LONDON'S GLOBAL UNIVERSITY

Could linked health and council data advance our understanding of the determinants of multimorbidity and inform service provision? A mixed methods study

Elizabeth May Ingram

UCL

A thesis submitted for the degree of Doctor of Philosophy

Declaration

I, Elizabeth Ingram confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

This PhD was funded by the National Institute for Health Research (NIHR) School for Public Health Research (SPHR) and the NIHR Applied Research Collaboration (ARC) North Thames.

I would like to thank my primary supervisor Dr Jessica Sheringham for her exceptional advice and expertise throughout this PhD. I am particularly grateful for her unwavering support over the last 18 months, whilst I conducted and completed this PhD remotely. I would also like to thank the rest of my supervisory panel, Dr Manuel Gomes, Dr Helen McDonald, Professor David Osborn, and Sue Hogarth for their constructive and invaluable guidance. I am indebted to colleagues at Care City and North Central London for their support with this piece of work. I would also like to thank the Care City Community board and the NIHR ARC North Thames patient and public involvement panel for their advice and input. I am grateful to my study participants who will remain anonymous to ensure confidentiality.

Thank you to my wonderful colleagues, friends, and family. I would particularly like to thank Alice Ingram, Olivia Hadjinicolaou, Irina Lut, Sarah Beardon, Dr Silvie Cooper, Tishya Venkatraman, my nana, and my parents for their unconditional support. And last, but by no means least, I am enormously grateful to Luca Tiratelli for his faith in me, for feeding me, and for proofreading this thesis.

Abstract

Background: Multimorbidity – the co-occurrence of multiple chronic conditions within a single individual – is a major public health problem influenced by social determinants of health. To meet challenges such as multimorbidity, whose causes and management transcend organisational boundaries, some areas have linked their administrative health and council records. This mixed-methods thesis aimed to investigate how knowledge from the analysis of linked health and council data ('analytics') could advance understanding of the determinants of multimorbidity (Aim 1) and inform the equitable provision of services for groups such as those with, or at risk of, multimorbidity (Aim 2).

Methods: Findings from my systematic review of literature examining household and area-level social determinants of multimorbidity informed a quantitative study. Using multi-level logistic regression, I analysed a linked health and council dataset to quantify associations between household tenure and multimorbidity amongst working age residents of Barking and Dagenham. Semi-structured interviews were conducted with 20 senior leaders of North London health and care organisations to explore barriers and facilitators of analytics use for strategic and equitable health and care decision-making.

Results: The review found that household-level social determinants of multimorbidity are often overlooked despite large effect sizes for household compared to area-level determinants. The quantitative analysis found that risk of multimorbidity was greater for social housing tenants and lower for private renters when both were compared to owner-occupiers. Interview findings indicated that leaders did not uniformly use this type of knowledge generated from analytics to inform decision-making due to barriers spanning their working environments, relationships, and data quality.

Conclusions: Linked health and council data can provide novel population health insights for local concerns like multimorbidity. However, improved data linkage alone will unlikely influence the use of these insights for more equitable service provision without efforts to address further barriers to analytics access and use.

Impact Statement

My research has provided a use case for creating and analysing linked health and council data to advance understanding of household-level social determinants of multimorbidity and generate knowledge that could be used to inform equitable service provision for groups such as those with, or at risk of, multimorbidity. However, this thesis challenges some of the policy assumptions behind the creation of such linked data, namely that knowledge generated from linked data *will* improve decision-making, care, and the equity of services. My research identified considerable and complex barriers senior leaders' face when trying to use such knowledge to inform decision-making across organisational boundaries. These findings have implications for research, practice, and policy.

Firstly, my systematic review and quantitative analysis of a linked health and council dataset illustrate the importance yet under exploration of household-level social determinants of multimorbidity. My quantitative analysis has generated new findings and highlights the strength of household tenure as an exposure for understanding household-level inequalities in multimorbidity. For other researchers, this work demonstrates the importance of investigating household-level variation in multimorbidity. Local system leaders have expressed interest in these findings to better understand where to target health prevention and promotion resources. Local leaders have also expressed an interest in this work more broadly, as these findings present a case study for how such linked data can be used to widen understanding of population health. Other local areas looking to link their administrative health and care records could be interested in this work as an example of the type of insights possible to generate when such data are linked.

My qualitative interviews have generated novel findings and illustrate barriers and facilitators leaders face when trying to use knowledge generated from administrative data to inform strategic and equitable health and care decision-making for groups such as those with, or at risk of, multimorbidity. Locally, system leaders are using my qualitative findings to inform their analytics strategies and to inform wider work aiming to improve analytics use for decision-making. These findings and the proposed recommendations may also have implications for national policy and for further local

areas looking to improve their analytical capacity. Recent UK policy often assumes that data linkage *will* improve decision-making, care, and the equity of health and care services. My findings suggest data linkage alone will be insufficient to realise this aspiration without strategies to address further organisational and relational barriers to analytics use. For those in practice and for other researchers, my research presents fruitful areas where further work is needed. Further research could identify ways to successfully address these wider barriers to analytics use.

At the time of submission, one peer-reviewed paper has been published, a second paper has been submitted for publication in a peer-reviewed journal, and a third paper is in preparation. During the course of this thesis, I have presented findings at various national conferences, in local system meetings, and in patient and public involvement meetings. My findings have also been disseminated via online blogs and other research forums.

Table of contents

List of Tables

List of Figures

Abbreviations

AIC	Akaike's Information Criteria
AMS	Academy of Medical Sciences
ARC	Applied Research Collaboration
BHR	Barking and Dagenham, Havering, and Redbridge
BMI	Body Mass Index
CCG	Clinical Commissioning Group
CKD	Chronic Kidney Disease
CLAHRC	Collaboration for Applied Leadership in Health Research and Care
COPD	Chronic Obstructive Pulmonary Disease
CSDH	Commission on SDoH
EHRs	Electronic Health Records
EHS	English Housing Survey
ESA	Employment Support and Allowance
GLA	Greater London Authority
GP	General practice
HIC	High-income country
HIV	Human immunodeficiency virus

ICP	Integrated Care Pilot
ICS	Integrated Care System
IMD	Index of Multiple Deprivation
JSA	Job Seeker's Allowance
LBBD	London Borough of Barking and Dagenham
LMIC	Low-middle income country
LSOA	Lower Super Output Area
NCD	Non-communicable disease
NCL	North Central London
NHS	National Health Service
NHS FYFV	NHS Five Year Forward View
NICE	National Institute for Health and Care Excellence
NIHR	National Institute of Health Research
ONS	Office for National Statistics
OR	Odds ratio
PPI	Patient and public involvement
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
SDoH	Social determinants of health
SPHR	School for Public Health Research

STP	Sustainability and Transformation Partnership
UK	United Kingdom
UPRN	Unique Property Reference Number
VPC	Variance partition coefficient
WHO	World Health Organization
WMC	Wales Multimorbidity e-Cohort

Chapter 1 Introduction

Multimorbidity – broadly defined as the co-occurrence of multiple long-term or chronic conditions within a single individual – is a major public health problem. The nature and extent of multimorbidity is influenced by social factors widely referred to as social determinants of health (SDoH). Efforts to prevent and treat multimorbidity, therefore, may need to address SDoH. To facilitate the formation of more integrated health and care in England, a handful of areas across the country have joined up their administrative health and council records. These novel linked datasets present opportunities to better understand the social determinants of multimorbidity. There are also opportunities for using knowledge generated from the analysis of linked data to improve the equity of efforts to prevent and manage multimorbidity.

This mixed-methods thesis explores whether linked health and council data could advance our understanding of the social determinants of multimorbidity and inform the provision of more equitable health and care services. In this chapter, I introduce key terminology and relevant literature to give an overview of the problem of multimorbidity, summarise the role of SDoH and describe the creation and potential of novel linked health and council datasets. In Chapter 2, I describe the aims and objectives for this thesis and give an outline of the remaining chapters in this thesis.

1.1 Epidemiological changes and chronic disease

Globally, countries are facing dramatic epidemiological transitions due to advances in medicine and improvements in living standards (Bunker, 2001). These changes are enabling populations to both grow and age. According to the Office for National Statistics (ONS), the total population in England increased by approximately 13.7% between mid-2000 and mid-2018 (Office of National Statistics, 2020a). Looking forward, ONS project that the proportion of people in England aged 65 years and over will increase from 18.2% in mid-2018, to an estimated 20.7% by mid-2028 (Office of National Statistics, 2020b). This equates to nearly two million additional people in this age bracket.

As we age, we are more likely to develop a chronic (or long-term) condition (Nolte and McKee, 2008). Medical advances, that have increased longevity and population growth by allowing us to *treat* rather than *cure* disease, are therefore enabling more people to live longer with chronic conditions. As such, the prevalence of many chronic conditions are rising (Nolte and McKee, 2008). This has led to a global shift in the conditions responsible for the majority of disease burden. Worldwide, non-communicable, chronic conditions such as stroke, diabetes and depression have replaced infectious diseases as the lead drivers of disease burden (Vos *et al.*, 2020). In England alone, approximately 15 million people were estimated to be living with a chronic condition in 2012 (Department of Health, 2012) and an estimated 60% of people aged 65 years and over reported having one of more chronic diseases between mid-2016 and mid-2017 (European Commission, 2019).

Chronic conditions are generally defined as "health problems that require ongoing management over a period of years or decades" (World Health Organization, 2002). As a result, individuals with chronic conditions typically exhibit long-term, fluctuating needs (Lawton *et al.*, 2005; Morris *et al.*, 2011). These present significant challenges to public health and health and care systems (Nolte and McKee, 2008). In many countries, healthcare systems have evolved to primarily treat discrete episodes of disease in isolation (Nolte and McKee, 2008). As such, many healthcare systems are ill-equipped to tackle the needs of those with chronic conditions as meeting these needs can require continual monitoring from different providers and settings (Nolte and McKee, 2008). The cost of delivering this care places huge strain on health and care resources. In England, care for individuals with chronic conditions was estimated to account for approximately £7 in every £10 of total health and social care expenditure in 2012 (Department of Health, 2012).

1.2 Multimorbidity as a growing public health challenge

> *"Multimorbidity has emerged as one of the greatest challenges facing health services, both presently and in the coming decades".* (Pearson-Stuttard, Ezzati and Gregg, 2019)

With age, individuals are not only more likely to develop single chronic conditions but accrue multiple chronic conditions over time. This state of health - widely referred to as 'multimorbidity' – is a considerable and growing public health challenge.

1.2.1 Defining multimorbidity

There is no universally accepted definition of multimorbidity, with one review identifying over 130 different definitions operationalised in the literature prior to 2013 (Le Reste *et al.*, 2013). This is compounded by the fact that different terms such as 'multiple long-term conditions' or 'multiple morbidity' are used to refer to the same concept. Despite this, it is generally agreed that multimorbidity is broadly defined as the co-occurrence of several (i.e., multi) long-term or chronic conditions (i.e., morbidity) within a single individual (Mercer, Salisbury and Fortin, 2014).

The term 'multimorbidity' is also used to indicate different concepts, however, it is important to note that multimorbidity is distinctly different from a concept coined by Feinstein in 1970, that of 'comorbidity'. Comorbidity was defined by Feinstein as "any distinct additional clinical entity that has existed or that may occur during the clinical course of a patient who has an index disease under study" (Feinstein, 1970). Unlike for comorbidity, the concept of multimorbidity does not give focus or precedence to a single condition (van den Akker, Buntinx and Knottnerus, 1996; Nicholson, Makovski, *et al.*, 2019). This key difference is reflected in one of the most widely cited definitions of multimorbidity: "the coexistence of two or more chronic conditions, where one is not necessarily more central than others" (Boyd and Fortin, 2010).

To reach an operational definition of multimorbidity, two steps typically occur. First, a decision is made about the minimum number of conditions required for an individual to be classified as having multimorbidity. Second, the conditions to be included on this list are selected. Given that Boyd and Fortin's definition is the most commonly used in the literature, studies most often require an individual to have a minimum number of two conditions to be classified as having multimorbidity. However, some researchers use cut-offs of three, four or more conditions and a number of studies fail to report the threshold used (Ho *et al.*, 2021). There is also considerable debate around the conditions that should be included in a definition of multimorbidity. Since 2002, 'chronic' conditions have been typically defined as non-communicable conditions that require continual management over a period of 12 months of more, in accordance with the World Health Organization's (WHO) definition (World Health Organization, 2002; Mercer, Salisbury and Fortin, 2014). However, debate remains around what conditions classify as 'chronic'. Many argue that definitions of multimorbidity should also include acute conditions and infectious diseases like

human immunodeficiency virus (HIV) (Nolte and McKee, 2008). Some also argue that risk factors for chronic diseases, such as obesity, are important to include in definitions that are employed in clinical practice (Willadsen *et al.*, 2016). In some cases, researchers have extended multimorbidity definitions even further to incorporate social problems associated with chronic conditions. For example, Watt's definition of multimorbidity is "the number, severity and complexity of health and social problems within families" (Watt, 2008). This debate has led to huge variation in the number and type of chronic conditions included in a multimorbidity definition, with certain conditions such individual mental health conditions, haematological conditions and skin conditions under-represented (Ho *et al.*, 2021). In some cases, the number and type of included conditions are simply not reported (Ho *et al.*, 2021). This variation can, unsurprisingly, impact study results (Fortin *et al.*, 2012).

In an attempt to address the lack of consensus when defining multimorbidity, The Academy of Medical Sciences (AMS) have released a definition that is more comprehensive than Boyd and Fortin's definition but consistent with that adopted by the WHO (The Academy of Medical Sciences, 2018):

The coexistence of two or more chronic conditions, each one of which is either:

- A physical non-communicable disease of long duration, such as a cardiovascular disease or cancer.
- A mental health condition of long duration, such as a mood disorder or dementia.
- An infectious disease of long duration, such as HIV or hepatitis C.

This definition has been adopted by the National Institute of Health Research (NIHR) (NIHR, 2021). However, many continue to argue that multimorbidity definitions like those adopted by the AMS are still too heterogeneous to be useful in research and continue to not adequately capture the complexity of the problem clinically (Ford and Ford, 2018). For example, within this definition, an individual with asthma and chronic obstructive pulmonary disease (COPD) will be grouped with someone who may have asthma, COPD, chronic pain, frailty, and depression. The latter individual will likely have more complex health needs and poorer outcomes. As such, slightly expanded versions of this definition that capture the implications for those living with

multimorbidity have been employed by others. For example, The National Institute for Health and Care Excellence (NICE) defines multimorbidity as the presence of two or more long-term health conditions, which can include:

- Defined physical or mental health conditions, such as diabetes or schizophrenia.
- Ongoing conditions, such as learning disability.
- Symptom complexes, such as frailty or chronic pain.
- Sensory impairment, such as sight or hearing loss.
- Alcohol or substance misuse.

The NICE definition and guidelines for multimorbidity suggest that healthcare professionals should account for multimorbidity when patients have frailty, struggle to manage their own care, are prescribed multiple regular medications, and who regularly access unplanned care (Kernick, Chew-Graham and O'Flynn, 2017).

The AMS acknowledges that their definition of multimorbidity could be extended as "given the heterogeneous nature of multimorbidity…. pooling individuals who may well have entirely different clusters of conditions is unlikely to provide generalizable evidence" (The Academy of Medical Sciences, 2018). They recommended that future multimorbidity research should take into account that different conditions can be grouped (or "clustered") if they originate in the same bodily system, have similar origins or similar treatment plans (The Academy of Medical Sciences, 2018; Willadsen *et al.*, 2018). The rationale behind this is that two or more diagnoses from the same cluster leads to a less complex clinical picture as treatments for each condition are more likely to be complementary and hence treatment plans should be easier to follow. They have termed this 'concordant multimorbidity' (The Academy of Medical Sciences, 2018). Conversely, two or more diagnoses from two different clusters are more likely to share different underlying causes and more likely require different treatment plans. The AMS termed this 'discordant multimorbidity' (The Academy of Medical Sciences, 2018). Discordant multimorbidity is thought to better capture more complex multimorbidity profiles and circumstances (The Academy of Medical Sciences, 2018).

Some researchers have taken this approach and extended it further to try and better capture more complex multimorbidity. For example, Harrison and colleagues have shown that prevalence estimates differ depending on how a disease entity is defined (e.g., whether diseases are grouped by bodily systems or kept separate) and by the minimum number of disease entities included. Given their findings, they recommended that:

- 'multimorbidity' be defined as the "co-occurrence of two or more chronic conditions within one person without defining an index chronic condition", and,
- 'complex multimorbidity' defined as the "co-occurrence of three or more chronic conditions affecting three or more different body systems within one person without defining an index chronic condition." (Harrison *et al.*, 2014).

1.2.2 Multimorbidity prevalence and incidence

In most countries multimorbidity is common and rising in prevalence. In 2013, 50 million individuals residing in the European Union were estimated to be living with multimorbidity (Rijken *et al.*, 2013). In England alone, an estimated one in four adults are currently living with multimorbidity, equating to just over 14 million people (Stafford *et al.*, 2018). Multimorbidity prevalence in England (defined as two or more chronic conditions) has also risen steadily over the last 15 years, from an estimated 30.8% of the population to an estimated 52.8% (Head, Fleming, Kypridemos, Schofield, *et al.*, 2021). However, it is important to note that multimorbidity prevalence estimates do vary considerably; in England alone, estimates have ranged between 15% and 30% depending on the source of data analysed and the conditions included in a definition of multimorbidity (The Richmond Group of Charities, 2018).

In most developed countries, multimorbidity prevalence is greatest amongst the elderly; in 2015, 54% of individuals in England aged 65 and over were estimated to have multimorbidity, and these figures are projected to increase two-fold over the coming decade (Kingston *et al.*, 2018). Nonetheless, the absolute number of those with multimorbidity is greater during mid-life than old age (Barnett *et al.*, 2012; Nicholson, Terry, *et al.*, 2019). Recent evidence from Head and colleagues also suggests that, since 2004, there has been a considerable decrease in the median age of multimorbidity onset, and hence multimorbidity incidence among people of working

age (between ages 16 and 65 years old) has steadily increased (Head, Fleming, Kypridemos, Schofield, *et al.*, 2021).

1.2.3 Outcomes for individuals with multimorbidity

Individuals with multimorbidity experience worse outcomes in multiple domains of life when compared to individuals with single chronic conditions. Those with multimorbidity experience higher mortality rates, a poorer quality of life and a lower quality of care (Boyd *et al.*, 2005; Walker, 2007; Marengoni *et al.*, 2011; Nunes *et al.*, 2016; Williams and Egede, 2016). From a health and care system perspective, these individuals also utilise more primary care consultations than individuals with single conditions, as well as exhibit higher rates of planned and unplanned hospitalisations and have higher costs (Mercer and Watt, 2007; Glynn *et al.*, 2011; Marengoni *et al.*, 2011; Naylor *et al.*, 2012; Kasteridis *et al.*, 2014; Chung *et al.*, 2016; Shand, Morris and Gomes, 2020; Stokes *et al.*, 2021). Polypharmacy - the use of multiple medications over a prolonged period of time – also remains particularly problematic for individuals with multimorbidity, and is associated with functional decline, adverse drug events and treatment burden (Boyd *et al.*, 2005; Akazawa *et al.*, 2010).

The picture worsens if the component chronic conditions span both physical and mental health (physical-mental multimorbidity) as opposed to solely physical conditions (physical-only multimorbidity; Buist-Bouwman *et al.*, 2005; Knowles *et al.*, 2013; Lackner *et al.*, 2013; Panagioti *et al.*, 2015; Naylor *et al.*, 2016). In England, approximately 4.5 million people are estimated to have physical-mental multimorbidity and this number is set to rise substantially over the next 15 years (Naylor *et al.*, 2016; Kingston *et al.*, 2018). In addition, physical-mental multimorbidity is costly. Poor mental health accounts for between 12 and 18 per cent of all of England's National Health Service (NHS) expenditure on long-term conditions, which equates to more than £8 billion each year (Naylor *et al.*, 2012).

1.2.4 Prevention as a priority for multimorbidity

Multimorbidity is a growing public health challenge. In 2018, the AMS stated that: "further research is urgently required to better understand the growing challenge of multimorbidity and improve the care of patients across the globe" (The Academy of Medical Sciences, 2018). Funding bodies such as the Medical Research Council, the

NIHR and the Wellcome Trust are increasingly collaborating to offer funding for research on and around the problem of multimorbidity (The Academy of Medical Sciences, 2020). Notwithstanding, the majority of research inevitably continues to take an economic, health system perspective - most research focuses on reducing the risk of costly and adverse experiences for older multimorbid patients such as preventing unplanned hospital admissions and tackling inappropriate polypharmacy. This is despite greater recognition of the problem of multimorbidity in mid-life (The Academy of Medical Sciences, 2018; Head, Fleming, Kypridemos, Pearson-Stuttard, *et al.*, 2021).

In their influential report, the AMS outlined six research priorities. One of these called for future research to develop a better understanding of strategies that prevent the accumulation of chronic conditions and the eventual development of multimorbidity (The Academy of Medical Sciences, 2018). These calls have been echoed by others (Head, Fleming, Kypridemos, Schofield, *et al.*, 2021). To develop these strategies, we need a better understanding of the causes and risk factors of developing multimorbidity, including the social determinants of multimorbidity (Head, Fleming, Kypridemos, Schofield, *et al.*, 2021).. This understanding is, at present, limited (Smith *et al.*, 2016).

1.3 Social determinants of health

1.3.1 Health as a biopsychosocial process

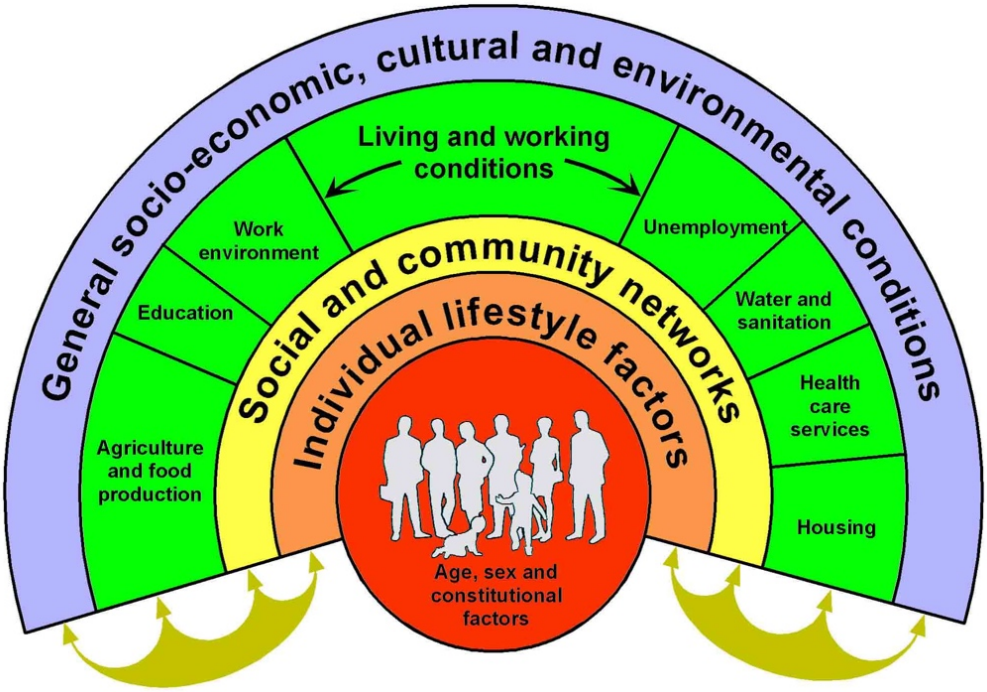
It is well acknowledged that social and environmental circumstances drive inequalities in health between different population groups, with those in less advantaged positions in society developing adverse health outcomes earlier in life, and spending a larger proportion of their lives living in ill-health (Marmot, 2010). These circumstances include our income, workplace, where we live and early childhood experiences (Marmot, 2010). These are broadly known as SDoH and are estimated to play a greater role in keeping people healthy than medical systems and healthcare services (McGinnis, Williams-Russo and Knickman, 2002).

The WHO defines SDoH as “the conditions in which people are born, grow, live, work and age” and the “fundamental drivers of these conditions” (Solar and Irwin, 2013). “Fundamental drivers” include social and economic policies enacted by governments.

The term ‘SDoH’ is therefore used to mean both the social and environmental factors influencing differences in individual health outcomes (for example differences in education levels) and the social processes (and policies) which cause the unequal distribution of such social factors between population groups of different socio-economic positions (Graham, 2004).

There are multiple conceptualisations of SDoH, however, a hierarchical division is common to many. These hierarchical divisions reflect distinctions between how areas and environments might influence health (contextual effects) compared to the characteristics of people within these areas (compositional effects). One widely known framework, the Dahlgren and Whitehead model, illustrates these different levels of influence (Dahlgren and Whitehead, 1991) (see Figure 1-1). Determinants positioned on the periphery, such as government policies, are often referred to as ‘upstream’ SDoH, or the “causes of the causes” (Marmot, 2010). These result in the unequal distribution of more immediate factors, such as unaffordable housing, which directly affect health via the unequal distribution of health behaviours or biological responses to living environments. This latter set of factors are often referred to as the ‘downstream’ effects of SDoH or the “causes of ill-health” (Braveman and Gottlieb, 2014).

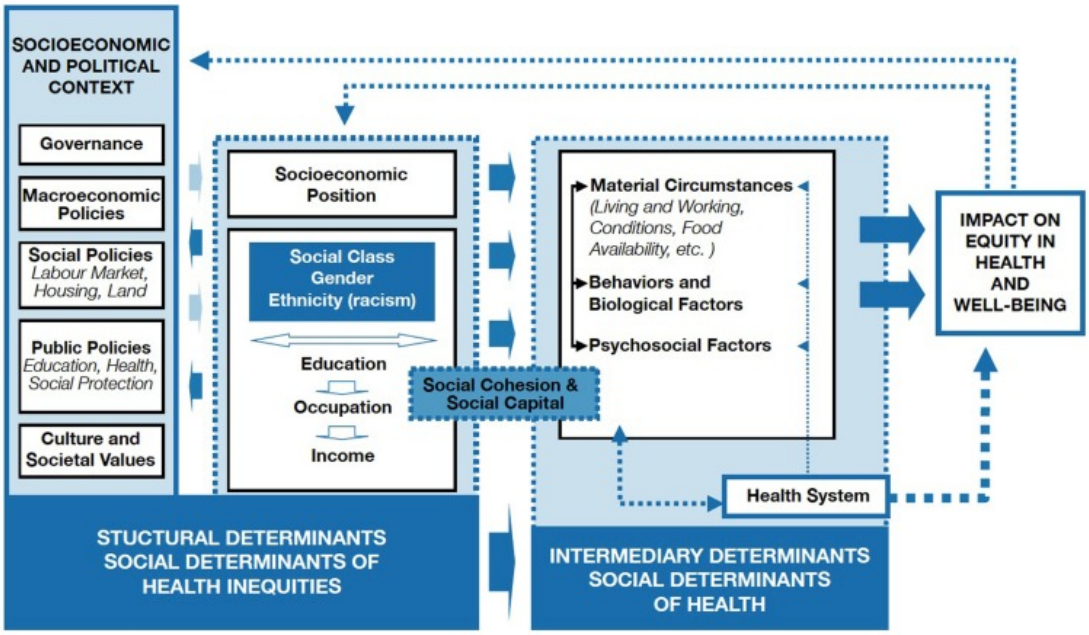
Figure 1-1: The Dahlgren and Whitehead 'rainbow model'



1.3.2 Conceptualising social determinants of health

To conceptualise SDoH, this thesis adapts the WHO's Commission on SDoH (CSDH) Framework (Solar and Irwin, 2013) (see Figure 1-2). The WHO's CSDH Framework summarises how socioeconomic and political mechanisms such as macroeconomic policies (structural determinants) can produce different socioeconomic positions where populations are stratified by characteristics such as social class, gender and/or ethnicity. According to this framework, socioeconomic positions then influence intermediary determinants of health such as material circumstances, with factors such as social cohesion and capital mediating these associations. Intermediary determinants then directly influence health, with the health system considered an additional intermediary determinant as healthcare access can influence individual health.

Figure 1-2: World Health Organization's Commission on Social Determinants of Health (CSDH) Framework



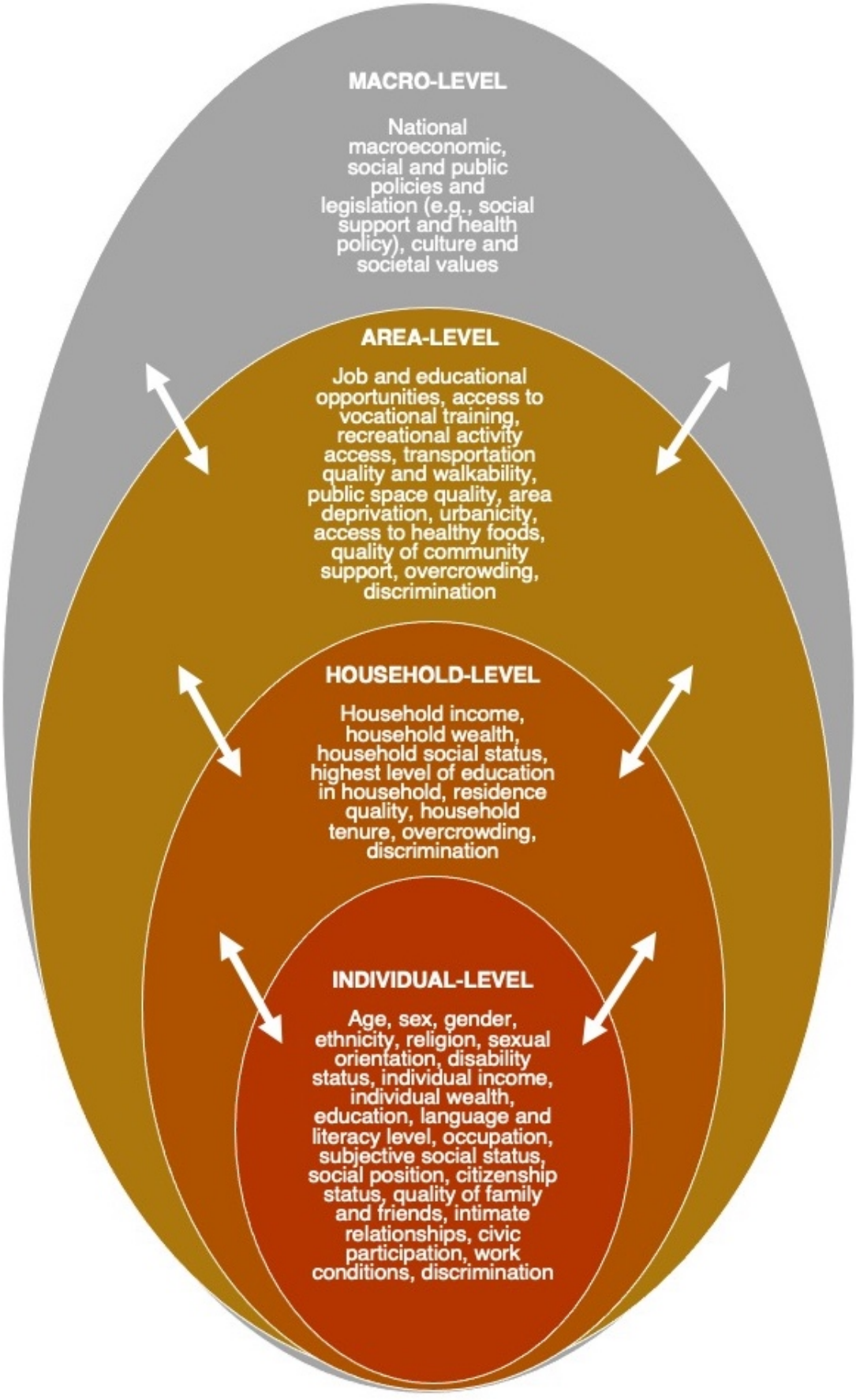
To guide my definition, I have selected upstream factors from the WHO framework that align with the idea that SDoH are the "causes of the causes" of ill-health (Marmot, 2010). These factors are the following structural determinants: governance, policies, cultural and societal values, socioeconomic position, social class, sex, gender, ethnicity, education, occupation, and income. I have also included material circumstances such as living and working conditions in my definition of upstream

SDoH as, whilst the WHO framework has classified these as intermediary determinants, I consider these to be structural SDoH. This is because considerable evidence shows that material circumstances can be directly influenced by government policies and, in turn, can directly influence health behaviours which then influence health. For example, access to healthy eating food outlets is a material circumstance shown to be directly influenced by government policies and that, in turn, influences eating behaviour, which in turn influences health (Pearce *et al.*, 2007).

In my conceptualisation of SDoH, I have excluded factors found in the WHO's CSDH Framework that are classified as the "causes of ill-health" to align with previous definitions of SDoH. These include health behaviours such as tobacco consumption, alcohol consumption and diet. This is because, whilst addressing health behaviours can realise direct benefits at an individual level, addressing health behaviours individually has limited influence on population health, with some arguing that there are distinct differences between the causes of individual variations in disease risk and the causes of disease differences at the population level (Rose, 1992). In addition, I have excluded factors related to the health system from my conceptualisation as many argue the health system itself is not a SDoH. This is because, whilst the health system can alleviate ill-health, the health system is not a major driver of ill-health in most circumstances (Illich, 2003). In this thesis health-related risk behaviours (such as smoking) and medical care factors (such as practitioners' decision-making) will, instead, be thought of as downstream effects of SDoH - the "causes" of ill health (Braveman and Gottlieb, 2014). By excluding these downstream factors this affords the opportunity to consider upstream factors in more detail.

In this thesis, SDoH are conceptualised as acting at four levels of influence: the individual, household, area, and macro-level (see Figure 1-3). Factors in my conceptualisation at the macro-level mirror those captured in the WHO Framework under the "socioeconomic and political context" (see Figure 1-2). Factors grouped under individual, household and area-levels are the remaining structural factors selected from the WHO's Framework, stratified according to the level at which they exert their influence on health. I have included arrows to illustrate the bidirectional relationships that can exist between SDoH and across levels. My conceptualisation is similar to others, both in terms of the dimensions included as well as the focus on relationships between levels (Centers for Disease Control and Preventation, 2021).

Figure 1-3: Conceptualisation of SDoH used in this thesis



1.3.3 Multimorbidity and social determinants of health

1.3.3.1 Compositional social determinants

One of the biggest drivers of multimorbidity is increasing age (Marengoni *et al.*, 2011; Violan *et al.*, 2014; The Academy of Medical Sciences, 2018). This has been demonstrated in many population sub-groups across a range of countries and contexts. However, as described in section 1.2.2, a multitude of research suggests that, whilst the prevalence of multimorbidity is higher in older populations, the absolute number of those with multimorbidity is greater amongst those 65 and younger (Barnett *et al.*, 2012; Rocca *et al.*, 2014; Bobo *et al.*, 2016). In addition, recent evidence suggests that the median age of multimorbidity onset is decreasing, and, thus, the incidence of multimorbidity among people of working age is steadily increasing (see section 1.2.2; Head, Fleming, Kypridemos, Schofield, *et al.*, 2021).

Multimorbidity prevalence and incidence is also greater for women, ethnic minorities, individuals with lower levels of education and social class, and those with lower incomes (Marengoni *et al.*, 2011; Salisbury *et al.*, 2011; Rocca *et al.*, 2014; Violan *et al.*, 2014; St Sauver *et al.*, 2015; Bobo *et al.*, 2016; Mokraoui *et al.*, 2016; The Academy of Medical Sciences, 2018; Khanolkar *et al.*, 2020). Adverse socioeconomic status is associated with both an earlier onset of multimorbidity, and a more rapid acceleration of multimorbidity accumulation (Khanolkar *et al.*, 2020). For these compositional SDoH, there are differences in reported findings depending on the country and context under study, highlighting how associations between some SDoH and multimorbidity are context dependent (The Academy of Medical Sciences, 2018).

1.3.3.2 Contextual social determinants

First investigated by Barnett and colleagues, considerable research examining social determinants of multimorbidity has explored associations between living in socioeconomically deprived areas and multimorbidity prevalence or incidence (Barnett *et al.*, 2012; The Academy of Medical Sciences, 2018). Barnett and colleagues found that equivalent prevalence of multimorbidity occurred 10 to 15 years earlier in those living in the most deprived compared to the least deprived areas, and that the co-occurrence of physical and mental health conditions was more common with increasing deprivation (Barnett *et al.*, 2012). Similar findings have since been reported, although the magnitude of associations differ depending on the measure of

area-level socioeconomic deprivation used as well as the population, country and context under study (The Academy of Medical Sciences, 2018). Of considerable concern is recent evidence suggesting that area-level socioeconomic inequalities in multimorbidity prevalence have steadily increased since 2004, particularly for working age adults below 65 years old (Head, Fleming, Kypridemos, Schofield, *et al.*, 2021).

It has been argued that higher multimorbidity prevalence in more socioeconomically deprived areas is a consequence of higher rates of lifestyle risk factors in these areas. These risk factors include smoking, diet, and Body Mass Index (BMI). However, one study by Katikireddi and colleagues found that five different risk factors (smoking, alcohol consumption, diet, BMI and physical activity) explained just over 40% of socioeconomic inequalities in multimorbidity development over a 20-year period, after accounting for age and sex (Katikireddi *et al.*, 2017). In addition, evidence from Chen and colleagues suggests that SDoH are associated with not only the prevalence of multimorbidity, but the impact multimorbidity has on an individual (Chen, Karimi and Rutten-Van Mölken, 2020). They found that individuals with multimorbidity reported poorer mental health and greater limitations to their activities of daily living if in the lowest education level group compared to the highest (Chen, Karimi and Rutten-Van Mölken, 2020).

1.3.3.3 Previous literature reviews in this area

Four previous studies have reviewed and synthesised evidence on the determinants of multimorbidity. These are limited for several reasons:

- Violan and colleagues (2014) reviewed studies on the prevalence, patterns and determinants of multimorbidity (Violan *et al.*, 2014). They identified five studies examining multimorbidity determinants. SDoH were measured at individual and area-levels of influence - they did not examine household SDoH such as household income. As such, relevant studies by Agborsangaya and colleagues were missed (Agborsangaya *et al.*, 2012).
- Northwood and colleagues (2018) published an integrative review examining how SDoH have been considered as dimensions of multimorbidity (Northwood *et al.*, 2018). They focused on older adults (>60 years) and included studies published between 2000 and 2015, missing subsequently published, relevant studies. By focusing on how SDoH have been considered,

they did not report actual associations between a SDoH under investigation and multimorbidity.

- Pathirana and Jackson (2018) published a systematic review examining existing literature on the associations between socioeconomic position and multimorbidity occurrence (Pathirana and Jackson, 2018). They primarily focused on individual education, deprivation (unclearly defined) and individual income. Their search strategy was limited to papers accessible via Medline and EMBASE pre-December 2014, missing key literature such as studies by Agborsangaya and colleagues, and McLean and colleagues (Agborsangaya *et al.*, 2013; McLean *et al.*, 2014). Restricting searches to Medline and EMBASE is problematic when exploring SDoH as these databases solely include medical, biomedical sciences, and drug and pharmacy journals. Pathirana and Jackson also combined findings from high- and low-income countries and did not formally assess study quality using an assessment tool. The former is an issue when exploring social determinants of multimorbidity as findings suggest the socioeconomic gradient of multimorbidity is reversed in low income countries (Garin *et al.*, 2016).
- The AMS (2018) examined literature investigating social determinants of multimorbidity as part of their report (The Academy of Medical Sciences, 2018). However, they did not conduct a systematic search which missed relevant literature.

1.3.3.4 Research gaps and priorities

Our understanding of social determinants of multimorbidity is incomplete and current literature reviews are limited. A multitude of research has investigated compositional SDoH (e.g., age, sex, and ethnicity), however, aside from area-level deprivation indices, the possible influence of household and other contextual characteristics (e.g., household income) have received little attention. In their report, The AMS stated that "research on the determinants of multimorbidity is sparse, conflicting, and limited to cross-sectional studies". They also concluded that most evidence focuses on "population or individual-level" determinants and that "it will be valuable to consider whether factors that operate at the household level can also influence multimorbidity". In addition to their research priority calling for the development of strategies to prevent multimorbidity, they also recommended that future research prioritises work to identify

and understand the determinants of the most common clusters of conditions in multimorbidity (The Academy of Medical Sciences, 2018).

Effective preventative strategies for multimorbidity will require efforts that tackle determinants of health outside of the control of health and care systems – including the environmental, social and economic determinants of health (NIHR, 2021). Understanding and addressing SDoH may prevent and postpone the accumulation of chronic diseases and the emergence of multimorbidity (Head, Fleming, Kypridemos, Pearson-Stuttard, *et al.*, 2021). This could minimise future strain on the health and care system. As stated by Head and colleagues:

> *"The importance of structural drivers, and crucially social determinants, in driving the NCD [non-communicable disease] burden and multimorbidity means preventative efforts are not limited to the health and social care sectors…..We need to challenge the common narrative that multimorbidity is inevitable in a modern ageing society. To do this, the focus on multimorbidity must shift from solely management of high-risk older individuals to include integrated population-level prevention strategies throughout the life-course to address the drivers of multimorbidity."* (Head, Fleming, Kypridemos, Pearson-Stuttard, *et al.*, 2021)

Understanding social variations in disease, and in public health issues like multimorbidity, is also needed for health and care leaders to ensure care is organised to be equitable.

1.4 Integrated care

The growing prevalence of chronic disease, coupled with an ageing population, is placing increasing demand and pressure on the health and care system. This is compounded by NHS funding deficits of approximately £22 billion, sustained cuts to social care funding in real-terms and wider budget cuts to local government, community services and mental health services over the last decade (NHS, 2014; Hastings *et al.*, 2015). The sustainability of the NHS is therefore under threat, and more efficient ways of working are needed (The King's Fund, 2020a).

To meet these challenges, there is a national drive to increase coordination across health and care, and place greater focus on disease prevention by addressing SDoH. In England, this approach is often referred to as delivering more 'integrated' health

and care. Integrated care has been proposed as a means to address rising multimorbidity and tackle health inequalities:

> *"Not only is our population growing in size, people are also living longer but suffering from more long-term conditions…… Faced with these challenges, as well as those from Covid-19, the case couldn't be clearer for joining up and integrating care around people rather than around institutional silos – care that focuses not just on treating particular conditions, but also on lifestyles, on healthy behaviours, prevention and helping people live more independent lives for longer."* (Department of Health and Social Care, 2021b)

1.4.1 Introducing "integrated care"

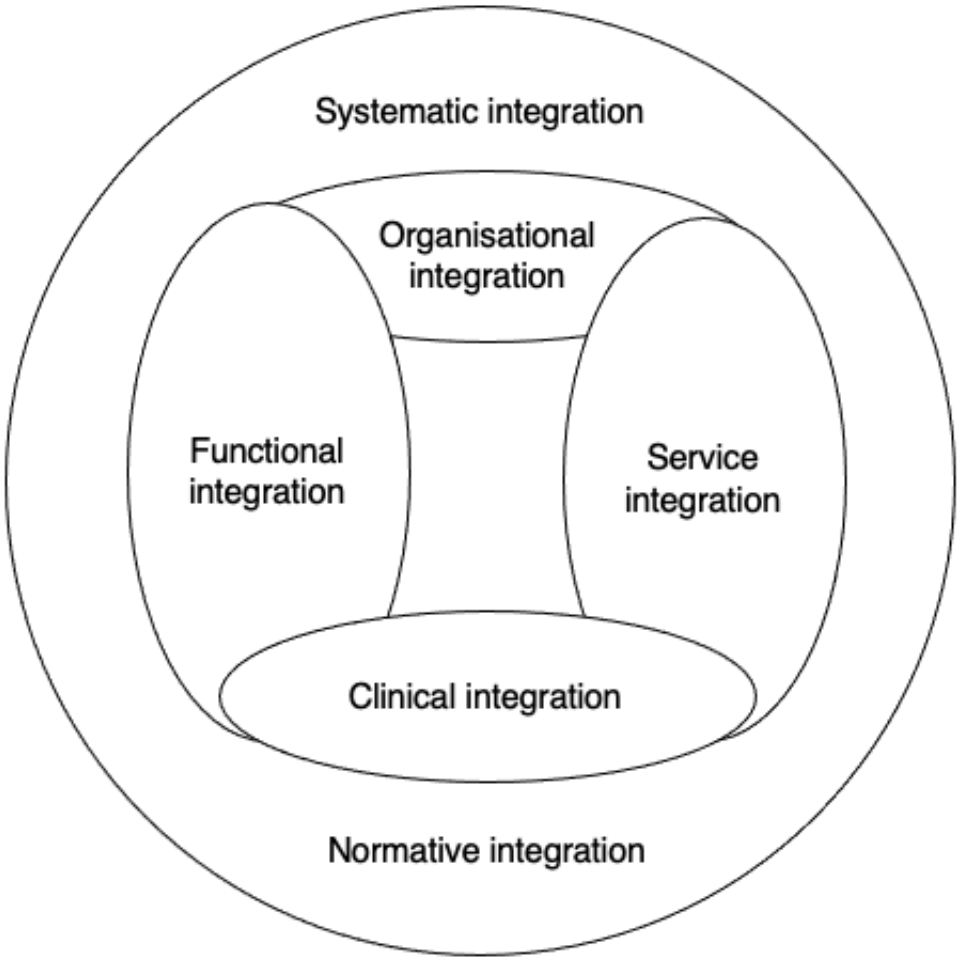
The concept of integration has its origins in organisational theory, where it is postulated that success of an organisation relies upon the ability of its constituent departments to functionally integrate in order to meet external demands (Axelsson and Axelsson, 2006). In health and care, the concept of integration can take on different meanings depending on the context and perspective taken (Leichsenring, 2004). Healthcare perspectives to integration typically emphasise increasing coordination between organisational interfaces within the health system, such as between primary and secondary care (Leichsenring, 2004). Broader perspectives encapsulate care delivered by other sectors such as social care and therefore take a more "patient-centric view of integrated care" (Kodner and Spreeuwenberg, 2002; Leichsenring, 2004). In the United Kingdom (UK), the concept of integrated care typically draws on the latter perspective, and often refers to the linking of health and social care sectors (Department of Health and Social Care, 2021b).

A plethora of terms are used to refer to integrated care given inconsistency and ambiguity around the concept of integration itself. These include 'coordinated care', 'collaborative care' and 'patient-centred care', with a recent review identifying 175 different definitions of integrated care in the literature (Kodner and Spreeuwenberg, 2002; Nolte and McKee, 2008; Armitage *et al.*, 2009). Kodner and Spreeuwenberg's (2002) definition of integrated care – one of the most cited definitions - will be employed throughout this thesis:

> *"[Integration is] a coherent set of methods and models of funding, administrative, organisational, service delivery and clinical levels designed to create connectivity, alignment and collaboration within and between the cure and care sectors…. [whose overall aim is to] enhance quality of care and quality of life, consumer satisfaction and system efficiency for patients with complex, long-term problems cutting across multiple services, providers and settings."* (Kodner and Spreeuwenberg, 2002)

Despite inconsistency in how the concept of integration is defined and operationalised, it is generally agreed that distinguished dimensions of integration exist. One taxonomy of integration, first proposed by Contandriopoulos and Denis (2001) and later adapted by Mowlam and Fulop (2005), is seen in Figure 1-4 (Contandriopoulos, Denis and Touati, 2001; Fulop, Mowlem and Edwards, 2005). Mowlam and Fulop's typology emphasises different types and processes of integration argued to be "key requirements for effective integration" (Fulop, Mowlem and Edwards, 2005; Nolte and McKee, 2008). In their taxonomy, organisational integration refers to the integration of formal organisational structures (e.g., through structural changes including mergers or contractual arrangements). Functional integration refers to the degree to which back-office functions such as human resources, financing and strategic planning are coordinated and integrated across organisations. Service integration refers to the coordination of clinical services at the organisational level, whereas clinical integration refers to the coordination of the care process for patients at the clinical team level. The convergence of values and approaches (normative integration) and organisational rules and policies (systemic integration) are also included and deemed "crucial in determining how successful integration is" (Fulop, Mowlem and Edwards, 2005).

Figure 1-4: Typology of healthcare integration, adapted from Mowlam and Fulop (2005)



1.4.2 Integrated care in England

By deduplicating work and streamlining services, health and care integration in England is viewed as an opportunity to achieve efficiency gains, whilst simultaneously improving the quality of care (Department of Health and Social Care, 2021b). As such, there has been a progressive movement towards greater integration in the last decade on account of the following milestones in national policy:

- **2010:** The 'Healthy Lives, Healthy People' report was published, which argued that local authorities should be responsible for public health (HM Government, 2010).
- **2010-12:** 16 areas in England were chosen to form Integrated Care Pilots (ICPs). Each area proposed a plan to integrate certain services around a given

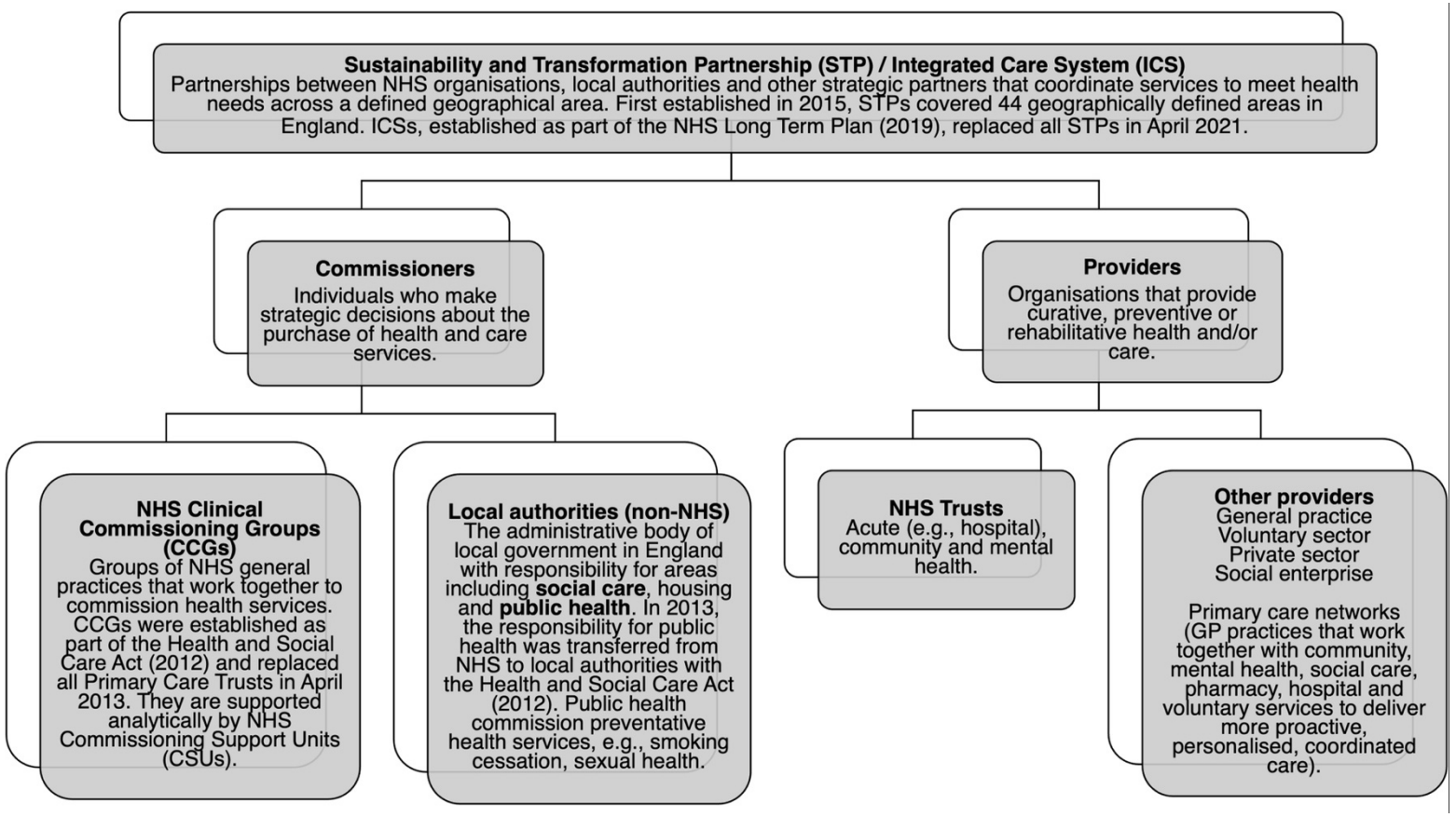
focus or client group (Europe R.A.N.D., 2012). The ICP programme lasted two years.

- **2012:** The Health and Social Care Act was passed. The Act sought to "encourage and enable more integration between services". It legislated the transfer of responsibility for public health from NHS to local authorities.
- **2014:** The NHS Five Year Forward View (FYFV) put forward a vision for reorganisation of the system around "new models of care", which emphasised increased co-ordination between local health and social care providers and a "radical upgrade in prevention" (NHS, 2014).
- **2015:** In response to the FYFV, NHS England selected 50 sites to act as 'vanguards' to lead the development of these new care models.
- **2015:** England was divided into 44 geographical 'footprints' forming 44 local Sustainability and Transformation Partnerships (STPs; previously known as sustainability and transformation plans). STPs were partnerships between NHS organisations, local authorities, provider organisations and other strategic partners. They were required to collectively develop long-term plans outlining how they would improve care in their areas and achieve financial balance. Separate financial contracts between care commissioners and each care provider remained.
- **2017:** NHS England's FYFW update boldly stated their aim to "use the next several years to make the biggest national move to integrated care of any major western country" (NHS England, 2017).
- **2018:** 14 STP areas were selected to gradually develop into Integrated Care Systems (ICSs; previously 'accountable care organisations') (King's Fund, 2018). Like STPs, ICSs are partnerships between NHS organisations, local authorities, provider organisations and other strategic partners. They have a contract that is shared across providers, and organisations in the collaboration are collectively responsible for managing resources. See Figure 1-5 for an overview of the partners in a typical ICS in England, accurate as of September 2021.
- **2019:** The NHS Long Term Plan stated that all local NHS organisations would increasingly focus on population health through a move towards ICSs everywhere (NHS, 2019).

- **February 2021:** The Department of Health and Social Care released legislation that required all 44 STPs in England form ICSs by April 2021 (Department of Health and Social Care, 2021b).

Inconsistency in how the concept of integrated care is defined and operationalised in practice, plus a lack of national frameworks or blueprints for integration, has, inevitably led to variation in how ICSs are developing (Charles *et al.*, 2018). ICSs vary in the population sizes they cover, the number of organisations involved and in their complexity (Charles *et al.*, 2018). They also differ in the specific activities taking place. For example, the South Yorkshire and Bassetlaw ICS are working to develop 'networks of care' in each hospital in their area. The aim is for each 'network' to specialise in a given service (e.g., stroke) in order to reduce the transfer of patients between hospitals and maximise expertise (Charles *et al.*, 2018). In contrast, Bedfordshire, Luton and Milton Keynes ICS have piloted a programme that aims to improve cardiovascular prevention by increasing screening for hypertension and atrial fibrillation (AF) in community pharmacies, who then refer patients to general practice (Charles *et al.*, 2018). Despite this variation, integration in most areas involves horizontal integration (between organisations on the same level of health and care delivery) rather than vertical integration (between organisations occupying different levels of a hierarchical structure) (Axelsson and Axelsson, 2006).

Figure 1-5: An overview of partners within a typical (NHS) Integrated Care System in England.



Note: Relationships between constituent Integrated Care System (ICS) organisations, formerly sustainability and transformation partnership (STPs), adapted from The King's Fund explainer: "The NHS: how providers are regulated and commissioned" (The King's Fund, 2020b).

1.4.3 Integrated care as a proposed solution for multimorbidity

Increasing integration between health and care services offers opportunities to address the public health challenge of multimorbidity and provide better care for people with multimorbidity (NHS, 2019; Nicholson, Makovski and Stranges, 2019). This is because our health and care system is currently designed around a single disease model and medical training, drug trials and treatment guidelines continue to be designed around single conditions (Boyd *et al.*, 2005; Naylor *et al.*, 2016; The Academy of Medical Sciences, 2018; Whitty *et al.*, 2020). This often means those with multimorbidity must navigate complex, uncoordinated care pathways, as care regularly crosses multiple health and care services including social care, and, as such, crosses organisational, professional and sectoral boundaries (Kasteridis *et al.*, 2014, 2015; Doessing and Burau, 2015; James Lind Alliance, 2018; The Academy of Medical Sciences, 2018; The Richmond Group of Charities, 2018; National Voices, 2019). This is particularly challenging for those with mental health conditions, as care in the UK is more fragmented if it spans physical and mental health services (Naylor *et al.*, 2016). People with multimorbidity and their carers consistently call for structural and clinical changes that afford more integrated, person-centred and holistic care (Bayliss *et al.*, 2008; James Lind Alliance, 2018; The Richmond Group of Charities, 2018; National Voices, 2019; NIHR, 2021).

Historically, addressing health issues has been within the remit of the NHS, and addressing social circumstances within the remit of local authorities. Integration of health and care, therefore, also presents opportunities to address SDoH and social determinants of multimorbidity, as the risk factors for multimorbidity are often social in nature and influenced by public services outside of typical health and care services (see section 1.3.3). Indeed, the government's 2021 White Paper stated that "integrating care…..enables greater ambition on tackling health inequalities and the wider determinants of health – issues which no one part of the system can address alone" (Department of Health and Social Care, 2021b).

1.4.4 Integrated care and data

Health and care services collect and hold a vast amount of administrative data about patients and residents (Groves *et al.*, 2013; Fontana *et al.*, 2020). This data is often collected for operational purposes, such as helping to plan effective day-to-day care,

and/or monitoring and improving service delivery. In some cases, collecting certain data is statutorily required. For example, as part of their statutory requirements, local authorities collect information on the social circumstances of their residents. This information includes administrative data around housing, social care provision, benefits receipt, and educational outcomes.

In England, the use of knowledge generated from the analysis of administrative data records (widely referred to as 'analytics') is seen as central to delivering more integrated care. This idea is echoed throughout government policy, for example the recent government White Paper (published February 2021) stated that "integrating care… relies on the power of digital and data to join up care and uses that power to drive transformation of care" (Department of Health and Social Care, 2021b). Amongst other things, analytics can aid assessments of local health and care needs to support the development of new, more integrated services or can be used to monitor the effectiveness, efficiency and quality of existing services (Beenstock *et al.*, 2014; Bardsley, 2016; Marshall *et al.*, 2016; Kneale *et al.*, 2017; Edwards, 2019). Senior leaders of health and care organisations are, therefore, increasingly expected to use analytics to inform decisions about the short and long-term structure and delivery of services. These decisions can have implications across organisational and sectoral boundaries and are, hereafter, called 'strategic health and care decisions'.

At present, health and care organisations collect, store, and process their respective data separately, reflecting divisions in the NHS and the wider care system. It is argued that this makes it difficult to deliver more integrated clinical care day-to-day at an individual-level, and plan more health and care service integration at a population-level. Separate data systems within and across organisations therefore present barriers to realising the opportunities of integrated care. As such, whilst service changes taking place with each ICS vary depending on area and context, one overall common goal that ICSs are working towards is to improve information-sharing across services, sectors, and organisations.

1.4.5 Integrated care and data linkage

Data linkage can be defined as "the process of identifying, matching and merging records that correspond to the same person from several [separate] datasets" (SA NT DataLink, 2021). It is increasingly seen as a way to overcome some of the limitations

of separate health and care records and thought to provide a more in-depth and complete understanding of local health and care needs and patient journeys across services. Many who argue for increased data linkage suggest that insights generated from the analysis of linked data *will* improve decisions, service delivery and, ultimately, care. For example, a recent UK government paper acknowledges that divisions of data across health and care organisations make it "very difficult for local and national leaders…to effectively plan, commission, and develop policy [and services]" (Department of Health and Social Care, 2021a), In addition, Administrative Data Research UK (ADRUK) – a partnership of UK-based organisations working to increase the linkage of administrative data - state that:

> *"By joining up the abundance of administrative data already being created by government and public bodies across the UK….we are enabling vital research that has the potential to lead to better informed policy decisions and more effective public services"* (ADR UK, 2021b)

In the England, each NHS patient is given a unique, individual NHS identifier shortly after birth that is recorded on their health records. The NHS number is typically used to identify which health records, across separate services, belong to the same patient, and the NHS number typically forms the basis for the linkage of health data. However, whilst NHS numbers are increasingly being used on social care records, this is not routinely done. In addition, few council departments, such as those responsible for housing and education, use NHS numbers in their administrative data - instead using their own different individual identifiers. This makes it hard to fully link all records across health and care. Therefore, whilst several local areas are working to develop, or have developed, datasets linking health and care administrative records, few contain social care and other council data (see Table 1-1). When NHS numbers are not available, records can be linked using sensitive data such as first and last name, address, and date of birth. There are a handful of examples where councils have used these methods to link data across departments and inform service delivery, although these examples are often small scale projects with discrete uses of linked data (Symons, 2016). As can be seen in Table 1-1, linked health and care datasets in the UK (that are possible to use for research) often link data from different services, depending on local priorities.

Table 1-1: Examples of linked health and care datasets in the UK possible to use for research, adapted from (Shand, 2020)

Dataset	Aim/Description	Sample size and geography	Time period	Settings included					
				Primary	Hospital	Community	Mental Health	Social Care	Council (other)
Clinical Practice Research Datalink (CPRD)	CPRD is a dataset containing de-identified data from a sample of UK primary care practices. Hospital data and other datasets including mental health data can be linked on a project-by-project basis.	~16 million active and 60 million ever recorded individuals across 674 UK primary care practices	1988-present (average of 5 years of data per person)	**X**	**/**	**--**	**/**	**--**	**--**
Clinical Record Interactive Search system (CRIS)	CRIS is a dataset that captures mental health activity and free text data from patient records from South London and Maudsley NHS Foundation Trust.	~1.2 million individuals who have had a mental health service contact in South London	2007-present	**--**	**X**	**--**	**X**	**--**	**/**
The Secure Anonymised Information Linkage for Wales (SAIL Databank)	The SAIL Databank is a dataset that contains multiple datasets from across health, disease registries, screening data and education data that can be linked at individual and household-levels using a linkage key.	~3 million individuals in Wales, representing entire population.	2007-present	**X**	**X**	**/**	**--**	**--**	**/**
Kent Integrated Dataset (KID)	KID is a dataset that contains linked data to provide system-wide population health and care utilisation insights. Data updated monthly.	~2 million individuals in Kent and Medway	2014-present	**X**	**X**	**X**	**X**	**X**	**--**
Care City Cohort	The Care City Cohort is a dataset linked at individual and household-levels. Data updated annually.	~250,000 residents of the London Borough of Barking and Dagenham	2011-present	**X**	**X**	**X**	**X**	**X**	**X**

Dataset	Aim/Description	Sample size and geography	Time period	Settings included					
				Primary	Hospital	Community	Mental Health	Social Care	Council (other)
Discover-NOW	Discover-NOW is a de-identified dataset created from the identifiable Whole Systems Integrated Care dataset used for direct care. Data updated monthly.	~2.4 million individuals in North West London	2014-present	**X**	**X**	**X**	**X**	**X**	**-**
Connected Yorkshire	Connected Yorkshire is a de-identified dataset created by the Connected Health Cities programme, which was a four-year pilot funded by the Department of Health. The programme created multiple datasets for different research projects.	~700,000 individuals from across Bradford, Leeds and Sheffield	1970-present	**X**	**X**	**X**	**X**	**X**	**X**
Scottish Longitudinal Study (SLS)	SLS is a dataset that contains longitudinally linked census, vital events, and education data. It is accessible to researchers on a project-by-project basis with health data added for specific projects.	~275,000 individuals representing a random 5% sample of the Scottish population.	1991-present	**--**	**X**	**--**	**--**	**--**	**X**

X = Data available; **/** = Data partially available**; --** = Data not available

As can be seen in Table 1-1, most datasets that link administrative data from health and council records, of which there are few, often link social care data extracted from council records. A handful of these linked datasets have been used to better understand how those with multimorbidity use health and care services (Kasteridis *et al.*, 2014; Mori *et al.*, 2019; Henderson *et al.*, 2021). In Scotland, health and council social care records have been linked nationally to quantify associations between multimorbidity and receipt of formal social care services (Henderson *et al.*, 2021). This work found that those with more severe multimorbidity profiles exhibit higher levels of social care receipt (Henderson *et al.*, 2021). A further example from Japan demonstrates that similar linkages can reveal how those with multimorbidity have greater medical care costs, social care costs and overall costs, suggesting that "the economic burden on society caused by multimorbidity can be evaluated better by considering both medical and [social care] expenditures, rather than medical expenditures alone" (Mori *et al.*, 2019). The inclusion of other types of council data such as data on housing or education are less frequently included in such data linkage initiatives.

Datasets that link information from across health and wider council services such as housing are *specifically* viewed as a potential way to "generate useful insights, shining a light on inequalities and their causes" (ADR UK, 2021a). This is because, whilst most of the drivers of health inequalities are social in nature (Marmot, 2010), health data has traditionally been held by health services independent of social information that is held by local authorities. Linked health and council datasets therefore present opportunities to investigate social determinants of important public health problems like multimorbidity. An example of a project with this aim (currently in progress) is the creation of the Wales Multimorbidity e-Cohort (WMC) (Lyons *et al.*, 2021). The WMC has been created and derived from data held in the Secure Anonymised Information Linkage (SAIL) Databank (see Table 1-1). It is intended that the WMC will elucidate understanding of the prevalence, trajectories, and determinants of multimorbidity, and identify clusters causing the greatest healthcare burden to help support service planning and preventative initiatives (Lyons *et al.*, 2021).

This type of knowledge generated from the analysis of linked health and council data therefore also presents opportunities to inform strategic and *equitable* decision-making if senior leaders investigate population groups with the highest levels of need

and plan care services accordingly. However, it is unclear if and how the analysis of such linked data will influence decision-making and the equity of resultant services. At present, our understanding of when and how analytics are used to inform more equitable strategic decision-making is incomplete, and it is unclear what facilitates and hinders data use in this context. The actual impact of such linked data on decision-making and care delivery also depends on how analytics are perceived and used by senior leaders in practice.

1.5 Chapter summary

In England, there is a drive to increase integration both within health and across health and local government. To try and facilitate integration, several areas are linking, or have linked, data records from across their different health and care services. Only a handful of these include council data because of various practical, technical, and legal barriers. As such, this raises one important overarching question: what knowledge can be generated when data are successfully linked across health and council records? In addition, it is often assumed that knowledge generated from the analysis of such linked data *will* improve decision-making and enable the delivery of more equitable health and care services. This raises a second important, overarching question: When these linked datasets are more readily available, will such knowledge inform the equity of decision-making and, if so, how?

Through the lens of the important public health problem of multimorbidity, this thesis explores these two related questions. This is because linked health and council datasets present opportunities to further our understanding of social determinants of multimorbidity as well as inform the delivery of strategic and more equitable health and care decision-making for groups such as those with, or at risk of, multimorbidity.

Chapter 2 Research questions, design, aims and objectives of this thesis

2.1 Research questions

This thesis has two related research questions:

1. Could linked health and council data advance our understanding of the social determinants of multimorbidity?
2. How might such linked data influence the equity of strategic health and care decision-making for groups such as those with, or at risk of, multimorbidity?

2.2 Overview of research design

Using mixed methods, this thesis will attempt to answer these research questions, addressing research gaps and priorities outlined in Chapter 1.

The first part of this thesis will include a systematic review of the literature examining household and area-level social determinants of multimorbidity (Chapter 3). Findings from this review will inform my quantitative study (Chapter 4 and Chapter 5). Through the lens of multimorbidity, the quantitative study I describe in Chapter 4 and Chapter 5 will act as a use case for creating, using and analysing linked health and council datasets to generate knowledge that could inform equitable decision-making.

