPD, chronic obstructive pulmonary disease

Bibliography

[1] Abraham, G., Havulinna, A. S., Bhalala, O. G., Byars, S. G., De Livera, A. M., Yetukuri, L., Tikkanen, E., Perola, M., Schunkert, H., Sijbrands, E. J., Palotie, A., Samani, N. J., Salomaa, V., Ripatti, S., & Inouye, M. (2016). Genomic prediction of coronary heart disease. *European Heart Journal*, 37(43), 3267–3278. 205

[2] Agniel, D., Kohane, I. S., & Weber, G. M. (2018). Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. *BMJ*, 361. 201

[3] Agusti, A. (2007). Systemic effects of chronic obstructive pulmonary disease. *Proceedings of the American Thoracic Society*, 4(7), 522–525. 3

[4] Agusti, A., Bel, E., Thomas, M., Vogelmeier, C., Brusselle, G., Holgate, S., Humbert, M., Jones, P., Gibson, P. G., Vestbo, J., Beasley, R., & Pavord, I. D. (2016). Treatable traits: toward precision medicine of chronic airway diseases. *Eur Respir J*, 47(2), 410–419. 8, 11

[5] Agusti, A., Edwards, L. D., Celli, B., Macnee, W., Calverley, P. M. A., Mullerova, H., Lomas, D. A., Wouters, E., Bakke, P., Rennard, S., Crim, C., Miller, B. E., Coxson, H. O., Yates, J. C., Tal-Singer, R., Vestbo, J., & ECLIPSE Investigators (2013). Characteristics, stability and outcomes of the 2011 GOLD COPD groups in the ECLIPSE cohort. *European Respiratory Society*, 42(3), 636–46. 21, 29

[6] Agusti, A., Edwards, L. D., Rennard, S. I., MacNee, W., Tal-Singer, R., Miller, B. E., Vestbo, J., Lomas, D. A., Calverley, P. M., Wouters, E., Crim, C., Yates, J. C., Silverman, E. K., Coxson, H. O., Bakke, P., Mayer, R. J., Celli, B., & ECLIPSE Investigators (2012).

Persistent systemic inflammation is associated with poor clinical outcomes in COPD: a novel phenotype. *PLOS One*, 7(5), e37483. 12, 29

[7] Agusti, A., Sobradillo, P., & Celli, B. (2011). Addressing the complexity of chronic obstructive pulmonary disease: from phenotypes and biomarkers to scale-free networks, systems biology, and P4 medicine. *Am J Respir Crit Care Med*, 183(9), 1129–37. 32

[8] Ai-Ping, C., Lee, K. H., & Lim, T. K. (2005). In-hospital and 5-year mortality of patients treated in the ICU for acute exacerbation of COPD: a retrospective study. *CHEST*, 128(2), 518–24. 129

[9] Altman, D. G. (1990). *Practical statistics for medical research*. CRC press. 102

[10] Altman, D. G., Vergouwe, Y., Royston, P., & Moons, K. G. (2009). Prognosis and prognostic research: validating a prognostic model. *BMJ*, 338, b605. 10, 102, 185

[11] American Thoracic Society and American College of Chest Physicians (2003). ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med*, 167(2), 211–77. 9

[12] Anderson, L., Oldridge, N., Thompson, D. R., Zwisler, A. D., Rees, K., Martin, N., & Taylor, R. S. (2016). Exercise-based cardiac rehabilitation for coronary heart disease cochrane systematic review and meta-analysis. *Journal of the American College of Cardiology*, 67(1), 1–12. 195

[13] Antonelli-Incalzi, R., Corsonello, A., Pedone, C., Trojano, L., Acanfora, D., Spada, A., Izzo, O., & Rengo, F. (2006). Drawing impairment predicts mortality in severe COPD. *CHEST*, 130(6), 1687–1694. 29, 34

[14] ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories (2002). ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*, 166(1), 111–7. 58

[15] Bafadhel, M., McKenna, S., Terry, S., Mistry, V., Reid, C., Haldar, P., McCormick, M., Haldar, K., Kebadze, T., Duvoix, A., Lindblad, K., Patel, H., Rugman, P., Dodson, P., Jenkins, M., Saunders, M., Newbold, P., Green, R. H., Venge, P., Lomas, D. A., Barer, M. R., Johnston, S. L., Pavord, I. D., & Brightling, C. E. (2011). Acute exacerbations of chronic obstructive pulmonary disease: identification of biologic clusters and their biomarkers. *Am J Respir Crit Care Med*, 184(6), 662–71. 29

[16] Bajekal, M., Scholes, S., Love, H., Hawkins, N., O'Flaherty, M., Raine, R., & Capewell, S. (2012). Analysing recent socioeconomic trends in coronary heart disease mortality in England, 2000–2007: A population modelling study. *PLOS Med*, 9(6), 1–14. 151

[17] Barnes, P. J. & Celli, B. R. (2009). Systemic manifestations and comorbidities of copd. *Eur Respir J*, 33(5), 1165–1185. 3, 21

[18] Bell, S., Daskalopoulou, M., Rapsomaniki, E., George, J., Britton, A., Bobak, M., Casas, J. P., Dale, C. E., Denaxas, S., Shah, A. D., & Hemingway, H. (2017). Association between clinically recorded alcohol consumption and initial presentation of 12 cardiovascular diseases: population based cohort study using linked health records. *BMJ*, 356, j909. 196

[19] Ben-Shlomo, Y., Spears, M., Boustred, C., May, M., Anderson, S. G., Benjamin, E. J., Boutouyrie, P., Cameron, J., Chen, C.-H., Cruickshank, J. K., Hwang, S.-J., Lakatta, E. G., Laurent, S., Maldonado, J., Mitchell, G. F., Najjar, S. S., Newman, A. B., Ohishi, M., Pannier, B., Pereira, T., Vasan, R. S., Shokawa, T., Sutton-Tyrell, K., Verbeke, F., Wang, K.-L., Webb, D. J., Willum Hansen, T., Zoungas, S., McEniery, C. M., Cockcroft, J. R., & Wilkinson, I. B. (2014). Aortic pulse wave velocity improves cardiovascular event prediction. *Journal of the American College of Cardiology*, 63(7), 636. 13

[20] Bernabeu-Mora, R., Medina-Mirapeix, F., & Llamazares-Herran, E. (2016). The short physical performance battery is a discriminative tool for identifying patients with COPD at risk of disability. *International Journal of COPD*, 11, 623–623. 58, 101

[21] Bhatnagar, P., Wickramasinghe, K., Wilkins, E., & Townsend, N. (2016). Trends in the epidemiology of cardiovascular disease in the uk. *Heart*, 102(24), 1945–1952. 151

[22] Blanchette, C. M., Gross, N. J., & Altman, P. (2014). Rising costs of COPD and the potential for maintenance therapy to slow the trend. *Am Health Drug Benefits*, 7(2), 98–106. 150

[23] Blumenthal, J. A., Smith, P. J., Durheim, M., Mabe, S., Emery, C. F., Martinu, T., Diaz, P. T., Babyak, M., Welty-Wolf, K., & Palmer, S. (2016). Biobehavioral prognostic factors in chronic obstructive pulmonary disease: Results from the INSPIRE-II trial. *Psychosomatic Medicine*, 78(2), 153–162. 29, 32, 34, 53

[24] Bone, A. E., Hepgul, N., Kon, S., & Maddocks, M. (2017). Sarcopenia and frailty in chronic respiratory disease. *Chron Respir Dis*, 14(1), 85–99. 206

[25] Bots, M. L., Grobbee, D. E., Hofman, A., & Witteman, J. C. (2005). Common carotid intima-media thickness and risk of acute myocardial infarction: the role of lumen diameter. *Stroke*, 36(4), 762–7. 13

[26] Bots, M. L., Hoes, A. W., Koudstaal, P. J., Hofman, A., & Grobbee, D. E. (1997). Common carotid intima-media thickness and risk of stroke and myocardial infarction: the Rotterdam study. *Circulation*, 96(5), 1432–7. 13

[27] Bouwmeester, W., Zuithoff, N. P., Mallett, S., Geerlings, M. I., Vergouwe, Y., Steyerberg, E. W., Altman, D. G., & Moons, K. G. (2012). Reporting and methods in clinical prediction research: a systematic review. *PLOS Med*, 9(5), 1–12. 10

[28] Briggs, A., Spencer, M., Wang, H., Mannino, D., & Sin, D. D. (2008). Development and validation of a prognostic index for health outcomes in chronic obstructive pulmonary disease. *Arch Intern Med*, 168(1), 71–9. 100

[29] British Thoracic Society and the Primary Care Respiratory Society UK (2012). *IMPRESS Guide to the relative value of COPD interventions*. Report, British Thoracic Society. 8, 20

[30] Bu, X. N., Yang, T., Thompson, M. A., Hutchinson, A. F., & Irving, L. B. (2011). Changes in the BODE index, exacerbation duration and hospitalisation in a cohort of COPD patients. *Singapore Medical Journal*, 52(12), 894–900. 29

[31] Budweiser, S., Hitzl, A. P., Joerres, R. A., Schmidbauer, K., Heinemann, F., & Pfeifer, M. (2007). Health-related quality of life and long-term prognosis in chronic hypercapnic respiratory failure: a prospective survival analysis. *Respiratory Research*, 8. 29, 34, 48

[32] Burrows, B. & Earle, R. H. (1969). Course and prognosis of chronic obstructive lung disease. *N Engl J Med*, 280(8), 397–404. 129

[33] Cano, N. J., Pichard, C., Roth, H., Court-Fortuné, I., Cynober, L., Gérard-Boncompain, M., Cuvelier, A., Laaban, J. P., Melchior, J. C., Raphaël, J. C., & Pison, C. M. (2004). C-reactive protein and body mass index predict outcome in end-stage respiratory failure. *CHEST*, 126(2), 540–546. 29, 34, 42

[34] Carter, R., Nicotra, B., Blevins, W., & Holiday, D. (1993). Altered exercise gas exchange and cardiac function in patients with mild chronic obstructive pulmonary disease. *CHEST*, 103(3), 745–50. 3

[35] Casanova, C., Cote, C., de Torres, J. P., Aguirre-Jaime, A., Marin, J. M., Pinto-Plata, V., & Celli, B. R. (2005). Inspiratory-to-total lung capacity ratio predicts mortality in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 171(6), 591–7. 124

[36] Casanova, C., Cote, C., Marin, J. M., Pinto-Plata, V., de Torres, J. P., Aguirre-Jaime, A., Vassaux, C., & Celli, B. R. (2008). Distance and oxygen desaturation during the 6-min walk test as predictors of long-term mortality in patients with COPD. *CHEST*, 134(4), 746–52. 29, 33

[37] Cavaillès, A., Brinchault-Rabin, G., Dixmier, A., Goupil, F., Gut-Gobert, C., Marchand-Adam, S., Meurice, J.-C., Morel, H., Person-Tacnet, C., Leroyer, C., et al. (2013). Comorbidities of COPD. *European Respiratory Review*, 22(130), 454–475. 207