The second part of this thesis will use qualitative methods to explore senior leaders' experiences of using knowledge generated from the analysis of administrative health or care records to inform strategic and equitable health and care decision-making (Chapter 6 and Chapter 7). This qualitative study will explore whether linked health and council data could or would be used by senior leaders to influence the equity of strategic decision-making.

2.3 Aims

Given my overarching research questions, the aims of this thesis are:

1. To illustrate how knowledge generated from the analysis of linked health and council data could advance our understanding of the social determinants of multimorbidity.
2. To explore how knowledge generated from the analysis of linked health and council data might influence strategic and equitable health and care decision-making.

2.4 Objectives

Aim 1: To illustrate how knowledge generated from the analysis of linked health and council data could advance our understanding of the social determinants of multimorbidity.

As illustrated in Chapter 1, the majority of literature examining social determinants of multimorbidity has focused on compositional characteristics (e.g., age, sex, and ethnicity) and area-level deprivation. No syntheses of literature examining associations for other household and contextual characteristics (e.g., household income) exist.

To address Aim 1, I will therefore:

Conduct a literature review (Chapter 3) to:

- systematically identify, critically appraise, and synthesise existing literature examining household and area-level social determinants of multimorbidity,
- inform the design of my quantitative study (Chapter 4 and Chapter 5).

Conduct a quantitative analysis of a linked health and council dataset extracted from administrative health and care records (Chapter 4 and Chapter 5) to:

- examine and quantify associations between selected household characteristics (informed by my systematic review) and multimorbidity,
- act as a use case for creating, using, and analysing such linked datasets to understand the social determinants of local public health concerns and generate knowledge that could inform equitable decision-making.

Aim 2: To explore how knowledge generated from the analysis of linked health and council data might influence strategic and equitable health and care decision-making.

As described in Chapter 1, it is often assumed that linked data *will* improve decision-making, care, and the equity of health and care services. However, it is unclear whether these aspirations will be realised as our understanding of when and how analytics are used to inform decision-making is incomplete, and it is unclear what facilitates and hinders data use in this context. In addition, the majority of senior leaders who make strategic decisions do not have access to data linked across health, council and other settings. I have therefore focused the second part of my thesis on how knowledge generated from the analysis of linked health and council data (such as knowledge generated in Chapter 4 and Chapter 5) *might* influence strategic and equitable decision-making.

To address Aim 2, I will therefore:

Conduct individual, semi-structured interviews with senior leaders of health and care organisations (Chapter 6 and Chapter 7) that:

- explore when and how knowledge generated from the analysis of residents' health *or* care records is used to inform strategic and equitable decision-making across health and care,
- identify barriers and facilitators of analytics use in this context,
- describe how leaders differ in their experiences of, and responses to, identified barriers and facilitators.

Construct a typology based on senior leaders' experiences of using analytics (Chapter 6 and Chapter 7) to:

- identify and define different types of analytics users,
- develop recommendations for those aiming to improve analytics use for decision-making.

To finish, I will summarise the implications of the thesis findings for policy and practice and identify where further research is needed (Chapter 8).

Chapter 3 Household and area-level social determinants of multimorbidity: a systematic review

3.1 Introduction

In Chapter 1, I described multimorbidity as a major public health problem that, to address, requires strategies focused on prevention. Preventative strategies necessitate a broader understanding of the factors associated with multimorbidity. In section 1.3.3.1, I described how compositional social determinants of multimorbidity (e.g., age, sex, and ethnicity) have been well examined in the literature. However, our understanding of household and area-level social determinants of multimorbidity (e.g., household income) is incomplete, with most primary research focusing on area-level deprivation indices.

In section 1.3.3.3, I also gave an overview of previous reviews of the literature that have been published in this area and outlined their limitations which include: missing relevant literature, restricting searches to older age populations, restricting searches to solely medical databases and including studies conducted in both high- and low-income countries. At present, no clear synthesis of evidence specifically examining household and area-level social determinants of multimorbidity exists.

In this chapter I will present the methodology and results of a review I have conducted that aimed to systematically identify, critically appraise, and synthesise existing literature examining household and area-level social determinants of multimorbidity. I will finish by considering how my findings compare to wider literature and the limitations of the literature to date. I have described how these findings have been used to inform my quantitative study (Chapter 4 and Chapter 5) in section 4.1.1.

3.1.1 Research questions

This review answered the following research questions:

1. What household and area-level SDoH have been included in studies exploring the association between SDoH and multimorbidity prevalence/incidence?

2. How are the considered SDoH associated with multimorbidity prevalence/incidence? (i.e., what have these studies found?)
3. How has multimorbidity been defined and measured?
4. How do reported associations differ with differences in how multimorbidity has been defined and measured?
5. How do associations differ with age, sex, and ethnicity?
6. What are the strengths and limitations of the identified literature?

Contributions of others to this study: I undertook all searches with second screening conducted by Sarah Ledden (SL), UCL Division of Psychiatry, and Sarah Beardon (SB), UCL Department of Applied Health Research and UCL Faculty of Laws. I assessed each study for risk of bias and SL assessed a subset of studies. Selected findings from this chapter have been published in a peer-reviewed journal article (Ingram *et al.*, 2021).

3.2 Methods

The methodological steps followed in this review were developed a priori and a summary of this protocol was registered on 13th May 2019 (PROSPERO number CRD42019135281). I adhered to guidelines from Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and the York Centre for Reviews and Dissemination when planning, conducting and writing this systematic review (Moher *et al.*, 2009; Tacconelli, 2010).

3.2.1 Eligibility criteria for inclusion

Studies meeting the criteria outlined in Table 3-1 were included. Studies were included if in English, conducted in high-income countries (HICs) and published between 1st January 2010 and 10th January 2019. HICs were defined according to the World Bank (The World Bank, 2020). The former date restriction coincides with the publication of The Marmot Report, which raised the profile of SDoH in England and led to increased efforts to examine and address SDoH (Marmot and Bell, 2012). The latter date restriction coincides with the date on which database searches were first ran. Studies were excluded if conducted solely with institutionalised individuals as social determinants of multimorbidity may differ between institutional and community settings (Moore *et al.*, 2014). Studies were excluded if conducted with

solely young people (<18 years) as the prevalence of multimorbidity is low for this group (Barnett *et al.*, 2012).

Table 3-1: Review inclusion and exclusion criteria

	Inclusion	Exclusion
Population	Study participants from the general population and assessed for the presence of multiple chronic conditions (multimorbidity).	Study participants initially selected based on the presence of index diseases (i.e., studies of comorbidity). Participants from institutionalised care settings (e.g., nursing homes). Participants solely young people (age < 18 years).
Exposure	Study exposure(s) included at least one household or area-level SDoH that aligns with my conceptualisation of SDoH (see section 1.3.2) and the idea that SDoH are "causes of the causes" of ill-health (Marmot, 2005) (e.g. household income or area-level deprivation).	Study exposure(s) include individual-level SDoH only (e.g., ethnicity). Study exposure(s) are direct "causes" of ill-health, such as health behaviours (e.g., smoking) and factors associated with the health system itself (e.g., access to services).
Comparator	Study reports comparator group for SDoH exposure(s) i.e., what is the prevalence of multimorbidity for those in the lowest versus the highest household income groups.	Study does not report a comparator group for SDoH exposure(s).
Outcome	Study assesses multimorbidity burden (prevalence or incidence studies).	Study assigns participants to multimorbidity patterns or trajectories, or measures multimorbidity severity (e.g., indices weighted by disease severity).

	Inclusion	Exclusion
Study Design	Peer-reviewed studies of quantitative research designs (cross-sectional and longitudinal).	Systematic reviews, meta-analyses, and qualitative research (citations of relevant reviews searched).

3.2.2 Developing and implementing the search strategy

The search terms used in this review were initially developed in MedLine (see Appendix 1) and adapted for a further five databases (EMBASE, PsychINFO, Web of Science, CINAHL Plus and Scopus). The search terms used were pre-defined and aimed to be exhaustive. I combined terms relating to multimorbidity, specific SDoH and households or areas using Boolean language. When developing the multimorbidity search terms, I decided to exclude the MeSH term 'comorbidity' and its linguistic variations as adding these returned an unfeasible number of references and a high proportion of irrelevant references. After running the initial searches, I added the MeSH term 'comorbidity' into my MedLine search to examine if any studies had been missed through excluding the term 'comorbidity' and its linguistic variations (see Appendix 2 for more details and the outcomes of this exercise). To develop my SDoH search terms, I drew on household and area-level factors included in my conceptualisation of SDoH (see section 1.3.2) as well as published frameworks and previous literature reviews that had searched for SDoH that aligned with my conceptualisation (Dahlgren and Whitehead, 1991; Nagata *et al.*, 2013; Solar and Irwin, 2013; Walker *et al.*, 2014; Canadian Council on Social Determinants of Health, 2015; Duan-Porter *et al.*, 2018). See Appendix 2 for further details on how I developed SDoH search terms.

Following database searches, I removed duplicated references and screened all titles and abstracts against the selection criteria. The entirety of these references were screened independently by a second reviewer (SL) and a subset (100 references) screened independently by a third (SB). References that could not be excluded at this stage were read in full. I then read all full texts and screened these against the selection criteria. The second reviewer screened a subset (20%). Any disagreement at each stage of screening was resolved by discussion or consultation with the third reviewer. Inter-rater reliability was calculated using Kappa statistics.

Following screening at full text, I extracted the relevant data for all studies using a pre-piloted form that included study characteristics, definitions of exposures and outcomes, and findings (see Appendix 3). The second reviewer checked the data extraction table for accuracy and completeness. If there was inadequate detail in the paper to enable a complete data extraction, study authors were contacted.

After running my initial searches in each database, I forward and backward citation searched all identified references and searched the citations of relevant reviews in order to identify further studies. I used the snowball method as it is an efficient, reliable and useful method when terminology is not consistently used in the literature (Greenhalgh and Peacock, 2005; Valderas *et al.*, 2009).

The first set of database searches were ran on 10th January 2019. Prior to submission of this thesis, all database searches were re-ran in May 2021 to identify if any further eligible references were published following the initial searches.

3.2.3 Quality assessment

Quality assessments were used to provide insight into the overall quality of evidence in this field. I did not make judgments about individual studies, rank studies on quality, or exclude based on quality. Rather, quality assessments were used to explore any associations between study results and quality assessments.

I developed my own quality assessment criteria that drew on categories seen in several published assessment checklists including the Newcastle Ottawa Scale and The Risk Of Bias In Non-randomized Studies of Interventions assessment tool (Appendix 4) (Sterne *et al.*, 2016; Wells *et al.*, 2019). I chose to do this as most widely used tools for assessing the quality of observational studies or the strength of evidence were not applicable. My criteria assessed study quality within four domains: selection bias, information bias for exposure and outcome, and confounding. Non-interventional studies are rarely at low overall risk of bias and reporting by domain allows comparison of the main sources of bias across studies (Sterne *et al.*, 2016). I also examined the strength of evidence reported and each study's applicability to this review, in keeping with advances in review methodologies (Viswanathan *et al.*, 2017). Information biases refer to biases that occur during data collection and can include errors misclassifying information (misclassification biases) which can occur for

information relating to both the exposures and outcomes measured (Delgado-Rodríguez and Llorca, 2004). Selection biases occur when the study population does not represent the target population (Delgado-Rodríguez and Llorca, 2004). Confounder biases are errors in accounting for other variables that may explain any observed associations (Delgado-Rodríguez and Llorca, 2004).

I assigned each study a high, medium, low or unclear quality rating for each domain to separate study quality from reporting quality, and as numerical or quality ratings combined across criteria can be hard to interpret (Tacconelli, 2010; Viswanathan *et al.*, 2017). I assessed risk of selection bias by comparing sample demographics to census data when possible. Studies where risk of bias was high across two or more domains were deemed low quality. Studies where risk of bias was mixed or medium across all domains were deemed moderate. Studies with a low risk of bias across two or more categories, with no high risk of bias across any domains, were deemed high quality. A second reviewer (SL) assessed the quality of a subset (20%) of included references.

3.2.4 Data synthesis

I chose to narratively synthesise study findings, structured per social determinant, as the exposures, outcomes, and study methodologies used across the included studies were diverse. I also decided that the studies were too heterogeneous to allow a meta-analysis of findings. For studies investigating associations between area-level deprivation and multimorbidity, I pooled available data to calculate overall multimorbidity prevalence in deprivation quintiles.

I have added the results of studies identified when the searches were re-ran prior to thesis submission to all results tables to give a thorough and complete overview of findings in the literature for each SDoH. However, I have separately synthesised the results of these studies from the results of the studies identified when initial searches were ran in January 2019. This is because the findings of these initial studies informed the design of the study described in Chapter 4 and Chapter 5, and these decisions were made prior to rerunning the searches.

3.3 Results

The following sections of this chapter (sections 3.3.1, 3.3.2, and 3.3.3) narratively synthesise the results of the studies identified when database searches were first ran on 10th January 2019. The results of additional studies identified when the searches were re-ran prior to the submission of this thesis are presented in section 3.3.4).

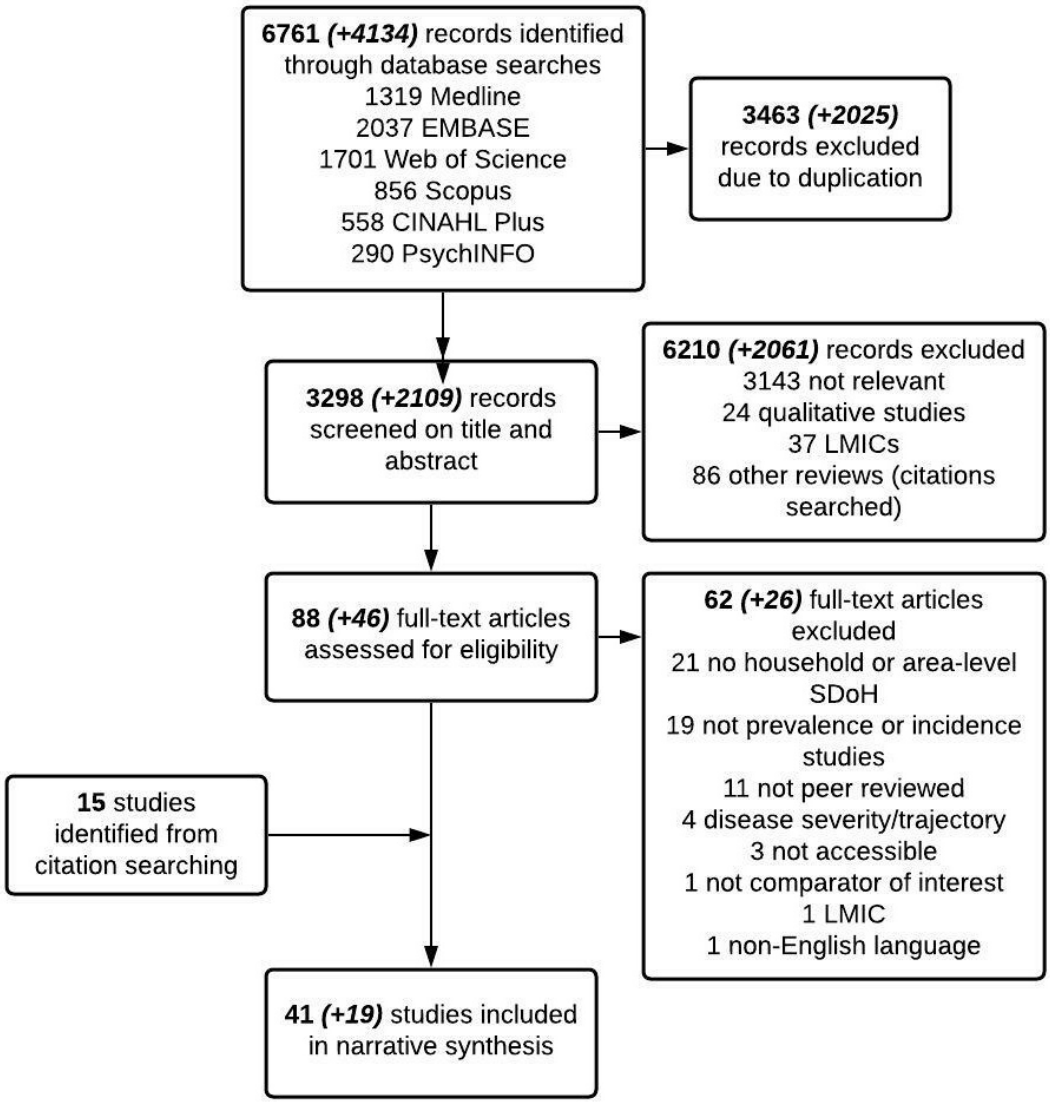
3.3.1 Study selection and characteristics

From the first set of database searches, 41 studies were eligible at full text screen to be included in this systematic review (see Figure 3-1). Inter-rater reliability was good for title and abstract screening ($\kappa = 0.71$), and full-text screening ($\kappa = 0.77$) (McHugh, 2012).

Table 3-2 gives an overview of key study characteristics. Studies were conducted in a range of countries - six studies were conducted in Canada, six in England, five in Spain, four in Scotland and the remaining conducted in countries including Australia, Sweden and Israel. Two studies were conducted across multiple countries. 27 studies were cross-sectional and 14 were cohort studies. Sample sizes ranged from 232 to over 13 and a half million, and the ages of participants varied greatly. 25 studies included participants from across the life-course, while 10 focused on older adults aged 50 and over and three focused on adults aged 30 and over. Three studies focused on middle-aged participants (30-64 years old) and one study was unclear about the age range of participants included (Johnston, Black, *et al.*, 2019).

Household SDoH included measures of household income, tenure, and composition, self-reported by participants in all studies. Area-level SDoH included measures of socioeconomic deprivation and rurality, the former measured using validated indices (16/17 studies) and polling data. Appendix 3 presents the full data extraction table for each of the initially identified studies.

Figure 3-1: PRISMA flow diagram



Note: Numbers in italics give the results of the second screen conducted in May 2021 to update the results of this review prior to submission of this thesis. LMIC = Low-middle income country; SDoH = social determinant of health.

Table 3-2: Key study characteristics

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Studies initially included in review (n=41)									
Agborsangaya (2012)	Canada	Cross-sectional	4980	≥18	Household: income, composition	Self-report	Presence of 2 or more chronic conditions	16	Self-report
Agborsangaya (2013)	Canada	Cross-sectional	4803	≥18	Household: income	Self-report	Concurrent occurrence of 2 or more chronic conditions in the same individual	16	Self-report
Arbelle (2014)	Israel	Cross-sectional	1972798	0-85+	Area-level: socioeconomic deprivation	Poverty index	2 or more of these morbidities in one patient	40	EHRs screened
Bahler (2015)	Switzerland	Cross-sectional	229493	≥65	Area-level: socioeconomic deprivation	Polling data	2 or more chronic conditions in one person	22	EHRs screened

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Barnett (2012)	Scotland	Cross-sectional	1751841	0-85+	Area-level: socioeconomic deprivation	Carstairs index	2 or more morbidities in one patient	40	EHRs screened
Cantarero-Prieto (2018)	Multiple	Longitudinal	31536	.≥50	Household: composition Area-level: rurality	Interviewed (no further details)	3 or more chronic diseases	14	Self-report
Cassell (2018)	England	Longitudinal	403985	≥18	Area-level: socioeconomic deprivation	IMD (year unclear)	2 or more currently active long-term conditions	36	EHRs screened
Charlton (2013)	England	Longitudinal	282887	≥30	Area-level: socioeconomic deprivation	IMD (2010)	Dual (2 conditions) and triple (3) morbidity	5	EHRs screened
Chung (2015)	Hong Kong	Cross-sectional	25780	≥15	Household: income, tenure	Self-report	2 or more chronic health conditions	46	Self-report

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Foguet-Boreu (2014)	Spain	Cross-sectional	1749710	≥19	Area-level: rurality	Assigned by researcher	Coexistence of 2 or more chronic diseases	146	EHRs screened
Hayek (2017)	Israel	Cross-sectional	4325	≥21	Household: income	Self-report	2 or more physician-diagnosed conditions	10	Self-report
Henchoz (2019)	Switzerland	Longitudinal	4055	65-70	Household: childhood financial hardship, composition	Self-report	Co-occurrence of 2 or more medical conditions	13	Self-report
Humphreys (2018)	England	Longitudinal	1979	64-68	Household: paternal social class at birth	Self-report	Total number of multi-morbid conditions	10	Self-report
Johnson-Lawrence (2017)	USA	Cross-sectional	115097	30-64	Household: income, tenure	Self-report	2 or more conditions	9	Self-report

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Johnston (2019)	Scotland	Longitudinal	6561	N/A	Household: paternal social class at birth	Self-report	2 or more self-reported conditions	N/A	Self-report
Katikireddi (2017)	Scotland	Longitudinal	10083	18-75	Household: income Area-level: socioeconomic deprivation	Self-report and Carstairs index	2 or more of the relevant conditions	40	Self-report
Ki (2017)	Korea	Longitudinal	9971	≥30	Household: relative poverty (based on income)	Self-report	Number of chronic diseases	66	Self-report
Laires (2018)	Portugal	Cross-sectional	15196	25-79	Household: income	Unclear	2 or more of these chronic conditions	13	Self-report

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Lebenbaum (2018)	Canada	Cross-sectional	288300	≥18	Household: income, tenure Area-level: rurality	Self-report	At least 2 chronic conditions	10	Self-report
Li (2016)	England	Cross-sectional	27806	16-85	Area-level: socioeconomic deprivation	IMD (2010)	At least 2 of the listed conditions	12 (+ 'other')	Self-report
Lujic (2017)	Australia	Longitudinal	90352	≥45	Household: income, language Area-level: rurality	Self-report	2 or more chronic conditions	8	Self-report + EHRs screened
Melis (2014)	Sweden	Longitudinal	390	≥75	Household: composition	Self-report	Co-occurrence of 2 or more chronic conditions	38	Clinician report + EHRs screened
McLean (2014)	Scotland	Cross-sectional	1272685	≥25	Area-level: socioeconomic deprivation	Carstairs index	Coexistence of 2 or more chronic conditions	40	EHRs screened

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Moin (2018)	Canada	Cross-sectional	12516587	≥50	Area-level: socioeconomic deprivation	ON-Marg Index	Co-occurrence of 2+ (and 3+) chronic conditions	18	EHRs screened
Mounce (2018)	England	Longitudinal	5564	≥50	Household: composition	Self-report	2 or more conditions	15	Self-report
Neilsen (2017)	Multiple	Cross-sectional	63842	≥50	Household: income	Self-report	Coexistence of 2 or more chronic conditions	12	Self-report
Orueta (2013)	Spain	Cross-sectional	452698	≥65	Area-level: socioeconomic deprivation	Deprivation index	Co-occurrence of 2 or more (or 3 or more) health problems	47	EHRs screened
Orueta (2013)	Spain	Cross-sectional	2262286	0-75	Area-level: socioeconomic inequality	Deprivation index	Number of chronic conditions	52	EHRs screened

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Orueta (2014)	Spain	Cross-sectional	2262698	0-85	Area-level: deprivation	Deprivation index	Coexistence of 2 or more conditions in the same patient	52	EHRs screened
Prazeres (2015)	Portugal	Cross-sectional	1993	≥18	Household: income, composition Area-level: rurality	Self-report	Presence of ≥2 or ≥3 chronic health problems	147[b]	Self-report + EHRs screened
Roberts (2015)	Canada	Cross-sectional	105416	≥20	Household: income, education level Area-level: rurality	Self-report (unclear for rurality)	2 or more, and 3 or more, chronic diseases	9	Self-report
Ryan (2018)	Canada	Cross-sectional	13581191	0-105	Area-level: socioeconomic deprivation	ON-Marg Index	Presence of 3 or more chronic conditions	17	EHRs screened

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Salisbury (2011)	England	Longitudinal	99997	≥18	Area-level: socioeconomic deprivation	Townsend index	More than 1 chronic condition	17 (+114[b])	EHRs screened
Schäfer (2012)	Germany	Longitudinal	3189	65-84	Household: income, tenure, composition	Self-report	Number of chronic conditions	29	EHRs screened
Sinnott (2015)	Ireland	Cross-sectional	2047	50-69	Household: dysfunction in childhood (e.g., divorce)	Self-report	2 or more chronic diseases	20	Self-report
Stanley (2018)	New Zealand	Cross-sectional	3489747	≥18	socioeconomic	NZDep index (2013)	At least 2 conditions from 2 different condition lists	61 and 30	EHRs screened
Stokes (2018)	New Zealand	Cross-sectional	232	≥35	Area-level: socioeconomic deprivation	NZDep index (year unclear)	Presence of 2 or more morbidities in one patient	31	EHRs screened

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Tomasdottir (2016)	Norway	Longitudinal	20365	20-59	Area-level: distrust in neighbours	Self-report	2 or more coinciding chronic diseases within the same individual	17	Self-report + EHRs screened
Tucker-Seeley (2011)	USA	Longitudinal	7305	50-75+	Household: childhood financial hardship	Self-report	Count of chronic conditions	6	Self-report
Verest (2019)	Netherlands	Cross-sectional	22362	18-70	Household: income	Self-report	2 or more chronic diseases	21	Self-report
Violan (2014)	Spain	Cross-sectional	1356761	≥19	Area-level: socioeconomic deprivation	Deprivation index	Coexistence of 2 or more chronic conditions	146[b]	EHRs screened

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Further studies identified from rerunning searches before thesis submission (n=19)									
Ashworth (2019)	*England*	*Longitudinal*	*332353*	*≥18*	*Area-level: socioeconomic deprivation*	*IMD (2015)*	*Three or more long term conditions*	*12*	*EHRs screened*
Basham (2019)	*Canada*	*Cross-sectional*	*110924*	*≥12*	*Household: income, education level*	*Self-report*	*3 or more chronic conditions*	*14*	*Self-report*
Dorrington (2020)	*England*	*Longitudinal*	*326415*	*16-60*	*Area-level: socioeconomic deprivation*	*IMD (2010)*	*Presence of two or more long term conditions*	*15*	*EHRs screened*
Chamberlain (2020)	*USA*	*Cross-sectional*	*198941*	*≥20*	*Area-level: socioeconomic deprivation*	*Area deprivation index*	*≥2 (and ≥5) chronic conditions*	*21*	*EHRs screened*
Chudasama (2019)	*England*	*Longitudinal*	*502611*	*38-73*	*Area-level: socioeconomic deprivation*	*Townsend deprivation index*	*Two or more chronic conditions*	*36*	*Self-report*

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Ferry (2020)	*Northern Ireland*	*Longitudinal*	*878345*	*25-64*	*Household: car availability, tenure and rateable property value* *Area-level: rurality*	*Self-report*	*At least two of the self-report conditions*	*11*	*Self-report*
Keats (2020)	*Canada*	*Cross-sectional and longitudinal*	*15215*	*35-69*	*Area-level: neighbourhood walkability*	*CAN-ALE Index*	*≥2 self-reported chronic conditions*	*5*	*Self-report*
Kone (2021)	*Canada*	*Cross-sectional*	*12929733[c]*	*0-65+*	*Area-level: socioeconomic deprivation*	*ON-Marg Index*	*Co-occurrence of 2 or more conditions*	*18*	*EHRs screened*
Kim (2020)	*South Korea*	*Cross-sectional*	*68950*	*≥19*	*Household: income*	*Self-report*	*Two or more chronic conditions in one person*	*28*	*Self-report*

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Low (2019)	*Singapore*	*Cross-sectional*	*1181024*	*0-85+*	*Household: income*	*Government subsidy scheme*	*Two or more chronic conditions concurrently in an individual*	*48*	*EHRs screened*
Mbuya-Bienge (2021)	*Canada*	*Longitudinal*	*5316830*	*≥18*	*Area-level: socioeconomic deprivation*	*Area-level deprivation index*	*2 or more chronic conditions and count of conditions*	*31*	*EHRs screened*
Newman (2019)	*USA*	*Cross-sectional*	*76186*	*≥18*	*Household: income*	*Self-report*	*Count of chronic conditions*	*12*	*Self-report*
Rolewicz (2020)	*England*	*Cross-sectional*	*199150*	*16-85+*	*Area-level: socioeconomic deprivation*	*IMD (2015)*	*Count of long term conditions*	*16*	*Self-report*
Singer (2019a)	*England*	*Longitudinal*	*56202*	*50+*	*Household: wealth*	*Self-report*	*2 or more (and 3 or more) morbidities and 3 or more body systems affected by disease*	*24*	*Self-report*

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Singer (2019b)	*England*	*Longitudinal*	*65386*	*50+*	*Household: wealth*	*Self-report*	*2 or more morbidities and 3 or more body systems affected by disease*	*24*	*Self-report*
Shang (2020)	*Australia*	*Longitudinal*	*53867*	*45-64*	*Area-level: rurality, socioeconomic deprivation*	*Accessibility Remoteness Index and Index of Relative Socioeconomic Disadvantage*	*Co-existence of ≥2, ≥3, and ≥ 4 chronic conditions*	*11*	*Self-report*
Sreedhar (2019)	*New Zealand*	*Cross-sectional*	*375*	*0-75+*	*Area-level: socioeconomic deprivation*	*NZDep (2013)*	*Two or more long-term health conditions*	*38*	*EHRs screened*
St John (2021)	*Canada*	*Cross-sectional*	*19971*	*45-85*	*Household: income*	*Self-report*	*3 or more chronic conditions and count of chronic conditions*	*31*	*Self-report*

First Author (Year)	Country	Study Design	Participants		SDoH Exposure(s)		Multimorbidity Outcome(s)		
			No.	Age range	SDoH Investigated	Method of data collection	Definition(s)[a]	No. of conditions	Method of data collection
Wister (2020)	*Canada*	*Cross-sectional*	*16313*	*45-85*	*Area-level: socioeconomic deprivation*	*VANDIX (2016)*	*Two or more chronic health conditions*	*37*	*Self-report*

Note: SDoH = social determinants of health; EHRs = electronic health records; IMD = Index of Multiple Deprivation; ON-Marg = Ontario marginalisation index; NZDep = New Zealand's deprivation index; VANDIX = Vancouver Area Neighbourhood Deprivation Index; CAN-ALE Index = Canadian Active Living Environments Index

[a]Definition(s) of multimorbidity are taken as direct quotes from each paper. [b]Defined using O'Halloran et al.'s criteria for chronicity (O'Halloran, Miller and Britt, 2004). [c]Mean sample size from across three repeated cross sections.

3.3.2 Defining and measuring multimorbidity

For each of the 41 studies initially identified in January 2019, Table 3-2 outlines the multimorbidity definition(s) and outcome(s) used, and the method used for ascertaining information on multimorbidity.

Most of the initially identified studies (34/41) defined multimorbidity as two or more conditions taken from a pre-specified list of "long-term" or "chronic" conditions. Two further studies defined multimorbidity as three or more conditions taken from a pre-specified list, whilst six of the 34 studies used both definitions and included two multimorbidity outcomes in their analyses. Seven studies used a count of chronic conditions as one outcome to define multimorbidity. One study also defined a 'complex multimorbidity' outcome, which they defined as "three or more chronic conditions affecting three or more bodily systems" (Lujic *et al.*, 2017). However, this study did not look at associations between their chosen SDoH and this multimorbidity outcome (Lujic *et al.*, 2017).

Across the 41 initially identified studies, the number of conditions included on the pre-specified list ranged from five to 146 diagnostic clusters defined using O'Halloran et al.'s criteria for chronicity (O'Halloran, Miller and Britt, 2004). 36 of the 41 studies included a mix of chronic physical and mental health conditions on their pre-specified list, whilst four studies included only physical conditions (Tucker-Seeley *et al.*, 2011; Hayek *et al.*, 2017; Johnson-Lawrence, Zajacova and Sneed, 2017; Cantarero-Prieto, Pascual-Sáez and Blázquez-Fernández, 2018) and one study was unclear about the conditions they included (Johnston, Black, *et al.*, 2019). 20 studies used health data self-reported by participants, 17 studies screened EHRs and three studies used a combination of the two methods (Prazeres and Santiago, 2015; Tomasdottir *et al.*, 2016; Lujic *et al.*, 2017). One study used a combination of physician determination and screening of EHRs to determine the presence of multimorbidity (Melis *et al.*, 2014).

3.3.3 Study results

Household income (n=15): From the 41 initially identified studies, 13 consistently found that multimorbidity prevalence or incidence was markedly and negatively associated with household income and, of all SDoH investigated, associations were

consistently strongest for household income (Agborsangaya *et al.*, 2012, 2013; Schäfer *et al.*, 2012; Roberts *et al.*, 2015; Chung *et al.*, 2015; Nielsen, Halling and Andersen-Ranberg, 2017; Hayek *et al.*, 2017; Johnson-Lawrence, Zajacova and Sneed, 2017; Katikireddi *et al.*, 2017; Ki *et al.*, 2017; Lujic *et al.*, 2017; Lebenbaum *et al.*, 2018; Laires and Perelman, 2019). Higher quality studies reported comparatively small estimated effect sizes, for example Agborsangaya et al. reported that an annual household income <$30,000 CAD was associated with a 2.39-fold increase in multimorbidity prevalence (95% CI 1.72-3.33) compared with >=$100,000 CAD, after multiple adjustments (Agborsangaya *et al.*, 2012). In contrast, Roberts et al. – a lower quality study – reported chances of multimorbidity 4.4 times higher for participants with the lowest level of income compared to the highest in multivariate analyses (OR 4.4, 95% CI 3.6-5.5). Roberts et al reported greater odds of multimorbidity amongst 35–49 year olds compared with over 65s for those with the lowest income vs the highest (OR 7.5, 95% CI 4.0-13.7 vs OR 2.5, 95% CI 1.8-3.5, respectively) (Roberts *et al.*, 2015).

Two further studies – of low and moderate quality, respectively - examined problems managing household income and reported mixed results (Prazeres and Santiago, 2015; Verest *et al.*, 2019). Verest et al. found that individuals who self-reported “lots of problems” were over five times likely to have multimorbidity compared to those with “no problems” (OR 5.36, 95% CI 4.88-5.88). Inequalities varied slightly by gender and ethnicity. After adjusting for age, Dutch men who self-reported “lots of problems” managing their income had nearly 4.5 times the odds of having multimorbidity than those reporting “no problems” (OR 4.48, 95% CI 2.76-7.29), and this figure increased to nearly 7 times the odds for Dutch women (OR 6.82, 95% CI 4.47-10.41). For Moroccan residents, ethnic minorities in the Netherlands, estimates were similar for men (OR 5.05, 95% CI 3.54-7.22) and lower for women (OR 3.96, 95% CI 2.89-5.44) (Verest *et al.*, 2019). In contrast, Prazeres et al. found no evidence of an association between problems managing household income and multimorbidity prevalence when screening EHRs (Prazeres and Santiago, 2015).

See Table 3-3 for key results and quality assessments for studies investing household income.

Table 3-3: Key results and quality assessments for studies investigating household income

First Author (Year)	Key Results				Risk of Bias[e]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Agborsangaya (2012)	Yes	OR 2.39[f] (1.72-3.33)	Annual income <$30k vs. ≥$100k CAD	Age, sex, education, living with children	H	M	M	L
Agborsangaya (2013)	Yes	OR 2.9 (2.2-3.7)	Annual income <$30k vs. ≥$100k CAD	Age, sex, education, obesity	H	H	M	L
Basham (2019)	*Yes*	*OR 0.39[h] (0.35-0.44)*	*Annual income ≥$80k vs. <$20k CAD*	*Age, sex, physical activity, smoking status, household education, alcohol consumption, fruit and vegetable consumption*	*L*	*M*	*M*	*L*
Chung (2015)	Yes	OR 1.52 (1.39-1.66, P<.001)	Monthly income <4k vs. >40k HKD	Age, gender, education, housing, employment	H	M	M	L
Hayek (2017)	Yes	PRR 1.7 (1.2-2.5, P=.005)	Monthly income ≤$2k vs. >$4k USD	Unclear	U	H	H	U

First Author (Year)	Key Results				Risk of Bias[e]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Johnson-Lawrence (2017)	Yes	OR 1.45 (1.38-1.53)	Lowest income tertile vs. highest	Age, gender, ethnicity, education, interview year, region, marital status, last doctor visit, employment, home ownership	U	M	H	L
Katikireddi (2017)	Yes	OR 1.53 (1.25-1.87, P<.05)	Lowest income[b] tertile vs. highest	Age, age^2, age^3, sex, cohort, prior multimorbidity, time between waves and sex*cohort	M	M	M	L
Ki (2017)	Yes	OR 3.48[a] (3.20-3.78)	"Poor" (less than half the median annual household income[b]) vs. "non-poor"	No adjustment	U	H	M	H
Kim (2020)	*Yes*	*35.4% (34.1-36.6) vs. 11.5% (10.9-12.1)*	*Lowest income quartile vs. highest*	*No adjustment*	*L*	*H*	*M*	*H*
Laires (2018)	Yes	OR 2.16[a] (1.95-2.40)	Lowest income[b] quintile vs. highest	No adjustment	L	H	M	H

First Author (Year)	Key Results				Risk of Bias[e]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Lebenbaum (2018)	Yes	OR 0.57 (0.52-0.62, P<.001)	Highest income[b] quintile vs. lowest	Age, age^2, sex, marital status, immigration status, education, rurality, homeownership, smoking, alcohol use	L	M	H	L
Low (2019)	*Yes*	*OR 2.83[a,i] (2.80-2.86)*	*Monthly income ≤$1100 vs. >$1800 SGD*	*No adjustment*	*L*	*L*	*L*	*H*
Lujic (2017)	Yes	OR 0.58[c] (95% CI 0.52-0.66)	Income >$70k vs. <$20k CAD	Age, sex	H	M	M	M
Neilsen (2017)	Yes	OR 1.44 (1.32-1.59, P<.05)	Lowest income tertile vs. highest	Age, sex, education	U	H	M	L
Newman (2019)	*Yes*	*61.9% vs. 43.9% (P<.001)*	*Annual income of <$25k vs. ≥$50k USD*	*No adjustment*	*H*	*H*	*M*	*H*
Prazeres (2015)	No	OR 0.8[d] (0.5-1.1, P=0.182)	"Some monthly income left over" vs. "Not enough to make ends meet"	Age, sex, marital status, education, profession, residence area, living arrangement	H	M	L	L

First Author (Year)	Key Results				Risk of Bias[e]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Roberts (2015)	Yes	OR 4.4[d] (3.6-5.5)	Lowest income quintiles vs. highest	Age, sex, household education, Aboriginal status, activity level smoking, stress, blood pressure, obesity	H	M	H	M
Schäfer (2012)	Yes	-0.27 conditions (-0.47 to -0.08, P=0.005)	Change per unit on income[b] scale (one unit = one of steps: €400 to €1,100 to €3,000 to €8,100 net income per month)	Age, gender, marital status, job autonomy, household composition, income	H	M	L	U
Singer (2019a)	*Yes*	*OR 1.47 (1.34-1.61)*	*Lowest wealth tertile vs. highest*	*Age, sex, wave, subjective social status, last occupation, education, social engagement, loneliness, social support, sense of control, physical activity, alcohol use, smoking status*	*U*	*H*	*M*	*L*
Singer (2019b)	*Yes*	*OR 1.9[j]*	*Lowest wealth quintile vs. highest*	*No adjustment*	*U*	*H*	*M*	*H*

First Author (Year)	Key Results				Risk of Bias[e]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
St John (2021)	*Yes*	*OR 3.77[a,k] (3.29-4.32)*	*Annual income of <$20k vs. ≥$100k CAD*	*No adjustment*	*H*	*H*	*M*	*H*
Verest (2019)	Yes	OR 5.36[a,g] (4.88-5.88)	"Lots of problems" managing money vs. "no problems"	No adjustment	H	H	M	H

Note: SDoH = social determinant of health; MM = multimorbidity; OR = odds ratio; PRR = prevalence ratio; information in italics gives the results for studies identified when searches were re-ran in May 2021 prior to submission of this thesis.

[a]OR calculated from data reported in paper. [b]Income equivalised to account for number and/or age of residents in household. [c]Based on self-reported health data. Findings consistent across hospital and medication health data. [d]Multimorbidity defined as ≥3 chronic conditions. [e]H = high, M = medium, L = low, U = unclear.

Differences by subgroup: [f]Inequalities greater for ages 25-44, [g]Inequalities greater for women and similar by ethnicity group, [h]Inequalities similar for men and women, [i]Inequalities greater <65 and ≥85 years of age, [j]Inequalities greater for age 50-54, [k]Inequalities greater <65 years of age.

Household composition (n=7): From the 41 initially identified studies, four measured household composition as living alone versus cohabiting (Melis *et al.*, 2014; Cantarero-Prieto, Pascual-Sáez and Blázquez-Fernández, 2018; Mounce *et al.*, 2018; Henchoz *et al.*, 2019) and three studies measured it as living alone, living with various family members, or living in other situations (including care homes) (Agborsangaya *et al.*, 2012; Schäfer *et al.*, 2012; Prazeres and Santiago, 2015).

Four cohort studies of older adults (aged 50-84 years old) reported mixed findings on the risk of living alone versus cohabiting (Melis *et al.*, 2014; Cantarero-Prieto, Pascual-Sáez and Blázquez-Fernández, 2018; Mounce *et al.*, 2018; Henchoz *et al.*, 2019). Two high quality studies found living alone increased chances of multimorbidity versus living with others (Cantarero-Prieto, Pascual-Sáez and Blázquez-Fernández, 2018; Henchoz *et al.*, 2019), for example Cantarero-Prieto et al. found living alone increased chances of multimorbidity by 20% (OR 1.20, 95% CI 1.04-1.39, P<0.05) (Cantarero-Prieto, Pascual-Sáez and Blázquez-Fernández, 2018). Whereas two other studies – high and moderate quality - found no evidence living alone was associated with multimorbidity incidence (Melis *et al.*, 2014; Mounce *et al.*, 2018). Differences in study characteristics such as methods of ascertaining multimorbidity presence could not explain these mixed findings.

Of the three studies with alternative measures of composition, one moderate quality cross-sectional study by Agborsangaya et al. (2012) found odds of multimorbidity were over two times greater if not living with children versus living with children (OR 2.11, 95% CI 1.60-2.78; adjusted for age, sex, education and household income) (Agborsangaya *et al.*, 2012). When stratified by age, odds of multimorbidity were 2 times higher amongst 25-44 year olds (OR 2.00, 95% CI 1.29-3.02) and 45-64 year olds (OR 1.96, 95% CI 1.30-2.95), and more than 8 times higher for those aged 65+ years (OR 8.45, 95% CI 2.02-35.41). Agborsangaya et al. (2012) found no evidence that the risk of multimorbidity differed for those living with adults compared to those not living with adults in univariate analyses (Agborsangaya *et al.*, 2012). Two further moderate quality studies (one of which included solely older adults) found no evidence of any associations with multimorbidity when living alone was compared with living as a couple, with family/others or living in situations such as care homes (Schäfer *et al.*, 2012; Prazeres and Santiago, 2015).

See Table 3-4 for key results and quality assessments for studies investigating household composition.

Household tenure (n=4): From the 41 initially identified studies, four moderate quality studies (three cross-sectional and one cohort) investigated associations between tenure and multimorbidity (Schäfer *et al.*, 2012; Chung *et al.*, 2015; Johnson-Lawrence, Zajacova and Sneed, 2017; Lebenbaum *et al.*, 2018). Findings from these studies were mixed and hard to compare given different reference groups and comparators; two studies compared homeowners and non-homeowners, one compared renters with homeowners and one compared social housing residents with homeowners, private renters and subsidised housing residents.

Of the two studies comparing homeowners and non-homeowners, Lebenbaum et al. found the odds of multimorbidity decreased by 18% for homeowners compared to non-homeowners (OR 0.82, 95% CI 0.78-0.87, P<0.001) (Lebenbaum *et al.*, 2018), whereas Schäfer et al. (2012) found no evidence of an association between homeownership and multimorbidity (Schäfer *et al.*, 2012). Johnson-Lawrence et al. reported 19% higher odds for renters versus homeowners after multiple adjustments (OR 1.19, 95% CI 1.15-1.24) (Johnson-Lawrence, Zajacova and Sneed, 2017). Differences in study characteristics such as multimorbidity definitions and measurements could not explain these mixed findings.

Chung et al. (2015) found weak evidence that, compared to participants living in public (social) housing in Hong Kong, the odds of multimorbidity were greater for those living in subsidized housing and privately rented accommodation by 11% and 19%, respectively (OR 1.11, 95% CI 1.05-1.18, P=0.070, and OR 1.19, 95% CI 1.09-1.29, P=0.041). The odds of multimorbidity for those who privately owned their properties were 17% greater compared to those in public (social) housing (OR 1.17, 95% CI 1.11-1.24, P=0.003). All models were adjusted for age, gender, education, household income, and employment (Chung *et al.*, 2015).

See Table 3-4 for key results and quality assessments for studies investigating household tenure.

Table 3-4: Key results and quality assessments for studies investigating household composition and tenure

First Author (Year)	Key Results				Risk of Bias[b]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Household composition								
Agborsangaya (2012)	Yes	OR 2.11[c] (1.60-2.78)	Living with children vs. not living with children	Age, sex, education, household income	H	M	M	L
	No	Data not available	Living with adults vs. not living with adults					
Cantarero-Prieto (2018)	Yes	OR 1.20 (1.04-1.39, P<.05)	Living alone vs. cohabits	Unclear	U	U	M	U
Henchoz (2019)	Yes	OR 1.40[a] (1.21-1.61)	Living alone vs. cohabits	No adjustment	U	M	M	M
Melis (2014)	No	OR 1.34 (0.60-3.01)	Living alone vs. cohabits	No adjustment	U	M	L	H
Mounce (2018)	No	HR 0.93 (0.71-1.21, P=.580)	Living alone vs. cohabits	Baseline age, sex, total wealth, education, health behaviours, social detachment, locus of control	U	M	M	L

First Author (Year)	Key Results				Risk of Bias[b]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Prazeres (2015)	No	OR 1.4 (0.9-2.3, P=0.182) OR 1.0 (0.6-1.7, P=0.985) OR 1.3 (0.7-2.6, P=0.410)	Living as a couple vs. alone Living as extended family vs. alone Living in other situation (inc. care home) vs. alone	Age, sex, marital status, education, professional status, residence area, living arrangement	H	M	L	L
Schäfer (2012)	No	-0.10 conditions (-0.42-0.23, P=0.562) 0.24 conditions (-0.14-0.62, P=0.210) -0.01 conditions (-0.59-0.57, P=0.231)	Living at home with spouse vs. home alone Living at home with family members or others vs. home alone Living in assisted living or retirement home vs. home alone	Age, gender, marital status, job autonomy, household composition, income	H	M	L	U

First Author (Year)	Key Results				Risk of Bias[b]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Household tenure								
Chung (2015)	Yes	OR 1.17 (1.11-1.24, P=.003)	Homeowner vs. public (social) housing	Age, gender, education, housing, employment	H	M	M	L
		OR 1.19 (1.09-1.29, P=0.041)	Private renting vs. public (social) housing					
		OR 1.11 (1.05-1.18, P=0.070)	Subsidized housing vs. public (social) housing					
Ferry (2020)	*Yes*	*OR 1.57[d] (1.52-1.63, P<.001)*	*Homeowners with properties worth <£75k vs. ≥£200k GBP*	*Age, sex, marital status, household car access, rurality, education*	*L*	*M*	*H*	*L*
		OR 2.07[d] (2.00-2.14, P<.001)	*Private renting vs. homeowners with properties worth ≥£200k GBP*					
		OR 3.26 (3.15-3.38, P<.001)	*Social housing vs. homeowners with properties worth ≥£200k GBP*					

First Author (Year)	Key Results				Risk of Bias[b]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Johnson-Lawrence (2017)	Yes	OR 1.19 (1.15-1.24)	Renters vs. homeowners	Age, gender, ethnicity, education, interview year, region, marital status, last doctor visit, employment, household income	U	M	H	L
Lebenbaum (2018)	Yes	OR 0.82 (0.78-0.87, P<.001)	Homeowners vs. non-homeowners	Age, age^2, sex, marital status, immigration status, education, rurality, homeownership, smoking, alcohol use	L	M	H	L
Schäfer (2012)	No	-0.13 conditions (-0.30-0.05, P=0.148	Homeowners vs. non-homeowners	Age, gender, marital status, job autonomy, household composition, income	H	M	L	U

Note: SDoH = social determinant of health; MM = multimorbidity; OR = odds ratio; HR = hazard ratio; information in italics gives the results for studies identified when searches were re-ran in May 2021 prior to submission of this thesis.
[a]OR calculated from data reported in paper. [b]H = high, M = medium, L = low, U = unclear.
Differences by subgroup: [c]Associations greater for 65+, [d]Inequalities greater for women.

Household determinants in childhood (n=5): From the 41 initially identified studies, two examined associations between paternal social class at birth and multimorbidity in adulthood. Findings were mixed. One higher quality study from Johnston et al. found lower paternal social class at birth was associated with increased multimorbidity in middle age (Johnston, Black, *et al.*, 2019). Conversely, one study of lower quality reported no association (Humphreys *et al.*, 2018).

Two studies investigated associations between self-reported childhood financial hardships and multimorbidity and, again, findings were mixed (Tucker-Seeley *et al.*, 2011; Henchoz *et al.*, 2019). One higher quality study by Henchoz et al. (2019) found no evidence of an association between the two (Henchoz *et al.*, 2019), whereas one fairly low-quality study from Tucker-Seeley and colleagues (2011) found strong evidence that the expected number of chronic conditions for those reporting childhood financial hardships was 1.19 times that of those not reporting such hardship (95% CI 1.07-1.32, P<.001) (Tucker-Seeley *et al.*, 2011).

One further, moderate quality study found that the odds of multimorbidity increased by 40% amongst those who had experienced a form of household dysfunction during childhood, such as parental divorce, compared to those who had not in multivariate analyses (OR 1.4, 95% 1.1 to 1.7, P<0.05) (Sinnott *et al.*, 2015).

Household primary language and education (n=2): From the 41 initially identified studies, one moderate quality Australian study from Lujic and colleagues (2017) found that people speaking a language other than English at home had 6% higher odds of having multimorbidity than those who speak English when medication records were screened (OR 1.06, 95% CI 1.01-1.10) and 32% higher odds when hospital data were screened (OR 1.32, 95% CI 1.22-1.42). However, the same participants had 20% lower odds of multimorbidity in self-report data (OR 0.80, 95% CI 0.76-0.84) (Lujic *et al.*, 2017). One lower quality study found higher odds of multimorbidity for participants living in households where all residents had not completed high school compared to those in households where someone had a post-secondary school education (OR 1.8, 95% CI 1.6-2.1, adjusting for age and sex) (Roberts *et al.*, 2015).

Trust in neighbours (n=1): One fairly low quality study found that participants who "somewhat distrusted" their neighbours had increased risk of developing

multimorbidity within 11 years compared with those who strongly disagreed with the statement "One cannot trust each other here" (RR 1.13, 95% CI 1.03-1.23). However, they found no evidence that risk differed for those strongly agreeing this statement versus those strongly disagreeing (Tomasdottir *et al.*, 2016).

See Table 3-5 key results and quality assessments for studies investigating these other household-level SDoH.

Household rurality (n=7): From the 41 initially identified studies, seven investigated associations between residing in rural areas and multimorbidity and these studies also reported mixed results (Foguet-Boreu *et al.*, 2014; Prazeres and Santiago, 2015; Roberts *et al.*, 2015; Lujic *et al.*, 2017; Cantarero-Prieto, Pascual-Sáez and Blázquez-Fernández, 2018; Lebenbaum *et al.*, 2018; Ryan *et al.*, 2018).

Two high quality studies that provided clear rurality definitions both suggested odds of multimorbidity decreased with increased rurality (Foguet-Boreu *et al.*, 2014; Ryan *et al.*, 2018). For example, Foguet-Boreu et al. (2014) found evidence suggesting that women aged 45-64 had 20% lower odds of multimorbidity if they live in rural areas (OR 0.80, 95% CI 0.78-0.82, P<0.001), and men of the same age had 13% lower odds (OR 0.87, 95% CI 0.85-0.89, P<0.001). However, the variables they adjusted for were unclear (Foguet-Boreu *et al.*, 2014).

Two further studies – one low and one moderate quality – reported greater odds of multimorbidity with increased rurality (Roberts *et al.*, 2015; Lujic *et al.*, 2017). Lujic et al. (2017) found that, after adjusting for age and sex, the odds of multimorbidity were higher for those living in remote/very remote areas compared to major cities. This was seen when data on multimorbidity was self-reported (OR 1.14, 95% CI 1.03-1.26), obtained from screening medication data (OR 1.11, 95% CI 1.00-1.23) and obtained from hospital data (OR 1.28, 95% CI 1.08-1.53) (Lujic *et al.*, 2017). Three further studies found no evidence of any association between rurality and multimorbidity (Prazeres and Santiago, 2015; Cantarero-Prieto, Pascual-Sáez and Blázquez-Fernández, 2018; Lebenbaum *et al.*, 2018). Aside from study quality, differences in study characteristics could not explain these mixed findings.

See Table 3-6 for key results and quality assessments for studies investigating rurality.

Table 3-5: Key results and quality assessments for studies investigating other household-level SDoH

First Author (Year)	Key Results				Risk of Bias[a]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Household education								
Basham (2019)	*Yes*	*OR 0.92 (0.86-0.99)*	*Highest household education post-secondary vs. non-post-secondary*	*Age, sex, physical activity, smoking status, household education, alcohol consumption, fruit and vegetable consumption*	*L*	*M*	*M*	*L*
Roberts (2015)	Yes	OR 1.8[b] (1.6-2.1)	No one completed high school vs. someone with post-secondary school education	Age and sex	H	M	H	M
Household primary language								
Lujic (2017)	Yes	OR 0.80[c] (0.76-0.84)	English not primary language at home vs. is primary language	Age and sex	H	M	M	M

First Author (Year)	Key Results				Risk of Bias[a]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Paternal social class at birth								
Humphreys (2018)	No	OR 1.15 (0.93-1.43, P>0.01)	Manual social class at birth vs. non-manual	Age, gender, health behaviours, time in cohort, year of recruitment	H	L	H	L
Johnston (2019)	Yes	OR 1.74 (1.11-2.72)	Father in a skilled manual occupation vs. unemployed/disabled/dead	Gender, education, cognition at age 7, school type				
Childhood financial hardships								
Henchoz (2019)	No	OR 0.94 (0.74-1.19)	Poor family economic environment: yes vs. not	Sex, cohort, socioeconomic status, behaviours, other stressful events in childhood and adulthood	U	M	M	M
Tucker-Seeley (2011)	Yes	IRR 1.19 (1.07-1.32, P<.001)	Childhood financial hardship: yes vs. no	Age, gender, race, education, lifetime earnings, lifetime earnings*childhood financial hardship	U	H	H	L

First Author (Year)	Key Results				Risk of Bias[a]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Childhood household dysfunction								
Sinnott (2015)	Yes	OR 1.4 (1.1-1.7, P<.05).	Childhood household dysfunction: yes vs. no	Age, gender, education, income, behaviour factors, depression and anxiety scores	H	M	M	L
Trust in neighbours								
Tomasdottir (2016)	No	RR 1.13 (0.98-1.32)	Strongly distrusted neighbours vs. strongly trusted	Age, gender, smoking, physical activity, education, and current depressive symptoms	H	H	M	L
Household car availability								
Ferry (2020)	*Yes*	*OR 1.84[d] (1.79-1.88, P<.001)*	*No vs. two or more cars accessible*	*Age, sex, marital status, household car access, rurality, education*	*L*	*M*	*H*	*L*

Note: SDoH = social determinant of health; MM = multimorbidity; OR = odds ratio; RR = rate ratio; IRR = incidence rate ratio; information in italics gives the results for studies identified when searches were re-ran in May 2021 prior to submission of this thesis.
[a]H = high, M = medium, L = low, U = unclear, [b]Multimorbidity defined as ≥3 chronic conditions, [c]Based on self-reported health data. Higher odds in medication/hospital data (e.g., hospital: OR 1.32, 95%CI 1.32-1.42).
Differences by subgroup: [d]Inequalities greater for men.

Table 3-6: Key results and quality assessments for studies investigating rurality

First Author (Year)	Key Results				Risk of Bias[c]			
	Association between SDoH and MM?	Value (95% CI, P value)	Comparator	Adjusted for…	Selection	Information (Exposure)	Information (Outcome)	Confounding
Cantarero-Prieto (2018)	No	OR 0.92 (0.93-1.03, P>0.1)	Living in rural vs. non-rural areas	Unclear	U	U	M	U
Ferry (2020)	*Yes*	*OR 0.97[f] (0.95-0.99, P<.005)*	*Living in rural vs. urban areas*	*Age, sex, marital status, household car access, rurality, education*	*L*	*M*	*H*	*L*
Foguet-Boreu (2014)	Yes	OR 1.04[a,e] (1.03-1.05)	Living in rural (<10,000 inhabitants and/or population density <150 people/km^2) vs. non-rural areas	Unadjusted	U	L	L	U
Lebenbaum (2018)	No	OR 0.98 (0.93-1.02, P=0.323)	Rural vs. non-rural areas	Age, age^2, sex, marital status, immigration status, education, rurality, homeownership, smoking, alcohol use	L	M	H	L
Lujic (2017)	Yes	OR 1.14[b] (1.03-1.26)	Living in remote/very remote areas (vs. major cities)	Age and sex	H	M	M	M

First Author (Year)	Key Results				Risk of Bias[c]			
	Association between SDoH and MM?	**Value (95% CI, P value)**	**Comparator**	**Adjusted for…**	**Selection**	**Information (Exposure)**	**Information (Outcome)**	**Confounding**
Prazeres (2015)	No	OR 1.0 (0.8-1.3, P=0.746)	Living in rural vs. urban areas	Age, sex, marital status, education, profession, residence area, living arrangement	H	M	L	L
Roberts (2015)	Yes	OR 1.1 (1.0-1.3)	Living in rural vs. urban areas	Age, sex, household education and income, Aboriginal status, activity level, smoking, stress, blood pressure, obesity	H	M	H	M
Ryan (2018)	Yes	OR 0.85[a] (0.85-0.86)	Living in rural (<10,000 inhabitants) vs. non-rural areas	Age-sex standardised	L	L	L	M
Shang (2020)	*Yes*	*PAR 1.2 and 0.8[d]*	*Living in outer regional or remote vs. urban areas*	*Unclear*	*H*	*L*	*L*	*U*

Note: SDoH = social determinant of health; MM = multimorbidity; OR = odds ratio; PAR = population attributable risk; information in italics gives the results for studies identified when searches were re-ran in May 2021 prior to submission of this thesis.

[a]OR calculated from data reported in paper. [b]Based on self-reported health data. Findings consistent across hospital and medication health data. [c]H = high, M = medium, L = low, U = unclear.

Differences by subgroup: [d]Men and women respectively (no further information provided), [e]Inequalities similar with gender and greater≥45 years, [f]Inequalities greater for women than men.

Area-level socioeconomic situation (n=17): From the 41 initially identified studies, 17 investigated how the socioeconomic situation of a participants' residential area was associated with multimorbidity prevalence or incidence (see Table 3-2 for key study characteristics). Findings from these studies were fairly consistent - participants residing in areas with greater socioeconomic deprivation were more likely to be multimorbid than those living in the most affluent areas. Appendix 3 shows the key results from each of these 17 studies. Odds of multimorbidity prevalence were 42% higher for participants residing in the most versus the least deprived areas when available data were pooled (OR 1.42, 95% CI 1.41-1.42). Differences in study quality could not explain differences in reported effect sizes across studies.

12 of these studies investigated differences in associations between measures of deprivation and multimorbidity by age, sex, and ethnicity (see Appendix 3). 10 studies found that differences in multimorbidity prevalence with area deprivation reduced in older age. Four studies found inequalities with area-level deprivation were greater for women than for men (Orueta, García-Álvarez, *et al.*, 2013; Orueta, Nuño-Solinís, *et al.*, 2013; Orueta *et al.*, 2014; Violán *et al.*, 2014). Only one low quality study investigated how associations differed by ethnicity and they found inequalities were greater for Pacific versus Maori New Zealand residents (Stokes, Azam and Noble, 2018).

Five studies specifically examined differences in associations between deprivation and different types of multimorbidity: physical-mental multimorbidity, physical-only multimorbidity and multimorbidity comprised of mental health conditions only (mental-only multimorbidity). Findings across these studies were fairly consistent; differences between the most and least deprived areas were greater in middle age for basic multimorbidity (Barnett *et al.*, 2012; Arbelle *et al.*, 2014; McLean *et al.*, 2014; Orueta *et al.*, 2014; Violán *et al.*, 2014; Li *et al.*, 2016; Cassell *et al.*, 2018; Stanley *et al.*, 2018), greater for physical-mental multimorbidity and mental-only multimorbidity in younger age groups (McLean *et al.*, 2014), and differences were greater in older age groups for physical-only multimorbidity (McLean *et al.*, 2014). More details can be found in Appendix 3.

3.3.4 Rerunning database searches prior to thesis submission

3.3.4.1 Study characteristics

The second set of database searches were ran in May 2021 and a further 19 studies were identified that would have been eligible for inclusion in this review (see Figure 3-1). Table 3-2 gives an overview of the key characteristics of these studies, in italics. Studies were conducted in a range of countries: six studies were conducted in England, six in Canada and two in the United States. Sample sizes ranged from 375 to over twelve million participants. 10 studies were cross-sectional, eight were cohort studies and one study used both cross-sectional and longitudinal analyses. 10 studies included participants from across the life-course, while six focused on adults spanning mid-life to older age and three focused on working age adults.

Of these 19 studies, 10 investigated how the socioeconomic situation of a participants' residential area was associated with multimorbidity prevalence or incidence (Ashworth *et al.*, 2019; Chudasama *et al.*, 2019; Sreedhar, Richard and Stokes, 2019; Chamberlain *et al.*, 2020; Dorrington *et al.*, 2020; Rolewicz *et al.*, 2020; Shang *et al.*, 2020; Wister *et al.*, 2020; Kone *et al.*, 2021; Mbuya-Bienge *et al.*, 2021). Seven studies investigated household income (Basham and Karim, 2019; Low *et al.*, 2019; Newman, Levine and Kishore, 2019; Singer, Green, Rowe, Ben-Shlomo and Morrissey, 2019; Singer, Green, Rowe, Ben-Shlomo, Kulu, *et al.*, 2019; Kim, Keshavjee and Atun, 2020; St John *et al.*, 2021), two rurality (Ferry *et al.*, 2020; Shang *et al.*, 2020) and one study investigated household education level (Basham and Karim, 2019). A further two studies examined SDoH not yet investigated by studies included in this review: household car availability, a combination of housing tenure and rateable value of a property, and area-level walkability of a neighbourhood (Ferry *et al.*, 2020; Keats *et al.*, 2020). See Table 3-2 for more details.

3.3.4.2 Defining and measuring multimorbidity

For each of the 19 studies identified in the second set of database searches, Table 3-2 outlines the multimorbidity definition(s) and outcome(s) used, and the method used for ascertaining information on multimorbidity.

14 out of the 19 studies included a multimorbidity outcome defined as two or more conditions taken from a pre-specified list of "long-term" or "chronic" conditions. Five

of these studies included additional multimorbidity outcomes, such as a count of chronic conditions, five or more chronic conditions or “three or more body systems affected by disease”. A further four studies solely defined multimorbidity as three or more conditions taken from a pre-specified list, or solely defined multimorbidity as a count of chronic conditions.

Across these 19 studies, the number of conditions included on the pre-specified list ranged from five to 48 conditions. 16 of the 19 studies included a mix of chronic physical and mental health conditions on their pre-specified list, whilst two studies included only physical conditions (Wister *et al.*, 2020; St John *et al.*, 2021) and one study was unclear about the conditions they included (Rolewicz *et al.*, 2020). 12 studies used health data self-reported by participants whilst seven studies screened EHRs.

3.3.4.3 Study results

Studies that investigated associations between multimorbidity and area-level socioeconomic deprivation, household income, rurality and household education level reported findings that were consistent with those reported in section 3.3.3.

Ferry and colleagues (2020) investigated associations between household car availability and multimorbidity prevalence (Ferry *et al.*, 2020). They found that individuals living in households without access to a car had 84% greater odds of multimorbidity compared to those with two or more cars accessible, after adjustment for age, sex, marital status, rurality, education and a combination of tenure and property value (OR 1.84, 95% CI 1.79-1.88, P<.001). Inequalities were greater for men than women (men: OR 2.02, 95% CI 1.95-2.09, P<.001, women: OR 1.69, 95% CI 1.64-1.75, P<.001). Inequalities in physical-mental multimorbidity with car accessibility were greater for both men and women compared to physical-only multimorbidity.

In separate analyses, Ferry and colleagues (2020) created eight categories to reflect a combination of housing tenure and rateable value of a property: social renting, private renting, ‘properties as yet unvalued’ and, for owner-occupiers, five categories to capture property prices ranging from less than £75000 to over £200000 (Ferry *et al.*, 2020). Individuals living in social housing and privately renting had over three and

two times the odds of multimorbidity, respectively, when compared to individuals who owned and lived in properties worth over £200000 GBP (OR 3.26, 95% CI 3.15-3.38, P<.001 and OR 2.07, 95% CI 2.00-2.14, P<.001, respectively). Homeowners reported greater multimorbidity prevalence with decreasing value of their property. For example, homeowners with properties worth less than £75000 reported over 50% greater odds of multimorbidity compared to homeowners whose properties were worth greater than or equal to £200000 (OR 1.57, 95% CI 1.52-1.63, P<.001). Inequalities were greater for women than men. Inequalities in physical-mental multimorbidity with tenure were greater for both men and women compared to physical-only multimorbidity.

Keats and colleagues (2020) investigated associations between the area-level walkability of a neighbourhood and multimorbidity and found no evidence of an association (Keats *et al.*, 2020).

3.4 Discussion

3.4.1 Summary of findings

This is the first study to systematically identify, critically appraise and synthesise existing literature on associations between household and area-level SDoH and multimorbidity. My review suggests that an array of household and area-level SDoH, and their associations with multimorbidity, have been investigated in the literature. Household income and area-level deprivation were the most explored social determinants, and findings for these were fairly consistent; odds of multimorbidity were up to 4.4 times higher for those within the lowest level of household income (versus the highest), and prevalence was 1.4 times higher in the most versus the least deprived areas. Findings relating to measures of household composition, tenure and rurality were more mixed. Following the initial databases searches, I concluded that, aside from household income and area-level measures of deprivation, other household and area-level SDoH had been underexplored in the literature. Findings from eligible studies published since 2019 suggest this could be somewhat improving.

3.4.2 Comparisons to existing literature

Previous research has proposed that household factors are often overlooked in studies exploring SDoH, despite households (or families) influencing physical and mental health through various material and psychosocial factors (McNeill, 2010; Vaezghasemi *et al.*, 2016).

In this systematic review, seven studies were identified that investigated household composition and these presented mixed results; living alone was associated with increased multimorbidity in two studies and not associated with multimorbidity in four. These studies included different reference groups and comparators, making them hard to compare. For example, ill-health greatly drives care home admissions (Bowman, Whistler and Ellerby, 2004) and therefore comparing “living alone” with either “not living alone” or “living in a care home” could be comparing groups in very different health, leading to differential associations between household composition and multimorbidity. One further study found living with children (versus not) was associated with increased chances of multimorbidity, and this effect was greater for individuals over 65 years old. Chronic illness may give rise to older individuals residing with family and may lead younger individuals unable to (or decide not to) have children. Interestingly, none of the included studies examining household composition adjusted for care provision, which can differ considerably for those living with a partner, family or alone (Hellström and Hallberg, 2004). Care provision could plausibly influence the relationship between household composition and multimorbidity - further research should gather data on care provision and adjust associations accordingly. Unpicking whether social circumstances drive multimorbidity, or vice versa, also requires better designed longitudinal studies. This could aid targeting of resources for prevention.