[38] Celli, B., Tetzlaff, K., Criner, G., Polkey, M. I., Sciurba, F., Casaburi, R., Tal-Singer, R., Kawata, A., Merrill, D., Rennard, S., & Consortium, C. B. Q. (2016). The 6-minute walk test as a COPD stratification tool: insights from the COPD biomarker qualification consortium. *Am J Respir Crit Care Med*. 14

[39] Celli, B., Vestbo, J., Jenkins, C. R., Jones, P. W., Ferguson, G. T., Calverley, P. M. A., Yates, J. C., Anderson, J. A., Willits, L. R., Wise, R. A., & TORCH Study Investigators (2011). Sex differences in mortality and clinical expressions of patients with chronic obstructive pulmonary disease the torch experience. *Am J Respir Crit Care Med*, 183(3), 317–322. 129

[40] Celli, B. R., Cote, C. G., Marin, J. M., Casanova, C., Montes de Oca, M., Mendez, R. A., Pinto Plata, V., & Cabral, H. J. (2004a). The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med*, 350(10), 1005–12. 10, 99, 100, 101, 183

[41] Celli, B. R., Locantore, N., Tal-Singer, R., Riley, J., Miller, B., Vestbo, J., Yates, J. C., Silverman, E. K., Owen, C. A., Divo, M., Pinto-Plata, V., Wouters, E. F. M., Faner, R., Agusti, A., & ECLIPSE Investigators (2018). Emphysema and extrapulmonary tissue loss in COPD: a multi-organ loss of tissue phenotype. *Eur Respir J*, 51(2). 207

[42] Celli, B. R., Locantore, N., Yates, J., Tal-Singer, R., Miller, B. E., Bakke, P., Calverley, P., Coxson, H., Crim, C., Edwards, L. D., Lomas, D. A., Duvoix, A., MacNee, W., Rennard, S., Silverman, E., Vestbo, J., Wouters, E., Agustí, A., Celli, B. R., & Locantore, N. (2012). Inflammatory biomarkers improve clinical prediction of mortality in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 185(10), 1065–1072. 29, 32, 34, 38, 48, 49, 100, 124

[43] Celli, B. R., MacNee, W., & Force, A. E. T. (2004b). Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J*, 23(6), 932–46. 3, 21

[44] Cesari, M., Landi, F., Vellas, B., Bernabei, R., & Marzetti, E. (2014). Sarcopenia and physical frailty: two sides of the same coin. *Front Aging Neurosci*, 6, 192. 207

[45] Chen, C. Z., Ou, C. Y., Yu, C. H., Yang, S. C., Chang, H. Y., & Hsiue, T. R. (2015a). Comparison of global initiative for chronic obstructive pulmonary disease 2013 classification and body mass index, airflow obstruction, dyspnea, and exacerbations index in predicting mortality and exacerbations in elderly adults with chronic obstructive pulmonary disease. *J Am Geriatr Soc*, 63(2), 244–50. 11

[46] Chen, W., Thomas, J., Sadatsafavi, M., & FitzGerald, J. M. (2015b). Risk of cardiovascular comorbidity in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Lancet Respir Med*, 3(8), 631–9. 129, 181

[47] Chen, Y. W., Leung, J. M., & Sin, D. D. (2016). A systematic review of diagnostic biomarkers of COPD exacerbation. *PLOS One*, 11(7), e0158843. 21, 205

[48] Chung, L. P., Lake, F., Hyde, E., McCamley, C., Phuangmalai, N., Lim, M., Waterer, G., Summers, Q., & Moodley, Y. (2016). Integrated multidisciplinary community service for chronic obstructive pulmonary disease reduces hospitalisations. *Intern Med J*, 46(4), 427–34. 8, 20

[49] Cinarka, H., Kayhan, S., Gumus, A., Durakoglugil, M. E., Erdogan, T., Ezberci, I., Yavuz, A., Ozkaya, S., & Sahin, U. (2013). Arterial stiffness measured by carotid femoral pulse wave velocity is associated with disease severity in chronic obstructive pulmonary disease. *Respiratory Care*. 52

[50] Claes, L., Recknagel, S., & Ignatius, A. (2012). Fracture healing under healthy and inflammatory conditions. *Nature Reviews Rheumatology*, 8(3), 133–143. 207

[51] Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Elsevier Science. 60, 83, 160, 185

[52] Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod). *Ann Intern Med*, 162(10), 735–6. 9, 11, 101

[53] Connors, A. F., J., Dawson, N. V., Thomas, C., Harrell, F. E., J., Desbiens, N., Fulkerson, W. J., Kussin, P., Bellamy, P., Goldman, L., & Knaus, W. A. (1996). Outcomes following acute exacerbation of severe chronic obstructive lung disease. The SUPPORT investigators (study to understand prognoses and preferences for outcomes and risks of treatments). *Am J Respir Crit Care Med*, 154(4 Pt 1), 959–67. 129

[54] Cooper, C. B., Barjaktarevic, I., Curtis, J., Martinez, C., Bowler, R. P., O'Neal, W., Hoffman, E. A., Carretta, E. E., Qibrera, M., & Barr, R. (2017). A103. a novel debility score predicts future exacerbations and death in SPIROMICS. *Am Thoracic Soc*, 1073-449X, A2721–A2721. 54

[55] Cote, C. G., Casanova, C., Marin, J. M., Lopez, M. V., Pinto-Plata, V., De Oca, M. M., Dordelly, L. J., Nekach, H., & Celli, B. R. (2008a). Validation and comparison of reference equations for the 6-min walk distance test. *Eur Respir J*, 31(3), 571–578. 29, 33

[56] Cote, C. G., Dordelly, L. J., & Celli, B. R. (2007a). Impact of COPD exacerbations on patient-centered outcomes. *CHEST*, 131(3), 696–704. 29, 33

[57] Cote, C. G., Pinto-Plata, V., Kasprzyk, K., Dordelly, L. J., & Celli, B. R. (2007b). The 6-min walk distance, peak oxygen uptake, and mortality in COPD. *CHEST*, 132(6), 1778–85. 29, 32, 34, 38

[58] Cote, C. G., Pinto-Plata, V. M., Marin, J. M., Nekach, H., Dordelly, L. J., & Celli, B. R. (2008b). The modified BODE index: validation with mortality in COPD. *Eur Respir J*, 32(5), 1269–74. 29, 100

[59] Crisafulli, E., Costi, S., Luppi, F., Cirelli, G., Cilione, C., Coletti, O., Fabbri, L. M., & Clini, E. M. (2008). Role of comorbidities in a cohort of patients with COPD undergoing pulmonary rehabilitation. *Thorax*, 63(6), 487–92. 8, 20

[60] Cruz-Jentoft, A. J., Baeyens, J. P., Bauer, J. M., Boirie, Y., Cederholm, T., Landi, F., Martin, F. C., Michel, J. P., Rolland, Y., Schneider, S. M., Topinkova, E., Vandewoude, M., Zamboni, M., & European Working Group on Sarcopenia in Older People (2010). Sarcopenia: European consensus on definition and diagnosis: Report of the european working group on sarcopenia in older people. *Age Ageing*, 39(4), 412–23. 207

[61] Csanyi, A. & Egervari, A. (1996). Simple clinical method of average intima-media thickness measurement in the common carotid artery. *Vasa-Journal of Vascular Diseases*, 25(3), 242–248. 183

[62] D'Agostino, R. B., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). General cardiovascular risk profile for use in primary care. *Circulation*, 117(6), 743–753. 181

[63] Dahl, M., Tybjœrg-Hansen, A., Vestbo, J., Lange, P., & Nordestgaard, B. G. (2001). Elevated plasma fibrinogen associated with reduced pulmonary function and increased risk of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 164(6), 1008–1011. 29

[64] Dahl, M., Vestbo, J., Lange, P., Bojesen, S. E., Tybjaerg-Hansen, A., & Nordestgaard, B. G. (2007). C-reactive protein as a predictor of prognosis in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 175(3), 250–255. 29

[65] Dahl, M., Vestbo, J., Zacho, J., Lange, P., Tybjaerg-Hansen, A., & Nordestgaard, B. G. (2011). C reactive protein and chronic obstructive pulmonary disease: a mendelian randomisation approach. *Thorax*, 66(3), 197–204. 29, 42

[66] Dajczman, E., Wardini, R., Kasymjanova, G., Prefontaine, D., Baltzan, M. A., & Wolkove, N. (2015). Six minute walk distance is a predictor of survival in patients with chronic obstruc-

tive pulmonary disease undergoing pulmonary rehabilitation. *Canadian Respiratory Journal*, 22(4), 225–9. 29, 33, 34

[67] Damen, J. A., Hooft, L., Schuit, E., Debray, T. P., Collins, G. S., Tzoulaki, I., Lassale, C. M., Siontis, G. C., Chiocchia, V., Roberts, C., Schlussel, M. M., Gerry, S., Black, J. A., Heus, P., van der Schouw, Y. T., Peelen, L. M., & Moons, K. G. (2016). Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ*, 353, i2416. 210

[68] Davies, S. (2018). Annual report of the chief medical officer 2017. 7

[69] De Martinis, M., Franceschi, C., Monti, D., & Ginaldi, L. (2005). Inflamm-ageing and lifelong antigenic load as major determinants of ageing rate and longevity. *FEBS Letters*, 579(10), 2035–2039. 3

[70] de Ruijter, W., Westendorp, R. G., Assendelft, W. J., den Elzen, W. P., de Craen, A. J., le Cessie, S., & Gussekloo, J. (2009). Use of framingham risk score and new biomarkers to predict cardiovascular mortality in older people: population based observational cohort study. *BMJ*, 338, a3083. 181

[71] de Torres, J. P., Cote, C. G., Lopez, M. V., Casanova, C., Diaz, O., Marin, J. M., Pinto-Plata, V., de Oca, M. M., Nekach, H., Dordelly, L. J., Aguirre-Jaime, A., & Celli, B. R. (2009). Sex differences in mortality in patients with COPD. *Eur Respir J*, 33(3), 528–35. 129

[72] de Torres, J. P., Pinto-Plata, V., Casanova, C., Mullerova, H., Cordoba-Lanus, E., de Fuentes, M. M., Aguirre-Jaime, A., & Celli, B. R. (2008). C-reactive protein levels and survival in patients with moderate to very severe COPD. *CHEST*, 133(6), 1336–1343. 29, 32, 34, 42

[73] Den Ruijter, H. M., Peters, S. A. E., Anderson, T. J., Britton, A. R., Dekker, J. M., Eijkemans, M. J., Engström, G., Evans, G. W., de Graaf, J., Grobbee, D. E., Hedblad, B.,