In this review, four studies were identified in the initial searches that investigated household tenure, as well as a further study when the databases searches were re-ran prior to the submission of this thesis. These five studies also reported contradicting results; homeownership and residing in social housing were associated with both increased and decreased chances of multimorbidity. Comparing these results was, again, complicated by different reference groups and comparators, however, study contexts and research settings may be more pertinent here. These studies were conducted in Hong Kong, Canada, USA, Germany, and Northern

Ireland. The degree of homeownership, and supply and conditions of social housing, may vary across these locations, for example approximately 45% of the Hong Kong population lived in public (social) housing in 2019 compared to 10% of the German population in 2017 (Vetter, 2019; Eurostat, 2020). This, plus other social circumstances, may profoundly influence the status, stigma and meaning associated with owning, renting, or residing in social housing across geographies and over time, differentially impacting health and associations between tenure and multimorbidity (Shaw, 2004; Reeves *et al.*, 2016; Clair and Hughes, 2019). In my conceptualisation of SDoH, the higher order structural factors that may influence these associations (such as housing policies) are classified as acting at the macro-level of influence and were not explicitly searched for in this review (see Figure 1-3). In addition, no studies were identified in this review that explored associations between household tenure and multimorbidity in the English context.

This review identified seven studies in the initial searches, and two in the subsequent searches, that investigated associations between the degree of rurality of one's home area and multimorbidity. Again, findings were contradictory; living in a rural area was associated with increased chances of multimorbidity in three studies, decreased chances of multimorbidity in three studies and not associated in three studies. Only three studies were clear about their definition of rurality, which made the study results hard to compare. Study contexts and therefore factors classified in my conceptualisation of SDoH as acting at the macro-level (see Figure 1-3) may also be pertinent here. Different countries may differ in the degree to which rurality is associated with deprivation and health care access, which may, in turn, directly influence health (Baird and Wright, 2006). Only four of the studies adjusted for individual-level measures of deprivation, such as education, whilst the other four were unadjusted, adjusted for age and sex, unclear about adjustments made or solely presented age-sex standardised results. These differences in analysis methods may have had profound effects on associations reported across the nine studies.

A minority of studies identified in this review examined whether associations differed by age or gender, and only two studies examined differences by ethnicity. Findings suggested women experience greater inequality in multimorbidity prevalence with area-level deprivation and a combination of household tenure and property value, in line with research highlighting an increase in life expectancy inequality for women in

the UK (Marmot *et al.*, 2020). Prevalence with area-level deprivation was greater for younger populations for physical-mental multimorbidity, unsurprising given the consistently high prevalence of mental ill-health amongst young, deprived communities (Bond *et al.*, 2012). Problems managing household income was also differentially associated with multimorbidity depending on the intersections of gender and ethnicity. Yet public health research rarely embraces intersectionality approaches to help better understand have social inequities manifest as health inequities. Future research should examine whether associations differ by key demographics and how these demographics intersect with each other and further SDoH to influence multimorbidity (Hankivsky and Christoffersen, 2008; Bowleg, 2012).

The initial database searches for my systematic review found that 10 studies specifically focused on older aged adults, 29 focused on individuals across the entire span of adult life and only two focused on working aged adults. When the database searches were re-ran prior to submission of this thesis, a further three studies that focused solely on working aged adults were identified. This is in line with the fact that multimorbidity research tends to focus on older populations, given the increased prevalence of multimorbidity in this age group and the associated health and social care implications and costs (Barnett *et al.*, 2012; The Academy of Medical Sciences, 2018). However, whilst the prevalence of multimorbidity increases with age, the absolute number of those with multimorbidity is greater amongst those 65 years and below (Barnett *et al.*, 2012; Rocca *et al.*, 2014; Wang *et al.*, 2015; Bobo *et al.*, 2016). As outlined in Chapter 1 (section 1.2.4), the AMS recommends that future research around multimorbidity places greater focus on younger populations (The Academy of Medical Sciences, 2018). My review findings support this recommendation and suggest that increased focus needs to be paid to working age adults in research examining associations between SDoH and multimorbidity.

A lack of consensus around a definition for multimorbidity is a consistently raised issue in the literature (Johnston, Crilly, *et al.*, 2019). Studies identified from both the initial and follow up searches most often defined multimorbidity as two or more chronic conditions, reflecting the most cited and used definition in the literature (Johnston, Crilly, *et al.*, 2019). However, several of these studies also used a cut-off point of three or more chronic conditions or a count of conditions. To ascertain multimorbidity presence, the included studies often used either self-reported data, data from EHRs

or a combination of the two. This, plus the wide variation in number and type of conditions included in multimorbidity definitions, hampered effective comparisons of study findings. Despite this, I found no evidence suggesting differences in findings could be explained by differences in multimorbidity definitions or measurement methods. There was also no variation in determinants of multimorbidity by measurement methods used. Consistent definitions of multimorbidity and consistent methods for ascertaining its presence are needed to improve comparability of findings, and the results of the second set of database searches suggest this recommendation remains relevant and pertinent. In addition, differences in the prevalence of physical-mental multimorbidity with area-level deprivation suggests studies that exclude mental health conditions from multimorbidity definitions, or that specify multimorbidity as specifically crossing physical and mental health, may report different associations than studies not. Future research should consider physical and mental dimensions of multimorbidity when examining associations between SDoH and multimorbidity.

3.4.3 Strengths and limitations of this study

One key strength of this study is that it is the first to systematically identify, critically appraise and synthesise existing literature on associations between household and area-level SDoH. This means that this study has systematically included household SDoH, which has captured studies missed by previous reviews in this area (Violan *et al.*, 2014; Northwood *et al.*, 2018; Pathirana and Jackson, 2018; The Academy of Medical Sciences, 2018). Further strengths include the investigation of how reported associations differ with differences in multimorbidity definitions and measurement methods, as well as the careful assessment of each study for risk of four dimensions of bias using pre-specified and tailored criteria.

There are several, however, several limitations of this review. First, the term ‘comorbidity’ (and its linguistic variations) was excluded from the search terms despite ‘comorbidity’ being used interchangeably in the literature with the term ‘multimorbidity’(van den Akker, Buntinx and Knottnerus, 1996). This was a pragmatic decision taken to make this review manageable, as described in detail in Appendix 2. Whilst this may have missed some relevant literature, a subsequent ad hoc search in Medline, that included the MeSH term ‘comorbidity’, did not identify any additional, relevant hits.

A further limitation is that a large proportion (15/41, 36.6%) of the initially included 41 studies were identified via citation searching. This is likely to be an intrinsic issue that arises when conducting these types of reviews. In the literature, a SDoH is rarely referred to as a 'social determinant of health' and, instead, is referred to by the name of the determinant of interest (e.g., 'rurality' or 'education level'). Search strategies, therefore, need to pre-specify terms to search for these specific determinants. This may have missed relevant studies if alternative terms that I am unfamiliar with have been used to describe a SDoH in a study. A considerable number of studies were also identified by my second set of database searches, which suggests that studies that would have been eligible for inclusion will likely have been published since the submission of this thesis.

Additional limitations of this review are that the searches were restricted to peer-reviewed, English-language studies conducted in HICs. By excluding grey literature and non-English language publications the search may have missed relevant studies and reports, such as work by Guy's and St Thomas' charity (Guy's and St Thomas' Charity, 2018). I chose to focus on HICs only given evidence suggestion the socioeconomic gradient in multimorbidity is reversed in low and middle-income countries (Kunna, San Sebastian and Stewart Williams, 2017). However, this latter restriction to the search may mean findings from this review are not necessarily generalisable to these settings.

3.5 Chapter summary

This chapter has identified, critically appraised, and synthesised existing literature examining household and area-level social determinants of multimorbidity. I found that, aside from measures of household income, household determinants of multimorbidity are often overlooked in the literature despite comparatively large effect sizes for household compared to area-level SDoH. In particular, none of the identified studies had examined associations between household tenure and multimorbidity in the English context. This review also found that few studies focused solely on working age adults and few studies considered how associations between SDoH and multimorbidity differ by age, gender, and ethnicity. Overall, this review suggests that strategies to better understand and prevent multimorbidity should consider household SDoH, younger populations and the influence of key demographics.

This systematic review has also highlighted how associations between SDoH and multimorbidity can differ depending on whether multimorbidity is defined as two or more chronic conditions taken from a pre-specified list of chronic physical and mental health conditions, defined to capture only physical conditions, or defined to capture both physical and mental health conditions specifically. I found that health inequalities were starker when the latter definition was operationalised. However, the vast majority of studies identified in this area did not consider the physical and mental dimensions of multimorbidity. In addition, few studies examined associations between SDoH and more complex multimorbidity profiles. Multimorbidity outcomes that capture more complex circumstances should be operationalised in future research to develop our understanding of how SDoH are associated with different types of multimorbidity.

The findings from this review, and the subsequent revisions I have made to my conceptualisation of SDoH (described in section 4.1.1.1), have been used to inform the design of the next study in this thesis. This study is a quantitative analysis of a linked health and council dataset and is presented in Chapter 4 and Chapter 5.

Chapter 4 A quantitative analysis of linked health and council data investigating associations between household tenure and multimorbidity: Introduction and Methods

4.1 Introduction

In Chapter 1, I described how multimorbidity is a major public health challenge that requires preventative strategies that address SDoH. In this chapter, I will describe the introduction and methods for the second study conducted as part of this thesis. Using linked health and council data, this study examines and quantifies associations between selected household characteristics (informed by my systematic review) and multimorbidity.

As outlined in Chapter 2, this study aims to illustrate how knowledge generated from the analysis of linked health and council data could advance our understanding of the social determinants of multimorbidity. This study will act as a use case for creating, using, and analysing linked health and council datasets to understand the social determinants of local public health concerns and generate knowledge that could inform equitable decision-making. Multimorbidity and SDoH are issues that should span the organisational boundaries of a health and care system. Both are therefore areas where our understanding could improve if health and council data are linked.

4.1.1 Learnings from my systematic review

In Chapter 3, I presented findings from my systematic review that identified, critically appraised, and synthesised existing literature that has examined associations between household and area-level SDoH and multimorbidity. I have used findings from my systematic review, and the subsequent revisions I have made to my conceptualisation of SDoH, to inform the design of this study.

Firstly, my systematic review found that, aside from measures of household income, household determinants of multimorbidity are often overlooked despite comparatively large effect sizes for household compared to area-level SDoH. My review identified only five studies that examined associations between household tenure and

multimorbidity. These studies reported contradictory results, and not one of these five studies examined associations between tenure and multimorbidity in the English context.

Second, my review found that few studies focused solely on working age adults (between 18 and 64 years old, inclusive). This is despite evidence suggesting that the absolute number of those with multimorbidity is greater in working age, and that the incidence of multimorbidity and socioeconomic inequalities in multimorbidity are rising amongst people of working age (Barnett *et al.*, 2012; Head, Fleming, Kypridemos, Schofield, *et al.*, 2021). This is also despite increasing calls to focus on younger populations (The Academy of Medical Sciences, 2018). For working age adults there may be greater opportunity for prevention of multimorbidity through addressing SDoH than amongst older adults.

Thirdly, my review found that the majority of studies examining associations between household and area-level SDoH and multimorbidity defined multimorbidity using the widely used cut off of two or more chronic conditions taken from a pre-specified list. In contrast, few studies specifically considered physical and mental dimensions of multimorbidity despite findings suggesting that associations between area-level deprivation and multimorbidity differed depending on whether a basic multimorbidity or physical-mental multimorbidity outcome was operationalised (Barnett *et al.*, 2012; McLean *et al.*, 2014). In addition, few studies examined associations between SDoH and more complex multimorbidity profiles, such as the co-occurrence of three or more chronic conditions affecting three or more different bodily systems (Harrison *et al.*, 2014).

This study aims to address some of these gaps identified by my review.

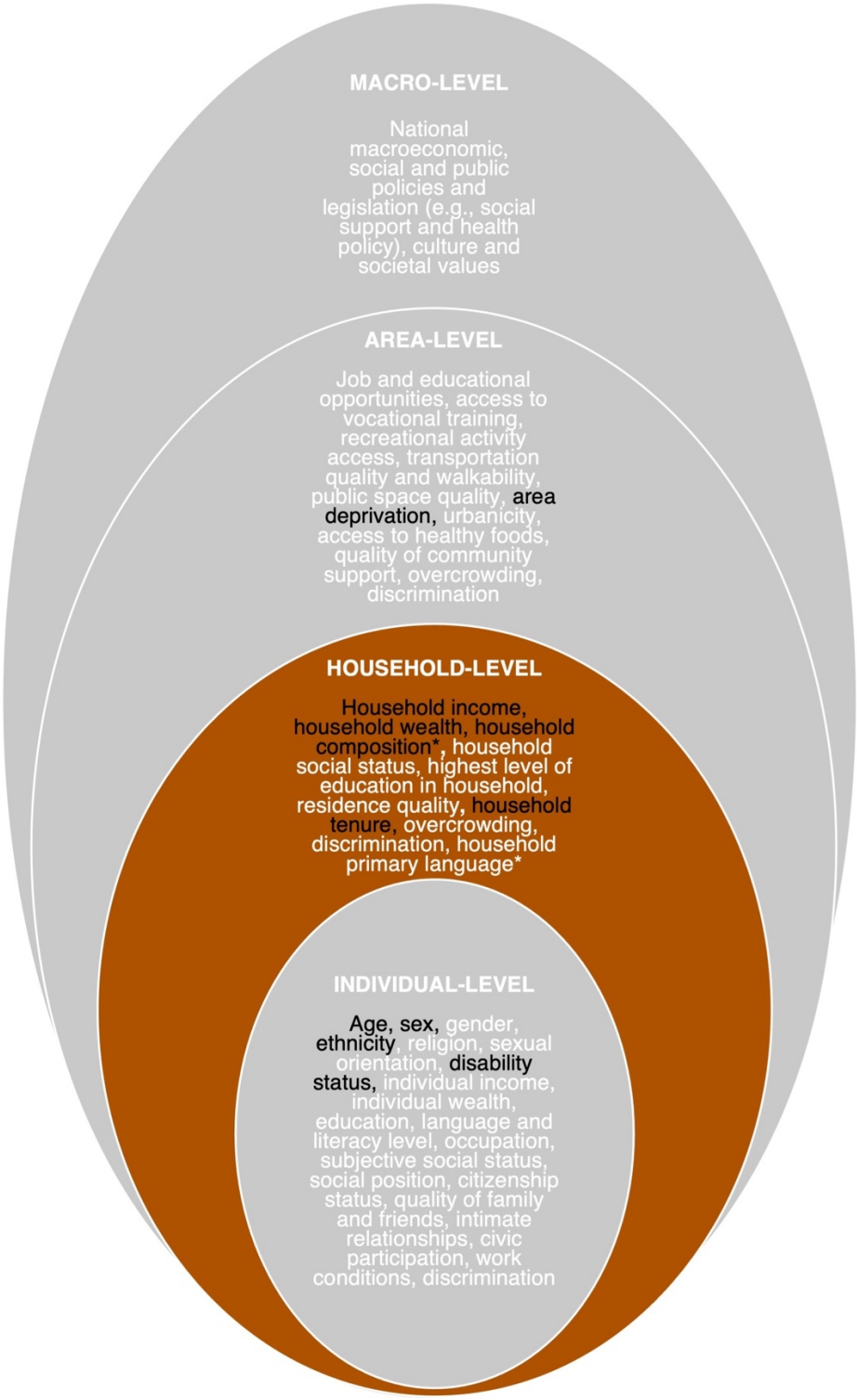
4.1.1.1 A revised conceptualisation of SDoH

In Chapter 1 (section 1.3.2) I presented my conceptualisation of SDoH as an adaption of the WHO's CSDH Framework. My conceptualisation grouped factors as acting at four levels of influence: the individual-level, household-level, area-level, and macro-level (see Figure 1-3). In my systematic review (Chapter 3), I focused on household and area-level social determinants of multimorbidity given that individual-level social determinants, such as age and sex, have been well examined in previous literature

(see section 1.3.3.1). I found that area-level social determinants of multimorbidity have also been well examined, and household-level factors such as household tenure are often overlooked. As such, I have revised my conceptualisation to reflect the focus of this study on household-level factors. My review also suggested that macro-level factors influence the interpretation of household and area-level SDoH in the context of multimorbidity (see section 3.4.2). However, macro-level factors may be less relevant to include in a conceptualisation of SDoH when conducting analyses within one defined and relatively small geographical context (as will be done in this study). As such, I have revised my conceptualisation to reflect my lack of focus on macro-level factors.

In my systematic review (Chapter 3), I found several studies that explored factors that were not included in my initial conceptualisation of SDoH. These have been added to my revised conceptualisation (see Figure 4-1). In addition, information is only collected and available in routine administrative health and care records for selected factors. As such, it was not possible to examine and quantify associations between certain household-level factors and multimorbidity in this study, despite evidence suggesting they are underexplored in the literature.

Figure 4-1: Revised conceptualisation of SDoH used in this thesis



Note: Conceptualisation revised following my systematic review. Factors followed by an asterisk (*) indicate newly identified factors. Factors in black and white indicate those that should and should not be captured in administrative health and care records, respectively.

4.1.2 Household tenure as a social determinant of health

I chose to focus this study on household tenure given gaps identified by my systematic review, however the relationships between household tenure and health are complex. On one hand, some argue that household tenure could be simply considered a marker of one's financial position within society, with those owning their property reflecting the most affluent and those renting from social housing landlords reflecting the most deprived (Hamnett, 2004; André, Dewilde and Muffels, 2019; Anderson, Han and Hisnanick, 2021). Wealth or income then, in turn, directly act as determinants of health.

Alternatively, many argue that household tenure is, in itself, a SDoH by explaining differences in exposure to various household and area-level stressors (Ellaway and Macintyre, 1998). Individuals may be differentially exposed to overcrowding, damp or mould, and other hazardous conditions depending on the tenure they reside in (Ellaway and Macintyre, 1998; Macintyre *et al.*, 2003). For example, in the 2011 English census 15% of all social housing tenants lived in overcrowded households compared to 3% of people who owner-occupied (Whitehead, 2014). In England, tenure types are also differentially distributed across areas depending on neighbourhood characteristics, with a greater proportion of social housing found in less affluent areas (Livingston, Kearns and Bailey, 2013). As a consequence, individuals living in different household tenures may be differentially exposed to stressors afforded by their neighbourhood environment, such as limited employment opportunities and high levels of crime.

Household tenure has been linked to poor physical and mental health. For example, living in rented housing (particularly socially rented) reduces a tenant's agency to prevent and fix issues such as damp and mould when compared to individuals who own their properties (Scanlon and Kochan, 2011). Exposure to different hazards such as damp, mould and cold, that may result from tenure type, has, in turn, been linked to various respiratory conditions, myocardial infarction, raised blood pressure and cholesterol, and excess winter mortality deaths (Shaw, 2004; Fowler *et al.*, 2015). Recent work also recognises the influence of psychological housing factors. Strong evidence has been found for a causal relationship between the affordability of housing and mental health in UK households (Reeves *et al.*, 2016) and evidence suggests the effects of tenure-type may be reflected in biological markers of inflammation or stress

(Clair and Hughes, 2019). Shaw refers to the former, more material, factors as “hard” factors and the latter as “soft”. Soft housing factors include feelings of security, social position and pride afforded by the home, which can, in turn, influence mental health (Shaw, 2004). These hard and soft factors may be differentially distributed across tenure types.

Whilst the WHO’s CSDH Framework recognises household tenure as a SDoH by stating that tenure can act as a proxy measure of socioeconomic status, it also states that tenure gives an indicator of living standards and describes how tenure indirectly affects health via neighbourhood effects (Solar and Irwin, 2013). This framework, therefore, reflects Ellaway and Macintyre’s arguments that tenure affects health through exposure to various household and area-level stressors (Ellaway and Macintyre, 1998). As a result, household tenure is now widely considered to be a social determinant that influences both physical and mental health through exposure to health-harming and health-protecting factors that act at multiple levels of influence. These factors could plausibly interact to cause or exacerbate different long-term conditions and, ultimately, multimorbidity. Household tenure is also a suitable SDoH to explore in this study as data on tenure is systematically recorded in council data and expected to have good completeness and validity. This is because councils administer council tax and are responsible for council-ran social housing. In addition, tenure is a SDoH that affects almost all members of a population.

4.1.3 Research questions and objectives

In Chapter 2, I stated that my objective for this study was to examine and quantify associations between selected household characteristics (informed by my systematic review) and multimorbidity. As described in section 4.1.2, I chose to focus on household tenure given the gaps I identified in the literature, given the potential importance of tenure as a SDoH in the English context, and given that data on tenure is available and likely to be well-recorded in council data.

This study, therefore, has the following research questions:

1. What is the prevalence of different definitions of multimorbidity amongst working age adults captured in a linked health and council dataset?

2. How is household tenure associated with multimorbidity amongst working age adults captured in this linked dataset?
3. How do associations differ given different multimorbidity definitions?

Contributions of others to this study: Dr Jenny Shand, Care City and UCL Partners, Amber Gibney, Care City, and Dr Dan Lewer, Care City and UCL Department of Epidemiology and Public Health, kindly provided access to the dataset. Dr Jenny Shand, Phil Canham, London Borough of Barking and Dagenham, and Simon Lam, NHS East London CCG, kindly provided information to help me understand the data origins, definitions, and flows. Dr Melvyn Jones (MJ), UCL Primary Care and Population Health, Dr Helen McDonald (HM), LSHTM Department of Infectious Disease Epidemiology, Professor David Osborn, UCL Division of Psychiatry, members of the Care City Community Board and ARC North Thames patient and public involvement panel, advised on my definitions of multimorbidity.

4.2 Methods

4.2.1 Context and background to the Care City Cohort

For this study, I used a linked dataset that covers residents of the London Borough of Barking and Dagenham (LBBD) – The Care City Cohort. LBBD is a densely populated and socially deprived borough in North East London, with a younger and more ethnically diverse population compared to the rest of England (London Borough of Barking and Dagenham, 2018). In mid-2020, the total population was estimated to be around 214107 residents, with 135749 people (63.4% of the total population) of working age (defined in this study as between 16 and 64 years old, inclusive) (Barking and Dagenham Council, 2020).

Care City, a partner to East London's health and care system, launched a linked dataset to the research community in January 2020. This dataset, called the Care City Cohort, was created as part of a wider project to better understand population service use and support the delivery of more integrated care locally (Shand, 2020). The Care City Cohort contains health information linked at the individual-level to social care and socio-demographic information from council records for residents of LBBD. The dataset contains individual characteristics (e.g., age and sex), health information

covering five settings of care – primary, hospital, community, mental health, and social care - and information about living situations (e.g., household tenure, household occupancy and levels of area deprivation).

The complete dataset was created by Barking and Dagenham, Havering, and Redbridge (BHR) Clinical Commissioning Group (CCG) staff. There are ten different datasets linked in the Care City Cohort and available for researchers to request access to. Each dataset contains a snapshot of variables on 1st April each data year to avoid in-year changes in key variables.

To create the completed Care City Cohort, all NHS datasets were linked at the individual-level by NHS number and accuracy checks were conducted by BHR CCG on age, sex, and address (Shand, 2020). For council information on tenure, occupancy, education and benefits, fuzzy logic matching was applied using first name, surname, date of birth and postcodes to identify individuals. These individuals were then assigned an NHS number and council records were linked to NHS records. Once linked, the dataset was de-identified: names and addresses were removed, and each individual was given two identification codes to replace their NHS number and their Unique Property Reference Number (UPRN). Dates of birth were also replaced with year of birth and dates of death were replaced with month and year of death. Records that could not be linked with confidence were retained in the dataset.

The BHR Information Governance group oversaw the linkage of the dataset and ensured all processes were in line with legal requirements. The relevant ethical approvals were obtained for the creation of the dataset (Shand, 2020). Individuals were able to opt out of data sharing via the BHR CCG website and their data were removed at source. The different datasets were stored and linked within the BHR accredited Data Safe Haven, which complies with all necessary information governance processes and has the infrastructure to store, manage and link the data. The de-identified linked dataset is stored and accessed within the BHR CCG accredited Data Safe Haven.

Researchers can only access the de-identified data required for their analyses, in line with governance confidentiality requirements. Different datasets are provided to researchers *unlinked* with linkage keys i.e., with the identification codes generated to

replace NHS numbers and UPRNs. These can then be used by researchers to link data at individual or household-levels (e.g., social information split across primary care and council records). This enables researchers to evaluate linkage error rates and potential biases as a result of linkage success.

To conduct my analyses, I was given an honorary contract with BHR CCG and access to the BHR accredited Data Safe Haven. BHR CCG staff deposited the data I had requested into a folder I had access to. The data variables I requested access to were informed by my systematic review (Chapter 3) and my knowledge of the multimorbidity literature. Practical consideration was given to the data available within the Care City Cohort and the quality of data available.

4.2.2 Creating the linked health and council dataset

I requested access to all five years of primary care and council data available (2015-2020). The 2018/19 and 2019/20 primary care cohorts contained a more extensive list of chronic conditions. Only the 2019/20 council cohort had data on household tenure. I had planned to create a cross section of the population for 2018/19 and 2019/20 to increase sample size, however Care City had not finished collating and cleaning the 2018/19 council cohort data when I started the study. I therefore solely utilised the 2019/20 primary care and council data.

First, I linked the individual and household-level council data on Household_ID (the household-level identification code created by Care City to replace UPRNs). Second, I linked the individual-level primary care data to the linked council data on Patient_ID (the individual-level identification code created by Care City to replace NHS numbers). Third, I linked a fourth dataset provided by Care City that detailed care homes in LBBD and their Household_IDs. I linked this to the cohort data on Household_ID. Finally, I linked a fifth dataset from ONS that contained area-level deprivation data from 2019. I linked this dataset to the cohort data on LSOA_code. All linkages were conducted in R software using the merge function from the R base package.

To assess whether there were any potential selection biases in the linkage results I calculated standardised differences in key variables for matched and unmatched primary care records (Harron *et al.*, 2017). Selection biases occur when the study population systematically differs from the population of interest (Zaccai, 2004). I was

able to assess potential biases in variables from the primary care data. I was not able to assess potential biases in social variables extracted from council records (i.e., in the household tenure variable and other household variables) as, by definition, unmatched primary care records did not have corresponding council data.

4.2.3 Data variables: reformatting variables and data cleaning

Administrative datasets were of variable quality in terms of their completeness, accuracy, and relevance to this study. The majority of data cleaning was conducted by Care City with the creation of the Care City Cohort. I reformatted different variables as appropriate for analyses and examined missing data. Below I describe the reformatting, standardisation, and handling of missingness undertaken by Care City and the further data cleaning I undertook for this study.

4.2.3.1 Multimorbidity outcomes

The 2019/20 primary care data contained flags to indicate the presence or absence of 38 chronic conditions. These flags were extracted from an individuals' primary care record, in which a frontline practitioner may have recorded the results of diagnostic assessments or other medical tests as part of care delivery. Each flag was derived by Care City using validated and publicly available code lists from the Primary Care Unit at University of Cambridge (Cambridge, 2018). I compared the definitions and code lists of each of the 38 conditions with other published definitions and code lists to assess the consistency in conditions captured in multimorbidity definitions across the literature (see Table 4-1).

As can be seen in Table 4-1, there is considerable variation in the number and type of conditions included definitions of multimorbidity. This is because there remains considerable debate around what should be included as a chronic condition in a definition of multimorbidity (see section 1.2.1). If we consider Care City's list of 38 conditions, an individual with constipation will typically present to primary care with symptoms that are relatively mild and can be easily managed. Conversely, an individual with diabetes will typically present with more serious symptoms that require increased consultation time and greater efforts to manage. Including both constipation and diabetes in a definition of multimorbidity that counts chronic conditions will not capture differences in disease severity, complexity or burden for individuals and

health services. To account for differences in disease severity and complexity, several indices have been developed that weight conditions based on an outcome of interest, for example on past healthcare utilisation or mortality risk (Charlson *et al.*, 1987; Payne *et al.*, 2020). However, none of the indices in the literature utilise outcomes relevant to this study.

Given this debate around conditions to include in a multimorbidity definition, I consulted with two clinicians (MJ and DO), a public health consultant (HM) and a panel of patient and public involvement (PPI) representatives regarding Care City's list of 38 conditions. I convened the PPI panel specifically for this study - it included members of the ARC North Thames Research Advisory Panel and members of the Care City Community Board. In these conversations, I sought their views on the meaning and value of being tagged with each disease label, what each flag means for a multimorbidity definition, and asked each individual to reflect on whether each of the 38 conditions would likely be:

- Poorly recorded in primary care data or recorded with bias
- Of less or little burden to individuals and health services relative to other conditions

Different individuals raised concerns about different conditions. Table 4-2 presents the outcomes of these consultations. Given the outcomes of these consultations, I decided to conduct a series of sensitivity analyses. These sensitivity analyses are described in section 4.2.5.1.

When planning this study, I had initially intended to group the 38 conditions available in the Care City data by bodily system and examine:

- Concordant multimorbidity, defined as two or more chronic conditions affecting a single bodily system
- Discordant physical multimorbidity, defined as two or more chronic physical conditions affecting two or more different bodily systems (e.g., respiratory conditions and gastrointestinal conditions)

- Discordant physical-mental multimorbidity, defined as two or more chronic conditions affecting two or more different bodily systems, one of which must be a mental health condition (Willadsen *et al.*, 2018).

This approach followed the premise set out in the AMS report, which reasons that it is easier to treat multimorbidity when the two or more conditions stem from the same bodily system than when they stem from different systems (see section 1.2.1).

I presented these study plans at an ARC North Thames Multimorbidity Theme meeting in February 2020, which was attended by academics and clinicians across North London who work on multimorbidity research. My plans were criticised by attendees with clinical backgrounds, who argued that a patient presenting with two or more conditions from the same cluster (such as COPD and bronchiectasis, which would be grouped in the "lung" bodily system) can be more difficult to treat and face poorer health and social outcomes than a patient presenting with two or more conditions across bodily systems (chronic sinusitis and constipation, for example). AMS's ideas about ease of treatment differing with multimorbidity that is concordant versus discordant were therefore not agreed upon in the group, however the importance of capturing complexity in my definitions was still recognised by all. I, therefore, decided to not pursue distinguishing concordant and discordant multimorbidity and instead sought to operationalise a definition of 'complex' multimorbidity. Ultimately, I chose to operationalise three multimorbidity outcomes in this study to attempt to capture different degrees of complexity in multimorbidity:

- Basic multimorbidity, defined as the co-occurrence of two or more chronic conditions within a single individual.
- Physical-mental multimorbidity, defined as the co-occurrence of two or more chronic conditions within a single individual, one of which must be depression or anxiety and one of which must be a physical condition.

I chose to focus on depression and anxiety as opposed to any mental health condition as the origins of these conditions may differ than for severe mental illness, as may the social determinants of these conditions.

- Complex multimorbidity, defined as the co-occurrence of three or more chronic conditions affecting three or more different bodily systems within a single individual.

Table 4-3 shows how I grouped each of the 38 chronic conditions by bodily system. My groupings were guided by previous research (Willadsen *et al.*, 2018; Singer, Green, Rowe, Ben-Shlomo, Kulu, *et al.*, 2019). Three binary variables were created in the dataset to indicate the presence or absence of each multimorbidity outcome for each individual.

Table 4-1: Comparison of the 38 chronic conditions used in this study and their code lists with lists published in the literature

Chronic condition category	Barnett *et al.*, 2012	Rocca *et al.*, 2014	Roberts *et al.*, 2015	Tonelli *et al.*, 2016	Willadsen *et al.*, 2018	Kingston *et al.*, 2018	Nicholson, *et al.*, 2019	Stokes *et al.*, 2021	Head, *et al.*, 2021
Alcohol problems	X	/	--	X	X	--	--	X	X
Anorexia or bulimia	X	--	--	--	X	--	--	--	/
Anxiety & other neurotic, stress related & somatoform disorders	X	--	--	--	X	--	/	--	X
Asthma (currently treated)	X	X	/	X	X	/	/	X	X
Atrial fibrillation (AF)	X	X	--	X	X	--	/	X	X
Blindness and low vision	X	--	--	--	X	X	--	--	X
Bronchiectasis	X	X	--	--	X	/	--	X	X
Cancer - [New] Diagnosis in last five years	X	X	/	X	X	X	X	X	X
Chronic kidney disease	/	/	--	X	X	--	X	X	X
Chronic Liver Disease and Viral Hepatitis	X	/	--	--	/	--	/	X	X
Chronic sinusitis	X	--	--	--	--	--	--	--	X
Constipation (Treated)	X	--	--	X	--	--	--	X	--
Chronic obstructive pulmonary disease (COPD)	X	X	/	X	X	/	/	X	X
Coronary heart disease	X	X	/	X	X	X	X	/	X
Dementia	X	X	/	X	X	X	X	X	X

Chronic condition category	Barnett *et al.*, 2012	Rocca *et al.*, 2014	Roberts *et al.*, 2015	Tonelli *et al.*, 2016	Willadsen *et al.*, 2018	Kingston *et al.*, 2018	Nicholson, *et al.*, 2019	Stokes *et al.*, 2021	Head, *et al.*, 2021
Depression	X	X	--	X	X	X	/	X	X
Diabetes	X	X	/	X	X	X	X	X	X
Diverticular disease of intestine	X	--	--	--	--	--	/	--	X
Epilepsy (currently treated)	X	--	--	X	X	--	--	X	X
Hearing loss	X	--	--	--	X	X	--	--	X
Heart failure	X	X	--	X	/	--	X	X	X
Hypertension	X	X	--	X	--	X	X	X	X
Inflammatory bowel disease	X	--	--	X	X	--	/	X	X
Irritable bowel syndrome	X	--	--	X	/	--	/	X	X
Learning disability	X	--	--	--	--	--	--	--	X
Migraine	X	--	--	--	X	--	--	--	X
Multiple sclerosis	X	--	--	X	X	--	--	X	X
Painful condition	X	--	--	X	--	--	--	X	--
Parkinson's disease	X	--	--	X	X	--	--	X	X
Peptic Ulcer Disease	--	--	--	X	--	--	/	X	--
Peripheral vascular disease	X	--	--	X	--	--	/	X	X
Prostate disorders	X	--	--	--	--	--	--	--	X
Psoriasis or eczema	X	--	--	/	/	--	--	/	X
Psychoactive substance misuse (not alcohol)	X	/	--	--	X	--	--	--	X

Chronic condition category	Barnett *et al.*, 2012	Rocca *et al.*, 2014	Roberts *et al.*, 2015	Tonelli *et al.*, 2016	Willadsen *et al.*, 2018	Kingston *et al.*, 2018	Nicholson, *et al.*, 2019	Stokes *et al.*, 2021	Head, *et al.*, 2021
Rheumatoid arthritis, other inflammatory polyarthropathies & systematic connective tissue disorders	X	/	/	X	X	/	X	X	X
Schizophrenia (and related non-organic psychosis) or bipolar disorder	X	/	--	/	X	--	--	/	X
Stroke & transient ischaemic attack	X	X	/	X	/	/	X	X	X
Thyroid disorders	X	X	--	--	X	--	X	/	X

Note: the conditions' definitions and/or code lists were similar or identical across studies (X); the conditions' definitions and/or code lists were somewhat similar across studies (/); no similar condition definitions and/or code lists were identified across studies (--)

Table 4-2: Results of consultations conducted with clinicians, experts and the patient and public panel

Condition(s) captured by a given flag	Issues/concerns raised through consultation	Who raised issue/concern
Chronic kidney disease (CKD)	Stages 1, 2 and 3 of CKD are often asymptomatic and mild when compared to stages 4 and 5.	MJ, HM, DO
	Testing for CKD is not universal and is indicated for people with some medications or conditions (such as diabetes). Therefore, an individual with one chronic condition may be more likely to be diagnosed with CKD.	MJ, HM
	Some individuals with stages 4 and 5 of CKD will not be captured in primary care data as most of these will be managed in hospital.	HM
Psoriasis or eczema	These are conditions that are often mild and, whilst burdensome for some individuals, will most often require little consultation in primary care and will be relatively easy to manage.	MJ, HM
	These are conditions that will likely exhibit recording bias due to difficulties recognising and diagnosing these on darker skin tones.	PPI
Chronic sinusitis	This is a condition that is often mild and that, whilst burdensome for some individuals, will most often require little consultation in primary care and will be relatively easy to manage.	MJ, HM
	This is a condition with limited treatments and, as such, diagnosis may be delayed.	PPI
Constipation (treated)	This is a condition that is often mild and that, whilst burdensome for some individuals, will most often require little consultation in primary care and will be relatively easy to manage.	MJ, HM
	This is a condition that may be caused by medications taken to treat other conditions.	PPI

Condition(s) captured by a given flag	Issues/concerns raised through consultation	Who raised issue/concern
Diverticular disease of intestine	This is a condition that is often mild and that, whilst burdensome for some individuals, will most often require little consultation in primary care and will be relatively easy to manage.	MJ
Prostate disorder	This is a condition that is often mild and that, whilst burdensome for some individuals, will most often require little consultation in primary care and will be relatively easy to manage.	MJ
Bronchiectasis and COPD	Many individuals with a flag of one of these conditions will have received a diagnosis of the other condition as the origins of both conditions are the similar. Including these conditions as separate flags may lead to misclassifying individuals as having multimorbidity.	MJ, HM
Anorexia, depression, and anxiety	Anorexia is often incorrectly first diagnosed as depression and/or anxiety and flags of depression and/or anxiety may not be subsequently removed from a person's record. Including these conditions as separate flags may lead to misclassifying individuals as having multimorbidity.	MJ, DO, HM
Atrial fibrillation (AF)	Most people with AF are asymptomatic. This is a risk factor for certain chronic conditions rather than a chronic condition in and of itself.	MJ MJ, DO
Hypertension	This is a risk factor for certain chronic conditions rather than a chronic condition in and of itself.	MJ, DO

Table 4-3: The 38 chronic conditions grouped by 10 bodily systems

Respiratory
Asthma (currently treated)
Bronchiectasis
Chronic obstructive pulmonary disorder
Sensory
Blindness and low vision
Chronic sinusitis
Hearing loss
Psoriasis or eczema
Cardiovascular
Atrial fibrillation
Coronary heart disease
Heart failure
Hypertension
Peripheral vascular disease
Endocrine
Diabetes
Thyroid disorders
Cancer
Cancer (in last 5 years)
Muscoskeletal
Painful conditions
Rheumatoid arthritis (or other inflammatory polyarthropathies & systematic connective tissue disorders)
Mental health
Alcohol problems
Anorexia & bulimia
Anxiety (& other neurotic, stress related & somatoform disorders)
Depression
Dementia
Psychoactive substance misuse
Schizophrenia & bipolar
Neurological
Epilepsy (currently treated)
Learning disability
Migraine
Stroke & transient ischaemic attack
Multiple sclerosis
Parkinson's disease
Genitourinary
Chronic kidney disease
Prostate disorders
Gastrointestinal
Chronic liver disease and viral hepatitis
Constipation (treated)
Diverticular disease of intestine
Irritable bowel syndrome
Inflammatory bowel disease
Peptic ulcer disease

4.2.3.2 Household status and household tenure

In the dataset, individuals were clustered within households. Data on household clustering were extracted from council records. Care City generated household identifiers based on each household's UPRN as part of the creation of the Care City cohort. Individuals from the same household shared the same Household_ID and this variable was complete in my analyses.

Data on household tenure was extracted from council property records. Care City captured individuals in privately rented properties if their households were recorded in the council's private rented licensing system. Individuals in social housing properties were captured if their households were recorded in the council's social housing data system. Individuals in the boroughs Reside programme - a LBBD-specific housing programme that provides those in employment with access to affordable rental properties (at 65 to 80% below the market rate) – were captured through the Reside programmes' data system. Care City assumed that any individuals who had not been grouped into the former three categories were owner-occupiers. Care City also created a fifth category of "unknown" when the council did not hold any information on the tenure status of a property and when the assumption of owner-occupying was not viable. To assess possible errors or misclassification biases in this variable, I compared the variable's breakdown to mid-2019 ONS tenure estimates for LBBD. ONS population estimates are based on annual population modelling of 2011 census data and account for changes in births, deaths, and national migration (Office for National Statistics, 2019b).

For this study, households were grouped into three categories for tenure: owner-occupied, social housing, and privately rented. Individuals in households in the Reside programme were grouped with social housing residents given the small number of observations in this category and the similarity between Reside residents and social housing residents. I excluded those with "unknown" household tenure.

4.2.3.3 Individual-level sociodemographic covariates

Age: Data on age are routinely recorded in primary care. For each individual, Care City extracted data on year of birth from their primary care record. This was instead of extracting data on date or month of birth for pseudonymisation purposes.

To calculate age, I subtracted the year of birth from 2020 giving the maximum age an individual could be at the time of data extraction. I created eight categories to code an individuals' age in years (<16, 16-29, 30-44, 45-54, 55-64, 65-74, 75-84, 85+). Those born between April and December 2004 would therefore be aged 15 at the time of data extraction yet categorised in the 16-29 years old age category. Those born between April and December 1990 would therefore be aged 29 at the time of data extraction yet categorised in the 30-44 years old age category. This remains true for the remaining age categories (see Table 4-4). To assess possible errors or misclassification biases in this variable I compared the variable's breakdown to mid-2019 ONS age estimates for LBBD.

Table 4-4: Age categories created in the dataset

Age category	Year of birth
<16	2005 - 2020
16-29	1991 - 2004
30-44	1976 - 1990
45-54	1966 - 1975
55-64	1956 - 1965
65-74	1946 - 1955
75-84	1936 - 1945
85+	1935 and before

Sex: Data on sex was extracted from primary care records and categorised by Care City as "female", "male" or "other". To assess possible errors or misclassification biases in this variable I compared the variable's breakdown to mid-2019 ONS gender estimates for LBBD.

Ethnicity: The linked dataset contained three variables coding ethnicity: one from the primary care dataset and two from the individual-level council dataset. The variable from the primary care dataset extracted data on ethnicity directly from an individuals' primary care record. Data on ethnicity is collected in primary care as part of day-to-day care provision. One of the variables on ethnicity in the council dataset was also extracted from an individuals' primary care record. The third variable from the council dataset reflected work LBBD have done to address missing ethnicity data in primary care.

To address missingness on ethnicity, LBBD have drawn on several council data sources to create this third ethnicity variable. First, LBBD conduct a termly school census of all state school students in the borough. Parents fill out the school census and data on ethnicity is collected. The school census is the first data source used to fill in missing ethnicity data. If individuals leave school, but stay in the borough, the ethnicity they recorded in the school census will remain on their council records. Where possible, remaining gaps in ethnicity coding are then filled in using primary care records, however these do not override the school census recordings. After these two steps, remaining gaps in the council ethnicity variable are filled using a programme called Origins (Webber Phillips, 2021). Origins uses a comprehensive compilation of datasets including telephone directories to identify international locations where first and last names are most frequently observed. Origins use this information to estimate the origin of an individuals' name and estimate their ethnicity. LBBD started using Origins in 2011 and compared its accuracy to estimate ethnicity with the 2011 national census and every LBBD school census conducted since 2011. In 2011 it performed well for most ethnic groups when compared with the national census, however overestimated "White-Other" ethnicities and underestimated "Black-Caribbean" ethnicities. LBBD accounted for this over-estimation in future iterations of the data they hold on ethnicity. Origins has been used by the council since 2011 as a last resort to estimate ethnicity. If new data is collected via the school census or via primary care records, these data override the Origins coding.

For this study, I examined the completeness of each ethnicity variable and possible biases in ethnicity coding to make a decision about which variable to use. The variable for ethnicity from the primary care dataset had 29.0% (67493/232671) missing data for linked primary care and council records, which is reasonably consistent with national ethnicity recording in primary care (Mathur *et al.*, 2014). Missingness on this variable was associated with all three multimorbidity outcomes and household tenure (see Table 4-5). The second ethnicity variable from the individual-level council dataset, and that was directly extracted from primary care records, similarly had 28.7% missing data (66769/232641). Care City stated that the data from these two variables originated from the same data source, however, despite a high correlation between the variables (>0.9), cross tabulation revealed that these first two variables were not identical.

For this study, I chose to use the more complete third variable coding ethnicity extracted from the individual-level council dataset. I recategorized this third variable into four categories: “White”, “Black”, “Asian” and “Other”. These reflected ONS categories used for the 2011 census (see

Table 4-6), omitting the category for “Mixed/Multiple ethnic groups” (Office for National Statistics, 2011b). A sixth category “unknown” was created to account for missing data. To assess possible errors or misclassification biases in this variable, I compared the variable’s breakdown to 2013 Greater London Authority (GLA) Ethnic Group Projections for LBBD for 2020 (Barking and Dagenham Council, 2001). GLA extract their data from ONS estimates.

Table 4-5: Results of analyses assessing whether missingness on the primary care ethnicity variable was associated with exposure and outcome variables

	Ethnicity from primary care records		
	Known N=165178	**Unknown N=67493**	**χ2***
Basic multimorbidity: N (%)			
Present	31660 (76.6)	9669 (23.4)	p<.001
Absent	133518 (69.8)	57824 (30.2)	
Physical-mental multimorbidity: N (%)			
Present	6855 (75.5)	2222 (24.5)	p<.001
Absent	158323 (70.8)	65271 (29.2)	
Complex multimorbidity: N (%)			
Present	13669 (77.1)	4052 (22.9)	p<.001
Absent	151509 (70.5)	63441 (29.5)	
Household tenure: N (%)			
Owner-Occupied	63529 (70.2)	26934 (29.8)	
Private Rented	57284 (74.3)	19779 (25.7)	p<.001

Social Housing	40852 (17.6)	191718 (82.4)
Unknown	3513 (76.8)	1062 (23.2)
*Pearson's Chi-squared test with Yates' continuity correction		

Table 4-6: Office for National Statistics list of ethnicity groups and categories

Ethnic groups	
White	English/Welsh/Scottish/Northern Irish/British Irish Gypsy or Irish Traveller Any other White background
Black/African/Caribbean/Black British	African Caribbean Any other Black/African/Caribbean background
Asian/Asian British	Indian Pakistani Bangladeshi Chinese Any other Asian background
Mixed/Multiple ethnic groups	White and Black Caribbean White and Black African White and Asian Any other Mixed/Multiple ethnic background
Other ethnic group	Arabic Any other ethnic group

4.2.3.4 Individual-level health behaviour covariates

BMI: Data on BMI was extracted by Care City from an individuals' primary care record. There were two variables that captured the most recent recording of BMI in the dataset: a raw value and a description. For this study, I used both variables.

First, I explored the completeness and accuracy of the raw value variable. In the raw value variable, I found several zeros and values below nine, as well as other implausible values (e.g., there were approximately 600 individuals with BMI values recorded as over 60). I recoded the BMI values of these records as missing.

Second, I explored information contained within the BMI description variable, which contained free text information. I found that there were records that recorded an individuals' BMI in the free text, but not in the raw value variable. In these circumstances, the raw value variable was coded as missing. To increase completeness of the raw value variable, I extracted data from the BMI description variable and coded these into the raw value for BMI variable in circumstances where these data were missing. For example, free text entries such as "Body mass index 25-29 – overweight" were coded as 27.5 in the raw value for BMI variable if no other information was recorded in this variable. I did this for just over 600 records.

For this study, I then recategorized the raw value variable for BMI into five categories defined by the NHS as follows: underweight (below 18.5), healthy (between 18.5 and 24.9), overweight (between 25 and 29.9), obese (between 30 and 39.9) and morbidly obese (over 40). A sixth category of "unknown" was created to account missing data and unknown BMI. To explore possible biases in this variable, I assessed whether missingness was associated with each multimorbidity outcome and the household tenure exposure. I also compared the proportion of adults classified as overweight and obese in the Care City Cohort to 2019 estimates for LBBD from Public Health England's Fingertips dashboard, which compiles indicators across a range of health and wellbeing themes to support local areas and commissioning (Public Health England, 2019).

Smoking: Smoking status for each individual was extracted by Care City from primary care records. Care City created four categorises to reflect the most recent recording of smoking status: non-smoker, smoker, ex-smoker, or unknown. To explore possible biases in this variable, I assessed whether missingness was associated with each multimorbidity outcome and the household tenure exposure. I also compared the proportion of adults classified as current smokers in the Care City Cohort to 2019 estimates for LBBD from Public Health England's Fingertips dashboard (Public Health England, 2019).

4.2.3.5 Household-level socioeconomic covariates

Benefits receipt: Data on household benefits receipt was extracted from the household-level council dataset from Care City. The data originated from council housing records. Since 2016, a national programme known as "passported benefits" has allowed all individuals or households in the UK in receipt of any type of benefit to be automatically entitled to Housing Benefit. The council records whether or not people are "passported" and, if so, what benefits they are in receipt of that make them eligible for Housing Benefit: Employment Support and Allowance (ESA), Pension Credit, Income Support and/or Job Seeker's Allowance (JSA). The council also record whether an individual is not passported and receiving only Housing Benefit (see Table 4-7). Therefore, in this variable, I only had access to benefits data for individuals in receipt of Housing Benefits. There will have been individuals who were in receipt of benefits such as JSA but who were not receiving Housing Benefit. Individuals such as these were not captured in this variable.

Data on household benefits receipt are recorded at household-level rather than at individual-level. As such, I could not identify individuals in a household that were in receipt of the given benefit(s). The variables available indicated whether or not an individual resided in a household where a household member, either themselves or someone else, was in receipt of the given benefit(s). As such, a greater number of individuals may be coded as living in households where someone is in receipt of benefits than individuals actually in receipt. In addition, it is possible households may have contained more than one individual in receipt of benefits. For example, households with more than one individual over the age of 65 may have contained more than one individual in receipt of Pension Credit. To assess the possible impact of this, I compared the number of individuals and households in the Care City Cohort coded as in receipt of each type of benefit to mid-2019 population estimates for LBBD obtained from ONS.

Table 4-7: An overview of each benefit type captured in the dataset

Benefit	Description
Employment and Support Allowance	"Passported" and therefore in receipt of Housing Benefit because a member of their household is in receipt of Employment and Support Allowance
Pension Credit	"Passported" and therefore in receipt of Housing Benefit because a member of their household is in receipt of Pension Credit

Income Support	"Passported" and therefore in receipt of Housing Benefit because a member of their household is in receipt of Income Support
Job Seeker's Allowance	"Passported" and therefore in receipt of Housing Benefit because a member of their household is in receipt of Job Seeker's Allowance
Housing Benefit Only	"Not passported" and therefore in receipt of Housing Benefit but not in receipt of any other benefits.
None	Not in receipt of Housing Benefits

For this study, six categories were used to indicate whether anyone in an individuals' household was in receipt of: ESA, Pension Credit, Income Support, JSA, Housing Benefit only or no benefits.

Household occupancy: Data on household occupancy were extracted from council records and originated from the council's housing department. This meant that individuals who resided with someone excluded from the dataset would have that person reflected in their household occupancy data. Individuals who lived with people that were registered with a general practice (GP) outside of the borough also had that person reflected in their household occupancy variable.

For this study, I categorised household occupancy data into four categories to reflect 1-2, 3-5, 6-10 and 11 or more people within a household. For modelling purposes, I did not create a category reflecting 1 person households in this variable.

Household type: Data on household type were extracted by Care City from council records and originated from the council's housing department. Care City grouped households into seven types: adults with children, adults with no children, single adult with children, single adult, older adults with no children, three-generations, and other.

4.2.3.6 Area-level sociodemographic covariates

Area-level clustering: In the dataset, individuals were clustered within small geographical areas known as Lower Super Output Areas (LSOAs). LSOAs are small areas created as part of the census (Office for National Statistics, 2011a). They contain a minimum and maximum population of 1000 and 3000, respectively, and a

minimum and maximum number of households of 400 and 1200, respectively. In LBBD there were 110 LSOAs in 2019/20.

There were two variables indicating an individuals' LSOA in the dataset, one originating from their primary care record and one from their council record. Data on an individuals' LSOA in the primary care data was generated by Care City from each residents' address captured in their primary care record. Data on an individuals' LSOA in the council data was generated by the council - the council had previously assigned each individual an LSOA code based on their postcode. Individuals within the same small geographical area shared the same LSOA code.

For this study, I examined the completeness of each LSOA variable to make a decision about which variable to use. The LSOA variable from the primary care records contained 476 individuals with missing LSOA codes (0.20% of successfully linked primary care and council records), and 145 individuals with LSOA codes from outside LBBD (0.06% of linked records). The LSOA variable extracted from the council records was more complete. I therefore chose to use LSOA codes extracted from council records. Where LSOA codes were missing, these were coded as "unknown".

Area-level deprivation: I obtained data on area-level deprivation from ONS. To measure deprivation, ONS calculate Index of Multiple Deprivation (IMD) scores for each LSOA in England as part of the census. IMD scores allows areas to be ranked from 1 (most deprived area) to 32844 (least deprived area). The score is constructed by combining seven domains of deprivation that cover a range of economic, social and housing issues. Each domain is given a weighting to reflect its perceived importance: Income Deprivation (22.5% weighting), Employment Deprivation (22.5%), Health Deprivation and Disability (13.5%), Education Skills and Training Deprivation (13.5%), Barriers to Housing and Services (9.3%), Living Environment Deprivation (9.3%) and Crime (9.3% weighting) (Office for National Statistics, 2019a).

For this study, I linked IMD (2019) scores to each individual's LSOA code in the data (extracted from their council records). This gave a marker of overall relative deprivation in each participants' residential area. IMD scores are usually categorised nationally into tertiles or quintiles. However, LBBD is a highly deprived borough and

most LSOAs are categorised within the two most deprived national IMD quintiles. To account for this, and to increase the spread of IMD scores within the linked dataset, I calculated IMD quintiles for LBBD based on raw IMD scores from 2019 for each LSOA in LBBD.

Given this study's focus on housing, I included the Barriers to Housing and Services domain of the IMD as a separate covariate. For this variable, I also calculated LBBD-specific quintiles based on raw scores on the Barriers to Housing and Services domain to increase the spread of scores.

4.2.3.7 Confirming LBBD residency

Care City used Mayhew and Harper's Residents' Matrix to identify confirmed LBBD residents. Confirmed residents are defined by the Matrix as those present on the address register, and on either another council data, the GP register, or both (Harper and Mayhew, 2012). Care City created a binary variable to indicate whether each individual was confirmed or not confirmed as a resident of LBBD.

4.2.4 Creating the study cohort

To create the cohort used in this study, I excluded:

- Individuals not of working age (16-64 years old, inclusive).
- Individuals not identified as residents of LBBD by Mayhew and Harper's Residents' Matrix (Harper and Mayhew, 2012).
- Care home residents.
- Individuals with missing data on household tenure.
- Individuals with missing LSOA codes extracted from council records.
- Individuals identifying as "other" gender.
- Individuals who lived in "other" types of households (excluded for modelling purposes due to too few numbers in this category).

4.2.5 Data analysis

I explored associations between household tenure and multimorbidity prevalence amongst working age LBBD residents using multilevel dichotomous logistic regression modelling.

Multilevel regression modelling allows social epidemiologists to model hierarchical data, where individuals are grouped (or 'clustered') within groups (Goldstein, Browne and Rasbash, 2002). Groups can include schools, households, or geographical areas. In hierarchical data, individuals may share characteristics attributable to these groups that go above and beyond individual-level characteristics in explaining the outcome of interest. As such, the assumption of independent observations does not hold for hierarchical data and multilevel regression methods are required (Goldstein, Browne and Rasbash, 2002).

In this study, the multilevel models included random effects at small geographical area (LSOA) level to a) account for clustering within areas and b) quantify the amount of variation in multimorbidity prevalence that can be attributed to area-level characteristics compared with individual-level and household-level characteristics. I explored adding a further level to account for household-level clustering in the data. In my study cohort, 35% of households had only one resident and most of the remaining households had only two or three residents. In multilevel modelling, too few individuals per cluster leads to less precise random intercepts and too small standard errors (Austin and Leckie, 2018). I, therefore, chose to exclude from all models a second level modelling household-level clustering.

For each model, variance partition coefficients (VPCs) were estimated to assess the proportion of the total variance in multimorbidity prevalence explained by area-level characteristics. When all individual-level and household-level characteristics are added to the model, the VPC represents the variance attributed to variation at the area-level.

To build the final models for each outcome, age and sex were first added a priori. The addition of ethnicity, BMI and smoking as covariates was informed by previous literature suggesting that ethnicity and health-related behaviours are associated with both multimorbidity and household tenure (Rocca *et al.*, 2014; Bobo *et al.*, 2016; Katikireddi *et al.*, 2017; Clair and Hughes, 2019; Office for National Statistics, 2020). Household-level covariates were then added to explore the additional benefit of adjusting for these variables. After fitting each model, model fit was assessed using Akaike's Information Criteria (AIC) and variables were retained in models if adding them improved model fit.

For each outcome, I fitted the following multilevel dichotomous logistic regression models:

- Model 1 – an unadjusted model with no covariates.
- Model 2 – a model adjusted for individual-level socio-demographic characteristics age, sex, and ethnicity.
- Model 3 – a model adjusted for model 2 covariates plus individual-level health behaviour covariates BMI and smoking status.
- Model 4 – a model adjusted for model 2 and model 3 covariates plus household-level covariates household receipt of benefits, household occupancy and household type.

To conduct these analyses, I used R software (version Rx64 4.1.1) and the glmer function from the lme4 R software package (Bates *et al.*, 2015). My models were fitted using penalised least squares and 95% confidence intervals calculated using the Wald test (Doerken *et al.*, 2019). I chose to use the lme4 package as an alternative R package - the MASS package (Finch, Bolin and Kelley, 2014) - uses penalized quasi-likelihood estimation (Wolfinger and O'Connell, 1993). As such, it is not possible to obtain goodness of fit parameters using the MASS package. In addition, I faced difficulties calculating 95% confidence intervals when using this second package.

4.2.5.1 Subgroup and sensitivity analyses

For each final model (Model 4), I conducted subgroup analyses to examine whether there was evidence of interactions between household tenure and other household-level variables: household benefits receipt, type, and occupancy. These household-level variables were key candidates for subgroup analyses as they are most likely to interact with household tenure, and act at the same level of influence as household tenure. For each multimorbidity outcome, I added three interaction terms one at a time to the final model (Model 4) to model interactions between household tenure and household receipt of benefits, household occupancy and household type. After fitting each model with each interaction term, model fit was assessed using AIC and evidence of an interaction implied from the p value.

For each multimorbidity outcome, I also conducted sensitivity analyses for each final model (Model 4) to examine whether model estimates and variance estimates

changed substantially when models were either single-level, accounted for household-level clustering, and accounted for household or area-level clustering. Given the results of my consultations with clinicians and the PPI panel, I conducted a further three sensitivity analyses for each final model (Model 4) that examined associations between household characteristics and each multimorbidity outcome when excluding certain conditions. These analyses included multimorbidity definitions that:

1. Excluded risk factors (AF and hypertension).
2. Excluded conditions where there were concerns around coding of the conditions in the data. Conditions excluded were:
    a. Those that are typically mild and often require relatively little consultation in primary care: chronic sinusitis, constipation, diverticular disease of intestine, prostate disorders, psoriasis, or eczema.
    b. CKD diagnosed at stages 1, 2 and 3.
    c. Double counted COPD and bronchiectasis (i.e., if an individual had only two chronic conditions and these were COPD and bronchiectasis, they were not classified as having multimorbidity).
    d. Double counted anorexia and depression and/or anxiety (i.e., if an individual had only two chronic conditions and these were anorexia and depression and/or anxiety, they were not classified as having multimorbidity).
3. Excluded risk factors and conditions where there were concerns around coding of the conditions in the data.

4.3 Chapter summary

In this chapter I have presented the introduction and methodology for my second study. I have described how findings from my systematic review (Chapter 3) have been used to inform the design of this study. Using a novel dataset that links health and council records for residents of LBBD, as well as multilevel logistic regression modelling, this study examines and quantifies associations between household tenure and different types of multimorbidity amongst working age adults. Through the lens of multimorbidity, this study acts as a use case for creating, using, and analysing linked health and council datasets to understand the social determinants of local

public health concerns and generate knowledge that could inform equitable decision-making. The results for this study are presented in Chapter 5.

Chapter 5 A quantitative analysis of linked health and council data investigating associations between household tenure and multimorbidity: Results and Discussion

In this chapter, I present the results of a study that, using linked health and council data, examines and quantifies associations between household tenure and different types of multimorbidity amongst working age adults in LBBD. The introduction and methodology for this study are presented in Chapter 4. Section 1 of this chapter presents the results of the data linkage. Section 2 describes the study cohort. Sections 3, 4 and 5 describe the results of the analyses conducted for each multimorbidity outcome. I finish this chapter by considering how findings compare to existing literature and the strengths and limitations of this study.

5.1 Results

5.1.1 Linked health and council dataset

Figure 5-1 illustrates the results of my linkage of the separate primary care and council datasets. A total of 232671 individuals were linked across primary care and council datasets (84.0% of the original primary care records).

5.1.1.1 Unmatched records

1.3% (3014/236658) of individual-level council records were unmatched to a household-level council record on Household_ID. 0.3% (639/236658) of individual-level council records were duplicate records and excluded. Care City examined these records and found they originated from children in a pathway for special educational needs. As such, these children received two education records meaning they were registered twice in the dataset. 0.2% (369/233005) of linked council records were unmatched to a primary care record. 16.0% (44269/276905) of primary care records were unmatched to a council record. 2.9% of these were because individuals had died in 2019/20 (1295/44269).

Table 5-1 shows the results of analyses conducted to assess whether there were any potential biases in the linkage results for matched and unmatched primary care

records for selected variables originating from the primary care data. Appendix 5 shows these results for each of the 38 chronic conditions included in the multimorbidity definitions. Standardised differences of 0.2, 0.5, and 0.8 indicate small, medium and large effect sizes, respectively (Harron *et al.*, 2017).

Figure 5-1: Results of data linkage

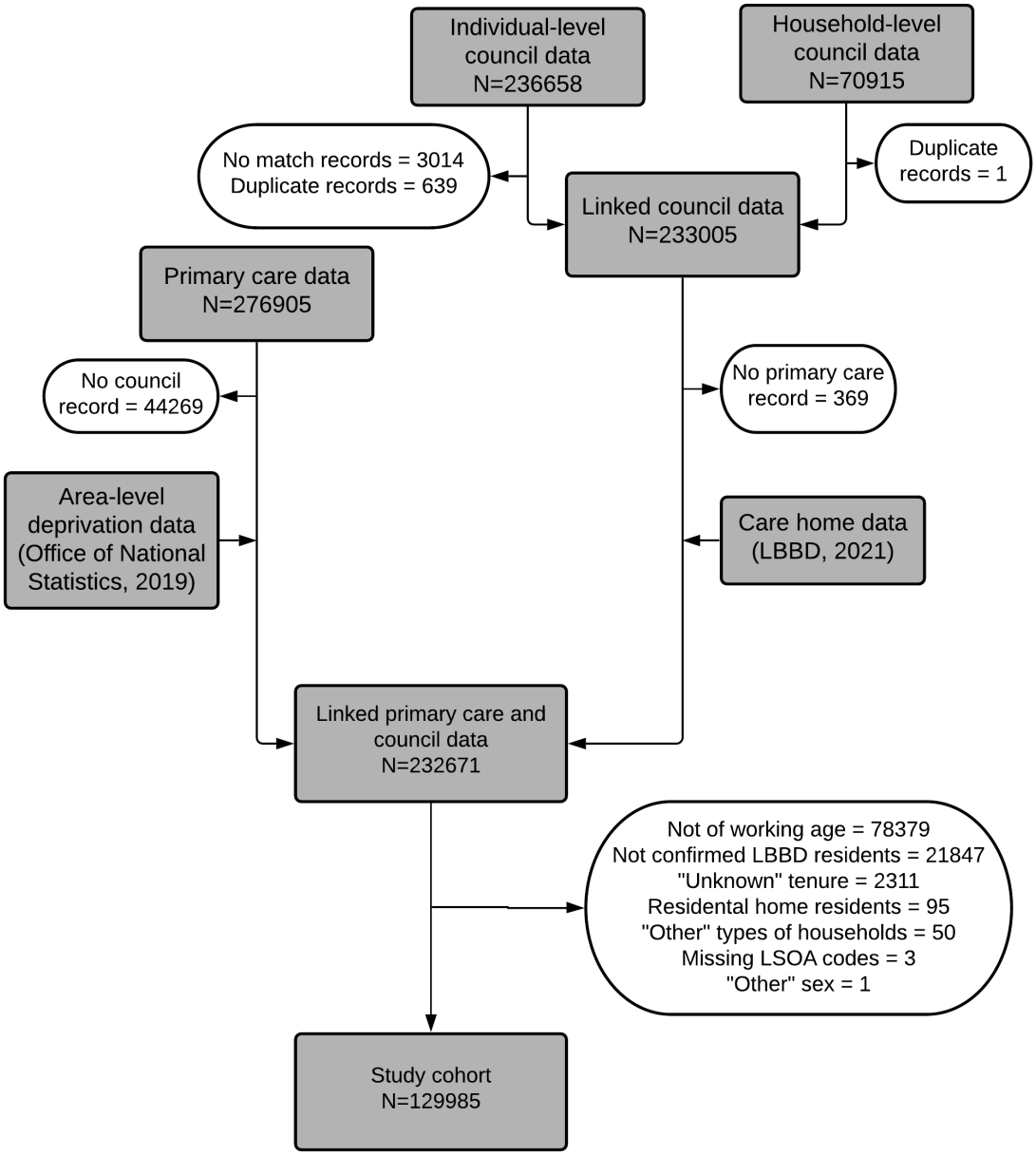


Table 5-1: Results of analyses to assess potential biases in the linkage results for matched (N=232671) and unmatched (N=44269) primary care records

	Primary care matched records N=232671	Primary care unmatched records N=44269	Standardised difference
Age: N (%)			
<16	57402 (24.7)	8877 (20.1)	0.150
16-29	42325 (18.2)	8593 (19.4)	
30-44	59891 (25.7)	11942 (27.0)	
45-54	30738 (13.2)	5679 (12.8)	
55-64	21338 (9.17)	4101 (9.26)	
65-74	11602 (4.99)	2461 (5.56)	
75-84	6366 (2.74)	1414 (3.19)	
85+	3009 (1.29)	1202 (2.72)	
Sex: N (%)			
Female	116186 (49.9)	21787 (49.2)	0.025
Male	116484 (50.1)	22472 (50.8)	
Other/Missing	1 (0.00)	10 (0.00)	
Ethnicity*: N (%)			
White	76524 (32.9)	13633 (31.4)	0.128
Black	32708 (14.1)	5029 (11.4)	
Asian	42222 (18.1)	9710 (21.9)	
Mixed	6285 (2.70)	1137 (2.57)	
Other	4309 (1.85)	831 (1.88)	
Unknown	67493 (29.0)	13629 (30.8)	
Basic multimorbidity: N (%)			
Present	41329 (17.8)	7931 (17.9)	0.004
Absent	191342 (82.2)	36338 (82.1)	
Physical-mental multimorbidity: N (%)			
Present	9077 (3.90)	1542 (3.48)	0.022
Absent	223594 (96.1)	42727 (96.5)	
Complex multimorbidity N (%):			0.016
Present	17721 (7.65)	3562 (8.09)	
Absent	214950 (92.4)	40707 (91.6)	
BMI categories: N (%)			
Underweight	11645 (5.00)	2115 (4.78)	0.077
Healthy weight	48101 (20.7)	10355 (23.4)	
Overweight	49180 (21.1)	9493 (21.4)	
Obese	37566 (16.1)	6612 (14.9)	
Morbidly obese	6077 (2.61)	934 (2.11)	
Unknown	80102 (34.4)	14760 (33.3)	
Smoking status: N (%)			
Non-smoker	107326 (46.1)	21247 (48.0)	0.043
Ex-smoker	24385 (10.5)	4620 (10.4)	
Smoker	33722 (14.5)	6372 (14.4)	
Unknown	67238 (28.9)	12030 (27.2)	

*variable taken from primary care records, unlike in the study analyses

5.1.1.2 Matched records

For the primary care matched records (N=232671), there were 69467 households. 39.7% of households were coded as owner-occupied properties, 29.2% as rented from a local authority or housing association, 27.8% as privately rented, 1.10% as in LBBD's Reside programme and 2.17% of households had "unknown" tenure. As seen in Table 5-2, the Care City Cohort data estimated a higher proportion of individuals resided in privately rented housing and a lower proportion lived in owner-occupied housing when compared to mid-2019 ONS estimates for LBBD.