Hofman, A., Holewijn, S., Ikeda, A., Kavousi, M., Kitagawa, K., Kitamura, A., Koffijberg, H., Lonn, E. M., Lorenz, M. W., Mathiesen, E. B., Nijpels, G., Okazaki, S., O'Leary, D. H., Polak, J. F., Price, J. F., Robertson, C., Rembold, C. M., Rosvall, M., Rundek, T., Salonen, J. T., Sitzer, M., Stehouwer, C. D. A., Witteman, J. C., Moons, K. G., & Bots, M. L. (2012). Common Carotid Intima-Media Thickness Measurements in Cardiovascular Risk Prediction: A Meta-analysisCarotid Intima-Media Thickness and Risk Prediction. *JAMA*, 308(8), 796–803. 181

[74] Deng, Z. C., Zhao, P., Cao, C., Sun, S. F., Zhao, F., Lu, C. Y., & Ma, H. Y. (2014). C-reactive protein as a prognostic marker in chronic obstructive pulmonary disease. *Experimental and Therapeutic Medicine*, 7(2), 443–446. 29

[75] Department of Health (2011). *An Outcomes Strategy for Chronic Obstructive Pulmonary Disease (COPD) and Asthma in England.* Report, Department of Health. 8, 20

[76] Department of Health (2013). *Cardiovascular Disease Outcomes Strategy. Improving outcomes for people with or at risk of cardiovascular disease.* Report, Cardiovascular Disease Team. 151

[77] Donaldson, A. V., Maddocks, M., Martolini, D., Polkey, M. I., & Man, W. D. (2012). Muscle function in COPD: a complex interplay. *Int J Chron Obstruct Pulmon Dis*, 7, 523–35. 13

[78] Donaldson, G. C., Seemungal, T. A. R., Patel, I. S., Bhowmik, A., Wilkinson, T. M. A., Hurst, J. R., MacCallum, P. K., & Wedzicha, J. A. (2005). Airway and systemic inflammation and decline in lung function in patients with COPD. *CHEST*, 128(4), 1995–2004. 29

[79] Dransfield, M. T., Rowe, S. M., Johnson, J. E., Bailey, W. C., & Gerald, L. B. (2008). Use of beta blockers and the risk of death in hospitalised patients with acute exacerbations of COPD. *Thorax*, 63(4), 301–5. 150

[80] Dreyse, J., Diaz, O., Repetto, P. B., Morales, A., Saldias, F., & Lisboa, C. (2015). Do frequent moderate exacerbations contribute to progression of chronic obstructive pulmonary disease in patients who are ex-smokers? *International Journal of COPD*, 10, 525–33. 30, 33, 34, 38, 42

[81] Drummond, M. B., Wise, R. A., John, M., Zvarich, M. T., & McGarvey, L. P. (2010). Accuracy of death certificates in COPD: analysis from the TORCH trial. *COPD*, 7(3), 179–85. 129

[82] Durheim, M. T., Smith, P. J., Babyak, M. A., Mabe, S. K., Martinu, T., Welty-Wolf, K. E., Emery, C. F., Palmer, S. M., & Blumenthal, J. A. (2015). Six-minute-walk distance and accelerometry predict outcomes in chronic obstructive pulmonary disease independent of global initiative for chronic obstructive lung disease 2011 group. *Annals of the American Thoracic Society*, 12(3), 349–356. 30, 32

[83] Duvoix, A., Dickens, J., Haq, I., Mannino, D., Miller, B., Tal-Singer, R., & Lomas, D. A. (2013). Blood fibrinogen as a biomarker of chronic obstructive pulmonary disease. *Thorax*, 68(7), 670–676. 195

[84] Edan, N. M. (2018). Proposed prototype for multimedia protocol in m-health to support a communication between patient and hospital. *Tikrit Journal of Pure Science*, 22(6), 144–150. 204

[85] Edwards, R. H., Young, A., Hosking, G. P., & Jones, D. A. (1977). Human skeletal muscle function: description of tests and normal values. *Clin Sci Mol Med*, 52(3), 283–90. 58

[86] Emerging Risk Factors Collaboration (2007). The emerging risk factors collaboration: analysis of individual data on lipid, inflammatory and other markers in over 1.1 million participants in 104 prospective studies of cardiovascular diseases. *European Journal of Epidemiology*, 22(12), 839–869. 196

[87] Esteban, C., Quintana, J. M., Aburto, M., Moraza, J., & Capelastegui, A. (2006). A simple score for assessing stable chronic obstructive pulmonary disease. *QJM*, 99(11), 751–9. 100

[88] European Medicines Agency (2018). *Reflection paper on physical frailty: instruments for baseline characterisation of older populations in clinical trials*. Report, European Medicines Agency. 124

[89] Faganello, M. M., Tanni, S. E., Sanchez, F. F., Pelegrino, N. R., Lucheta, P. A., & Godoy, I. (2010). BODE index and GOLD staging as predictors of 1-year exacerbation risk in chronic obstructive pulmonary disease. *Am J Med Sci*, 339(1), 10–4. 30, 32, 38

[90] Farmer, A., Toms, C., Hardinge, M., Williams, V., Rutter, H., & Tarassenko, L. (2014). Self-management support using an internet-linked tablet computer (the EDGE platform)-based intervention in chronic obstructive pulmonary disease: protocol for the EDGE-COPD randomised controlled trial. *BMJ Open*, 4(1). 175, 204

[91] Fermont, J. M., Masconi, K. L., Jensen, M. T., Ferrari, R., Di Lorenzo, V. A. P., Marott, J. L., Schuetz, P., Watz, H., Waschki, B., Müllerova, H., Polkey, M. I., Wilkinson, I. B., & Wood, A. M. (2019). Biomarkers and clinical outcomes in COPD: a systematic review and meta-analysis. *Thorax*. 182

[92] Ferrari, R., Tanni, S. E., Caram, L. M. O., Correa, C., Correa, C. R., & Godoy, I. (2013). Three-year follow-up of interleukin 6 and c-reactive protein in chronic obstructive pulmonary disease. *Respiratory Research*, 14. 30, 32, 34, 42

[93] Ferrari, R., Tanni, S. E., Caram, L. M. O., Naves, C. R., & Godoy, I. (2011). Predictors of health status do not change over three-year periods and exacerbation makes difference in chronic obstructive pulmonary disease. *Health and Quality of Life Outcomes*, 9, 112. 30, 33, 38

[94] Fisk, M., McEniery, C. M., Gale, N., Maki-Petaja, K., Forman, J. R., Munnery, M., Woodcock-Smith, J., Cheriyan, J., Mohan, D., Fuld, J., Tal-Singer, R., Polkey, M. I., Cock-

croft, J. R., Wilkinson, I. B., & ERICA Consortium and ACCT Investigators (2018). Surrogate markers of cardiovascular risk and chronic obstructive pulmonary disease: a large case-controlled study. *Hypertension*, 71(3), 499–506. 52, 181

[95] Folsom, A., Kronmal, R., Detrano, R., et al. (2008). Coronary artery calcification compared with carotid intima-media thickness in the prediction of cardiovascular disease incidence: The multi-ethnic study of atherosclerosis (MESA). *Archives of Internal Medicine*, 168(12), 1333–1339. 13

[96] Frei, A., Siebeling, L., Wolters, C., Held, L., Muggensturm, P., Strassmann, A., Zoller, M., ter Riet, G., & Puhan, M. A. (2016). The inaccuracy of patient recall for COPD exacerbation rate estimation and its implications: Results from central adjudication. *CHEST*, 150(4), 860–868. 94, 175

[97] Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., & Allen, N. E. (2017). Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol*, 186(9), 1026–1034. 149

[98] Fuso, L., Incalzi, R. A., Pistelli, R., Muzzolon, R., Valente, S., Pagliari, G., Gliozzi, F., & Ciappi, G. (1995). Predicting mortality of patients hospitalized for acutely exacerbated chronic obstructive pulmonary disease. *Am J Med*, 98(3), 272–7. 129

[99] Gaki, E., Papatheodorou, G., Pappa, I., Papiris, S., & Loukides, S. (2011). Correlation between frequency of hospitalization of patients with severe COPD and severity indices. *Pneumon*, 24(2), 164–171. 30

[100] Gale, N. S., Albarrati, A., Munnery, I., Munnery, M., Tal-Singer, R., Hubbard, R., Cockcroft, J. R., & Shale, D. (2012). P75 physical impairment and frailty in patients with chronic obstruction pulmonary disease (COPD). *Thorax*, 67(Suppl 2), A96–A96. 206

[101] Gale, N. S., Albarrati, A. M., Munnery, M. M., Munnery, I. C., Irfan, M., Bolton, C. E., Rambaran, C. N., Singer, R. M. T., Cockcroft, J. R., & Shale, D. J. (2014). Assessment of risk in chronic airways disease evaluation (arcade): Protocol and preliminary data. *Chronic Respiratory Disease*, 11(4), 199–207. 15, 56

[102] Gan, W. Q., Man, S., Senthilselvan, A., & Sin, D. (2004). Association between chronic obstructive pulmonary disease and systemic inflammation: a systematic review and a meta-analysis. *Thorax*, 59(7), 574–580. 205

[103] Ganna, A. & Ingelsson, E. (2015). 5 year mortality predictors in 498103 UK Biobank participants: a prospective population-based study. *The Lancet*, 386(9993), 533–540. 150

[104] Garcia-Aymerich, J., Lange, P., Benet, M., Schnohr, P., & Anto, J. M. (2006). Regular physical activity reduces hospital admission and mortality in chronic obstructive pulmonary disease: a population based cohort study. *Thorax*, 61(9), 772–778. 155

[105] Ghobadi, H., Ahari, S. S., Kameli, A., & Lari, S. M. (2012). The relationship between COPD assessment test (CAT) scores and severity of airflow obstruction in stable COPD patients. *Tanaffos*, 11(2), 22–26. 59

[106] Global Burden of Disease Chronic Respiratory Disease Collaborators (2017). Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the global burden of disease study 2015. *Lancet Respir Med*, 5(9), 691–706. 5, 8, 98

[107] Global Initiative for Chronic Obstructive Lung Disease (2016). *Global Strategy for the Diagnosis, Management and Prevention of COPD*. Report, GOLD. 2, 7, 8, 31, 34, 58, 60, 101, 131, 156, 157, 182, 183

[108] Goldstein, B. A., Navar, A. M., Pencina, M. J., & Ioannidis, J. P. (2017). Opportunities

and challenges in developing risk prediction models with electronic health records data: a systematic review. *J Am Med Inform Assoc*, 24(1), 198–208. 10

[109] Gosselink, R., Troosters, T., & Decramer, M. (1996). Peripheral muscle weakness contributes to exercise limitation in COPD. *Am J Respir Crit Care Med*, 153(3), 976–80. 13

[110] Graat-Verboom, L., van den Borne, B. E., Smeenk, F. W., Spruit, M. A., & Wouters, E. F. (2011). Osteoporosis in COPD outpatients based on bone mineral density and vertebral fractures. *Journal of Bone and Mineral Research*, 26(3), 561–568. 207

[111] Groenewegen, K. H., Postma, D. S., Hop, W. C., Wielders, P. L., Schlösser, N. J., & Wouters, E. F. (2008). Increased systemic inflammation is a risk factor for COPD exacerbations. *CHEST*, 133(2), 350–357. 30, 42, 49