Table 5-2: Breakdown of household tenure variable in the primary care matched records (N=232671) compared to 2019 mid-year ONS tenure estimates for LBBD

	Households in primary care matched records (N=69467)		**Households in mid-2019 ONS estimates (N=75829)**	
	N	%	N	%
Tenure				
OOC*	27576	39.7	38178	50.3
Social Housing+	20312	29.2	22140	29.2
Private Rented	19304	27.8	15511	20.5
Reside	767	1.10	-	-
Unknown	1508	2.17	-	-

Note: ONS population estimates are mid-year estimates based on annual population modelling of 2011 census data and account for changes in births, deaths, and national migration (Office for National Statistics, 2019b).
*OOC = Owner-occupied.
+Rented from local authority or housing association.

For the primary care matched records (N=232671), 66.3% of participants were working age (between 16 and 64 years old, inclusive) and 49.9% of participants were female (see Table 5-3). Table 5-3 shows that the age and sex breakdown of these matched records was similar to mid-2019 ONS population estimates for LBBD. For the primary care matched records (N=232671), the variable that captured an individuals' ethnicity and that was used in this study was comprised of 51.8% White, 21.3% Black and 23.9% Asian participants. In this variable, 2.2% of participants were classified as having "Other" ethnicities and 0.7% of participants had "unknown" ethnicities. This ethnic breakdown is compared to 2011 census data and GLA 2020 projections in Table 5-4. These comparisons suggest that the council ethnicity

variable may have overestimated White ethnicities and underestimated Black ethnicities in 2019/2020 (see Table 5-4).

Table 5-3: Breakdown of the age and sex variables in the primary care matched records (N=232671) compared to 2019 mid-year population estimates for LBBD

		Primary care matched records (N=232671)	Mid-2019 ONS population estimates (N=211185)
Age: N (%)			
	<16	57402 (24.7)	57528 (27.2)
	16-29	42325 (18.2)	40050 (19.0)
	30-44	59891 (25.7)	50034 (23.7)
	45-54	30738 (13.2)	26541 (12.6)
	55-64	21338 (9.17)	17896 (8.47)
	65-74	11602 (4.99)	10485 (4.96)
	75-84	6366 (2.74)	5975 (2.83)
	85+	3009 (1.29)	2676 (1.27)
Sex: N (%)			
	Female	116186 (49.9)	105877 (50.1)
	Male	116484 (50.1)	105308 (49.9)
	Other/Missing	1 (0.0)	-

Note: ONS population estimates are mid-year estimates based on annual population modelling of 2011 census data and account for changes in births, deaths, and national migration (Office for National Statistics, 2019b).

Table 5-4: Breakdown of the council ethnicity variable in the primary care matched records (N=232671) compared to 2011 census data and 2020 ethnic projections for LBBD

	Ethnicity variable from council records (%)	Census data (2011) (%)	Ethnic group projections (2020)* (%)
White ("White")	51.8	58.3	44.8
Black/African/Caribbean/Black British ("Black")	21.3	20.0	30.3
Asian/Asian British ("Asian")	23.9	15.9	21.4
Mixed/Multiple ethnic groups ("Mixed")	-	4.24	-
Other ethnic group ("Other")	2.2	1.56	3.5
Unknown	0.7	-	-

*2013 GLA Ethnic Group Projections for LBBD for 2020 (Barking and Dagenham Council, 2001). GLA extract their data from ONS estimates.

For the primary care matched records (N=232671), participants with missing BMI or smoking status were less likely to have all three multimorbidity outcomes recorded as present in their primary care data, and less likely to be in owner-occupied housing (see Table 5-5 and Table 5-6, respectively). The proportion of adults classified as overweight or obese in the Care City data, and as currently smoking are compared to PHE Fingertips data for LBBD in Table 5-7. These comparisons suggested that the matched primary care records may have underestimated the number of overweight or obese adults and underestimated the number of adults who smoked (see Table 5-7).

For the primary care matched records (N=232671), the variable that captured household receipt of benefits was comprised of 7756 individuals living in households in receipt of ESA, 6082 for pension credit, 5109 individuals in households in receipt of income support and 990 for JSA (see Table 5-8). Table 5-8 also shows that 2990 households were in receipt of ESA, 3045 in receipt of pension credit, 1295 in receipt of income support and 362 households in receipt of JSA. Discussions with LBBD suggested that differences between these levels in the dataset and ONS mid-2019 estimates were in line with expectations. This is because ONS estimates are based on modelling of 2011 census data which leaves room for error in the modelling. These analyses indicated that receipt of benefits may have strongly clustered within households in LBBD in 2019/20.

Table 5-5: Results of analyses investigating whether missingness on BMI was associated with exposure and outcome variables for primary care matched records (N=232671)

	BMI from primary care records		
	Known N=152569	Unknown N=80102	$\chi 2^*$
Basic multimorbidity: N (%)			
Present	38078 (25.0)	3251 (4.06)	p<.001
Absent	114491 (75.0)	76851 (95.9)	
Physical-mental multimorbidity: N (%)			
Present	8569 (5.62)	508 (0.63)	p<.001
Absent	144000 (94.4)	79594 (99.4)	
Complex multimorbidity: N (%)			
Present	16763 (11.0)	958 (1.20)	p<.001
Absent	135806 (89.0)	79144 (98.8)	
Household tenure: N (%)			
Owner-Occupied	63380 (41.5)	27083 (33.8)	p<.001
Private Rented	47089 (30.9)	29974 (37.4)	
Social Housing	38874 (25.5)	21696 (27.1)	
Unknown	3226 (2.11)	1349 (1.68)	
*Pearson's Chi-squared test with Yates' continuity correction			

Table 5-6: Results of analyses investigating whether missingness on smoking status was associated with exposure and outcome variables for primary care matched records (N=232671)

	Smoking status from primary care records		
	Known N=165433	Unknown N=674238	$\chi 2^*$
Basic multimorbidity: N (%)			
Present	40095 (24.2)	1234 (1.84)	p<.001
Absent	125338 (75.8)	66004 (98.2)	
Physical-mental multimorbidity: N (%)			
Present	9019 (5.45)	58 (0.09)	p<.001
Absent	156414 (94.5)	67180 (99.9)	
Complex multimorbidity: N (%)			
Present	17565 (10.6)	156 (0.23)	p<.001
Absent	147868 (89.4)	67082 (99.8)	
Household tenure: N (%)			
Owner-Occupied	69146 (41.8)	21317 (31.7)	p<.001
Private Rented	50261 (30.4)	26802 (39.9)	
Social Housing	42593 (25.7)	17977 (26.7)	
Unknown	3433 (2.08)	1142 (1.70)	
*Pearson's Chi-squared test with Yates' continuity correction			

Table 5-7: Comparisons of BMI and smoking status recording in the Care City Cohort with 2017/18 Public Health England Fingertips data

	Care City Cohort data (%)	2019 PHE Fingertips data* (%)
Proportion of adults classified as overweight or obese	39.9	64.4
Proportion of current adult smokers	14.5	22.4

*2019 estimates for LBBD from Public Health England's Fingertips dashboard (Public Health England, 2019).

Table 5-8: Breakdown of household benefits receipt variable in the primary care matched records (N=232671) compared to mid-2019 ONS estimates

Benefit	ONS mid-2019 estimates for LBBD	Individuals in Care City Cohort (2019/20)	Individuals in Care City Cohort as a % of ONS figures	Households in Care City Cohort (2019/20)	Households in Care City Cohort as a % of ONS figures
ESA*	6530	7756	119%	2990	46%
Pension Credit	5170	6082	118%	3045	59%
Income Support	2520	5109	203%	1295	51%
JSA*	1200	990	83%	362	30%

Note: ONS population estimates are mid-year estimates based on annual population modelling of 2011 census data and account for changes in births, deaths, and national migration (Office for National Statistics, 2019b).
ESA = Employment Support Allowance; JSA = Job Seeker's Allowance.

5.1.2 Study cohort

The cohort used in this study was comprised of 129985 working age LBBD residents. This number represented 55.9% (129985/232671) of the records that were successfully linked across the 2019/20 primary care and council datasets. 78379 records (33.7%) were excluded as individuals were not of working age (between 16-64 years old, inclusive). 21847 records (9.38%) were excluded as individuals were not confirmed as residents of LBBD by Mayhew and Harper's Residents' Matrix (Harper and Mayhew, 2012). 2311 records were excluded as individuals had "unknown" household tenure (0.99%), 95 excluded as individuals resided in

residential homes (0.04%), three excluded due to missing LSOA codes and one excluded as they were recorded as “other” sex. 50 records (0.02%) were also excluded as these individuals resided in “other” types of households. See Figure 5-1.

The 129985 working age LBBD residents included in these analyses resided in 58201 households and 110 LSOAs. Table 5-9 gives an overview of the study cohort. 41.9% (54324/129985) of the cohort resided in owner-occupied properties, 30.7% (39885/129985) in privately rented properties and 27.5% (35776/129985) in social housing. 65.4% (87031/129985) of participants were 16-44 years old, 52.8% (68593/129985) were White and 52.0% (67487/129985) were overweight or obese.

Table 5-9: An overview of the study cohort

		Total cohort (N=129985)	Basic multimorbidity (N=23425)		Physical-mental multimorbidity (N=6214)		Complex multimorbidity (N=7857)	
		N (%)	N (%)	Odds ratio (95% CI)	N (%)	Odds ratio (95% CI)	N (%)	Odds ratio (95% CI)
Individual-level variables								
Age								
	16-29	36900 (28.4)	2475 (10.6)	-	574 (9.24)	-	327 (4.16)	-
	30-44	48131 (37.0)	5617 (24.0)	1.84 (1.75-1.93)	1650 (26.6)	2.25 (2.04-2.47)	1262 (16.1)	3.01 (2.67-3.41)
	45-54	26198 (20.2)	6939 (29.6)	5.01 (4.77-5.26)	1880 (30.3)	4.89 (4.45-5.38)	2342 (29.8)	11.0 (9.78-12.4)
	55-65	18756 (14.4)	8394 (35.8)	11.3 (10.7-11.8)	2110 (34.0)	8.02 (7.31-8.82)	3926 (50.0)	29.6 (26.5-33.3)
Sex								
	Female	66848 (51.4)	13162 (56.2)	-	4076 (65.6)	-	4583 (58.3)	-
	Male	63137 (48.6)	10263 (43.8)	0.79 (0.77-0.81)	2138 (34.4)	0.54 (0.51-0.57)	3274 (41.2)	0.74 (0.71-0.78)
Ethnicity								
	White	68593 (52.8)	14333 (61.2)	-	4714 (75.9)	-	5135 (65.4)	-
	Black	27983 (21.5)	4142 (17.7)	0.66 (0.63-0.68)	616 (9.91)	0.31 (0.28-0.33)	1164 (14.8)	0.54 (0.50-0.57)
	Asian	31032 (23.9)	4746 (20.3)	0.68 (0.66-0.71)	839 (13.5)	0.38 (0.35-0.41)	1508 (19.2)	0.63 (0.59-0.67)
	Other	1877 (1.44)	169 (0.72)	0.37 (0.32-0.44)	40 (0.64)	(0.30 (0.21-0.40)	38 (0.48)	0.26 (0.18-0.35)
	Unknown	500 (0.38)	35 (0.15)	0.28 (0.20-0.40)	5 (0.08)	0.14 (0.05-0.30)	12 (0.15)	0.30 (0.16-0.51)

	Total cohort (N=129985)	Basic multimorbidity (N=23425)		Physical-mental multimorbidity (N=6214)		Complex multimorbidity (N=7857)	
	N (%)	N (%)	Odds ratio (95% CI)	N (%)	Odds ratio (95% CI)	N (%)	Odds ratio (95% CI)
BMI categories							
Underweight	4263 (3.28)	516 (2.20)	0.85 (0.77-0.93)	142 (2.29)	0.88 (0.73-1.04)	119 (1.51)	0.73 (0.60-0.87)
Healthy weight	33166 (25.5)	4646 (19.8)	-	1251 (20.1)	-	1262 (16.1)	-
Overweight	35182 (27.1)	6776 (28.9)	1.46 (1.41-1.53)	1664 (26.8)	1.27 (1.18-1.37)	2123 (27.0)	1.62 (1.51-1.74)
Obese	27537 (21.2)	7813 (33.4)	2.43 (2.33-2.53)	2123 (34.2)	2.13 (1.98-2.29)	3012 (38.3)	3.10 (2.90-3.32)
Morbidly obese	4768 (3.67)	2089 (8.92)	4.79 (4.49-5.11)	705 (11.3)	4.43 (4.01-4.88)	995 (12.7)	6.67 (6.09-7.29)
Unknown	25069 (19.3)	1585 (6.77)	0.41 (0.39-0.44)	329 (5.29)	0.34 (0.30-0.38)	346 (4.40)	0.35 (0.31-0.40)
Smoking status							
Non-smoker	76430 (58.8)	12815 (54.7)	-	2821 (45.4)	-	4071 (51.8)	-
Ex-smoker	15575 (12.0)	4492 (19.2)	2.01 (1.93-2.09)	1356 (21.8)	2.49 (2.33-2.66)	1795 (22.8)	2.32 (2.18-2.45)
Smoker	25095 (19.3)	5795 (24.7)	1.49 (1.44-1.54)	1998 (32.2)	2.26 (2.13-2.39)	1970 (25.1)	1.51 (1.43-1.60)
Unknown	12885 (9.91)	323 (1.38)	0.13 (0.11-0.14)	39 (0.63)	0.08 (0.06-0.11)	21 (0.27)	0.03 (0.02-0.04)
Household-level variables							
Tenure							
Owner-Occupied	54324 (41.8)	9278 (39.6)	-	1801 (29.0)	-	2853 (36.3)	-
Private Rented	39885 (30.7)	5143 (22.0)	0.72 (0.69-0.75)	1328 (21.4)	1.00 (0.93-1.08)	1554 (19.8)	0.73 (0.69-0.78)
Social Housing	35776 (27.5)	9004 (38.4)	1.63 (1.58-1.69)	3085 (49.6)	2.75 (2.59-2.92)	3450 (43.9)	1.93 (1.83-2.03)

	Total cohort (N=129985)	Basic multimorbidity (N=23425)		Physical-mental multimorbidity (N=6214)		Complex multimorbidity (N=7857)	
	N (%)	N (%)	Odds ratio (95% CI)	N (%)	Odds ratio (95% CI)	N (%)	Odds ratio (95% CI)
Benefit Receipt							
None	102460 (78.8)	15690 (67.0)	-	3212 (51.7)	-	4501 (57.3)	-
ESA	5634 (4.33)	2924 (12.5)	5.97 (5.65-6.30)	1499 (24.1)	11.2 (10.5-12.0)	1652 (21.0)	9.03 (8.46-9.63)
Pension	1839 (1.41)	489 (2.09)	2.00 (1.80-2.22)	130 (2.09)	2.35 (1.95-2.81)	183 (2.33)	2.41 (2.05-2.80)
Income Support	2486 (1.91)	756 (3.23)	2.42 (2.21-2.64)	242 (3.89)	3.33 (2.90-3.81)	298 (3.79)	2.96 (2.61-3.35)
JSA	638 (0.49)	168 (0.72)	1.98 (1.65-2.35)	41 (0.66)	2.12 (1.52-2.88)	63 (0.80)	2.38 (1.82-3.07)
Housing benefit only	16928 (13.0)	3398 (14.5)	1.39 (1.33-1.45)	1090 (17.5)	2.13 (1.98-2.28)	1160 (14.8)	1.60 (1.50-1.71)
Occupancy							
1-2	28110 (21.6)	7750 (33.1)	-	2526 (40.7)	-	3075 (39.1)	-
3-5	75198 (57.9)	12337 (52.7)	0.52 (0.50-0.53)	3020 (48.6)	0.42 (0.40-0.45)	3828 (48.7)	0.44 (0.42-0.46)
6-10	25037 (19.3)	3180 (13.6)	0.38 (0.37-0.40)	635 (10.2)	0.26 (0.24-0.29)	908 (11.6)	0.31 (0.28-0.33)
11+	1640 (1.26)	158 (0.67)	0.28 (0.24-0.33)	33 (0.53)	0.21 (0.14-0.29)	46 (0.59)	0.23 (0.17-0.31)
Household type							
Adults with children	59754 (46.0)	7653 (32.7)	-	1758 (28.3)	-	2101 (26.7)	-
Adults with no children	41331 (31.8)	8936 (38.1)	1.88 (1.82-1.94)	2320 (37.3)	1.96 (1.84-2.09)	3233 (41.1)	2.33 (2.20-2.46)
Single adult with children	7087 (5.45)	1085 (4.63)	1.23 (1.15-1.32)	351 (5.65)	1.72 (1.53-1.93)	252 (3.21)	1.01 (0.88-1.15)
Single adult	10205 (7.85)	3274 (14.0)	3.22 (3.06-3.37)	1196 (19.2)	4.38 (4.06-4.73)	1384 (17.6)	4.31 (4.01-4.62)
Older cohabiting adults	7156 (5.50)	1849 (7.89)	2.37 (2.24-2.51)	465 (7.48)	2.29 (2.06-2.54)	690 (8.78)	2.93 (2.68-3.20)
Three-generations	4452 (3.42)	628 (2.68)	1.12 (1.02-1.22)	124 (2.00)	0.95 (0.78-1.13)	197 (2.51)	1.27 (1.09-1.47)

	Total cohort (N=129985)	Basic multimorbidity (N=23425)		Physical-mental multimorbidity (N=6214)		Complex multimorbidity (N=7857)	
	N (%)	N (%)	Odds ratio (95% CI)	N (%)	Odds ratio (95% CI)	N (%)	Odds ratio (95% CI)
LBBD IMD Quintile*							
1 (least deprived)	26767 (20.6)	4469 (19.1)	-	995 (16.0)	-	1444 (18.4)	-
2	25835 (19.9)	4331 (18.5)	1.00 (0.96-1.05)	1060 (17.1)	1.11 (1.01-1.21)	1351 (17.2)	0.97 (0.90-1.04)
3	25989 (20.0)	4666 (19.9)	1.09 (1.04-1.14)	1261 (20.3)	1.32 (1.21-1.44)	1579 (20.1)	1.13 (1.05-1.22)
4	25505 (19.6)	4750 (20.3)	1.14 (1.09-1.19)	1321 (21.3)	1.41 (1.30-1.54)	1610 (20.5)	1.18 (1.10-1.27)
5 (most deprived)	25889 (19.9)	5209 (22.2)	1.26 (1.20-1.31)	1577 (25.4)	1.68 (1.55-1.82)	1873 (23.8)	1.37 (1.27-1.47)
LBBD IMD Housing Quintile*							
1 (least deprived)	26489 (20.4)	4435 (18.9)	-	1021 (16.4)	-	1457 (18.5)	-
2	26319 (20.2)	5151 (22.0)	1.21 (1.16-1.26)	1411 (22.7)	1.41 (1.30-1.53)	1791 (22.8)	1.25 (1.17-1.35)
3	26287 (20.2)	4750 (20.3)	1.10 (1.05-1.15)	1340 (21.6)	1.34 (1.23-1.46)	1579 (20.1)	1.10 (1.02-1.18)
4	25235 (19.4)	4931 (21.1)	1.21 (1.15-1.26)	1378 (22.2)	1.44 (1.33-1.57)	1715 (21.8)	1.25 (1.17-1.35)
5 (most deprived)	25655 (19.7)	4158 (17.8)	0.96 (0.92-1.01)	1064 (17.1)	1.08 (0.99-1.18)	1315 (16.7)	0.93 (0.86-1.00)

Note: the denominator for all characteristics (individual and household) is the number of individuals

*Calculated for LBBD based on raw Index of Multiple Deprivation (IMD) scores (2019) and the IMD Barriers to Housing and Services domain scores (2019) (Office for National Statistics, 2019a)

5.1.3 Outcome 1: Basic multimorbidity

The crude prevalence of basic multimorbidity amongst working age residents of LBBD was 18.0% (23425/129985) in 2019/20. Results from bivariate, unadjusted analyses found strong evidence that the odds of basic multimorbidity increased with age (OR 11.3, 95% CI 10.7-11.8 for 55–65-year-old residents compared to 16–29-year-olds) and were lower for males (OR 0.79, 95% CI 0.77-0.81). There was also strong evidence that unadjusted odds of basic multimorbidity were lower for Black and Asian working age residents compared to White participants (OR 0.66, 95% CI 0.63-0.68 for participants grouped as Black compared to White, for example). These results found strong evidence that unadjusted odds of multimorbidity were greater with increased BMI and for smokers and ex-smokers compared to non-smokers (see Table 5-9).

Table 5-10 presents the results from the multilevel models built for the basic multimorbidity outcome. In Table 5-10, model 1 shows results of a model that found strong evidence that the unadjusted odds of basic multimorbidity were 75% higher for working age individuals residing in social housing and 25% lower for working age residents of privately rented properties, when both groups were compared to those in owner-occupied properties (OR 1.75; 95% CI 1.69-1.82; p<.001 and OR 0.75, 95% CI 0.72-0.78, p<.001, respectively). In the unadjusted model, 3.25% of the variance in basic multimorbidity was attributable to differences between areas.

After adjusting for age, sex, and ethnicity, model 2 found strong evidence that the odds of basic multimorbidity were two times greater for working age individuals residing in rented social housing compared to owner-occupied properties (OR 2.04, 95% CI 1.96-2.12, p<.001). There was no longer evidence of an association between tenure and basic multimorbidity for those in privately rented compared to owner-occupied properties (p=0.321). After adjusting for age, sex, and ethnicity, 2.16% of the variance in basic multimorbidity was attributable to differences between areas.

After adjusting for age, sex, ethnicity, BMI, and smoking status (model 3), the odds of basic multimorbidity were reduced for working age residents of rented social housing properties compared to individuals in owner-occupied properties, although strong evidence of an association remained (OR 1.91; 95% CI 1.83-1.99; p<.001). No evidence of an association remained for private renters (p=0.758).

In Table 5-10, model 4 presents the results of the final model adjusted for all individual-level variables included in models 2 and 3, plus household benefits receipt, occupancy, and type. This model found strong evidence that, when compared to working age residents of owner-occupied properties, the odds of basic multimorbidity were 36% higher for working age individuals residing in rented social housing properties and 21% lower for those in privately rented properties (OR 1.36; 95% CI 1.31-1.42; p<.001 and OR 0.79, 95% CI 0. 75-0.83, p<.001, respectively). LBBD-specific IMD quintiles and LBBD-specific IMD housing quintiles were not included in model 4 for the basic multimorbidity outcome as adding these resulted in poorer model fit (indicated by a higher AIC). In this final, fully adjusted model, 1.44% of the variance in basic multimorbidity was attributable to differences between areas.

5.1.3.1 Subgroup and sensitivity analyses

Appendix 6 presents the results of all subgroup analyses conducted for the basic multimorbidity outcome. Evidence was found for interactions between household tenure and household benefits receipt (p<.001), occupancy (p<.001), and household type (p<.001) when terms for each interaction were separately added to model 4. For example, the impact of privately renting compared to owner-occupying on multimorbidity was 71% greater for individuals in households where someone was in receipt of ESA compared to individuals in households receiving no benefits (OR 1.71, 95% CI 1.34-2.19, p<.001), and the impact of social housing compared to owner-occupying on multimorbidity was 15% lower for individuals in households with 3-5 compared to 1-2 occupants (OR 0.85, 95% CI 0.78-0.93, p<.001). See Appendix 6.

Appendix 7 presents the results of all sensitivity analyses conducted for the basic multimorbidity outcome. For the final, fully adjusted model (model 4), associations between household tenure and basic multimorbidity remained similar when the model was either single-level, accounted for household-level clustering or accounted for household and area-level clustering (see Appendix 7). For the final, fully adjusted model (model 4), associations between household tenure and basic multimorbidity also remained similar when selected conditions were excluded from this definition of multimorbidity (see Appendix 7).

Table 5-10: Estimated odds ratios of basic multimorbidity with household tenure for working age adults residing in LBBD in 2019/20 (N=129985)

	Multimorbidity prevalence N (%)	Model 1[a] unadjusted OR (95% CI)	P value	Model 2[b] adjusted OR (95% CI)	P value	Model 3[c] adjusted OR (95% CI)	P value	Model 4[d] adjusted OR (95% CI)	P value
Household tenure									
OOC* (ref)	9278 (39.6)	-		-		-		-	
Social housing	9004 (38.4)	1.75 (1.69-1.82)	<.001	2.04 (1.96-2.12)	<.001	1.91 (1.83-1.99)	<.001	1.36 (1.31-1.42)	<.001
Privately rented	5143 (22.0)	0.75 (0.72-0.78)	<.001	1.02 (0.98-1.06)	0.321	0.99 (0.95-1.04)	0.758	0.79 (0.75-0.83)	<.001
Variance Partition Coefficient (%)		3.25		2.16		1.73		1.44	

Basic multimorbidity = the co-occurrence of two or more chronic conditions within a single individual
*OOC = owner-occupied
[a]Model 1 – an unadjusted model with no covariates
[b]Model 2 - model adjusted for individual-level sociodemographic characteristics: age, sex, and ethnicity
[c]Model 3 – model adjusted for model 2 covariates plus individual-level behavioural characteristics: BMI and smoking status
[d]Model 4 – model adjusted for model 2 and 3 covariates plus household-level sociodemographic characteristics: household benefits receipt, household occupancy and household type

5.1.4 Outcome 2: Physical-mental multimorbidity

The crude prevalence of physical-mental multimorbidity amongst working age residents of LBBD was 4.8% (6214/129985) in 2019/20. Results from bivariate, unadjusted analyses for this outcome are presented in Table 5-9. These results found strong evidence that the odds of physical-mental multimorbidity increased with age (OR 8.02, 95% CI 7.31-8.82 for 55–65-year-old residents compared to 16–29-year-olds) and were lower for males compared to females (OR 0.54, 95% CI 0.51-0.57). There was also strong evidence that the unadjusted odds of physical-mental multimorbidity were lower for Black and Asian working age residents compared to White participants (OR 0.31, 95% CI 0.28-0.33 for participants grouped as Black compared to White, for example). These results found strong evidence that the unadjusted odds of physical-mental multimorbidity were also greater with increased BMI and for working age residents living in households in receipt of benefits (see Table 5-9).

Table 5-11 presents the results from the multilevel models built for the physical-mental multimorbidity outcome. In Table 5-11, model 1 shows the results of a model that found strong evidence that the unadjusted odds of physical-mental multimorbidity were nearly three times higher for working age residents of social housing compared to owner-occupied properties (OR 2.94; 95% CI 2.76-3.13; p<.001). No evidence was found for an association between tenure and physical-mental multimorbidity for those who privately rented (p=0.173). In the unadjusted model, 6.91% of the variance in physical-mental multimorbidity was attributable to differences between areas.

After adjusting for age, sex, and ethnicity, model 2 found strong evidence that the odds of physical-mental multimorbidity remained three times higher for working age individuals residing in rented social housing properties and were 34% higher for those in privately rented properties, when both groups were compared to residents of owner-occupied properties (OR 2.91, 95% CI 2.72-3.10, p<.001 and OR 1.34, 95% CI 1.25-1.45, p<.001, respectively). After adjusting for age, sex, and ethnicity (model 2), 2.93% of the variance in physical-mental multimorbidity was attributable to differences between areas.

After adjusting for age, sex, ethnicity, BMI, and smoking status (model 3), the odds of physical-mental multimorbidity remained higher for working age individuals residing

in rented social housing (OR 2.62; 95% CI 2.45-2.79; p<.001) and in privately rented properties (OR 1.28; 95% CI 1.19-1.38; p<.001), when both groups were compared to residents of owner-occupied properties.

In Table 5-11, model 4 presents the results of the final model adjusted for all individual-level variables included in models 2 and 3, plus household benefits receipt, occupancy, and type. This model found strong evidence that, when compared to working age residents of owner-occupied properties, the odds of physical-mental multimorbidity were 47% higher for residents of social housing and 15% lower for private renters (OR 1.47; 95% CI 1.37-1.58; p<.001 and OR 0.85, 95% CI 0.78-0.92, p<.001, respectively). LBBD-specific IMD quintiles and LBBD-specific IMD housing quintiles were not included in model 4 as adding these resulted in poorer model fit (indicated by a higher AIC). In this final, fully adjusted model, 1.66% of the variance in physical-mental multimorbidity was attributable to differences between areas.

5.1.4.1 Subgroup and sensitivity analyses

Appendix 6 presents the results of all subgroup analyses conducted for the physical-mental multimorbidity outcome. Evidence was found for interactions between household tenure and household benefits receipt (p<.001), occupancy (p=0.010), and household type (p=0.045) when terms for each interaction were separately added to model 4. For example, the impact of privately renting compared to owner-occupying on multimorbidity was 36% greater for individuals in single adult households compared to individuals in households with adults and children (OR 1.46, 95% CI 1.06-1.74, p=0.015), and 23% lower for individuals in households with 6-10 compared to 1-2 occupants (OR 0.77, 95% CI 0.60-0.99, p=0.038). See Appendix 6.

Appendix 7 presents the results of all sensitivity analyses conducted for the physical-mental multimorbidity outcome. For the final, fully adjusted model (model 4), associations between household tenure and physical-mental multimorbidity remained similar when the model was either single-level, accounted for household-level clustering or accounted for household and area-level clustering (see Appendix 7). For the final, fully adjusted model (model 4), associations between household tenure and physical-mental multimorbidity also remained similar when selected conditions were excluded from this definition of multimorbidity (see Appendix 7).

Table 5-11: Estimated odds ratios of physical-mental multimorbidity with household tenure for working age adults residing in LBBD in 2019/20 (N=129985)

	Multimorbidity prevalence N (%)	Model 1[a] unadjusted OR (95% CI)	P value	Model 2[b] adjusted OR (95% CI)	P value	Model 3[c] adjusted OR (95% CI)	P value	Model 4[d] adjusted OR (95% CI)	P value
Household tenure									
OOC* (ref)	1801 (29.0)	-		-		-		-	
Social housing	3085 (49.6)	2.94 (2.76-3.13)	<.001	2.91 (2.72-3.10)	<.001	2.62 (2.45-2.79)	<.001	1.47 (1.37-1.58)	<.001
Privately rented	1328 (21.4)	1.05 (0.98-1.13)	0.173	1.34 (1.25-1.45)	<.001	1.28 (1.19-1.38)	<.001	0.85 (0.78-0.92)	<.001
Variance Partition Coefficient (%)		6.91		2.93		2.34		1.66	

Physical-mental multimorbidity = the co-occurrence of two or more chronic conditions within a single individual, one of which must be depression or anxiety and one of which must be a physical condition
*OOC = owner-occupied
[a]Model 1 – an unadjusted model with no covariates
[b]Model 2 - model adjusted for individual-level sociodemographic characteristics: age, sex, and ethnicity
[c]Model 3 – model adjusted for model 2 covariates plus individual-level behavioural characteristics: BMI and smoking status
[d]Model 4 – model adjusted for model 2 and 3 covariates plus household-level sociodemographic characteristics: household benefits receipt, household occupancy and household type

5.1.5 Outcome 3: Complex multimorbidity

The crude prevalence of complex multimorbidity amongst working age residents of LBBD was 6.0% (7857/129985) in 2019/20. Results from bivariate, unadjusted analyses for this outcome are presented in Table 5-9. These results found strong evidence that the odds of complex multimorbidity also increased markedly with age (OR 29.6, 95% CI 26.5-33.3 for 55–65-year-old residents compared to 16–29-year-olds) and were lower for males compared to females (OR 0.74, 95% CI 0.71-0.78). There was also strong evidence that the unadjusted odds of complex multimorbidity were lower for Black and Asian working age residents compared to White participants (OR 0.54, 95% CI 0.50-0.57 for participants grouped as Black compared to White, for example). The results found strong evidence that the unadjusted odds of complex multimorbidity were greater with increased BMI, and for smokers and ex-smokers compared to non-smokers (see Table 5-9).

Table 5-12 presents the results from the multilevel models built for the complex multimorbidity outcome. In Table 5-12, model 1 shows the results of a model that found strong evidence that the unadjusted odds of complex multimorbidity were just over two times higher for working age residents of social housing and 24% lower for residents of privately rented properties, when both groups were compared to those in owner-occupied properties (OR 2.07; 95% CI 1.96-2.18; p<.001 and OR 0.76, 95% CI 0.72-0.82, p<.001, respectively). In this unadjusted model, 4.30% of the variance in complex multimorbidity was attributable to differences between areas.

After adjusting for age, sex, and ethnicity, model 2 found strong evidence that the odds of complex multimorbidity were nearly two and a half times higher for working age residents of social housing and 19% higher for those in privately rented properties, when both groups were compared to residents of owner-occupied properties (OR 2.40, 95% CI 2.26-2.54, p<.001 and OR 1.19, 95% CI 1.12-1.28, p<.001, respectively). After adjusting for age, sex, and ethnicity (model 2), 1.95% of the variance in complex multimorbidity was attributable to differences between areas.

After adjusting for age, sex, ethnicity, BMI, and smoking status (model 3), the odds of complex multimorbidity remained higher for working age residents of social housing (OR 2.24; 95% CI 2.12-2.37; p<.001) and privately rented properties (OR 1.17; 95%

CI 1.09-1.25; p<.001), when both groups were compared to residents of owner-occupied properties.

In Table 5-12, model 4 presents the results of the final model adjusted for all individual-level variables included in models 2 and 3, plus household benefits receipt, occupancy, and type. This model found strong evidence that, when compared to working age residents of owner-occupied properties, the odds of complex multimorbidity were 34% higher for working age individuals residing in social housing and 19% lower for those in privately rented properties (OR 1.34; 95% CI 1.26-1.44; p<.001 and OR 0.81, 95% CI 0.75-0.87, p<.001, respectively). LBBD-specific IMD quintiles and LBBD-specific IMD housing quintiles were not included in model 4 for the complex multimorbidity outcome as adding these resulted in poorer model fit (indicated by a higher AIC). In the final, fully adjusted model, 1.03% of the variance in complex multimorbidity was attributable to differences between areas.

5.1.5.1 Subgroup and sensitivity analyses

Appendix 6 presents the results of all subgroup analyses conducted for the complex multimorbidity outcome. Evidence was found for interactions between household tenure and household benefits receipt (p<.001), occupancy (p<.001), and household type (p<.001) when terms for each interaction were separately added to model 4. For example, the impact of social housing compared to owner-occupying on multimorbidity was 38% lower for individuals in households with 6-10 compared to 1-2 occupants (OR 0.62, 95% CI 0.50-0.75, p<.001). In addition, the impact of privately renting compared to owner-occupying on multimorbidity was 69% greater for individuals in older cohabiting adult households compared to households with adults and children (OR 1.69, 95% CI 1.26-2.27, p<.001). See Appendix 6.

Appendix 7 presents the results of all sensitivity analyses conducted for the complex multimorbidity outcome. For the final, fully adjusted model (model 4), associations between household tenure and complex multimorbidity remained similar when the model was either single-level, accounted for household-level clustering or accounted for household and area-level clustering (see Appendix 7). For the final, fully adjusted model (model 4), associations between household tenure and complex multimorbidity also remained similar when selected conditions were excluded from this definition of multimorbidity (see Appendix 7).

Table 5-12: Estimated odds ratios of complex multimorbidity with household tenure for working age adults residing in LBBD in 2019/20 (N=129985)

	Multimorbidity prevalence N (%)	Model 1[a] unadjusted OR (95% CI)	P value	Model 2[b] adjusted OR (95% CI)	P value	Model 3[c] adjusted OR (95% CI)	P value	Model 4[d] adjusted OR (95% CI)	P value
Household tenure									
OOC* (ref)	2853 (36.3)	-		-		-		-	
Social housing	3450 (43.9)	2.07 (1.96-2.18)	<.001	2.40 (2.26-2.54)	<.001	2.24 (2.12-2.37)	<.001	1.34 (1.26-1.44)	<.001
Privately rented	1554 (19.8)	0.76 (0.72-0.82)	<.001	1.19 (1.12-1.28)	<.001	1.17 (1.09-1.25)	<.001	0.81 (0.75-0.87)	<.001
Variance Partition Coefficient (%)		4.30		1.95		1.54		1.03	

Complex multimorbidity = the co-occurrence of three or more chronic conditions affecting three or more different body systems within a single individual

*OOC = owner-occupied

[a]Model 1 – an unadjusted model with no covariates

[b]Model 2 - model adjusted for individual-level sociodemographic characteristics: age, sex, and ethnicity

[c]Model 3 – model adjusted for model 2 covariates plus individual-level behavioural characteristics: BMI and smoking status

[d]Model 4 – model adjusted for model 2 and 3 covariates plus household-level sociodemographic characteristics: household benefits receipt, household occupancy and household type

5.2 Discussion

5.2.1 Summary of study findings

This study used linked health and council data to illustrate how knowledge generated from the analysis of such data could advance our understanding of household-level social determinants of multimorbidity (addressing Aim 2 of this thesis). Risk of multimorbidity was greater for working age residents of social housing and lower for residents of privately rented properties when both groups were compared to owner-occupiers. These associations remained after adjusting for a range of individual and household-level sociodemographic, economic, and behavioural characteristics, and were consistent across different definitions of multimorbidity. Other household-level variables – household benefits receipt, type, and occupancy – were important explanatory factors for the reported associations between tenure and multimorbidity but did not explain all of the additional risk experienced by those in social housing nor all of the decreased risk experienced by those who privately rent.

5.2.2 Comparisons to existing literature

5.2.2.1 Prevalence estimates

In this study, prevalence estimates of basic multimorbidity (18.0%), physical-mental multimorbidity (4.8%) and complex multimorbidity (6.0%) amongst working age residents of LBBD were in keeping with previous multimorbidity prevalence estimates for this age group (Taylor *et al.*, 2011; Barnett *et al.*, 2012; Johnson-Lawrence, Zajacova and Sneed, 2017; Head, Fleming, Kypridemos, Schofield, *et al.*, 2021). Prevalence of all types of multimorbidity increased dramatically with age and amongst females compared to males, consistent with previous literature (Barnett *et al.*, 2012; Rocca *et al.*, 2014; Bobo *et al.*, 2016; The Academy of Medical Sciences, 2018). However, multimorbidity prevalence was lower for all ethnic minority groups when compared to White participants, which contradicts many previous studies (Rocca *et al.*, 2014; Bobo *et al.*, 2016; The Academy of Medical Sciences, 2018).

5.2.2.2 Possible explanations for associations between household tenure and multimorbidity

The main objective for this study was to examine and quantify *associations* between selected household characteristics (informed by my review) and multimorbidity (see

Chapter 2). I found that the risk of multimorbidity was higher for social housing tenants compared to owner-occupiers, which aligns with findings from a recent study conducted in Northern Ireland that was identified in my review (Ferry *et al.*, 2020). However, my findings disagree with those from another study identified in my review that was conducted in Hong Kong (Chung *et al.*, 2015). As such, the higher risk of multimorbidity found for those in social housing may be because characteristics intrinsic to social housing *in the UK* cause the development of multiple chronic conditions. Using the UK Household Longitudinal Study, Clair and Hughes (2019) found strong evidence that, when compared to owner-occupiers, social housing tenants have higher levels of the biomarker C-reactive protein, which is a marker of inflammation and is associated with increased risk of various chronic conditions (Ansar and Ghosh, 2013; Wium-Andersen *et al.*, 2013; Clair and Hughes, 2019). Previous work has also shown that residents of social housing in Scotland are more likely to be exposed to health-damaging factors such as overcrowding and dampness, and less likely to be exposed to health-protecting factors such as gardens (Ellaway and Macintyre, 1998; Macintyre *et al.*, 2003). Indeed, 8.7% of social housing households surveyed in the 2019/20 English Housing Survey (EHS) reported overcrowding compared to 1.2% of owner-occupied households. On top of this, it has been suggested that social housing tenants have less control and choice over their properties, and are less able to access and leave their property, whilst owner-occupying affords ontological security – the sense of security and control afforded when owning your own home (Hiscock *et al.*, 2001; Scanlon and Kochan, 2011). Some evidence therefore suggests that social housing may expose individuals to multiple "hard" (material) and "soft" (psychological) factors that could interact to cause or exacerbate various chronic physical and mental health conditions and lead to multimorbidity (Shaw, 2004).

Cross-sectional studies are useful for establishing associations between exposures and outcomes at a given point in time and therefore my study design allowed me to address the main objective of this study. They are, however, unable to explore causal relationships or mediators of a relationship. As such, whilst it is plausible that characteristics related to different tenure types may cause multimorbidity, it is also likely that multimorbidity causes residents with certain socioeconomic and health characteristics to cluster by tenure type. For example, the 1988 Housing Act requires social housing to be allocated to those most in need based on selected criteria, one

of which is ill-health. This may explain the poorer relative health of social housing tenants as ill-health and multimorbidity status may directly lead to individuals being allocated social housing. Indeed, the 2019/20 EHS found that 53.8% of respondents in socially rented households had a long-term illness or disability compared to 25.0% in privately rented properties and 30.8% in owner-occupied properties (see Table 5-13).

Table 5-13: Household characteristics by tenure in 2019/20 English Housing Survey

	Owner-occupied	Privately rented	Socially rented	All households
Long term illness or disability (%)				
Yes	30.8	25.0	53.8	33.6
No	69.2	75.0	46.2	66.4
Economically inactive[a] (%)	3.3	11.0	28.9	9.1
Household income[b]				
Quintile 1 (lowest)	13.0	20.0	47.3	20.0
Quintile 2	18.4	21.8	24.0	20.0
Quintile 3	20.2	22.6	16.3	20.0
Quintile 4	23.0	19.6	8.9	20.0
Quintile 5 (highest)	25.5	15.9	3.5	20.0

[a]For household reference person
[b]Weekly gross household income

Changes to the English housing landscape since the 1980s may then explain the strength of the observed associations between social housing and multimorbidity. In 1980, the UK government introduced a Right to Buy policy allowing some social housing tenants to legally buy the properties they live in at a discount. This policy, and its future iterations, enabled wealthier residents to buy their properties and saw those unable to buy concentrated in a diminishing number of social housing properties, leading to tenure types that were more segregated by economic status and social class (see Table 5-13). It is, therefore, highly plausible that reverse causality explains the associations found in my analyses. In LBBD, over 7,000 households are on the waiting list for council housing yet only approximately 600 council homes become available each year, and these are allocated to the most in need (Barking and Dagenham Council, 2019).

In this study, I found that the impact of tenure on multimorbidity was greater for those in households where someone was in receipt of certain benefits. This aligns with findings from Ferry and colleagues that suggest associations between tenure and multimorbidity differ with property value (those in lower value properties reporting greater multimorbidity risk) (Ferry *et al.*, 2020). I also found that the impact of tenure on multimorbidity was greater for single adult households, aligning with findings from two high quality studies identified by my review (Cantarero-Prieto, Pascual-Sáez and Blázquez-Fernández, 2018; Henchoz *et al.*, 2019). However, the lower risk of multimorbidity found for those who privately rent compared to owner-occupy contradicts previous evidence suggesting that private renters report poorer overall health than owner-occupiers (Macintyre *et al.*, 2003; Gibson *et al.*, 2011), as well as contradicts findings reported by several studies identified in my review (Johnson-Lawrence, Zajacova and Sneed, 2017; Ferry *et al.*, 2020). Clair and Hughes (2019) found that, when compared to owner-occupiers, private renters have higher levels of C-reactive protein after adjusting for various sociodemographic and health-related variables, which also contradicts the findings in this study (Clair and Hughes, 2019). This may be because the data I analysed (The Care City Cohort) overestimated the number of privately rented households in 2019/2020 and underestimated the number of owner-occupying households, leading to misclassification biases in the tenure variable (see 5.1.1.2). Possible misclassification biases and their influences on model estimates will be discussed in more detail in section 5.2.2.3.

5.2.2.3 Error and bias in measurement

Multimorbidity outcomes: My findings suggested that selection biases were not introduced in selected variables originating from primary care records as a result of the success of data linkages, which is in keeping with previous analyses of this data (Shand, 2020). However, whilst this study used well-established, publicly available code lists that captured conditions missed by previous multimorbidity definitions (see Table 4-1), it is possible that some individuals were misclassified as living without multimorbidity and that the proportion misclassified varied by tenure type. In primary care, mental health conditions are commonly underdiagnosed (Cepoiu *et al.*, 2007) and conditions managed in secondary care (such as chronic liver disease) may be under-recorded (Walker *et al.*, 2021), yet these conditions are more common in more deprived populations (Barnett *et al.*, 2012; Williams *et al.*, 2020). It is also well acknowledged that when a patient is recognised to be living with a chronic condition,

this can lead to further tests that identify additional conditions. Patients with early stages of CKD are more likely to be diagnosed in primary care if they have other risk factors for CKD such as diabetes (Mcdonald *et al.*, 2016) and hypertension is more likely to be diagnosed earlier for patients with other comorbidities (Tran *et al.*, 2021). CKD, diabetes and hypertension are, in turn, more prevalent in more deprived areas (Lyratzopoulos *et al.*, 2006; Cox *et al.*, 2007; Hossain *et al.*, 2012), where social housing is the most common tenure (Macintyre *et al.*, 2003). On top of this, individuals in different tenures may have differential health care access for diagnosis. Overall, these issues may have introduced biases in the classification of those with multimorbidity by tenure type and biased study estimates towards the null.

Despite these potential misclassification errors, excluding selected conditions from multimorbidity definitions did not substantially change final model estimates for each multimorbidity outcome (see Appendix 7). This is surprising as considerable debate remains around what conditions should be included or excluded in a definition of multimorbidity (see section 1.2.1). This suggests that, whilst misclassification errors in primary care data may be problematic for accurately estimating multimorbidity prevalence (Fortin *et al.*, 2012), they may make minimal difference when examining associations between household tenure and multimorbidity.

Tenure exposure: Misclassification biases in the tenure variable may explain why I found lower risk of multimorbidity for private renters compared to owner occupiers. Comparisons of the Care City Cohort data to mid-2019 tenure estimates from ONS suggest LBBD may have underestimated the number of owner-occupying households and overestimated the number of privately rented households in 2019/2020 (see section 5.1.1.2). However, it is important to note that ONS estimates are based on modelling of 2011 census data, which leaves considerable room for error and makes it difficult to meaningfully interpret ONS estimates. Misclassification biases in the tenure variable may have also arisen if owner-occupiers in LBBD privately rented rooms in 2019/2020. In these circumstances, a household would have been recorded by the council as an owner-occupied property but would have included both owner-occupying and privately renting residents. Private renters who co-reside with their owner-occupying landlords may differ systematically in their health compared to private renters who do not. This may explain my observed associations.

Ethnicity: Comparisons of the council ethnicity variable with 2011 census data and 2020 GLA ethnic projections for LBBD suggested the variable may have overestimated White and underestimated Black ethnicities, leading to misclassification errors in the ethnicity variable used in this study (see section 5.1.1.2). This is a common challenge in council data where the missingness of ethnicity data can differ depending on ethnic group (Office for National Statistics, 2019c). This may be due to methods used by Origins – a company LBBD work with to address ethnicity missingness (see section 4.2.3.3). LBBD had previously found that, in 2011, Origin's data performed well for most ethnic groups when compared with the national census yet overestimated "White-Other" ethnicities and underestimated "Black-Caribbean" ethnicities. Another reason may be because the variable did not code "Mixed/Multiple" ethnic categories and these individuals were misclassified as White. This could explain why I found unexpected results suggesting lower multimorbidity prevalence for all ethnic minority groups compared to White participants.

BMI, and smoking status: Missingness on BMI and smoking status variables extracted from primary care records was associated with all multimorbidity outcomes and household tenure, indicating non-response biases in the completeness of these variables (see section 5.1.1.2). This is in keeping with evidence suggesting that completeness and misclassification of BMI can be a challenge in primary care (Bhaskaran *et al.*, 2013), with recording more common in those with comorbidities (Nicholson *et al.*, 2019). This is also in keeping with evidence suggesting that recording of smoking status is more common for those with chronic conditions and greater social deprivation (Taggar *et al.*, 2012). Overall, this suggests that adjusting for BMI and smoking status when investigating associations between tenure and multimorbidity may have falsely decreased risk estimates.

Household benefits receipt: In this study, financial circumstances were crudely adjusted for using a marker of household benefits receipt. It is also likely that misclassification biases were present in this variable as some individuals may have lived in households where a resident was in receipt of benefits, but not in receipt of housing benefit. These individuals would have been misclassified as not living in a household in receipt of benefits. In addition, this variable only captured benefits receipt and not eligibility, which could have missed important information about a

household's financial circumstances. As such, it is likely that this variable will not have fully captured information on individual or household financial circumstances and some residual confounding remains.

In the 2019/20 EHS, 28.9% of respondents living in socially rented households were economically inactive, compared to 11.0% in privately rented properties and 3% in owner-occupied properties. In addition, 47.3% of socially rented households were in the lowest quintile for household income compared to 20.0% of privately rented households and 13.0% of owner-occupied households (see Table 5-13). Lower household income is, in turn, consistently associated with increased multimorbidity, as found in my review (Chapter 3), and financial worries are particularly associated with poor mental health (Bisgaier and Rhodes, 2011). Information on household income or savings could have better controlled for household financial circumstances and could plausibly explain the observed associations between tenure type and multimorbidity reported in this study.

Ecological fallacy and area-level measures of deprivation: A further important consideration in this study is ecological fallacy – a problem that occurs when factors related to a group are incorrectly associated with individual-level outcomes (Freedman, 1999). Attributing household-level characteristics such as benefits receipts to individuals and their multimorbidity status may incorrectly assume that all individuals in a household are equally affected by household benefits receipt. This may have led to estimates that suggest living in a household where someone is in receipt of benefits confounds associations between tenure and multimorbidity with greater strength than in reality.

Adding two area-level measures of relative deprivation led to worse model fit for all three multimorbidity outcomes, which suggests that area-level characteristics are not associated with multimorbidity over and above individual and household-level characteristics. On one hand, this is surprising as evidence suggests that social housing is typically concentrated in the most deprived areas where exposure to health-harming area-level stressors such crime is more likely and access to local health-protecting amenities less likely (Macintyre *et al.*, 2003). The reverse is true for owner-occupied properties, whilst private rental properties are typically found in areas with deprivation levels between the two (Macintyre *et al.*, 2003). However, this is

perhaps unsurprising as most areas in LBBD are within the two most deprived national IMD quintiles. Evidence suggests tenure mix is higher in the tighter housing markets of London and, as such, the area-level deprivation measures used in this study may cover geographical areas that are too large to capture enough variation to model socioeconomic inequalities in multimorbidity at the area-level (Livingston, Kearns and Bailey, 2013). This may explain why area-deprivation measures did not explain associations between tenure and multimorbidity in this study.

5.2.3 Strengths and limitations of this study

One strength of this study is that it is the first to explore associations between household tenure and multimorbidity prevalence in the English context. This is particularly important as evidence suggests that context-specific factors such as housing policy, the degree of homeownership, and the supply and conditions of rented housing may profoundly influence the status, stigma and meaning associated with residing in different tenure types across geographies and over time (Shaw, 2004).

A further strength is the use of novel health and council data linked at individual and household-levels. Without these linkages, it would not have been possible to examine associations between tenure and multimorbidity. My use of this data also means the likelihood that my reported associations are due to chance is minimal as the data includes participants from almost the entirety of one local area and, as such, the sample size afforded is large. A large sample size achieves this by reducing uncertainty in model estimates, decreasing confidence intervals and increasing precision (Zaccai, 2004). This study is also not hampered by selection and attrition biases to the same extent as previous studies in this area that utilise survey data (see Chapter 3).

In this study, I used well-established, publicly available code lists to capture the presence or absence of each condition included in my definitions of multimorbidity. I operationalised three different definitions of multimorbidity that included the most used definition in the literature as well as more complex definitions, one of which tried to capture physical and mental health dimensions of multimorbidity. However, whilst researchers are encouraged to examine the most frequently seen and burdensome disease clusters and combinations (The Academy of Medical Sciences, 2018), my definition of physical-mental multimorbidity included only depression and anxiety,

which may have missed any important associations between tenure and SMI that clusters with physical health conditions. In addition, these flags did not account for disease severity or symptom burden on the patient, or other important dimensions of multimorbidity such as frailty (Hanlon *et al.*, 2018; Rogers *et al.*, 2021). These could have captured important information about more complex multimorbidity profiles.

One of the key limitations of this study is generalisability, which refers to the extent to which study findings can be applied to other settings and contexts (Zaccai, 2004). The individuals under study here resided in a densely populated and socially deprived borough in North East London, with a younger and more ethnically diverse population compared to the rest of England (London Borough of Barking and Dagenham, 2018). The extent to which these findings can be applied to less deprived, more rural settings that contain a different tenure profile is under question. However, these findings could be relevant to other urban areas with young and deprived populations outside of LBBD. These findings also act as a use case for creating, using, and analysing such linked datasets to understand the social determinants of local public health concerns and generate knowledge that could inform equitable decision-making.

Another limitation of this study is that some residual confounding may remain. First, my analyses suggest that there may be some errors and misclassification biases in the measurement of key variables such as tenure and household benefits receipt that may contribute to residual confounding. I also did not have information on certain characteristics such as household income and overcrowding that could have plausibly confounded associations. Second, I could not account for household-level clustering in my multi-level modelling which is likely to be more important than confounding variables related to clustering within areas. As such, participants that were clustered in households will likely have shared similar household-specific characteristics that were not controlled for in the analyses. Nevertheless, my subsequent sensitivity analyses suggest that cluster-constant variables had little influence on the reported associations between tenure and multimorbidity (see Appendix 7).

Finally, this study is cross-sectional. Whilst this allowed me to explore associations between household tenure and multimorbidity and address the objective of this study, I was not able to explore possible causal relationships. As described in section 5.2.2.2, tenure type may causally contribute to multimorbidity status, and/or

multimorbidity may cause residents with certain socioeconomic and health characteristics to cluster by tenure type. More longitudinal analyses are needed to assess causal relationships between tenure and multimorbidity by establishing temporality (the risk factor precedes the outcome) and specificity (changes in a risk factor are linked to changes in an outcome).

5.3 Chapter summary

This chapter presented the results of my quantitative study that used linked health and council data to examine and quantify associations between household tenure and multimorbidity amongst working age residents of LBBD in 2019/2020. I found that the risk of multimorbidity was greater for working age adults living in social housing and lower for residents of privately rented properties when both groups were compared to owner-occupiers. Differences in risk could not be explained by area-level measures of deprivation that are commonly used in the literature (see Chapter 3) and were only partly explained by other individual and household-level sociodemographic, behavioural, and economic characteristics.

This study acts as a use case for creating, using, and analysing linked health and council datasets to understand the social determinants of local public health concerns. In particular, it would not have been possible to examine associations between tenure and multimorbidity without the linkage of individuals at individual and household-levels (based on a shared UPRN). Findings were consistent across all multimorbidity outcomes and different definitions of multimorbidity (see Appendix 7), illustrating the strength of household tenure as an exposure for multimorbidity and the importance of capturing data on, and understanding, household-level SDoH. This knowledge could be used by commissioners and planners to inform strategic and equitable health and care decision-making for those with, or at risk of, multimorbidity.

Chapter 6 Senior leaders' experiences of using analytics to inform strategic and equitable health and care decision-making: Introduction and Methods

6.1 Introduction

As I described in Chapter 1, it is often assumed that linked data *will* improve decision-making, care, and the equity of health and care services. However, it is unclear whether these aspirations will be realised. In this chapter, I present the introduction and methodology for my third study, which investigated how knowledge generated from the analysis of linked health and council data might inform strategic and equitable health and care decision-making.

6.1.1 Origins of this study

One of my research funders, the NIHR ARC North Thames, partners with North Central London's (NCL) ICS. NCL covers the five North London boroughs of Islington, Camden, Haringey, Barnet, and Enfield. It is a digitally engaged and innovative ICS that has a history of encouraging leaders to use data for decision-making. For example, NCL and their industry partner Cerner are in the process of creating a platform that contains data linked across NHS and council records for residents of NCL to facilitate their formation of an ICS (North London Partners in Health and Care, 2021). The intention is for the platform, HealtheIntent, to also be made available to health and care professionals to support them in providing more proactive health and care (North London Partners in Health and Care, 2021). As part of the partnership between the NIHR ARC North Thames and NCL ICS, my primary supervisor (Dr Jessica Sheringham) is invited to sit on NCL's Analytics Board, and I have been invited to attend several of their meetings.

The idea for this study was borne out of an expressed NCL need to understand how leaders use data for strategic decision-making from the perspective of senior leaders. As described in section 1.4.4, senior leaders utilise various sources of evidence to inform strategic decision-making and one possible source is knowledge generated from the analysis of administrative data extracted from health or care records of

residents ('analytics') (Clarke *et al.*, 2013). In the first instance, NCL's Analytics Board wanted to better understand local senior leaders' skills, knowledge, and capacity to obtain and interpret analytics outputs. NCL then wanted to develop a training plan to address any local training needs around analytics.

The wider study questions arose from subsequent conversations around how local leaders currently use or would use knowledge generated from the analysis of linked data, such as data linked across health and council settings in HealtheIntent. For example, in previous literature, operational barriers to generating high-quality analytics have been well described and some findings suggest leaders do not always value and use analytics for decision-making (Shaw *et al.*, 2013; Beenstock *et al.*, 2014; Van Panhuis *et al.*, 2014; Bardsley, 2016; Kneale *et al.*, 2017; Bardsley, Steventon and Fothergill, 2019). NCL's Analytics Board were interested in understanding if and how these barriers, plus further and wider barriers and facilitators of analytics use, influenced decision-making locally. In addition, one of the explicit aims for HealtheIntent is for it to enable local health and care professionals and leaders to capture and address local health inequalities. Members of NCL's Analytics Board were therefore also interested in understanding barriers and facilitators of analytics use for local strategic health and care decision-making *that considers health inequalities*.

This study aimed to address some of these gaps in local knowledge rather than gaps in the literature. To address the latter would first require a systematic exploration of the literature in this area, and a systematic review of this kind has not been conducted as part of this thesis. In addition, at present, the majority of senior leaders across NCL who make strategic health and care decisions do not have access to data linked across health, council and other care settings. I have therefore focused this study on how knowledge generated from the analysis of health and care records *might* be used to inform decision-making in this context.

I presented the study ideas to the Analytics Board in early 2019 and they agreed to support the study and, in particular, support study recruitment. More details on how NCL have been involved with this study can be found in sections 6.2.1 and 6.2.5 of this chapter.

6.1.2 Study overview

Qualitative research methods are useful when investigating new or underexplored topics. For this study, semi-structured interviews were chosen as the most suitable qualitative method as they allow a comprehensive and fairly systematic exploration of a specific topic of interest (Jamshed, 2014). I chose to interview participants individually as I wished to discuss participants' activities in their workplace, their professional relationships, and their organisational cultures and/or practices. I felt that an individual interview could facilitate an environment where participants felt more comfortable and able to talk about these topics compared to in a focus group, for example.

As little is known in this area and this work was fairly exploratory, I conducted a pilot study before starting the full study in order to inform the wider research questions, develop the topic guide and inform an application for ethical approval. Findings from this pilot study were not incorporated into the full study findings (known as an 'external' pilot study). Pilot studies are valuable as they also allow one to test the feasibility of study procedures, tools and recruitment, and ensure the safety of participants prior to a full study (Sampson, 2004). The process of conducting the external pilot study also enabled me to develop my competence and confidence as a qualitative researcher (Wray, Archibong and Walton, 2017). This was my first experience with qualitative research, and I valued this learning experience.

For the external pilot study, I recruited seven senior leaders from my study sites of interest (see section 6.2.1). I asked participants for feedback on the interview and recruitment process to determine the acceptability and feasibility of study procedures. I analysed study interviews following the methodological steps described in section 6.2.4. In September 2019, I presented the preliminary findings and themes from this pilot study at NCL's Analytics Board meeting. The board found the findings helpful and requested the full study findings to inform their analytics strategy.

Several changes were made for the full study. These were informed by feedback from the board, the notes I recorded in my reflexive journal and my personal reflections when analysing the pilot study findings. These reflections and changes are detailed in the relevant sections of this chapter.

6.1.3 Research questions

Prior to the pilot study, one of the study aims had been to understand senior leaders' perceived value of analytics for informing strategic and equitable health and care decision-making. In an attempt to address this aim, I directly asked pilot study participants about their perceived value of analytics. On reflection, this encouraged participants to discuss analytics in an abstract manner. As the pilot study progressed, I therefore started to ask participants to give examples of when they had used analytics to inform a strategic health and care decision. I thought that centring the discussion on a specific example and asking questions around it would tap into their perceived value of analytics and encourage fewer abstract discussions. On reflection, this led to examples that were too disparate and technical for me to be able to understand, ask probing questions and analyse collectively.

In light of these reflections, I no longer sought to understand senior leaders' perceived value of analytics for the full study. I instead asked senior leaders if they could describe an example of a recent strategic health and care decision they had made *around health inequalities*. This was in an effort to obtain examples that were more similar and therefore easier to collectively analyse. Health inequalities was chosen as a topic as it was of interest to NCL, likely touches many leaders' roles, and is of interest to this thesis as a whole (see section 6.1.1).

The research questions for the full study were:

1. When and how is knowledge generated from the analysis of residents' health or care records used to inform strategic and equitable health and care decision-making?
2. What are the barriers and facilitators of analytics use in this context?
3. How do leaders differ in their experiences of and responses to identified barriers and facilitators?
4. Are there distinct types of analytics users amongst this sample of senior leaders?
5. What should local areas consider when developing programmes of work that aim to improve senior leaders' analytics use for the explored purposes?

Contributions of others to this study: Dr Sarah Dougan, London Borough of Camden, Will Huxter, NCL, and other members of NCL's Analytics Board helped to develop the research questions and supported study recruitment. Sarah Beardon (SB), UCL Department of Applied Health Research and UCL Faculty of Laws, double coded a subset of my interview transcripts and checked the final coding frame. Dr Silvie Cooper (SC), a senior qualitative researcher in UCL's Department of Applied Health Research, provided advice and guidance during the interview, analysis, and write-up stages of the study. Katherine Körner (KK), London Borough of Islington, reviewed emergent themes.

6.2 Methods

6.2.1 Sampling and recruitment

For this study, I initially planned to recruit senior leaders from one of two study sites in North and East London, acting as two case studies (Yin, 2003):

1. NCL, which covers the London Boroughs of Islington, Camden, Barnet, Haringey, and Enfield (as described in section 6.1.1).
2. Care City – a collaboration of North East London health, care and third sector organisations that support the local health and care system in LBBD (as described in section 4.2.1).

Care City was chosen as a second study site as it is also a partner organisation with the ARC North Thames. NCL and Care City also correspond to the areas where I explored obtaining linked data. NCL is actively pursuing linkage of health and council records and Care City is post-linkage. Organisations in these collaborations include CCGs, councils, and a range of health and care providers including NHS foundation trusts. Participants were included if they met the inclusion criteria:

1. A senior leader working in any of the study site's constituent organisations,
2. A senior leader responsible for strategic decision-making for their own organisation or local health and care system.

For the pilot study, I mainly recruited participants through NCL's Analytics Board. Because of this, I felt that several pilot study participants spoke more towards the board's views on analytics in light of their planned workstreams, rather than towards their own personal experiences of using analytics within a specific decision-making context. When I presented the pilot study preliminary findings to NCL's Analytics Board, it was raised that the full study should also aim to recruit a more representative balance of participants. In particular, participants from the London Boroughs of Haringey and Barnet were underrepresented in the pilot study, as well as female participants, ethnic minority participants, and clinical, social care and public health leads. My initial strategy to recruit through NCL's Analytics Board for the pilot study may have contributed to the lack of representativeness in my sample and may have missed less heard views, including views of those less interested in analytics.

For the full study, I used a combination of purposive and snowball sampling to recruit a more representative sample of participants (Palinkas *et al.*, 2015; Naderifar, Goli and Ghaljaie, 2017). I chose to continue to recruit via the Analytics Board and also recruited participants through one of my thesis supervisors who works with local senior leaders who do not sit on the board. I asked stakeholders (individuals sitting on the Analytics Board and not sitting on the board) to put me in contact with eligible colleagues who did not sit on the board and could improve the representativeness of my sample (Naderifar, Goli and Ghaljaie, 2017). I started the full study by recruiting from my first case study site, NCL.

With the onset of the Covid-19 pandemic in the UK, logistical reasons prevented expanding this study to Care City, my second case study site. Care City had significantly reduced capacity and an increased workload at the onset of the pandemic. In addition, decision-making in health and care organisations became very reactive and less strategic, instead focusing on adapting health and care services so they could function in some capacity whilst keeping staff and patients safe. At this point in time, I had also reached data saturation with the interviews I had conducted. I therefore felt it was appropriate to suspend recruitment at Care City and not pursue recruitment from another study site. All participants for the full study were recruited from my first study site – NCL.

In both the pilot and full study, all potential participants were first identified and contacted by key leaders at NCL. Potential participants were first approached by key leaders via email and given the information sheet. I then contacted potential participants via email and answered any questions. If interested in the study, I checked participants met the inclusion criteria and arranged an interview date and time. Following each interview, I asked participants if they could pass on details of the study to further eligible colleagues who might have been interested in participating. At the end of each pilot study interview, I asked participants to reflect on the study materials and my recruitment strategy. All pilot study participants responded positively to these and therefore changes to these were minimal for the full study.

Pilot study participants were not included in the full study sample. Recruitment ended when I had recruited a minimum of 20 interviews and reached data saturation (Saunders *et al*., 2018).

6.2.2 Data collection

The topic guide I used for my interviews contained open-ended questions that reflected the study objectives (see Appendix 8). The guide was developed following guidance from Britten for conducting interviews in medical research, the most appropriate guidelines for this context (Britten, 1995). Participants were asked to describe their use of analytics as part of a strategic health and care decision they had made that had considered health inequalities, factors that had facilitated or hindered their use, and their training needs with respect to their use of analytics (see Appendix 8).

The topic guide was adapted as the study progressed to reflect changes made to the study objectives. Following the pilot study, I also had several reflections:

- As described in section 6.1.3, asking participants to describe their perceived value of analytics led to abstract discussions. Subsequently asking them to give examples of when they had used analytics to inform strategic health and care decision-making, in an attempt to capture their perceived value of analytics, led to examples that were too disparate and technical for me to be able to understand, ask probing questions and analyse collectively.

- I felt that some clarification around participants' meaning of their 'system' would have also helped me collectively analyse their examples.
- Interviews tended to primarily focus on perceived factors (facilitators and barriers) influencing their current use of analytics, rather than on their perceived training needs.
- I found that pilot study participants did not respond particularly well to questions around their own skills and knowledge when commissioning and interpreting analytical outputs to inform their decision-making. Instead, they tended to suggest any issues in these areas would be found outside of their immediate teams, and they would often redirect the focus of the interview if I asked about their own skills and knowledge.

For the full study, I asked senior leaders if they could describe an example of a recent strategic health and care decision they had made around health inequalities (as described in section 6.1.3). This was in an effort to obtain examples that were more similar and therefore easier to collectively analyse. I also asked full study participants to describe the organisations they were working with in their specific examples, in order to gain an understanding of the 'system' they were referring to.

We discussed some of the challenges I faced when asking participants about skills and knowledge and, therefore, training needs when I presented preliminary findings from the pilot study to NCL's Analytics Board. There were no suggestions as to how I could better tap into this in the full study, however I decided to change the way in which I asked about training needs. Instead, I asked full study participants to discuss 'support' they felt they needed to facilitate their use of analytics and I only used the word 'training' when it felt appropriate in the interview.