[112] Grolimund, E., Kutz, A., Marlowe, R. J., Vogeli, A., Alan, M., Christ-Crain, M., Thomann, R., Falconnier, C., Hoess, C., Henzen, C., Zimmerli, W., Mueller, B., & Schuetz, P. (2015). Long-term prognosis in COPD exacerbation: Role of biomarkers, clinical variables and exacerbation type. *Journal of Chronic Obstructive Pulmonary Disease*, 12(3), 295–305. 30, 34, 42, 48

[113] Guerra, B., Gaveikaite, V., Bianchi, C., & Puhan, M. A. (2017). Prediction models for exacerbations in patients with COPD. *Eur Respir Rev*, 26(143). 155, 210

[114] Guevara, J. P., Berlin, J. A., & Wolf, F. M. (2004). Meta-analytic methods for pooling rates when follow-up duration varies: a case study. *BMC Medical Research Methodology*, 4(1), 17. 53

[115] Guo, S. J., Sun, Z. T., Liu, E. S., Feng, J. H., Fu, M., Li, Y. C., & Wu, Q. (2014). Effect of bufei granule on stable chronic obstructive pulmonary disease: a randomized, double blinded, placebo-controlled, and multicenter clinical study. *Journal of Traditional Chinese Medicine*, 34(4), 437–444. 30

[116] Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., Scherr, P. A., & Wallace, R. B. (1994). A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol*, 49(2), M85–94. 14

[117] Hamblin, A., Wordsworth, S., Fermont, J. M., Page, S., Kaur, K., Camps, C., Kaisaki, P., Gupta, A., Talbot, D., Middleton, M., Henderson, S., Cutts, A., Vavoulis, D. V., Housby, N., Tomlinson, I., Taylor, J. C., & Schuh, A. (2017). Clinical applicability and cost of a 46-gene panel for genomic analysis of solid tumours: Retrospective validation and prospective audit in the UK National Health Service. *PLOS Med*, 14(2), e1002230. 211

[118] Hansell, A., Walk, J., & Soriano, J. (2003). What do chronic obstructive pulmonary disease patients die from? A multiple cause coding analysis. *Eur Respir J*, 22(5), 809–814. 129

[119] Hansen, T. W., Staessen, J. A., Torp-Pedersen, C., Rasmussen, S., Li, Y., Dolan, E., Thijs, L., Wang, J.-G., O'Brien, E., Ibsen, H., & Jeppesen, J. (2006a). Ambulatory arterial stiffness index predicts stroke in a general population. *Journal of Hypertension*, 24(11), 2247–2253. 181

[120] Hansen, T. W., Staessen, J. A., Torp-Pedersen, C., Rasmussen, S., Thijs, L., Ibsen, H., & Jeppesen, J. (2006b). Prognostic value of aortic pulse wave velocity as index of arterial stiffness in the general population. *Circulation*, 113(5), 664–670. 13

[121] Herrett, E., Thomas, S. L., Schoonen, W. M., Smeeth, L., & Hall, A. J. (2010). Validation and validity of diagnoses in the general practice research database: a systematic review. *Br J Clin Pharmacol*, 69(1), 4–14. 196

[122] Hersh, C. P., DeMeo, D. L., Al-Ansari, E., Carey, V. J., Reilly, J. J., Ginns, L. C., & Silverman, E. K. (2004). Predictors of survival in severe, early onset COPD. *CHEST*, 126(5), 1443–51. 8, 20

[123] Hicks, K. A., Mahaffey, K. W., Mehran, R., Nissen, S. E., Wiviott, S. D., Dunn, B., Solomon, S. D., Marler, J. R., Teerlink, J. R., Farb, A., Morrow, D. A., Targum, S. L., Sila, C. A., Hai, M. T. T., Jaff, M. R., Joffe, H. V., Cutlip, D. E., Desai, A. S., Lewis, E. F., Gibson, C. M., Landray, M. J., Lincoff, A. M., White, C. J., Brooks, S. S., Rosenfield, K., Domanski, M. J., Lansky, A. J., McMurray, J. J. V., Tcheng, J. E., Steinhubl, S. R., Burton, P., Mauri, L., O'Connor, C. M., Pfeffer, M. A., Hung, H. M. J., Stockbridge, N. L., Chaitman, B. R., Temple, R. J., & Standardized Data Collection for Cardiovascular Trials Initiative (2018). 2017 Cardiovascular and stroke endpoint definitions for clinical trials. *Circulation*, 137(9), 961–972. 196

[124] Higgins, J. P. & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions*, volume 4. John Wiley and Sons. 24

[125] Hillas, G., Perlikos, F., & Tzanakis, N. (2016). Acute exacerbation of COPD: is it the "stroke of the lungs"? *Int J Chron Obstruct Pulmon Dis*, 11, 1579–86. 155

[126] Hogg, J. C., Chu, F., Utokaparch, S., Woods, R., Elliott, W. M., Buzatu, L., Cherniack, R. M., Rogers, R. M., Sciurba, F. C., Coxson, H. O., & Paré, P. D. (2004). The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med*, 350(26), 2645–2653. 3

[127] Holland, A. E., Spruit, M. A., Troosters, T., Puhan, M. A., Pepin, V., Saey, D., McCormack, M. C., Carlin, B. W., Sciurba, F. C., Pitta, F., Wanger, J., MacIntyre, N., Kaminsky, D. A., Culver, B. H., Revill, S. M., Hernandes, N. A., Andrianopoulos, V., Camillo, C. A., Mitchell, K. E., Lee, A. L., Hill, C. J., & Singh, S. J. (2014). An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. *Eur Respir J*, 44(6), 1428–46. 99

[128] Hopkinson, N. S., Tennant, R. C., Dayer, M. J., Swallow, E. B., Hansel, T. T., Moxham,

J., & Polkey, M. I. (2007). A prospective study of decline in fat free mass and skeletal muscle strength in chronic obstructive pulmonary disease. *Respiratory research*, 8, 25. 30

[129] Hosmer, D. W. & Lemeshow, S. (2000). *Applied logistic regression (Wiley Series in probability and statistics)*. Wiley-Interscience Publication, 2nd edition. 10

[130] Hunninghake, G. M., Cho, M. H., Tesfaigzi, Y., Soto-Quiros, M. E., Avila, L., Lasky-Su, J., Stidley, C., Melen, E., Soderhall, C., Hallberg, J., Kull, I., Kere, J., Svartengren, M., Pershagen, G., Wickman, M., Lange, C., Demeo, D. L., Hersh, C. P., Klanderman, B. J., Raby, B. A., Sparrow, D., Shapiro, S. D., Silverman, E. K., Litonjua, A. A., Weiss, S. T., & Celedon, J. C. (2009). MMP12, lung function, and COPD in high-risk populations. *N Engl J Med*, 361(27), 2599–608. 209

[131] Hunter, L. C., Lee, R. J., Butcher, I., Weir, C. J., Fischbacher, C. M., McAllister, D., Wild, S. H., Hewitt, N., & Hardie, R. M. (2016). Patient characteristics associated with risk of first hospital admission and readmission for acute exacerbation of chronic obstructive pulmonary disease (COPD) following primary care COPD diagnosis: a cohort study using linked electronic patient records. *BMJ Open*, 6(1), e009121. 155

[132] Hurst, J. R., Vestbo, J., Anzueto, A., Locantore, N., Mullerova, H., Tal-Singer, R., Miller, B., Lomas, D. A., Agusti, A., Macnee, W., Calverley, P., Rennard, S., Wouters, E. F., Wedzicha, J. A., & ECLIPSE Investigators (2010). Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med*, 363(12), 1128–38. 30, 38, 48, 49, 157, 160, 176

[133] Husebo, G. R., Bakke, P. S., Aanerud, M., Hardie, J. A., Ueland, T., Gronseth, R., Persson, L. J. P., Aukrust, P., & Eagan, T. M. (2014). Predictors of exacerbations in chronic obstructive pulmonary disease–results from the Bergen COPD cohort study. *PLOS One*, 9(10). 30, 34, 38, 42, 48

[134] Institute Metrics and Evaluation (2013). Global Burden of Disease Compare. Chronic obstructive pulmonary disease. 5

[135] Iwamoto, H., Yokoyama, A., Kitahara, Y., Ishikawa, N., Haruta, Y., Yamane, K., Hattori, N., Hara, H., & Kohno, N. (2009). Airflow limitation in smokers is associated with subclinical atherosclerosis. *Am J Respir Crit Care Med*, 179(1), 35–40. 52

[136] Jenkins, C. R., Chapman, K. R., Donohue, J. F., Roche, N., Tsiligianni, I., & Han, M. K. (2017). Improving the management of COPD in women. *CHEST*, 151(3), 686–696. 151

[137] Jennings, J. H., Digiovine, B., Obeid, D., & Frank, C. (2009). The association between depressive symptoms and acute exacerbations of COPD. *Lung*, 187(2), 128–35. 30, 38, 42

[138] Jensen, H. H., Godtfredsen, N. S., Lange, P., & Vestbo, J. (2006). Potential misclassification of causes of death from COPD. *Eur Respir J*, 28(4), 781–785. 129, 197

[139] Jensen, M. T., Marott, J. L., Lange, P., Vestbo, J., Schnohr, P., Nielsen, O. W., Jensen, J. S., & Jensen, G. B. (2013). Resting heart rate is a predictor of mortality in COPD. *Eur Respir J*, 42(2), 341–349. 30, 34, 42

[140] Jinks, R. C., Royston, P., & Parmar, M. K. (2015). Discrimination-based sample size calculations for multivariable prognostic models for time-to-event data. *BMC Med Res Methodol*, 15, 82. 202

[141] Johannessen, A., Nilsen, R. M., Storebo, M., Gulsvik, A., Eagan, T., & Bakke, P. (2013). Comparison of 2011 and 2007 global initiative for chronic obstructive lung disease guidelines for predicting mortality and hospitalization. *Am J Respir Crit Care Med*, 188(1), 51–9. 9

[142] Jones, P. W., Quirk, F. H., & Baveystock, C. M. (1991). The St George's Respiratory Questionnaire. *Respir Med*, 85 Suppl B, 25–31; discussion 33–7. 59

[143] Kannel, W. B. & McGee, D. L. (1979). Diabetes and glucose tolerance as risk factors for cardiovascular disease: the Framingham Study. *Diabetes Care*, 2(2), 120–126. 181

[144] Kappen, T. H., van Klei, W. A., van Wolfswinkel, L., Kalkman, C. J., Vergouwe, Y.,

& Moons, K. G. M. (2018). Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagnostic and Prognostic Research*, 2(1), 11. 10

[145] Kennedy, C. C., Novotny, P. J., LeBrasseur, N. K., Wise, R. A., Sciurba, F. C., Benzo, R. P., & Group, N. R. (2018). Frailty and clinical outcomes in chronic obstructive pulmonary disease. *Ann Am Thorac Soc*. 207