I adapted the topic guide for the full study to reflect these changes in approach. All participants were given prior information on the topics to be discussed. I asked for consent to audio record interviews and interviews were transcribed by an external transcription agency. Audio recordings were de-identified by the transcription agency and I subsequently checked them for accuracy.

6.2.3 Ethics

My pilot study informed a protocol for the full study. This was reviewed by UCL's Research Ethics Committee and ethical approval for the full study was obtained on 29th August 2019. Participants gave their formal consent to participate in the study. Anonymised quotes were used in data analyses and results to ensure participant confidentiality. Data were stored securely.

6.2.4 Data analysis procedures

6.2.4.1 Thematic analysis

Interview transcripts were analysed using the Framework Method (Smith and Firth, 2011; Gale *et al.*, 2013). This method is a form of thematic analysis, whereby qualitative data are analysed to generate common patterns or themes (Braun and Clarke, 2006, 2014). The framework method was chosen as it is particularly suited to this study:

- It is the most suitable form of thematic analysis for interview data where candidate themes have been identified before the data analysed.
- It enables one to gather perspectives from a specific group of senior leaders.
- It allows the systematic generation of cross-cutting themes across interviews.

Unlike other qualitative methods, the Framework Method is not aligned with a particular epistemological, philosophical, or theoretical approach. Instead, it is a flexible approach that can be adapted for use with various qualitative methods (Braun and Clarke, 2006).

My analysis generally followed steps outlined by Smith and Firth (2011), Gale et al. (2013) and Braun and Clarke (2006) (Braun and Clarke, 2006; Smith and Firth, 2011; Gale *et al.*, 2013). To ensure this process stayed true to the data, these steps were highly iterative:

1. Codes were first generated deductively based on previously identified barriers to high-quality analytics (Shaw *et al.*, 2013; Bardsley, 2016; Bardsley, Steventon and Fothergill, 2019). These deductive codes included data quality, analysts' skills and knowledge, and data sharing.

2. Salient phrases were then coded inductively using constant comparison between transcripts and researcher interpretations. This stage of inductive coding was thorough, inclusive and comprehensive. Each data item had the opportunity to be coded multiple times.
3. When sufficient codes were generated, similar codes were grouped to form categories. A miscellaneous category was created for codes that did not align with any particular category or sub-category.
4. The generated categories formed a working analytical framework. As the analysis progressed, this framework was indexed, and codes and categories refined to represent a robust theme across participants. I used Patton's dual criteria for judging and refining categories (Patton, 1990). As part of refinement, generated codes and themes were compared to the study aims and research questions, and the coding frame adapted to fit with these.
5. The coding frame ultimately formed a Framework Matrix that included participants as rows and emergent categories as columns, including illustrative quotations. See Appendix 9 for a simplified version of the coding frame used in this study.

To minimise biases and improve the reliability of a qualitative study, double coding – the process whereby two researchers independently code transcripts using a draft of the coding frame and subsequently revise the frame – is encouraged (Berends and Johnston, 2005; Braun and Clarke, 2014). I was fortunate enough to be able to work with another PhD student in my department (SB) who had extensive experience of qualitative research and who double coded a proportion (20%) of the transcripts. There are no guidelines on double coding in the literature. The final coding frame (Appendix 9) was checked by SB, my supervisors and one practice partner (KK).

My primary supervisor (JS) and the aforementioned senior qualitative researcher (SC) looked at my work at all stages (Berends and Johnston, 2005). To foster reflexive research practice, I also kept a reflexive journal and I referred to this throughout the analysis and write-up (Malterud, 2001; Smith and Firth, 2011).

6.2.4.2 Typology

Since the 1980s, grouping objects, participants or ideas on the basis of one or more attributes has become increasingly popular in qualitative social research (Bailey,

2011). The constructed groups are known as 'types' and form a typology, which can facilitate comprehension and explanation of complex social realities (Bailey, 2011). In this study, one of my research questions was: are there distinct types of analytics users amongst this sample of senior leaders? To address this question, facilitate my analysis and better investigate these types, I grouped participants according to aspects of their interview responses and constructed a typology. I generally followed Kluge's methodological steps for constructing typologies (Kluge, 2000):

1. *Development of relevant analysing dimensions:* Typologies are generally multidimensional in that groups are formed on the basis of more than one characteristic. I first decided that the relevant analysing dimensions were participants' use of analytics and their experience of important factors that facilitate or hinder analytics use for strategic health and care decision-making. These were revised through reflection and discussion. I subsequently added a third relevant analysing dimension – the participants' response and actions in light of experiencing a given facilitator or barrier.
2. *Grouping of cases by defined dimension(s):* Participants were grouped according to the initially identified analysing dimensions and revised upon addition of the third analysing dimension.
3. *Analysis of meaningful relationships and type construction:* According to Bailey (2011), groups formed should be exhaustive, in that there must be a group for each case (or participant), and mutual, in that there is only one correct group for each participant (Bailey, 2011). Different barriers or facilitators of analytics use that were discussed by participants in their interviews were explored as alternative categorising dimensions and groups reorganised accordingly. Similarities and differences between different participants were considered when reorganising groups. Groups were refined to align with Bailey's methodological guidelines when constructing typologies.
4. *Characterisation of the constructed types:* The constructed types were described and compared in detail. The description of each 'type' was based on the most commonly found characteristics (a 'constructed type') rather than an extreme type in the group (an 'ideal type'). Each type is conceptual in that the resultant groups represent combinations of people and their context rather than empirical cases. In the final typology, participants were grouped according to: their own skills, knowledge and interests in relation to analytics,

their experiences of barriers and facilitators of analytics use for health and care decision-making, and their actions in light of these experiences.

6.2.5 Participant involvement and engagement

As described in section 6.1.1, the idea for this study was borne out of an expressed stakeholder need to understand analytics use for strategic decision-making from the perspective of senior leaders. Whilst I worked collaboratively with NCL throughout the research process, I developed the main study idea, research questions and study materials. Key staff at NCL reviewed all study documents and plans. These were adapted to reflect their priorities, where appropriate. Emergent themes were discussed with a key staff at NCL (KK) during analysis and write-up.

As with the pilot study, I shared findings from the full study with stakeholders at NCL. My aims for seeking feedback on study findings were two-fold: to support validation of my analysis and to aid development of study recommendations and, in turn, knowledge translation. These strategies are forms of 'member validation', a term which denotes techniques to validate qualitative findings by demonstrating correspondence between researcher findings and understandings of the members being analysed (Bloor, 1997). Any feedback was incorporated into the study findings and subsequent recommendations.

6.3 Chapter summary

In this chapter I have presented the introduction and methodology for my third study, which investigated how knowledge generated from the analysis of linked health and council data might inform strategic and equitable health and care decision-making. Using individual, semi-structured interviews, I interviewed a group of senior leader participants recruited from study sites in North London in an attempt to address some of the current gaps in local knowledge in this area. I analysed these interviews using the Framework Method and created a typology. The results of my analyses of these interviews are described in Chapter 7.

Chapter 7 Senior leaders' experiences of using analytics to inform strategic and equitable health and care decision-making: Results and Discussion

In Chapter 6, I presented the introduction and methodology for my third study, which investigated how knowledge generated from the analysis of linked health and council data might inform strategic and equitable health and care decision-making. Given that most senior leaders do not have access to data linked across health, council and other care settings, I focused this study on how knowledge generated from the analysis of linked health and council data *might* be used to inform decision-making in this context.

In this chapter, I present the results of this study. Section 7.1.1 presents participants characteristics. Section 7.1.2 gives an overview of the themes generated from the study interviews. Section 7.1.3 presents the typology created to identify and define different types of analytics users based on participant interviews. Sections 7.1.4, 7.1.5, 7.1.6, 7.1.7 give a comprehensive description of the themes identified in participant interviews and how they related to the types of analytics users captured in my typology.

7.1 Results

7.1.1 Study participants

Interviews were conducted with 20 senior leaders recruited from constituent organisations of one case study site – NCL – prior to its formation of an ICS. NCL included CCGs, councils, hospitals, and other service providers. Participants were in health and/or care commissioning, provider and public health roles.

Table 7-1 gives an overview of participant characteristics. Whilst I attempted to recruit a diverse sample, 60% of participants were male. This generally reflects the UK's public sector senior leader, management and decision-making workforce (Healy, Bradley and Forson, 2011; Kline, 2014).

7.1.2 Overview of themes

In this study, participants primarily described using pseudonymised, "higher-level" data and/or analytics to understand population health needs, monitor and evaluate services, inform planning of new services, or inform the redesigning of existing services for their health and care systems (**Theme 1**). Participant responses captured when and how knowledge generated from the analysis of residents' health or care records is used to inform strategic and equitable decision-making that has implications across health and care (addressing research question 1).

Overall, participants described the process of attempting to obtain data and/or analytics for decision-making as uncoordinated, *"ad-hoc"* or *"random"*. Interviewee responses suggested that, because of this, various factors greatly influenced: if and how analytics were obtained, the type and utility of analytics obtained and the subsequent use of analytics for informing decision-making. These factors were grouped according to whether they were macro factors related to the context and **environment** in which individuals were working (**Theme 2**), micro factors related to the **people** involved in the process and decision (**Theme 3**) or meso factors related to **data** quality (**Theme 4**). These factors generally facilitated and/or hindered analytics use depending on the circumstance, setting or people involved. Participant responses enabled me to identify barriers and facilitators of analytics use for informing strategic and equitable health and care decision-making (addressing research question 2). Participant responses also enabled me to describe how leaders differed in their experiences of, and responses to, the identified barriers and facilitators (addressing research question 3).

An overview of themes 2-4, and their subthemes, can be seen in Table 7-2. Table 7-3 gives a comprehensive summary of themes 2-4 and summarises their general impact on strategic health and care decision-making. If the factor is followed by a (+) then it had a positive impact on decision-making (was a facilitating factor), by a (-) then it had a negative impact (was a challenge or a barrier) and if the factor is followed by a (+/-) then it acted as a facilitator or barrier depending on circumstance or setting. Figure 7-1 depicts the process of obtaining and using analytics for a given strategic health and care decision, as well as the factors affecting each stage of this process, as described by study participants.

Table 7-1: Participant characteristics (N=20)

Characteristic	N (%)
Gender: Male	12 (60)
Geography:	
Inner London Borough	8 (40)
Outer London Borough	4 (20)
Inner and Outer London Boroughs*	8 (40)
Generic Organisation and Role:	
Health – Provider	6 (30)
Health – Commissioner	4 (20)
Local Authority – Social Care Commissioner	4 (20)
Local Authority – Public Health Lead	2 (10)
Health and Local Authority – Health and Social Care Commissioner	4 (20)
*Split role across inner and outer boroughs. Includes NCL leads.	

Table 7-2: Summary of cross-cutting themes 2-4 (and their subthemes) identified from interviews

Factors that facilitate and hinder use of analytics:		
Macro factors related to the working environment (**Theme 2**)	Micro factors related to people (**Theme 3**)	Meso factors related to data quality (**Theme 4**)
System structures Organisational fragmentation Alignment of priorities	**Personal relationships** Between leaders and analysts Between leaders and leaders	**Data availability and accuracy**
Top-down constraints Resources pressures Policy priorities	**Skills and knowledge** Leaders' skills and knowledge Analysts' skills and knowledge	**Data richness and linkage**

Table 7-3: A comprehensive summary of cross-cutting themes 2-4 identified from interviews and their impact on strategic health and care decision-making

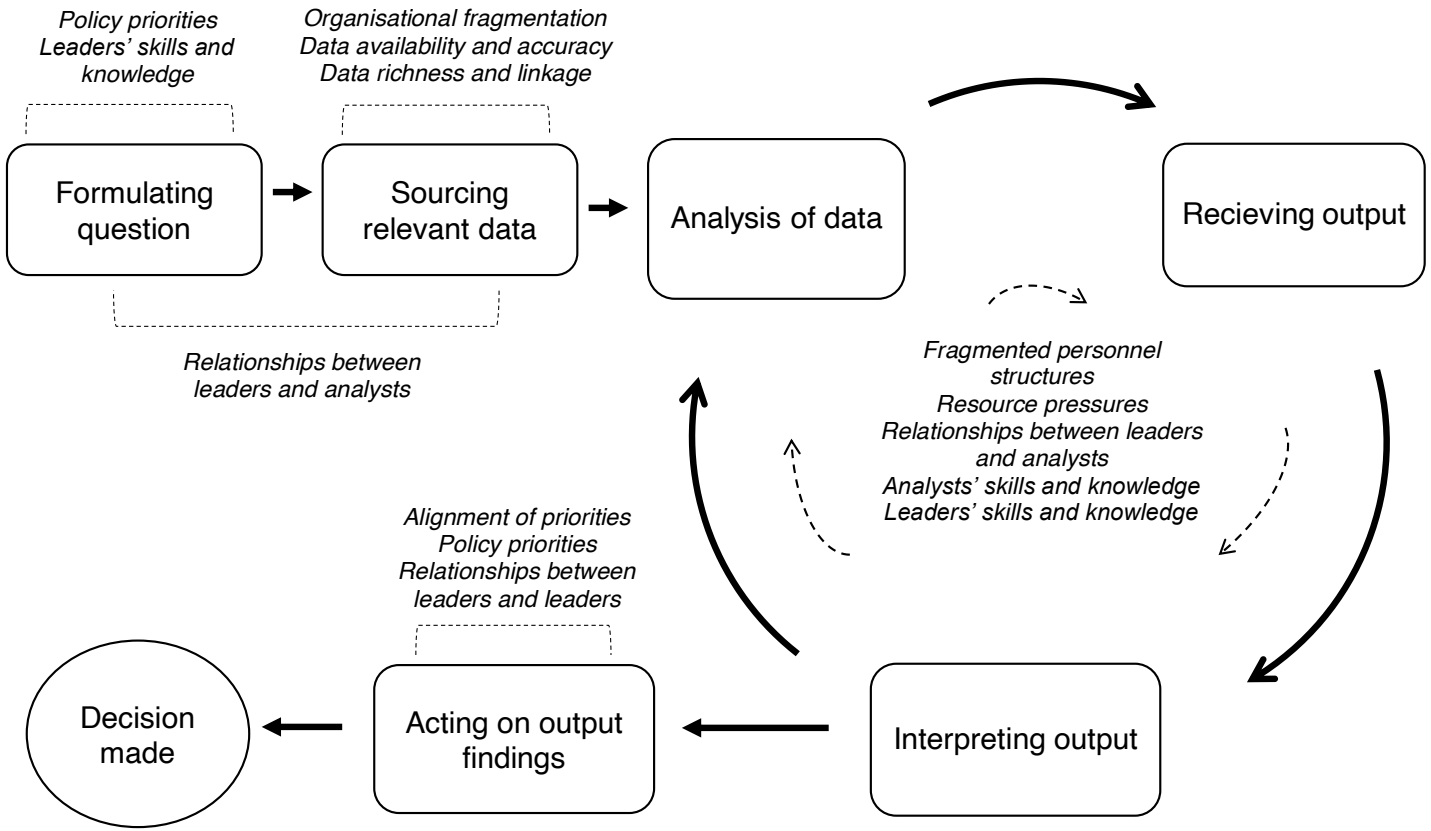
Macro factors related to the working environment (Theme 2)	
System structures	
Organisational fragmentation	Divisions within, and between, organisations created siloed data systems, fragmented personnel structures, and barriers to data sharing (-)
Alignment of priorities	Organisational fragmentation led to aligned, different or competing priorities and timelines across different organisations and teams involved in decision (+/-)
Top-down constraints	
Resource pressures	Resource pressures influenced process when attempting to obtain, interpret, and use analytics; influenced size and capacity of analyst teams; influenced speed at which some decisions needed to be made (-)
Policy priorities	Local and national policy priorities shaped and constrained health and care decisions possible to make given resource pressures e.g., decisions often needed to show short-term financial savings (-)
Micro factors related to the people involved (Theme 3)	
Personal relationships	
Between leaders and analysts	Type of relationships influenced data sharing and process that took place when data/analytics were requested e.g., how questions were asked of data, and how output content was developed and used (+/-)
Between leaders and leaders	Nature of relationships influenced priority and timeline setting, and collaborative decision-making, for health and care system (+/-)
Skills and knowledge	
Leaders' skills and knowledge	Leaders possessed a broad range of skills and knowledge that did not correlate with views on training for leaders; several leaders had a background in data and/or analytics (+)
Analysts' skills and knowledge	Analysts possessed a range of skills and knowledge, however, were predominantly constrained by resource pressures (+/-)
Meso factors related to data quality (Theme 4)	
Data availability and accuracy	Data availability limited and accuracy poor for certain population groups, services, and organisations (-)
Data richness and linkage	Data lacked adequate detail in part because data is collected for different purposes, and in part because of limited access to data linked across health and care services (-)

(-) Factor or process negatively impacted decision-making (was a challenge or a barrier)

(+) Factor or process positively impacted decision-making (was a facilitating factor)

(+/-) In some circumstances or settings, factor acted as a facilitator or negatively impacted decision-making (was a barrier)

Figure 7-1: An overview of the process that takes place when leaders attempt to obtain and use analytics for a given decision, as well as the factors (in italics) affecting each stage of this process, as described by study participants



7.1.3 Types of analytics users

As described in section 6.2.4.2, a typology of analytics users was developed to facilitate understanding of senior leaders' experiences of using analytics for decision-making in this context. The typology acted as an explanatory mechanism for participants' described experiences. Types of analytics users were reached inductively, and five idealistic types were identified: the 'Advanced' analytics user, the 'Hands-On' user, the 'Waiting' user, the 'Challenged' user and the 'Reluctant' analytics user (see Table 7-4 for a brief summary of each type). In this study, three participants were classified as 'Advanced' analytics users, five as 'Hands-On' users and three as 'Waiting' analytics users. Five participants were classified as 'Challenged' users and three participants as 'Reluctant' analytics users.

Each type of analytics user differed in their experiences shared during the interviews. In particular, users differed in their experiences of the factors identified in themes 2-4: in their working contexts and environments (**Theme 2**), in their relationships, their skills and knowledge with analytics and their confidence in the skills of analysts (**Theme 3**) and in factors related to data quality (**Theme 4**). Users also, in turn, differed in their responses to these experiences and, therefore, in how these experiences impacted their ability to engage with, and use, analytics to inform strategic and equitable health and care decision-making.

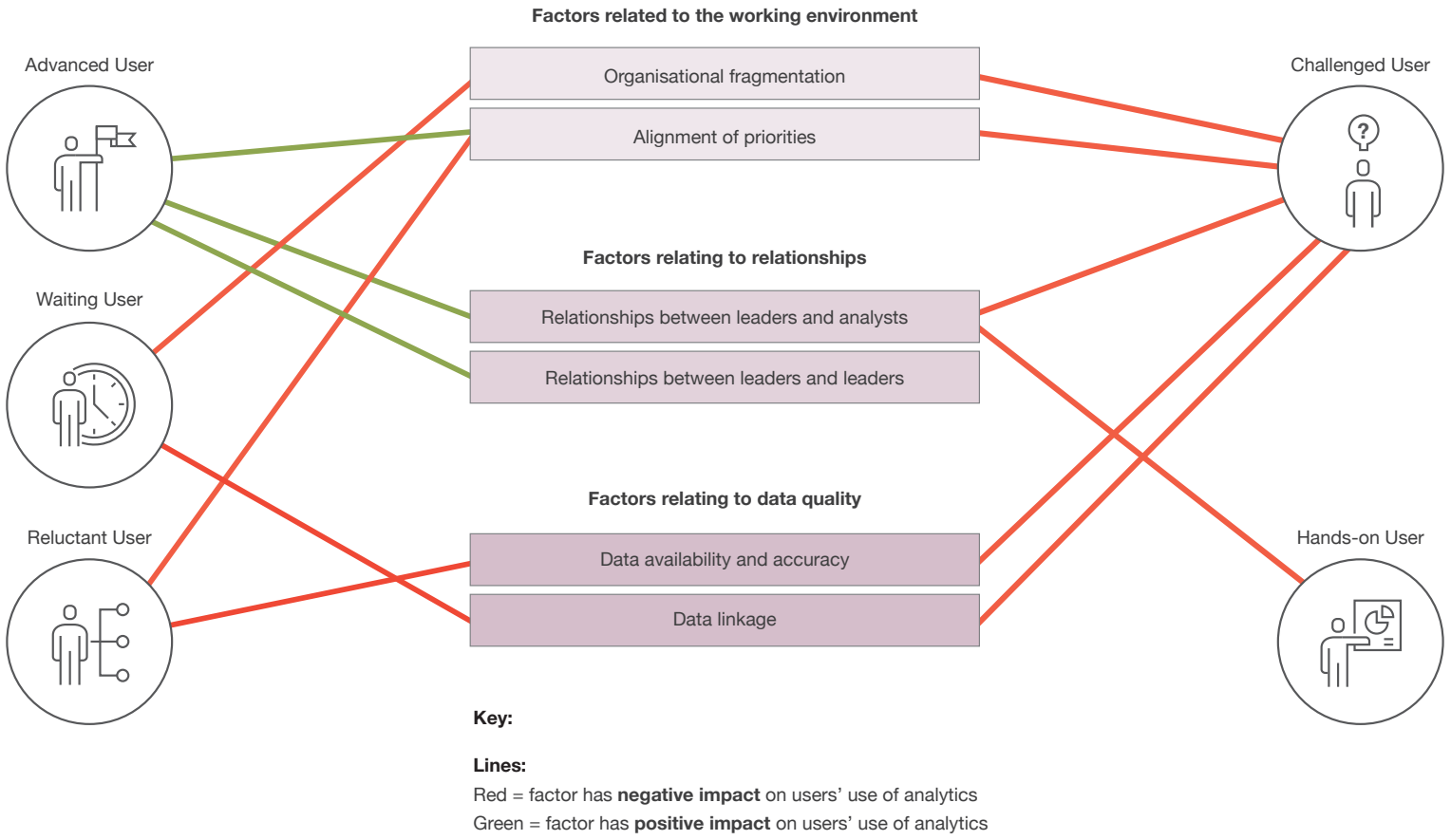
Each type of analytics user will now be discussed in relation to each identified theme and subtheme.

Table 7-4: Brief descriptions of each type of analytics user identified

Type of Analytics User	Description
The 'Advanced' analytics user	A regular user of analytics for strategic health and care decision-making. Does not do their own analysis of data to generate analytical outputs, but, rather, works collaboratively with analysts to obtain analytics support. Analysts often lead the work, and these users rely on analysts' interpretations of analyses to inform their own decision-making. This user experiences few challenges when they try and use analytics, and several facilitating factors.
The 'Hands-On' analytics user	A regular user of analytics for strategic health and care decision-making. Prefers to do their own analyses; rather than work with analysts, they chose to request and obtain raw data and conduct their own analysis to produce the analytical outputs and interpretations of the data. This user has previously experienced several challenges when they had tried to obtain and use analytics, and these past experiences drive their hands-on approach to data and analytics.
The 'Waiting' analytics user	A seldom user of analytics for strategic health and care decision-making. Heavily involved in setting up data systems they believe could one day facilitate such use. This user experiences several challenges when they try and use analytics, which primarily centre around data systems, and few facilitating factors.
The 'Challenged' analytics user	An inconsistent user of analytics for strategic health and care decision-making. Does not collaboratively work with analysts to obtain analytics. Experiences many, various and wide-ranging barriers, and few facilitating factors, when trying to use analytics. These challenges can span several stages of the process that occurs when they seek analytics, from sourcing and obtaining, to attempting to use data and/or analytics for these types of decisions. This user attempts to overcome the barriers they face and use analytics to inform their decision in some capacity.
The 'Reluctant' analytics user	A hesitant and inconsistent user of analytics and does not typically make strategic health and care decisions. Experiences many, various and wide-ranging challenges, and few facilitating factors, when trying to use analytics. This user often halts their use of analytics to inform decision-making when faced with barriers, and does not attempt overcome barriers.

Figure 7-2 captures the main factors that shaped each users' use of analytics according to participant interviews, and whether the factor had a positive or negative impact on decision-making.

Figure 7-2: Main factors impacting each users' use of analytics



7.1.4 Theme 1: Uses of data and/or analytics by senior leaders

'Advanced' and 'Hands-On' analytics users regularly used analytics for strategic health and care decision-making, whilst 'Challenged' and 'Reluctant' users inconsistently used analytics and 'Waiting' users seldom used analytics. In this study, a small minority of participants described using analytics for clinical decision-making, monitoring and evaluating services and for payment for performance. However, the majority of participants described principally using pseudonymised data, and its subsequent analysis, to help inform strategic choices of investment and disinvestment, plan new services or redesign existing services, and understand the impacts of changing or implementing new services models.

Around half of participants described using analytics to better understand health inequalities in their populations and plan more equitable services accordingly. There were no obvious patterns by type of analytics user in the extent to which they sought to reduce inequalities. Several participants described practical and logistical reasons why they did not make decisions around health inequalities, whilst others felt this responsibility was not part of their job role. For example, one Health Provider said that a decision around health inequalities is:

> *"More of a sort of public health policy-level [decision].....We kind of just do our job. So, our patient population [for our organisation] is referred to us, not from basis of inequalities, but on the basis of they've got a problem. And inequality is a feature of our patient population, but it's not a driver of how we negotiate things....If we get [someone] referred [to us], we deal with it, and that's it. So, no I am personally not involved in any of these policy decisions, it's politics, it's high-level stuff. In this organization we don't represent a political view at all....We don't actually focus on deprivation or inequality." (ID011, Health Provider, 'Reluctant' analytics user)*

This interviewee strongly emphasised that if they, and by extension their organisation, made decisions that considers health inequalities, they would be seen as representing a political view or party, which they cannot do. Other interviewees more generally described how making decisions around health equalities was not part of their job role or that answers to health inequalities were *"outside of our control" (ID019, Health Provider, 'Reluctant' analytics user).* For example, one Social Care Commissioner said: *"We don't think of the decisions [we make] being about health inequalities, when*

they're about services… [although] we are required, politically, to consider the sort of inequalities impact." *(ID008, Social Care Commissioner, 'Advanced' analytics user)*

One Health Provider described feelings of discomfort with implementing different services tailored to different parts of their borough to address variations in need: *"as a provider, we would want to provide [a service] to everybody. But it has felt quite uncomfortable to have a service for part of the borough and not have that service for another part of the borough….not offering that service feels wrong." (ID012, Health Provider, 'Challenged' analytics user)*

Participants were, therefore, divided in whether they felt it was legitimate for them to consider health inequalities when making strategic health and care decisions. This is despite one of NCL's aims being to challenge and address health inequalities collectively across NCL. The remaining results described in this chapter, therefore, describe barriers and facilitators of analytics use for strategic health and care decision-making *in general*. In *some* instances, these strategic decisions considered a health inequalities angle.

7.1.5 Theme 2: Factors relating to the working environment

All participants described how their working environments drastically influenced their use of analytics for strategic health and care decision-making. These influences have been grouped as factors relating to system structures (subtheme 1) and top-down constraints being placed on participants (subtheme 2). Whilst these factors affected all types of analytics users, users differed in their responses in light of these constraints and in the impact they described these factors as having on their decision-making.

7.1.5.1 Subtheme 1: System structures

Issues stemming from system structures were described by all participants as key barriers to their use of analytics for strategic health and care decision-making. Analytics users faced different and various challenges stemming from organisational fragmentation, including siloed data systems, barriers to data sharing and difficulties aligning strategic priorities across organisations.

Organisational fragmentation: Health and care organisations are often discrete entities with separate financial structures and data systems, though integrated systems do exist. All participants interviewed in this study belonged to the former type of organisations. Those who recounted facing challenges when they had attempted to use analytics described how divisions between, and within, organisations created siloed data systems. These meant a patient or residents' records could be stored in different data systems if they came into contact with more than one service or organisation, such as a GP service and social care.

The majority of participants described how siloed data systems had made it difficult to access data as they had to actively request data and/or analytics from individuals in other departments or organisations. Divisions in data systems across organisations meant senior leaders did not always know who held certain data, whether the data they held would be relevant to inform decision-making, or how to contact key individuals to request access to data. 'Challenged' analytics users typically described experiencing difficulties accessing data, which limited their ability to understand population needs (particularly for vulnerable populations) and measure the impact of strategic health and care decisions. For example, one Health and Social Care Commissioner, who was classified as a 'Challenged' analytics user, stated that individuals *"jealously guard"* data and this made it difficult for them to access necessary analytics support *(ID025, Health and Social Care Commissioner, 'Challenged' analytics user).* These barriers were aptly described by another 'Challenged' analytics user who had tried to access analytics to better understand and plan for social care accommodation needs:

"[RES]: We need housing data, we need social care data, we need some health data, but it's proving difficult to get those data sources and then when we do eventually get them it's a lot of work to then bind them together trying to paint that coherent story. But there's also issues around [asking] "where does the data sit?". So, I had a meeting with [another internal team] asking for some data [for understanding social care accommodation needs]. They're like, "But this sits here, this doesn't sit with us". It's unclear who owns certain pieces of data and how best to extract it. [INT]: Is that the reason that you had issues accessing it in the first place? [RES]: Definitely. So, housing data in particular, where it sits [is] in a completely different department, a different team. We have no right to access any of that data, so it will take quite a lot of time to get it." (ID023, Social Care Commissioner, 'Challenged' analytics user)

It's important to note that a lack of data sharing did not simply imply a lack of willingness to share and, rather, participants in this study suggested various motivations, drivers and reasons for withholding data. For some, information governance requirements contributed to these barriers to data sharing across departmental and organisational boundaries.

Overcoming barriers to data sharing often involved a time-consuming process, where participants had to identify who to request data from and justify their need. The participant planning for social care accommodation needs in their area continued:

"Having to explain the rationale as to why we need data is always the start of it and can always be a bit of a challenge [in] trying to make them understand why I need access to this data and what it will be used for….But I think the biggest thing is, everyone's busy…. It's never a priority when someone else comes saying, "Do you have this data source? I need it for X", because I think, "I've got twenty other things on my plate." (ID023, Social Care Commissioner, 'Challenged' analytics user)

This time-consuming process requiring continuous justification was, therefore, described as an additional organisational barrier to data sharing, analytics access and analytics use – with other priorities and work often taking precedence. When participants could not access data held on siloed systems, some made decisions without all the *"necessary information" (ID022, Health and Social Care Commissioner, 'Challenged' analytics user),* whilst others relied more on expert opinion (such as the opinion of single practitioners) or halted their analytics use.

Alignment of priorities: Many participants described how organisational fragmentation across their health and care system, at times, led to different or competing

organisational priorities and timelines. In more extreme circumstances, this hindered collective priority setting for 'Challenged' and 'Reluctant' analytics users, despite organisations being encouraged to align priorities locally to facilitate more integrated ways of working. For example, one participant – classified as a 'Challenged' user of analytics – primarily spoke about competing priorities and organisational cultures as delaying collective decision-making. This participant was a Health Provider, and described working with their local authority partners to redesign and streamline intermediate care services to reduce variation in care:

> *"We started talking about this two years ago, having all partners sitting in the room, it was facilitated, you know, we've got some data, we've got everything, we've got a high-level model. And then we've got a new Director come in [on the local authority side], who very much sees their problem (they've got to sort out this little corner of bringing the social care bits together), as a separate project, rather than doing it all at once. Which has delayed the togetherness of the project. It's caused frustrations, you know. And, so, we feel that we could have got further along the road. We were talking 18 months ago. We'd got the model ready and yet we're still sitting here, now, talking about it… …With this particular project, with the culture and with social care, there is a definite, different pace." (ID016, Health Provider, 'Challenged' analytics user)*

'Reluctant' users of analytics typically experienced conflicting financial drivers and role priorities. For example, one 'Reluctant' user, also a Health Provider, expressed little motivation to engage in collective decision-making with other organisations because of conflicting financial priorities between their organisation and their system. This was because, whilst changing their model would realise savings across their system as a whole, it would lead to financial losses for their own individual organisation:

"If we have a patient who we see in the hospital we get paid £70 or something for a follow-up patient. If we work out a new model of care where this patient can be seen in the community or virtually, we would get paid £10 or £15 or something. What on earth would we want to do that for?…. If you're saying let's [in a] wholesale [manner] move half of our patients into the community, let's lose all of that revenue, then suddenly the fixed costs that we have in this building and others become overwhelming. Our sort of model is predicated on getting the type of revenue in from these types of patients…..It is a very huge risk to us." (ID011, Health Provider, 'Reluctant' analytics user)

One further 'Reluctant' user of analytics felt that, whilst their organisation was geographically located in NCL, they served a much broader population, too broad to meaningfully engage in local health and care system decision-making and justify the resources that would require. This influenced their ability to share data and participate in system decision-making:

"[Health and care system decision-making] doesn't quite work for us because of our demographic as an organization, because the vast majority of our patients are not local. That makes that sort of "system" quite complicated…. We will come to the party with North Central London and share, again with the right governance processes, share our data and share our information and be part of that process. But that's only 5% of what we do. We can share the high-level and the concepts and the rest of it, but actually that's [another area's] data or [another area's] data." (ID010, Health Provider, 'Reluctant' analytics user)

Indeed, several participants described how reservations around sharing data often stemmed from conflicting priorities. In addition, some participants stated they were more likely to share their data if they trusted that recipients had priorities aligned to their own and, as such, would use their data as they had specified. This was particularly relevant for data sharing between commissioners and providers, where providers were hesitant to share data in case commissioners used it to justify disinvestment in the health and care services they provided.

Participants in this study also observed they were often competing for analysts' time against other priorities such as the extensive mandatory reporting requirements analysts faced from external public bodies like NHS England:

"The structures that sit across us, there are data requirements placed upon us which are often at short notice and unexpected or slightly different or very similar to one that we did previously. The time and energy and resources that it takes for [analysts] to keep changing that information and updating it and translating it into the latest format is time consuming, it's energy sapping… So, yeah it's not [the analysts'] priority to respond to our [analytics] requests immediately." (ID022, Health and Social Care Commissioner, 'Challenged' analytics user)

Externally mandated requirements that occurred frequently, unexpectedly and at short notice were, therefore, described as creating *"time consuming, energy sapping" (ID022, Health and Social Care Commissioner, 'Challenged' analytics user)* work that needed to be prioritised over requests from leaders for analytics support. This was described as a further organisational barrier to analytics access, which, in turn, hindered senior leaders' analytics use for strategic health and care decision-making.

7.1.5.2 Subtheme 2: Top-down constraints

The majority of participants described top-down constraints being placed on their system as impacting their use of analytics. These included resource (financial and time) pressures, as well as policy priorities.

Resource pressures: Financial and time pressures were described as affecting multiple stages of the decision-making process, from obtaining analytics to interpreting and using analytical outputs. Almost all participants – and all types of analytics users – described shortages in funding as influencing the size and capacity of teams of analysts in their organisation and system. This was described by some as limiting the availability of good quality data and by others as influencing their ability to iteratively and collaboratively work with analysts to reach an analytical output that answered their questions and suited their needs.

In light of resources pressures, 'Advanced' and 'Challenged' analytics users typically attempted to work within the constraints these created and took the best course of action, however this shaped decision-making if limited resources constrained the range of decisions possible or contributed to mistakes being made. For example, one Health Commissioner explained how an error occurred within a report they had received:

"When I got [the analytical output] back it was a massive pivot table that showed me what I wanted. But I didn't go into the sheets behind it and see what [data] was selected in the pivot. Then we did this whole analysis and about five to six months later we realised that one thing was selected in the pivot that shouldn't have been selected in the pivot. That then [influenced what was possible to achieve with the given decision]. And like that should have been picked up somewhere, like by someone….There's multiple points of failure, but because it's quite complex and like things are moving fast there's people who aren't really checking details of things, which maybe they should." (ID018, Health Commissioner, 'Advanced' analytics user)

This participant then described how incorrect information was used to contribute to a larger, more detailed analysis, which fed into their decision. However, when more scrutiny was applied further down the line, this mistake was identified, and this greatly influenced their decision-making process.

'Hands-On' analytics users described requesting raw data and conducting their own analysis when faced with resource pressures. This was because they typically had experience working in or managing teams of analysts and were perhaps more aware of how analysts' limited capacity influenced their work.

A handful of participants described how their need to quickly make decisions had not previously aligned with the time it took to source relevant analytics. In these circumstances, 'Challenged' analytics users sometimes described seeking analytics support elsewhere, with one Health Provider describing how, for a given decision, an analyst had "*took weeks"* when they *"needed [the analytical output] in days, or minutes, ideally". ID012 (Health Provider, 'Challenged' analytics user).* This participant then chose to hire their own analyst. Other 'Challenged' analytics users instead made decisions without all the information they deemed necessary. This can be seen in the following exchange with a Social Care Commissioner where they described a decision they had made around relocating and centralising dementia day services:

"[INT]: Were there questions you would've wanted to ask of the data that could have influenced the way in which the services were designed? [RES]: …I guess there would be something around if a lot of other people at the [care] centre had other health needs, could we have catered for them more effectively? …But we just didn't have that data readily available to make sure we could do it. I think we probably could've gotten it, but within the timescales of the project it would've been a challenge to get and honestly there were other priorities at that time and because the data would've been a challenge to get it wasn't seen as the main priority." (ID023, Social Care Commissioner, 'Challenged' analytics user)

In contrast, 'Reluctant' users were more likely to describe having made decisions based on *"gut feeling"* or *"anecdotal" information (ID019, Health Provider, 'Reluctant' analytics user),* rather than attempting to obtain analytical support via a process that could be lengthy and resource intensive.

Policy priorities: In this study, local and national policy priorities were described by a minority of participants – 'Advanced' and 'Challenged' analytics users – as constraining the health and care decisions they were able to make given resource pressures. For example, one of the main aims of the integration agenda is to reduce duplication of work across the system and, in turn, realise financial savings. Participants commented on how this limited their ability to do "*proper health and care planning"* that could show longer term savings as *"it's much harder for commissioners to continue to make the case to commission something that isn't showing an immediate return on investment" (ID020, Health Commissioner, 'Advanced' analytics user).* Decisions they made, therefore, needed to demonstrate short term impacts on their organisations' financial activity. This view was not universally shared across participants.

7.1.6 Theme 3: Factors relating to the individuals involved

Participants described how factors related to the individuals involved in the process of obtaining and using analytics for strategic health and care decision-making can greatly influence the process and resultant decision. Personal relationships (subtheme 1) and individual skills and knowledge around analytics (subtheme 2) were discussed repeatedly. Different types of analytics users identified in my typology differed considerably in their relationships with analysts and other leaders, which appeared to influence their use of analytics in subsequent decision-making. However,

it was less possible to distinguish different types of analytics users based on levels of analytical skills and knowledge.

7.1.6.1 Subtheme 1: Personal relationships

Most participants described how building trust and relationships between leaders and analysts, and between individual leaders of different organisations, can help circumvent barriers stemming from organisational fragmentation and facilitate cross-organisational alignment of priorities. While these relationships were described as important, they were also described as challenging to build and sustain.

Relationships between leaders and analysts: In this study, participants suggested that the uncoordinated, *"random"* way in which leaders obtained analytics meant relationships between leaders and analysts greatly influenced analytics access and use. Pre-existing relationships with analysts were described as influencing the type of data and/or analytics sourced, and the utility of outputs received. As one participant explained:

> *"We weren't sure who the right team [was] to ask for certain bits of data. So, some people were asking public health for stuff, some people were asking a [local] analyst for data and information. Some people were asking people from [another organisation] for data and information. It was like a scatter gun, like whoever you know, you ask….I feel like if you ask one team a question, they might say like, "Oh you can't do that, we don't have that information". And then you would just be like, "Oh we don't have that information". But if you went to that [other] team and you said, "We need to do this, do you have this information?", they'd be like, "yeah, yeah, here you go". Do you know? But if you didn't know to even ask those people, like you might just report back and be like "we can't gather this information, it doesn't exist". I think personal relationships with different analysts in different teams definitely has an impact. I think if you can get something from someone quickly and you trust the information, and you build that [relationship], you would probably easily go back to that person again." (ID018, Health Commissioner, 'Advanced' analytics user)*

'Advanced' users of analytics described having a *"good dialogue"* with analysts, *"trusted"* analysts, felt they are *"part of the team"*, even if not physically co-located. 'Advanced' analytics users described working collaboratively to iteratively obtain and use analytics for decision-making. When the process was described in this way, these analytics users described few, if any, issues with obtaining data and receiving, reviewing, and using outputs from analysts:

"[RES]: We [leader and analyst] talked about the scope, they took the lead, so [the analyst] was fab, as [they are]. We kind of described the scope of the strategy, and what we'd intended it to do, and then [the analyst] went off and led [the work]. We had a couple of meetings to check in every so often and she went off and led a team. [INT]: So how did the process work in terms of developing the questions asked of the data? [RES]: I don't know actually. [The analyst] and I have worked together on and off for years, so maybe it was that. Also, I just inherently trust [the analyst] to know what [they're] doing." (ID020, Health Commissioner, 'Advanced' analytics user)

The benefits of having good working relationships between leaders and analysts appeared to be crucial, so much so that leaders *"attach themselves to good analysts"*, even if external to their own organisation:

"There's a better analyst in the XXX. [And so,] I would nick [them] sometimes. I would trust [their] judgement around [how the analysis should be conducted], because I know after a few words of briefing from me, [they] would pick out the themes and I wouldn't have to look over [their] shoulder every five minutes." (ID021, Health and Social Care Commissioner, 'Hands-On' analytics user)

Conversely, participants who faced barriers when trying to make analytics-informed decisions typically stated that, while they wanted collaborative working relationships with analysts, these were not currently available. These participants were all classified as 'Challenged', 'Hands-On' or 'Reluctant' users of analytics. When participants described not knowing who to contact, they seemed more likely to request raw data and do their own analysis ('Hands-On' analytics users) or described issues where data-based decisions were not made or were delayed ('Challenged' and 'Reluctant' analytics users). When 'Challenged' and 'Reluctant' analytics users knew who to go to for data, they often described developing questions without analysts' input. This appeared to lead to *"insufficient"* outputs which: did *"not fit"* participants' needs (i.e., did not answer the questions they had wanted answering), lacked extra detail around how to interpret and use the output, or recommended unfeasible actions.

Organisational fragmentation and resource pressures were also described by participants as creating physical disconnect between leaders and analysts, meaning that good, cross-organisational relationships were even more salient. For example, several participants believed they had a better understanding of how services operated and the surrounding decision-making contexts than analysts, because analysts where not co-located in decision-making teams and they had struggled to

form relationships with analysts. For instance, one Health Provider – classified as a 'Challenged' analytics user – faced difficulties working with external analysts as outputs did not contain details and information necessary for their decision-making around diabetes care provision. They felt this was because analysts were not *"part of the team"* and, therefore *"didn't know what [the leaders were] talking about and leading on"* with respect to a decision, and did not have access to the right, necessary levels of data *(ID012, Health Provider, 'Challenged' analytics user).* This participant eventually hired an internal analyst. They went on to say that working with these external analysts had been a costly process and, instead, hiring their own analyst was both more financially viable and better suited their analytic needs. This participants' experience was fairly common.

'Hands-On' analytics users used barriers related to relationships between leaders analysts to justify their choice to request raw data and conduct their own analyses. For example, one Health and Social Care Commissioner, classified as a 'Hands-On' analytics user, described how they chose to revise an analytical output to include additional financial information because, without this, they could not make a successful business case to convince relevant stakeholders to implement the outputs' recommendations. They stated that an output they had received from an analyst only contained information on how the recommendations would benefit *"social value"* in the long-term and did not include information on possible short- and longer-term financial savings. They conducted additional analyses with the raw data and "*put a business case together to say this is going to really deliver big bucks in three years' time and, by the way, it'll be a small cost in the first year but actually, you know, probably a gross saving as well." (ID021, Health and Social Care Commissioner, 'Hands-On' analytics user).* They went on to describe how, often, outputs they received did not contain this information, but that this additional information was crucial, as they had to *"go through the finance person in the CCG and then the council and explain how much money [implementing the outputs' recommendations was] going to save from the respective budgets".*

Whilst the majority of 'Challenged' and 'Hands-on' analytics users described this physical disconnect between leaders and analysts as inhibiting the formation of good and collaborative working relationships, having co-located analysts did not always appear to facilitate use of analytics. One participant, also a 'Hands-On' user of

analytics, described never going to their co-located analyst because they were *"unapproachable"*, and they did not *"know what they do" (ID026, Health and Social Care Commissioner, 'Hands-On' analytics user)*. When this participant conducted their own analysis, they, however, faced difficulties when trying to conduct more advanced analyses needed to support their decision-making. They ultimately chose to seek analytical support externally, and this delayed their decision-making.

Relationships between leaders and leaders: In this study, some participants described how, when making strategic health and care decisions, building trust and relationships between individual leaders from different organisations was key. 'Advanced' users of analytics, who described few issues using analytics in such decisions, had typically known other leaders they collaborated with for a period of time, which had allowed a degree of trust to develop:

> *"We've worked with the team [of leaders] for a couple of years now as well, so there's some trust there as well between us, like that we wouldn't try to implement something that wasn't robust." (ID018, Health Commissioner, 'Advanced' analytics user)*

For 'Advanced' analytics users, these relationships were described as facilitating the alignment of strategic cross-organisational priorities and timelines, fostering a *"sense of shared ownership"* and facilitating collaborative decision-making processes.

Conversely, on top of struggling to obtain and use analytics, 'Challenged' users of analytics described being poorly networked with other leaders and faced difficulties forming relationships with other leaders. This appeared to impact the progress and implementation of decisions and their use of analytics as part of these, as explained by one 'Challenged' analytics user:

"For a project to be successful you've got to have the buy in of everybody who's leading that project. And unfortunately, in [our area], that is a challenge, especially when you've got a project such as this which is quite a large project, which doesn't happen overnight. You've got different members of staff coming and going, and that handoff isn't necessarily always there. You might have an agreement at the offset of the project and of course that would be a high-level of agreement in the initial stages. But when you get down to the nitty gritty and then you're dealing with a new set of people maybe sitting around a room….. Relationships are absolutely imperative to getting that done, which, I think, in [our area] is a challenge, because of the throughput of staff. For those of us around it's absolutely key to getting those relationships, learning how to develop those relationships very quickly, and getting trust. Trust is really important." (ID016, Health Provider, 'Challenged' analytics user)

High turnover of senior leaders was, therefore, described as an additional barrier to developing and sustaining leader relationships, stalling project delivery and analytics use as part of strategic health and care decision-making.

7.1.6.2 Subtheme 2: Skills and knowledge

Interview participants suggested that this group of leaders possessed a range of skills and knowledge in relation to analytics. Several participants suggested training for leaders could facilitate use of analytics for decision-making, although this view was not universally shared. A minority of participants suggested analysts had gaps in their skills and knowledge in relation to analytics, which influenced their ability to obtain sufficient analytical outputs, and interpret and use these. Again, this view was not universally shared.

<u>Leaders' skills and knowledge:</u> Several participants felt that issues around the term 'data' and 'analytics' hindered the use of analytics for decision-making. The terms were often used interchangeably and described to have multiple meanings. Participants broadly suggested that the term 'data' referred to raw numerical information and 'analytics' to analysis of that information for interpretation. Nevertheless, I did not consistently ask about meaning derived from these terms, and it was unclear from participants' responses how any ambiguity around these terms influenced the process of obtaining and using data and/or analytics for decision-making.

In this study, some participants discussed complicated concepts and statistical techniques when discussing analytics, with 'Hands-on' analytics users classified, in part, due to their own skills and knowledge obtained in previous analytical roles enabling them to conduct their own data analyses. This was described by most 'Hands-On' analytics users as being the reason why they faced few issues when they had attempted to use their own analytical outputs to inform subsequent decision-making:

> *"Having had a bit of a data analytical background in a previous life, [more complicated analytical work] is the kind of stuff that I always enjoyed…. So, yeah, it was basically myself doing it." (ID013, Health Commissioner, 'Hands-On' analytics user)*

However, some participants – including one participant classified as a 'Hands-On' analytics user – did appear to find it difficult to recall basic statistical terms (such as P values, confidence intervals and statistical significance). How this impacted their use of analytics was not clear. This can be seen in the following exchange:

> *"I've got a science degree, a long time ago now, so I guess I did some stuff around err you know erm what do you call it, I've forgotten the word, erm, significance and all those things.. data. So, I guess I think we are very poor generally, I think, at knowing what is actually, you know, what's a significant change in data and what is just a change." (ID024, Social Care Commissioner, 'Hands-On' analytics user)*

As such, different types of analytics users did not clearly differ in their own skills and knowledge relating to analytics. Participants' views on the value of training around analytics for senior leaders were also similarly split and did not correlate with skills and knowledge (those with less skills desiring more training, for example). Those not in favour of training were typically 'Reluctant' users of analytics, and they often described their limited time and capacity as a key reason for this view. Others perceived analytics training as unhelpful. For example, one 'Reluctant' user, a Health Provider, said that any training aiming to improve statistical knowledge amongst senior leaders would be futile because they had "*already decided [they were] not trusting the data, because the data is not clean enough. So, [they] wouldn't bother training people up, so they get funny about whether [they] should've used a chi-squared test." (ID019, Health Provider, 'Reluctant' analytics user).*

Pro-training participants were typically 'Challenged' users of analytics and generally suggested that training should include elements that described where different data are held across data systems and its limitations, who to contact to access data, how to interrogate data and interpret analytical outputs, what key questions should be asked when receiving analytical outputs and how to conduct their own analyses. Some of these participants also described a lack of training around analytics as delaying some of their previous decisions as they needed time to understand, interpret and use an output.

Analysts' skills and knowledge: A minority of participants described analysts' skills and knowledge as having influenced their ability to obtain and use analytics, however these participants also often reflected on how top-down resource pressures and numerous mandatory reporting tasks impacted the capacity and skill levels of analysts. Amongst those who did discuss analysts' skills and knowledge with relation to analytics, views were again divided.

'Advanced' users of analytics viewed analysts as the *"experts"* who could provide *"rigour"* to the work and experienced few, if any, issues with analysts' skills *(ID020, Health Commissioner, 'Advanced' analytics user)*. In contrast, participants classified as 'Hands-On', 'Challenged' and 'Reluctant' analytics users described more negative experiences. For example, 'Hands-On' users of analytics typically chose to request raw data and do their own analysis because they felt they possessed better expertise than analysts. This can be seen in this exchange with a Health and Social Care Commissioner:

> *"I think the skills of the [analysts] vary quite a bit and I'm not convinced, being an ex-analyst myself and having run an analytics team, that they get the help that they need. So I find myself knowing what I need quicker than they do and being able to dictate terms a little bit better than some of my commissioning colleagues.....I think the [analysts] aren't great, with all due respect to them, at the analytics part, so I think they're good at the data preparation part. And that's why I was saying I kind of commissioned them to do the informatics [data] piece, but not the analytics piece. Because I don't think they would have the insight I would have as a commissioner into the data." (ID021, Health and Social Care Commissioner, 'Hands-On' analytics user)*

This view regarding analysts' skills and knowledge is echoed in the following exchange with a 'Reluctant' analytics user:

"I've got an informatics department. I've got a tentative commercial relationship with a bunch of shiny suited sharks [able to provide informatics support]. That's my choices. I have one or both or none of them. What I don't really have is a real good data analysis service, you know.... Health data scientists. That's what we need." (ID019, Health Provider, 'Reluctant' analytics user).

In contrast to 'Hands-On' analytics users, 'Challenged' and 'Reluctant' users often described situations where they had sought and obtained *"insufficient"* analytical outputs. The 'Reluctant' analytics user quoted above went on to state that this happened because analysts *"just put [data] in expediently just to get the job done".* In these circumstances, 'Challenged' analytics users often described circumstances where they had tried to take the output and make decisions based on its content, filling in knowledge gaps with *"gut feeling"* and *"anecdotal information" (ID022, Health and Social Care Commissioner, 'Challenged' analytics user)*. Instead, 'Reluctant' analytics users often described halting their decision-making when faced with barriers around analysts' skills, knowledge, and capacity.

All types of analytics users supported training around analytics for analysts. However, all participants also believed that increased resources could more readily increase analysts' capacity and their ability to support senior leaders when they request analytics to inform decision-making.

7.1.7 Theme 4: Factors relating to data quality

The fourth theme in this study centres on data quality, which was described as hindering senior leaders' use of analytics to inform strategic health and care decision-making. The term 'data quality' was used as an umbrella term to signify a range of issues, namely: data availability and accuracy (subtheme 1) and richness and linkage (subtheme 2). 'Challenged' and 'Reluctant' users of analytics typically described considerable concerns around, and issues stemming from, poor quality data, whilst 'Advanced' analytics users described few data quality concerns or issues.

7.1.7.1 Subtheme 1: Data availability and accuracy

Several participants, particularly 'Challenged' analytics users, described circumstances where data they required for a decision did not exist because certain groups had little or sporadic contact with services or recording of certain information was not mandatory. For example, when discussing autism service provision, one

Health and Social Care Commissioner said that they *"simply don't know how many children have autism because there are whole cohorts that are just not recorded. So, if you're at a private school, it's not recorded. If you don't have an educational health and care plan, it's not recorded."* This participant was classified as a 'Challenged' user of analytics partly because they went on to make comments such as: *"you felt like how on earth could we possibly do accurate service planning because we will never know the extent [of the number of people with autism in their borough]" (ID022, Health and Social Care Commissioner, 'Challenged' analytics user).* They subsequently described how they retrospectively collected data in the absence of this necessary information, which was a resource-intensive and *"frustrating"* task, and how they relied more on *"professional judgement"* and out-of-date information. For example, they described how they "*end[ed] up having to just look at studies that were carried out quite a long time ago",* but they were concerned about relying on these too heavily because "*the whole autism diagnosis, that whole field has changed so much, that actually are [they] basing [service planning] on completely inaccurate assumptions or diagnostic techniques that were undertaken ten years ago?" (ID022, Health and Social Care Commissioner, 'Challenged' analytics user).* This idea that the absence of necessary data made it difficult for senior leaders to understand needs and plans services was mirrored by other participants, particularly 'Challenged' analytics users. Again, these participants sometimes went on to describe how they relied on out-of-date information or other sources of information in an attempt to fill these gaps in understanding.

Several participants described how their perceived inaccuracy of data also hindered their ability to use analytics to design services, as well as their ability to use analytics to measure the impact of service changes. This issue of perceived inaccuracy of data spanned different types of analytics users. Some participants described how data accuracy issues led to considerable time and resources being spent trying to determine the *"source of the truth"* i.e., trying to determine the *"correct"* data *(ID021, Health and Social Care Commissioner, 'Hands-On' analytics user).* For example, when discussing implementing intermediate care services in their borough, this Health and Social Care Commissioner said the following:

"We were spending all of our time arguing about data between [analysts] and the provider teams. Why the hell are we doing that? We could just agree to disagree or whatever, but surely you want your analysts to do higher-order analytics stuff, not to piss round (pardon my French) arguing about a million records, which I was at [my last organisation], about whether they're right or not, that's obviously annoying….it was like do you really want to be doing that or do you want to be doing a simulation model to understand what the level of resource is we need. (ID021, Health and Social Care Commissioner, 'Hands-On' analytics user)

In the absence of accurate data several participants – principally 'Challenged' analytics users – again described using *"anecdotal information"* and *"gut feeling"* to make decisions and/or *"heavily caveating" decisions and "evaluating the hell"* out of the resultant services *(ID022, Health and Social Care Commissioner, Struggling User).* In some cases, participants halted decision-making due to perceived data inaccuracies, and again relied more on expert opinion. 'Advanced' analytics users rarely communicated data availability and accuracy as barriers to analytics use.

7.1.7.2 Subtheme 2: Data richness and linkage

A minority of participants described data quality as an issue because data lacked adequate detail, in part because data they had available was collected for one purpose (for example, to determine payment for performance) and used for another (for example, to plan service delivery). However, the majority of participants – crossing all types of analytics users – described the absence of data linked across health and care services and organisations as limiting their ability to fully understand population health needs and plan or evaluate services accordingly. Most participants believed linked data would help overcome barriers stemming from siloed data systems or poor working relationships.

'Challenged' analytics users typically described desiring access to linked data that allowed them to obtain simple, rudimentary analyses and gave them basic insights into the impact of a strategic health and care decision. For example, one 'Challenged' analytics user described being unable to understand the impact of providing mobility equipment for residents, because of data system limitations as well as a lack of linked data across systems:

"Equipment is ordered by health and social care professionals and working out the impact of that is, you know, impossible. So, even Mosaic [the social care data system] doesn't have data about someone getting equipment, there's no way you can search that. Might be in the notes but there's no searchable way. So, you're tracking people who might have equipment prescribed by health or by social care, you can't see the wider impact of that and whether they go to hospital or not." (ID024, Social Care Commissioner, 'Challenged' analytics user)

In contrast, 'Hands-On' users of analytics typically wanted linked data to obtain or conduct more advanced analyses to better understand the wider impacts of a given decision. For example, when describing the implementation of the National Diabetes Programme locally, one Public Health Lead – classified as a 'Hands-On' analytics user – described how they were able to obtain basic levels of data to understand engagement and dropout rates, but desired more advanced linked data to understand the wider impacts of the programme. This can be seen in the following exchange:

"What would be amazing obviously in terms of evaluation [of the National Diabetes Programme locally] is that you could link those people back into their health record and then get some kind of long-term outcome for people with diabetes who'd been in the prevention programme. But what we would currently look at [with unlinked programme data] is just engagement with the programme and dropout rates." (ID017, Public Health Lead, 'Hands-On' analytics user)

One commissioner – classified as a 'Hands-On' user of analytics – described linking primary care, acute and publicly available data to inform their decision-making:

"We looked at primary care data, [and] prevalence [of various health conditions]. Then we looked at some acute data, and we managed to link the acute and primary care data through the pseudonymised NHS number. But because we had all this [geographical] mapping in our data, we said, actually, well we can link to [area-level deprivation data]….We then looked at it, and what we ended up with was six very different projects, so not this blanket one size fits all." (ID013, Health Commissioner, Hands-On User)

They described how this linked data enabled them to see the *"fuller picture"* of service use for residents who accessed care across organisational boundaries. They felt more able to holistically understand health needs and more efficiently make strategic and more equitable health and care decisions. However, they felt unable to make decisions that considered residents' individual social circumstances or social care

use, as local authority records (containing such information) did not contain NHS numbers. NHS numbers were seen as necessary enablers of data linkage:

> *"[RES]: [With] our local authority data, unfortunately, they didn't use NHS number at all. So normally where you might get say a 65% to 70% match, or even a 50%/60%, we had nothing. [INT]: Do you think the local authority data could have added anything to this planning model? [RES]: So, if we were looking at something, let's just say we had an area where there was high COPD attendances and non-elective admissions, and then we looked at the IMD [Index of Multiple Deprivation] and we saw that there were housing barriers…. The local authority data could have added value, you know, well how many council houses are there, what are the estates, what is the state of them?" (ID013, Health Commissioner, 'Hands-On' analytics user)*

This participant was fairly exceptional as they conducted their own data linkage, whilst other participants did not have access to data linked across services. However, this positive view of data linkage was not universally shared. One participant expressed concerns that linking data across health and care would be futile given data inaccuracy across the system and limited comparability of data across the system.

'Waiting' analytics users were characterised by their work setting up new, more advanced data systems that aimed to enable leaders' access to linked data and overcome issues of siloed data across health and care. For example, another Public Health Lead – classified as a 'Waiting' analytics user – also described another instance of trying to capture the number of people with autism in their borough and how this was not possible, *"because some of them [the lists and data on people with autism] are held on local authority systems, some of them are held within primary care and some of them are held within the community learning disability service." (ID014, Public Health Lead, 'Waiting' analytics user).* Their work aimed to *"identify which data sources would have information on people with autism and link them up"* as, without data linkage, the data sources that contained this information were numerous and stored across different, isolated data systems. This inhibited their use of analytics for decision-making around services for residents with autism as they simply did not know how many residents had autism. By linking the data, they believed they would "*be able to describe, more accurately, [the] population with autism locally, and use that to then analyse and advocate for service provision." (ID014, Public Health Lead, 'Waiting' analytics user).* However, 'Waiting' analytics users also described spending considerable time and resources encouraging and supporting individuals and

organisations to share data in order to create these new, more advanced data systems.

7.2 Discussion

7.2.1 Summary of study findings

This study used individual qualitative interviews to explore how knowledge generated from the analysis of linked health and council data might influence strategic and equitable decision-making (addressing Aim 2 of this thesis). Findings suggested that the interplay between factors related to three areas – the environments in which leaders worked, the people involved in the process and decision, and data quality – produced different outcomes in relation to analytics use for decision-making.

Many of these factors were identified as barriers to analytics use, some of which were expected in the context of using analytics to address health inequalities, and some of which were unexpected. Familiar and expected barriers included a lack of trust in data quality, challenges accessing necessary data and the importance of good working relationships as a way to circumvent barriers stemming from organisational fragmentation. However, unexpectedly, around half of participants did not believe it legitimate to consider health inequalities as part of their roles. In addition, several participants actively resisted data sharing due to conflicting organisational priorities and participants challenged NCL's position that senior leaders' skills and knowledge was the main determinant of productive analytics use. Overall, my study findings challenge the widely held assumption that knowledge generated from increased data sharing and linkage will improve decision-making, care, and the equity of services *without* strategies to address further, wider barriers to analytics use.

The data collected during this study also allowed me to characterise five different types of analytics users that differed in their experiences of, and responses to, the identified barriers and facilitators of analytics use. The five types were leaders as: 'Advanced', 'Hands-On', 'Waiting', 'Challenged' and 'Reluctant' users. This suggests that different approaches might be needed to optimise analytics use for informing strategic and equitable decision-making, depending on the type of analytics user.

7.2.2 Comparisons to existing literature

7.2.2.1 Using analytics to inform strategic and equitable decision-making

In this study, around half of participants described using analytics to better understand health inequalities in their populations and plan more equitable services accordingly. The remaining half of participants did not consider this to be a legitimate part of their roles. This is surprising as one of the aims of NCL is to challenge and address health inequalities, which is in line with one aim of health and care integration in general (Department of Health and Social Care, 2021b). Leaders of all constituent organisations of NCL are expected to collectively meet this aim. Participants gave various reasons for not considering health inequalities in their decision-making: health inequalities were seen as a political and, as such, untouchable issue; addressing inequalities was viewed as outside of the control of individuals and organisations; and/or addressing inequalities was simply not viewed as a part of leaders' job roles. These findings suggest that linking health and council data and making the outputs of analysis of such data more readily available to senior leaders will not *uniformly* influence decision-making nor the equity of health and care services, even if wider barriers to analytics access and use are addressed.

7.2.2.2 Factors relating to working environments

At the time of interviews, constituent organisations of NCL were structurally independent. This hindered senior leaders' access to and use of analytics by creating siloed data systems and barriers to data sharing. As a result, most participants could not follow patient or resident journeys across services, plan services effectively using data that might be linked across this journey, nor gain a broader understanding of the context in which residents were living.

Previous literature has shown that siloed data systems consistently create barriers to UK health and care integration (Edwards, 2019; Erens *et al.*, 2019; Billings *et al.*, 2020). My study found that siloed systems meant senior leaders had to go through a time-consuming and resource intensive process to request access to data, with 'Challenged' users of analytics often not knowing who to contact. This created additional barriers as other work and priorities often took precedence. My findings agree with previous literature, which has found that different datasets collected and stored by separate health and care organisations are often not comparable in quality

or coverage and not up-to-date, making it challenging to use analytics for decision-making (Shaw *et al.*, 2013). In addition, previous literature has shown that, without information-sharing and data linkage across organisations, senior leaders cannot fully capture a resident's context, their related service use and its associations with SDoH (Boyd *et al.*, 2005; Bower *et al.*, 2011; Shaw *et al.*, 2013; Sinnott *et al.*, 2013; Kasteridis *et al.*, 2014). Whilst Erens et al. (2019) suggest these data sharing barriers are reducing over time, my findings suggest that organisational fragmentation continues to create barriers to data sharing and that increased data sharing and linkage across organisations could *partly* improve analytics use for strategic and equitable decision-making (Erens *et al.*, 2019). This is particularly true for 'Challenged' analytics users. However, more is also needed to address the wider barriers that all analytics users face in order to realise the government's aims for data to transform health and care and address inequalities (Department of Health and Social Care, 2021b, 2021c). These include barriers relating to the relational aspects of analytics access, and data quality.

7.2.2.3 Factors relating to relationships between leaders and analysts

This study found that 'Advanced' analytics users were those who described having good working relationships with trusted analysts that allowed them to overcome barriers stemming from organisational fragmentation and obtain suitable analytical support. Previous work has described how organisational fragmentation isolates analysts from other analysts, limiting their ability to share experiences and learn new analytical methods (Bardsley, 2016). However, there has been no research examining how organisational fragmentation influences relationships between leaders and analysts and how this, in turn, influences analytics use for informing strategic and equitable decision-making. I found that 'Challenged', 'Reluctant' and 'Hands-On' analytics users typically did not have collaborative working relationships with analysts and that this was a significant barrier influencing their access to and use of analytics.

These findings suggest that, although necessary, upcoming data-related policy changes on their own that aim to facilitate data sharing will be insufficient to realise the White Paper's aspiration for data to transform care (Department of Health and Social Care, 2021b, 2021a). This is because, when these reforms come into force in 2022, leaders may continue to struggle accessing and using data and/or analytics if they do not know where different data are held, who to contact to request certain data

or believe analysts do not understand decision-making contexts. Efforts to develop and sustain relationships between leaders and analysts across organisations are therefore crucial and could better support leaders to use analytics for informing strategic and equitable decision-making. These efforts could include analyst secondments that provide analysts' greater proximity to decision-makers and foster shared understanding of priorities, and decision-making contexts. Such programmes could afford senior leaders – who have been previously 'Challenged', 'Reluctant' and 'Hands-On' users of analytics – opportunities to foster relationships with analysts similar to those found with 'Advanced' analytics users.

7.2.2.4 Factors relating to relationships between leaders and leaders

There has been considerable research exploring if and how relationships between leaders of different health and care organisations can facilitate or hinder integration in the UK (Jones and Barry, 2011; The Health Foundation, 2012; Shaw *et al.*, 2013; Timmins, 2015; Hulks *et al.*, 2017; Edwards, 2019; Erens *et al.*, 2019; Billings *et al.*, 2020). In particular, leader-leader relationships have been identified as important for helping create a shared vision and shared priorities across organisational boundaries (The Health Foundation, 2012; Shaw *et al.*, 2013; Maruthappu and Keogh, 2015; Hulks *et al.*, 2017; Edwards, 2019). However, both the considerable time needed to form successful leader-leader relationships, and the need for stable leadership to do this, continue to be identified as barriers to successful health and care integration (Ritchie *et al.*, 2008; Timmins, 2015; Erens *et al.*, 2019).

My study findings confirm this previous research. I found that good working relationships between leaders of different organisations circumvented organisational barriers by facilitating shared priority setting for 'Advanced' users of analytics. For 'Reluctant' users of analytics, cross-organisational strategic priorities were described as, at times, conflicting. This is because financial structures often disincentivised cross-sectoral working, particularly in acute settings where investments in system-wide priorities conflicted with the priorities of individuals' own organisations. This has also been identified as a barrier to integration in another recent study (Shand and Turner, 2019). Aspects of the government's 2021 White Paper included reforms that aim to facilitate shared priority setting across organisational boundaries, however separate financial budgets will remain for NHS and local government once the reforms come into force (Department of Health and Social Care, 2021b). My findings suggest

these reforms will be insufficient if we want to effectively include certain health providers (such as those providing acute care) in strategic and equitable health and care decision-making.

7.2.2.5 Factors relating to the skills and knowledge

Previous research by The Health Foundation has suggested that analysts do not always possess the skill levels needed to conduct and explain complex analyses to senior leaders, which negatively impacts analytical capacity across UK healthcare (Bardsley, 2016). In contrast, I did not find that analysts' skills and knowledge was a major barrier to senior leaders' use of analytics. In addition, participants challenged NCL's position that senior leaders' skills and knowledge around data and/or analytics was the main determinant of productive analytics use and that formal training would be beneficial.