[146] Kerr, K. F., Wang, Z. Y., Janes, H., McClelland, R. L., Psaty, B. M., & Pepe, M. S. (2014). Net reclassification indices for evaluating risk prediction instruments a critical review. *Epidemiology*, 25(1), 114–121. 208

[147] Knowles, J. W. & Ashley, E. A. (2018). Cardiovascular disease: The rise of the genetic risk score. *PLOS Med*, 15(3), e1002546. 209

[148] Kon, S. S., Jones, S. E., Schofield, S. J., Banya, W., Dickson, M. J., Canavan, J. L., Nolan, C. M., Haselden, B. M., Polkey, M. I., Cullinan, P., & Man, W. D. (2015). Gait speed and readmission following hospitalisation for acute exacerbations of COPD: a prospective study. *Thorax*, 70(12), 1131–7. 52, 155

[149] Kon, S. S., Patel, M. S., Canavan, J. L., Clark, A. L., Jones, S. E., Nolan, C. M., Cullinan, P., Polkey, M. I., & Man, W. D. (2013). Reliability and validity of 4-metre gait speed in COPD. *Eur Respir J*, 42(2), 333–40. 125

[150] Kostianev, S. S., Hodgev, V. A., & Iluchev, D. H. (2008). Multidimensional system for assessment of COPD patients. comparison with BODE index. *Folia Med (Plovdiv)*, 50(4), 29–38. 100

[151] Koutsokera, A., Stolz, D., Loukides, S., & Kostikas, K. (2012). Systemic biomarkers in exacerbations of COPD: the evolving clinical challenge. *CHEST*, 141(2), 396–405. 21, 95

[152] Lacasse, M., Maltais, F., Poirier, P., Lacasse, Y., Marquis, K., Jobin, J., & LeBlanc,

P. (2005). Post-exercise heart rate recovery and mortality in chronic obstructive pulmonary disease. *Respiratory medicine*, 99(7), 877–86. 30, 34

[153] Lahousse, L., Ziere, G., Verlinden, V. J., Zillikens, M. C., Uitterlinden, A. G., Rivadeneira, F., Tiemeier, H., Joos, G. F., Hofman, A., Ikram, M. A., et al. (2016). Risk of frailty in elderly with COPD: a population-based study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 71(5), 689–695. 207

[154] Lange, P., Celli, B., Agusti, A., Boje Jensen, G., Divo, M., Faner, R., Guerra, S., Marott, J. L., Martinez, F. D., Martinez-Camblor, P., Meek, P., Owen, C. A., Petersen, H., Pinto-Plata, V., Schnohr, P., Sood, A., Soriano, J. B., Tesfaigzi, Y., & Vestbo, J. (2015). Lung-function trajectories leading to chronic obstructive pulmonary disease. *N Engl J Med*, 373(2), 111–22. 21

[155] Laupacis, A., Sekar, N., & Stiell, l. (1997). Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA*, 277(6), 488–494. 10

[156] Lee, H. M., Lee, J., Lee, K., Luo, Y. T., Sin, D. D., & Wong, N. D. (2012). Relation between COPD severity and global cardiovascular risk in US adults. *CHEST*, 142(5), 1118–1125. 181

[157] Legrand, D., Vaes, B., Mathei, C., Adriaensen, W., Van Pottelbergh, G., & Degryse, J. M. (2014). Muscle strength and physical performance as predictors of mortality, hospitalization, and disability in the oldest old. *Journal of the American Geriatrics Society*, 62(6), 1030–1038. 99

[158] Leuzzi, G., Galeone, C., Taverna, F., Suatoni, P., Morelli, D., & Pastorino, U. (2017). C-reactive protein level predicts mortality in COPD: a systematic review and meta-analysis. *Eur Respir Rev*, 26(143). 21, 95

[159] Liu, S. F., Wang, C. C., Chin, C. H., Chen, Y. C., & Lin, M. C. (2011). High value of

combined serum C-reactive protein and BODE score for mortality prediction in patients with stable COPD. *Archivos de Bronconeumologia*, 47(9), 427–432. 30, 34, 38

[160] Loeckx, M., Rabinovich, R. A., Demeyer, H., Louvaris, Z., Tanner, R., Rubio, N., Frei, A., De Jong, C., Gimeno-Santos, E., Rodrigues, F. M., Buttery, S. C., Hopkinson, N. S., Busching, G., Strassmann, A., Serra, I., Vogiatzis, I., Garcia-Aymerich, J., Polkey, M. I., & Troosters, T. (2018). Smartphone-based physical activity telecoaching in chronic obstructive pulmonary disease: Mixed-methods study on patient experiences and lessons for implementation. *JMIR Mhealth Uhealth*, 6(12), e200. 204

[161] Lord, S. R. & Menz, H. B. (2002). Physiologic, psychologic, health predictors of 6-minute walk performance in older people. *Archives of Physical Medicine and Rehabilitation*, 83(7), 907–911. 195

[162] Lorenz, M. W., Polak, J. F., Kavousi, M., Mathiesen, E. B., Völzke, H., Tuomainen, T.-P., Sander, D., Plichart, M., Catapano, A. L., Robertson, C. M., Kiechl, S., Rundek, T., Desvarieux, M., Lind, L., Schmid, C., DasMahapatra, P., Gao, L., Ziegelbauer, K., Bots, M. L., & Thompson, S. G. (2012). Carotid intima-media thickness progression to predict cardiovascular events in the general population (the prog-imt collaborative project): a meta-analysis of individual participant data. *The Lancet*, 379(9831), 2053–2062. 181

[163] Luo, W., Nguyen, T., Nichols, M., Tran, T., Rana, S., Gupta, S., Phung, D., Venkatesh, S., & Allender, S. (2015). Is demography destiny? application of machine learning techniques to accurately predict population health outcomes from a minimal demographic dataset. *PLOS One*, 10(5), 1–13. 206

[164] Mackenzie, I. S., Wilkinson, I. B., & Cockcroft, J. R. (2002). Assessment of arterial stiffness in clinical practice. *QJM*, 95(2), 67–74. 197

[165] MacNee, W. (2006). Pathology, pathogenesis, and pathophysiology. *BMJ*, 332(7551), 1202–1204. 3

[166] Maddocks, M., Kon, S. S., Canavan, J. L., Jones, S. E., Nolan, C. M., Labey, A., Polkey, M. I., & Man, W. D. (2016). Physical frailty and pulmonary rehabilitation in COPD: a prospective cohort study. *Thorax*, 71(11), 988–995. 206

[167] Man, S. F., Xing, L., Connett, J. E., Anthonisen, N. R., Wise, R. A., Tashkin, D. P., Zhang, X., Vessey, R., Walker, T. G., Celli, B. R., & Sin, D. D. (2008). Circulating fibronectin to C-reactive protein ratio and mortality: a biomarker in COPD? *Eur Respir J*, 32(6), 1451–7. 30

[168] Mannino, D. M., Homa, D. M., Akinbami, L. J., Ford, E. S., & Redd, S. C. (2002). Chronic obstructive pulmonary disease surveillance–United States, 1971-2000. *Respiratory care*, 47 10, 1184–99. 96

[169] Mannino, D. M., Valvi, D., Mullerova, H., & Tal-Singer, R. (2012). Fibrinogen, COPD and mortality in a nationally representative US cohort. *COPD*, 9(4), 359–66. 30, 33, 34, 38

[170] Marino, D. M., Marrara, K. T., Arcuri, J. F., Candolo, C., Jamami, M., & Di Lorenzo, V. A. P. (2014). Determination of exacerbation predictors in patients with COPD in physical therapy - a longitudinal study. *Brazilian Journal of Physical Therapy*, 18(2), 127–136. 30, 38

[171] Marquis, K., Debigaré, R., Lacasse, Y., LeBlanc, P., Jobin, J., Carrier, G., & Maltais, F. (2002). Midthigh muscle cross-sectional area is a better predictor of mortality than body mass index in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 166(6), 809–813. 14, 21

[172] Mathers, C. D. & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLOS Med*, 3(11), e442. 20, 123, 197

[173] McEniery, C. M., Yasmin, Hall, I. R., Qasem, A., Wilkinson, I. B., & Cockcroft, J. R. (2005). Normal vascular aging: Differential effects on wave reflection and aortic pulse wave velocity. *Journal of the American College of Cardiology*, 46(9), 1753. 13

[174] McGarvey, L. P., John, M., Anderson, J. A., Zvarich, M., & Wise, R. A. (2007). Ascertainment of cause-specific mortality in COPD: operations of the TORCH clinical endpoint committee. *Thorax*, 62(5), 411–415. 21, 123, 129, 149, 181

[175] Mehrotra, N., Freire, A. X., Bauer, D. C., Harris, T. B., Newman, A. B., Kritchevsky, S. B., Meibohm, B., & Health, A. B. C. S. (2010). Predictors of mortality in elderly subjects with obstructive airway disease: the pile score. *Ann Epidemiol*, 20(3), 223–32. 100

[176] Mendoza, L., Horta, P., Espinoza, J., Aguilera, M., Balmaceda, N., Castro, A., Ruiz, M., Díaz, O., & Hopkinson, N. S. (2015). Pedometers to enhance physical activity in COPD: a randomised controlled trial. *European Respiratory Society*, 45(2), 347–354. 204

[177] Metz, C. E. (1978). Basic principles of roc analysis. *Semin Nucl Med*, 8(4), 283–98. 10

[178] Mieno, M. N., Tanaka, N., Arai, T., Kawahara, T., Kuchiba, A., Ishikawa, S., & Sawabe, M. (2016). Accuracy of death certificates and assessment of factors for misclassification of underlying cause of death. *J Epidemiol*, 26(4), 191–8. 129

[179] Miller, B. E., Tal-Singer, R., Rennard, S. I., Furtwaengler, A., Leidy, N., Lowings, M., Martin, U. J., Martin, T. R., Merrill, D. D., Snyder, J., Walsh, J., & Mannino, D. M. (2016). Plasma fibrinogen qualification as a drug development tool in chronic obstructive pulmonary disease perspective of the chronic obstructive pulmonary disease biomarker qualification consortium. *Am J Respir Crit Care Med*, 193(6), 607–613. 12, 155

[180] Miller, J., Edwards, L. D., Agusti, A., Bakke, P., Calverley, P. M. A., Celli, B., Coxson, H. O., Crim, C., Lomas, D. A., Miller, B. E., Rennard, S., Silverman, E. K., Tal-Singer, R., Vestbo, J., Wouters, E., Yates, J. C., Macnee, W., & ECLIPSE Investigators (2013). Comorbidity, systemic inflammation and outcomes in the ECLIPSE cohort. *Respiratory medicine*, 107(9), 1376–84. 3, 30, 53

[181] Mills, N. L., Miller, J. J., Anand, A., Robinson, S. D., Frazer, G. A., Anderson, D., Breen, L., Wilkinson, I. B., McEniery, C. M., Donaldson, K., Newby, D. E., & Macnee, W.