Instead, the majority of participants believed that increased resources and less mandatory reporting to central bodies such as NHS England could more readily increase analysts' capacity and their ability to support senior leaders, as well as their own capacity to use the outputs of analytics to inform their decision-making. Indeed, these findings confirm findings previously reported by The Health Foundation and others, for example, in 2016, The Health Foundation reported that "there are too few analysts and those that are there are too busy working on mundane data manipulation" (Timmins, 2015; Bardsley, 2016). This suggests that policy changes aiming to increase analysts' capacity by reducing duplicative data requests so they can better support leaders could potentially improve the use of analytics for informing strategic and equitable decision-making (Department of Health and Social Care, 2020). This is particularly true for Hands-On users, who often chose to conduct their own analyses because of concerns around analysts' capacity to support their data needs.

7.2.2.6 Factors related to data quality

Previous research has found that poor data availability and quality creates barriers to delivering integrated care in England (Shaw *et al.*, 2013; Bardsley, 2016; NHS England, 2016; Billings *et al.*, 2020). For example, in their report, The Health Foundation also found that analysts' can have limited access to the right level of data needed to support decision-making (Bardsley, 2016). I found that data quality

continues to be a barrier for 'Challenged' and 'Reluctant' users of analytics when they attempted to use analytics for decision-making. In some cases, certain data necessary for a decision around health inequalities simply did not exist. My findings therefore suggest that, even with increased data linkage across health and council records, there will continue to be missing information around important circumstances that might help leaders better understand and plan services around health inequalities, simply because that information is not collected. Amongst other things, this includes information on sexual orientation, ethnicity, and religion. Researchers are exploring the plausibility of capturing this missing information using social media interactions (Kosinski, Stillwell and Graepel, 2013). Whilst my interviews suggest these efforts would be welcomed, there remains considerable ethical and information governance concerns that need to be addressed in order to realise the potential of these methods (Kosinski *et al.*, 2015).

As outlined in Chapter 1, increased data linkage across organisational boundaries is viewed as a potential enabler of more integrated care by helping overcome barriers stemming from organisational fragmentation (Bardsley, 2016; Charles *et al.*, 2018; Edwards, 2019). However, whilst technical solutions that collect, harmonise, integrate and share complex data have been develop in the private sector, considerable work is still needed to develop these solutions in the public sector (Van Panhuis *et al.*, 2014). Despite these challenges, my findings suggest that an increase in the availability of data linked across health and council records *could* influence strategic and equitable decision-making if further barriers to analytics use are also addressed. One participant was able to link data across primary and acute care records and, as a result, felt better able to understand local health inequalities, and tailored the subsequent services they commissioned to these. However, they still faced difficulties understanding wider determinants of health that would require council data. This is because linking council data to health records remains a challenge due to a lack of shared identifiers. It is currently unclear how upcoming policy reforms propose to improve data sharing, never mind data linkage, with local government (Local Government Association, 2021). The national government's forthcoming Data Strategy for Health and Care should consider how to improve data sharing and linkage with local government, which could facilitate integration and help realise aims to tackle health inequalities (Department of Health and Social Care, 2021c). This is

an important challenge to overcome as the vast majority of information on SdoH is held by constituent council departments.

7.2.3 Reflections on the full study fieldwork and analysis

My reflections on the pilot study, and any corresponding changes made for the full study, have been detailed in the relevant subsections of Chapter 6. Here, I will describe my reflections on the full study fieldwork and analysis.

Reflexivity is a process whereby a researcher examines their own beliefs and practices during the research process and considers how these may have influenced the research (Hennick, Hutter and Bailey, 2020). Reflexivity is encouraged in qualitative research as it is thought to legitimise and validate the research process, as well as query it (Hennick, Hutter and Bailey, 2020). During the full study fieldwork and analysis, I reflected on how my background, choices, position as a researcher and experiences may have influenced this study.

7.2.3.1 Sampling and recruitment strategy

For the full study, I successfully recruited my target sample size of 20 senior leaders with little difficulty. However, the nature of my pre-existing contacts influenced the participants' that I was subsequently put in contact with. For example, participants from organisations that operated in the London Borough of Enfield were harder to access, and these difficulties were compounded by a high turnover of senior leaders of Enfield council. To address this challenge, I relied heavily on purposive sampling strategies to recruit participants from health and care organisations within Enfield (Palinkas *et al.*, 2015). This was a challenge at the time of recruitment, though ultimately led to a good spread of participants from across the five boroughs and across the constituent organisations of NCL.

I recruited all participants via contacts I had at either NCL's Analytics Board or Islington and Camden councils. Participant's responses may have been shaped by whether they were recruited via contacts on the board or via other routes. For example, if I recruited a participant through the Analytics Board, that participant may have wanted to show themselves in a favourable light as they knew study findings were being relayed back to the board (Shattell and Thomas, 2005). To minimise the risk of this occurring, I assured participants that their participation would be kept

confidential, and all quotes used in dissemination materials kept anonymous. Participants may have also been more positive or selective about certain things they discussed (for example, they may have focused on barriers to analytics use they thought the board could more readily address). To minimise the risk of this, I reiterated that I was trying to capture the wide range of barriers and facilitators of analytics use. I also structured my topic guide to reflect this.

7.2.3.2 Interview process

Prior to each interview, I gave all participants information on the topics I would like to discuss. This may have introduced social desirability bias as participants may have considered their answers prior to interviews and crafted them in a favourable light (Shattell and Thomas, 2005). This was, however, a requirement as part of my ethics application. In addition, pilot study participants said that they had welcomed this information and that it encouraged them to participate. I, therefore, retained this approach when arranging interviews throughout the full study.

As expected with individual semi-structured interviews, I played an active role in shaping and guiding interview conversations. At times, I found myself nodding and smiling in response to some of their comments, particularly during the first few interviews. I think this was due to my lack of experience conducting qualitative interviews. This may have encouraged participants to talk more on certain topics if they felt I was indicating they were giving the 'right' answers and less about topics that they would have otherwise discussed without my presence (Shattell and Thomas, 2005). As the study progressed, I became more aware of how my non-verbal social cues may have been inadvertently shaping interviews and increasingly tried to mitigate the risk of this by making increased effort to keep my body language and facial expressions neutral.

As described in section 6.2.2, I made changes following the pilot study so that, for the full study, I asked participants to clarify their meaning of the word 'system'. I chose to do this to try and help me collectively analyse their examples of decisions. However, this change in approach was not as successful as I had hoped. At the start of each interview, I gave a brief introduction of the study origins, in which I mentioned NCL. On reflection, I think this led to most participants describing the structure of NCL when I asked them their meaning of the word 'system'. I tried to further unpick this by asking

them to describe the teams and organisations they worked with in their specific examples. Whilst this partly helped, their examples of strategic health and care decisions remained fairly unique and disparate, with a huge variety of practice partners involved in each decision. I, therefore, cannot define what participants meant by the 'system' from my results.

Throughout the full study, I found it difficult to ask participants about their own skills and knowledge with relation to analytics. For example, if a participant responded negatively to the first question I asked on this topic, I often struggled to follow it up as I was concerned I would offend participants. My position as a researcher, coming from an academic partner of NCL and studying 'linked health and council data', may have compounded this problem. Interview participants may have felt pressure to present themselves as knowledge with regards to data and/or analytics and as more 'evidence based' and data-driven in their decision-making (Shattell and Thomas, 2005). I tried to mitigate the risks of possible social desirability biases by reiterating that the study was exploratory in nature and by, again, trying to maintain neutral body language and facial expressions. I also repeatedly reiterated that the overall aim of this study was to find ways to help and support senior leaders with their analytics use. Video recording interviews could have captured additional data around social cues, such as voice, intonation and body language which could have added nuance to study results around contentious topics such as skills and knowledge in relation to analytics (for example, by allowing me to explore how comfortable participants appeared when discussing this topic). However, this approach would have been resource intensive and the data difficult to analyse. Interviews were, therefore, not recorded for pragmatic reasons.

My background as a non-clinical student, who has never worked in the NHS, in a local authority, or in any job role similar to those in which participants were working, may have shaped my study results. For example, I may have missed opportunities to probe certain topics further or misunderstood some of their responses due to my limited experience in these areas. To minimise the risk of missing anything of importance during analysis, I followed an extensive process of second coding, ensured two more knowledgeable researchers looked at my work at all stages, and discussed emergent themes with key staff at NCL during analysis and write-up.

7.2.3.3 Data analysis

Conducting this study was my first experience of collecting and analysing qualitative data, which I believe had its advantages and disadvantages. My approach to analysing the full study interviews may have been shaped by my approach to, and results from, analysing my pilot study interviews. The Framework Method is flexible tool that allows one to incorporate pre-existing ideas for possible themes, as identified from the literature or one's own prior work and understanding (Smith and Firth, 2011; Gale *et al.*, 2013). The method I chose to analyse data, and the way in which I deductively and inductively generated codes, should have mitigated any risks of my results being shaped by prior work. During the thematic analysis, I also found it difficult to group the codes I had generated into coherent, meaningful, and robust themes. The process of second coding that I followed and the support I received from more experienced qualitative researchers helped to refine my thinking and address this challenge.

7.2.4 Strengths and limitations

One of the strengths of this study is that it fills some of the gaps in local knowledge on the use of analytics by senior leaders for decision-making. Whilst this study has identified familiar factors that continue to facilitate and hinder health and care integration in England, it also offers novel and rich insights into the complexity of barriers and facilitators of analytics use for strategic and/or equitable decision-making that has implications across health and care. Furthermore, this study describes how these can impact decision-making for five different and distinct types of analytics users. As such, findings present practical ways in which organisations and local areas can support senior leaders to become more 'Advanced' users of analytics for strategic and/or equitable health and care decision-making.

A further strength of this study it that is explores these questions at a time when areas were on the cusp of transitioning from local models of integration (STPs) to statutory organisations (ICSs). As described in Chapter 1 of this thesis, the UK government views data as a key enabler of this transition. This study, therefore, presents timely findings that have implications for the governments' aim for data to transform care. These implications are discussed in more detail in Chapter 8.

A final strength of this study is that participants were from a wide range of health and care roles and organisations across a single STP (now ICS) case study site. This has provided nuanced, empirically-rich and context-specific findings. Previous studies in this area have primarily focused on the perspectives of NHS and public health leaders (Shaw *et al.*, 2013; Van Panhuis *et al.*, 2014; Bardsley, 2016). By including the perspectives of social care leaders and leaders of provider organisations such as acute care settings, this study offers a wider breadth of perspectives and a more detailed understanding of experiences of analytics use for strategic health and care decision-making.

There are, however, several limitations to this study. First, I recruited from one digitally engaged and digitally innovative London-based STP (now ICS) and was unable to recruit from my second study site – Care City. At the time of recruitment, NCL were actively pursuing a data linkage programme to facilitate their formation of an ICS and to better support health and care professionals in delivering more proactive care (as described in section 6.1.1). This kind of work is not common across other ICSs. Therefore, whilst it is likely that most of my study findings are transferable to other settings, this may not be the case for *all* findings and my findings may be not be transferable to less digitally engaged ICSs (Kuper, Lingard and Levinson, 2008). For example, the 'Waiting' analytics users, who were primarily characterised by their work setting up new, more advanced data systems that aimed to enable leaders' access to linked data, may not be as readily identified in less digitally innovative areas. In addition, there may be further and wider barriers that I have not captured in this study and that are important for better understanding the use of analytics for decision-making in these ICSs. However, despite NCL's focus on data and analytics and their digital innovation, this study still identified extensive barriers to analytics use for strategic and equitable health and care decision-making and it is likely these barriers have a greater impact in less digitally engaged and less digitally innovative ICSs.

Second, whilst NCL were actively pursuing a data linkage programme, I was exploring my research questions at a time when linked data were abstract for the vast majority of participants. As such, participant responses may reflect the hopes and aspirations for linked data that are echoed throughout national policy; namely, that linked data on its own *will* overcome most of the barriers to analytics use and *will* improve decision-making, care, and the equity of health and care services. When complete, a further

study that explores whether their data linking programme successfully improves decision-making in the ways hoped is therefore needed.

A third limitation of this study is that I found it difficult to ask participants about their own skills, knowledge and training needs with relation to analytics, as described in section 7.2.3.2. This was, in part, because of concerns that I would offend participants if I pursued a line of questioning they were not comfortable with. This difficulty was compounded by that fact that some participants found it hard to reflect on if and how training may be beneficial to their analytics use and how a lack of training may have impacted their decision-making. For example, as one participant put it, it was hard for participants to describe their training needs as "you don't know what you don't know". This may have meant I did not capture the range of opinions on this topic across participants. A data collection method that allowed participants to remain anonymous, such as a survey, may have better captured participants' views on this topic.

A fourth limitation is that this study presents a snapshot of analytics use prior to the onset of the Covid-19 pandemic and at a time when national policy priorities included shifting care out of acute settings. Following the onset of the Covid-19 pandemic in the UK, senior leaders were primarily and collectively focused on the pandemic response, which included adapting health and care services so that they could function in some capacity at a remote distance. Higher-level strategic decision-making that would see changes in the longer term, such as planning new services, was halted. As such, a minority of the barriers I have identified, like those around alignment of strategic cross-organisational priorities, may not be relevant in such times of crisis. In addition, sharing of certain data across organisations was mandated as part of the UK Covid-19 response. This may have mitigated some structural barriers related to data sharing during the pandemic. Despite this, barriers related to alignment of cross-organisation priorities and data sharing remain important for future partnership working. These findings also provide insight into possible strategies that could facilitate analytics use.

One final limitation of this study relates to my method of choice. Individual interview methods capture what people *say* their actions and experiences were in a given scenario, rather than their *actual* actions and experiences in practice. Observational research methods may be better suited to capture senior leaders' use of analytics and

how barriers and facilitators of analytics use influence their decision-making. However, these can be equally influenced by researcher presence (Mays and Pope, 1995). More importantly, observational research methods were not possible to utilise in this study as some strategic decision-making processes can last months if years. Participant and organisational acceptance of an external researcher observing decision-making processes may have also been less welcomed than individual interview methods.

7.3 Chapter summary

This chapter presents the results of my qualitative study, which investigated how knowledge generated from the analysis of linked health and council data *might* inform strategic and equitable decision-making that has implications across organisational boundaries. From the interview data I collected, I found that not all participants made strategic health and care decisions that considered health inequalities (addressing research question 1). When attempting to obtain data and/or analytics for strategic health and care decision-making, participants described the process as uncoordinated and random, which allowed various factors to influence this process (addressing research question 2). The factors spanned three areas: the contexts and environments in which individuals were working, the people involved in the process and decision and data quality. Across the senior leaders interviewed, participants differed in their experiences of, and responses to, these factors. This allowed me to identify five types of analytics users: 'Advanced' users, 'Hands-On' users, 'Waiting' users, 'Challenged' users and 'Reluctant' users of analytics (addressing research questions 3 and 4).

Overall, findings from this study challenge the assumption that knowledge generated from increased data sharing and linkage *will* improve decision-making, care, and the equity of services without strategies to address further, wider barriers to analytics use (Department of Health and Social Care, 2021b, 2021a). Most importantly, more is needed to better integrate organisations, align organisational priorities, and build and sustain cross-organisational relationships between leaders and analysts, and leaders of different organisations. Efforts to ensure that leaders view addressing health inequalities as part of their roles are also needed as, without these, increased data

linkage will unlikely uniformly influence the equity of strategic decision-making for health and care.

From this study, there are key take away messages for public health policy and practice (addressing research question 5). In Chapter 8, I will summarise these implications and identify where further research is needed.

Chapter 8 Discussion

As described in Chapter 2, the aims of this thesis were two-fold. First, I aimed to illustrate how knowledge generated from the analysis of linked health and council data could advance our understanding of the social determinants of multimorbidity (Aim 1). This aim was addressed by firstly undertaking a review that systematically identified, critically appraised, and synthesised existing literature examining household and area-level social determinants of multimorbidity (Chapter 3). Using findings from this review, I then conducted a quantitative analysis of a linked health and council dataset to examine and quantify associations between selected household characteristics and multimorbidity (Chapter 4 and Chapter 5). This study acts as a use case for creating, using, and analysing such linked datasets to understand social determinants of local public health concerns like multimorbidity and generate knowledge that could inform equitable decision-making for those with, or at risk of, multimorbidity.

The second aim of this thesis was to explore how such knowledge generated from the analysis of linked health and council data might influence strategic and equitable health and care decision-making (Aim 2). As described in Chapter 1, it is often assumed that knowledge generated from the analysis of such linked data (like that presented in Chapter 4 and Chapter 5) *will* improve decision-making and enable the delivery of more equitable health and care services. This aim was addressed by conducting individual, semi-structured interviews with senior leaders of health and care organisations (Chapter 6 and Chapter 7). I asked participants about their experiences of using analytics for strategic and equitable health and care decision-making, and the barriers and facilitators of analytics use they had faced in this context.

8.1 Key contributions of this thesis to the wider field

In Chapter 1 of this thesis, I described how our understanding of household and area-level social determinants of multimorbidity is incomplete, with most primary research focusing on area-level deprivation indices and previous reviews hampered by issues such as the exclusion of relevant literature. With the progressive move towards increased health and care integration, linked health and council datasets present opportunities to better understand the social determinants of important public health

problems such as multimorbidity. However, the value of such data linkages for this purpose, and for generating knowledge that could inform more equitable decision-making, remained unclear.

As described in Chapter 1, it is also often assumed that linked data *will* improve decision-making, care, and the equity of health and care services, yet this is a forgone conclusion. Prior to this thesis, there were gaps in local knowledge around understanding when and how analytics are used to inform decision-making, and it was unclear what facilitates and hinders analytics use for informing local strategic and equitable health and care decision-making.

The research undertaken as part of this thesis attempted to address these gaps in literature and knowledge, and has made several contributions to the wider field.

8.1.1 Household-level social determinants of multimorbidity

My systematic review described in Chapter 3 found that household income and area-level deprivation were the most explored contextual social determinants of multimorbidity, and findings for these were fairly consistent; multimorbidity risk was higher for those within the lowest level of household income versus the highest, and higher for individuals living in the most versus the least deprived areas. However, aside from measures of household income, my review found that household social determinants of multimorbidity are often overlooked in the literature despite comparatively large effect sizes for household compared to area-level SDoH. In particular, no previous studies had examined associations between household tenure and multimorbidity in the English context, despite information on tenure being systematically recorded in council data and expected to have good completeness and validity.

My quantitative study described in Chapter 4 and Chapter 5 partly addressed this gap and found that household tenure was an important social factor associated with multimorbidity amongst working age adults residing in LBBD in 2019/20. I found that, compared to owner-occupiers, the risk of multimorbidity was greater for residents of social housing and lower for those who privately rent. These associations were consistent across three different definitions of multimorbidity as well as several

sensitivity analyses, demonstrating the strength of household tenure as an exposure for understanding household-level inequalities in multimorbidity.

8.1.2 Use of linked data to better understand social determinants of health

My quantitative analysis of a linked health and council dataset described in Chapter 4 and Chapter 5 illustrates that collecting data on, and investigating, household-level social determinants of multimorbidity could generate useful and valuable knowledge that could inform more equitable decision-making for those with, or at risk of, multimorbidity. This research demonstrates the value of investigating SDoH other than measures of area-level deprivation to advance understanding of population health, and thus better target health prevention or promotion resources. My findings could be used to identify and support groups vulnerable to multimorbidity that would not have been possible to identify without the individual and household-level linkages of the separate health and council datasets.

8.1.3 Use of linked data to inform decision-making

My qualitative study described in Chapter 6 and Chapter 7 is the first to explore whether linked health and council data could or would be used by senior leaders to influence the equity of strategic decision-making. This study has addressed major gaps in local knowledge. For example, despite the knowledge on SDoH that can be generated from linked health and council data, my interviews found that leaders did not always feel it was legitimate to consider health inequalities as part of their roles. As such, these leaders did not uniformly use the outputs of analytics to inform strategic and equitable health and care decision-making.

This study has also added to current knowledge by demonstrating the considerable and complex barriers that senior leaders faced when attempting to access and use analytics for decision-making. These barriers spanned the environments in which they worked, the people involved in the process and decision, and data quality. Good working relationships between leaders and analysts, and leaders of different organisations, were described as key facilitators of analytics access and use by enabling leaders to circumvent barriers stemming from organisational fragmentation. Across the senior leaders interviewed, participants differed in their experiences of, and responses to, these factors, and this allowed me to identify five novel types of

analytics users. These results suggested that knowledge generated from the analysis of linked health and council data could inform strategic and equitable decision-making for more 'Advanced' analytics users, but more is needed to address further, wider barriers to analytics use to enable all senior leaders to use analytics from linked data for the explored purposes.

8.2 Strength and limitations of this thesis

The strengths and limitations of each study are described in detail in the respective chapters of this thesis. Here, I describe the overarching strengths and limitations of this thesis.

8.2.1 Research setting

The linked health and council dataset analysed to address Aim 1 of this thesis was created by linking administrative records for residents of one borough in East London – LBBD. The interviews conducted to address Aim 2 of this thesis recruited senior health and care leaders working for organisations within a single ICS, covering a defined geography in North London. As such, both my quantitative and qualitative study findings provided rich, nuanced, and context-specific population health insights. As a result, local system leaders in North and East London were particularly interested in study findings, which facilitated the dissemination of findings locally and maximised the impact of this research.

There were, however, disadvantages afforded by restricting my thesis to select geographical areas in London. Both settings in North and East London are socially deprived, with younger and more ethnically diverse populations compared to the rest of England. As such, the extent to which findings from each study can be generalised and transferred to less deprived, more rural settings in England, and other settings internationally, is under question (Zaccai, 2004; Kuper, Lingard and Levinson, 2008). However, these findings could be relevant to other urban areas with young and deprived populations. The overarching learnings from this thesis could also be applicable and useful to other areas looking to create linked datasets and use the knowledge generated from analysis of these datasets to inform decision-making.

8.2.2 Mixed methods

One strength of this thesis is that it has used a mixed methods research design which enabled me to answer two related research questions (Creswell, Fetters and Ivankova, 2004; Tariq and Woodman, 2013). The value and utility of linked health and council datasets covering a defined geography are, in part, determined by their ability to answer questions of interest locally that are not possible to answer if this data is held separately. A mixed methods design allowed me to use quantitative research methods to illustrate how knowledge generated from the analysis of linked health and council data could advance our understanding of the social determinants of multimorbidity. This study acted as a use case for creating and using such data to better understand important public health problems and generate knowledge that could inform equitable decision-making. However, the value and utility of locally linked health and council datasets are also dependent on whether they realise hopes of improving decision-making, care, and the equity of health and care services for populations such as those with, or at risk of, multimorbidity. To address this required qualitative research methods to gather in-depth data on analytics use for informing strategic and equitable health and care decision-making.

Nevertheless, I was unable to align my quantitative and qualitative studies both geographically and temporally. In late 2017, the London Borough of Islington, Islington CCG, North East London Commissioning Support Unit, and the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC; now ARC) North Thames were granted a Health Foundation Advancing Applied Analytics award for a project aiming to create a dataset linking NHS and local government data for households across Islington. I had initially planned to use this dataset for my quantitative study and conduct my interviews amongst the individuals that had access to this data to explicitly link what I asked in interviews to my quantitative findings. This would have, arguably, enabled me to better answer my two related research questions. However, the application process stalled before reaching the formal Independent Group Advising on the Release of Data review process at NHS Digital due to a combination of factors including frequent staff changes at NHS Digital and variable interpretations of data protection regulations. These delays meant I had to find a suitable linked dataset from another source and setting (The Care City Cohort from LBBD). To mitigate this, I had planned to conduct interviews with leaders across LBBD. However, the onset of the Covid-19 pandemic meant I had to stop recruitment

and all of my interviews were conducted with senior leaders of NCL, a different geographical area.

There were, however, advantages to collaborating with two areas. As mentioned in section 8.2.1, local system leaders in North and East London were interested in both my quantitative and qualitative findings, and I was invited to present my findings at NCL, Care City, and Camden and Islington Councils. This not only helped disseminate findings more widely but allowed me to gain feedback from a broader and more diverse pool of leaders, some of whom were trying to link data and some of whom were using linked data to inform decision-making. For example, I presented my quantitative findings to colleagues from the Camden and Islington Public Health Team, who stated that these findings underlined the importance of what they were trying to do locally (linking household-level data via UPRNs), and provided a strong use case that demonstrated the opportunities available when data is linked in this way. Collaborating with two areas at different stages of data linkage also allowed me to gain a thorough understanding of some of the challenges of working in applied health research when trying to create and use linked datasets for practice and research purposes. I was able to learn approaches to managing these challenges including maintaining a degree of flexibility, adapting research plans to mitigate risks, resilience, and the importance of discussing the possibility of failure upfront in order to maintain good research partnerships despite project setbacks.

8.2.3 Collaboration with local partners

Another key strength of this thesis is that, throughout, I worked collaboratively with local health and care leaders. For the quantitative study, I worked with staff at the LBBD and staff at Care City. For the qualitative study, I worked with NCL's Analytics Board and staff across the London Boroughs of Camden and Islington. This meant that the research questions for this thesis were tailored to address local priorities and interests. This also meant that local leaders put me in contact with potential participants for my qualitative study and, as such, I was easily able to recruit my desired sample size. As described in sections 8.2.1 and 8.2.2, my local collaborations facilitated the dissemination of findings and increased the impact of my research.

8.2.4 Covid-19

One strength of this thesis is that the findings are timely and may be of greater relevance given the Covid-19 pandemic. Senior decision-makers and policy makers have positioned data and analytics at the forefront of the UK's pandemic response, and senior leaders are particularly engaged in data and analytics at this current point in time. This means that my findings exploring the value and utility of linked health and council datasets could be of greater interest to leaders both locally and nationally than they may have been prior to the pandemic.

A limitation of this thesis is that both the quantitative and qualitative studies were conducted prior to the onset of Covid-19 – the quantitative study was conducted using a cohort of individuals that resided in LBBD between April 2019 and April 2020, and interviews were conducted between January and March 2020. Associations between household tenure and multimorbidity may have changed as a result of Covid-19 – local and national lockdowns have seen the majority of the population confined to their households, which may have increased the impact of household-level factors on health. In addition, data sharing was identified as a barrier to analytics use in my qualitative study yet sharing of certain data across organisations was mandated as part of the UK Covid-19 response (Department of Health and Social Care, 2020). This, plus increased focus on the value of data for managing the pandemic, may have mitigated some of the structural and cultural barriers related to data that I identified in my interviews. However, despite these limitations, my findings still have key take home messages that are relevant for current policy and practice.

8.3 Implications

8.3.1 For policy and practice

8.3.1.1 Social determinants of multimorbidity

Recent policy changes legislating the integration of health and care in England explicitly state that one aim of joining up and integrating care is to meet the rising challenge of multimorbidity (Department of Health and Social Care, 2021b). Findings from Chapter 5 of this thesis support this vision by illustrating how potential risk factors for multimorbidity are often social in nature and influenced by public services outside of typical health services such as housing. Local councils are responsible for providing

and managing council housing, and for monitoring the standards of other tenure types. Improved integration between local councils and health organisations therefore presents opportunities to better understand and, in turn, address SDoH and social determinants of multimorbidity.

Most interventions for those with multimorbidity focus on older adults or those who have experienced or are at risk of numerous admissions to secondary care. This is despite evidence suggesting that the absolute number of those with multimorbidity is greater amongst those 65 years old and below, and the incidence of multimorbidity and socioeconomic inequalities in multimorbidity are increasing amongst people of working age (Barnett *et al.*, 2012; Head, Fleming, Kypridemos, Schofield, *et al.*, 2021). Findings from Chapter 5 of this thesis indicate that working age adults are an important population to consider when aiming to address multimorbidity. There is currently a gap in models of care or interventions aimed at working age adults, for whom there may be greater opportunity for prevention of multimorbidity through addressing SDoH than amongst older adults. Policies or initiatives that target preventative resources at working age adults with multimorbidity could curtail the progression of multimorbidity and the emergence of more complex multimorbidity profiles, ultimately saving future costs for the health and care system (Head, Fleming, Kypridemos, Pearson-Stuttard, *et al.*, 2021; Head, Fleming, Kypridemos, Schofield, *et al.*, 2021).

The analysis presented in Chapter 5 suggests that it is important for those designing and delivering services to consider drivers of multimorbidity beyond measures of area-level socioeconomic deprivation. This is particularly important for deprived areas like LBBD, where most of the borough is within the top two quintiles of the English IMD. My findings demonstrate that resources to tackle multimorbidity could be targeted differentially by tenure type. Social housing tenants could be targeted with resources aiming to prevent and help them better manage their multimorbidity. This group is also more likely to already be in contact with council services, presenting possible avenues for identifying, targeting, and supporting those most at risk. My findings also suggest that targeting resources at all private renters in deprived areas would miss those most at risk of multimorbidity, although private renters that receive benefits and live alone could be a subgroup of the population that would benefit from the targeting of health prevention and promotion resources.

8.3.1.2 Linked health and council data

Findings from Chapter 5 of this thesis demonstrate that snapshots of static, linked health and council data can provide a deeper understanding of population health and social determinants of health. Developing this understanding became a key requirement for all ICSs in April 2021 (Department of Health and Social Care, 2021b). However, despite the opportunities these linked datasets present, my interviews suggest that progress to link health and care datasets across organisational boundaries is slow and challenging due to difficulties around information governance and the considerable resources needed to support data sharing. This suggests that efforts to create these linked datasets should be focused at the local level, where there are increased opportunities to overcome information governance challenges and develop the relationships needed to facilitate data sharing.

Whilst recent government policy places great focus on data as a driving force of integration and improved care delivery (see Chapter 1), data linkage across organisations receives little attention in national policy and there is little guidance for local areas aiming to create linked datasets. For example, the 2021 White Paper legislated reforms aiming to continue increased data sharing after the pandemic yet did not discuss data linkage, instead generally committing to improving data availability and quality (Department of Health and Social Care, 2021b, 2021a). In addition, the government's "Data saves lives" policy paper released in June 2021 simply states that "ICSs will help the NHS join up data" and, throughout the paper, this was the only point at which data linkage was mentioned (Department of Health and Social Care, 2021a). Findings from Chapter 7 of this thesis suggest that policy reforms from national government should include structural changes that facilitate data sharing and, in turn, linkage. These should include fundamental changes in financial incentives so that acute care settings are better able to share their data for linked data initiatives to help support the integration agenda (Edwards, 2019; Erens *et al.*, 2019).

8.3.1.3 Using analytics to inform strategic and equitable health and care decision-making

As described in Chapter 1 of this thesis, much of the rhetoric around data linkage often assumes that linked data *will* improve decision-making, care, and the equity of health and care services. However, findings from Chapter 7 suggest that, whilst widely

welcomed, programmes linking administrative data across health and care may not uniformly improve the equity of care delivery, if successful (HSR UK, 2017; Charles *et al.*, 2018). This is because leaders may continue to believe that any consideration of health inequalities during their decision-making is not a legitimate part of their job role, even if their umbrella organisation (ICS) explicitly aims to challenge and address health inequalities. More is needed to ensure that ICS priorities align with the priorities of their constituent organisations, particularly if shared priority setting is viewed as a cornerstone of integration (Department of Health and Social Care, 2021b).

Findings from my interviews also suggest that *solely* linking health and council data will be insufficient to realise the White Paper's aspiration for data to transform care without strategies to address the relational barriers to analytics access and use. Whilst a recent government policy paper places considerable focus on the development of analytical skills for analysts (Department of Health and Social Care, 2021a), my findings suggest that it would be more productive to target resources that develop and sustain relationships and shared understanding between leaders and analysts and leaders of different organisations. This could facilitate shared priority setting and the use of data and analytics to inform strategic and equitable health and care decision-making. As previously commented by Scott et al. (2021): "The strategy rightly aims to recruit more analysts and data scientists. However, a learning health and care system is sociotechnical—it includes people working in interdisciplinary teams, not just technology and data science." (Scott, Emerson and Henderson-Reay, 2021).

8.3.2 For research

8.3.2.1 Social determinants of multimorbidity

My systematic review (Chapter 3) and quantitative analysis (Chapter 5) suggest that social determinants of multimorbidity operating at the household-level, such as household tenure, warrant further exploration and may explain associations above and beyond individual-level factors. This is in keeping with evidence that suggests households (or families) influence physical and mental health through various material and psychosocial factors – households and families often exhibit similar health behaviours, financial circumstances, and other common risk factors for multimorbidity (McNeill, 2010; Vaezghasemi *et al.*, 2016).

Further research that employs longitudinal methods is also needed to assess causal relationships between household SDoH such as tenure and multimorbidity, as the temporal associations between housing and multimorbidity are likely to have long gestation periods. However, adopting a lifecourse epidemiological approach will be a challenging due to data availability and changes in household circumstances, amongst other things. Whilst longitudinally linked health and council data may enable one to better explore these associations, further information on potential confounders of the relationships between household SDoH and multimorbidity may also be needed to establish causality. This includes data on household income, and markers of poor housing conditions including overcrowding and damp. Further research should also explore whether the observed associations I have found hold for children and older adults.

8.3.2.2 Linked health and council data

Data linkage is well known as a powerful way to improve the completeness and accuracy of patient and resident information for population health research. One of the key advantages of using linked health and council data, such as the dataset presented in Chapter 4, is it codes UPRNs and, therefore, allows researchers to group individuals within households. In theory, this should allow one to account for shared characteristics attributable to these groups that go above and beyond individual-level characteristics. However I was unable to accurately model and account for clustering of individuals within households as, in multilevel regression modelling, too few individuals per cluster leads to less precise random intercepts and too small standard errors (Austin and Leckie, 2018). Yet my findings suggest that the exclusion of a household-level may not necessarily and substantially impact model estimates, and the addition of UPRNs can still add value by enabling the characterisation of households in terms of tenure, occupancy, type, and benefits receipt. Many areas should, therefore, now be able to further their understanding of household-level population health as the inclusion and usage of UPRNs in administrative data increases, even it is not possible to account for household-level clustering (GeoPlace LLP, 2020; Scottish Centre for Adminstrative Data Research, 2020).

My analyses presented in Chapter 5 suggest that selection biases were not introduced in selected variables originating from primary care records (including age, sex, and multimorbidity status) when primary care records were linked to council

records. This is particularly relevant for researchers looking to utilise similar types of linked data. However, I was unable to assess whether any selection biases were introduced in council variables for matched compared to unmatched primary care records as, by definition, unmatched primary care records did not have corresponding council data. It is plausible that there were biases introduced at this stage of linkage, for example, councils may underestimate owner-occupiers and privately rented households that have little contact with council services. Further research is needed to assess whether selection biases are introduced in council variables for matched compared to unmatched primary care records.

8.3.2.3 Using analytics to inform strategic and equitable health and care decision-making

The majority of participants interviewed for this research did not have access to linked health and council data and, instead, described instances of using analytics derived from unlinked datasets to inform decision-making. One of the aims of linking health and council data is to facilitate the use of analytics by senior leaders for strategic decision-making, *in general*, but particularly to identify and better serve populations where service provision across organisational boundaries has been inadequate. However, few, if any, of these linked health and council datasets have been evaluated for such purposes. This presents outstanding research questions. Have these linked datasets better informed decision-making when compared to unlinked datasets? Have they been used to inform strategic and *equitable* decision-making across organisational boundaries, as intended? Further research is needed to evaluate the adoption and success of creating these linked datasets for informing decision-making. This is imperative as considerable time and resources, often funded by the taxpayer, can be expended to create and implement these datasets in practice.

Findings from my interviews suggest that for more leaders to become 'Advanced' analytics users, more is needed to better integrate organisations, align organisational priorities, and build and sustain cross-organisational relationships between leaders and analysts, and leaders of different organisations. *How* to best address these barriers within a resource-constrained environment is currently unclear. Further research is needed that examines the successes and challenges of implementing programmes that aim to develop relationships between leaders and analysts or aim to facilitate conversations between leaders of different organisations. It is likely that

such programmes will be more successful if developed locally, with local priorities and needs incorporated from the onset.

8.4 Concluding remarks

Datasets that link health and council records to support the integration of health and care present opportunities to generate knowledge that can advance our understanding of the social determinants of health and, in turn, be used to inform strategic and equitable health and care decision-making. Through the lens of the important public health problem of multimorbidity, this thesis explores these two related areas.

This thesis has demonstrated how linked health and council data can be used to provide novel and actionable population health insights for local concerns like multimorbidity. These insights have revealed how household tenure data extracted from council records is associated with multimorbidity and the strength of household tenure as an exposure for understanding household-level inequalities in multimorbidity. These analyses would not have been possible without the individual and household-level linkages of the separate health and council datasets.

This thesis has challenged some of the policy assumptions behind the creation of such linked data, namely that knowledge generated from linked data *will* improve decision-making, care, and the equity of health and care services. My findings suggest that without efforts to address the wider relational and organisational barriers to analytics access and use, building linked data systems will not lead to substantial changes in the equity of health and care provision for populations such as those with, or at risk of, multimorbidity.

Bibliography

1. ADR UK (2021a) *Administrative Data*. Available at: https://www.adruk.org/our-mission/administrative-data/ (Accessed: 28 April 2021).
2. ADR UK (2021b) *Our Mission*. Available at: https://www.adruk.org/our-mission/our-mission/ (Accessed: 27 April 2021).
3. Agborsangaya, C. B. *et al.* (2012) 'Multimorbidity prevalence and patterns across socioeconomic determinants: a cross-sectional survey.', *BMC public health*, 12(1), p. 201. doi: 10.1186/1471-2458-12-201.
4. Agborsangaya, C. B. *et al.* (2013) 'Multimorbidity prevalence in the general population: the role of obesity in chronic disease clustering', *BMC Public Health*, 13(1), p. 1161. doi: 10.1186/1471-2458-13-1161.
5. Akazawa, M. *et al.* (2010) 'Potentially inappropriate medication use in elderly japanese patients', *American Journal Geriatric Pharmacotherapy*, 8(2), pp. 146–160. doi: 10.1016/j.amjopharm.2010.03.005.
6. Van den Akker, M. *et al.* (1998) 'Multimorbidity in general practice: Prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases', *Journal of Clinical Epidemiology*, 51(5), pp. 367–375. doi: 10.1016/S0895-4356(97)00306-5.
7. van den Akker, M., Buntinx, F. and Knottnerus, J. A. (1996) 'Comorbidity or multimorbidity', *European Journal of General Practice*, 2(2), pp. 65–70. doi: 10.3109/13814789609162146.
8. Anderson, D. A., Han, H.-S. and Hisnanick, J. (2021) 'The Effect of Household Debt and Wealth on Subsequent Housing Tenure Choice', *City and Community*, 00(0), pp. 1–29. Available at: www.census.gov/programs-surveys/sipp/tech-documentation/nonresponse-reports.html (Accessed: 4 September 2019).
9. André, S., Dewilde, C. and Muffels, R. (2019) 'Housing Tenure, Housing Wealth and Subjective Wellbeing in Australia: The Case of Unemployment', in *Wealth(s) and Subjective Well-Being*. Springer, pp. 281–304. doi: 10.1007/978-3-030-05535-6_13.
10. Ansar, W. and Ghosh, S. (2013) 'C-reactive protein and the biology of disease', *Immunologic Research*, 56(1), pp. 131–142. doi: 10.1007/s12026-013-8384-0.
11. Arbelle, J. E. *et al.* (2014) 'Multiple chronic disorders - health care system's modern challenge in the Maccabi Health Care System', *Israel Journal of Health Policy Research*, 3(1), p. 29. doi: 10.1186/2045-4015-3-29.
12. Armitage, G. D. *et al.* (2009) 'Health systems integration: State of the evidence', *International Journal of Integrated Care*, 9(2). doi: 10.5334/ijic.316.
13. Ashworth, M. *et al.* (2019) 'Journey to multimorbidity: Longitudinal analysis exploring cardiovascular risk factors and sociodemographic determinants in an urban setting', *BMJ Open*, 9(12). doi: 10.1136/bmjopen-2019-031649.
14. Austin, P. C. and Leckie, G. (2018) 'The effect of number of clusters and cluster size on statistical power and Type I error rates when testing random effects variance components in multilevel linear and logistic regression models', *Journal of Statistical Computation and Simulation*, 88(16), pp. 3151–3163. doi: 10.1080/00949655.2018.1504945.
15. Axelsson, R. and Axelsson, S. B. (2006) 'Integration and collaboration in public health - A conceptual framework', *International Journal of Health Planning and Management*, 21(1), pp. 75–88. doi: 10.1002/hpm.826.

16. Bailey, K. (2011) *Typologies and Taxonomies*, *Typologies and Taxonomies*. SAGE Publications, Inc. doi: 10.4135/9781412986397.
17. Baird, A. G. and Wright, N. (2006) 'Poor access to care: Rural health deprivation?', *British Journal of General Practice*, 56(529), pp. 567–568. Available at: http://www.bma.org.uk/ap.nsf/Content/healthcarerural (Accessed: 6 May 2021).
18. Bardsley, M. (2016) *Understanding analytical capability in health care: Do we have more data than insight?*, *The Health Foundation*. Available at: https://www.health.org.uk/sites/default/files/UnderstandingAnalyticalCapabilityInHealthCare.pdf (Accessed: 23 September 2019).
19. Bardsley, M., Steventon, A. and Fothergill, G. (2019) *Untapped potential: Investing in health and care data analytics*, *The Health Foundation*. Available at: https://www.health.org.uk/sites/default/files/upload/publications/2019/Untapped potential.pdf.
20. Barking and Dagenham Council (2001) *Growth and changes in our local population*. Available at: https://www.lbbd.gov.uk/sites/default/files/attachments/1.2-Growth-and-changes-in-population.pdf (Accessed: 13 October 2021).
21. Barking and Dagenham Council (2019) *Applying for council housing | LBBD*. Available at: https://www.lbbd.gov.uk/applying-for-council-housing (Accessed: 5 September 2019).
22. Barking and Dagenham Council (2020) *Population and demographic data | LBBD*. Available at: https://www.lbbd.gov.uk/population-and-demographic-data (Accessed: 10 October 2021).
23. Barnett, K. *et al.* (2012) 'Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross-sectional study', *The Lancet*, 380(9836), pp. 37–43. doi: 10.1016/S0140-6736(12)60240-2.
24. Basham, C. A. and Karim, M. E. (2019) 'Multimorbidity prevalence in Canada: a comparison of Northern Territories with Provinces, 2013/14', *International Journal of Circumpolar Health*, 78(1). doi: 10.1080/22423982.2019.1607703.
25. Bates, D. *et al.* (2015) 'Fitting linear mixed-effects models using lme4', *Journal of Statistical Software*, 67(1). doi: 10.18637/jss.v067.i01.
26. Bayliss, E. A. *et al.* (2008) 'Processes of care desired by elderly patients with multimorbidities', *Family Practice*, 25(4), pp. 287–293. doi: 10.1093/fampra/cmn040.
27. Beenstock, J. *et al.* (2014) 'Are health and well-being strategies in England fit for purpose? A thematic content analysis', *Journal of Public Health*, 37(3), pp. 461–469. doi: 10.1093/pubmed/fdu073.
28. Berends, L. and Johnston, J. (2005) 'Using multiple coders to enhance qualitative analysis: The case of interviews with consumers of drug treatment', *Addiction Research and Theory*, 13(4), pp. 373–381. doi: 10.1080/16066350500102237.
29. Bhaskaran, K. *et al.* (2013) 'Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD)', *BMJ Open*, 3(9). doi: 10.1136/bmjopen-2013-003389.
30. Billings, J. *et al.* (2020) 'Supporting integrated care in practice: Perceptions of a national support programme for the development of new care models in England', *Journal of Integrated Care*, 28(1), pp. 27–40. doi: 10.1108/JICA-05-2019-0018.
31. Bisgaier, J. and Rhodes, K. V. (2011) 'Cumulative adverse financial circumstances: Associations with patient health status and behaviors', *Health*

and Social Work, 36(2), pp. 129–137. doi: 10.1093/hsw/36.2.129.
32. Bloor, M. (1997) 'Techniques of Validation in Qualitative Research: a Critical Commentary In: Context and Method in Qualitative Research'. doi: 10.4135/9781849208758.
33. Bobo, W. V. *et al.* (2016) 'Prevalence of Combined Somatic and Mental Health Multimorbidity: Patterns by Age, Sex, and Race/Ethnicity', *J Gerontol A Biol Sci Med Sci*, 71(11), pp. 1483–1491. doi: 10.1093/gerona/glw032.
34. Bond, L. *et al.* (2012) 'Exploring the relationships between housing, neighbourhoods and mental wellbeing for residents of deprived areas', *BMC Public Health*, 12(1), pp. 1–14. doi: 10.1186/1471-2458-12-48.
35. Bower, P. *et al.* (2011) 'Multimorbidity, service organization and clinical decision making in primary care: A qualitative study', *Family Practice*, 28(5), pp. 579–587. doi: 10.1093/fampra/cmr018.
36. Bowleg, L. (2012) 'The problem with the phrase women and minorities: Intersectionality-an important theoretical framework for public health', *American Journal of Public Health*, 102(7), pp. 1267–1273. doi: 10.2105/AJPH.2012.300750.
37. Bowman, C., Whistler, J. and Ellerby, M. (2004) 'A national census of care home residents', *Age and Ageing*, 33(6), pp. 561–566. doi: 10.1093/ageing/afh177.
38. Boyd, C. M. *et al.* (2005) 'Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: Implications for pay for performance', *Journal of the American Medical Association*, 294(6), pp. 716–724. doi: 10.1001/jama.294.6.716.
39. Boyd, C. M. and Fortin, M. (2010) 'Future of Multimorbidity Research: How Should Understanding of Multimorbidity Inform Health System Design?', *Public Health Reviews*, 32(2), pp. 451–474. doi: 10.1007/BF03391611.
40. Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3(2), pp. 77–101. doi: 10.1191/1478088706qp063oa.
41. Braun, V. and Clarke, V. (2014) 'What can "thematic analysis" offer health and wellbeing researchers?', *International Journal of Qualitative Studies on Health and Well-being*. Taylor & Francis, p. 26152. doi: 10.3402/qhw.v9.26152.
42. Braveman, P. and Gottlieb, L. (2014) 'The Social Determinants of Health: It's Time to Consider the Causes of the Causes', *Public Health Reports*, 129(1_suppl2), pp. 19–31. doi: 10.1177/00333549141291S206.
43. Britten, N. (1995) 'Qualitative Research: Qualitative interviews in medical research', *BMJ*, 311(6999), p. 251. doi: 10.1136/bmj.311.6999.251.
44. Buist-Bouwman, M. A. *et al.* (2005) 'Comorbidity of physical and mental disorders and the effect on work-loss days', *Acta Psychiatrica Scandinavica*, 111(6), pp. 436–443. doi: 10.1111/j.1600-0447.2005.00513.x.
45. Bunker, J. P. (2001) 'The role of medical care in contributing to health improvements within societies', *International Journal of Epidemiology*, pp. 1260–1263. doi: 10.1093/ije/30.6.1260.
46. Cambridge, C. @ (2018) *Codes Lists Version 1.1 - October 2018 - Primary Care Unit*. Available at: https://www.phpc.cam.ac.uk/pcu/cprd_cam/codelists/v11/ (Accessed: 6 November 2019).
47. Canadian Council on Social Determinants of Health (2015) *A Review of Frameworks on the Determinants of Health*. Available at: http://ccsdh.ca/images/uploads/Frameworks_Report_English.pdf (Accessed: 28 February 2019).

48. Cantarero-Prieto, D., Pascual-Sáez, M. and Blázquez-Fernández, C. (2018) 'Social isolation and multiple chronic diseases after age 50: A European macro-regional analysis', *PLoS ONE*, 13(10), p. e0205062. doi: 10.1371/journal.pone.0205062.
49. Cassell, A. *et al.* (2018) 'The epidemiology of multimorbidity in primary care: A retrospective cohort study', *British Journal of General Practice*, 68(669), pp. e245–e251. doi: 10.3399/bjgp18X695465.
50. Centers for Disease Control and Prevention (2021) *The Social-Ecological Model: A Framework for Prevention*. Available at: https://www.cdc.gov/violenceprevention/about/social-ecologicalmodel.html (Accessed: 8 October 2021).
51. Cepoiu, M. *et al.* (2007) 'Recognition of Depression by Non-psychiatric Physicians-A Systematic Literature Review and Meta-analysis', *J Gen Intern Med*, 23(1), pp. 25–36. doi: 10.1007/s11606-007-0428-5.
52. Chamberlain, A. M. *et al.* (2020) 'Neighborhood socioeconomic disadvantage is associated with multimorbidity in a geographically-defined community', *BMC Public Health*, 20(1). doi: 10.1186/s12889-019-8123-0.
53. Charles, A. *et al.* (2018) *A year of integrated care systems: reviewing the journey so far*, *The King's Fund*. Available at: https://www.kingsfund.org.uk/sites/default/files/2018-09/Year-of-integrated-care-systems-reviewing-journey-so-far-full-report.pdf (Accessed: 18 January 2019).
54. Charlson, M. E. *et al.* (1987) 'A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation', *Journal of Chronic Diseases*, 40(5), pp. 373–383. doi: 10.1016/0021-9681(87)90171-8.
55. Chen, Y. H., Karimi, M. and Rutten-Van Mölken, M. P. M. H. (2020) 'The disease burden of multimorbidity and its interaction with educational level', *PLoS ONE*, 15(12: e0243275). doi: 10.1371/journal.pone.0243275.
56. Chudasama, Y. V *et al.* (2019) 'Physical activity, multimorbidity, and life expectancy: A UK Biobank longitudinal study', *BMC Medicine*, 17(1). doi: 10.1186/s12916-019-1339-0.
57. Chung, R. Y. *et al.* (2015) 'Socioeconomic determinants of multimorbidity: A population-based household survey of Hong Kong Chinese', *PLoS ONE*, 10(10), p. e0140040. doi: 10.1371/journal.pone.0140040.
58. Chung, R. Y. *et al.* (2016) 'The association between types of regular primary care and hospitalization among people with and without multimorbidity: A household survey on 25,780 Chinese', *Scientific Reports*, 6, pp. 1–9. doi: 10.1038/srep29758.
59. Clair, A. and Hughes, A. (2019) 'Housing and health: new evidence using biomarker data', *J Epidemiol Community Health*, 73, pp. 256–262. doi: 10.1136/jech-2018-211431.
60. Clarke, A. *et al.* (2013) 'Evidence-based commissioning in the English NHS: who uses which sources of evidence? A survey 2010/2011', *BMJ Open*, 3: e002714. doi: doi:10.1136/bmjopen-2013- 002714.
61. Contandriopoulos, A.-P., Denis, J.-L. and Touati, N. (2001) 'Intégration des soins: concepts et mise en oeuvre', *Ruptures, revue transdisciplinaire en santé*, 8(2), pp. 38–52. Available at: www.medsp.umontreal.ca/getos/pdf/intgrsoin.pdf (Accessed: 5 October 2021).
62. Cox, M. *et al.* (2007) 'Locality deprivation and Type 2 diabetes incidence: A local test of relative inequalities', *Social Science & Medicine*, 65(9), pp. 1953–1964. doi: 10.1016/J.SOCSCIMED.2007.05.043.

63. Creswell, J. W., Fetters, M. D. and Ivankova, N. (2004) 'Designing A Mixed Methods Study In Primary Care', *Annals of Family Medicine*, 2(1). doi: 10.1370/afm.104.
64. Dahlgren, G. and Whitehead, M. (1991) *Policies and strategies to promote social equity in health*. Available at: https://www.researchgate.net/publication/5095964 (Accessed: 28 February 2019).
65. Delgado-Rodríguez, M. and Llorca, J. (2004) 'Bias', *J Epidemiol Community Health*, 58, pp. 635–641. doi: 10.1136/jech.2003.008466.
66. Department of Health (2012) *Long Term Conditions Compendium of Information: Third Edition*. Available at: http://www.dh.gov.uk/publications (Accessed: 13 February 2019).
67. Department of Health and Social Care (2020) *Busting Bureaucracy Empowering frontline staff by reducing excess bureaucracy in the health and care system in England*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/944045/20112020_Busting_Bureaucracy_FINAL_PDF_VERSION_.pdf (Accessed: 19 March 2021).
68. Department of Health and Social Care (2021a) *Data saves lives: reshaping health and social care with data (draft) - GOV.UK*. Available at: https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data-draft/data-saves-lives-reshaping-health-and-social-care-with-data-draft#supporting-local-and-national-decision-makers-with-data (Accessed: 19 July 2021).
69. Department of Health and Social Care (2021b) *Integration and Innovation: working together to improve health and social care for all*. Available at: https://www.gov.uk/government/publications/working-together-to-improve-health-and-social-care-for-all/integration-and-innovation-working-together-to-improve-health-and-social-care-for-all-html-version (Accessed: 17 March 2021).
70. Department of Health and Social Care (2021c) *New review into use of health data for research and analysis*. Available at: https://www.gov.uk/government/news/new-review-into-use-of-health-data-for-research-and-analysis (Accessed: 19 March 2021).
71. Doerken, S. *et al.* (2019) 'Penalized logistic regression with low prevalence exposures beyond high dimensional settings', *PLoS ONE*, 14(5), p. e0217057. doi: 10.1371/journal.pone.0217057.
72. Doessing, A. and Burau, V. (2015) 'Care coordination of multimorbidity: a scoping study.', *Journal of comorbidity*, 5, pp. 15–28. doi: 10.15256/joc.2015.5.39.
73. Dorrington, S. *et al.* (2020) 'Demographic variation in fit note receipt and long-term conditions in south London', *Occupational and Environmental Medicine*, 77(6), pp. 418–426. doi: 10.1136/oemed-2019-106035.
74. Duan-Porter, W. *et al.* (2018) 'Evidence Review—Social Determinants of Health for Veterans', *Journal of General Internal Medicine*, pp. 1785–1795. doi: 10.1007/s11606-018-4566-8.
75. Edwards, N. (2019) 'Lessons from integrated care pilots in England', *Journal of Integrated Care*, 28(1), pp. 2–6. doi: 10.1108/JICA-10-2019-0044.
76. Ellaway, A. and Macintyre, S. (1998) 'Does housing tenure predict health in the UK because it exposes people to different levels of housing related hazards in the home or its surroundings?', *Health & Place*, 4(2), pp. 141–150. doi: 10.1016/S1353-8292(98)00006-9.

77. Erens, B. *et al.* (2019) 'Can health and social care integration make long-term progress? Findings from key informant surveys of the integration Pioneers in England', *Journal of Integrated Care*, 28(1), pp. 14–26. doi: 10.1108/JICA-05-2019-0020.
78. Europe R.A.N.D. (2012) *National Evaluation of the Department of Health's Integrated Care Pilots*, *Cambridge: RAND Corporation*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215103/dh_133127.pdf (Accessed: 13 April 2021).
79. European Commission (2019) *State of Health in the EU .United Kingdom. The Country Health Profile series.* Available at: http://www.who.int/bulletin/disclaimer/en/ (Accessed: 14 April 2021).
80. Eurostat (2020) *Archive:Housing statistics - Statistics Explained*. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Housing_statistics&oldid=497245#Type_of_dwelling (Accessed: 6 May 2021).
81. Feinstein, A. R. (1970) 'The pre-therapeutic classification of co-morbidity in chronic disease', *Journal of Chronic Diseases*, 23(7), pp. 455–468. doi: 10.1016/0021-9681(70)90054-8.
82. Ferry, F. R. *et al.* (2020) 'Multimorbidity among persons aged 25–64 years: a population-based study of social determinants and all-cause mortality', *Journal of Public Health*. doi: 10.1093/pubmed/fdaa209.
83. Finch, W. H., Bolin, J. and Kelley, K. (2014) *Multilevel Modeling Using R*. Second Edi. Taylor & Francis Group, LLC.
84. Foguet-Boreu, Q. *et al.* (2014) 'Impact of multimorbidity: Acute morbidity, area of residency and use of health services across the life span in a region of south Europe', *BMC Family Practice*, 15(1), p. 55. doi: 10.1186/1471-2296-15-55.
85. Fontana, G. *et al.* (2020) 'Ensuring that the NHS realises fair financial value from its data', *The Lancet Digital Health*. Elsevier Ltd, pp. e10–e12. doi: 10.1016/S2589-7500(19)30225-0.
86. Ford, J. C. and Ford, J. A. (2018) 'Multimorbidity: Will it stand the test of time?', *Age and Ageing*, 47(1), pp. 6–8. doi: 10.1093/ageing/afx159.
87. Fortin, M. *et al.* (2012) 'A systematic review of prevalence studies on multimorbidity: Toward a more uniform methodology', *Annals of Family Medicine*, 10(2), pp. 142–151. doi: 10.1370/afm.1337.
88. Fowler, T. *et al.* (2015) 'Excess winter deaths in Europe: A multi-country descriptive analysis', *European Journal of Public Health*, 25(2), pp. 339–345. doi: 10.1093/eurpub/cku073.
89. France, E. F. *et al.* (2012) 'Multimorbidity in primary care: A systematic review of prospective cohort studies', *British Journal of General Practice*, (April), pp. e297-307. doi: 10.3399/bjgp12X636146.e297.
90. Freedman, D. A. (1999) 'Ecological Inference and the Ecological Fallacy', *International Encyclopedia of the Social & Behavioral Sciences*, 6(4027–4030), pp. 1–7. Available at: http://statistics.berkeley.edu/tech-reports/549.pdf (Accessed: 1 July 2021).
91. Fulop, N., Mowlem, A. and Edwards, N. (2005) *Building integrated care: Lessons from the UK and Elsewhere*, *The NHS Confederation: London*. Available at: https://www.semanticscholar.org/paper/Building-integrated-care-Lessons-from-the-UK-and-of/0802a075618541d1293351d56b0d8bc2d7032862 (Accessed: 18 March 2021).
92. Gale, N. K. *et al.* (2013) *Using the framework method for the analysis of qualitative data in multi-disciplinary health research*, *BMC Medical Research*

Methodology. doi: 10.1186/1471-2288-13-117.
93. Garin, N. *et al.* (2016) 'Global Multimorbidity Patterns: A Cross-Sectional, Population-Based, Multi-Country Study', *J Gerontol A Biol Sci Med Sci*, 71(2), pp. 205–214. doi: 10.1093/gerona/glv128.
94. GeoPlace LLP (2020) *Central government and the NHS must now use UPRNs and USRNs to unlock the power of place*. Available at: https://www.geoplace.co.uk/press/2020/central-government-and-the-nhs-must-now-use-uprns-and-usrns-to-unlock-the-power-of-place (Accessed: 11 October 2021).
95. Gibson, M. *et al.* (2011) 'Housing and health inequalities: A synthesis of systematic reviews of interventions aimed at different pathways linking housing and health', *Health and Place*, 17(1), pp. 175–184. doi: 10.1016/j.healthplace.2010.09.011.
96. Glynn, L. *et al.* (2011) 'The prevalence of multimorbidity in primary care and its effect on health care utilization and cost', *Family Practice*, 28(5), pp. 516–523. doi: 10.1093/fampra/cmr013.
97. GoInvo (2018) *Determinants of Health Visualized*. Available at: https://www.goinvo.com/vision/determinants-of-health/ (Accessed: 28 September 2021).
98. Goldstein, H., Browne, W. and Rasbash, J. (2002) 'Multilevel modelling of medical data', *Statistics in Medicine*, 21(21), pp. 3291–3315. doi: 10.1002/sim.1264.
99. Graham, H. (2004) 'Social determinants and their unequal distribution: Clarifying policy understandings', *Milbank Quarterly*, 82(1), pp. 101–124. doi: 10.1111/j.0887-378X.2004.00303.x.
100. Greenhalgh, T. and Peacock, R. (2005) 'Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources', *British Medical Journal*, pp. 1064–1065. doi: 10.1136/bmj.38636.593461.68.
101. Groves, P. *et al.* (2013) *The 'big data' revolution in healthcare*. Available at: https://www.ghdonline.org/uploads/Big_Data_Revolution_in_health_care_2013_McKinsey_Report.pdf (Accessed: 27 April 2021).
102. Guy's and St Thomas' Charity (2018) *From One to Many: Exploring people's progression to multiple long-term health conditions in an urban environment*. Available at: https://www.gsttcharity.org.uk/sites/default/files/GSTTC_MLTC_Report_2018.pdf.
103. Hamnett, C. (2004) *Unequal City: London in the Global Arena*. Taylor & Francis Ltd. doi: 10.4324/9780203580240.
104. Hankivsky, O. and Christoffersen, A. (2008) 'Intersectionality and the determinants of health: A Canadian perspective', *Critical Public Health*, 18(3), pp. 271–283. doi: 10.1080/09581590802294296.
105. Hanlon, P. *et al.* (2018) 'Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants', *The Lancet Public Health*, 3(7), pp. e323–e332. doi: 10.1016/S2468-2667(18)30091-4.
106. Harper, G. and Mayhew, L. (2012) 'Using Administrative Data to Count an Local Applications', *Applied Spatial Analysis and Policy*, 5(4), pp. 97–122. doi: 10.1007/s12061-015-9162-2.
107. Harrison, C. *et al.* (2014) 'Examining different measures of multimorbidity, using a large prospective cross-sectional study in Australian

general practice.', *BMJ Open*, 4(7), p. e004694. doi: 10.1136/bmjopen-2013-004694.
108. Harron, K. *et al.* (2017) 'Challenges in administrative data linkage for research', *Big Data and Society*, 4(2), pp. 1–12. doi: 10.1177/2053951717745678.
109. Hastings, A. *et al.* (2015) *The Cost of Cuts: The Impact on Local Government and Poorer Communities*, *Joseph Rowntree Foundation*. Available at: https://www.jrf.org.uk/report/cost-cuts-impact-local-government-and-poorer-communities (Accessed: 4 May 2021).
110. Hayek, S. *et al.* (2017) 'Prevalence, correlates, and time trends of multiple chronic conditions among Israeli adults: Estimates from the Israeli National Health Interview Survey, 2014-2015', *Preventing Chronic Disease*, 14(8), p. E64. doi: 10.5888/pcd14.170038.
111. Head, A., Fleming, K., Kypridemos, C., Schofield, P., *et al.* (2021) 'Inequalities in incident and prevalent multimorbidity in England, 2004–19: a population-based, descriptive study', *The Lancet Healthy Longevity*. doi: 10.1016/S2666-7568(21)00146-X.
112. Head, A., Fleming, K., Kypridemos, C., Pearson-Stuttard, J., *et al.* (2021) 'Multimorbidity: The case for prevention', *Journal of Epidemiology and Community Health*, 75(3), pp. 242–244. doi: 10.1136/jech-2020-214301.
113. Healy, G., Bradley, H. and Forson, C. (2011) 'Intersectional sensibilities in analysing inequality regimes in public sector organizations', *Gender, Work and Organization*, 18(5), pp. 467–487. doi: 10.1111/j.1468-0432.2011.00557.x.
114. Hellström, Y. and Hallberg, I. R. (2004) 'Determinants and characteristics of help provision for elderly people living at home and in relation to quality of life', *Scandinavian Journal of Caring Sciences*, 18(4), pp. 387–395. doi: 10.1111/j.1471-6712.2004.00291.x.
115. Henchoz, Y. *et al.* (2019) 'Childhood adversity: A gateway to multimorbidity in older age?', *Archives of Gerontology and Geriatrics*, 80, pp. 31–37. doi: 10.1016/j.archger.2018.10.003.
116. Henderson, D. A. G. *et al.* (2021) 'Linkage of national health and social care data: a cross-sectional study of multimorbidity and social care use in people aged over 65 years in Scotland', *Age and Ageing*, 50, pp. 176–182. doi: 10.1093/ageing/afaa134.
117. Hennick, M., Hutter, I. and Bailey, A. (2020) 'Qualitative Research Methods', in. SAGE. Available at: https://books.google.co.uk/books?hl=en&lr=&id=_InCDwAAQBAJ&oi=fnd&pg=PP1&dq=(Hennink,+Hutter,+%26+Bailey,+2010)&ots=3u9KmWo3jy&sig=cSA9LTw8Z9Sje5tA1WLTdBWNUWo#v=onepage&q=(Hennink%2C Hutter%2C %26 Bailey%2C 2010)&f=false (Accessed: 20 May 2021).
118. Hiscock, R. *et al.* (2001) 'Ontological security and psycho-social benefits from the home: Qualitative evidence on issues of tenure', *Housing, Theory and Society*, 18(1–2), pp. 50–66. doi: 10.1080/14036090120617.
119. HM Government (2010) *Healthy Lives, Healthy People: Our strategy for public health in England*. Available at: http://www.official-documents.gov.uk/ (Accessed: 13 April 2021).
120. Ho, I. S.-S. *et al.* (2021) 'Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies', *The Lancet Public Health*, 6(8), pp. e587–e597. doi: 10.1016/s2468-2667(21)00107-9.
121. Hossain, M. P. *et al.* (2012) 'Social deprivation and prevalence of chronic kidney disease in the UK: workload implications for primary care',