(2008). Increased arterial stiffness in patients with chronic obstructive pulmonary disease: a mechanism for increased cardiovascular risk. *Thorax*, 63(4), 306–11. 13, 52

[182] Mitchell, T. M. (1997). *Machine Learning*. WCB McGraw-Hill. 206

[183] Mohan, D., Forman, J. R., Allinder, M., McEniery, C. M., Bolton, C. E., Cockcroft, J. R., MacNee, W., Fuld, J., Marchong, M., Gale, N. S., Fisk, M., Nagarajan, S., Cheriyan, J., Lomas, D. A., Calverley, P. M. A., Miller, B. E., Tal-Singer, R., Wilkinson, I. B., & Polkey, M. I. (2018). Fibrinogen does not relate to cardiovascular or muscle manifestations in COPD: cross-sectional data from the ERICA study. *Thorax*, 73(12), 1182–1185. 195

[184] Mohan, D., Gale, N. S., McEniery, C. M., Bolton, C. E., Cockcroft, J. R., MacNee, W., Fuld, J., Lomas, D. A., Calverley, P. M. A., Shale, D. J., Miller, B. E., Wilkinson, I. B., Tal-Singer, R., Polkey, M. I., & on behalf of the ERICA Consortium (2014). Evaluating the role of inflammation in chronic airways disease: The ERICA study. *COPD: Journal of Chronic Obstructive Pulmonary Disease*, 11(5), 552–559. 15, 57, 58, 93, 94, 101, 156, 182, 202

[185] Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., Group, P.-P., Altman, D. G., Booth, A., Chan, A.-W., Chang, S., Clarke, M., Clifford, T., Dickersin, K., Egger, M., Ghersi, D., Gotzsche, P. C., Grimshaw, J. M., Groves, T., Helfand, M., Higgins, J., Lasserson, T., Lau, J., Liberati, A., Lohr, K., McGowan, J., Moher, D., Mulrow, C., Norton, M., Page, M., Petticrew, M., Sampson, M., Schunemann, H., Shamseer, L., Shekelle, P., Simera, I., Stewart, L. A., Summerskill, W., Tetzlaff, J., Trikalinos, T. A., Tovey, D., Turner, L., & Whitlock, E. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews*, 4, 1. 22, 26

[186] Monninkhof, E., van der Valk, P., van der Palen, J., van Herwaarden, C., & Zielhuis, G. (2003). Effects of a comprehensive self-management programme in patients with chronic obstructive pulmonary disease. *Eur Respir J*, 22(5), 815–820. 30, 33, 38, 42

[187] Montgomery, P. S. & Gardner, A. W. (1998). The clinical utility of a six-minute walk test in peripheral arterial occlusive disease patients. *J Am Geriatr Soc*, 46(6), 706–11. 14

[188] Moore, A. J., Soler, R. S., Cetti, E. J., Amanda Sathyapala, S., Hopkinson, N. S., Roughton, M., Moxham, J., & Polkey, M. I. (2010). Sniff nasal inspiratory pressure versus IC/TLC ratio as predictors of mortality in COPD. *Respiratory medicine*, 104(9), 1319–25. 30, 50, 99, 124

[189] Moy, M. L., Teylan, M., Danilack, V. A., Gagnon, D. R., & Garshick, E. (2014). An index of daily step count and systemic inflammation predicts clinical outcomes in chronic obstructive pulmonary disease. *Annals of the American Thoracic Society*, 11(2), 149–57. 30, 42

[190] Moy, M. L., Teylan, M., Weston, N. A., Gagnon, D. R., & Garshick, E. (2013). Daily step count predicts acute exacerbations in a US cohort with COPD. *PLOS One*, 8(4). 30

[191] Mullerova, H., Agusti, A., Erqou, S., & Mapel, D. W. (2013). Cardiovascular comorbidity in COPD systematic literature review. *CHEST*, 144(4), 1163–1178. 195

[192] Mullerova, H., Maselli, D. J., Locantore, N., Vestbo, J., Hurst, J. R., Wedzicha, J. A., Bakke, P., Agusti, A., Anzueto, A., Ivanov, Y., Kostov, K., Bourbeau, J., Fitzgerald, M., Hernandez, P., Killian, K., Levy, R., Maltais, F., O'Donnell, D., Krepelka, J., Wouters, E. F. M., Quinn, D., Kosnik, M., Feschenko, Y., Gavrisyuk, V., Yashina, L., MacNee, W., Singh, S. D., Braman, S., Casaburi, R., Celli, B., Giessel, G., Gotfried, M., Greenwald, G., Hanania, N., Mahler, D., Make, B., Rennard, S., Rochester, C., Scanlon, P., Schuller, D., Sciurba, F., Sharafkhaneh, A., Siler, T., Silverman, E., Wanner, A., Wise, R., & Wallack, R. Z. (2015). Hospitalized exacerbations of COPD: Risk factors and outcomes in the ECLIPSE cohort. *CHEST*, 147(4), 999–1007. 31, 32, 42, 48

[193] Murray, C. J., Atkinson, C., Bhalla, K., Birbeck, G., Burstein, R., Chou, D., et al. (2013).

The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. *JAMA*, 310(6), 591–608. 3

[194] Nabel, E. G. (2003). Cardiovascular disease. *N Engl J Med*, 349(1), 60–72. 209

[195] Nambi, V., Chambless, L., Folsom, A. R., He, M., Hu, Y., Mosley, T., Volcik, K., Boerwinkle, E., & Ballantyne, C. M. (2010). Carotid intima-media thickness and presence or absence of plaque improves prediction of coronary heart disease risk: the ARIC (atherosclerosis risk in communities) study. *J Am Coll Cardiol*, 55(15), 1600–7. 13

[196] National Health Services (2017). Chronic obstructive pulmonary disease (COPD) pathway. 7

[197] National Health Services Digital (2016). Hospital admitted patient care activity, 2016-17. all diagnoses: chronic obstructive pulmonary disease. 154

[198] National Health Services Scotland (2013). European standard population 2013 by sex. 132

[199] National Institute for Health and Care Excellence (2013). Guide to the methods of technology appraisal 2013. *NICE*. 211

[200] National Institute for Health and Care Excellence (2018). *Chronic obstructive pulmonary disease in over 16s: diagnosis and management (NG115)*. Technical report, NICE. 99

[201] Newman, A. B., Simonsick, E. M., Naydeck, B. L., Boudreau, R. M., Kritchevsky, S. B., Nevitt, M. C., Pahor, M., Satterfield, S., Brach, J. S., Studenski, S. A., & Harris, T. B. (2006). Association of long-distance corridor walk performance with mortality, cardiovascular disease, mobility limitation, and disability. *JAMA*, 295(17), 2018–2026. 195

[202] Newson, R. B. (2010). Comparing the predictive powers of survival models using Harrell's C or Somers' D. *Stata Journal*, 10(3), 339–358. 10, 102, 185

[203] Norrish, A., North, D., Kirkman, P., & Jackson, R. (1994). Validity of self-reported hospital admission in a prospective study. *Am J Epidemiol*, 140(10), 938–42. 176

[204] Odden, M. C., Beilby, P. R., & Peralta, C. A. (2015). Blood pressure in older adults: the importance of frailty. *Curr Hypertens Rep*, 17(7), 55. 194

[205] Office for National Statistics (2015a). *Expenditure on Healthcare in the UK: 2013*. Report, Office for National Statistics. 5

[206] Office for National Statistics (2015b). National statistics. Number of deaths where the underlying cause was chronic obstructive pulmonary disease by month of occurrence, England and Wales, deaths that occurred between 2001 and 2014. 20

[207] Office for National Statistics (2018). National life tables, UK: 2015 to 2017. trends in the average number of years people will live beyond their current age measured by period life expectancy, analysed by age and sex for the UK and its constituent countries. 94

[208] Ong, K. C., Earnest, A., & Lu, S. J. (2005). A multidimensional grading system (BODE index) as predictor of hospitalization for COPD. *CHEST*, 128(6), 3810–6. 11

[209] Organisation for Economic Co-operation and Development (2017). *Health at a Glance 2017*. OECD. 197

[210] Ozgur, E. S., Nayci, S. A., Ozge, C., & Tasdelen, B. (2012). An integrated index combined by dynamic hyperinflation and exercise capacity in the prediction of morbidity and mortality in COPD. *Respiratory care*, 57(9), 1452–9. 31, 33, 34

[211] Palange, P. & Ward, S. A. (2007). Recommendations on the use of exercise testing in clinical practice. *Eur Respir J*, 29(5), 1066–1067. 9

[212] Parker, D. R., Liu, J., Roberts, M. B., & Eaton, C. B. (2014). Is inflammatory chronic obstructive pulmonary disease a coronary heart disease risk equivalent? A longitudinal analysis of the third national health and nutrition examination survey (NHANES III), 1988-1994. *BMC Pulmonary Medicine*, 14(1), 195. 31

[213] Patel, A. R., Kowlessar, B. S., Donaldson, G. C., Mackay, A. J., Singh, R., George, S. N., Garcha, D. S., Wedzicha, J. A., & Hurst, J. R. (2013). Cardiovascular risk, myocardial injury, and exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 188(9), 1091–9. 52

[214] Patel, M. S., Mohan, D., Andersson, Y. M., Baz, M., Kon, S. S., Canavan, J. L., Jackson, S. G., Clark, A. L., Hopkinson, N. S., Natanek, S. A., Kemp, P. R., Bruijnzeel, P. L., Man, W. D., & Polkey, M. I. (2014a). Phenotypic characteristics associated with reduced short physical performance battery score in COPD. *CHEST*, 145(5), 1016–24. 59

[215] Patel, M. S., Natanek, S. A., Stratakos, G., Pascual, S., Martínez-Llorens, J., Disano, L., Terzis, G., Hopkinson, N. S., Gea, J., Vogiatzis, I., Maltais, F., & Polkey, M. I. (2014b). Vastus lateralis fiber shift is an independent predictor of mortality in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 190(3), 350–352. 21

[216] Pavasini, R., Guralnik, J., Brown, J. C., di Bari, M., Cesari, M., Landi, F., Vaes, B., Legrand, D., Verghese, J., Wang, C., Stenholm, S., Ferrucci, L., Lai, J. C., Bartes, A. A., Espaulella, J., Ferrer, M., Lim, J. Y., Ensrud, K. E., Cawthon, P., Turusheva, A., Frolova, E., Rolland, Y., Lauwers, V., Corsonello, A., Kirk, G. D., Ferrari, R., Volpato, S., & Campo, G. (2016). Short physical performance battery and all-cause mortality: systematic review and meta-analysis. *BMC Med*, 14(1), 215. 52, 58, 101, 155

[217] Pepe, M. S., Fan, J., Feng, Z., Gerds, T., & Hilden, J. (2015). The net reclassification index (NRI): a misleading measure of prediction improvement even with independent test data sets. *Stat Biosci*, 7(2), 282–295. 208

[218] Petty, T. L. (2002). COPD in perspective. *CHEST*, 121(5 Suppl), 116S–120S. 5

[219] Pinto-Plata, V., Casanova, C., Mullerova, H., de Torres, J. P., Corado, H., Varo, N., Cordoba, E., Zeineldine, S., Paz, H., Baz, R., Divo, M., Cortopassi, F., & Celli, B. R. (2012).

Inflammatory and repair serum biomarker pattern: association to clinical outcomes in COPD. *Respiratory Research*, 13, 71. 31

[220] Pinto-Plata, V. M., Cote, C., Cabral, H., Taylor, J., & Celli, B. R. (2004). The 6-min walk distance: change over time and value as a predictor of survival in severe COPD. *European Respiratory Society*, 23(1), 28–33. 155

[221] Pitta, F., Troosters, T., Probst, V. S., Spruit, M. A., Decramer, M., & Gosselink, R. (2006). Physical activity and hospitalization for exacerbation of COPD. *CHEST*, 129(3), 536–44. 13

[222] Polkey, M. I., Kyroussis, D., Hamnegard, C. H., Mills, G. H., Green, M., & Moxham, J. (1996). Diaphragm strength in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 154(5), 1310–7. 124

[223] Powrie, D. J., Wilkinson, T. M., Donaldson, G. C., Jones, P., Scrine, K., Viel, K., Kesten, S., & Wedzicha, J. A. (2007). Effect of tiotropium on sputum and serum inflammatory markers and exacerbations in COPD. *Eur Respir J*, 30(3), 472–8. 31

[224] Public Health England (2017). Health profile for England 2017. Chapter 2: major causes of death and how they have changed. 151

[225] Puhan, M. A., Garcia-Aymerich, J., Frey, M., ter Riet, G., Anto, J. M., Agusti, A. G., Gomez, F. P., Rodriguez-Roisin, R., Moons, K. G., Kessels, A. G., & Held, U. (2009). Expansion of the prognostic assessment of patients with chronic obstructive pulmonary disease: the updated BODE index and the ADO index. *Lancet*, 374(9691), 704–11. 11, 100, 117, 122, 123

[226] Puhan, M. A., Gimeno-Santos, E., Scharplatz, M., Troosters, T., Walters, E. H., & Steurer, J. (2011). Pulmonary rehabilitation following exacerbations of chronic obstructive pulmonary disease. *Cochrane Database Syst Rev*, 1(10), CD005305. 8, 20

[227] Quint, J., Donaldson, G., Hurst, J., Goldring, J., Seemungal, T., & Wedzicha, J. (2011). Predictive accuracy of patient-reported exacerbation frequency in COPD. *Eur Respir J*, 37(3), 501–507. 94

[228] Raymakers, A. J. N., Sadatsafavi, M., Sin, D. D., De Vera, M. A., & Lynd, L. D. (2017). The impact of statin drug use on all-cause mortality in patients with COPD a population-based cohort study. *CHEST*, 152(3), 486–493. 150

[229] Riley, D. J., Thakker-Varia, S., Wilson, F. J., Poiani, G. J., & Tozzi, C. A. (2000). Role of proteolysis and apoptosis in regression of pulmonary vascular remodeling. *Physiol Res*, 49(5), 577–85. 3

[230] Rothnie, K. J., Mullerova, H., Thomas, S. L., Chandan, J. S., Smeeth, L., Hurst, J. R., Davis, K., & Quint, J. K. (2016). Recording of hospitalizations for acute exacerbations of COPD in UK electronic health care records. *Clin Epidemiol*, 8, 771–782. 157

[231] Rothnie, K. J., Yan, R., Smeeth, L., & Quint, J. K. (2015). Risk of myocardial infarction (MI) and death following mi in people with chronic obstructive pulmonary disease (COPD): a systematic review and meta-analysis. *BMJ Open*, 5(9), e007824. 129

[232] Royston, P., Moons, K. G., Altman, D. G., & Vergouwe, Y. (2009). Prognosis and prognostic research: developing a prognostic model. *BMJ*, 338, b604. 101

[233] Salpeter, S. R., Ormiston, T. M., & Salpeter, E. E. (2004). Cardiovascular effects of beta-agonists in patients with asthma and COPD: a meta-analysis. *CHEST*, 125(6), 2309–21. 150

[234] Santibanez, M., Garrastazu, R., Ruiz-Nunez, M., Helguera, J. M., Arenal, S., Bonnardeux, C., Leon, C., & Garcia-Rivero, J. L. (2016). Predictors of hospitalized exacerbations and mortality in chronic obstructive pulmonary disease. *PLOS One*, 11(6), e0158727. 155

[235] Seemungal, T., Harper-Owen, R., Bhowmik, A., Moric, I., Sanderson, G., Message, S., MacCallum, P., Meade, T. W., Jeffries, D. J., Johnston, S. L., et al. (2001). Respiratory viruses, symptoms, and inflammatory markers in acute exacerbations and stable chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 164(9), 1618–1623. 31

[236] Seidl, H., Meisinger, C., Kirchberger, I., Burkhardt, K., Kuch, B., & Holle, R. (2016). Validity of self-reported hospital admissions in clinical trials depends on recall period length and individual characteristics. *J Eval Clin Pract*, 22(3), 446–54. 176

[237] Seymour, J. M., Spruit, M. A., Hopkinson, N. S., Natanek, S. A., Man, W. D., Jackson, A., Gosker, H. R., Schols, A. M., Moxham, J., Polkey, M. I., & Wouters, E. F. (2010). The prevalence of quadriceps weakness in COPD and the relationship with disease severity. *Eur Respir J*, 36(1), 81–8. 14, 207

[238] Similowski, T., Yan, S., Gauthier, A. P., Macklem, P. T., & Bellemare, F. (1991). Contractile properties of the human diaphragm during chronic hyperinflation. *N Engl J Med*, 325(13), 917–23. 124

[239] Simpson, C. R., Hippisley-Cox, J., & Sheikh, A. (2010). Trends in the epidemiology of chronic obstructive pulmonary disease in England: a national study of 51 804 patients. *Br J Gen Pract*, 60(576), 277–84. 5

[240] Sin, D. D. & Man, S. F. P. (2005). Chronic obstructive pulmonary disease as a risk factor for cardiovascular morbidity and mortality. *Proceedings of the American Thoracic Society*, 2(1), 8–11. 195

[241] Sin, D. D. & Man, S. P. (2003). Why are patients with chronic obstructive pulmonary disease at increased risk of cardiovascular diseases? *Circulation*, 107(11), 1514–1519. 3

[242] Singh, D., Edwards, L., Tal-Singer, R., & Rennard, S. (2010). Sputum neutrophils as a biomarker in COPD: Findings from the ECLIPSE study. *Respiratory Research*, 11(77). 31

[243] Singh, S., Loke, Y. K., & Furberg, C. D. (2008). Inhaled anticholinergics and risk of major adverse cardiovascular events in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis. *JAMA*, 300(12), 1439–50. 150

[244] Smith, C., Hill, S., & Amos, A. (2018). *Stop Smoking Inequalities: A systematic review of socioeconomic inequalities in experiences of smoking cessation interventions in the UK.* Report, University of Edinburgh. 8

[245] Smith, G. C., Seaman, S. R., Wood, A. M., Royston, P., & White, I. R. (2014). Correcting for optimistic prediction in small data sets. *Am J Epidemiol*, 180(3), 318–24. 102, 185, 196

[246] Soler-Cataluna, J. J., Martinez-Garcia, M. A., Sanchez, L. S., Tordera, M. P., & Sanchez, P. R. (2009). Severe exacerbations and BODE index: two independent risk factors for death in male COPD patients. *Respir Med*, 103(5), 692–9. 100, 124

[247] Song, S., Yang, P. S., Kim, T. H., Uhm, J. S., Pak, H. N., Lee, M. H., & Joung, B. (2017). Relation of chronic obstructive pulmonary disease to cardiovascular disease in the general population. *Am J Cardiol*, 120(8), 1399–1404. 129

[248] Spruit, M. A., Polkey, M. I., Celli, B., Edwards, L. D., Watkins, M. L., Pinto-Plata, V., Vestbo, J., Calverley, P. M. A., Tal-Singer, R., Agusti, A., Coxson, H. O., Lomas, D. A., MacNee, W., Rennard, S., Silverman, E. K., Crim, C. C., Yates, J., & Wouters, E. F. M. (2012). Predicting outcomes from 6-minute walk distance in chronic obstructive pulmonary disease. *Journal of the American Medical Directors Association*, 13(3), 291–297. 31, 32, 33, 99, 155

[249] Stefan, M. S., Bannuru, R. R., Lessard, D., Gore, J. M., Lindenauer, P. K., & Goldberg, R. J. (2012). The impact of COPD on management and outcomes of patients hospitalized with acute myocardial infarction: a 10-year retrospective observational study. *CHEST*, 141(6), 1441–1448. 150

[250] Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, 339. 197

[251] Steyerberg, E. W. (2008). *Clinical prediction models: a practical approach to development, validation, and updating*. Springer Science and Business Media. 101

[252] Stolz, D., Kostikas, K., Blasi, F., Boersma, W., Milenkovic, B., Lacoma, A., Louis, R., Aerts, J. G., Welte, T., Torres, A., Rohde, G. G., Boeck, L., Rakic, J., Scherr, A., Hertel, S., Giersdorf, S., & Tamm, M. (2014). Adrenomedullin refines mortality prediction by the BODE index in COPD: the "BODE-A" index. *Eur Respir J*, 43(2), 397–408. 100

[253] Stone, I. S., Barnes, N. C., & Petersen, S. E. (2012). Chronic obstructive pulmonary disease: a modifiable risk factor for cardiovascular disease? *Heart*, 98(14), 1055–62. 3

[254] Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., Moher, D., Becker, B. J., Sipe, T. A., & Thacker, S. B. (2000). Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA*, 283(15), 2008–12. 22

[255] Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., & Collins, R. (2015). UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Med*, 12(3), e1001779. 131

[256] Suissa, S. (2014). Will SUMMIT reach the peak in COPD? *Thorax*, 69(5), 405–7. 123

[257] Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., Burgess, S., Jiang, T., Paige, E., Surendran, P., Oliver-Williams, C., Kamat, M. A., Prins, B. P., Wilcox, S. K., Zimmerman, E. S., Chi, A., Bansal, N., Spain, S. L., Wood, A. M.,

Morrell, N. W., Bradley, J. R., Janjic, N., Roberts, D. J., Ouwehand, W. H., Todd, J. A., Soranzo, N., Suhre, K., Paul, D. S., Fox, C. S., Plenge, R. M., Danesh, J., Runz, H., & Butterworth, A. S. (2017). Consequences of natural perturbations in the human plasma proteome. *bioRxiv*. 209

[258] Sutton-Tyrrell, K., Najjar, S. S., Boudreau, R. M., Venkitachalam, L., Kupelian, V., Simonsick, E. M., Havlik, R., Lakatta, E. G., Spurgeon, H., Kritchevsky, S., Pahor, M., Bauer, D., & Newman, A. (2005). Elevated aortic pulse wave velocity, a marker of arterial stiffness, predicts cardiovascular events in well-functioning older adults. *Circulation*, 111(25), 3384–3390. 181