QJM: An International Journal of Medicine, 105(2), pp. 167–175. doi: 10.1093/QJMED/HCR153.
122. HSR UK (2017) *Our hubs*. Available at: https://www.afklcargo.com/US/en/common/about_us/our_hubs.jsp (Accessed: 2 March 2021).
123. Hulks, S. *et al.* (2017) *Leading across the health and care system: Lessons from experience*, *The King's Fund*. Available at: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/Leading across the health and care system.pdf (Accessed: 3 March 2021).
124. Humphreys, J. *et al.* (2018) 'Early-life predictors of future multimorbidity: results from the Hertfordshire Cohort.', *Age and ageing*, 47(3), pp. 474–478. doi: 10.1093/ageing/afy005.
125. Illich, I. (2003) 'Medical nemesis', *Journal of Epidemiology & Community Health*, 57(12), pp. 919–922. doi: 10.1136/JECH.57.12.919.
126. Ingram, E. *et al.* (2021) 'Household and area-level social determinants of multimorbidity: a systematic review', *Journal of Epidemiology and Community Health*, pp. 232–241. doi: 10.1136/jech-2020-214691.
127. James Lind Alliance (2018) *Multiple Conditions in Later Life*. Available at: https://www.jla.nihr.ac.uk/priority-setting-partnerships/health-with-multiple-conditions-in-old-age/downloads/Multiple-Conditions-in-Later-Life-PSP-Top-10-report.pdf (Accessed: 26 April 2021).
128. Jamshed, S. (2014) 'Qualitative research method-interviewing and observation', *Journal of Basic and Clinical Pharmacy*, 5(4), p. 87. doi: 10.4103/0976-0105.141942.
129. Johnson-Lawrence, V., Zajacova, A. and Sneed, R. (2017) 'Education, race/ethnicity, and multimorbidity among adults aged 30–64 in the National Health Interview Survey', *SSM - Population Health*, 3, pp. 366–372. doi: 10.1016/j.ssmph.2017.03.007.
130. Johnston, M. C., Crilly, M., *et al.* (2019) 'Defining and measuring multimorbidity: A systematic review of systematic reviews', *European Journal of Public Health*, 29(1), pp. 182–189. doi: 10.1093/eurpub/cky098.
131. Johnston, M. C., Black, C., *et al.* (2019) 'Impact of educational attainment on the association between social class at birth and multimorbidity in middle age in the Aberdeen Children of the 1950s cohort study', *BMJ Open*, 9(1), p. e024048. doi: 10.1136/bmjopen-2018-024048.
132. Jones, J. and Barry, M. M. (2011) 'Exploring the relationship between synergy and partnership functioning factors in health promotion partnerships', *Health Promotion International*, 26(4), pp. 408–420. doi: 10.1093/heapro/dar002.
133. Kasteridis, P. *et al.* (2014) *The Importance of Multimorbidity in Explaining Utilisation and Costs Across Health and Social Care Settings: Evidence from South Somerset's Symphony Project*. Available at: http://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP96_multimorbidity_utilisation_costs_health_social care.pdf.
134. Kasteridis, P. *et al.* (2015) 'Who would most benefit from improved integrated care? Implementing an analytical strategy in South Somerset.', *International Journal of Integrated Care*, 15(28), pp. 1–15.
135. Katikireddi, S. V. *et al.* (2017) 'The contribution of risk factors to socioeconomic inequalities in multimorbidity across the lifecourse: a longitudinal analysis of the Twenty-07 cohort', *BMC Medicine*, 15(1), p. 152. doi: 10.1186/s12916-017-0913-6.
136. Keats, M. R. *et al.* (2020) 'Associations between neighborhood

walkability, physical activity, and chronic disease in Nova Scotian adults: An Atlantic path cohort study', *International Journal of Environmental Research and Public Health*, 17(22), pp. 1–16. doi: 10.3390/ijerph17228643.
137. Kernick, D., Chew-Graham, C. A. and O'Flynn, N. (2017) 'Clinical assessment and management of multimorbidity: NICE guideline', *The British Journal of General Practice*, 67(658), p. 235. doi: 10.3399/BJGP17X690857.
138. Khanolkar, A. R. *et al.* (2020) 'Socioeconomic inequalities in prevalence and development of multimorbidity across adulthood: Findings from the MRC 1946 national survey of health & development', *medRxiv*, pp. 1–21. doi: 10.1101/2020.08.14.20174037.
139. Ki, M. *et al.* (2017) 'Socioeconomic inequalities in health in the context of multimorbidity: A Korean panel study', *PLOS ONE*, 12(3), p. e0173770. doi: 10.1371/journal.pone.0173770.
140. Kim, J., Keshavjee, S. and Atun, R. (2020) 'Trends, patterns and health consequences of multimorbidity among South Korea adults: Analysis of nationally representative survey data 2007-2016', *Journal of Global Health*, 10(2), pp. 020426–020426. doi: 10.7189/JOGH.10.020426.
141. King's Fund (2018) *Making sense of integrated care systems, integrated care partnerships and accountable care organisations in the NHS in England | The King's Fund*. Available at: https://www.kingsfund.org.uk/publications/making-sense-integrated-care-systems (Accessed: 17 May 2019).
142. Kingston, A. *et al.* (2018) 'Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model', *Age and Ageing*, 47, pp. 374–380. doi: 10.1093/ageing/afx201.
143. Kline, R. (2014) *The 'snowy white peaks' of the NHS: a survey of discrimination in governance and leadership and the potential impact on patient care in London and England*. Available at: https://www.mdx.ac.uk/__data/assets/pdf_file/0015/50190/The-snowy-white-peaks-of-the-NHS.pdf.pdf (Accessed: 16 October 2020).
144. Kluge (2000) *Empirically Grounded Construction of Types and Typologies in Qualitative Social Research*. Available at: https://www.qualitative-research.net/index.php/fqs/article/view/1124/2499 (Accessed: 24 September 2020).
145. Kneale, D. *et al.* (2017) 'The use of evidence in English local public health decision-making: a systematic scoping review', *Implementation Science*, 12(1), p. 53. doi: 10.1186/s13012-017-0577-9.
146. Knowles, S. E. *et al.* (2013) 'Better together? A naturalistic qualitative study of inter-professional working in collaborative care for co-morbid depression and physical health problems', *Implementation Science*, 8(1), pp. 1–9. doi: 10.1186/1748-5908-8-110.
147. Kodner, D. L. and Spreeuwenberg, C. (2002) 'Integrated care: meaning, logic, applications, and implications – a discussion paper', *International Journal of Integrated Care*, 2(4). doi: 10.5334/ijic.67.
148. Kone, A. P. *et al.* (2021) 'Rising burden of multimorbidity and related socio-demographic factors: a repeated cross-sectional study of Ontarians', *Canadian Journal of Public Health*, 112(4), pp. 737–747. doi: 10.17269/s41997-021-00474-y.
149. Kosinski, M. *et al.* (2015) 'Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines', *American Psychologist*, 70(6), pp. 543–556. doi:

10.1037/a0039210.
150. Kosinski, M., Stillwell, D. and Graepel, T. (2013) 'Private traits and attributes are predictable from digital records of human behavior', *Proceedings of the National Academy of Sciences of the United States of America*, 110(15), pp. 5802–5805. doi: 10.1073/pnas.1218772110.
151. Kunna, R., San Sebastian, M. and Stewart Williams, J. (2017) 'Measurement and decomposition of socioeconomic inequality in single and multimorbidity in older adults in China and Ghana: Results from the WHO study on global AGEing and adult health (SAGE)', *International Journal for Equity in Health*, 16(1), p. 79. doi: 10.1186/s12939-017-0578-y.
152. Kuper, A., Lingard, L. and Levinson, W. (2008) 'Critically appraising qualitative research', *BMJ*, 337(7671), pp. 687–689. doi: 10.1136/bmj.a1035.
153. Lackner, J. M. *et al.* (2013) 'Type, Rather Than Number, of Mental and Physical Comorbidities Increases the Severity of Symptoms in Patients With Irritable Bowel Syndrome', *Clinical Gastroenterology and Hepatology*, 11(9), pp. 1147–1157. doi: 10.1016/j.cgh.2013.03.011.
154. Laires, P. A. and Perelman, J. (2019) 'The current and projected burden of multimorbidity: a cross-sectional study in a Southern Europe population', *European Journal of Ageing*, 16(2), pp. 181–192. doi: 10.1007/s10433-018-0485-0.
155. Lawton, J. *et al.* (2005) 'Diabetes service provision: A qualitative study of newly diagnosed Type 2 diabetes patients' experiences and views', *Diabetic Medicine*, 22(9), pp. 1246–1251. doi: 10.1111/j.1464-5491.2005.01619.x.
156. Lebenbaum, M. *et al.* (2018) 'Trends in obesity and multimorbidity in Canada', *Preventive Medicine*, 116, pp. 173–179. doi: 10.1016/j.ypmed.2018.08.025.
157. Leichsenring, K. (2004) 'Developing integrated health and social care services for older persons in Europe', *International Journal of Integrated Care*, 4(3). doi: 10.5334/ijic.107.
158. Li, J. *et al.* (2016) 'Patterns of multimorbidity and their association with health outcomes within Yorkshire, England: Baseline results from the Yorkshire Health Study', *BMC Public Health*, 16(1), p. 649. doi: 10.1186/s12889-016-3335-z.
159. Livingston, M., Kearns, A. and Bailey, N. (2013) 'Delivering Mixed Communities: The Relationship between Housing Tenure Mix and Social Mix in England's Neighbourhoods', *Housing Studies*, 28(7), pp. 1056–1080. doi: 10.1080/02673037.2013.812723.
160. Local Government Association (2021) *LGA briefing: Health and Social Care Bill White Paper*. Available at: https://www.local.gov.uk/parliament/briefings-and-responses/lga-briefing-health-and-social-care-bill-white-paper (Accessed: 19 March 2021).
161. London Borough of Barking and Dagenham (2018) *Joint strategic needs assessment*. Available at: https://www.lbbd.gov.uk/joint-strategic-needs-assessment-jsna (Accessed: 7 November 2019).
162. Low, L. L. *et al.* (2019) 'Epidemiologic Characteristics of Multimorbidity and Sociodemographic Factors Associated with Multimorbidity in a Rapidly Aging Asian Country', *JAMA Network Open*, 2(11). doi: 10.1001/jamanetworkopen.2019.15245.
163. Lujic, S. *et al.* (2017) 'Multimorbidity in Australia: Comparing estimates derived using administrative data sources and survey data', *PLoS ONE*, 12(8), p. e0183817. doi: 10.1371/journal.pone.0183817.
164. Lyons, J. *et al.* (2021) 'Protocol for the development of the Wales

Multimorbidity e-Cohort (WMC): Data sources and methods to construct a population-based research platform to investigate multimorbidity', *BMJ Open*, 11(1). doi: 10.1136/bmjopen-2020-047101.
165. Lyratzopoulos, G. *et al.* (2006) 'Deprivation and trends in blood pressure, cholesterol, body mass index and smoking among participants of a UK primary care-based cardiovascular risk factor screening programme: both narrowing and widening in cardiovascular risk factor inequalities', *Heart*, 92(9), pp. 1198–1206. doi: 10.1136/HRT.2005.081042.
166. Macintyre, S. *et al.* (2003) 'What features of the home and the area might help to explain observed relationships between housing tenure and health? Evidence from the west of Scotland', *Health and Place*, 9(3), pp. 207–218. doi: 10.1016/S1353-8292(02)00040-0.
167. Malterud, K. (2001) 'Qualitative research: Standards, challenges, and guidelines', *Lancet*, pp. 483–488. doi: 10.1016/S0140-6736(01)05627-6.
168. Marengoni, A. *et al.* (2011) 'Aging with multimorbidity: A systematic review of the literature', *Ageing Research Reviews*, 10(4), pp. 430–439. doi: 10.1016/j.arr.2011.03.003.
169. Marmot, M. (2005) 'Social determinants of health inequalities', *Lancet*, 365, pp. 1099–104. doi: 10.1007/978-3-319-71812-5_36.
170. Marmot, M. (2010) *Fair Society, Healthy Lives*, *The Marmot Review*. Available at: http://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review/fair-society-healthy-lives-full-report-pdf.pdf (Accessed: 14 December 2018).
171. Marmot, M. *et al.* (2020) *Health equity in England: The Marmot review 10 years on*. doi: 10.1136/bmj.m693.
172. Marmot, M. and Bell, R. (2012) 'Fair society, healthy lives', *Public Health*, 126, pp. S4–S10. doi: 10.1016/J.PUHE.2012.05.014.
173. Marshall, D. A. *et al.* (2016) 'Transforming Healthcare Delivery: Integrating Dynamic Simulation Modelling and Big Data in Health Economics and Outcomes Research', *PharmacoEconomics*, 34(2), pp. 115–126. doi: 10.1007/s40273-015-0330-7.
174. Maruthappu, M. and Keogh, B. (2015) 'The future of health system leadership', *The Lancet*, 385(9985), pp. 2325–2326. doi: 10.1016/S0140-6736(15)60990-4.
175. Mathur, R. *et al.* (2014) 'Completeness and usability of ethnicity data in UK-based primary care and hospital databases', *Journal of Public Health*, 36(4), pp. 684–692. doi: 10.1093/pubmed/fdt116.
176. Mays, N. and Pope, C. (1995) 'Rigour and qualitative research', *British Medical Journal*, 311(6997), pp. 109–112. doi: 10.1136/bmj.311.6997.109.
177. Mbuya-Bienge, C. *et al.* (2021) 'Does socio-economic status influence the effect of multimorbidity on the frequent use of ambulatory care services in a universal healthcare system? A population-based cohort study', *BMC Health Services Research*, 21(1), p. 202. doi: 10.1186/s12913-021-06194-w.
178. Mcdonald, H. I. *et al.* (2016) 'Methodological challenges when carrying out research on CKD and AKI using routine electronic health records', *Kidney International*, 90, pp. 943–949. doi: 10.1016/j.kint.2016.04.010.
179. McGinnis, J. M., Williams-Russo, P. and Knickman, J. R. (2002) 'The case for more active policy attention to health promotion', *Health Affairs*, 21(2), pp. 78–93. doi: 10.1377/hlthaff.21.2.78.
180. McHugh, M. L. (2012) 'Interrater reliability: The kappa statistic', *Biochemia Medica*, 22(3), pp. 276–282. doi: 10.11613/bm.2012.031.
181. McLean, G. *et al.* (2014) 'The influence of socioeconomic deprivation

on multimorbidity at different ages: A cross-sectional study', *British Journal of General Practice*, 64(624), pp. 440–447. doi: 10.3399/bjgp14X680545.
182. McNeill, T. (2010) *Family as a social determinant of health: implications for governments and institutions to promote the health and well-being of families.*, *Healthcare Quarterly*. doi: 10.12927/hcq.2010.21984.
183. Melis, R. *et al.* (2014) 'Incidence and predictors of multimorbidity in the elderly: a population-based longitudinal study.', *PloS one*, 9(7), p. e103120. doi: 10.1371/journal.pone.0103120.
184. Mercer, S., Salisbury, C. and Fortin, M. (2014) *ABC of Multimorbidity*. Wiley Blackwell BMJ Books. Available at: http://as.wiley.com/WileyCDA/WileyTitle/productCd-EHEP003257.html (Accessed: 9 September 2019).
185. Mercer, S. W. and Watt, G. (2007) 'The inverse care law: Clinical primary care encounters in deprived and affluent areas of Scotland', *Annals of Family Medicine*, 5(6), pp. 503–510. doi: 10.1370/afm.778.
186. Moher, D. *et al.* (2009) 'Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement', *PLoS Medicine*, p. e1000097. doi: 10.1371/journal.pmed.1000097.
187. Mokraoui, N. M. *et al.* (2016) 'Prevalence of self-reported multimorbidity in the general population and in primary care practices: A cross-sectional study', *BMC Research Notes*, 9(1), p. 314. doi: 10.1186/s13104-016-2121-4.
188. Moore, K. L. *et al.* (2014) 'Patterns of chronic co-morbid medical conditions in older residents of U.S. nursing homes: Differences between the sexes and across the agespan', *Journal of Nutrition, Health and Aging*, 18(4), pp. 429–436. doi: 10.1007/s12603-014-0001-y.
189. Mori, T. *et al.* (2019) 'The associations of multimorbidity with the sum of annual medical and long-term care expenditures in Japan', *BMC Geriatrics*, 19(1). doi: 10.1186/s12877-019-1057-7.
190. Morris, R. L. *et al.* (2011) 'Shifting priorities in multimorbidity: A longitudinal qualitative study of patient's prioritization of multiple conditions', *Chronic Illness*, 7(2), pp. 147–161. doi: 10.1177/1742395310393365.
191. Mounce, L. T. A. *et al.* (2018) 'Predicting incident multimorbidity', *Annals of Family Medicine*, 16(4), pp. 322–329. doi: 10.1370/afm.2271.
192. Naderifar, M., Goli, H. and Ghaljaie, F. (2017) 'Snowball Sampling: A Purposeful Method of Sampling in Qualitative Research', *Strides in Development of Medical Education*, 14(3), p. 67670. doi: 10.5812/sdme.67670.
193. Nagata, J. M. *et al.* (2013) 'Social determinants of health and seasonal influenza vaccination in adults ≥65 years: A systematic review of qualitative and quantitative data', *BMC Public Health*, p. 388. doi: 10.1186/1471-2458-13-388.
194. National Voices (2019) *Multiple Long-Term Conditions. Exploring the priorities of people working age: workshop report for NIHR*. Available at: https://www.nationalvoices.org.uk/publications/our-publications/multiple-long-term-conditions (Accessed: 26 April 2021).
195. Naylor, C. *et al.* (2012) *Long-term condition and mental health*, *The King's Fund*. Available at: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/long-term-conditions-mental-health-cost-comorbidities-naylor-feb12.pdf (Accessed: 26 April 2021).
196. Naylor, C. *et al.* (2016) *Bringing together physical and mental health:*

A new frontier for integrated care, *The King's Fund*. Available at: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/Bringing-together-Kings-Fund-March-2016_1.pdf (Accessed: 3 June 2019).
197. Newman, D., Levine, E. and Kishore, S. P. (2019) 'Prevalence of multiple chronic conditions in New York State, 2011–2016', *PLoS ONE*, 14(2). doi: 10.1371/journal.pone.0211965.
198. NHS (2014) *FIVE YEAR FORWARD VIEW*. Available at: https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf (Accessed: 26 April 2019).
199. NHS (2019) *The NHS Long Term Plan*. Available at: www.longtermplan.nhs.uk (Accessed: 9 July 2019).
200. NHS England (2016) *Data Quality: Guidance for providers and commissioners*, *Data Services for Commissioners*. Available at: https://www.england.nhs.uk/wp-content/uploads/2016/03/local-commissioning-data-quality-guidance.pdf (Accessed: 3 March 2021).
201. NHS England (2017) *NEXT STEPS ON THE NHS FIVE YEAR FORWARD VIEW*. Available at: https://www.england.nhs.uk/wp-content/uploads/2017/03/NEXT-STEPS-ON-THE-NHS-FIVE-YEAR-FORWARD-VIEW.pdf (Accessed: 18 January 2019).
202. Nicholson, B. D. *et al.* (2019) 'Determinants and extent of weight recording in UK primary care: an analysis of 5 million adults' electronic health records from 2000 to 2017', *BMC Medicine 2019 17:1*, 17(1), pp. 1–11. doi: 10.1186/S12916-019-1446-Y.
203. Nicholson, K., Makovski, T. T., *et al.* (2019) 'Multimorbidity and comorbidity revisited: refining the concepts for international health research', *Journal of Clinical Epidemiology*, 105, pp. 142–146. doi: 10.1016/j.jclinepi.2018.09.008.
204. Nicholson, K., Terry, A. L., *et al.* (2019) 'Prevalence, characteristics, and patterns of patients with multimorbidity in primary care: A retrospective cohort analysis in Canada', *British Journal of General Practice*, 69(686), pp. E647–E656. doi: 10.3399/bjgp19X704657.
205. Nicholson, K., Makovski, T. T. and Stranges, S. (2019) 'The integration of primary care and public health to improve population health: tackling the complex issue of multimorbidity', *International Journal of Public Health*, pp. 983–984. doi: 10.1007/s00038-019-01278-1.
206. Nielsen, C. R., Halling, A. and Andersen-Ranberg, K. (2017) 'Disparities in multimorbidity across Europe – Findings from the SHARE Survey', *European Geriatric Medicine*, 8(1), pp. 16–21. doi: 10.1016/j.eurger.2016.11.010.
207. NIHR (2021) *Multiple long-term conditions (multimorbidity): making sense of the evidence*. doi: 10.3310/collection_45881.
208. Nolte, E. and McKee, M. (2008) *Caring for people with chronic conditions: a health system perspective*. Open University Press. Available at: https://books.google.co.uk/books?hl=en&lr=&id=FudEBgAAQBAJ&oi=fnd&pg=PP1&dq=Nolte+and+McKee,+2008&ots=406BMsgie3&sig=w2jeVuSdmrAgWD_tmbH8TDJTHSA (Accessed: 22 October 2019).
209. North London Partners in Health and Care (2021) *HealtheIntent*. Available at: https://www.northlondonpartners.org.uk/ourplan/Areas-of-work/Digital/health-information-exchange.htm (Accessed: 21 December 2021).
210. Northwood, M. *et al.* (2018) 'Integrative review of the social determinants of health in older adults with multimorbidity', *Journal of*

Advanced Nursing, 74(1), pp. 45–60. doi: 10.1111/jan.13408.
211. Nunes, B. P. *et al.* (2016) 'Multimorbidity and mortality in older adults: A systematic review and meta-analysis', *Archives of Gerontology and Geriatrics*, 67, pp. 130–138. doi: 10.1016/j.archger.2016.07.008.
212. O'Halloran, J., Miller, G. C. and Britt, H. (2004) 'Defining chronic conditions for primary care with ICPC-2', *Family Practice*, 21(4), pp. 381–386. doi: 10.1093/fampra/cmh407.
213. Office for National Statistics (2011a) *Census geography*. Available at: https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeography (Accessed: 13 September 2021).
214. Office for National Statistics (2011b) *List of ethnic groups*. Available at: https://www.ethnicity-facts-figures.service.gov.uk/style-guide/ethnic-groups (Accessed: 13 September 2021).
215. Office for National Statistics (2019a) *National Statistics - English Indices of Deprivation 2019*. Available at: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 (Accessed: 13 September 2021).
216. Office for National Statistics (2019b) *Population estimates for the UK, mid-2019: methods guide*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/methodologies/methodologyguideformid2015ukpopulationestimatesenglandandwalesjune2016 (Accessed: 16 September 2021).
217. Office for National Statistics (2019c) *Research report on population estimates by ethnic group and religion*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/researchreportonpopulationestimatesbyethnicgroupandreligion/2019-12-04.
218. Office for National Statistics (2020) *Home ownership - Ethnicity facts and figures*. Available at: https://www.ethnicity-facts-figures.service.gov.uk/housing/owning-and-renting/home-ownership/latest (Accessed: 20 September 2021).
219. Office of National Statistics (2020a) *England population mid-year estimate*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/timeseries/enpop/pop (Accessed: 14 April 2021).
220. Office of National Statistics (2020b) *Subnational population projections for England*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/subnationalpopulationprojectionsforenglan d/2018based (Accessed: 14 April 2021).
221. Orueta, J. F., Nuño-Solinís, R., *et al.* (2013) 'Prevalence of multimorbidity according to the deprivation level among the elderly in the Basque Country', *BMC Public Health*, 13(1), p. 918. doi: 10.1186/1471-2458-13-918.
222. Orueta, J. F., García-Álvarez, A., *et al.* (2013) 'Socioeconomic variation in the burden of chronic conditions and health care provision - Analyzing administrative individual level data from the Basque Country, Spain', *BMC Public Health*, 13(1), p. 870. doi: 10.1186/1471-2458-13-870.
223. Orueta, J. F. *et al.* (2014) 'Prevalence and costs of multimorbidity by deprivation levels in the Basque Country: A population based study using health administrative databases', *PLoS ONE*, 9(2), p. e89787. doi: 10.1371/journal.pone.0089787.

224. Palinkas, L. A. *et al.* (2015) 'Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research', *Administration and Policy in Mental Health and Mental Health Services Research*, 42(5), pp. 533–544. doi: 10.1007/s10488-013-0528-y.
225. Panagioti, M. *et al.* (2015) 'Multimorbidity and patient safety incidents in primary care: A systematic review and meta-analysis', *PLoS ONE*, 10(8), pp. 1–30. doi: 10.1371/journal.pone.0135947.
226. Van Panhuis, W. G. *et al.* (2014) 'A systematic review of barriers to data sharing in public health', *BMC Public Health*. doi: 10.1186/1471-2458-14-1144.
227. Pathirana, T. I. and Jackson, C. A. (2018) 'Socioeconomic status and multimorbidity: a systematic review and meta-analysis', *Australian and New Zealand Journal of Public Health*, 42(2), pp. 186–194. doi: 10.1111/1753-6405.12762.
228. Pati, S. *et al.* (2015) 'Prevalence, Correlates, and Outcomes of Multimorbidity Among Patients Attending Primary Care in Odisha, India', *Annals of family medicine*, 13(4). doi: 10.1370/afm.1843.
229. Patton, M. (1990) *Qualitative evaluation and research methods*. Beverly Hills, CA: Sage. Available at: https://legacy.oise.utoronto.ca/research/field-centres/ross/ctl1014/Patton1990.pdf (Accessed: 3 September 2019).
230. Payne, R. A. *et al.* (2020) 'Development and validation of the Cambridge Multimorbidity Score', *CMAJ*, 192(5), pp. E107–E114. doi: 10.1503/cmaj.190757.
231. Pearce, J. *et al.* (2007) 'Neighborhood Deprivation and Access to Fast-Food Retailing: A National Study', *American Journal of Preventive Medicine*, 32(5), pp. 375–382. doi: 10.1016/J.AMEPRE.2007.01.009.
232. Pearson-Stuttard, J., Ezzati, M. and Gregg, E. W. (2019) 'Multimorbidity—a defining challenge for health systems', *The Lancet Public Health*. Elsevier Ltd, pp. e599–e600. doi: 10.1016/S2468-2667(19)30222-1.
233. Prazeres, F. and Santiago, L. (2015) 'Prevalence of multimorbidity in the adult population attending primary care in Portugal: a cross-sectional study', *BMJ Open*, 5(e009287). Available at: https://bmjopen.bmj.com/content/bmjopen/5/9/e009287.full.pdf (Accessed: 11 September 2019).
234. Public Health England (2019) *Fingertips - Local Authority Health Profiles*. Available at: https://fingertips.phe.org.uk/profile/health-profiles/data#page/1/ati/102/are/E09000002 (Accessed: 14 September 2021).
235. Reeves, A. *et al.* (2016) 'Reductions in the United Kingdom's Government Housing Benefit and Symptoms of Depression in Low-Income Households', *American Journal of Epidemiology*, 184, pp. 421–9. doi: 10.1093/aje/kww055.
236. Le Reste, J. Y. *et al.* (2013) 'The European General Practice Research Network Presents a Comprehensive Definition of Multimorbidity in Family Medicine and Long Term Care, Following a Systematic Review of Relevant Literature', *Journal of the American Medical Directors Association*, 14(5), pp. 319–325. doi: 10.1016/j.jamda.2013.01.001.
237. Rijken, M. *et al.* (2013) *ICARE4EU: Improving Care for People with Multiple Chronic Conditions in Europe*, *Eurohealth*. Available at: https://apps.who.int/iris/handle/10665/333078 (Accessed: 14 April 2021).
238. Ritchie, D. *et al.* (2008) '"People pull the rug from under your feet":

Barriers to successful public health programmes', *BMC Public Health*, 8(1), p. 173. doi: 10.1186/1471-2458-8-173.
239. Roberts, K. C. *et al.* (2015) 'Prevalence and patterns of chronic disease multimorbidity and associated determinants in Canada', *Health Promotion and Chronic Disease Prevention in Canada*, 35(6), pp. 87–94. doi: 10.24095/hpcdp.35.6.01.
240. Rocca, W. A. *et al.* (2014) 'The prevalence of multimorbidity in a geographically defined American population: patterns by age, sex, and ethnicity', *Mayo Clin Proc.*, 89(10), pp. 1336–1349. doi: 10.1016/j.mayocp.2014.07.010.The.
241. Rogers, N. T. *et al.* (2021) 'Early-Life Socioeconomic Position and the Accumulation of Health-Related Deficits by Midlife in the 1958 British Birth Cohort Study', *American Journal of Epidemiology*, 00(00), pp. 1–11. doi: 10.1093/aje/kwab038.
242. Rolewicz, L. *et al.* (2020) 'Are the needs of people with multiple long-term conditions being met? Evidence from the 2018 General Practice Patient Survey', *BMJ Open*, 10(11). doi: 10.1136/bmjopen-2020-041569.
243. Rose, G. (1992) *The Strategy of Preventive Medicine*. Oxford University Press. Available at: https://books.google.co.uk/books?hl=en&lr=&id=BQDXQI8n7qQC&oi=fnd&pg=PR17&dq=the+strategy+of+preventive+medicine&ots=Rp4jJHVPeL&sig=Np8RweGL_berTqcwPiCDhiTaV5Y#v=onepage&q=the strategy of preventive medicine&f=false (Accessed: 24 October 2019).
244. Ryan, B. L. *et al.* (2018) 'Beyond the grey tsunami: a cross-sectional population-based study of multimorbidity in Ontario', *Canadian Journal of Public Health*, 109(5–6), pp. 845–854. doi: 10.17269/s41997-018-0103-0.
245. SA NT DataLink (2021) *What is Data Linkage?* Available at: https://www.unisa.edu.au/santdatalink/about-us/what-is-data-linkage/ (Accessed: 7 September 2021).
246. Salisbury, C. *et al.* (2011) 'Epidemiology and impact of multimorbidity in primary care: A retrospective cohort study', *British Journal of General Practice*, 61(582), pp. 12–21. doi: 10.3399/bjgp11X548929.
247. Sampson, H. (2004) 'Navigating the waves: The usefulness of a pilot in qualitative research', *Qualitative Research*, 4(3), pp. 383–402. doi: 10.1177/1468794104047236.
248. Saunders, B. *et al.* (2018) 'Saturation in qualitative research: exploring its conceptualization and operationalization', *Quality & Quantity*, 52, pp. 1893–1907. doi: 10.1007/s11135-017-0574-8.
249. Scanlon, K. and Kochan, B. (2011) *Towards a sustainable private rented sector. The lessons from other countries*. Available at: http://eprints.lse.ac.uk/56070/1/Towards_a_sustainable_private_rented_sector.pdf (Accessed: 29 June 2021).
250. Schäfer, I. *et al.* (2012) 'The influence of age, gender and socio-economic status on multimorbidity patterns in primary care. first results from the multicare cohort study', *BMC Health Services Research*, 12(1), p. 89. doi: 10.1186/1472-6963-12-89.
251. Scott, P., Emerson, K. and Henderson-Reay, T. (2021) 'Data saves lives', *The BMJ*, p. n1694. doi: 10.1136/bmj.n1694.
252. Scottish Centre for Adminstrative Data Research (2020) *A guide to CHI-UPRN Residential Linkage (CURL) File*. Available at: https://www.scadr.ac.uk/sites/default/files/CURLreport2311 - A guide to CHI-UPRN Residential Linkage.pdf (Accessed: 11 October 2021).

253. Shand, J. (2020) *Towards integrated care: using linked data to explore health and social care utilisation for adult residents of Barking and Dagenham in 2016/17*, *Doctoral Thesis*. UCL (University College London).
254. Shand, J., Morris, S. and Gomes, M. (2020) 'Understanding health and care expenditure by setting – who matters to whom?', *Journal of Health Services Research and Policy*, 26(2), pp. 77–84. doi: 10.1177/1355819620936721.
255. Shand, J. and Turner, S. (2019) 'System wide collaboration? Health and social care leaders' perspectives on working across boundaries', *Journal of Integrated Care*, 27(1), pp. 83–94. doi: 10.1108/JICA-06-2018-0042.
256. Shang, X. *et al.* (2020) 'Leading determinants for multimorbidity in middle-aged Australian men and women: A nine-year follow-up cohort study', *Preventive Medicine*, 141, p. 106260. doi: 10.1016/j.ypmed.2020.106260.
257. Shattell, M. and Thomas, M. (2005) 'An exploration of problematic interviewee behaviors in qualitative research', *Western Journal of Nursing Research*, 27, pp. 188–199. doi: 10.1177/0193945904268068.
258. Shaw, M. (2004) 'Housing and public health', *Annual Review of Public Health*, 25, pp. 397–418. doi: 10.1146/annurev.publhealth.25.101802.123036.
259. Shaw, S. E. *et al.* (2013) 'The work of commissioning: A multisite case study of healthcare commissioning in England's NHS', *BMJ Open*, 3(9), p. 3341. doi: 10.1136/bmjopen-2013-003341.
260. Singer, L., Green, M., Rowe, F., Ben-Shlomo, Y. and Morrissey, K. (2019) 'Social determinants of multimorbidity and multiple functional limitations among the ageing population of England, 2002-2015', *SSM - Population Health*, 8, p. 100413. doi: 10.1016/j.ssmph.2019.100413.
261. Singer, L., Green, M., Rowe, F., Ben-Shlomo, Y., Kulu, H., *et al.* (2019) 'Trends in multimorbidity, complex multimorbidity and multiple functional limitations in the ageing population of England, 2002–2015', *Journal of Comorbidity*, 9, pp. 1–9. doi: 10.1177/2235042x19872030.
262. Sinnott, C. *et al.* (2013) 'GPs' perspectives on the management of patients with multimorbidity: systematic review and synthesis of qualitative research', *BMJ Open*, 3(9), p. e003610. doi: 10.1136/bmjopen-2013-003610.
263. Sinnott, C. *et al.* (2015) 'Psychosocial complexity in multimorbidity: The legacy of adverse childhood experiences', *Family Practice*, 32(3), pp. 269–275. doi: 10.1093/fampra/cmv016.
264. Smith, J. and Firth, J. (2011) 'Qualitative data analysis: the framework approach', *Nurse Researcher*, pp. 52–65. Available at: http://eprints.hud.ac.uk/id/eprint/18884/1/RepositoryFrameworkNResearcher.pdf (Accessed: 6 June 2019).
265. Smith, S. M. *et al.* (2016) 'Interventions for improving outcomes in patients with multimorbidity in primary care and community settings', *Cochrane Database of Systematic Reviews*, (3). doi: 10.1002/14651858.CD006560.pub3.
266. Solar, O. and Irwin, A. (2013) *A Conceptual Framework for Action on the Social Determinants of Health*, *Social Determinants of Health Discussion Paper 2 (Policy and Practice)*. doi: ISBN 978 92 4 150085 2.
267. Sreedhar, S., Richard, L. and Stokes, T. (2019) 'Multimorbidity and multiple social disadvantage in a New Zealand high-needs free primary healthcare clinic population: A cross-sectional study', *New Zealand Medical Journal*, 132(1490), pp. 42–51. Available at: www.nzma.org.nz/journal (Accessed: 3 May 2021).

268. St John, P. D. *et al.* (2021) 'Multimorbidity in Canadians living in the community Results from the Canadian Longitudinal Study of Aging', *Canadian Family Physician*, 67(3), pp. 187–197. doi: 10.46747/cfp.6703187.
269. St Sauver, J. L. *et al.* (2015) 'Risk of developing multimorbidity across all ages in an historical cohort study: Differences by sex and ethnicity', *BMJ Open*, 5(2). doi: 10.1136/bmjopen-2014-006413.
270. Stafford, M. *et al.* (2018) *Briefing: Understanding the health care needs of people with multiple health conditions*, *The Health Foundation*. Available at: https://reader.health.org.uk/understanding-health-care-needs-people-multiple-health-conditions (Accessed: 22 October 2019).
271. Stanley, J. *et al.* (2018) 'Epidemiology of multimorbidity in New Zealand: a cross-sectional study using national-level hospital and pharmaceutical data.', *BMJ open*, 8(5), p. e021689. doi: 10.1136/bmjopen-2018-021689.
272. Sterne, J. *et al.* (2016) *Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I): detailed guidance*. Available at: http://www.riskofbias.info (Accessed: 28 October 2019).
273. Stokes, J. *et al.* (2021) 'Multimorbidity combinations, costs of hospital care and potentially preventable emergency admissions in England: A cohort study', *PLoS Medicine*, 18(1). doi: 10.1371/journal.pmed.1003514.
274. Stokes, T., Azam, M. and Noble, F. D. (2018) 'Multimorbidity in Māori and Pacific patients: Cross-sectional study in a Dunedin general practice', *Journal of Primary Health Care*, 10(1), pp. 39–43. doi: 10.1071/HC17046.
275. Symons, T. (2016) *Wise council: Insights from the cutting edge of data-driven local government*, *Local Government Association and Nesta*. Available at: http://www.nesta.org.uk/publications/wise-council-insights-cutting-edge-data-driven-local-government (Accessed: 27 April 2021).
276. Tacconelli, E. (2010) *Systematic reviews: CRD's guidance for undertaking reviews in health care*, *The Lancet Infectious Diseases*. doi: 10.1016/s1473-3099(10)70065-7.
277. Taggar, J. S. *et al.* (2012) 'The impact of the Quality and Outcomes Framework (QOF) on the recording of smoking targets in primary care medical records: Cross-sectional analyses from the Health Improvement Network (THIN) database', *BMC Public Health*, 12(1), pp. 1–11. doi: 10.1186/1471-2458-12-329.
278. Tariq, S. and Woodman, J. (2013) 'Using mixed methods in health research', *JRSM Short Reports*, 4(6), p. 204253331347919. doi: 10.1177/2042533313479197.
279. Taylor, A. W. *et al.* (2011) 'Multimorbidity: not just an older person's issue. Results from an Australian biomedical study.', *Social psychiatry and psychiatric epidemiology*, 46(4), p. 351. doi: 10.1007/s00127-010-0309-9.
280. The Academy of Medical Sciences (2018) *Multimorbidity: a priority for global health research*. Available at: https://acmedsci.ac.uk/file-download/82222577.
281. The Academy of Medical Sciences (2020) *Cross-funder multimorbidity research framework*. Available at: https://acmedsci.ac.uk/file-download/49628715 (Accessed: 23 April 2021).
282. The Health Foundation (2012) *Cross sector working to support large-scale change.* Available at: http://www.health.org.uk/publication/cross-sector-working-support-large-scale-change.
283. The King's Fund (2020a) *Social care services: funding cuts are biting hard*. Available at: https://www.kingsfund.org.uk/blog/2020/01/social-care-

funding-cuts-are-biting-hard (Accessed: 15 April 2021).
284. The King's Fund (2020b) *The NHS: how providers are regulated and commissioned*. Available at: https://www.kingsfund.org.uk/sites/default/files/2020-03/NHS_structure_March_2020.pdf (Accessed: 10 May 2021).
285. The Richmond Group of Charities (2018) *Multimorbidity: Understanding the challenge*. Available at: https://richmondgroupofcharities.org.uk/sites/default/files/multimorbidity_-_understanding_the_challenge.pdf (Accessed: 14 April 2021).
286. The World Bank (2020) *High income | Data*. Available at: https://data.worldbank.org/country/XD (Accessed: 6 October 2021).
287. Timmins, N. (2015) *The practice of system leadership Being comfortable with chaos*, *The King's Fund*. Available at: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/System-leadership-Kings-Fund-May-2015.pdf (Accessed: 24 May 2021).
288. Tomasdottir, M. O. *et al.* (2016) 'Does existential unease' predict adult multimorbidity? Analytical cohort study on embodiment based on the Norwegian HUNT population', *BMJ Open*, 6(11), p. e012602. doi: 10.1136/bmjopen-2016-012602.
289. Tonelli, M. *et al.* (2016) 'Methods for identifying 30 chronic conditions: Application to administrative data Healthcare Information Systems', *BMC Medical Informatics and Decision Making*, 15(1), p. 31. doi: 10.1186/s12911-015-0155-5.
290. Tran, J. *et al.* (2021) 'Multi-morbidity and blood pressure trajectories in hypertensive patients: A multiple landmark cohort study', *PLoS Medicine*, 18(6), p. e1003674. doi: 10.1371/journal.pmed.1003674.
291. Tucker-Seeley, R. D. *et al.* (2011) 'Lifecourse socioeconomic circumstances and multimorbidity among older adults', *BMC Public Health*, 11. doi: 10.1186/1471-2458-11-313.
292. Vaezghasemi, M. *et al.* (2016) 'Households, the omitted level in contextual analysis: Disentangling the relative influence of households and districts on the variation of BMI about two decades in Indonesia', *International Journal for Equity in Health*. doi: 10.1186/s12939-016-0388-7.
293. Valderas, J. *et al.* (2009) 'Defining comorbidity: implications for understanding health and health services', *Annals Of Family Medicine*, 7, pp. 357–363. doi: 10.1370/afm.983.Martin.
294. Verest, W. J. G. M. *et al.* (2019) 'Do ethnic inequalities in multimorbidity reflect ethnic differences in socioeconomic status? The HELIUS study', *European Journal of Public Health*. doi: 10.1093/eurpub/ckz012.
295. Vetter, D. (2019) *How Hong Kong's public housing system works: costs, waiting times and sales*, *South China Morning Post*. Available at: https://www.scmp.com/news/hong-kong/society/article/2182106/why-public-housing-shortfall-will-remain-thorn-hong-kongs (Accessed: 6 May 2021).
296. Violan, C. *et al.* (2014) 'Prevalence, determinants and patterns of multimorbidity in primary care: A systematic review of observational studies', *PLoS ONE*, 9(7), pp. 3–11. doi: 10.1371/journal.pone.0102149.
297. Violán, C. *et al.* (2014) 'Burden of multimorbidity, socioeconomic status and use of health services across stages of life in urban areas: A cross-sectional study', *BMC Public Health*, 14(1), p. 530. doi: 10.1186/1471-2458-14-530.
298. Viswanathan, M. *et al.* (2017) 'Assessing the Risk of Bias in Systematic Reviews of Health Care Interventions. Methods Guide for Comparative

Effectiveness Reviews.', *AHRQ Publication*, 17(18)-EHC. doi: 10.23970/AHRQEPCMETHGUIDE2.
299. Vos, T. *et al.* (2020) 'Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019', *Global Health Metrics*, 396, pp. 1135–59. doi: 10.1016/S0140-6736(20)30925-9.
300. Walker, A. E. (2007) 'Multiple chronic diseases and quality of life: Patterns emerging from a large national sample, Australia', *Chronic Illness*, 3(3), pp. 202–218. doi: 10.1177/1742395307081504.
301. Walker, J. L. *et al.* (2021) 'UK prevalence of underlying conditions which increase the risk of severe COVID-19 disease: a point prevalence study using electronic health records', *BMC Public Health 2021 21:1*, 21(1), pp. 1–14. doi: 10.1186/S12889-021-10427-2.
302. Walker, R. J. *et al.* (2014) 'Impact of social determinants of health on outcomes for type 2 diabetes: A systematic review', *Endocrine*, pp. 29–48. doi: 10.1007/s12020-014-0195-0.
303. Wang, S. B. *et al.* (2015) 'Prevalence and patterns of multimorbidity in northeastern China: A cross-sectional study', *Public Health*, 129(11), pp. 1539–1546. doi: 10.1016/j.puhe.2015.06.013.
304. Watt, G. (2008) 'What is the NHS for?', *British Journal of General Practice*, pp. 727–728. doi: 10.3399/bjgp08X342381.
305. Webber Phillips (2021) *Origins*. Available at: https://webberphillips.com/ (Accessed: 7 April 2021).
306. Wells, G. *et al.* (2019) *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (Accessed: 28 October 2019).
307. Whitehead, C. (2014) *Social Housing in Europe. Chapter 7: Social Housing in England*. Wiley Blackwell. Available at: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118412367#page=124 (Accessed: 4 September 2019).
308. Whitty, C. J. M. *et al.* (2020) 'Rising to the challenge of multimorbidity', *The BMJ*. doi: 10.1136/bmj.l6964.
309. Willadsen, T. *et al.* (2018) 'Multimorbidity and mortality: A 15-year longitudinal registry-based nationwide Danish population study', *Journal of Comorbidity*, 8(1). doi: 10.1177/2235042X18804063.
310. Willadsen, T. G. *et al.* (2016) 'The role of diseases, risk factors and symptoms in the definition of multimorbidity – a systematic review', *Scandinavian Journal of Primary Health Care*, 34(2), pp. 112–121. doi: 10.3109/02813432.2016.1153242.
311. Williams, J. S. and Egede, L. E. (2016) 'The Association Between Multimorbidity and Quality of Life, Health Status and Functional Disability', *American Journal of the Medical Sciences*, 352(1), pp. 45–52. doi: 10.1016/j.amjms.2016.03.004.
312. Williams, R. *et al.* (2020) 'Unacceptable failures: the final report of the Lancet Commission into liver disease in the UK', *The Lancet*, 395(10219), pp. 226–239. doi: 10.1016/S0140-6736(19)32908-3.
313. Wister, A. *et al.* (2020) 'Multimorbidity and Socioeconomic Deprivation among Older Adults: A Cross-sectional Analysis in Five Canadian Cities Using the CLSA', *Journal of Aging and Environment*, 34(4), pp. 435–454. doi: 10.1080/26892618.2020.1734138.
314. Wium-Andersen, M. K. *et al.* (2013) 'Elevated C-reactive protein levels,

psychological distress, and depression in 73131 individuals', *JAMA Psychiatry*, 70(2), pp. 176–184. doi: 10.1001/2013.jamapsychiatry.102.
315. Wolfinger, R. and O'Connell, M. (1993) 'Generalized Linear Mixed Models: A Pseudo-Likelihood Approach', *Journal of Statistical Computation and Simulation*, 48(3–4), pp. 233–243. doi: 10.1080/00949659308811554.
316. World Health Organization (2002) *Innovative Care for Chronic Conditions*. Available at: https://www.who.int/chp/knowledge/publications/icccglobalreport.pdf.
317. Wray, J., Archibong, U. and Walton, S. (2017) 'Why undertake a pilot in a qualitative PhD study? Lessons learned to promote success', *Nurse Researcher*, 24(3), pp. 31–35. doi: 10.7748/nr.2017.e1416.
318. Yin, R. K. (2003) *Case study research: design and methods*. Vol. 5. Sage.
319. Zaccai, J. H. (2004) 'How to assess epidemiological studies', *Postgraduate Medical Journal*, 80(941), pp. 140–147. doi: 10.1136/PGMJ.2003.012633.

Appendices

Appendix 1 MedLine search terms

1. MULTIMORBIDITY/

2. multimorbid*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

3. multi-morbid*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

4. "multiple morbidity".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

5. "intercurrent morbidity".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

6. "coexisting condition*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

7. "coexisting diagnos*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

8. "coexisting disease*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

9. "coexisting illness*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

10. "concurrent condition*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

11. "concurrent diagnos*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

12. "concurrent disease*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

13. "concurrent illness*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

14. "multiple chronic disease*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

15. "several chronic disease*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

16. "multiple chronic condition*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

17. "multiple chronic medical conditions".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

18. "several chronic condition*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

19. "multiple health problem*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

20. "several health problem*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

21. "one or more chronic conditions".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

22. "multiple illness*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

23. "multiple diagnos*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

24. "multiple disease*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

25. "multiple pathology*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

26. "multiple condition*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

27. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26

28. neighbourhood*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

29. neighborhood*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

30. area*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

31. region*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

32. small-area.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

33. context*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

34. household.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

35. residen*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

36. place*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

37. communit*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

38. "census tracts".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

39. municipalit*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

40. 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39

41. SOCIOECONOMIC FACTORS/

42. INCOME/

43. socioeconomic.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

44. socio-economic.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

45. "social capital".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

46. inequalit*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

47. disparit*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

48. inequit*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

49. income.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

50. wealth.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

51. "financial problem*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

52. "financial difficulties".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word,

protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

53. vocation*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

54. "training opportunit*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

55. "employment opportunit*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

56. "occupation* opportunit*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

57. "job opportunities".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

58. "job insecurit*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

59. "education* opportunit*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word,

protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

60. "education* achiev*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

61. "education* quality".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

62. fast-food.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

63. "fast food".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

64. "healthy food".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

65. "social determinant*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

66. "wider determinant*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word,

protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

67. 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66

68. 40 and 67

69. POVERTY AREAS/ or POVERTY/

70. Public Housing/

71. Residence Characteristics/

72. overcrowding.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

73. "public housing".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

74. "residence characteristics".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

75. "population density".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

76. neighbourhood.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

77. neighborhood.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

78. urban*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

79. rural.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

80. ghetto*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

81. slum*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

82. estate*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

83. "poverty area*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

84. "physical environment".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

85. "built environment".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

86. "living standard*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

87. "family structure".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

88. "family breakdown".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

89. "family disintegrat*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

90. "single parent*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

91. "housing tenure".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

92. Townsend.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

93. carstairs.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

94. "super profile*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

95. "Index of Multiple Deprivation".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

96. Nam-Powers.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

97. "Hollingshead Index".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word,

protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

98. "Breadline Britain Index".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

99. "inverse care law".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

100. depriv*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

101. impoverish*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

102. disadvantag*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

103. "sensitive population*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

104. "community cohesion".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word,

protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

105. "community connect*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

106. "community network*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

107. "community support".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

108. "transport* quality".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

109. "transport* services".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

110. walkability.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

111. "public space*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

112. "green space*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

113. "greenspace*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

114. "open space*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

115. recreation*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

116. facilit*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

117. leisure activit*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

118. "food outlet*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

119. "food environment".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

120. "air pollution".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

121. "particulate matter".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

122. "nitrogen dioxide".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

123. "sulphur dioxide".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

124. "sulfur dioxide".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

125. "carbon monoxide".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word,

protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

126. crime.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

127. 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100 or 101 or 102 or 103 or 104 or 105 or 106 or 107 or 108 or 109 or 110 or 111 or 112 or 113 or 114 or 115 or 116 or 117 or 118 or 119 or 120 or 121 or 122 or 123 or 124 or 125 or 126

128. 68 or 127

129. 27 and 128

130. limit 129 to (english language and yr="2010 -Current"

Appendix 2 Developing review search terms

Developing the multimorbidity search terms

One key decision I had to make when developing the search strategy for my systematic review was whether to include the term 'comorbidity' and its linguistic variations in my search terms. This was a pertinent issue as the terms 'multimorbidity' and 'comorbidity' are used interchangeably in the literature, however including the comorbidity terms returned an impractically high number of hits (Van den Akker *et al.*, 1998).

To make this decision, I used three systematic reviews that have investigated different aspects of multimorbidity and included search terms relating to comorbidity in their search strategies (Fortin *et al.*, 2012; France *et al.*, 2012; Pati *et al.*, 2015). These reviews were arbitrarily selected. I created tables that included details of the studies these reviews had identified and included (below). Two details of interest I noted were: 1) is the term 'multimorbidity' included in the title or abstract of the paper, and, therefore, would this study have been identified by a strategy that includes 'multimorbidity' search terms, and 2) if not, what other terms were used. For each review, I then calculated the proportion of studies that would have been identified using search terms that included 'multimorbidity' and these identified alternative search terms, but not the term 'comorbidity' and its linguistic variations.

First, I found that, across the three reviews, 40%, 45% and 80% of papers did not include the term 'multimorbidity' in their titles or abstracts. Instead, these studies used several alternative terms to describe the concept (for example "several chronic conditions", "multiple chronic conditions", "one or more chronic conditions", "co-occurring chronic conditions", "multiple morbidity"). I calculated that searching for 'multimorbidity' search terms and these alternative terms would capture 89% (31/35) of the papers included across all three reviews. For example, in 2012 Fortin et al. published a systematic review of 21 studies investigating multimorbidity prevalence. Of these, 55% used the term 'multimorbidity' in their title and/or abstract. Of the studies that did not, various other terms were used to describe the concept of multimorbidity ("multiple chronic conditions", "co-occurring chronic conditions"

etcetera). Using the term ‘multimorbidity’ and these various alternative terms captured 80% of the included studies.

As a result of this exercise, I chose to exclude the search term ‘comorbidity’ and its linguistic variations from my search terms. This made the review practically easier to conduct. I added some of the alternative terms identified in this exercise to my search terms to identify as many relevant hits as possible.

See tables below for the results of this exercise.

Exercise results for Fortin et al., (2012)

First Author (Year)	Study Location	Multimorbidity in title and/or abstract? (Y/N)	If no, what term(s) used?	Further Notes
Verbrugge (1989)	US	N	Comorbidity; Several chronic conditions	
Newacheck (1991)	US	N	Multiple chronic conditions	
Schellevis (1993)	Netherlands	N	Comorbidity; Chronic diseases	
Hoffman (1996)	US	N	One or more chronic conditions	
Fuchs (1998)	Israel	N	Comorbidity	
Van den Akker (1998)	Netherlands	Y		
Menotti (2001)	Finland, Netherlands, Italy	Y		
Rapoport (2004)	Canada	N	Several chronic conditions; Co-occurring chronic conditions	
Partnership for Solutions (2004)	US	N	Multiple chronic conditions	Grey literature, excluded from this exercise
Macleod (2004)	Scotland, UK	N	Comorbidity; Multiple morbidity	

First Author (Year)	Study Location	Multimorbidity in title and/or abstract? (Y/N)	If no, what term(s) used?	Further Notes
Fortin (2005)	Canada	Y		
Naughton (2006)	Ireland	N	Chronic disease	
Kadam (2007)	UK	Y		
Nagel (2008)	Germany	Y		
Marengoni (2008)	Sweden	Y		
Schram (2008)	Netherlands	Y		
Cazale (2008)	Canada			In French and full text not available, excluded from this exercise
Uijen (2008)	Netherlands	Y		
Britt (2008)	Australia	Y		
Loza (2009)	Spain	Y		
Minas (2010)	Greece	N	Chronic diseases	
Proportion with term 'multimorbidity' in title and/or abstract:		**55%**		
Proportion that should be captured by search strategy inc. multimorbidity terms, alternative terms but not 'comorbidity' terms:			**80%**	

Exercise results for France et al., (2012)

First Author (Year)	Study Location	Multimorbidity in title and/or abstract? (Y/N)	If no, what term(s) used?	Further Notes
Schellevis (1994)	The Netherlands	N	Intercurrent morbidity; Common chronic diseases	
Parkerson (1995)	US	N	Comorbidity	
Van Den Akker (2001)	The Netherlands	Y		
Bayliss (2004)	US	N	Comorbid chronic illness; Multiple chronic medical conditions	
Perkins (2004)	US	N	Comorbidity	
Van Den Akker (2006)	The Netherlands	Y		
Proportion with term "multimorbidity" in title and/or abstract:		**20%**		
Proportion that should be captured by search strategy inc. multimorbidity terms, alternative terms but not 'comorbidity' terms:			**66%**	

Exercise results for Pati et al., (2015)

First Author (Year)	Study Location	Multimorbidity in title and/or abstract? (Y/N)	If no, what term(s) used?	Further Notes
Joshi (2003)	India	N	Morbidity; Co-morbidity	
Purty (2006)	India	N	Morbidity	
Khanam (2011)	Bangladesh	Y		
Chakraborty (2004)	India	N		Grey literature, excluded from this exercise
Bhojani (2013)	India	N	Comorbid chronic conditions	
Van Minh (2008)	India, Bangladesh	N	2 or more chronic conditions	
Banjare (2014)	India	Y		
Pati (2014)	India	Y		

First Author (Year)	Study Location	Multimorbidity in title and/or abstract? (Y/N)	If no, what term(s) used?	Further Notes
Arokiasamy (2014)	India, China, Mexico, South Africa, Russia, Ghana	Y		
Vadrevu (2015)	India	Y		
Arokiasamy (2015)	India	Y		
Proportion with term "multimorbidity" in title and/or abstract:		**60%**		
Proportion that should be captured by search strategy inc. multimorbidity terms, alternative terms but not 'comorbidity' terms:			**90%**	

Developing SDoH search terms

In the literature, a SDoH is rarely referred to as a "social determinant of health", for example, whilst considered to be a SDoH, area deprivation is simply referred to as "area deprivation". To develop search terms that would capture all relevant SDoH I, therefore, had to identify specific factors considered to be SDoH that operated at household or area-levels of influence, and then include words that specifically searched for these in my terms.

To do this, I consulted and drew on several frameworks and previously conducted literature reviews (Dahlgren and Whitehead, 1991; Nagata *et al.*, 2013; Solar and Irwin, 2013; Walker *et al.*, 2014; Canadian Council on Social Determinants of Health, 2015; Duan-Porter *et al.*, 2018). I also consulted the GoInvo visualisation, which summarises SDoH across eight key frameworks and identifies 95 factors that could be considered determinants of health (GoInvo, 2018). Within the visualisation, GoInvo have identified determinants that fall under "social circumstances" and an "environment" category that would constitute SDoH. Those relevant to my review are:

- Access to healthy foods
- Crowding conditions
- Job opportunities
- Public space quality
- Educational opportunities
- Access to vocational training
- Recreational activity access
- Walkability
- Residence quality
- Quality of community support
- Work conditions

I used the search strategies employed in the conducted reviews and the factors identified in the frameworks to develop my search terms. My search terms were deliberately over-inclusive and tried to capture all the various factors considered to be household and area-level SDoH. I also ensured that my terms aligned with my conceptualisation of SDoH (see section 1.3.2).

Appendix 3 Data extraction table for initially included studies (n=41)

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Agborsangaya 2012 Canada	Cross-sectional	*Source of sample:* Alberta Health Quality Council of Alberta 2010 Patient Experience Survey *Characteristics:* ≥18 years, 52.3% female, N=4980	Annual household income Household composition (living with children vs. not, and living with adults vs. not) *Data collection*: self-reported via telephone	Multimorbidity prevalence *Data collection*: self-reported via telephone *Definition:* "Presence of two or more chronic conditions"; *No. of conditions:* 16	Having an annual household income < $30,000 CAD associated with 2.39-fold increase in multimorbidity prevalence (95% CI 1.72-3.33) compared with those >=$100,000 CAD, after adjustment for age, sex, education and living with children. Association greater for age 25-44. Not living with children associated with 2.11-fold increase in multimorbidity prevalence (95% CI 1.60-2.78) compared to those with children, after adjustment for age, sex, education and household income. Association greater for age 18-24 (although wide CIs) and age 65+. No evidence living with adults associated with multimorbidity e.g., for between 25-44, those not living with adults vs. with adults (OR 1.25, 95% CI 0.77-2.05).	H	M	M	L
Agborsangaya 2013 Canada	Cross-sectional	*Source of sample:* Alberta Health Quality Council of Alberta 2012 Patient Experience Survey *Characteristics*: ≥18 years, 55.8% female, N=4803	Annual household income *Data collection*: self-reported via telephone	Multimorbidity prevalence *Data collection*: self-reported via telephone *Definition:* "Concurrent occurrence of two or more chronic conditions in the same individual"; *No. of conditions:* 16	Having an annual household income < $30,000 CAD associated with 2.9-fold increase in multimorbidity prevalence (95% CI 2.2-3.7) compared with those >=$100,000 CAD, after adjustment for age, sex, education and obesity status.	H	H	M	L

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Arbelle 2014 Israel	Cross-sectional	*Source of sample:* EHRs of Macabi Healthcare Service, who are legally obliged to insure every citizen. Anyone alive and member of MHS on 6th August 2012 included. *Characteristics:* 0-85+ years, 51.2% female, N=1,972,798	Area socioeconomic deprivation *Data collection:* Participants' postcodes assigned to deciles of poverty index defined by parameters of 1995 national census	Multimorbidity prevalence *Data collection*: EHRs screened for conditions in clinical coding and prescription data *Definition:* "Two or more of these morbidities in one patient"; *No. of conditions:* 40	Residing in lowest SES area associated with higher prevalence of multimorbidity, particularly between 35 and 65, compared to those in highest SES areas. Between 45-49, multimorbidity present in 42.1% of those in lowest SES and 30.6% in highest. No substantial differences in older age groups (70+), whilst between 10-14 years multimorbidity was 3.8% in lowest SES level and 4.3% in highest.	L	M	H	H
Bahler 2015 Switzerland	Cross-sectional	*Source of sample:* Helsana group, the leading health insurer in the country. People included if insured in 2013. *Characteristics:* ≥65 years, 57.2% female, N=229,493	Area socioeconomic situation *Data collection:* Polling data from GFK used as a proxy of purchasing power (available net income of population) corresponding to zip code of participants	Prevalence of multiple chronic conditions *Data collection*: EHRs screened for conditions defined measure based on ATC classification system *Definition:* "Two or more chronic conditions in one person"; *No. of conditions:* 22	76.7% of those residing in areas with lowest purchasing power classified as having multimorbidity and 74.8% of those residing in area with highest purchasing power.	L	L	L	H

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Barnett 2012 Scotland	Cross-sectional	*Source of sample:* Clinical data from 314 GPs. Had to be alive and permanently registered with a participating practice on 31st March 2007. *Characteristics:* 0-85+ years, 50.5% female, N=1,751,841	Area socioeconomic deprivation *Data collection:* Carstairs deciles assigned to area in which patient lived	Multimorbidity prevalence; physical-mental multimorbidity prevalence *Data collection*: EHRs screened for conditions defined using Read Codes and prescription data *Definition:* "Two or more morbidities in one patient"; *No. of conditions:* 40	24.1% (23.9-24.4) of those residing in areas with highest level of deprivation had multimorbidity compared to 19.5% (19.3-19.6) of those in most affluent areas. Difference seen at all ages apart from those 85 and over. Equivalent prevalence of multimorbidity occurs 10-15 years earlier in most deprived vs. most affluent areas. 11.0% (10.9-11.2) of those residing in areas with highest level of deprivation had physical-mental multimorbidity compared to 5.9% (5.8-6.0) of those in most affluent areas.	L	L	L	H
Cantarero-Prieto 2018 Multi-country	Prospective cohort	*Source of sample:* 5 panel waves from Survey on Health, Ageing and Retirement in Europe. Excluded individuals who did not respond in consecutive waves *Characteristics:* ≥50 years, 56.3% female, N=31,536	Household composition (living alone vs. not) Rurality of household (definition unclear) *Data collection:* interviewed (no further details)	Prevalence of multiple chronic conditions *Data collection*: interviewed (no further details) *Definition:* "Three or more chronic diseases"; *No. of conditions:* 14	Strong evidence that 20% higher odds of multimorbidity amongst those living alone vs. those living with others (OR=1.20, 95% CI 1.04-1.39, P<.05). Variables adjusted for unclear. No evidence of an association between rurality of household and multimorbidity prevalence (OR 0.92, 95% CI 0.93-1.03, P>0.1). Variables adjusted for unclear.	U	U	M	U

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Cassell 2018 England	Retrospective cohort	*Source of sample:* CPRD database linked to deprivation quintiles and HES data. Included if up-to-standard registration data for at least 1 year prior to (April 16) and during study. Random subsample included. *Characteristics:* ≥18 years, 50.7% female, N=403,985	Area socioeconomic deprivation *Data collection:* IMD quintiles (year unclear) assigned to patient postcodes	Multimorbidity prevalence; physical-mental multimorbidity prevalence *Data collection*: EHRs screened for conditions defined using Read Codes and product codes. 4-year lookback. *Definition:* "Two or more currently active long-term conditions"; *No. of conditions:* 36	30.0% (29.6-30.4) of those residing in areas with highest level of deprivation had multimorbidity compared to 25.8% (25.5-26.0) of those in the most affluent areas. Difference greater for middle aged individuals (between 45 and 74 years). 14.0% (13.7-14.2) of those residing in areas with highest level of deprivation had physical-mental multimorbidity compared to 7.5% (7.2-7.7) of those in most affluent areas. Difference greater for middle aged individuals (between 35 and 84 years).	H	L	L	H
Charlton 2013 England	Prospective cohort	*Source of sample:* CPRD database linked to deprivation quintiles. Patients with complete data on deprivation included and followed up from 1 Jan 05 - 30 April 12. *Characteristics:* ≥ 30 years, 50% female, N=282,887	Area socioeconomic deprivation *Data collection:* IMD quintiles (2010) assigned to patient postcode	Incidence of multiple morbidity; prevalence of depression at different levels of morbidity *Data collection*: EHRs screened for presence of condition defined using Read Codes *Definition:* "Dual (2 conditions) and triple (3) morbidity"; *No. of conditions:* 5	Incidence of dual and triple morbidity associated with deprivation (e.g. highest deprivation accounted for 26%, and lowest deprivation 16%, of dual condition incidences, adjusted for age and sex). Relative risk of triple morbidity was 5.51 (4.70-6.47) for most deprived quintile and 4.76 (3.81-5.96) for the least versus those developing no conditions in the least deprived quintile. Depression was associated with deprivation at all levels of multimorbidity.	H	L	M	M

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Chung 2015 Hong Kong	Cross-sectional	*Source of sample:* Hong Kong Government's Thematic Household Survey (Oct 11-Jan 12) *Characteristics:* ≥15 years, 52.2% female, N=25,780	Monthly household income Household tenure *Data collection:* self-reported using structured questionnaires given in face-to-face home interviews	Multimorbidity prevalence *Data collection*: self-reported using structured questionnaires given in face-to-face home interviews *Definition:* "Two or more chronic health conditions"; *No. of conditions:* 46	Reporting an income of <4,000HKD associated with 52% increased odds of multimorbidity versus reporting income of >40,000HKD (OR 1.52, 95% CI 1.39-1.66, P<.001) after adjusting for demographics, education, housing and employment status. Compared to public (social) housing residents, homeowners, private renters and those in subsidized housing had 17% (OR 1.17, 95% CI 1.11-1.24, P=0.003), 19% (OR 1.19, 95% CI 1.09-1.29, P=0.041) and 11% (OR 1.11, 95% CI 1.05-1.18, P=0.070) higher odds of multimorbidity, respectively, in multivariate analyses.	H	M	M	L
Foguet-Boreu 2014 Spain	Cross-sectional	*Source of sample:* EHRs collected by The Catalan Health Institute. 40% of these meet the highest quality criteria and a 2010 subsample of these used *Characteristics:* ≥19 years, 50.7% female, N=1,749,710	Rurality of household *Data collection:* Assigned to the participants 'area of residence' (rural if <10,000 inhabitants and/or population density <150 people/km^2, otherwise urban)	Multimorbidity prevalence *Data collection*: EHRs screened for conditions for ICPC-2 codes classified as chronic according to O'Halloran criteria *Definition:* "Coexistence of two or more chronic diseases"; *No. of conditions:* 146 diagnostic clusters	47.6% of those living in rural areas classified as having multimorbidity and 46.6% of those not in rural areas. Differences in prevalence similar for women and men. For example, ORs (95% CIs) for women and men (45-64 years) were 0.80 (0.78-0.82) and 0.87 (0.85-0.89), respectively (P values<.001). When stratified by age categories and adjusted for covariates, odds of multimorbidity consistently lower for those living in rural locations versus not across all age groups, although variables adjusted for unclear. Inequality in multimorbidity prevalence with area rurality greater those 45 and over for men and women.	U	L	L	U

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Hayek 2017 Israel	Cross-sectional	*Source of sample:* Israeli National Health Interview Survey (2014-2015) *Characteristics:* ≥21 years, 49.6% female, N=4,325	Monthly household income *Data collection:* Self-reported via questionnaire delivered via phone	Prevalence of multiple chronic conditions *Data collection*: Self-reported over telephone if physician diagnosed them *Definition:* "Two or more self-reported physician-diagnosed conditions"; *No. of conditions:* 10	Strong evidence that the proportion of people with multiple chronic conditions was 1.7 times higher amongst those with a monthly household income ≤$2,000 than those with >$4,000 (PRR 1.7, 95% CI 1.2-2.5, P=.005). Variables adjusted for unclear.	U	H	H	U
Henchoz 2019 Switzerland	Retrospective cohort	*Source of sample:* Lausanne cohort 65+ study - 3 samples of population (04, 09 and 14) *Characteristics:* 65-70 years, 58% female, N=4,055	Poor family economic environment (in childhood) Household composition (living alone vs. not) *Data collection:* Self-reported in baseline questionnaire	Multimorbidity prevalence *Data collection*: Self-reported using questionnaire (at 2-year follow-up) *Definition:* "Co-occurrence of two or more medical conditions"; *No. of conditions:* 13	No association between poor family economic environment in childhood and multimorbidity in older age (OR=0.94, 95% CI 0.74-1.19) after adjustment for sex, cohort, socioeconomic status, behaviours, other stressful events in childhood and in adulthood. Strong evidence of an association between household composition and multimorbidity in univariate analyses (P<.001). 31.4% of those living alone classified as having multimorbidity and 24.7% of those living with others.	U	M	M	M
Humphreys 2018 England	Prospective cohort	*Source of sample:* Hertfordshire Cohort Study, participants linked to birth records *Characteristics:* 64-68 years, 49% female, N=1,979	Paternal social class (at birth) *Data collection:* Nurse-administered questionnaires given during home visit at birth	Multimorbidity count *Data collection*: Follow-up postal questionnaire asked for disease information *Definition:* "Total number of multi-morbid conditions"; *No. of conditions:* 10	No association found between paternal manual social class at birth and multimorbidity count at follow-up when compared to non-manual social class categories and after adjusting for baseline age, gender, health behaviours, time in cohort and year of recruitment (OR 1.15, 95% CI 0.93, 1.43, P>0.01).	H	L	H	L

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Johnson-Lawrence 2017 USA	Cross-sectional	*Source of sample:* National Health Interview Surveys (02-14). Those with education information and on or more chronic conditions included. *Characteristics:* 30-64 years, % female N/A, N=115,097	Household income Household tenure *Data collection:* Self-reported in face-to-face interview	Multimorbidity prevalence *Data collection*: Self-reported in face-to-face interview *Definition:* "Two or more conditions"; *No. of conditions:* 9	Odds of multimorbidity increased by 45% amongst those in the bottom tertile of household income versus the highest tertile (OR 1.45, 95% 1.38-1.53) after adjusting for age, gender, ethnicity, education, interview year, region of residence, marital status, last doctor visit, employment and home ownership. Those who rent their properties had 19% higher odds of multimorbidity compared to homeowners (OR 1.19, 95% CI 1.15-1.24) in multivariable analyses.	U	M	H	L
Johnston 2019 Scotland	Prospective cohort	*Source of sample:* Aberdeen Children of the 50s, a cohort of individuals born in Aberdeen between 1950 and 56. *Characteristics:* age range N/A, 52.3% female, N=6,561	Paternal social class (at birth) *Data collection:* Participants' linked to birth records containing paternal occupation, coded using General Register Office's Occupational classification (1950)	Multimorbidity prevalence *Data collection*: In postal questionnaire, asked to list up to six 'long-term illnesses, health problems or disabilities which limit (their) daily activities or work (they) can do' *Definition:* "Two or more self-reported conditions"; *No. of conditions:* N/A	After adjustment for gender, educational attainment, cognition at age 7 and school type, strong evidence paternal social class at birth associated with multimorbidity in older age (P<.001). Compared to individuals whose father was in skilled manual occupations, individuals whose fathers were unemployed/disabled/dead or their occupation unknown had 74% higher odds of multimorbidity (OR 1.74, 95% CI 1.11-2.72).	M	L	H	L