[259] Swallow, E. B., Reyes, D., Hopkinson, N. S., Man, W. D., Porcher, R., Cetti, E. J., Moore, A. J., Moxham, J., & Polkey, M. I. (2007). Quadriceps strength predicts mortality in patients with moderate to severe chronic obstructive pulmonary disease. *Thorax*, 62(2), 115–20. 21, 31, 50, 99, 155

[260] Thomsen, M., Ingebrigtsen, T. S., Marott, J. L., Dahl, M., Lange, P., Vestbo, J., & Nordestgaard, B. G. (2013). Inflammatory biomarkers and exacerbations in chronic obstructive pulmonary disease. *JAMA*, 309(22), 2353–61. 31, 38, 48, 155

[261] Tonnesen, P. (2013). Smoking cessation and COPD. *Eur Respir Rev*, 22(127), 37–43. 8, 20

[262] Touboul, P. J., Hennerici, M. G., Meairs, S., Adams, H., Amarenco, P., Bornstein, N., Csiba, L., Desvarieux, M., Ebrahim, S., Fatar, M., Hernandez Hernandez, R., Jaff, M., Kownator, S., Prati, P., Rundek, T., Sitzer, M., Schminke, U., Tardif, J. C., Taylor, A., Vicaut, E., Woo, K. S., Zannad, F., & Zureik, M. (2007). Mannheim carotid intima-media thickness consensus (2004-2006). An update on behalf of the advisory board of the 3rd and 4th watching the risk symposium, 13th and 15th european stroke conferences, Mannheim, Germany, 2004, and Brussels, Belgium, 2006. *Cerebrovasc Dis*, 23(1), 75–80. 13

[263] Troosters, T., Bourbeau, J., Maltais, F., Leidy, N., Erzen, D., De Sousa, D., Korducki, L., & Hamilton, A. (2016). Enhancing exercise tolerance and physical activity in copd with combined pharmacological and non-pharmacological interventions: Physacto randomised, placebo-controlled study design. *BMJ Open*, 6(4). 204

[264] Usher-Smith, J. A., Silarova, B., Sharp, S. J., Mills, K., & Griffin, S. J. (2018). Effect of interventions incorporating personalised cancer risk information on intentions and behaviour: a systematic review and meta-analysis of randomised controlled trials. *BMJ Open*, 8(1). 210

[265] Valvi, D., Mannino, D. M., Mullerova, H., & Tal-Singer, R. (2012). Fibrinogen, chronic obstructive pulmonary disease (COPD) and outcomes in two United States cohorts. *International Journal of COPD*, 7, 173–82. 31, 34, 38

[266] Van Belle, G. (2011). *Statistical rules of thumb*, volume 699. John Wiley and Sons. 60, 160

[267] van Dijk, W. D., van den Bemt, L., van den Haak-Rongen, S., Bischoff, E., van Weel, C., in 't Veen, J. C. C. M., & Schermer, T. R. J. (2011). Multidimensional prognostic indices for use in COPD patient care. A systematic review. *Respiratory Research*, 12(1), 151–151. 124, 210

[268] van Giessen, A., Peters, J., Wilcher, B., Hyde, C., Moons, C., de Wit, A., & Koffijberg, E. (2017). Systematic review of health economic impact evaluations of risk prediction models: Stop developing, start evaluating. *Value in health*, 20(4), 718–726. 211

[269] Vass, C. M. & Payne, K. (2017). Using discrete choice experiments to inform the benefit-risk assessment of medicines: Are we ready yet? *PharmacoEconomics*, 35(9), 859–866. 208

[270] Vestbo, J., Anderson, W., Coxson, H. O., Crim, C., Dawber, F., Edwards, L., Hagan, G., Knobil, K., Lomas, D. A., MacNee, W., Silverman, E. K., Tal-Singer, R., & ECLIPSE Investigators (2008). Evaluation of COPD longitudinally to identify predictive surrogate endpoints (ECLIPSE). *Eur Respir J*, 31(4), 869–73. 15, 56, 96, 176

[271] Vickers, A. J. & Elkin, E. B. (2006). Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*, 26(6), 565–74. 205

[272] Vivodtzev, I., Tamisier, R., Baguet, J. P., Borel, J. C., Levy, P., & Pepin, J. L. (2014). Arterial stiffness in COPD. *CHEST*, 145(4), 861–75. 3, 4, 21

[273] Vlachopoulos, C., Aznaouridis, K., & Stefanadis, C. (2010). Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-analysis. *J Am Coll Cardiol*, 55(13), 1318–27. 195, 196

[274] Volpato, S., Cavalieri, M., Sioulis, F., Guerra, G., Maraldi, C., Zuliani, G., Fellin, R., & Guralnik, J. M. (2011). Predictive value of the short physical performance battery following hospitalization in older patients. *J Gerontol A Biol Sci Med Sci*, 66(1), 89–96. 99

[275] von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., Vandenbroucke, J. P., & Strobe Initiative (2008). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*, 61(4), 344–9. 103, 156, 186

[276] Walter, R. E., Wilk, J. B., Larson, M. G., Vasan, R. S., Keaney, John F., J., Lipinska, I., O'Connor, G. T., & Benjamin, E. J. (2008). Systemic inflammation and COPD: the Framingham Heart Study. *CHEST*, 133(1), 19–25. 181

[277] Wang, G., Liu, B., Cao, Y., Du, Y., Zhang, H., Luo, Q., Li, B., Wu, J., Lv, Y., Sun, J., Jin, H., Wei, K., Zhao, Z., Kong, L., Zhou, X., Miao, Q., Wang, G., Zhou, Q., & Dong, J. (2014a). Effects of two chinese herbal formulae for the treatment of moderate to severe stable chronic obstructive pulmonary disease: a multicenter, double-blind, randomized controlled trial. *PLOS One*, 9(8), 1–12. 31

[278] Wang, M., Li, J., Li, S., & Xie, Y. (2014b). Effects of comprehensive therapy based on traditional chinese medicine patterns on older patients with chronic obstructive pulmonary

disease: a subgroup analysis from a four-center, randomized, controlled study. *Fronteras en Medicina*, 8(3), 368–75. 31

[279] Wannamethee, S. G., Shaper, A. G., Lennon, L., & Morris, R. W. (2005). Metabolic syndrome vs Framingham risk score for prediction of coronary heart disease, stroke, and type 2 diabetes mellitus. *Arch Intern Med*, 165(22), 2644–50. 181

[280] Warnier, M. J., Rutten, F. H., de Boer, A., Hoes, A. W., & De Bruin, M. L. (2014). Resting heart rate is a risk factor for mortality in chronic obstructive pulmonary disease, but not for exacerbations or pneumonia. *PLOS One*, 9(8). 31, 34

[281] Waschki, B., Kirsten, A., Holz, O., Muller, K. C., Meyer, T., Watz, H., & Magnussen, H. (2011). Physical activity is the strongest predictor of all-cause mortality in patients with COPD: a prospective cohort study. *CHEST*, 140(2), 331–342. 31, 33, 34, 42, 48, 155

[282] Wedzicha, J. A., Seemungal, T. A., MacCallum, P. K., Paul, E. A., Donaldson, G. C., Bhowmik, A., Jeffries, D. J., & Meade, T. W. (2000). Acute exacerbations of chronic obstructive pulmonary disease are accompanied by elevations of plasma fibrinogen and serum IL-6 levels. *Thrombosis and haemostasis*, 84(2), 210–5. 31, 38, 48

[283] Weiner, D. E., Tighiouart, H., Elsayed, E. F., Griffith, J. L., Salem, D. N., Levey, A. S., & Sarnak, M. J. (2007). The Framingham predictive instrument in chronic kidney disease. *Journal of the American College of Cardiology*, 50(3), 217–224. 181

[284] Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLOS One*, 12(4), 1–14. 206

[285] Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M., Sterne, J. A., Bossuyt, P. M., & QUADAS-Group (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*, 155(8), 529–36. 24

[286] Wilkinson, I. B., Fuchs, S. A., Jansen, I. M., Spratt, J. C., Murray, G. D., Cockcroft, J. R., & Webb, D. J. (1998). Reproducibility of pulse wave velocity and augmentation index measured by pulse wave analysis. *J Hypertens*, 16(12 Pt 2), 2079–84. 12, 58, 183

[287] Wood, A. M., Kaptoge, S., Butterworth, A. S., Willeit, P., Warnakula, S., Bolton, T., Paige, E., Paul, D. S., Sweeting, M., Burgess, S., Bell, S., Astle, W., Stevens, D., Koulman, A., Selmer, R. M., Verschuren, W. M. M., Sato, S., Njolstad, I., Woodward, M., Salomaa, V., Nordestgaard., et al. (2018). Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies. *Lancet*, 391(10129), 1513–1523. 210

[288] Wordsworth, S., Doble, B., Payne, K., Buchanan, J., Marshall, D. A., McCabe, C., & Regier, D. A. (2018). Using "big data" in the cost-effectiveness analysis of next-generation sequencing technologies: Challenges and potential solutions. *Value Health*, 21(9), 1048–1053. 211

[289] World Health Organisation (2014). The top 10 causes of death. 20

[290] World Health Organisation (2015). Mortality indicator database. Mortality indicators by 67 causes of death, age and sex (Health for All-European Mortality Database). 5

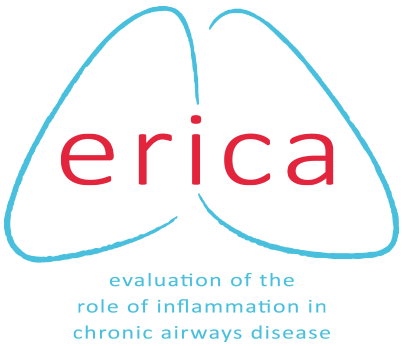
[291] World Health Organisation (2016). The international classification of adult underweight overweight and obesity according to BMI. 58, 157

[292] Wust, R. C. & Degens, H. (2007). Factors contributing to muscle wasting and dysfunction in COPD patients. *Int J Chron Obstruct Pulmon Dis*, 2(3), 289–300. 14

[293] Young, R. P. & Hopkins, R. J. (2013). Update on the potential role of statins in chronic obstructive pulmonary disease and its co-morbidities. *Expert Review of Respiratory Medicine*, 7(5), 533–544. 4

[294] Young, R. P. & J. Hopkins, R. (2014). Primary and secondary prevention of chronic obstructive pulmonary disease: Where to next? *Am J Respir Crit Care Med*, 190(7), 839–840. 3

[295] Zhang, W., Zhang, Y., Li, C. W., Jones, P., Wang, C., & Fan, Y. (2017). Effect of statins on COPD: a meta-analysis of randomized controlled trials. *CHEST*, 152(6), 1159–1168. 150


erica
evaluation of the
role of inflammation in
chronic airways disease