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Katikireddi 2017 Scotland	Prospective cohort	*Source of sample:* West of Scotland Twenty-07 cohort, respondents from 3 cohorts born in early 1930s, 1950s and 1970s. All cohorts and waves used in analysis apart from 1970s cohort. *Characteristics:* 18-75 years, % female N/A, N=3,466	Household income (equivalised) Area socioeconomic deprivation *Data collection:* Self-reported income and weighted for no. and age of residents; tertiles of Carstairs scores assigned to postcodes for deprivation	Multimorbidity prevalence *Data collection*: Self-reported conditions in face-to-face interviews for all waves apart from wave 3 (postal questionnaire) *Definition:* "Two or more (or three or more) of the relevant conditions"; *No. of conditions:* 40	Strong evidence found for higher odds of multimorbidity amongst those with lowest level of household income compared to highest (OR 1.53, 95% CI 1.25-1.87, P<.05). Adjusted for age, age^2, age^3, sex, cohort, prior multimorbidity, time between waves and sex*cohort interaction. Strong evidence found for higher odds of multimorbidity amongst those living in the most deprived compared to the least deprived areas (OR 1.46, 95% CI 1.26-1.68, P<.05). Adjusted for same variables. Difference greater for those between 50 and 70 years and relationship stronger when multimorbidity defined as three or more conditions.	M	M	M	L
Ki 2017 Korea	Longitudinal panel	*Source of sample:* Korea Health Panel Study (2009-2011, 2nd-4th waves) *Characteristics:* ≥30 years, 53.7% female, N=9,971	Relative household poverty *Data collection:* Self-reported income (poverty = less than half the median annual household income, equivalised to account for number of residents)	Number of diseases *Data collection*: Self-reported in face-to-face interview/computer assisted interview. Checked using health records. *Definition:* N/A; *No. of conditions:* 66	33% of those classified as "poor" had ≥ 3 diseases compared to 12.6% of "non-poor" participants (P<.001).	U	H	M	H

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Laires 2018 Portugal	Cross-sectional	*Source of sample:* Portuguese National Health Survey (2014) *Characteristics:* 25-79 years, 56% female, N=15,196	Household income *Method of data collection:* N/A	Multimorbidity prevalence *Method of data collection*: Self-reported (no further details) *Definition:* "Two or more of these self-reported chronic conditions"; *No. of conditions:* 13	51.2% of those with the lowest household income level were classified as having multimorbidity and 32.7% of those with the highest household income level.	L	H	M	H
Lebenbaum 2018 Canada	Pooled cross-sectional	*Source of sample:* Pooled data from 96-97 National Population Health Survey and 05 and 12-13 Canadian Community Health Surveys *Characteristics:* ≥18 years, 49.8% female in 96-97, 49.7% in 05, 49.6% in 12-13, N=288,300	Household income (equivalised) Household tenure Rurality of household (definition unclear) *Data collection:* Self-reported via computer questionnaire. Income adjusted for no. in household.	Multimorbidity prevalence *Data collection*: Self-reported conditions using computer assisted interview methods *Definition:* "At least two chronic conditions"; *No. of conditions:* 10	Participants with the highest income had 43% less odds of multimorbidity compared to those with the lowest (OR 0.57, 95% CI 0.52-0.62, P<.001) after adjusting for demographic, behavioural and socioeconomic variables. Homeownership associated with 18% lower odds of multimorbidity (OR 0.82, 95% 0.78-0.87, P<.001) in multivariate analyses. No evidence rurality associated with multimorbidity in multivariate analyses (OR 0.98, 95% CI 0.93-1.02, P=0.323).	L	M	H	L

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Li 2016 England	Cross-sectional	*Source of sample:* Baseline data from Yorkshire Health Study. *Characteristics:* 16-85 years, 56.3% female, N=27,806	Area socioeconomic deprivation *Data collection:* Quintiles of IMD scores (2010) assigned to participant postcodes	Multimorbidity prevalence *Data collection*: Self-reported conditions in questionnaire (postal or online) *Definition:* "At least two of the listed conditions"; *No. of conditions:* 12 (plus 'other' category)	45.7% of those residing in areas with the highest level of deprivation had multimorbidity compared to 26.8% of those in the most affluent areas.	M	L	M	H
Lujic 2017 Australia	Retrospective cohort	*Source of sample:* Linked data: 45 and Up Study - a random sample from Medicare data (05-09), Pharmaceutical Benefits Scheme - subsidised prescriptions (05-11), Hospital admissions data (00-13) *Characteristics:* ≥45 years, 55.7% female, N=90,352	Household income Speaks language other than English at home Rurality of household (definition unclear) *Data collection:* Self-reported in baseline questionnaire	Multimorbidity prevalence *Data collection*: Data obtained differently depending on dataset. Self-reported in 45 and Up, EHRs screened in medication data for ICD-10-AM codes and in hospital data for ATC codes. 2-year lookback. *Definition:* "Two or more chronic conditions"; *No. of conditions:* 8	Consistently lower odds of multimorbidity with higher income across datasets e.g. 42% lower odds if income $70k+ vs. <$20k when self-reporting health data (OR 0.58, 95% CI 0.52-0.66). Adjusted for age and sex. Not speaking English as primary language associated with lower odds of multimorbidity in self-report data (OR 0.80, 95% CI 0.76-0.84) and higher odds in medication/hospital data (e.g. hospital: OR 1.32, 95%CI 1.32-1.42). Adjusted for age and sex. Living in remote/very remote areas associated with increased odds of multimorbidity versus living in a major city when health information was self-reported (OR 1.14, 95% CI 1.03-1.26), obtained from medication data (OR 1.11, 95% CI 1.00-1.23) or hospital data (OR 1.28, 95% CI 1.08-1.53). Adjusted for age and sex.	H	M	M	M

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
McLean 2014 Scotland	Cross-sectional	*Source of sample:* Clinical data from 314 GPs. Had to be alive and permanently registered with practice on 31st March 07. *Characteristics:* ≥25 years, % female N/A, N=1,272,685	Area socioeconomic deprivation *Data collection:* Carstairs deciles assigned to participant postcodes	Prevalence of physical-only multimorbidity, physical-mental multimorbidity, mental-only multimorbidity *Data collection*: EHRs screened for presence of conditions defined using Read Codes and prescription data *Definition:* "Coexistence of two or more chronic conditions"; *No. of conditions:* 40	14.9% of those living in most deprived areas had physical-only multimorbidity compared to 16.8% of those in the least deprived. 2.0% of those living in most deprived areas had mental-only multimorbidity compared to 0.7% of those in the least deprived. Differences greater for ages 25-54, prevalence similar ≥65 years. 17.0% of those living in most deprived areas had physical-mental multimorbidity compared to 9.0% of those in the least deprived. Differences seen in all age groups <75 years.	L	L	L	H
Melis 2014 Sweden	Prospective cohort	*Source of sample:* Kungsholmen Project (91-93). Included those living independently and with no multimorbidity. *Characteristics:* ≥75 years, % female N/A, N=390	Household composition (living alone vs. living with others) *Data collection:* Self-reported in baseline social interview via standardised protocol	Multimorbidity incidence *Data collection*: Physicians determined conditions using medical history, inpatient registry and clinical examination *Definition:* "Co-occurrence of two or more chronic conditions"; *No. of conditions:* 38	In univariate analyses, no evidence found an association between living alone (versus with others) and multimorbidity incidence at follow-up, regardless of whether participants have no or one chronic disease at baseline (e.g. no disease at baseline OR 1.34, 95% CI 0.60-3.01).	U	M	L	H

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Moin 2018 Canada	Retrospective cross-sectional	*Source of sample:* EHRs linked to insurance data Sample contains all residents in Ontario eligible for health insurance and alive in 2015. *Characteristics:* 0-85 years, 51.3% female, N=12,516,587	Area material deprivation Area residential instability Ethnic concentration of area Area dependency (no. adults out of work or unpaid) *Data collection:* ON-Marg Index scores assigned to postcodes	Multimorbidity prevalence *Data collection*: EHRs screened for conditions based on ICD-10 codes. 10-year lookback. *Definition:* "Co-occurrence of two+ (and three+) chronic conditions"; *No. of conditions:* 18	20.4% of those in the most materially deprived areas classified as having multimorbidity (2+) compared to 15.7% of those in the least. These estimates change to 9.2% and 6.3%, respectively, for multimorbidity (3+). No differences seen by gender. Differences between most and least deprived greatest for ages 55-74. Relationship between residential instability and multimorbidity prevalence similar to material deprivation (data plotted visually but not reported in numerical form). Ethnic concentration of residential area and dependency of residents had no association with multimorbidity prevalence (data plotted visually but not reported in numerical form).	L	L	L	H
Mounce 2018 England	Prospective cohort	*Source of sample*: English Longitudinal Study of Ageing. Included participants in all 6 waves from 02-03 to 12-13. *Characteristics:* ≥50 years, 56.3% female, N=5,564	Household composition (living alone vs. not) Data collection: self-reported	Multimorbidity incidence *Data collection*: Self-reported conditions. Ascertained at each follow-up whether mental health condition(s) in remission. *Definition:* Two or more conditions; *No. of conditions:* 15	Living alone at baseline (versus cohabits) was not found to be associated with multimorbidity incidence after 11 years follow up (HR 0.93, 95% CI 0.71-1.21, P=.580) and after adjusting for baseline age, sex, total wealth, educational attainment, health behaviours, social detachment and locus of control.	U	M	M	L

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Neilsen 2017 Multi-country	Cross-sectional	*Source of sample:* Wave 5 of Survey on Health, Ageing and Retirement in Europe. *Characteristics:* ≥50 years, 55.4% female, N=63,842	Monthly household income *Data collection:* Self-reported in face-to-face interview	Multimorbidity prevalence *Data collection*: Self-reported face-to-face *Definition:* "Coexistence of two or more chronic conditions"; *No. of conditions:* 12	Participants with the lowest level of household income had 44% increased odds of multimorbidity compared to those with the highest (OR 1.44, 95% CI 1.32-1.59, P<.05), after adjusting for age, sex, and education level.	U	H	M	L
Orueta 2013 Spain	Retrospective cross-sectional	*Source of sample:* EHRs from Population Stratification Programme. Included those covered by health insurance on 31st Aug 11 and for 6 months in previous year. *Characteristics:* ≥65 years, 57.5% female, N=452,698	Area socioeconomic deprivation D*ata collection:* Participants' postcode assigned quintile of deprivation index based on census tract	Prevalence of multimorbidity (any), physical-mental, and physical only, multimorbidity *Data collection*: EHRs screened using ACG classification system *Definition:* "Co-occurrence of two or more (or three or more) health problems"; *No. of conditions:* 47	69.9% (69.6-70.3) of those in most deprived areas classified as having any multimorbidity vs. 60.2% (59.9-60.5) of those in the least. Inequalities greater for women and younger ages. Results similar if multimorbidity defined as 3 or more health problems. Living in the most deprived areas (vs. the least deprived) associated with higher prevalence of physical-mental multimorbidity and physical multimorbidity (78.1% vs. 71.8%, and 62.0% vs. 51.7%, respectively).	H	L	L	H

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Orueta 2013 Spain	Retrospective cross-sectional	*Source of sample:* EHRs from Population Stratification Programme. Included those covered by health insurance on 31st Aug 11 and for 6 months in previous year. *Characteristics:* 0-75+ years, 50.9% female, N=2,262,686	Area socioeconomic inequality *Data collection:* Participants' postcode assigned quintile of deprivation index based on census tract. Concentration index as the measure of socioeconomic-related inequality.	Prevalence of chronic diseases *Data collection*: EHRs screened using ACG classification system. 4-year lookback. *Definition:* "Number of chronic conditions"; *No. of conditions:* 52	After controlling for age, individuals living in more deprived areas had disproportionately more conditions than those living in the least deprived areas. Degree of inequality increased with increasing number of conditions. Inequality was greater for females than males for all numbers of conditions.	U	L	L	H
Orueta 2014 Spain	Retrospective cross-sectional	*Source of sample:* EHRs from Population Stratification Programme. Included those covered by health insurance on 31st Aug 11 and for 6 months in previous year. *Characteristics:* 0-85+ years, 50.9% female, N=2,262,698	Area socioeconomic deprivation *Method of data collection:* Participants' postcode assigned quintile of deprivation index based on census tract	Multimorbidity prevalence *Method of data collection*: EHRs screened for presence of conditions using ACG classification system. 4-year lookback period used. *Definition:* "Coexistence of two or more conditions in the same patient"; *No. of conditions:* 52	26.1% of those living in the most deprived areas classified as having multimorbidity compared to 20.5% of those in the least deprived. Differences greater for women than men (for women, 29.4% vs. 22.2% in most vs. least deprived areas; for men equivalent crude %s are 22.7% vs. 18.7%). Differences in prevalence as a function of area deprivation are negligible <34 years of age and most prominent between 55 and 79 years of age.	U	L	L	H

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Prazeres 2015 Portugal	Cross-sectional	*Source of sample:* Enrolled GPs who invited all adults attending consultations to participate in study during 3 days on 3 consecutive weeks. *Characteristics:* ≥18 years, 64.2% female, N=1,993	Perceived problems managing monthly household income Household composition (living as couple, with extended family, alone or other inc. care home) Rurality of household (definition unclear) *Data collection:* Self-reported using questionnaire	Multimorbidity prevalence *Data collection*: GPs recorded conditions using own knowledge, patient's self-report and medical records *Definition:* "Presence of ≥two or ≥three chronic health problems"; *No. of conditions:* 147 diagnostic clusters	No association between problems managing income and multimorbidity when defined as ≥two conditions (e.chances of multimorbidity for those self-reporting "Some monthly income left over" vs. "Not enough monthly income to make ends meet" were OR 0.8, 95% CI 0.5-1.1, P=0.182). Adjusted for age, sex, marital status, education, professional status, residence area, living arrangement. Data not reported for ≥three. No association found between household composition and multimorbidity in multivariate analysis. E.g. vs. living alone, ORs (95% CIs, P values) for those living as a couple were 1.4 (0.9-2.3 P=0.182 and 0.9 (0.6-1.5, P=0.778) when multimorbidity defined as 2+ and 3+ conditions, respectively, in multivariate analyses. In multivariate analysis, residing in rural areas versus urban not associated with multimorbidity when defined as ≥two (p=0.746) or ≥three (p=0.157) conditions in multivariate analyses.	H	M	L	L
Roberts 2015 Canada	Cross-sectional	*Source of sample:* Canadian Community Health Survey 2011/12 *Characteristics:* ≥20 years, % female N/A, N=105,416	Household income Highest level of education in household Rurality of household (definition unclear) *Data collection:* Self-reported income and education in interview	Multimorbidity prevalence *Data collection*: Self-reported conditions on questionnaire "expected to last or have already lasted 6 months or more and that have been diagnosed by a health professional". *Definition:* 2 or more, and 3 or more, chronic diseases (3 or more used in multivariable analyses); *No. of conditions:* 9	Those in the lowest income quintile had over 4 times for odds of multimorbidity than those in the highest (OR 4.4, 95% CI 3.6-5.5), after adjusting for age, sex, household education, Aboriginal status, activity level smoking, stress, blood pressure and obesity. Difference remained across age categories, but reduced for those 65+ (OR 2.5, 95% CI 1.8-3.5) Those living in households were no one completed high school had over 4 times odds of multimorbidity (OR 1.8, 95% CI 1.6-2.1), adjusting for age and sex. Living in rural areas associated with 10% increase in multimorbidity (OR 1.1, 95% CI 1.0-1.3), adjusted for age and sex.	H	M	H	M

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Ryan 2018 Canada	Cross-sectional	*Source of sample:* Linked EHRs. Participants required to be alive, have had contact with health service in 7 years and have health insurance (on 1st July 13) *Characteristics:* 0-105 years, 50.9% female, N=13,581,191	Area material deprivation Rurality of household (town <10,000) *Data collection:* Quintiles of urban material deprivation-based domain of ON-Marg index assigned to participants' postcodes	Multimorbidity prevalence *Data collection*: Presence determined if recorded in cohort and/or EHRs screened for ICD-9 or ICD-10 codes *Definition:* "Presence of three or more chronic conditions"; *No. of conditions:* 17	Age-sex standardised rate of multimorbidity 12.3% (12.1-12.5) for those living in the most deprived urban areas and 10.3% (10.2-10.3) for those in the least deprived urban areas. Age-sex standardised rate of multimorbidity 11.0% (11.0-11.1) for those in rural areas.	L	L	L	M
Salisbury 2011 England	Retrospective cohort	*Source of sample:* GPRD database linked to deprivation data. Included if registered at one of practices on 1st April 05. *Characteristics:* ≥18 years, % female N/A, N=99,997	Area socioeconomic deprivation *Data collection:* Quintiles of Townsend calculated using census (01) data and assigned to participants' postcodes	Multimorbidity prevalence *Data collection*: EHRs screened *Definition:* "More than one chronic condition"; *No. of conditions:* 17 (plus ACG/EDC approach of 114 clusters)	Those in most deprived quintile for deprivation were more than twice as likely to have multimorbidity as those in the least deprived quintile (OR 1.91, 95% CI 1.78-2.04) after adjusting for age and sex. Similar results found for ACG/EDC approach, although relationship less marked and results not shown.	U	L	L	M

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Schäfer 2012 Germany	Prospective cohort	*Source of sample:* EHRs from 158 GP practices. Included regular patients with 3 or more chronic conditions only. Exclusion criteria inc. unable to be interviewed, nursing home resident and had severe illness probably lethal in three months. *Characteristics:* 65-84 years, 59.3% female, N=3,189	Monthly household income (equivalised) Household tenure (owner vs. not) Household composition (living at home alone, with spouse, with family members/others, living in assisted living/retirement home) *Data collection:* Self-reported via questionnaire. Income weighted for no. and age of residents.	Multimorbidity prevalence *Data collection*: EHRs screened for diagnoses and open questions in baseline GP interviews ("Which additional diagnoses does that patient have?"). *Definition:* Number of chronic conditions; *No. of conditions:* 29	Evidence that the number of conditions individuals have decreases by 0.27 (-0.47 to -0.08) per unit on the logarithmic scale of income (p=0.005; one step on scale equates to one of following steps: €400 to €1,100 to €3,000 to €8,100 net income per month). Adjusted for age, gender, marital status, job autonomy, household composition and tenure. In multivariate analyse, no evidence number of chronic conditions differs with homeownership (vs. not homeowner) (-0.13 conditions, 95% CI -0.30-0.05, P=0.148) or different types of household composition (e.g. living at home with spouse vs. living alone associated with -0.10 conditions, 95% CI -0.42-0.23, P=0.562).	H	M	L	U
Sinnott 2015 Ireland	Retrospective cross-sectional	*Source of sample:* Baseline data from Mitchelstown cohort (patients from single GP). *Characteristics:* 50-69 years, 51% female, N=2,047	Household dysfunction *Data collection:* Self-reported during interview using ACE questionnaire	Multimorbidity prevalence, prevalence of psychiatric disease with multimorbidity *Data collection*: Self-reported in questionnaire *Definition:* "Two or more chronic diseases"; *No. of conditions:* 20	Higher odds of multimorbidity found for those reporting history of household dysfunction in childhood compared to those not after adjustment for age, gender, education, income, behaviour factors, depression and anxiety scores (OR 1.4, 95% CI 1.1-1.7, P<.05). Higher odds of psychiatric disease in those with multimorbidity for those reporting household dysfunction in childhood compared to those not, after adjusting for same variables (OR 1.6, no 95% CIs).	H	M	M	L

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Stanley 2018 New Zealand	Cross-sectional	*Source of sample:* EHRs (covering all publicly funded hospital discharges, and some private, and community-dispensed prescriptions). Included individuals with health insurance (Jan 2014). *Characteristics:* ≥18 years, 51.8% female, N=3,489,747	Area socioeconomic deprivation *Data collection:* Quintiles of NZDep index (2013) based on NZ census and tagged to participants addresses	Multimorbidity prevalence *Data collection*: EHRs screened for conditions. 5-year lookback for hospital data and 1 year for pharmaceutical *Definition:* "At least two conditions from two different condition lists"; *No. of conditions:* 61 in hospital data, 30 in pharmaceutical data	Multimorbidity was more common among those in higher socioeconomic deprivation areas, with age and sex standardised prevalence based on hospital diagnoses rising from 5.8% (least deprived quintile) to 10.8% (most deprived quintile); and for pharmaceutical-based definitions from 25.1% (least deprived) to 30.9% (most deprived). Difference in prevalence with levels of deprivation greater for those aged 35-74 years old.	L	L	L	M
Stokes 2018 New Zealand	Cross-sectional	*Source of sample:* EHRs of Maori and Pacific patients at a large urban GP in an island of NZ *Characteristics:* ≥35 years, % female N/A, N=232	Area socioeconomic deprivation *Data collection:* Quintiles of NZDep index tagged to participants addresses	Multimorbidity prevalence *Data collection*: EHRs screened for conditions *Definition:* "Presence of two of more morbidities in one patient"; *No. of conditions:* 31	61.4% of those in areas with highest level of deprivation were classified as having multimorbidity compared to 47.2% of those living in the least deprived areas. Difference in raw percentages of multimorbidity prevalence in most versus least deprived areas greater for Pacific patients than Maori patients - Pacific: 65.0% (40.8-84.6) in most deprived and 44.4% (13.7-78.8) in least, Maori: 59.5% (42.1-75.3) in most and 48.5% (28.7-68.0) in least.	H	L	L	H

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Tomasdottir 2016 Norway	Prospective cohort	*Source of sample:* Second and third waves - 95-97 and 06-08 - of the Nord-Trondelag Health Study. 11 years follow-up. *Characteristics:* 20-59 years, 53.7% female, N=20,365	Distrusting neighbours *Data collection:* Self-reported using questionnaire, asked to rate agreement with "Answer with regard to your environment i.e. neighbourhood/group of farms: One cannot trust each other here"	Multimorbidity prevalence *Data collection*: Self-reporting in face-to-face interview and clinical examination *Definition:* "Two or more coinciding chronic diseases coinciding within the same individual"; *No. of conditions:* 17	After adjustment for age, gender, smoking, physical activity, education and current depressive symptoms, no evidence of an association between distrusting neighbours at baseline and risk of developing multimorbidity within 11 years. RR for those who "strongly agree" with statement 1.13 (95% CI 0.98-1.32) compared to those who "strongly disagree".	H	H	M	L
Tucker-Seeley 2011 USA	Retrospecti ve cohort	*Source of sample:* 2004 wave of The Health and Retirement Study, linked to records of lifetime earnings. *Characteristics:* 50-75+ years, 53.6% female, N=7,305	Childhood financial hardship *Data collection:* Self-reported in interview, asked "While you were growing up, before age 16, did financial difficulties ever cause you or your family to move to a different place?"	Multimorbidity prevalence *Data collection:* Asked if a doctor had ever told them if they have one of the diseases *Definition: "*Count of chronic conditions"; *No. of conditions*: 6	In the unadjusted model, the expected number of chronic conditions for those reporting childhood financial hardship was 1.11 (95% CI 1.04-1.19) times that of those not reporting childhood financial hardship. After adjustment for age, gender, race and educational attainment, this estimated number of chronic conditions reduced to 1.08 (95% CI 1.02-1.14) times greater for those reporting childhood financial hardships versus those not.	U	H	H	L

First Author Year Country	Study Design	Participants	Exposure(s)	Outcome(s)[a]	Key Findings[b]	Risk of Biases[c]			
						Selection	Information (Exposure)	Information (Outcome)	Confounding
Verest 2019 Netherlands	Cross-sectional	*Source of sample:* Baseline data of HELIUS study (2011-2015) *Characteristics:* 18-70 years, 42.3% female, N=22,362	Problems managing household income *Data collection:* Self-reported using questionnaire	Multimorbidity prevalence *Data collection*: Self-reported using questionnaire, depression ≥10 on PHQ *Definition:* "Two or more chronic diseases"; *No. of conditions:* 21	73.6% of those with "lots of problems" and 34.2% of those with "no problems" were classified as having multimorbidity. Consistent patterns of higher odds of multimorbidity in lower SES groups for men and women in all ethnic groups, after adjustment for age. E.g. odds ratio of multimorbidity for Dutch male participants reporting lots of problems was 4.48 (2.76-7.29) and for Ghanaian males was 2.79 (1.77-4.38), when compared to those with "no problems". In women, equivalent estimates were 6.82 (4.47-10.41) and 2.60 (1.79-3.77), respectively.	H	H	M	H
Violan 2014 Spain	Cross-sectional	*Source of sample:* EHRs collected by The Catalan Health Institute. 40% of these meet the highest quality criteria and a 2010 subsample of these used *Characteristics:* ≥19 years, 51% female, N=1,356,761	Area socioeconomic deprivation *Data collection:* Participants' postcode assigned quintile of deprivation index based on census tract	Multimorbidity prevalence *Data collection*: EHRs screened for conditions based on ICPC-2 codes considered chronic *Definition:* "Coexistence of two or more chronic conditions"; *No. of conditions:* 146 diagnostic clusters	In multivariate analysis, odds of multimorbidity prevalence were greater for those in most deprived compared to the least deprived areas (OR 1.07, 95% CI 1.05-1.09). Adjusted for age, sex, number of visits home and primary care health visits during previous 12 months and quartiles of attended population. After adjustment for number of home and primary care health visits, and quartiles of attended population, women of all ages and men aged 25 to 65 showed a significant association (i.e. increasing deprivation associated with greater multimorbidity). For under 65s, greater variation in multimorbidity for women than men across all deprivation quintiles.	L	L	L	M

[a]EHRs=electronic health records; [b]Findings reported as in paper (i.e., my own conversions of data into ORs are not included); [c]H=high, M=Medium, L=Low, U=Unclear.

Appendix 4 Risk of Bias tool used in review

Risk of Bias	High	Medium	Low	Unclear
Selection bias				
Study population not representative of the population of interest	*(1) Sample not representative of general population e.g., convenience sampling, or low response rate with large difference sample characteristics from general population* *OR* *(2) (For cohort) Attrition rates depend on exposed status/different levels of exposure and no statistical analysis demonstrating low-moderate effect of differential loss to follow up on effect estimates*	*(1) Random sampling from general population with baseline inclusion criteria/sampling procedure that may be moderately associated with exposures or outcome e.g. moderate response rate with small-medium difference of sample characteristics from general population.* *OR* *(2) (For cohort) Attrition rates moderately differ according to exposed status/different levels of exposure, or statistical analysis demonstrates only moderate effect of differential loss to follow up on effect estimates* *OR* *(3) (For cohort) Attrition rates depend on exposed status/different levels of exposure but statistical analysis methods to account for attrition*	*(1) Random sampling from general population with baseline inclusion criteria/sampling procedure that should not be associated with exposures or outcome. E.g., high response rate and sample seemingly representative of general population* *OR* *(2) Weighted sampling strategy appropriately handled in analysis* *(3) (For cohort) Attrition rates low and similar for exposed and unexposed cohorts or different levels of exposure, or statistical analysis demonstrates only small effect of differential loss to follow up on effect estimates*	*No/inadequate information (e.g., no inclusion criteria)*

Risk of Bias	High	Medium	Low	Unclear
Information bias (Exposure)				
Potential information biases due to ascertainment of exposure(s) that is likely to be differential according to the outcome (multimorbidity)	*Self-report in circumstances likely to be strongly associated with outcome, such as retrospective recall of childhood circumstances, without appropriate sensitivity analyses or statistical adjustments to mitigate or quantify bias*	*(1) Self-report in circumstances unlikely to be associated with outcome, such as prospective data collection or objective measurement* *AND/OR* *(2) Self-report method validated using administrative or official data* *AND/OR* *(3) Administrative or official data used for all exposures but misclassification or missingness of exposure is likely to be associated with outcome*	*Validated tool used to collect information via administrative or official data with low risk of misclassification or missingness that is associated with the outcome (e.g., Index of Multiple Deprivation scores)*	*No/inadequate information (e.g., questionnaire used unavailable)*

Risk of Bias	High	Medium	Low	Unclear
Information bias (Outcome)				
Potential biases due to ascertainment of outcome (multimorbidity) that is likely to be differential according to the exposure(s).	*(1) Self-report based on a list of fewer than 12 diseases* *OR* *(2) Tool to measure multimorbidity not validated or validated in a very different population than study sample*	*(1) Assessment using objective tool validated in a similar population with disease list including fewer than 12 diseases* *OR* *(2) Administrative data with disease list including fewer than 12 diseases* *OR* *(3) Self-reported outcome based on a list of at least 12 diseases, with validation using administrative/official data*	*(1) Assessment using objective tool validated in a similar population, with specified list including at least 12 diseases* *OR* *(2) Multimorbidity ascertained from administrative data, with specified list including at least 12 diseases*	*No/inadequate information to assess risk of bias*
Confounding				
Potential biases due to insufficient control for confounding	*Does not address confounding by age and sex*	*Addresses confounding by age and sex (e.g., matching by age and sex; age and sex demonstrated not to be associated with exposure of interest; age and sex considered for inclusion in the final model)*	*Addresses confounding by age and sex, and at least one other potential confounder (e.g., age, sex and at least one other potential confounder eligible for inclusion in the final model)*	*No/inadequate information to assess risk of confounding*

Appendix 5 Results of analyses to assess potential biases in the linkage results for matched (N=232671) and unmatched (N=44269) primary care records for each of the 38 chronic conditions

	Primary care matched records N=232671	Primary care unmatched records N=44269	St. diff.
Alcohol problems: N (%)			
Present	1587 (0.68)	341 (0.77)	0.010
Absent	231084 (99.3)	43928 (99.2)	
Anorexia or bulimia: N (%)			
Present	1697 (0.73)	436 (0.98)	0.028
Absent	230974 (99.3)	43833 (99.0)	
Anxiety & other related disorders: N (%)			
Present	5163 (2.22)	958 (2.16)	0.004
Absent	227508 (97.8)	5168 (97.8)	
Asthma (currently treated): N (%)			
Present	10684 (4.59)	1686 (3.81)	0.039
Absent	221987 (95.4)	42583 (96.2)	
Atrial fibrillation: N (%)			
Present	2230 (0.96)	584 (1.32)	0.034
Absent	230441 (99.0)	43685 (98.7)	
Blindness and low vision: N (%)			
Present	1709 (0.73)	317 (0.72)	0.002
Absent	230962 (99.3)	43953 (99.3)	
Bronchiectasis: N (%)			
Present	341 (0.15)	73 (0.16)	0.005
Absent	232330 (99.9)	44196 (99.8)	
Cancer: N (%)			
Present	2566 (1.10)	761 (1.72)	0.052
Absent	230105 (98.9)	98.3 (43508)	
Chronic kidney disease: N (%)			
Present	2061 (0.89)	417 (0.94)	0.006
Absent	230610 (99.1)	43852 (0.89)	
Chronic liver disease and viral hepatitis: N (%)			
Present	1837 (0.79)	364 (0.82)	0.004
Absent	230834 (99.2)	43905 (99.2)	
Chronic sinusitis: N (%)			
Present	2373 (1.02)	433 (0.98)	0.004
Absent	230298 (99.0)	43836 (99.0)	
Constipation (treated): N (%)			
Present	2134 (0.92)	582 (1.31)	0.038
Absent	230537 (99.1)	43687 (98.7)	
COPD: N (%)			
Present	3851 (1.66)	772 (1.74)	0.007
Absent	228820 (98.3)	43497 (98.3)	
Coronary heart disease: N (%)			
Present	4347 (1.87)	1070 (2.42)	0.038
Absent	228324 (98.1)	43199 (97.6)	

		Primary care matched records N=232671	Primary care unmatched records N=44269	St. diff.
Dementia: N (%)				
	Present	959 (0.41)	511 (1.15)	0.084
	Absent	231712 (99.6)	43758 (98.8)	
Depression: N (%)				
	Present	11760 (5.05)	1949 (4.40)	0.031
	Absent	220911 (94.9)	42320 (95.6)	
Diabetes: N (%)				
	Present	15139 (6.51)	3139 (7.09)	0.023
	Absent	217532 (93.5)	41130 (92.9)	
Diverticular disease of intestine: N (%)				
	Present	2790 (1.20)	541 (1.22)	0.002
	Absent	229881 (98.8)	43728 (98.8)	
Epilepsy (currently treated): N (%)				
	Present	1191 (0.51)	215 (0.49)	0.004
	Absent	231480 (99.5)	44054 (99.5)	
Hearing loss: N (%)				
	Present	9704 (4.17)	1795 (4.05)	0.006
	Absent	222967 (95.8)	42474 (95.4)	
Heart failure: N (%)				
	Present	1120 (0.48)	357 (0.81)	0.041
	Absent	231551 (99.5)	43912 (99.2)	
Hypertension: N (%)				
	Present	28200 (12.1)	5433 (12.3)	0.005
	Absent	204471 (87.9)	38836 (87.7)	
Inflammatory bowel disease: N (%)				
	Present	1029 (0.44)	217 (0.49)	0.007
	Absent	231642 (99.6)	44052 (99.5)	
Irritable bowel syndrome: N (%)				
	Present	5121 (2.20)	981 (2.22)	0.001
	Absent	227550 (97.8)	43288 (97.8)	
Learning disability: N (%)				
	Present	1858 (0.80)	236 (0.53)	0.033
	Absent	230813 (99.2)	44033 (99.5)	
Migraine: N (%)				
	Present	385 (0.17)	57 (0.13)	0.010
	Absent	232286 (99.8)	44212 (99.9)	
Multiple sclerosis: N (%)				
	Present	238 (0.10)	43 (0.10)	0.002
	Absent	232433 (99.9)	44226 (99.9)	
Painful condition: N (%)				
	Present	13346 (5.74)	2304 (5.20)	0.023
	Absent	219325 (94.3)	41965 (94.8)	
Parkinson's disease: N (%)				
	Present	221 (0.09)	62 (0.14)	0.013
	Absent	232450 (99.9)	44207 (99.9)	
Peptic ulcer disease: N (%)				
	Present	1487 (0.64)	311 (0.70)	0.008
	Absent	231184 (99.4)	43958 (99.3)	

	Primary care matched records N=232671	Primary care unmatched records N=44269	St. diff.
Peripheral vascular disease: N (%)			
Present	668 (0.29)	152 (0.34)	0.010
Absent	232003 (99.7)	44117 (99.7)	
Prostate disorders: N (%)			
Present	1954 (0.84)	422 (0.95)	0.012
Absent	230717 (99.2)	43847 (99.0)	
Psoriasis or eczema: N (%)			
Present	1692 (0.73)	277 (0.63)	0.012
Absent	230979 (99.3)	43992 (99.4)	
Psychoactive substance misuse: N (%)			
Present	1949 (0.84)	503 (1.14)	0.030
Absent	230722 (99.2)	43766 (98.9)	
Rheumatoid arthritis & other related disorders: N (%)			
Present	5321 (2.29)	1093 (2.47)	0.012
Absent	227350 (97.7)	43176 (97.5)	
Schizophrenia or bipolar disorder: N (%)			
Present	11969 (5.14)	2069 (4.67)	0.022
Absent	220702 (94.9)	42200 (95.3)	
Stroke & transient ischaemic attack: N (%)			
Present	2305 (0.99)	603 (1.36)	0.034
Absent	230366 (99.0)	43666 (98.6)	
Thyroid disorders: N (%)			
Present	7405 (3.18)	1516 (3.42)	0.014
Absent	225266 (96.8)	42753 (96.6)	

Standardised differences of 0.2, 0.5, and 0.8 indicate small, medium and large effect sizes respectively (Harron *et al.*, 2017).
*percentages may sum to greater than 100% due to rounding

Appendix 6 Results of subgroup analyses for the general, physical-mental, and complex multimorbidity outcomes

Estimated odds ratios of basic, physical-mental, and complex multimorbidity with household tenure when the final models tested for interactions between tenure and household benefits receipt for working age adults residing in B&D in 2019/20 (N=129985)

Independent variables		Basic multimorbidity		Physical-mental multimorbidity		Complex multimorbidity	
		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Tenure	OOC	-	-	-	-	-	-
	Privately rented	0.76 (0.72-0.80)	<.001	0.76 (0.69-0.85)	<.001	0.76 (0.69-0.83)	<.001
	Social housing	1.38 (1.32-1.45)	<.001	1.55 (1.43-1.68)	<.001	1.36 (1.26-1.47)	<.001
Household benefits receipt	No benefits	-	-	-	-	-	-
	ESA	4.05 (3.27-5.02)	<.001	7.51 (5.84-9.67)	<.001	6.98 (5.48-8.90)	<.001
	Pension credit	1.45 (1.15-1.84)	0.002	1.58 (1.03-2.43)	0.037	1.72 (1.22-2.42)	0.002
	Income support	3.37 (2.37-4.80)	<.001	2.84 (1.62-4.97)	<.001	2.77 (1.71-4.48)	<.001
	JSA	0.89 (0.39-2.04)	0.777	0.68 (0.09-5.09)	0.711	1.54 (0.52-4.60)	0.436
	Housing benefit only	1.91 (1.63-2.23)	<.001	2.45 (1.91-3.15)	<.001	1.97 (1.57-2.48)	<.001

Independent variables		Basic multimorbidity		Physical-mental multimorbidity		Complex multimorbidity	
		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Tenure*Household benefits receipt	Privately rented*no benefits	-	-	-	-	-	-
	Privately rented*ESA	1.71 (1.34-2.19)	<.001	1.43 (1.06-1.93)	0.018	1.35 (1.01-1.80)	0.042
	Privately rented* pension credit	1.53 (1.11-2.12)	0.010	1.96 (1.11-3.46)	0.021	1.63 (1.01-2.62)	0.044
	Privately rented* income support	1.00 (0.67-1.49)	0.995	1.34 (0.71-2.52)	0.365	1.58 (0.90-2.76)	0.111
	Privately rented* JSA	2.38 (0.95-5.95)	0.065	1.96 (0.22-17.2)	0.544	1.65 (0.48-5.70)	0.428
	Privately rented* housing benefit only	0.93 (0.78-1.10)	0.405	0.96 (0.72-1.28)	0.775	1.03 (0.79-1.35)	0.803
	Social housing*no benefits	-	-	-	-	-	-
	Social housing*ESA	1.10 (0.87-1.38)	0.432	0.76 (0.58-1.00)	0.051	0.90 (0.69-1.17)	0.415
	Social housing* pension credit	1.04 (0.78-1.38)	0.787	0.97 (0.59-1.58)	0.888	0.88 (0.58-1.33)	0.546
	Social housing* income support	0.81 (0.55-1.17)	0.259	0.92 (0.51-1.66)	0.790	1.36 (0.82-2.28)	0.234
	Social housing*JSA	1.74 (0.73-4.15)	0.209	2.04 (0.27-15.7)	0.493	1.09 (0.35-3.44)	0.877
	Social housing* housing benefit only	0.97 (0.82-1.16)	0.767	1.01 (0.77-1.33)	0.924	1.16 (0.90-1.49)	0.254

Estimated odds ratios of basic, physical-mental, and complex multimorbidity with household tenure when the final models tested for interactions between tenure and household occupancy for working age adults residing in B&D in 2019/20 (N=129985)

Independent variables		Basic multimorbidity		Physical-mental multimorbidity		Complex multimorbidity	
		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Tenure	OOC	-	-	-	-	-	-
	Privately rented	0.99 (0.91-1.08)	0.819	1.05 (0.92-1.21)	0.452	1.03 (0.91-1.17)	0.638
	Social housing	1.54 (1.43-1.65)	<.001	1.60 (1.42-1.79)	<.001	159 (1.43-1.76)	<.001
Occupancy categories	1-2 occupants	-	-	-	-	-	-
	3-5 occupants	1.02 (0.96-1.09)	0.496	0.98 (0.87-1.10)	0.679	1.03 (0.93-1.14)	0.568
	6-10 occupants	1.02 (0.93-1.11)	0.707	0.80 (0.66-0.97)	0.022	1.13 (0.97-1.31)	0.114
	11+ occupants	0.90 (0.67-1.20)	0.468	0.63 (0.29-1.35)	0.236	1.28 (0.80-2.03)	0.301
Tenure*Occupancy	Privately rented* 1-2 occupants	-	-	-	-	-	-
	Privately rented* 3-5 occupants	0.75 (0.68-0.83)	<.001	0.71 (0.60-0.84)	<.001	0.74 (0.63-0.87)	<.001
	Privately rented* 6-10 occupants	0.69 (0.61-0.78)	<.001	0.77 (0.60-0.99)	0.038	0.62 (0.50-0.77)	<.001
	Privately rented* 11+ occupants	0.70 (0.47-1.03)	0.068	0.97 (0.39-2.39)	0.949	0.34 (0.17-0.69)	0.003

Independent variables		Basic multimorbidity		Physical-mental multimorbidity		Complex multimorbidity	
		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Tenure*Occupancy (continued)	Social housing* 1-2 occupants	-	-	-	-	-	-
	Social housing* 3-5 occupants	0.85 (0.78-0.93)	<.001	0.88 (0.76-1.01)	0.060	0.81 (0.72-0.92)	0.002
	Social housing* 6-10 occupants	0.82 (0.73-0.93)	0.002	0.96 (0.76-1.20)	0.697	0.62 (0.50-0.75)	<.001
	Social housing* 11+ occupants	0.40 (0.22-0.74)	0.003	0.85 (0.29-2.49)	0.762	0.40 (0.16-0.98)	0.045

Estimated odds ratios of basic, physical-mental, and complex multimorbidity with household tenure when the final models tested for interactions between tenure and household type for working age adults residing in B&D in 2019/20 (N=129985)

Independent variables		Basic multimorbidity		Physical-mental multimorbidity		Complex multimorbidity	
		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Tenure	OOC	-		-	-	-	-
	Privately rented	0.73 (0.69-0.78)	<.001	0.80 (0.70-0.91)	<.001	0.67 (0.59-0.75)	<.001
	Social housing	1.40 (1.31-1.49)	<.001	1.56 (1.37-1.77)	<.001	1.22 (1.09-1.37)	<.001
Household type	Adults with children	-	-	-	-	-	-
	Adults with no children	1.23 (1.16-1.31)	<.001	1.27 (1.12-1.43)	<.001	1.14 (1.03-1.26)	0.014
	Single adult with children	0.91 (0.76-1.09)	0.320	0.83 (0.57-1.22)	0.349	0.72 (0.49-1.05)	0.091
	Single adult	1.14 (1.02-1.27)	0.017	1.44 (1.19-1.75)	<.001	1.01 (0.85-1.20)	0.939
	Older cohabiting adults	1.42 (1.29-1.55)	<.001	1.34 (1.13-1.60)	<.001	1.33 (1.15-1.54)	<.001
	Three generations	1.06 (0.93-1.20)	0.402	1.06 (0.80-1.40)	0.702	1.12 (0.90-1.39)	0.308
Tenure*Household type	Privately rented*adults with children	-	-	-	-	-	-
	Privately rented*adults with no children	1.09 (0.99-1.20)	0.090	1.05 (0.87-1.26)	0.616	1.27 (1.08-1.49)	0.004
	Privately rented*single adult with children	1.17 (0.94-1.45)	0.157	1.28 (0.83-1.97)	0.270	1.17 (0.74-1.85)	0.497
	Privately rented*single adult	1.59 (1.37-1.83)	<.001	1.36 (1.06-1.74)	0.015	1.83 (1.46-2.31)	<.001

Independent variables		Basic multimorbidity		Physical-mental multimorbidity		Complex multimorbidity	
		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Tenure*Household type (continued)	Privately rented*older cohabiting adults	1.25 (1.01-1.53)	0.033	1.31 (0.93-1.85)	0.125	1.69 (1.26-2.27)	<.001
	Privately rented*three generations	0.99 (0.78-1.25)	0.912	0.98 (0.61-1.57)	0.924	1.16 (0.79-1.72)	0.451
	Social housing*adults with children	-	-	-	-	-	-
	Social housing*adults with no children	0.90 (0.83-0.99)	0.025	0.89 (0.76-1.04)	0.131	1.11 (0.96-1.28)	0.145
	Social housing*single adult with children	1.22 (0.98-1.50)	0.071	1.29 (0.85-1.96)	0.224	1.48 (0.97-2.27)	0.070
	Social housing*single adult	1.33 (1.17-1.51)	<.001	1.08 (0.87-1.33)	0.481	1.47 (1.21-1.78)	<.001
	Social housing*older cohabiting adults	0.75 (0.65-0.88)	<.001	0.83 (0.65-1.07)	0.152	0.86 (0.69-1.08)	0.204
	Social housing*three generations	0.92 (0.72-1.17)	0.483	0.78 (0.49-1.23)	0.282	0.94 (0.63-1.41)	0.776

Appendix 7 Results of sensitivity analyses for the general, physical-mental, and complex multimorbidity outcomes

Estimated odds ratios of basic multimorbidity with household tenure when the final model was single-level only, accounted for household-level clustering only, and accounted for household and area-level clustering for working age adults residing in B&D in 2019/20 (N=129985)

	Final, fully adjusted model* OR (95% CI)	Single-level model OR (95% CI)	Household-level model OR (95% CI)	Household and area-level model OR (95% CI)
Household tenure				
Owner-occupied (ref)	-	-	-	-
Social housing	1.36 (1.31-1.42)	1.33 (1.27-1.38)	1.33 (1.28-1.39)	1.37 (1.31-1.43)
Privately rented	0.79 (0.75-0.83)	0.78 (0.74-0.81)	0.78 (0.74-0.81)	0.79 (0.75-0.83)

*Area-level only model
All models adjusted for age, gender, ethnicity, BMI, smoking status, household benefits receipt, household occupancy and household type
OR = odds ratio

Estimated odds ratios of physical-mental multimorbidity with household tenure when the final model was single-level only, accounted for household-level clustering only, and accounted for household and area-level clustering for working age adults residing in B&D in 2019/20 (N=129985)

	Final, fully adjusted model* OR (95% CI)	Single-level model OR (95% CI)	Household-level model OR (95% CI)	Household and area-level model OR (95% CI)
Household tenure				
Owner-occupied (ref)	-	-	-	-
Social housing	1.47 (1.37-1.58)	1.46 (1.36-1.57)	1.47 (1.36-1.58)	1.48 (1.37-1.60)
Privately rented	0.85 (0.78-0.92)	0.84 (0.78-0.91)	0.84 (0.77-0.92)	0.84 (0.78-0.92)

*Area-level only model
All models adjusted for age, gender, ethnicity, BMI, smoking status, household benefits receipt, household occupancy and household type, and interactions between household tenure and benefits receipt
OR = odds ratio

Estimated odds ratios of complex multimorbidity with household tenure when the final model was single-level only, accounted for household-level clustering only, and accounted for household and area-level clustering for working age adults residing in B&D in 2019/20 (N=129985)

	Final, fully adjusted model* OR (95% CI)	Single-level model OR (95% CI)	Household-level model OR (95% CI)	Household and area-level model OR (95% CI)
Household tenure				
Owner-occupied (ref)	-	-	-	-
Social housing	1.34 (1.26-1.44)	1.34 (1.25-1.43)	1.34 (1.25-1.43)	1.35 (1.26-1.44)
Privately rented	0.81 (0.75-0.87)	0.80 (0.75-0.87)	0.80 (0.74-0.87)	0.81 (0.75-0.87)

*Area-level only model
All models adjusted for age, gender, ethnicity, BMI, smoking status, household benefits receipt, household occupancy and household type, and interactions between household tenure and benefits receipt household type and household occupancy
OR = odds ratio

Estimated odds ratios of different definitions of basic multimorbidity with household tenure for working age adults residing in B&D in 2019/20 (N=129985)

	Model 4 OR (95% CI)	P value	Model 4.1 OR (95% CI)	P value	Model 4.2 OR (95% CI)	P value	Model 4.3 OR (95% CI)	P value
Household tenure								
Owner-occupied (ref)	-		-		-		-	
Social housing	1.36 (1.31-1.42)	<.001	1.44 (1.37-1.50)	<.001	1.36 (1.30-1.42)	<.001	1.45 (1.38-1.52)	<.001
Privately rented	0.79 (0.75-0.83)	<.001	0.79 (0.75-0.83)	<.001	0.80 (0.77-0.84)	<.001	0.80 (0.76-0.84)	<.001

OR = odds ratio
Model 4 – model fully adjusted for individual-level sociodemographic characteristics (age, gender and ethnicity), individual-level behavioural characteristics (BMI and smoking status) and household-level sociodemographic characteristics (household benefits receipt, household occupancy and household type) and interaction terms
Model 4.1 – model fully adjusted for model 4 covariates with multimorbidity definition excluding "risk factor" conditions (atrial fibrillation and hypertension)
Model 4.2 – model fully adjusted for model 4 covariates with multimorbidity definition excluding conditions flagged as having coding concerns/meaningfulness of flags (stages 1-3 chronic kidney disease, psoriasis or eczema, chronic sinusitis, constipation, diverticular disease of intestine, prostate disorder, double counted chronic obstructive pulmonary disease and bronchiectasis and double count anorexia and depression and/or anxiety)
Model 4.3 – model fully adjusted for model 4 covariates with multimorbidity definition excluding conditions excluded in models 4.1 and 4.2

Estimated odds ratios of different definitions of physical-mental multimorbidity with household tenure for working age adults residing in B&D in 2019/20 (N=129985)

	Model 4 OR (95% CI)	P value	Model 4.1 OR (95% CI)	P value	Model 4.2 OR (95% CI)	P value	Model 4.3 OR (95% CI)	P value
Household tenure								
Owner-occupied (ref)	-		-		-		-	
Social housing	1.47 (1.37-1.58)	<.001	1.52 (1.40-1.64)	<.001	1.47 (1.36-1.58)	<.001	1.52 (1.40-1.64)	<.001
Privately rented	0.85 (0.78-0.92)	<.001	0.85 (0.78-0.93)	<.001	0.85 (0.78-0.93)	<.001	0.87 (0.79-0.95)	0.001

OR = odds ratio
Model 4 – model fully adjusted for individual-level sociodemographic characteristics (age, gender and ethnicity), individual-level behavioural characteristics (BMI and smoking status) and household-level sociodemographic characteristics (household benefits receipt, household occupancy and household type) and interaction terms
Model 4.1 – model fully adjusted for model 4 covariates with multimorbidity definition excluding "risk factor" conditions (atrial fibrillation and hypertension)
Model 4.2 – model fully adjusted for model 4 covariates with multimorbidity definition excluding conditions flagged as having coding concerns/meaningfulness of flags (stages 1-3 chronic kidney disease, psoriasis or eczema, chronic sinusitis, constipation, diverticular disease of intestine, prostate disorder, double counted chronic obstructive pulmonary disease and bronchiectasis and double count anorexia and depression and/or anxiety)
Model 4.3 – model fully adjusted for model 4 covariates with multimorbidity definition excluding conditions excluded in models 4.1 and 4.2

Estimated odds ratios of different definitions of complex multimorbidity with household tenure for working age adults residing in B&D in 2019/20 (N=129985)

	Model 4 OR (95% CI)	P value	Model 4.1 OR (95% CI)	P value	Model 4.2 OR (95% CI)	P value	Model 4.3 OR (95% CI)	P value
Household tenure								
Owner-occupied (ref)	-		-		-		-	
Social housing	1.34 (1.26-1.44)	<.001	1.42 (1.31-1.53)	<.001	1.37 (1.28-1.47)	<.001	1.48 (1.36-1.60)	<.001
Privately rented	0.81 (0.75-0.87)	<.001	0.80 (0.73-0.88)	<.001	0.84 (0.77-0.91)	<.001	0.86 (0.78-0.94)	0.002

OR = odds ratio
Model 4 – model fully adjusted for individual-level sociodemographic characteristics (age, gender and ethnicity), individual-level behavioural characteristics (BMI and smoking status) and household-level sociodemographic characteristics (household benefits receipt, household occupancy and household type) and interaction terms
Model 4.1 – model fully adjusted for model 4 covariates with multimorbidity definition excluding "risk factor" conditions (atrial fibrillation and hypertension)
Model 4.2 – model fully adjusted for model 4 covariates with multimorbidity definition excluding conditions flagged as having coding concerns/meaningfulness of flags (stages 1-3 chronic kidney disease, psoriasis or eczema, chronic sinusitis, constipation, diverticular disease of intestine, prostate disorder, double counted chronic obstructive pulmonary disease and bronchiectasis and double count anorexia and depression and/or anxiety)
Model 4.3 – model fully adjusted for model 4 covariates with multimorbidity definition excluding conditions excluded in models 4.1 and 4.2

Appendix 8 Topic guide used for individual interviews

Version 3 – 27.01.2020

Introduction: introduce self, the project, give information sheet and consent form, ask about audio recording

Job role/setting the scene:

- **To start, could you tell me a bit about what your job involves?**
    - How do you collaborate with others either within or outside of your organisation? How does this look/work in practice?
    - **You've mentioned / others have described this as working across a 'system'/working across a 'health and care system', just so we're on the same page could you describe what you mean by this?**
- Thinking about the STP, the STP aims to bring together North London health and social care services to provide the **entire** local population with access to the best possible health and care, and to make **North London a place where no-one is left behind**. This topic of health inequalities is of particular interest to this work and my PhD, and I'd like to try and structure this conversation around this issue.
- So on this topic, and for me to also be able to get a better idea about your role, could you please give me an example of a recent service planning decision you have made **around health inequalities** that has implications for your H&C system?
    - What was the aim of this decision?
    - If no, could you give me an example of a recent service planning decision you have made a) around health inequalities or b) for the system you just described?

Use of analytics in that decision

Moving forward, I'd like to try and centre this conversation on that decision you've just described.

- How did you make that decision?
    - Did you collaborate with others either within or outside of your organisation to make that decision?
    - What types of evidence did you draw on?
        - **You've mentioned data/analytics a few times, just to ensure we're on the same page what do you mean by that term?**
        - Where did you obtain this piece of analytics from?
        - Did you obtain analytics from any other source (e.g. other NHS/Local Authority)?
    - **How did you use analytics from X to inform your decision around Y?**
        - How did you initially intend to use analytics when you commissioned it?
        - Were you able to do that?
        - Why not?
        - How was analytics incorporated with the other types of evidence (e.g., academic literature, expert opinions) that you've mentioned?

Process of sourcing and obtaining analytics

- **How did you go about obtaining your analytics?**
    - Who did you contact? Did you face any difficulties trying to find who to contact?
    - How did you come up with the research question(s)?
    - Can you tell me about your involvement in the sourcing of the analytics from X?
    - How did you find this process?
    - Did you face any issues with receiving this analytical output from X?
    - Did you face any issues with using this analytical output from X i.e. with taking it to other colleagues not involved in the commission day-to-day?

Training needs/support

- **Would it influence your decision-making around X if you had access to data linked across health and council records?**
- **Was there anything that would have made it easier for you to use analytics in this decision you're describing?**
- **In the future, would you like anything to change?**
    - How comfortable would you say you were using this piece of analytics for the decision you have described?
    - Did you have support to use analytics in this decision?
    - What kind of support could help you better use analytics for service planning decisions around health inequalities?
    - Do you think any training (statistical or otherwise) would have made a difference to your use of analytics (from NHS or council) in this context? Could any training of this kind help you to make future service planning decisions around health inequalities?
    - Aside from what we've already mentioned, are there any other things that could have been done differently that could have helped you use that piece of analytics for this decision you're describing?

Appendix 9 Coding frame for individual, semi-structured interviews

Name	Description	Example quote(s)
Theme 1: Ways in which analytics are used		
Planning service delivery	Uses "higher level" pseudonymised data and/or analytics to help inform strategic choices of investment and disinvestment, plan new services, or redesign existing services, and understand the impacts of changing or implementing new services model.	"We only had one other nursing home in the borough, we're reliant on out of borough placements, which is fine, because we operate as a system across XXX for nursing care. But there's pressure across the whole of XXX for nursing care, so if we don't play our part in providing some capacity, it puts pressure elsewhere in the system. So, based on that, we did quite a lot more detailed financial modelling data, running data, on future demand, on the future provision picture. And came out with a view that we needed to expand the capacity in the home. Not to close it, not to reopen on the same basis, but to double the capacity." (ID008, Social Care Commissioner, 'Advanced' analytics user)
Monitoring and evaluation	Uses "higher level" pseudonymised data and/or analytics to monitor whether a service is meeting certain outcomes, performing as expected and/or to compare service/organisational performance with others.	"So, we would do a lot of sort of benchmarking. So, you know, how does XXX compare to you know, if a [patient] comes in here for [surgery], how long would [they] stay at XXX compared to if they'd had that somewhere else." (ID010, Health Provider, 'Reluctant' analytics user)
Payment for performance	Uses "higher level" pseudonymised data and/or analytics to determine how much providers of services should be paid for their activity.	"For frailty we say look at the electronic frailty index, use your clinical judgement about whether you think this [unidentifiable] person is frail or not. Put them onto a frailty register, and then can you call them in to do a care plan with multidisciplinary team input, and can you review that care plan as the year goes on and provide them a year of care. If you do that, we'll pay you £100." (ID018, Health Commissioner, 'Advanced' analytics user)

Name	Description	Example quote(s)
Clinical decision making	Uses patient identifiable data and/or analytics to inform clinical decisions i.e., how they manage patients in their services.	"What we've been doing more recently is supporting practices…we actually provide them with lists of patients that they need to call in, because they need to do a review with them and that sort of thing." (ID012, Health Provider, 'Challenged' analytics user)
Theme 2: Macro factors related to the working environment		
System structures		
Organisational fragmentation	Divisions between, and within, organisations creates siloed data systems, barriers to data sharing and fragmented personnel structures (particularly between analysts and decision-making teams).	"[RES]: I think one real difficult thing is around, again, the interoperability between different systems and the same data but the systems can't talk to each other so taking the needs analysis as an example we need housing data, we need social care data, we need some health data, but it's proving difficult to get those data sources and then when we do eventually get them it's a lot of work to then bind them together trying to paint that coherent story, but there's also issues around [asking] where does the data sit. So, I had a meeting with [another internal team] asking for some data. They're like but this sits here, this doesn't sit with us. It's unclear where … who owns certain pieces of data and how best to extract it. [INT]: Is that the reason that you have issues accessing it in the first place? [RES]: Definitely. So, housing data, in particular, where it sits, in a completely different department, a different team. We have no right to access any of that data, so it will take quite a lot of time to get it." (ID023, Social Care Commissioner, 'Challenged' analytics user) "I think I've got better insight [than analysts when it comes to] actually synthesising all of that available data and actually putting it into a logical pattern for further exploration and investigation. So, at best, the reports I've seen from XXX tend to be just slightly superficial in terms of that analytics piece, the insight piece, because their analysts are too remote from the policy and strategy bit of it." (ID021, Health and Social Care Commissioner, 'Hands-On' analytics user)

Name	Description	Example quote(s)
Alignment of priorities	To get organisations to work together collaboratively, priorities across the system need to be aligned, agreed upon and all organisations within the system decision need to feel ownership over the decision being taken. This is particularly an issue for some providers.	"The balance for developing new models of care, the impetus is not really there [for us]. Because if we, today, have a patient who we see in the hospital we get paid £70 or something for a follow-up patient. If we work out a new model of care where this patient can be seen in the community or virtually, we would get paid £10 or £15 or something. What on earth would we want to do that for? It doesn't make any sense at all…. If you're saying let's [in a] wholesale [manner] move half of our patients into the community, let's lose all of that revenue, then suddenly the fixed costs that we have in this building and others become overwhelming. Our sort of model is predicated on getting the type of revenue in from these types of patients…..It is a very huge risk to us." (ID011, Health Provider, 'Reluctant' analytics user)
Top-down constraints		
Resource pressures	Influences process when obtaining, interpreting, and using analytics. Limited funding and resources can influence analytics use as decisions sometimes need to be made in short time frames and resource pressures influence the size and capacity of teams of analysts across the system.	"I feel slightly silly saying this, even though it's a small number [the proportion of the population with learning disabilities] it seems really difficult to get it right and I think that's because we haven't, we don't have a dedicated team locally, there's a transforming care team in XXX, but we don't have a dedicated team that are focused on keeping the data clean and accurate. It's a split responsibility across teams, unfortunately, and that does mean that it can become a bit like messy and muddy" (ID022, Health and Social Care Commissioner, 'Challenged' analytics user).
Policy priorities	Some decisions made because of (or heavily shaped by) policy priorities, both locally and nationally, influencing where/how data can be used.	"But the first thing was just about, and I hate to say this, it felt like ticking a box and getting it green to say that we've got all these networks. Then phase two was well what can we drop in, and it just felt very robotic or artificial in terms of a process." (ID013, Health Commissioner, 'Hands-On' analytics user) "It's much harder for commissioners to continue to make the case to commission something that isn't showing an immediate return on investment" (ID020, Health Commissioner, 'Advanced' analytics user)

Name	Description	Example quote(s)
Theme 3: Micro factors related to the individuals involved		
Personal relationships		
Relationships between leaders and analysts	Working relationships between leaders and analysts influence the process of requesting/obtaining analytics e.g., how questions were asked of data, and how output content was developed and used. Relationships vary considerably with different individuals.	"[RES]: We [leader and analyst] talked about the scope, they took the lead, so [the analyst] was fab, as [they are]. We kind of described the scope of the strategy, and what we'd intended it to do, and then [the analyst] went off and led [the work]. We had a couple of meetings to check in every so often and she went off and led a team, , did some kind of very nice heat map, showing levels of the deprivation and kind of different population needs, the different long-term conditions for example. [INT]: So how did the process work in terms of developing the questions asked of the data? [RES]: I don't know actually. [The analyst] and I have worked together on and off for years, so maybe it was that. Also, I just inherently trust [the analyst] to know what [they're] doing and [they have] done it loads before." (ID020, Health Commissioner, 'Advanced' analytics user) "[RES]: I would always want an open dialogue [with analysts] but sometimes the other party might want to just tell me, be told what to do. So, I've had varying levels of interest in analytics teams around the extent to which they're really genuinely interested in it, which has been disappointing for me because when I've worked as a head of analytics elsewhere most of my analysts have been people who are curious. [INT]: Would you ever go to kind of analytical teams in other parts of the system? [RES]: Well erm, yes. I sometimes go to the, there's a better analyst in [another organisation], for example, I would nick [them] sometimes. I would trust [their] judgement around it, because I know after a few words of briefing from me, [they] would pick out the themes and I wouldn't have to look over [their] shoulder every five minutes". (ID021, Health and Social Care Commissioner, 'Hands-On' analytics user)

Name	Description	Example quote(s)
Relationships between leaders and leaders	Nature of relationships influence priority and timeline setting, and collaborative decision making, for a health and care system. Building trust between individuals/teams from different organisations is a challenge but a necessity.	"And we've worked with the team for a couple of years now as well, so there's some trust there as well between, like that we wouldn't try to implement something that wasn't robust." (ID018, Health Commissioner, 'Advanced' analytics user) "For a project to be successful you've got to have the buy in of everybody who's leading that project. And unfortunately, in [our area], that is a challenge, especially when you've got a project such as this which is quite a large project, which doesn't happen overnight. You've got different members of staff coming and going, and that handoff isn't necessarily always there. You might have an agreement at the offset of the project and of course that would be a high level of agreement in the initial stages. But when you get down to the nitty gritty and then you're dealing with a new set of people maybe sitting around a room….. Relationships are absolutely imperative to getting that done, which, I think, in [our area] is a challenge, because of the throughput of staff. For those of us around it's absolutely key to getting those relationships, learning how to develop those relationships very quickly, and getting trust. Trust is really important." (ID016, Health Provider, 'Challenged' analytics user)
Skills and knowledge		
Leaders' skills and knowledge	There is ambiguity around the meaning of the terms "data" and "analytics" despite them being used regularly by commissioners/data requesters and analysts. Participant responses suggested there is considerable variation in leaders' abilities to understand and interpret analytical outputs or do their own workarounds. A handful of participants acknowledged that they may have training needs, particularly around interpreting data correctly. This view is not universally shared.	"We're probably not very good at um, or we're not in fact I remember talking to [certain] teams about you know, randomising, or not randomising, what's the word, um, oh statistical significance and all that sort of stuff. We'd be fairly rubbish at that I think." (ID012, Health Provider, 'Challenged' analytics user) "Having had a bit of a data analytical background in a previous life, [more complicated analytical work] is the kind of stuff that I always enjoyed…. So, yeah, it was basically myself doing it." (ID013, Health Commissioner, 'Hands-On' analytics user)

Name	Description	Example quote(s)
Analysts' skills and knowledge	Analytics skills can vary across organisations, although often leaders describe this variation as stemming from resource pressures affecting size and capacity of teams of analysts.	"I think we should actually have statisticians employed by XXX who can make sense of the data for other clinicians…I think there are people from our side who will produce graphs and data and so on but not to the extent or a regression analysis or building algorithms or so on at the moment.…it's the more of the predictor stuff rather than just the post-hoc stuff." (ID009, Health Provider, 'Waiting' analytics user) "I think the skills of the [analysts] vary quite a bit and I'm not convinced, being an ex-analyst myself and having run an analytics team, that they get the help that they need. So I find myself knowing what I need quicker than they do and being able to dictate terms a little bit better than some of my commissioning colleagues.....I think the [analysts] aren't great, with all due respect to them, at the analytics part, so I think they're good at the data preparation part. And that's why I was saying I kind of commissioned them to do the informatics [data] piece, but not the analytics piece. Because I don't think they would have the insight I would have as a commissioner into the data." (ID021, Health and Social Care Commissioner, 'Hands-On' analytics user)
Theme 4: Meso factors related to data quality		
Data availability and accuracy	Data (and subsequent analytics) availability is influenced by whether the necessary data are initially recorded and the perceived "correctness" of the data.	"Our population projections are around seven to eight thousand adults with a learning disability in the borough. We've got about fifteen hundred registered with the GPs…one of the reasons could be how people with a learning disability are actually recorded on systems." (ID022, Health and Social Care Commissioner, 'Challenged' analytics user). "We were spending all of our time arguing about data between [analysts] and the provider teams. Why the hell are we doing that? We could just agree to disagree or whatever, but surely you want your analysts to do higher order analytics stuff, not to piss round (pardon my French) arguing about a million records … Do you really want to be doing that or do you want to be doing a simulation model to understand what the level of resource is we need for A&E." (ID021, Health and Social Care Commissioner, 'Hands-On' analytics user)

Name	Description	Example quote(s)
Data richness and linkage	The detail and richness of information recorded can influence utility of data (and subsequent analytics). Without linking data from different systems, data is siloed and disconnected. This limits leaders' to fully understand population health needs and plan or evaluate services accordingly	"The actual data that's held by GPs is, essentially, a contract reward. So, it's every health check that's carried out by a GP for somebody with a learning disability - they will qualify for an additional payment. So, we're using information which is kind of contract related rather than the individual. We're not getting anything around the quality of that health check, we know it's been done because they've been paid for it, but we don't know how long it took, we don't know the outcomes from that individual health check. So, drilling down into that information is really I think the next, you know, if it's developing the quality of the information, and what comes out of the health check should be out next kind of priority really. So, it's not just numbers." (ID022, Health and Social Care Commissioner, 'Challenged' analytics user). "What would be amazing obviously in terms of evaluation [of the National Diabetes Programme locally] is that you could link those people back into their health record and then get some kind of long-term outcome for people with diabetes who'd been in the prevention programme. But what we would currently look at [with unlinked programme data] is just engagement with the programme and dropout rates." (ID017, Public Health Lead, 'Hands-On' analytics user) "We looked at primary care data, [and] prevalence [of various health conditions]. Then we looked at some acute data, and we managed to link the acute and primary care data through the pseudonymised NHS number. But because we had all this [geographical] mapping in our data, we said, actually, well we can link to [area-level deprivation data]….We then looked at it, and what we ended up with was six very different projects, so not this blanket one size fits all." (ID013, Health Commissioner, Hands-On User)